

# Prediction and Mapping Physical and Chemical Soil Properties in Semi-Arid Regions: A Comparative Analysis of Machine Learning Algorithms

Mohammad Rastin <sup>a, \*</sup>, Ali Salehi <sup>b, \*</sup>, Ali Asghar Zolfaghari <sup>c</sup>, Nafiseh Yaghmaeian <sup>d</sup>

<sup>a</sup> PhD Student, Department of Forest Science and Engineering, Faculty of Natural Resources, University of Guilan, Soumeh Sara, Iran

<sup>b, \*</sup> Associate Professor, Department of Forest Science and Engineering, Faculty of Natural Resources, University of Guilan, Soumeh Sara, Iran

<sup>c</sup> Associate Professor, Department of Natural Resources, Faculty of Desert Studies, University of Semnan, Semnan, Iran

<sup>d</sup> Associate Professor, Department of Soil Engineering Sciences, Faculty of Agricultural Sciences, University of Guilan, Rasht, Iran

Research Full Paper

Article History (Received: 2024/12/22

Accepted: 2025/03/30)

## Extended abstract

### 1- Introduction

In recent years, with the rise of challenges such as food security, land degradation, water resource scarcity, climate change, and biodiversity conservation, the importance of understanding the physical and chemical properties of soils has increased significantly. These properties have a direct impact on food production, carbon storage, water filtration, and climate regulation, making accurate information on these characteristics essential for environmental and agricultural management. Traditional methods of measuring soil properties have limited applicability over large areas due to their cost and time requirements. However, the emergence of machine learning algorithms and remote sensing data has facilitated more precise mapping. Digital Soil Mapping (DSM) has emerged as an innovative approach that combines advanced algorithms with environmental variables to provide high-quality estimates of soil properties. Recent studies indicate that utilizing models such as Random Forest (RF) and Cubist, paired with topographic and remote sensing data, is effective for predicting soil variables and can significantly enhance the management of natural resources, agriculture, and watersheds. In this study, sampling locations were determined using the Cubist Latin Hypercube Sampling (cLHS) method.

### 2- Material and methods

Auxiliary variables in the Latin Hypercube Model included bands and indices from Sentinel-2 satellite imagery, encompassing visible spectrum bands, short-wave infrared, vegetation indices, and properties derived from a Digital Elevation Model (DEM), including elevation, slope, aspect, and valley flatness index with high resolution. The Boruta method was employed to select influential variables. The machine learning models used in this study included RF, Cubist, and RF-Cubist, with the data divided into two sections: training (70%) and testing (30%). Accordingly, all data were evaluated as testing data based on the aforementioned method.

### 3- Results

The results indicated that for the depth of 0-30 cm, the Random Forest (RF) model demonstrated the best performance for sand, with RMSE, Bias, Correlation (Cor), and Relative Root Squared Error (RRSE) values of 10.19%, -0.30, 0.92, and 0.41, respectively. Additionally, the Cubist model provided the best estimates for organic carbon, with RMSE, Bias, Cor, and RRSE values of %0.23, 0.005, 0.71, and 0.70, respectively. The RF-Cubist model achieved satisfactory results in predicting silt, with RMSE, Bias, Cor, and RRSE values of 7.80, -0.87, 0.87, and 0.28, respectively. At the depth of 30-60 cm, the RF model again performed best for sand, yielding RMSE, Bias, Cor, and RRSE values of %11.42, -0.23, 0.88, and

\* Corresponding Author: [asalehi@guilan.ac.ir](mailto:asalehi@guilan.ac.ir)

0.47, respectively. The Cubist model was noted for its performance in estimating organic carbon, achieving RMSE, Bias, Cor, and RRSE values of %0.26, 0.05, 0.53, and 0.85, respectively. Furthermore, the RF-Cubist model yielded good results in predicting silt, with RMSE, Bias, Cor, and RRSE values of %9.28, -0.58, 0.83, and 0.30, respectively. Uncertainty maps for the various parameters indicated that the RF model exhibited lower uncertainty at many locations compared to other models.

#### 4- Discussion & Conclusions

A specific model should provide the best results with the least error. For instance, at a depth of 0-30 cm, the Random Forest (RF) model exhibited higher accuracy in estimating soil physical properties, characterized by a high  $R^2$  and lower RMSE, compared to the Cubist and RF-Cubist models. Conversely, the Cubist model demonstrated superior accuracy in estimating soil chemical properties. At a depth of 30-60 cm, the RF-Cubist model outperformed the RF and Cubist models, achieving the best performance with a high  $R^2$  and lower RMSE. According to results from other studies, the variability of soil properties is attributed to environmental factors, diversity of parent materials, and ecomorphological processes, leading to the formation of different landforms and influencing soil particles, especially clay. The low average soil organic carbon at both surface and subsurface depths (0.36 and 0.33) is affected by factors such as poor vegetation cover, limited use of organic fertilizers, and reduced soil fertility. The use of the Boruta model for selecting environmental variables has reduced prediction errors compared to other methods, and variables derived from the Digital Elevation Model (DEM) have been recognized as effective in class modeling and soil property estimation. In comparing predictive models, the RF model demonstrated superior performance due to its lower error values (RMSE and Bias) and higher correlation coefficients compared to the Cubist and RF-Cubist models. Nevertheless, the Cubist model also displayed high accuracy in modeling due to its ability to capture the complexities between soil characteristics. Multiple studies have further confirmed that the RF model provides reliable results with low uncertainty when estimating properties such as organic carbon percentage and clay content. Ultimately, to improve results in future studies, it is recommended that model selection be tailored to the sampling depth and type of soil characteristics, data quality be enhanced, and the use of hybrid models be further developed.

**Key Words:** Climate change, Climate disasters, Erosion projection, Minab basin, RUSLE.

---

**Cite this article:** Rastin, M., Salehi, A., Zolfaghari, A.A., & Yaghmaeian, N. (2025). Prediction and Mapping Physical and Chemical Soil Properties in Semi-Arid Regions: A Comparative Analysis of Machine Learning Algorithms. *Journal of Environmental Erosion Research*. 2025; 15 (2):82-109. <http://doi.org/10.61186/jeer.15.2.82>



© The Author(s).

DOI: : <http://doi.org/10.61186/jeer.15.2.82>

Published by Hormozgan University Press.

URL: <http://magazine.hormozgan.ac.ir>

---

## پیش‌بینی و نقشه‌برداری خصوصیات فیزیکی و شیمیایی خاک در مناطق نیمه‌خشک: تحلیل

## مقایسه‌ای الگوریتم‌های یادگیری ماشین

محمد راستین، دانشجوی دکتری، گروه علوم و مهندسی جنگل، دانشکده منابع طبیعی، دانشگاه گیلان، صومعه سرا

علی صالحی\*، دانشیار گروه علوم و مهندسی جنگل، دانشکده منابع طبیعی، دانشگاه گیلان، صومعه سرا

علی اصغر ذوالفقاری، دانشیار دانشکده کویرشناسی، دانشگاه سمنان، سمنان

نقیسه یغماییان، دانشیار گروه علوم مهندسی خاک، دانشکده علوم کشاورزی، دانشگاه گیلان، رشت

نوع مقاله: پژوهشی

تاریخچه مقاله (تاریخ دریافت: ۱۴۰۳/۱۰/۰۲ تاریخ پذیرش: ۱۴۰۴/۰۱/۱۰)

DOI: <http://doi.org/10.61186/jeer.15.2.82>

## چکیده

درک تغییرات مکانی ویژگی‌های خاک برای نظارت بر قابلیت‌های زمین و مدیریت پایدار منابع خاک بسیار حائز اهمیت است. این مطالعه با هدف ارزیابی دقت و کارایی مدل‌های یادگیری ماشین در برآورد خصوصیات فیزیکی و شیمیایی خاک انجام شده است. برای این منظور، تعداد ۱۱۰ نمونه خاک از دو عمق ۰-۳۰ سانتی‌متر و ۳۰-۶۰ سانتی‌متر جمع‌آوری و مورد بررسی قرار گرفت. مدل‌های یادگیری ماشین مورد استفاده شامل RF، Cubist، و RF-Cubist بودند و داده‌ها به دو بخش آموزش (۷۰٪) و آزمون (۳۰٪) تقسیم شدند. نتایج نشان داد که برای عمق ۰-۳۰ سانتی‌متر، مدل RF برای شن با RMSE، Bias، Cor و RRSE به ترتیب معادل ۱۰/۱۹، ۰/۳۰، ۰/۹۲ و ۰/۴۱ بهترین عملکرد را داشت. همچنین، مدل Cubist برای کربن آلی با RMSE، Bias، Cor و RRSE به ترتیب معادل ۲۳/۰۲، ۰/۷۱، ۰/۷۰ و ۰/۷۰ بهترین برآورد را داشته است. مدل RF-Cubist نیز در پیش‌بینی سیلت با RMSE، Bias، Cor و RRSE به ترتیب معادل ۷/۸۰، ۰/۸۷، ۰/۸۷ و ۰/۲۸ نتایج قابل قبولی به دست آورد. در عمق ۳۰-۶۰ سانتی‌متر، مدل RF برای شن با RMSE، Bias، Cor و RRSE به ترتیب معادل ۱۱/۴۲، ۰/۲۳، ۰/۸۸ و ۰/۴۷ بهترین عملکرد را داشت، و مدل Cubist برای کربن آلی با RMSE، Bias، Cor و RRSE به ترتیب معادل ۲۶/۰، ۰/۵۳، ۰/۸۵ و ۰/۵۳ بهترین بود. همچنین، مدل RF-Cubist در پیش‌بینی سیلت با RMSE، Bias، Cor و RRSE به ترتیب معادل ۹/۲۸، ۰/۵۸، ۰/۸۳ و ۰/۳۰ عملکرد خوبی داشت. نقشه‌های عدم قطعیت برای پارامترهای مختلف نیز نشان دادند که مدل RF در بسیاری از نقاط دارای عدم قطعیت کمتری نسبت به سایر مدل‌ها است. این پژوهش نشان می‌دهد که مدل‌های یادگیری ماشین، به‌ویژه RF-Cubist و Cubist، ابزارهای مؤثری برای پیش‌بینی خصوصیات خاک بوده و می‌توانند به مدیریت بهتر منابع خاک و کشاورزی پایدار کمک کنند.

واژگان کلیدی: یادگیری ماشین، جنگل تصادفی، Cubist، سنجش از دور، خصوصیات خاک، حوضه آبخیز

دامغان‌رود.

## ۱- مقدمه

در سال‌های اخیر با افزایش جمعیت، حفظ امنیت غذایی، تخریب سرزمین، کمبود منابع آب شیرین، تهدید تنوع زیستی، تغییرات اقلیمی و توسعه پایدار به عنوان چالش‌های اساسی در زمینه زیست محیطی شناخته می‌شوند (Smith, 2011 & Fao, 2022). خصوصیات فیزیکی و شیمیایی خاک ارتباط مستقیمی با این چالش‌ها داشته و اهمیت آنها برای تولید غذا، تنظیم و سازگاری با اقلیم، ذخیره کربن، فیلتراسیون آب و حفظ تنوع زیستی بسیار مشهود است (Jordan & Fröhle, 2022). اطلاعات دقیق از خصوصیات خاک در زمینه محیط زیست، کشاورزی و مدیریت منابع آب ابزار حیاتی محسوب می‌شوند (Schoen & Johnson, 2021). این خصوصیات تأثیر قابل توجهی بر عملکرد خاک، رشد گیاهان، جریان آب و حفظ تنوع زیستی دارند (Kibblewhite, et al, 2007). در کشاورزی، شناخت خصوصیات خاک امکان بهینه‌سازی کوددهی، آبیاری و انتخاب مناسب‌ترین گیاهان را فراهم می‌کند (Ayaz, et al, 2019). در حوضه‌های آبخیز، شناخت ویژگی‌های می‌تواند به پیش‌بینی رواناب، ذخیره آب، جریان زیرزمینی و تعیین مناطق حساس هیدرولوژیکی کمک کرده و در نتیجه، به مدیریت بهینه منابع آب و پیشگیری از سیلاب‌ها و خشکسالی‌ها کمک نماید (Cover, et al, 2022). همچنین اطلاعات دقیق از خصوصیات خاک می‌تواند به شناخت بهتر از رفتار مواد آلی خاک و مسایل آلودگی خاک و آب کمک کند و در نتیجه، به حفاظت از محیط زیست و حفظ تنوع زیستی کمک نماید (Tudi, et al, 2021). به همین دلیل، تأمین اطلاعات دقیق و جامع از خصوصیات فیزیکی و شیمیایی خاک برای تصمیم‌گیری‌های مؤثر در زمینه مدیریت محیطی و توسعه پایدار بسیار اهمیت دارد.

روش‌های تعیین خصوصیات فیزیکی و شیمیایی خاک از جمله مهمترین ابزارهایی هستند که در تحلیل و تفسیر ویژگی‌های خاک به کار می‌روند؛ بدین منظور از روش‌های مختلفی استفاده می‌شود که شامل روش‌های مستقیم و روش‌های غیرمستقیم است (Alves, et al., 2024 & Bek, et al., 2024). در مناطق با وسعت بالا و حوضه‌های آبخیز بزرگ که نیاز به طیف وسیعی از اطلاعات است، استفاده از روش‌های اندازه‌گیری مستقیم و آزمایشگاهی اغلب زمان‌بر بوده و نیاز به صرف هزینه بالایی دارد و به همین دلیل نیاز به استفاده از روش‌های غیر مستقیم و جدید را افزایش می‌یابد. ظهور روش‌های پیشرفته، الگوریتم‌های یادگیری ماشین و داده‌های با قدرت تفکیک بالا باعث ایجاد یک تغییر بزرگ در شیوه‌های ارزیابی خصوصیات خاک شده است که می‌تواند در مدیریت سلامت خاک نقش مهمی داشته باشد.

نقشه‌برداری رقومی خاک (DSM<sup>۱</sup>) یکی از روش‌های نوآورانه در این زمینه است که در چند دهه اخیر پیشرفت زیادی در زمینه برآورد خصوصیات خاک داشته است (McBratney, et al, 2003 & Malone, et al, 2009 & Khosravi et al, 2022 Minasny et al, 2016 & Vanmaercke, et al, 2021 & Cheshmberah, et al, 2022). این روش یک ابزار قدرتمند برای تحلیل و برآورد مکانی ویژگی‌های خاک در مقیاس‌های بزرگ است که الگوریتم‌های یادگیری ماشین و متغیرهای محیطی را ادغام کرده و نقشه‌های خصوصیات خاک را با کیفیت بالا تولید می‌کند (Vanmaercke, et al, 2021 & Cheshmberah, et al, 2022). در دهه گذشته مطالعات مختلفی با هدف استفاده از نقشه برداری رقومی به منظور برآورد خصوصیات خاک انجام شده است. پژوهشگران متعدد از روش نقشه برداری رقومی خاک برای برآورد خصوصیات مختلف خاک مانند خصوصیات ثانویه خاک (Reddy, et al, 2023)، آب در دسترس (AWC)،

1 Digital Soil Mapping

نقطه پژمرگی (PWP)، ظرفیت زراعی (Fc) (Zare, et al, 2021 & Somarathna, et al, 2021 & Toth, et al, 2017) (Fc) (Nauman and ) (Malone, et al, 2017 & Romano, et al, 2022 & Gomez, et al, 2023)، هدایت الکتریکی (Duniway, 2020 & Dharumarajan, 2022 & Nabiollahi, et al, 2021 & Khongnawang, et al, 2022)، کربن آلی خاک (Duarte, et al, 2022 & Dharumarajan, et al, 2021 & Yang, et al, 2022 & Novzari, et al, 2024)، تخریب‌پذیری اراضی (Zeraatpisheh, et al, 2020 & Farah, et al, 2021) و توزیع اندازه ذرات خاک (Keshavarzi, et al, 2022 & Zhang, et al, 2022 & Dharumarajan, et al, 2022 & Arshad, et al, 2020) استفاده کرده‌اند.

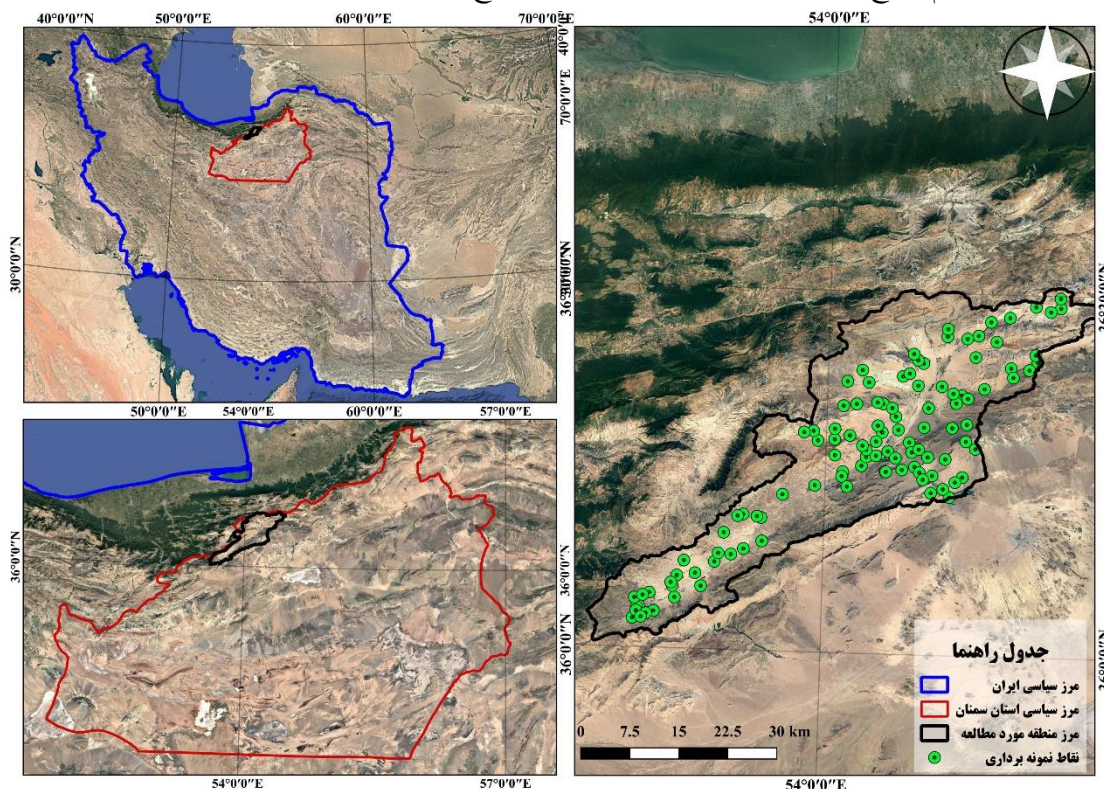
در نقشه‌برداری رقومی خاک، روش‌های مختلف مدل‌سازی با هدف برقراری ارتباط بین خصوصیات خاک و متغیرهای محیطی مانند داده‌های سنجش از دور توسط پژوهشگران متعدد مورد بررسی قرار گرفته‌اند. روش‌های یادگیری ماشین مانند روش جنگل تصادفی، رگرسیون خطی و رگرسیون‌های درختی برای پیش‌بینی خصوصیات خاک با استفاده از متغیرهای محیطی به کار رفته‌اند (Hounkpatin, et al, 2022 & Khaledian and Miller, 2020 & Wadoux, et al, 2020 & Taghizadeh-Mehrjardi, 2021 & Bao, et al, 2024). Suleymanov و همکاران (۲۰۲۱)، در مطالعه‌ای به بررسی نقشه‌برداری رقومی مواد مغذی خاک برای کشاورزی دقیق در منطقه استپی ترانس اورال روسیه با استفاده از ویژگی‌های توپوگرافی پرداختند. این تحقیق بر نقشه‌برداری دیجیتال خاک (DSM) و توزیع مکانی خصوصیات اصلی خاک‌های زراعی، شامل کربن آلی خاک (SOC)، نیتروژن، پتاسیم، کلسیم، منیزیم، سدیم، فسفر، pH و ضخامت افق هوموسی (AB) تمرکز دارد. از روش‌های رگرسیون خطی چندگانه (MLR) برای پیش‌بینی توزیع مکانی و تغییرات مواد مغذی خاک استفاده شده است. در این مطالعه، ۱۷ شاخص توپوگرافی محاسبه شده با مدل ارتفاعی دیجیتال SRTM به کار گرفته شد. در این پژوهش، ارتفاع، شیب و شاخص MMRTF (پهنای سطح قله چندرزلوشنی) مهم‌ترین متغیرها بودند. این روش می‌تواند برای مطالعه توزیع مکانی مواد مغذی خاک و نقشه‌برداری بزرگ‌مقیاس در مناطق مشابه استفاده شود. Mosleh و همکاران (۲۰۱۶)، در پژوهشی به بررسی توانایی روش‌های مختلف نقشه‌برداری رقومی خاک (DSM) در پیش‌بینی برخی از ویژگی‌های فیزیکی و شیمیایی خاک سطحی در دشت شهرکرد استان چهارمحال و بختیاری پرداختند، که در بین مدل‌های مورد مطالعه، GLM<sup>۱</sup> بالاترین عملکرد را برای پیش‌بینی pH، EC، رس، سیلت، شن و کربنات کلسیم معادل نشان داد. Mousavi و همکاران (۲۰۲۳)، با هدف نقشه‌برداری رقومی خصوصیات خاک، مطالعه‌ای در دشت مشهد به وسعت ۱۵۰۰ کیلومتر مربع انجام دادند. بدین منظور، با نمونه‌برداری از ۱۸۰ نقطه خاک از عمق ۰ تا ۱۰ سانتی‌متر و به کارگیری مدل جنگل تصادفی (RF) در ترکیب با کریجینگ ساده (OK) و کریجینگ رگرسیونی (RK) به بررسی ارتباط میان متغیرهای محیطی و ویژگی‌های خاک پرداختند. نتایج نشان داد که مدل RF-OK با دقت بالا، بهترین پیش‌بینی را برای خصوصیات خاک مانند سیلت، کربنات کلسیم، رس، کربن آلی، شن و pH ارائه داد و شاخص‌های سنجش از دور (RS) و برخی ویژگی‌های توپوگرافی نیز به عنوان مهم‌ترین متغیرها در مدل‌سازی نقش داشتند. Khosravi و همکاران (۲۰۲۲)، خصوصیات هیدرولیکی خاک را برای مدیریت بهتر حوضه‌های آبخیز با ادغام نقشه‌برداری رقومی خاک و توابع انتقالی مورد ارزیابی قرار دادند. نتایج آنها نشان داد که نقشه‌برداری رقومی خاک قادر به بررسی تغییرات خصوصیات هیدرولیکی حوضه آبخیز است.

<sup>1</sup> Generalized Linear Model

بررسی مطالعات پیشین نشان می‌دهد که نقشه‌برداری رقومی خاک به‌عنوان یک رویکرد نوین، دارای پتانسیل بالایی برای برآورد دقیق خصوصیات خاک است. در کنار این روش، به‌کارگیری تکنیک‌های یادگیری ماشین نیز دقت این برآوردها را به‌طور قابل توجهی افزایش می‌دهد. هدف این مطالعه ارزیابی اثر بخشی شاخص‌های پوشش گیاهی، تصاویر سنجش از دور و ویژگی‌های زمین در برآورد خصوصیات خاک در منطقه دامغان‌رود با استفاده از روش‌های یادگیری ماشین در چارچوب نقشه‌برداری رقومی خاک (DSM) است. این پژوهش با استفاده از مدل‌های پیشرفته‌ای نظیر روش‌های جنگل تصادفی (RF<sup>۱</sup>)، و کیوبیست (Cubist) به تحلیل و برآورد خصوصیات خاک در دو عمق ۰-۳۰ و ۳۰-۶۰ سانتی‌متر می‌پردازد. نتایج این تحقیق می‌تواند به بهبود مدیریت پایدار منابع خاک و اتخاذ تصمیمات علمی و کارآمد در حوضه‌های آبخیز، کشاورزی و حفاظت از منابع طبیعی کمک نماید.

## ۲- منطقه مورد مطالعه

منطقه مورد مطالعه حوضه آبخیز دامغان‌رود، واقع در شهرستان دامغان و در فاصله ۳۰ کیلومتری شمال شهر دامغان در استان سمنان است. مساحت این حوضه حدود ۱۳۰۰ کیلومترمربع است و در موقعیت عرض شمالی  $36^{\circ} 48' 22''$  تا  $36^{\circ} 21' 36''$  و طول شرقی  $48^{\circ} 21' 58''$  تا  $49^{\circ} 18' 36''$  واقع شده است (شکل-۱). بیشترین ارتفاع حوضه ۲۸۸۱ متر در کوه‌های جاران و کم‌ارتفاع‌ترین نقطه آن در محل دشت با ارتفاع ۱۰۵۰ متر است (Ebrahimi, 2016).



شکل ۱: موقعیت جغرافیایی حوضه آبخیز دامغان‌رود و موقعیت نقاط نمونه‌برداری در منطقه

این حوضه دارای ۳ زیرحوضه به نام‌های دیباج، کلاته رودبار و نمکه است و از نظر زمین‌شناسی در در واحد البرز جنوبی واقع شده است (Maghami moghim, et al, 2023). از نظر زمین‌شناسی، دارای چین‌های تاقدیسی و ناودیدی، گسل‌های راندگی و سنگ‌های رسوبی، آذرین و دگرگونی است. از مهم‌ترین سازندهای این منطقه می‌توان به سازند

<sup>1</sup> Random Forest

شمشک (شیل، ماسه سنگ، زغال سنگ)، سازند الیکا (آهک‌های کوره ای) و سازند پادگانه (آبرفتی قدیم) اشاره کرد (Jamshidi, 2010 & Jamshidi and Ghasemi, 2011 & Aghanbati, 2004 & Rahimpour and Ismaili, 2007 &). این حوضه بر اساس طبقه‌بندی اقلیمی کوپن دارای آب و هوای نیمه خشک است. متوسط دما و بارندگی سالانه به ترتیب ۲۳/۵۴ درجه سانتی گراد و ۱۵۰ میلی‌متر است (Maghami moghim, et al, 2023). از نظر کاربری اراضی حدود ۴۱٪ از مساحت مورد استفاده شامل زمین‌های قابل کشت است. در مجموع مساحت مراتع منطقه، ۶۶٪ به عنوان مراتع فقیر با تراکم کم طبقه‌بندی می‌شوند، در حالی که ۳۱٪ نشان دهنده مراتع متوسط، و ۲٪ را مراتع متراکم را تشکیل می‌دهد (Rezaei Moghaddam, et al, 2022).

### ۳- مواد و روش

#### ۳-۱- نمونه برداری

تعیین نقاط نمونه‌برداری در این مطالعه با استفاده از روش نمونه‌برداری مکعب لاتین (cLHS) انجام شد. این روش با استفاده از بررسی تغییرات توزیع تجمعی متغیرهای کمکی، نقاط نمونه برداری را تعیین می‌کند. در این مطالعه متغیرهای کمکی در برنامه مکعب لاتین، باندها و شاخص‌های تصویر ماهواره‌ای سنتینل-۲ شامل طیف باندهای مرئی، مادون قرمز کوتاه، شاخص‌های پوشش گیاهی و خصوصیات مستخرج از مدل رقومی ارتفاع از جمله نقشه رقومی ارتفاع، شیب، جهت و شاخص همواری دره با درجه تفکیک بالا بودند. با استفاده از روش cLHS تعداد ۱۱۰ نقطه نمونه‌برداری انتخاب و نمونه‌برداری از عمق ۳۰-۶۰ سانتی‌متر و ۳۰-۶۰ سانتی‌متر انجام شد (شکل-۱).

#### ۳-۲- آنالیزهای آزمایشگاهی

ذرات شن، سیلت و رس با استفاده از روش هیدرومتری سه قرائته انجام شد (Gee and Bauder, 1986). کربن آلی با استفاده از روش والکی بلک اندازه‌گیری شد (Walkey & Black, 1934). هدایت الکتریکی خاک (EC) با استفاده از EC متر و با عصاره‌گیری از گل اشباع انجام شد.

#### ۳-۳- نقشه‌برداری رقومی خصوصیات خاک (DSM)

برای تهیه نقشه خصوصیات خاک از روش نقشه‌برداری رقومی خاک (DSM) استفاده شد که روشی مبتنی بر داده‌کاوی و سنجش از دور است و به طور گسترده برای مدل‌سازی و پیش‌بینی خصوصیات خاک به کار می‌رود. این روش با بهره‌گیری از متغیرهای محیطی مانند توپوگرافی و کاربری اراضی، الگوهای مکانی خاک را پیش‌بینی می‌کند.

#### ۳-۴- متغیرهای محیطی

در مطالعه حاضر براساس مفاهیم روش اسکورپن، از داده‌های سنجش از دور (RS)، مدل رقومی ارتفاع و متغیرهای مستخرج از آن به عنوان متغیرهای کمکی (جدول-۱) در برآورد خصوصیات خاک استفاده شد. داده‌های سنجش از دور از تصاویر و شاخص‌های سنتینل-۲ مربوط به ماه‌هایی که منطقه دارای پوشش گیاهی است (فروردین، اردیبهشت و خرداد) در بازه زمانی ۲۰۱۹ تا ۲۰۲۳ استخراج گردید. اطلاعات یک تصویر در یک تاریخ مشخص از منطقه ممکن است تحت تأثیر خصوصیات دینامیک مانند بارش طی چند روز گذشته، دما، پوشش گیاهی و رطوبت خاک قرار بگیرد، لذا در برآورد دقیق‌تر خصوصیات خاک بهتر است از میانه یا میانگین تصاویر در بازه‌های زمانی مشخص و در یک دوره متوالی از منطقه برای یافتن ارتباط بین خصوصیات خاک و تصاویر استفاده کرد. در این مطالعه برای یافتن این ارتباط از

میان‌ه تصاویر به‌عنوان متغیرهای کمکی استفاده شد. پردازش و استخراج تصاویر در پلتفرم گوگل ارث انجام و میان‌ه تصاویر دانلود شده و مورد استفاده قرار گرفت. متغیرهای زمین با استفاده از نقشه رقومی ارتفاع با درجه وضوح ۱۵ متر برای منطقه مورد مطالعه استخراج شدند. با توجه به اینکه تصاویر داده‌های سنتیل ۲ دارای درجه وضوح ۱۰ و ۲۰ متر هستند، لذا در این مطالعه تمامی تصاویر به درجه وضوح ۲۰ متر تبدیل شدند و پردازش داده‌ها برای نقشه برداری رقومی خاک انجام شد.

جدول ۱: متغیرهای محیطی استفاده شده در مطالعه حاضر

| نوع شاخص                                             | اختصار              | نام باند/فرمول                                |
|------------------------------------------------------|---------------------|-----------------------------------------------|
| مقدار بازتاب دهنده‌گی در ریز گرد                     | Coastal and aerosol | Band 1                                        |
| مقدار بازتاب دهنده‌گی باند ۲ (آبی مرئی)              | Blue                | Band 2                                        |
| مقدار بازتاب دهنده‌گی باند ۳ (سبز مرئی)              | Green               | Band 3                                        |
| مقدار بازتاب دهنده‌گی باند ۴ (قرمز مرئی)             | Red                 | Band 4                                        |
| بازتاب پوشش گیاهی در محدوده مادون قرمز نزدیک         | Vegetation red edge | Band 5                                        |
| بازتاب پوشش گیاهی در محدوده مادون قرمز نزدیک         | Vegetation red edge | Band 6                                        |
| بازتاب پوشش گیاهی در محدوده مادون قرمز نزدیک         | Vegetation red edge | Band 7                                        |
| مقدار بازتاب دهنده‌گی در محدوده مادون قرمز           | NIR                 | Band 8                                        |
| مقدار بازتاب دهنده‌گی در محدوده مادون قرمز نزدیک     | Narrow NIR          | Band 8A                                       |
| مقدار بازتاب دهنده‌گی در بخار آب                     | Water vapor         | Band 9                                        |
| مقدار بازتاب در شرایط ابری                           | SWIR - Cirrus       | Band 10                                       |
| مقدار بازتاب برای اندازه‌گیری رطوبت خاک و پوشش گیاهی | SWIR                | Band 11                                       |
| مقدار بازتاب برای اندازه‌گیری رطوبت خاک و پوشش گیاهی | SWIR                | Band 12                                       |
| شاخص‌های طیفی                                        | شاخص‌های طیفی       | شاخص‌های طیفی                                 |
| شاخص پوشش گیاهی نرمال شده                            | (NDVI)              | $(B08-B04)/(B08+B04)$                         |
| شاخص نسبت پوشش گیاهی                                 | (RVI)               | $(B04/B08)$                                   |
| شاخص رس                                              | (CI)                | -                                             |
| شاخص کربنات                                          | (CrI)               | -                                             |
| شاخص روشنایی                                         | (BI)                | -                                             |
| شاخص شدت                                             | (IN)                | -                                             |
| شاخص شوری نرمال شده                                  | (NDSI)              | $(B3-B11)/(B3+B11)$                           |
| شاخص شوری                                            | (SI)                | -                                             |
| شاخص گیاهی تعدیل کننده اثر خاک                       | (SAVI)              | $[(B08 - B04) / (B08 + B04 + 0.5)] / (0.5+1)$ |

|                        |                        |                               |
|------------------------|------------------------|-------------------------------|
| شاخص‌های مستخرج از DEM | شاخص‌های مستخرج از DEM | شاخص‌های مستخرج از DEM        |
| شاخص‌های مستخرج از DEM | (AS)                   | جهت                           |
| شاخص‌های مستخرج از DEM | (CA)                   | حوضه آبریز                    |
| شاخص‌های مستخرج از DEM | (CI)                   | شاخص همگرایی                  |
| شاخص‌های مستخرج از DEM | (DEM)                  | نقشه رقومی ارتفاع             |
| شاخص‌های مستخرج از DEM | (FA)                   | جریان تجمعی                   |
| شاخص‌های مستخرج از DEM | (S.LF)                 | طول شیب                       |
| شاخص‌های مستخرج از DEM | (MSP)                  | شیب میانی حوضه                |
| شاخص‌های مستخرج از DEM | (MRVBF)                | همواری دره با درجه تفکیک بالا |
| شاخص‌های مستخرج از DEM | (PLC)                  | انحنای خطوط توپوگرافی         |
| شاخص‌های مستخرج از DEM | (RSP)                  | موقعیت نسبی شیب               |
| شاخص‌های مستخرج از DEM | (SL)                   | شیب                           |
| شاخص‌های مستخرج از DEM | (TWI)                  | خیسی توپوگرافی                |
| شاخص‌های مستخرج از DEM | (VD)                   | عمق دره                       |
| شاخص‌های مستخرج از DEM | (WE)                   | اثر باد                       |

### ۵-۳- انتخاب مؤثرترین متغیرهای محیطی

یکی از روش‌های مؤثر برای کاهش تعداد متغیرهای ورودی مدل، استفاده از روش انتخاب ویژگی‌ها است. در این روش، با استفاده از الگوریتم‌های مختلف، ویژگی‌های مهم و تأثیرگذارتر بر خروجی مدل، شناسایی و انتخاب شده و ویژگی‌های غیرضروری حذف می‌شوند (Sharma et al, 2022). این امر باعث کاهش پیچیدگی مدل، افزایش دقت و عملکرد مدل و حتی افزایش توجه به ویژگی‌های مهم‌تر می‌شود. روش‌های مختلفی برای انتخاب ویژگی‌های تأثیرگذار وجود دارد که از جمله آن‌ها می‌توان به بروتا<sup>۱</sup>، روش درختی، روش‌های مبتنی بر فیلتر، روش‌های مبتنی بر دسته‌بندی و ... اشاره کرد (Cheng, 2024). در این مطالعه از روش بروتا برای انتخاب متغیرهای تأثیرگذار استفاده شد. الگوریتم Boruta یک الگوریتم پوششی است که برای انتخاب ویژگی‌های قوی و مرتبط از نظر آماری استفاده می‌شود. این الگوریتم بر اساس ویژگی‌های HCR و DLR، از الگوریتم Boruta مبتنی بر جنگل تصادفی (RF) استفاده می‌کند. الگوریتم Boruta به کاربر این امکان را می‌دهد تا مهم‌ترین ویژگی‌های یک سیستم اطلاعاتی را بر اساس معیارهای بی‌طرفانه انتخاب کند و از روش‌های دستی برای انتخاب ویژگی‌ها بدون تعصب نسبت به ویژگی‌های مهم و دیگر استفاده می‌شود (Huda, et al, 2017, Nahar, et al, 2021, Haubold, et al, 2021).

### ۶-۳- مدل‌سازی و برآورد خصوصیات خاک

#### ۶-۳-۱- مدل جنگل تصادفی (RF)

در این مطالعه از روش جنگل تصادفی برای مدل‌سازی و برآورد خصوصیات خاک با استفاده از متغیرهای محیطی استفاده شده است. جنگل تصادفی (RF) الگوریتمی یادگیری ماشین ناپارامتری است که از مجموعه‌ای از درختان

1 Bruta

تصمیم‌گیری برای دسته‌بندی، رگرسیون و دیگر وظایف یادگیری ماشین استفاده می‌کند. این الگوریتم توسط بریمن<sup>۱</sup> (۲۰۰۱) معرفی شد و به دلیل دقت و کارایی بالا در طیف گسترده‌ای از مسائل، به یکی از محبوب‌ترین روش‌های یادگیری ماشین تبدیل شده است (Khosravani, et al, 2023). این الگوریتم براساس طبقه‌بندی درختی عمل می‌کند. هر درخت براساس زیرمجموعه‌ای تصادفی از داده‌ها آموزش داده می‌شود و می‌توان گفت که این مدل از همه داده‌های موجود برای بررسی دقت مدل استفاده می‌کند. در این مطالعه از الگوریتم مدل جنگل تصادفی آموزش دیده با متغیرهای محیطی، برای برآورد خصوصیات خاک استفاده شد. ابتدا داده‌ها به سه گروه تقسیم گردید و سپس در هر بار اجرای مدل یک گروه (حدود ۳۰٪ داده‌ها) به عنوان داده آزمون و دو گروه دیگر داده‌ها (حدود ۷۰٪ داده‌ها) برای آموزش استفاده شد؛ بنابراین، براساس روش مذکور تمامی داده‌ها به عنوان داده آزمون مورد ارزیابی قرار گرفتند. منابع متعدد نشان داده‌اند که الگوریتم RF با دقت بالا و توانایی مدیریت داده‌های پیچیده و حجیم، یکی از بهترین گزینه‌ها برای پیش‌بینی خصوصیات مختلف خاک است (Khosravani, et al, 2023 & Zhou, et al, 2019).

### ۳-۶-۲- مدل کوئیست Cubist

Cubist، الگوریتمی پیشرفته مبتنی بر درخت رگرسیون ناپارامتریک، به منظور آشکارسازی و مدل‌سازی روابط غیرخطی میان متغیرها، به ویژه روابطی که ماهیت پیچیده‌ای دارند، طراحی شده است. این الگوریتم که در مقاله Malone و همکاران (۲۰۱۷)، معرفی شده، از ترکیب قدرت نقشه‌برداری خود سازمان‌دهنده (SOM) و مدل‌های درخت تصمیم‌گیری برای خلق مدل‌های پیش‌بینی دقیق و قابل تفسیر بهره می‌برد. Cubist به‌طور خاص برای شناسایی و مدل‌سازی روابط پیچیده بین متغیرها، حتی در مواردی که ماهیت غیرخطی دارند، طراحی شده است. این الگوریتم در طیف وسیعی از وظایف مدل‌سازی، دقت پیش‌بینی قابل توجهی دارد. این امر به دلیل توانایی آن در مدل‌سازی پیچیدگی‌های روابط بین متغیرها حاصل می‌شود. مدل در دو مرحله اصلی به شرح زیر عمل می‌کند. در مرحله اول، Cubist از SOM برای سازماندهی داده‌ها در قالب خوشه‌هایی مجزا استفاده می‌کند. این خوشه‌ها به طور خودکار بر اساس شباهت بین نقاط داده شکل می‌گیرند و به Cubist در شناسایی روابط غیرخطی بین متغیرها کمک می‌کنند. در مرحله بعد، Cubist برای هر خوشه SOM یک درخت تصمیم‌گیری مجزا ایجاد می‌کند. این درختان تصمیم‌گیری برای مدل‌سازی رابطه بین متغیرهای پیش‌بینی‌کننده و متغیر هدف در هر خوشه به کار گرفته می‌شوند (Cover and Hart, 1967 & Kuhn and Johnson, 1992 & Quinlan, 2013).

### ۳-۶-۳- مدل RF-Cubist

مدل RF-Cubist ترکیبی از دو روش یادگیری ماشین جنگل تصادفی (Random Forest) و Cubist است که در مطالعات مختلف به عنوان یک ابزار قدرتمند برای پیش‌بینی خصوصیات خاک شناخته شده است. جنگل تصادفی با استفاده از تجمیع چندین درخت تصمیم‌گیری، دقت پیش‌بینی را افزایش می‌دهد و از بروز بیش‌برازش جلوگیری می‌کند. از سوی دیگر، مدل Cubist با ترکیب مدل‌های خطی محلی و الگوریتم‌های یادگیری ماشینی قادر به مدیریت داده‌های پیچیده و غیرخطی است. ترکیب این دو روش، دقت و پایداری بالاتری را در برآورد خصوصیات خاک فراهم می‌کند. نتایج مطالعات نشان می‌دهد که RF-Cubist می‌تواند با دقت بالا و عدم قطعیت پایین، خصوصیات مختلف خاک را برآورد

1 Berriman

کند (Taghizadeh-Mehrjardi, et al, 2020 & Chen, et al, 2019) این ویژگی‌ها مدل RF-Cubist را به ابزاری مناسب برای مدیریت پایدار منابع خاک تبدیل می‌کند.

#### ۳-۶-۴- ارزیابی عملکرد مدل‌ها

ارزیابی عملکرد مدل یک بخش مهم از مدل‌سازی است. در این مطالعه برای هر کدام از خصوصیات خاک از آماره‌های ضریب تبیین ( $R^2$ ) و ریشه دوم میانگین مربعات خطا (RMSE)، اریب (Bias)، ضریب همبستگی (Cor) و جذر میانگین مربعات خطاهای نسبی (RRSE) استفاده گردید. RMSE به عنوان یک معیار عمومی برای اندازه‌گیری دقت مدل، اریب نشان‌دهنده میزان بیش‌برآورد یا کم‌برآورد مدل، Cor میزان وابستگی خطی بین مقادیر واقعی و پیش‌بینی شده و RRSE خطای مدل را نسبت به مدل ساده‌ای که فقط از میانگین مقادیر واقعی استفاده می‌کند، ارزیابی می‌کند. که روابط ریاضی آن‌ها به شکل زیر است:

رابطه ۱

$$R^2 = \frac{[\sum_{i=1}^n (O_i - \bar{O}_i)(P_i - \bar{P}_i)]^2}{\sum_{i=1}^n (O_i - \bar{O}_i)^2 \sum_{i=1}^n (P_i - \bar{P}_i)^2}$$

رابطه ۲

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (P_i - O_i)^2}{n}}$$

رابطه ۳

$$Bias = \sum_{i=1}^n \frac{1}{n} (P_i - O_i)^2$$

رابطه ۴

$$Cor = \frac{\sum_{i=1}^n (\sum (P_i O_i) - (P_i)(O_i))}{\sqrt{[n \sum O_i^2 - (O_i^2)][n \sum P_i^2 - (P_i^2)]}}$$

رابطه ۵

$$RRSE = \sqrt{\frac{\sum_{i=1}^n (O_i - P_i)^2}{(O_i - \bar{O})}}$$

در این روابط،  $P_i$  و  $O_i$  به ترتیب نشان‌دهنده مقادیر مشاهداتی و پیش‌بینی شده،  $n$  تعداد داده‌ها و  $\bar{O}$  میانگین مقادیر مشاهداتی است. RMSE معمولاً عددی مثبت است و مقدار آن می‌تواند از ۰ تا بی‌نهایت تغییر کند. مقدار ۰ نشان‌دهنده دقت کامل مدل است. مقادیر مثبت اریب به معنای بیش‌برآورد و مقادیر منفی آن به معنای کم‌برآورد مدل است. اریب نزدیک به صفر نشان‌دهنده دقت بالای مدل در متوسط پیش‌بینی‌ها است. Cor نزدیک به ۱ نشان‌دهنده همبستگی مثبت قوی و Cor نزدیک به -۱ نشان‌دهنده همبستگی منفی قوی است. Cor نزدیک به صفر نشان‌دهنده همبستگی ضعیف است. مقادیر کمتر RRSE نشان‌دهنده عملکرد بهتر مدل نسبت به مدل پایه است. مقدار RRSE می‌تواند از ۰ تا بی‌نهایت تغییر کند، و مقدار ۱ نشان‌دهنده عملکرد برابر مدل با مدل پایه است.

#### ۴- یافته‌ها

##### ۴-۱- خلاصه آماری خصوصیات خاک

خلاصه آماری خصوصیات خاک در دو عمق ۰-۳۰ و ۳۰-۶۰ سانتی متری خاک در جدول ۲ ارائه شده است. مقادیر حداقل و حداکثر ذرات سیلت، رس و شن در عمق ۰-۳۰ سانتی متر به ترتیب برابر با ۱-۶۰/۹۹، ۱-۵۲/۸۲ و ۸۸-۱ درصد و در عمق ۳۰-۶۰ سانتی متر نیز برابر با ۲-۷۱/۷، ۳۱-۵۰/۹۶ و ۱-۸۱/۶ درصد و میانگین آن‌ها نیز به ترتیب در عمق ۰-۳۰ معادل ۲۸/۱۷، ۱۳/۳۶ و ۴۵/۹۸ درصد و عمق ۳۰-۶۰ معادل ۳۰/۳۵، ۱۴/۳۳ و ۴۱/۶۸ درصد بدست آمد. مقادیر کربن آلی خاک در منطقه از ۰/۰۰۱ تا ۱/۶۷ درصد در عمق ۰-۳۰ سانتی متر و صفر تا ۱/۶۳ درصد در عمق ۳۰-۶۰ سانتی متری متغیر بود. میانگین این پارامتر در عمق‌های ۰-۳۰ و ۳۰-۶۰ سانتی متری به ترتیب معادل ۰/۳۴ و ۰/۳۱ درصد تعیین شد که نشان‌دهنده مقادیر کم کربن آلی در خاک‌های منطقه است. هدایت الکتریکی خاک به ترتیب در دو عمق ۰-۳۰ و ۳۰-۶۰ سانتی متر بین ۰/۳۹ تا ۷۴/۴ و ۰/۰۹ تا ۵۶/۷ دسی زیمنس بر متر متغیر بوده است و مقدار آن به طور میانگین در سطح حوضه ۳/۱۶ و ۲/۵۵ دسی زیمنس بر متر است که نشان دهنده شوری کم خاک در سطح منطقه و مدیریت جهت جلوگیری از شوری شدید خاک است. البته برخی از مناطق منطقه دارای خاک های مارنی بسیار شور که شوری در بسیاری از خاک ها تا حدود ۷۴ دسی زیمنس بر متر نیز است.

جدول ۲- آمار توصیفی داده‌های اندازه‌گیری شده در ۱۱۰ نقطه نمونه برداری

| عمق (cm) | ویژگی (واحد)          | حداقل (%) | حداکثر (%) | میانگین (%) | انحراف معیار (%) | ضریب تغییرات |
|----------|-----------------------|-----------|------------|-------------|------------------|--------------|
| ۰-۳۰     | شن (%)                | ۱         | ۸۸         | ۴۵/۹۸       | ۲۴/۶۲            | ۰/۵۳         |
|          | رس (%)                | ۱         | ۵۲/۸۲      | ۱۳/۳۶       | ۱۰/۴۶            | ۰/۷۸         |
|          | سیلت (%)              | ۱         | ۶۰/۹۹      | ۲۸/۱۷       | ۱۴/۷۰            | ۰/۵۲         |
|          | کربن آلی (%)          | ۰/۰       | ۱/۶۷       | ۰/۳۶        | ۰/۳۴             | ۰/۹۱         |
|          | هدایت الکتریکی (dS/m) | ۰/۳۹      | ۷۴/۴       | ۳/۱۶        | ۹/۷۹             | ۳/۰۹         |
| ۳۰-۶۰    | شن (%)                | ۱         | ۸۱/۶       | ۴۱/۶۸       | ۲۳/۹۴            | ۰/۵۷         |
|          | رس (%)                | ۰/۳۱      | ۵۰/۹۶      | ۱۴/۳۳       | ۱۲/۱۱            | ۰/۸۵         |
|          | سیلت (%)              | ۲         | ۷۱/۷       | ۳۰/۳۵       | ۱۶/۶۵            | ۰/۵۴         |
|          | کربن آلی (%)          | ۰/۰۰۱     | ۱/۶۳       | ۰/۳۳        | ۰/۳۱             | ۰/۹۳         |
|          | هدایت الکتریکی (dS/m) | ۰/۰۹      | ۵۶/۷       | ۲/۵۵        | ۶/۴۸             | ۲/۵۳         |

Carter و Gregorich (۲۰۰۷)، بیان کردند که ضریب تغییرات بالای ۰/۶۴ نشان‌دهنده تنوع بالا در خصوصیات خاک است. با در نظر گرفتن این موضوع، کلاس ضریب تغییرات اجزای بافت خاک نشان دهنده تغییرات بالای رس (۰/۷۸) و ۰/۸۵ و تغییرات متوسط شن (۰/۵۳ و ۰/۵۷) و سیلت (۰/۵۲ و ۰/۵۴) به ترتیب در هر دو عمق ۰-۳۰ و ۳۰-۶۰ است. دو ویژگی کربن آلی و هدایت الکتریکی خاک نیز به ترتیب با ضریب تغییرات ۰/۹۱ و ۳/۰۹ و ۰/۹۳ و ۲/۵۳ به ترتیب در هر دو عمق ۰-۳۰ و ۳۰-۶۰ سانتی متر، نشان‌دهنده تغییرپذیری بسیار بالایی در منطقه می‌باشند. (Khosravi et al, 2022) تغییرپذیری بالای ذرات بافت خاک را متأثر از عوامل محیطی و دلیل بالا بودن تغییرپذیری رس نسبت به پارامترهای دیگر را وجود سازندهای رسی در منطقه دانسته است. Mousavi و همکاران (۲۰۲۳)، نیز تنوعپذیری خصوصیات را تحت تأثیر تنوع مواد اولیه مادری و فرآیندهای ژئومورفولوژیکی دانستند که منجر به تشکیل لندفرم

های مختلف با اجزای مختلف می‌شود. میانگین کربن آلی خاک در دو عمق سطحی و زیر سطحی به ترتیب معادل ۰/۳۶ و ۰/۳۳ است که این مقدار نشان‌دهنده مقادیر پایین کربن آلی در خاک است و براساس مطالعه Mousavi و همکاران (۲۰۲۳)، این امر متأثر از پوشش گیاهی ضعیف، استفاده کم از کودهای آلی و کاهش حاصلخیزی خاک است.

#### ۲-۴- انتخاب متغیرهای محیطی

مناسب‌ترین متغیرهای محیطی با استفاده از رویکرد انتخاب متغیر بروتا انتخاب شدند. جدول ۳ مهم‌ترین متغیرهای مؤثر در تغییرات هریک از پارامترهای خاک را نشان می‌دهد. داده‌های سنجش از دور در سطوح محلی و جهانی در مطالعات نقشه‌برداری رقومی خاک نقش مؤثری داشته‌اند (Hengl, et al, 2017). استفاده از مدل بروتا برای انتخاب متغیرهای محیطی نسبت به سایر مدل‌های انتخاب متغیر منجر به کاهش خطا در پیش‌بینی خصوصیات خاک است (Chen et al, 2022). متغیرهای مستخرج از مدل رقومی ارتفاع به‌عنوان متغیرهای محیطی مناسب در مدل‌سازی کلاس و خصوصیات خاک می‌باشند (Mousavi et al, 2023).

جدول ۳- متغیرهای محیطی تأثیرگذار در مدل‌سازی هریک از پارامترها

| متغیرهای محیطی | پارامتر |       |       |      |       |
|----------------|---------|-------|-------|------|-------|
|                | Clay    | Sand  | Silt  | EC   | OC    |
| NDVI           |         | B2    | B11   | B2   | B4    |
| SAVI           |         | B3    | B12   | B3   | NDVI  |
| NDSI           |         | B4    | NDVI  | B4   | SAV   |
| CI             |         | B11   | SI4   | B5   | NDSI  |
| CrI            |         | B12   | SAV   | B6   | SAVI  |
| RVI            |         | NDVI  | NDSI  | B7   | CrI   |
|                |         | SAV   | SAVI  | B8   | RVI   |
| FA             |         | NDSI  | CI    | B8A  |       |
|                |         | SAVI  | CrI   | NDVI | DEM   |
| DEM            |         | CI    | RVI   | SI4  |       |
|                |         | IN    | DEM   | SAV  | LS    |
| LS             |         | CrI   | LS    | NDSI | MRVBF |
|                |         | RVI   |       | VSSI |       |
| MRVBF          |         | DEM   | MRVBF | SAVI | RSP   |
|                |         | LSF   |       | IN   |       |
| RSP            |         | MRVBF | RSP   | BI   | TWI   |
|                |         | PLC   |       |      |       |
| TWI            |         | RSP   | TWI   | CrI  | WE    |
|                |         | TWI   |       |      |       |

راهنما: NDVI: شاخص پوشش گیاهی نرمال‌شده، SAVI: شاخص گیاهی با تعدیل اثر خاک، NDSI: شاخص شوری نرمال‌شده، CI: شاخص رس، CrI: شاخص کربنات، RVI: شاخص نسبت پوشش گیاهی، IN: شاخص شدت، SI4: شاخص شوری، BI: شاخص روشنایی، VSSI: شاخص تغییرات روشنایی و شوری، B2: بازتاب باند آبی مرئی، B3: بازتاب باند سبز مرئی، B4: بازتاب باند قرمز مرئی، B5, B6, B7: بازتاب مادون قرمز نزدیک، B8, B8A: بازتاب مادون قرمز، B11, B12: بازتاب مادون قرمز میانه.

#### ۲-۴- عملکرد مدل‌ها

براساس نتایج آماری (جدول-۴) عملکرد مدل‌های (Random Forest (RF، Cubist و RF-Cubist در پیش‌بینی خصوصیات خاک شامل شن (Sand)، سیلت (Silt)، رس (Clay)، هدایت الکتریکی (EC) و کربن آلی (OC) در دو عمق ۰-۳۰ و ۳۰-۶۰ سانتی‌متر، با استفاده از آماره‌های RMSE، RRSE، Cor و اریب، مورد بررسی قرار گرفت.

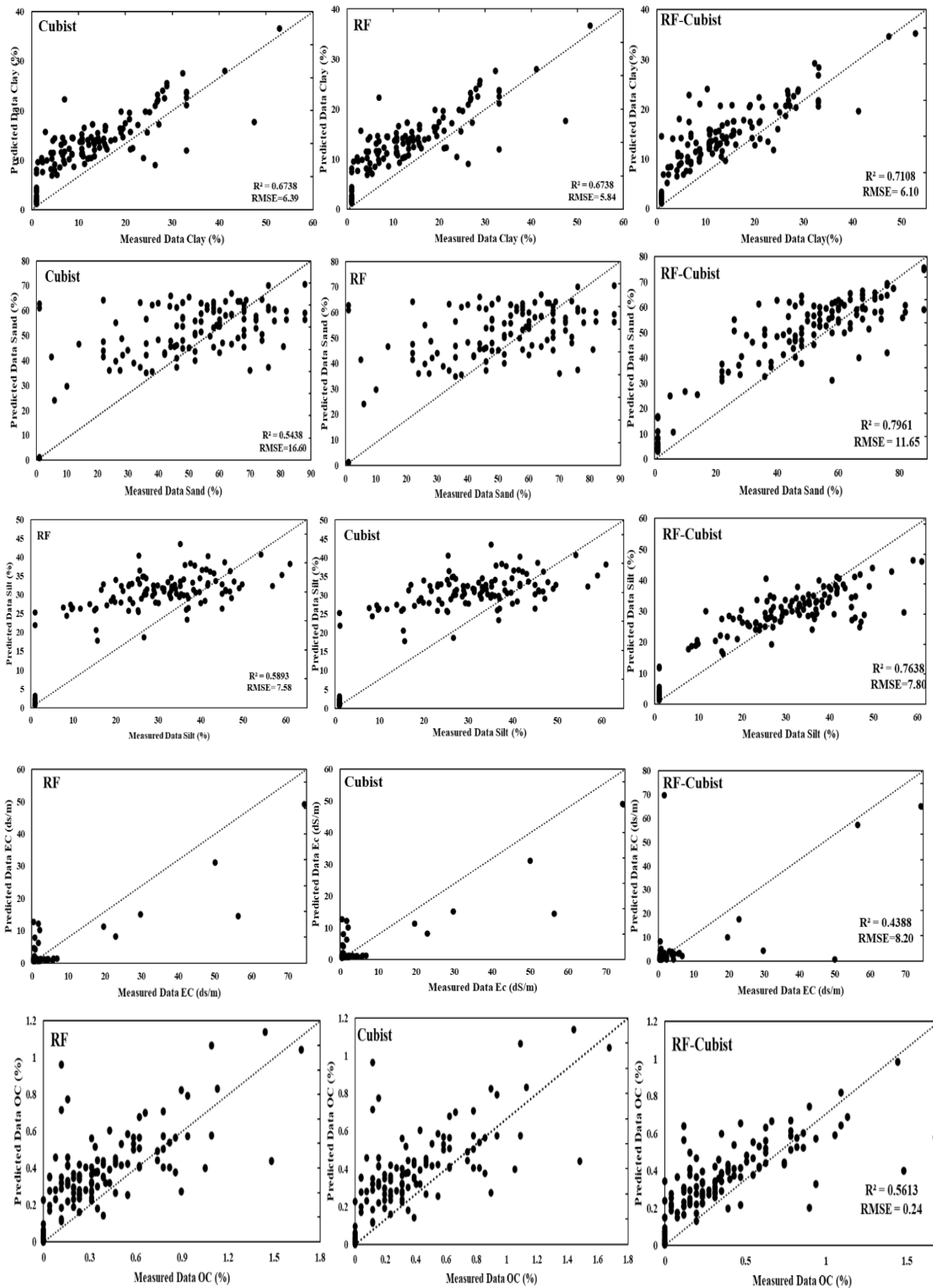
جدول ۴- نتایج آماری مدل‌های استفاده شده در برآورد خصوصیات خاک در دو عمق ۰-۳۰ و ۳۰-۶۰ سانتی‌متر

| عمق (cm) | مدل       | خصوصیات خاک | RRSE | Cor  | Bias  | RMSE  |
|----------|-----------|-------------|------|------|-------|-------|
| ۰-۳۰     | Cubist    | Sand        | ۰/۶۷ | ۰/۷۳ | -۰/۵۶ | ۱۶/۶۰ |
|          |           | Silt        | ۰/۶۴ | ۰/۷۶ | ۰/۲۲  | ۹/۶۹  |
|          |           | Clay        | ۰/۶۱ | ۰/۸۲ | ۰/۱۹  | ۶/۳۹  |
|          |           | EC          | ۰/۵۶ | ۰/۸۸ | ۰/۸۴  | ۵/۵۰  |
|          |           | OC          | ۰/۷۰ | ۰/۷۱ | ۰/۰۰۵ | ۰/۲۳  |
|          | RF        | Sand        | ۰/۴۱ | ۰/۹۲ | -۰/۳۰ | ۱۰/۱۹ |
|          |           | Silt        | ۰/۵۰ | ۰/۸۸ | ۰/۳۷  | ۷/۵۸  |
|          |           | Clay        | ۰/۵۶ | ۰/۸۵ | -۰/۰۳ | ۵/۸۴  |
|          |           | EC          | ۰/۵۰ | ۰/۹۲ | -۰/۴۹ | ۴/۹۰  |
|          |           | OC          | ۰/۵۹ | ۰/۸۱ | -۰/۰۲ | ۰/۲۰  |
| ۳۰-۶۰    | RF-Cubist | Sand        | ۰/۲۵ | ۰/۸۹ | ۰/۰۰۳ | ۱۱/۶۵ |
|          |           | Silt        | ۰/۲۸ | ۰/۸۷ | -۰/۸۷ | ۷/۸۰  |
|          |           | Clay        | ۰/۴۵ | ۰/۸۴ | ۰/۴۵  | ۶/۱۰  |
|          |           | EC          | ۲/۵۹ | ۰/۶۶ | -۰/۰۷ | ۸/۲۰  |
|          |           | OC          | ۰/۶۵ | ۰/۷۴ | -۰/۰۲ | ۰/۲۴  |
|          | Cubist    | Sand        | ۰/۵۳ | ۰/۸۴ | -۱/۰۴ | ۱۲/۸۷ |
|          |           | Silt        | ۰/۶۸ | ۰/۷۳ | ۰/۲۷  | ۱۱/۲۹ |
|          |           | Clay        | ۰/۶۲ | ۰/۷۸ | -۰/۱۵ | ۷/۵۴  |
|          |           | EC          | ۰/۶۲ | ۰/۷۹ | ۰/۷۳  | ۴/۰۱  |
|          |           | OC          | ۰/۸۵ | ۰/۵۳ | ۰/۰۵  | ۰/۲۶  |
| ۳۰-۶۰    | RF        | Sand        | ۰/۴۷ | ۰/۸۸ | -۰/۲۳ | ۱۱/۴۲ |
|          |           | Silt        | ۰/۵۴ | ۰/۸۴ | ۰/۸۳  | ۸/۹۹  |
|          |           | Clay        | ۰/۵۹ | ۰/۸۳ | -۰/۷۳ | ۷/۱۸  |
|          |           | EC          | ۰/۶۳ | ۰/۸۱ | -۰/۳۵ | ۴/۰۸  |
|          |           | OC          | ۰/۷۰ | ۰/۷۳ | ۰/۰۲  | ۰/۲۲  |
|          | RF-Cubist | Sand        | ۰/۲۹ | ۰/۸۶ | ۰/۱۴  | ۱۲/۱۵ |
|          |           | Silt        | ۰/۳۰ | ۰/۸۳ | -۰/۵۸ | ۹/۲۸  |
|          |           | Clay        | ۰/۴۷ | ۰/۸۵ | ۰/۲۰  | ۶/۸۴  |
|          |           | EC          | ۱/۸۰ | ۰/۷۶ | -۰/۵۵ | ۴/۶۱  |
|          |           | OC          | ۰/۶۷ | ۰/۶۸ | ۰/۰۰۳ | ۰/۲۲  |

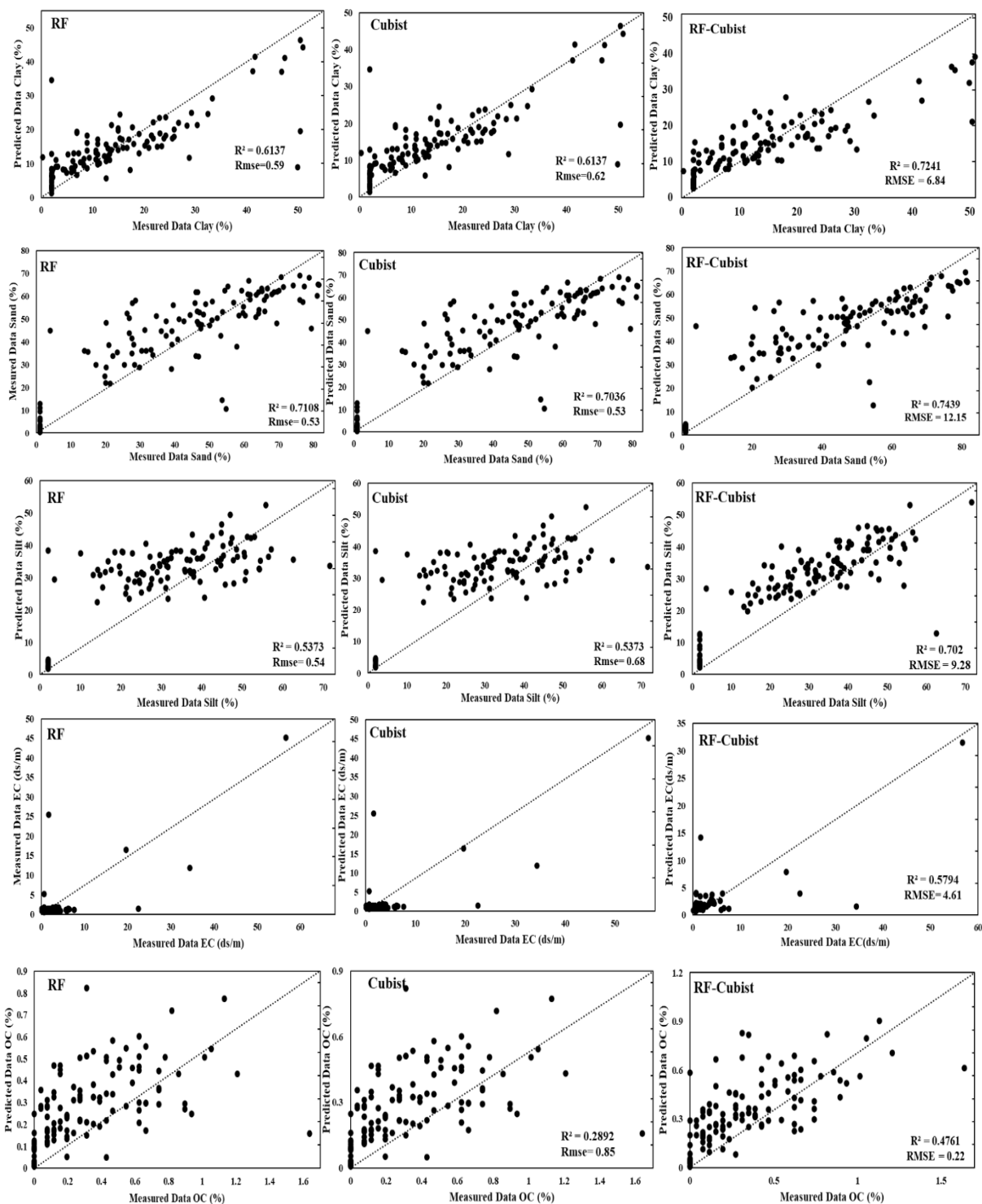
در عمق ۰-۳۰ سانتی‌متر، نتایج نشان داد که مدل RF بهترین عملکرد را در برآورد پارامترهای خاک داشته است. این مدل برای شن (Sand) با RMSE برابر با ۱۰/۱۹ درصد و ضریب همبستگی (Cor) معادل ۰/۹۲، برای سیلت (Silt) با RMSE برابر با ۷/۵۸ درصد و ضریب همبستگی ۰/۸۸، برای رس (Clay) با RMSE برابر با ۵/۸۴ درصد و ضریب همبستگی ۰/۸۵، برای هدایت الکتریکی (EC) با RMSE برابر با ۴/۹۰ دسی‌زیمنس بر متر و ضریب همبستگی ۰/۹۲ و برای کربن آلی (OC) با RMSE برابر با ۰/۲۰ درصد و ضریب همبستگی ۰/۸۱ عملکرد بهتری را نسبت به سایر مدل‌ها از خود نشان داد. مقادیر آماره اریب برای مدل RF در این عمق نیز به ترتیب برای شن ۰/۳۰-، سیلت ۰/۳۷-، رس ۰/۰۳- درصد، هدایت الکتریکی ۰/۴۹- دسی‌زیمنس بر متر و برای کربن آلی ۰/۰۲- درصد بوده که این مقادیر بیانگر دقت بالای این مدل در پیش‌بینی پارامترهای خاک است. نتایج مطالعات مشابه نیز بیانگر دقت بالای مدل جنگل تصادفی در برآورد خصوصیات خاک در عمق ۰-۳۰ سانتی‌متر خاک است (Chelar Awany, et al, 2024 & Musaab, et al, 2023).

در عمق ۳۰-۶۰ سانتی‌متر، مدل RF بهترین عملکرد را برای اکثر پارامترهای خاک داشته است. این مدل برای شن (Sand) با RMSE برابر با ۱۱/۴۲ درصد و ضریب همبستگی ۰/۸۸، برای سیلت (Silt) با RMSE برابر با ۸/۹۹ درصد و ضریب همبستگی ۰/۸۴، برای رس (Clay) با RMSE برابر با ۷/۱۸ درصد و ضریب همبستگی ۰/۸۳، برای هدایت الکتریکی (EC) با RMSE برابر با ۴/۰۸ دسی‌زیمنس بر متر و ضریب همبستگی ۰/۸۱ و برای کربن آلی (OC) با RMSE برابر با ۰/۲۲ درصد و ضریب همبستگی ۰/۷۳ نتایج بهتری نسبت به سایر مدل‌ها ارائه داد. مقادیر اریب در مدل RF برای این عمق نیز به ترتیب برای شن ۰/۲۳-، سیلت ۰/۸۳، رس ۰/۷۳- درصد، هدایت الکتریکی ۰/۳۵- دسی‌زیمنس بر متر و برای کربن آلی ۰/۰۲ درصد بود، که نشان‌دهنده دقت مناسب این مدل در برآورد خصوصیات خاک در این عمق است. براساس مطالعات صورت گرفته، این مقایسه نشان می‌دهد که مدل RF به دلیل مقادیر کمتر RMSE و اریب و مقادیر بالاتر ضریب همبستگی در برآورد ویژگی‌های خاک نسبت به مدل‌های Cubist و RF-Cubist عملکرد برتری دارد و نتایج مطمئن‌تری ارائه می‌دهد (Raúl, et al, 2019 & Muhammad noor, et al, 2022).

شکل‌های ۲ و ۳ نمودارهای پراکندگی داده‌های اندازه‌گیری و پیش‌بینی شده برای هر یک از خصوصیات خاک بر اساس مدل‌های RF و Cubist و RF-Cubist در دو عمق ۰-۳۰ و ۳۰-۶۰ سانتی‌متری خاک منطقه مورد مطالعه را نشان می‌دهد. براساس نمودارهای پراکندگی، در عمق ۰-۳۰ سانتی‌متر در برآورد خصوصیات خاک مدل RF نسبت به مدل Cubist دقت بالاتری داشته است. این در حالی‌ست که مدل RF-Cubist خصوصیات فیزیکی خاک (رس، سیلت و شن) را با دقت بالاتر و هدایت الکتریکی و کربن آلی را با دقت کمتری برآورد کرده است. در عمق ۳۰-۶۰ سانتی‌متر مدل Cubist نسبت به مدل RF دقت بالاتری در برآورد سیلت و مدل RF دقت بالاتری در برآورد شن، رس، هدایت الکتریکی و کربن آلی خاک داشته است. مدل RF-Cubist تمامی پارامترها به جز هدایت الکتریکی را با دقت بالاتری برآورد کرده است. براساس نتایج مطالعه (Clingsmith, & Grunwald, 2022) مدل کیویست دقت بالایی در مدل‌سازی خصوصیات خاک به دلیل پیچیدگی بالای بین ویژگی‌های خاک داشته است.



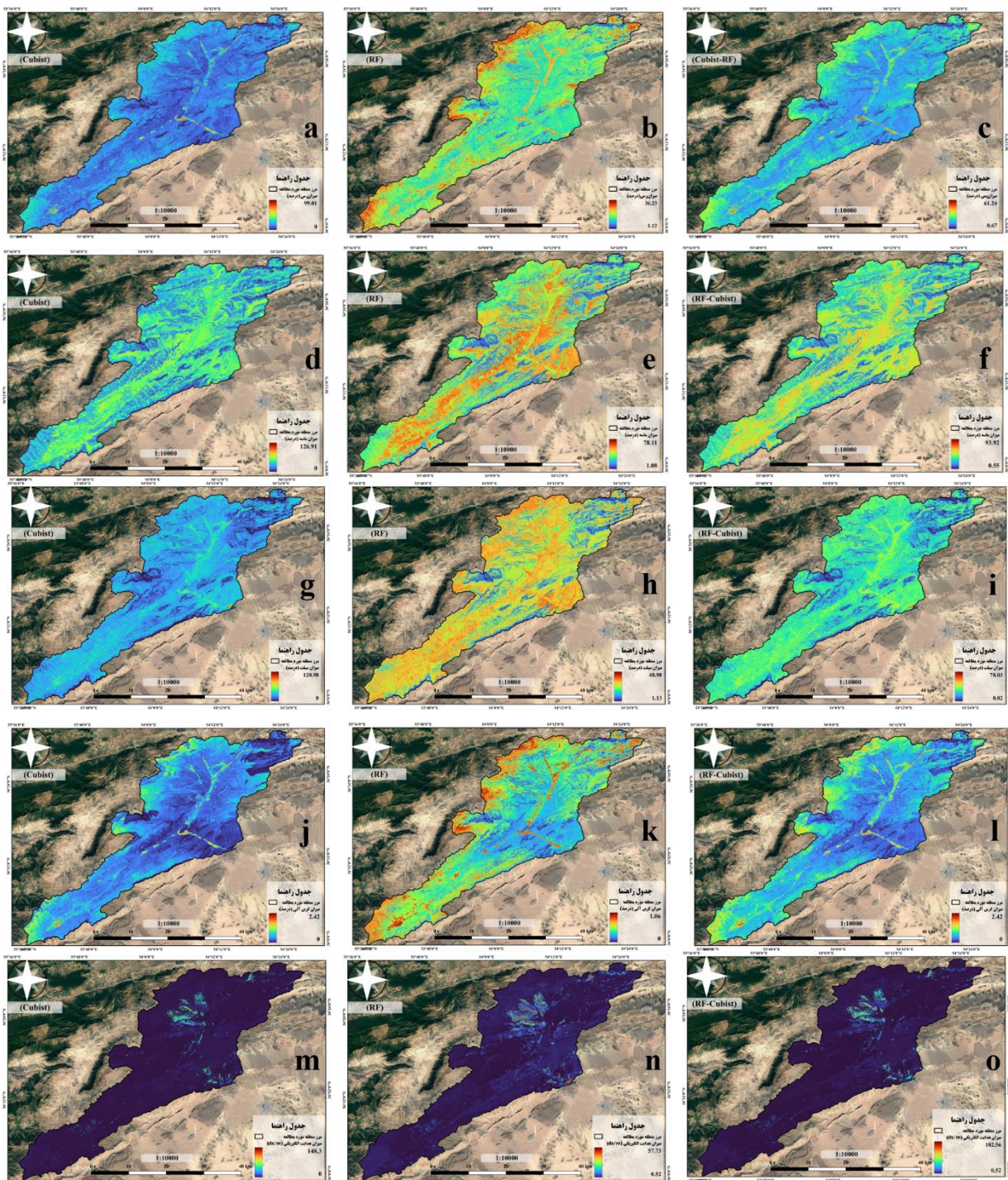
شکل ۲- نمودارهای پراکندگی داده‌های اندازه‌گیری و پیش‌بینی شده براساس مدل‌های RF و Cubist در عمق ۰-۳۰ سانتی متر



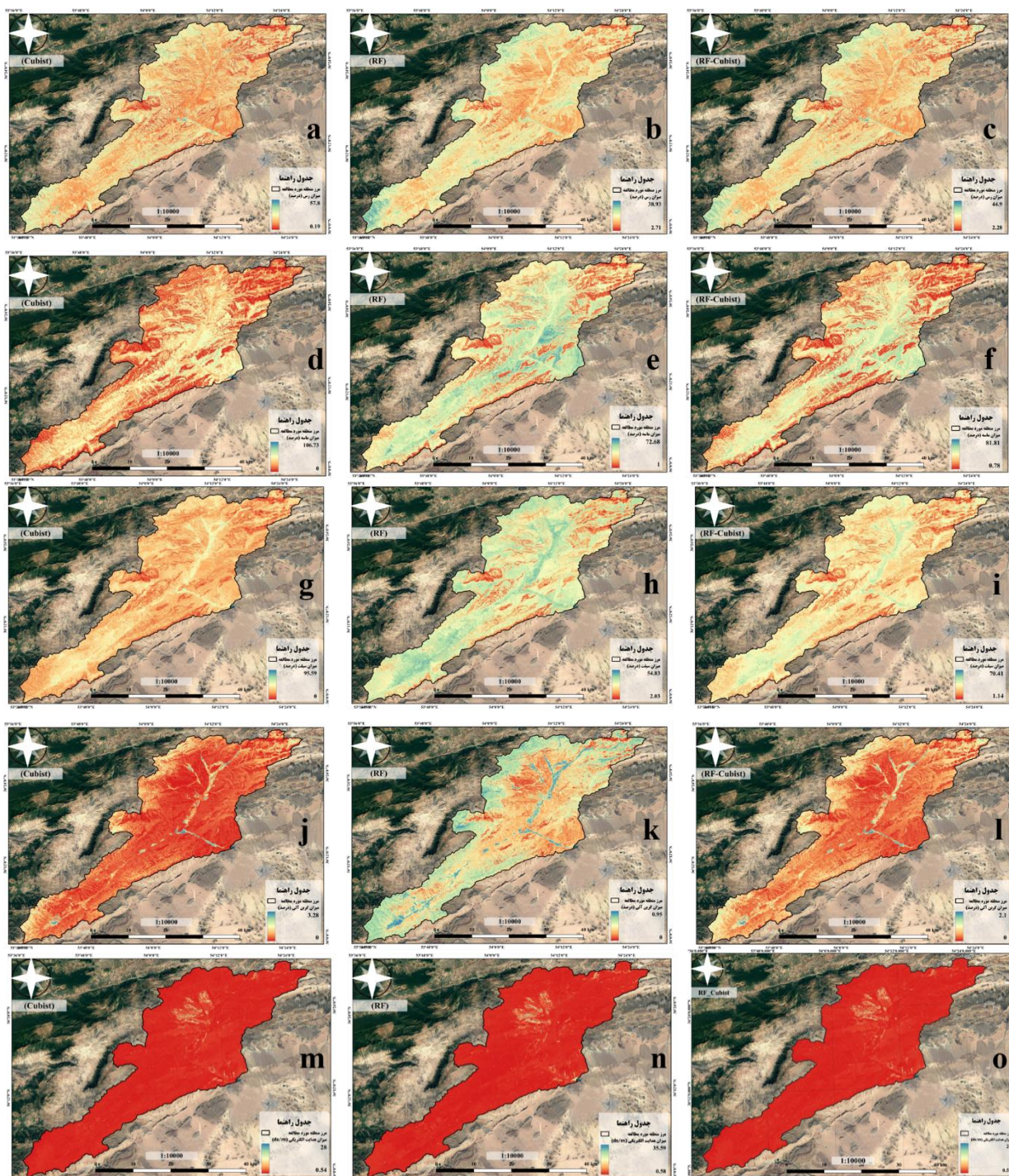
شکل ۳- نمودارهای پراکندگی داده‌های اندازه‌گیری و پیش‌بینی شده براساس مدل‌های RF و Cubist در عمق ۳۰-۶۰ سانتی متر

#### ۴-۳- نقشه توزیع مکانی خصوصیات خاک

با توجه به دقتی که مدل‌های RF و Cubist و RF-Cubist برای برآورد خصوصیات خاک منطقه داشتند، نقشه‌های برآورد توزیع مکانی خصوصیات خاک در عمق‌های ۰-۳۰ و ۳۰-۶۰ سانتی‌متر خاک منطقه توسط مدل‌های RF و Cubist و RF-Cubist در شکل‌های ۴ و ۵ نشان داده شده است.



شکل ۴- نقشه توزیع خصوصیات خاک بر اساس مدل‌های RF و Cubist و RF-Cubist در عمق ۰-۳۰ سانتی متر از سطح خاک



شکل ۵- نقشه توزیع خصوصیات خاک بر اساس مدل‌های RF و Cubist و RF-Cubist در عمق ۳۰-۶۰ سانتی متر از سطح خاک با توجه به اینکه در عمق ۰-۳۰ سانتی متر، هدایت الکتریکی و کربن آلی خاک با استفاده از مدل Cubist و شن، سیلت و رس خاک با استفاده از مدل RF و در عمق ۳۰-۶۰ سانتی متر رس، سیلت و هدایت الکتریکی با استفاده از مدل Cubist، شن و کربن آلی خاک با استفاده از مدل RF با دقت قابل قبولی برآورد شده‌اند. در ادامه توزیع مکانی خصوصیات خاک بر اساس نقشه این مدل‌ها ارائه خواهد شد. شکل ۴-ا توزیع مکانی رس در منطقه مورد مطالعه را نشان می‌دهد. مقادیر رس از حداقل ۱/۱۲ درصد تا حداکثر ۲۳/۳۶ درصد متغیر است و نشان‌دهنده تأثیر عوامل محلی و موضعی مانند نوع سنگ مادر و توپوگرافی بر توزیع رس است. مقادیر درصد شن از حداقل ۰/۶۵ درصد تا حداکثر

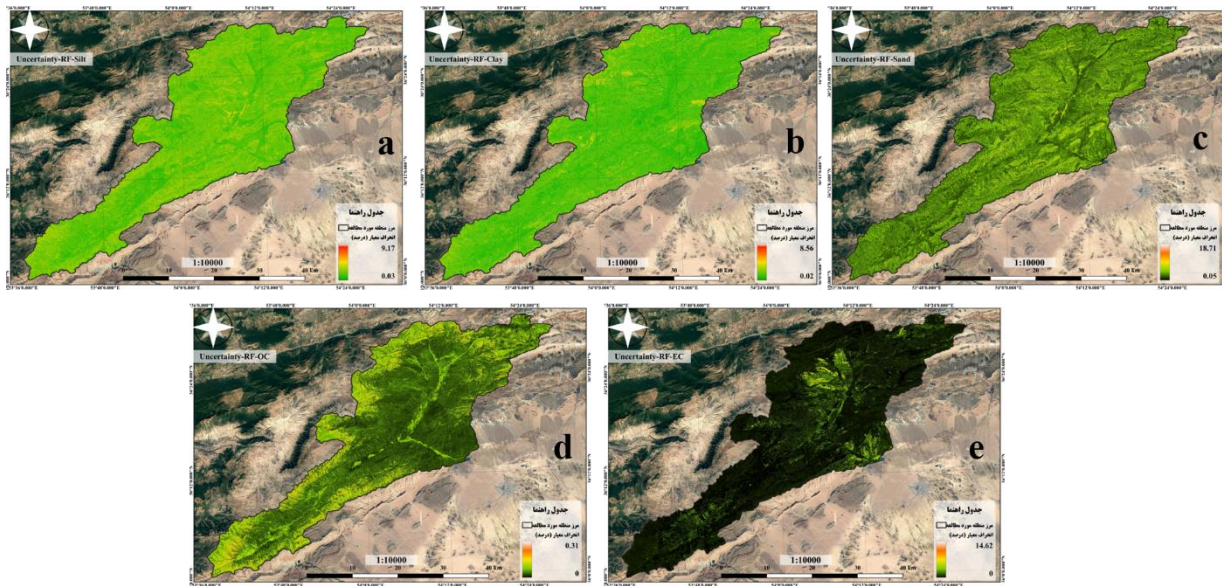
۷۷/۶۳ درصد متغیر است (شکل ۴-d). حداکثر شن در مناطق مرتفع مشاهده شده است و مناطق مرتفع و کوهستانی به دلیل فرسایش بیشتر، و سرد بودن هوا و تأثیر کمتر عوامل خاکسازی، ذرات خاک درشت تر بوده و معمولاً خاک‌ها درصد شن بالاتری دارند. براساس نقشه ۴-f توزیع سیلت در سطح منطقه به صورت ناهمگن است و از حداقل ۰/۸۲ تا حداکثر ۴۱/۷۲ درصد متغیر است. مناطق با درصد سیلت بالا به صورت لکه‌های پراکنده در سراسر منطقه مشاهده می‌شوند که می‌تواند به دلیل وجود مناطق پست و آبرفتی و رسوبگذاری ذرات ریز سیلت باشد. با توجه به تصویر ۴-g هدایت الکتریکی خاک در عمق ۰-۳۰ سانتی‌متری از حداقل صفر (dS/m) تا حداکثر ۱۴۸/۵۳ (dS/m) متغیر است که نشان می‌دهد توزیع هدایت الکتریکی در منطقه به صورت ناهمگن است و مقادیر آن در قسمت‌های جنوبی و شرقی منطقه بالاتر بوده و به سمت شمال و غرب کاهش می‌یابد. مناطق با شوری بالا به صورت لکه‌های پراکنده در برخی نقاط منطقه مشاهده می‌شود. این نشان می‌دهد که عوامل محلی و موضعی مانند نوع خاک، توپوگرافی، پوشش گیاهی و فعالیت‌های انسانی نقش مهمی در توزیع هدایت الکتریکی دارند. در عمق ۳۰-۶۰ سانتی‌متر توزیع رس (شکل ۵-a) به صورت ناهمگن است و مناطق با درصد رس بیشتر به صورت لکه‌ای پراکنده در سطح منطقه و در محل‌های کم شیب با رسوبگذاری بالا مشاهده شده است. مناطق با درصد رس کمتر که اغلب مساحت حوضه را در برمی‌گیرد، دامنه‌های پرشیب و کوهستانی با احتمال فرسایش بالا مشاهده شده است. وجود درصد رس بالا در منطقه می‌تواند به دلیل سنگ بستر رسی و کاهش شیب در منطقه باشد. براساس شکل ۵-e توزیع سیلت نیز در سطح منطقه ناهمگن بوده و مناطق با درصد سیلت خیلی بالا به صورت لکه‌ای در مناطق پست حوضه مشاهده است که دلیل آن کاهش شیب منطقه و عمل رسوبگذاری است. توزیع مکانی شن در منطقه در عمق ۳۰-۶۰ سانتی‌متری بر اساس شکل ۵-d نشان می‌دهد که در مناطق مرتفع با شیب زیاد که میزان فرسایش نیز بالا است بیشترهای خاک‌های دانه درشت شنی مشاهده شده است و با کاهش شیب و ارتفاع از میزان آن کاسته شده است. بررسی توزیع مکانی هدایت الکتریکی در منطقه مورد مطالعه نشان می‌دهد که تغییرات قابل توجهی در میزان این پارامتر مشاهده نمی‌شود و نقاط با هدایت الکتریکی بالا به صورت لکه‌ای در برخی از نقاط به صورت پراکنده مشاهده می‌گردد (شکل ۵-g). کربن آلی خاک بر اساس شکل ۵-j دارای توزیع ناهمگن در سطح است به طوریکه در مناطق مرکزی و شمالی حوضه مقادیر بالای کربن آلی و مناطق جنوبی آن مقادیر کمتر مشاهده می‌گردد که می‌تواند به دلیل وجود پوشش گیاهی انبوه در این مناطق و افزایش درصد کربن آلی خاک باشد.

نتایج مطالعات پیشین نشان می‌دهد که ترکیب معدنی سنگ مادر تأثیر مستقیمی بر تشکیل رس در خاک دارد. به طور خاص، سنگ‌هایی مانند سنگ آهک و بازالت، بسته به میزان هوازدگی، می‌توانند خاک‌هایی با درصد رس متفاوت ایجاد کنند. برای مثال، هوازدگی سنگ آهک به دلیل سرعت بالاتر فرآیندهای تجزیه، منجر به تشکیل خاک‌های توسعه یافته‌تر با مقدار رس بیشتر در مناطق مقعر می‌شود (Oliveira, et al., 2024). در مناطق مرتفع و شیب‌دار، به دلیل فعالیت‌های فرسایشی شدید، درصد شن موجود در خاک افزایش می‌یابد. این مسئله در مطالعات مختلف مورد تأیید قرار گرفته است و نشان می‌دهد که فرآیندهای فرسایش، ذرات درشت‌تر را در این مناطق غالب کرده و مانع از تجمع مواد ریزتر مانند رس و سیلت می‌شوند (Chandrakala, et al., 2022; Begum, et al., 2023). در مقابل، در مناطق پست‌تر، فرصت بیشتری برای رسوب گذاری فراهم بوده و در نتیجه، ذرات ریزتر مانند سیلت تجمع

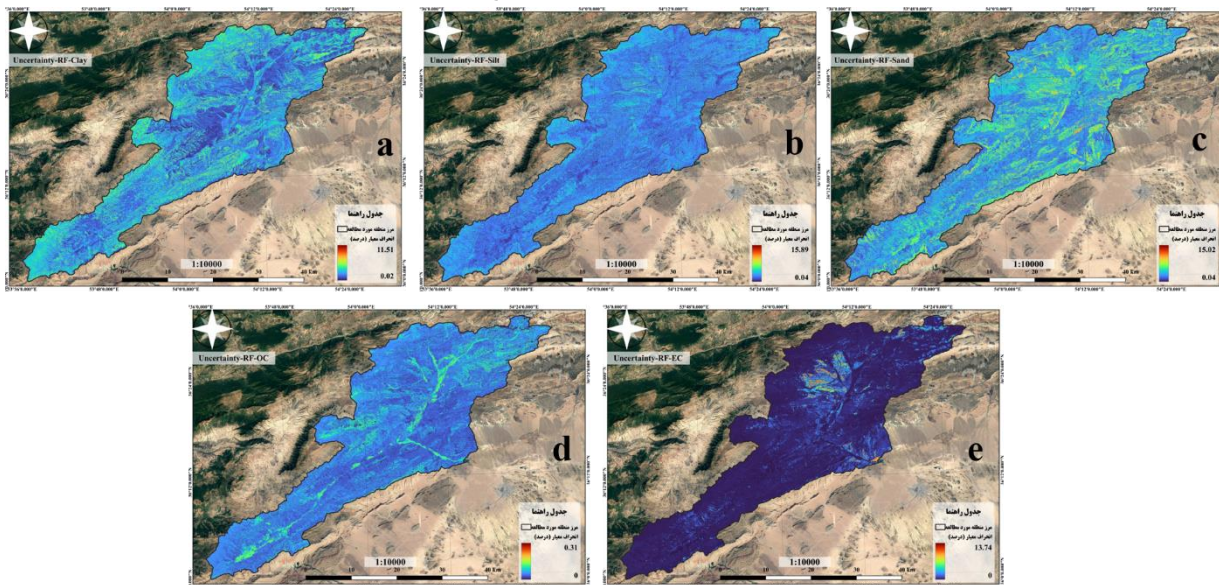
بیشتری دارند که این یافته مطابق با نتایج پژوهش‌های قبلی است (Melnik, 2022; Chandrakala, et al., 2022). علاوه بر این، وضعیت توپوگرافی، فعالیت‌های انسانی و پوشش گیاهی از عوامل تأثیرگذار بر شوری خاک هستند. بررسی‌ها نشان می‌دهند که در مناطق کم‌شیب، به دلیل کاهش زهکشی و افزایش تجمع املاح، شوری خاک افزایش می‌یابد. همچنین، فعالیت‌های انسانی مانند کشاورزی و گسترش مناطق شهری، به‌ویژه در مناطق خشک، می‌توانند باعث افزایش شوری خاک شوند (Zhang, 2010; Triantafilis & Santos, 2010; Beamish & White, 2012). پوشش گیاهی نیز نقش مهمی در بهبود کیفیت خاک دارد. پژوهش‌ها نشان داده‌اند که خاک‌های دارای پوشش گیاهی متراکم، به‌ویژه در مناطق کم‌ارتفاع و کوهستانی، میزان کربن آلی بیشتری نسبت به زمین‌های باز و بدون پوشش دارند. این امر می‌تواند به دلیل افزایش ورود مواد آلی و بهبود شرایط رطوبتی در این مناطق باشد (Yoni & Senou, 2024; Kramarczuk, et al., 2024).

#### ۴-۴- نقشه عدم قطعیت برآورد خصوصیات خاک

هدف از بررسی عدم قطعیت در نقشه برداری رقومی خاک، تعیین نقش داده‌های در دقت مدل‌ها است. به عنوان مثال عدم قطعیت بالا در مدل‌ها بدین مفهوم است که مدل‌ها نسبت به داده‌های آموزش حساس بود و با تغییر داده‌ها، دقت مدل تغییر می‌کند. در این مطالعه برای به‌دست آوردن نقشه عدم قطعیت خصوصیات خاک در هر سلول برای هر یک از خصوصیات خاک ۵۰ مجموعه داده تصادفی تهیه شد و مدل‌سازی با استفاده از هر سری از داده‌ها انجام شد و در نهایت انحراف معیار ۵۰ مدل برای متغیر مورد مطالعه در هر سلول محاسبه شد. نقشه‌های عدم قطعیت خصوصیات خاک در دو عمق ۰-۳۰ و ۳۰-۶۰ سانتی‌متر تهیه شد. برای عمق ۰-۳۰ سانتی‌متر، عدم قطعیت پارامتر رس بین ۰/۰۲ تا ۸/۵۶ درصد در مناطق مرکزی و شمالی، شن بین ۰/۰۵ تا ۱۸/۷۱ درصد در مناطق شمالی و شمال شرقی، سیلت بین ۰/۰۳ تا ۹/۱۷ درصد در مناطق شمالی و شمال شرقی، کربن آلی خاک بین ۰ تا ۰/۳۱ درصد در مناطق شمال شرقی و جنوب غربی و هدایت الکتریکی خاک بین ۰ تا ۱۴/۶۲ درصد در مناطق شمال شرقی، مرکز و جنوب متغیر بود. به طور کلی، مناطق مرکزی و جنوبی دارای کمترین عدم قطعیت و مناطق شمالی و شمال شرقی دارای بیشترین عدم قطعیت بودند. اما به طور کلی در بیشتر سلول‌ها مقادیر عدم قطعیت به مقدار صفر نزدیک بود. برای عمق ۳۰-۶۰ سانتی‌متر، عدم قطعیت پارامتر رس بین ۰/۰۲ تا ۱۱/۵۱ درصد در مناطق شمال شرقی، مرکزی و جنوب غربی، سیلت بین ۰/۰۴ تا ۱۵/۸۹ درصد در مناطق مرکزی و شمال شرقی، شن بین ۰/۰۴ تا ۱۵/۰۲ درصد در مناطق مرکزی، شمالی و جنوبی، کربن آلی خاک بین ۰ تا ۰/۳۱ درصد در مناطق با پوشش گیاهی بیشتر، و هدایت الکتریکی خاک بین ۰ تا ۱۳/۷۴ درصد در مناطق جنوب، مرکز و شمال شرقی دارای بیشترین عدم قطعیت بود. براساس نتایج به‌دست آمده مدل RF دارای عدم قطعیت پایینی برای همه خصوصیات خاک ذکر شده در این مطالعه است. نتایج مطالعات متعدد از جمله (Zhang, et al, 2022 & Zhou, et al, & 2021 De Anta, et al, 2020)، نیز نشان دهنده عدم قطعیت پایین مدل RF در برآورد خصوصیات خاک مانند درصد کربن آلی و رس است.



شکل ۶- نقشه عدم قطعیت توزیع خصوصیات خاک (a: سیلت، b: رس، c: شن، d: کربن آلی خاک، e: هدایت الکتریکی خاک) بر اساس مدل‌های RF در عمق ۰-۳۰ سانتی متر از سطح خاک



شکل ۷- نقشه عدم قطعیت توزیع خصوصیات خاک (a: رس، b: سیلت، c: شن، d: کربن آلی خاک، e: هدایت الکتریکی خاک) بر اساس مدل‌های RF در عمق ۳۰-۶۰ سانتی متر از سطح خاک

## ۵- بحث و نتیجه‌گیری

در این مطالعه، توزیع مکانی خصوصیات فیزیکی و شیمیایی خاک شامل درصد شن، رس، سیلت، کربن آلی و هدایت الکتریکی خاک با استفاده از مدل‌های یادگیری ماشین RF-Cubist و Cubist RF حوضه آبخیز دامغان رود مورد بررسی قرار گرفت. متغیرهای کمکی استفاده شده برای برآورد خصوصیات خاک داده‌های سنجش از دوری و شاخص‌های مستخرج از مدل رقومی ارتفاع بوده که در چارچوب روش DSM و با استفاده از روش‌های یادگیری ماشین صورت پذیرفت. نتایج نشان داد که مدل‌های یادگیری ماشین ابزارهای بسیار مناسبی برای برآورد خصوصیات خاک هستند، با این حال نتایج بیانگر این است که برای هر ویژگی خاک، مدل خاصی می‌تواند نقشه‌های دقیق با خطای کمتر ارائه دهد، به عنوان

مثال در عمق ۰-۳۰ سانتی‌متری مدل RF با داشتن  $R^2$  بالا و RMSE کمتر دقت بالایی در برآورد خصوصیات فیزیکی خاک نسبت به مدل‌های Cubist و RF-Cubist و مدل Cubist دقت بالاتری در برآورد خصوصیات شیمیایی خاک داشته است. مدل RF-Cubist در عمق ۳۰-۶۰ سانتی‌متری خاک با داشتن  $R^2$  بالا و RMSE کمتر بهترین مدل در برآورد خصوصیات خاک در مقایسه با مدل‌های RF و Cubist بوده است. در نهایت برای بهبود نتایج در آینده پیشنهاد می‌شود با توجه به عمق نمونه‌برداری و نوع خصوصیت خاک مدل‌ها انتخاب گردد و کیفیت داده‌ها افزایش یافته و استفاده از مدل‌های ترکیبی توسعه یابد. تعیین الگوهای کشت مناسب بر اساس نقشه‌ها و همچنین تغییر کشت سنتی به کشت گلخانه‌ای می‌تواند از تخریب خاک جلوگیری کرده و به طور بالقوه با استفاده بهینه از منابع، بهره‌وری را افزایش دهد که ممکن است در حوضه آبخیز دامغان رود ضروری باشد.

## منابع

1. Aghanbati, A. (2004). Geology of Iran, Geological Survey of Iran, first edition.
2. Alves, M. C. G. P., Barros, S. V. A., & Oliveira, F. N. de. (2024). Physical and chemical characterization of soil for use in ecological bricks.
3. <https://doi.org/10.56238/sevened2024.018-037>.
4. Arshad, M., Li, N., Bella, L. D., & Triantafilis, J. (2020). Field - scale digital soil mapping of clay: Combining different proximal sensed data and comparing various statistical models. *Soil Science Society of America Journal*, 84(2), 314-330.
5. Ayaz, M., Uddin, M. A., Sharif, Z., Mansour, A., & Aggoune, E. M. (2019). Internet-of-Things (IoT)-Based Smart Agriculture: Toward making the fields talk. *IEEE Access*, 7, 129551-129583. <https://doi.org/10.1109/access.2019.2932609>.
6. Bao, Y., Yao, F., Meng, X., Wang, J., Liu, H., Wang, Y., & Mouazen, A. M. (2024). A fine digital soil mapping by integrating remote sensing-based process model and deep learning method in Northeast China. *Soil and Tillage Research*, 238, 106010.
7. Beamish, D., & White, J. C. (2012). Mapping and predicting electrical conductivity variations across southern England using airborne electromagnetic data. *Quarterly Journal of Engineering Geology and Hydrogeology*, 45(1), 99-110. <https://doi.org/10.1144/1470-9236/11-026>
8. Begum, M., Hossain, M. S., Aziz, M. A., Choudhury, M. A. R., & Jahan, I. (2023). Variability of major soil properties of a fallow-acidic-level upland with high and multiple spatial resolutions. *Eurasian Journal of Soil Science*, 12(2), 177-189.
9. Bek, A., Jug, J., Strelec, S., & Dodigović, F. (2024). Application of geotechnical and geophysical research for soil characterization. 103-117. [https://doi.org/10.35123/geo-expo\\_2024\\_12](https://doi.org/10.35123/geo-expo_2024_12).
10. Breiman, L. (2001). Random forests. *Machine Learning*, 45(1), 5-32. <https://doi.org/10.1023/A:1010933404324>.
11. Carter MR, Gregorich EG. (2007) Soil sampling and methods of analysis. CRC press. Boca Raton, Carter, & E.G. Gregorich [Eds.], 1264 pp. Boca Raton. ISBN: 9780429126222. <https://doi.org/10.1201/9781420005271>.
12. Chandrakala, M., KS, A. K., & SUJATHA, K. (2022). Soil heterogeneity: A comparative assessment of soils from two different AESR, southern India. *Annals of Plant and Soil Research*, 24(1), 29-35.
13. Chelar, A. A., Aidi Shareef, M., & Al Nuaimy, Q. A. M. (2024). Estimating soil parameters using C-band synthetic aperture radar in Laylan, Iraq. *Ecological Engineering & Environmental Technology*. <https://doi.org/10.12912/27197050/189889>.
14. Chen, S., Arrouays, D., Angers, D. A., Chenu, C., Barré, P., Martin, M. P., Saby, N. P. A., & Walter, C. (2019). National estimation of soil organic carbon storage potential for arable soils:

- A data-driven approach coupled with carbon-landscape zones. *Science of The Total Environment*, 666, 355–367.
15. Cheng, X. (2024). A Comprehensive Study of Feature Selection Techniques in Machine Learning Models. *Insights in Computer, Signals and Systems.*, 1(1), 65–78. <https://doi.org/10.70088/xpf2b276>.
  16. Cheshmberah, F., Zolfaghari, A. A., Taghizadeh-Mehrjardi, R., & Scholten, T. (2022). Evaluation of mathematical models for predicting particle size distribution using digital soil mapping in semiarid agricultural lands. *Geocarto International*, 37(26), 13016-13038.
  17. Clingensmith, C. M., & Grunwald, S. (2022). Predicting Soil Properties and Interpreting Vis-NIR Models from across Continental United States. *Sensors*, 22(9), 3187. <https://doi.org/10.3390/s22093187>.
  18. Cosenza, P., Prêt, D., & Zamora, M. (2015). Effect of the local clay distribution on the effective electrical conductivity of clay rocks. *Journal of Geophysical Research*, 120(1), 145–168. <https://doi.org/10.1002/2014JB011429>
  19. Cover T, Hart P. Nearest neighbor pattern classification. *IEEE Trans Inf Theor.* 1967;13(1):21-7.
  20. Danial, Abdorahimi., Ali, M., Sadeghioon. (2019). Comparison of radio frequency path loss models in soil for wireless underground sensor networks. *Journal of Sensor and Actuator Networks*, doi: 10.3390/JSAN8020035.
  21. de Oliveira, D. P., Bovi, R. C., Castilho, Y. F. P., de Souza Júnior, V. S., Cooper, M., Queiroz, H. M., ... & Ferreira, T. O. (2024). Lithotype composition and micro-topographic features: drivers of soil variability on semiarid limestone plateaus. *Geoderma Regional*, 37, e00814.
  22. De Anta, R. C., Luís, E., Febrero-Bande, M., Galiñanes, J., Macías, F., Ortíz, R., & Casás, F. (2020). Soil organic carbon in peninsular Spain: Influence of environmental factors and spatial distribution. *Geoderma*, 370, 114365.
  23. Dharumarajan, S., & Hegde, R. (2022). Digital mapping of soil texture classes using Random Forest classification algorithm. *Soil Use and Management*, 38(1), 135-149.
  24. Dharumarajan, S., Kalaiselvi, B., Suputhra, A., Lalitha, M., Hegde, R., Singh, S. K., & Lagacherie, P. (2020). Digital soil mapping of key GlobalSoilMap properties in Northern Karnataka Plateau. *Geoderma Regional*, 20, e00250. <https://doi.org/10.1016/j.geodrs.2019.e00250>.
  25. Dharumarajan, S., Kalaiselvi, B., Suputhra, A., Lalitha, M., Vasundhara, R., Kumar, K. A., ... & Lagacherie, P. (2021). Digital soil mapping of soil organic carbon stocks in Western Ghats, South India. *Geoderma Regional*, 25, e00387.
  26. Dobarco, M. R., Bourennane, H., Arrouays, D., Saby, N. P., Cousin, I., & Martin, M. P. (2019). Uncertainty assessment of GlobalSoilMap soil available water capacity products: A French case study. *Geoderma*, 344, 14-30.
  27. Duarte, E., Zagal, E., Barrera, J. A., Dube, F., Casco, F., & Hernandez, A. J. (2022). Digital mapping of soil organic carbon stocks in the forest lands of Dominican Republic. *European journal of remote sensing*, 55(1), 213-231.
  28. FAO, 2011. FAO in the 21st century.
  29. Gee, G. W., & Bauder, J. W. (1986). Particle Size Analysis. In: *Methods of Soil Analysis. Part 1. Physical and Mineralogical Methods*.
  30. Gómez, A. M., van Lier, Q. D. J., Silvero, N. E., Inforsato, L., de Melo, M. L. A., Rodríguez-Albarracín, H. S., ... & Demattê, J. A. (2023). Digital mapping of the soil available water capacity: tool for the resilience of agricultural systems to climate change. *Science of The Total Environment*, 882, 163572.
  31. Haubold, J., Hosch, R., Parmar, V., Glas, M., Guberina, N., Catalano, O. A., ... & Umutlu, L. (2021). Fully automated MR-based virtual biopsy of cerebral gliomas. *Cancers*, 13(24), 1-13.
  32. Hengl, T., Mendes de Jesus, J., Heuvelink, G.B., Ruiperez Gonzalez, M., Kilibarda, M., Blagotić, A., & Kempen, B. (2017). SoilGrids250m: Global gridded soil information based on machine learning. *PLoS One*, 12(2), e0169748.

33. Hounkpatin, K. O., Bossa, A. Y., Yira, Y., Igue, M. A., & Sinsin, B. A. (2022). Assessment of the soil fertility status in Benin (West Africa)–Digital soil mapping using machine learning. *Geoderma Regional*, 28, e00444.
34. <https://doi.org/10.1371/journal.pone.0169748>
35. Huda, S., Yearwood, J., Jelinek, H. F., Hassan, M. M., Fortino, G., & Buckland, M. (2017). A hybrid feature selection with ensemble classification for imbalanced healthcare data: A case study for brain tumor diagnosis. *IEEE Access*, 4, 1-13.
36. Jamshidi, k. (2010). Study of Magmatism of the Shemshak Formation Basis in the Eastern Alborz Zone, M.Sc. Thesis, Shahroud University of Technology, p 159.
37. Jamshidi, K., & Ghasemi, H. (2011). Geology and geochemistry of basalt rocks at the base of Shemshak Formation in Tazreh, Kalateh Rudbar and Cheshmeh Ali areas (north and northwest of Damghan), 15th Iranian Geological Society Conference, Tehran, <https://civilica.com/doc/135552>.
38. Jordan, P., & Fröhle, P. (2022). Bridging the gap between coastal engineering and nature conservation? *Journal of Coastal Conservation*, 26(2). <https://doi.org/10.1007/s11852-021-00848-x>.
39. Keshavarzi, A., del Árbol, M. Á. S., Kaya, F., Gyasi - Agyei, Y., & Rodrigo - Comino, J. (2022). Digital mapping of soil texture classes for efficient land management in the Piedmont plain of Iran. *Soil Use and Management*, 38(4), 1705-1735.
40. Khaledian, Y., & Miller, B. A. (2020). Selecting appropriate machine learning methods for digital soil mapping. *Applied Mathematical Modelling*, 81, 401-418.
41. Khongnawang, T., Zare, E., Srihabun, P., Khunthong, I., & Triantafilis, J. (2022). Digital soil mapping of soil salinity using EM38 and quasi - 3d modelling software (EM4Soil). *Soil Use and Management*, 38(1), 277-291.
42. Khosravani, P., Baghernejad, M., Moosavi, A. A., et al. (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, 1367. <https://doi.org/10.1007/s10661-023-11980-6>.
43. Khosravi, M., Zolfaghari, A., Kaboli, S. H., & Ghafari, H. (2022). Application of digital soil mapping in soil particle size zonation and estimation of saturated soil hydraulic conductivity for optimal management of watersheds (case study: Damghanrood Watershed). *Iranian Journal of Soil and Water Research*, 53(2), 245-261.
44. Kibblewhite, M., Ritz, K., & Swift, M. J. (2007). Soil health in agricultural systems. *Philosophical Transactions - Royal Society. Biological Sciences*, 363(1492), 685–701. <https://doi.org/10.1098/rstb.2007.2178>
45. Kramarczuk, P., Musielok, Ł., Stolarczyk, M., Jelonkiewicz, Ł., Nikorych, V. A., & Szymański, W. (2024). Impact of vegetation type on the content and spectroscopic properties of soil organic matter in the subalpine zone of the Bieszczady Mountains (Eastern Carpathians). *Plant and Soil*. <https://doi.org/10.1007/s11104-024-07085-9>.
46. Kuhn M, Johnson K. Regression trees and rule-based models. In: *Applied predictive modeling*. Springer. 2013;173-220.
47. Liu, F., Wu, H., Zhao, Y., Li, D., Yang, J. L., Song, X., ... & Zhang, G. L. (2022). Mapping high resolution national soil information grids of China. *Science Bulletin*, 67(3), 328-340.
48. López-Castañeda, A., Zavala-Cruz, J., Palma-López, D. J., Rincón-Ramírez, J. A., & Bautista, F. (2022). Digital mapping of soil profile properties for precision agriculture in developing countries. *Agronomy*, 12(2), 353.
49. Maghami Moghim, Gh; Horasani, Y; Binaian, Sh. (2023). Estimation of Sedimentation and Erosion Potential of the Damghan Rud Watershed Using the Hydrophysical Model (Csy). *Hydrogeomorphology*, 10(35): 47-65.
50. Malone, B. P., Luo, Z., He, D., Rossel, R. V., & Wang, E. (2020). Bioclimatic variables as important spatial predictors of soil hydraulic properties across Australia's agricultural region. *Geoderma Regional*, 23, e00344.

51. Malone, B. P., McBratney, A. B., Minasny, B., & Laslett, G. M. (2009). Mapping continuous depth functions of soil carbon storage and available water capacity. *Geoderma*, 154(1-2), 138-152.
52. Malone, B. P., Minasny, B., & McBratney, A. B. (2017). Using R for digital soil mapping. *Progress in Soil Science*. Springer International Publishing, Cham. <https://doi.org/10.1007/978-3-319-44327-0>.
53. McBratney, A.B., Mendonca-Santos, L. and Minasny, B. (2003). On digital soil mapping. *Geoderma* 117, 3-52.
54. 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>.
55. Mosleh, Z., Salehi, M. H., Jafari, A., Borujeni, I.E., & Mehnatkesh, A. (2016). The effectiveness of digital soil mapping to predict soil properties over low-relief areas. *Environmental Monitoring and Assessment*, 188, 1-13. <https://doi.org/10.1007/s10661-016-5204-8>
56. Mousavi, A., Karimi, A., & Maleki, S. (2023). Digital mapping of selected soil properties using machine learning and geostatistical techniques in Mashhad Plain, northeastern Iran. *Environmental Earth Sciences*, 82, 234. <https://doi.org/10.1007/s12665-023-10919-x>.
57. Mousavi, A., Karimi, A., Maleki, S., Safari, T., & Taghizadeh-Mehrjardi, R. (2023). Digital mapping of selected soil properties using machine learning and geostatistical techniques in Mashhad plain, northeastern Iran. *Environmental Earth Sciences*, 82(9), 234.
58. Muhammad, Noor, Hazwan, Jusoh., Deprizon, Syamsunur., N., Abd, Rahman., Emmanuel, Olisa., Abdur, Rahim., Nur, Izzi, Md., Yusoff. (2022). 2. Soil Parameters Model Prediction via Resistivity Value Limit to Shallow Subsurface Areas. *Shock and Vibration*, doi: 10.1155/2022/3251250
59. Musaab, Sabah, Abed., Firas, Jawad, Kadhimi., Jwad, K., Almusawi., Hamza, Imran., L., Bernardo., Sadiq, N., Henedy. (2023). 4. Utilizing Multivariate Adaptive Regression Splines (MARS) for Precise Estimation of Soil Compaction Parameters. *Applied Sciences*, doi: 10.3390/app132111634.
60. Nabiollahi, K., Taghizadeh-Mehrjardi, R., Shahabi, A., Heung, B., Amirian-Chakan, A., Davari, M., & Scholten, T. (2021). Assessing agricultural salt-affected land using digital soil mapping and hybridized random forests. *Geoderma*, 385, 114858.
61. Nahar, N., Ara, F., Neloy, M. A. I., Biswas, A., Hossain, M. S., & Andersson, K. (2021). Feature selection-based machine learning to improve prediction of Parkinson's disease. In *International Conference on Brain Informatics* (pp. 496-508). Springer.
62. Nauman, T. W., & Duniway, M. C. (2019). Relative prediction intervals reveal larger uncertainty in 3D approaches to predictive digital soil mapping of soil properties with legacy data. *Geoderma*, 347, 170-184.
63. Poppiel, R. R., Coelho, M. P. L., Safanelli, J. L., Rizzo, R., Oliveira, M. P., Novais, J. J., & Demattê, J. A. M. (2019). Mapping at 30 m resolution of soil attributes at multiple depths in Midwest Brazil. *Remote Sensing*, 11(24), 2905. <https://doi.org/10.3390/rs11242905>.
64. Quinlan, J.R, 1992. Learning with continuous classes. *Proceedings of the 5th Australian Joint Conference On Artificial Intelligence*, pp. 343 - 348.
65. Reddy, N. N., & Das, B. S. (2023). Digital soil mapping of key secondary soil properties using pedotransfer functions and Indian legacy soil data. *Geoderma*, 429, 116265.
66. Rezaei Moghaddam, M., Mokhtari, D., Samandar, N. (2022). Detection of the effect of land use change trends on discharge and sediment Simulated with SWAT model in Ojan Chay watershed. *Hydrogeomorphology*, 9(30), 24-1. doi: 10.22034/hyd.2021.42595.1557.
67. Rhoades, J. D. (1996). Salinity: Electrical Conductivity and Total Dissolved Solids. In: *Methods of Soil Analysis. Part 3. Chemical Methods*.
68. Romano, E., Bergonzoli, S., Bisaglia, C., Picchio, R., & Scarfone, A. (2022). The correlation between proximal and remote sensing methods for monitoring soil water content in agricultural applications. *Electronics*, 12(1), 127.
69. Schoen, R. A., & Johnson, A. F. (Eds.). (2021). *Creating the Future Workforce in Food, Agriculture, and Natural Resources: Proceedings of a Workshop [E-book]*. Washington, DC:

- The National Academies Press. DOI: 10.17226/26065. Retrieved from <https://nap.nationalacademies.org/catalog/26065/creating-the-future-workforce-in-food-agriculture-and-natural-resources>
70. Searle, R., McBratney, A., Grundy, M., Kidd, D., Malone, B., Arrouays, D., ... & Andrews, K. (2021). Digital soil mapping and assessment for Australia and beyond: A propitious future. *Geoderma Regional*, 24, e00359.
  71. Sharma, A., & Mishra, P. K. (2022). Performance analysis of machine learning based optimized feature selection approaches for breast cancer diagnosis. *International Journal of Information Technology*, 14(4), 1949-1960.
  72. Smith, D., Bell, N., Faller, J., Galaz, V., Norström, A. V., Pattison, C., & Queiroz, C. (2022). Elements of a planetary Emergency: Environment of Peace (Part 1). <https://doi.org/10.55163/mdeb4357>.
  73. Somarathna, P. D. S. N., Searle, R., & Gladish, D. W. (2021). Mapping available soil water capacity in New South Wales, Australia using sparse data—An inverse Bayesian approach. *Geoderma Regional*, 25, e00396.
  74. Suleymanov, A., Abakumov, E., Suleymanov, R., Gabbasova, I., & Komissarov, M. (2021). The soil nutrient digital mapping for precision agriculture cases in the Trans-Ural steppe zone of Russia using topographic attributes. *ISPRS International Journal of Geo-Information*, 10, 243.
  75. Suleymanov, A., Tuktarova, I., Belan, L. et al. Spatial prediction of soil properties using random forest, k-nearest neighbors and cubist approaches in the foothills of the Ural Mountains, Russia. *Model. Earth Syst. Environ.* 9, 3461–3471 (2023). <https://doi.org/10.1007/s40808-023-01723-4>.
  76. Taghizadeh-Mehrjardi, R., Hamzehpour, N., Hassanzadeh, M., Heung, B., Goydaragh, M. G., Schmidt, K., & Scholten, T. (2021). Enhancing the accuracy of machine learning models using the super learner technique in digital soil mapping. *Geoderma*, 399, 115108.
  77. Taghizadeh-Mehrjardi, R., Nabiollahi, K., Rasoli, L., Kerry, R., Scholten, T. (2020). Land suitability assessment and agricultural production sustainability using machine learning models. *Agronomy* 10 (573), 1–20.
  78. Toth, B., Weynants, M., Pasztor, L., & Hengl, T. (2017). 3D soil hydraulic database of Europe at 250 m resolution. *Hydrological Processes*, 31, 2662–2666. <https://doi.org/10.1002/hyp.11203>.
  79. Triantafyllis, J., & Santos, F. A. M. (2010). Resolving the spatial distribution of the true electrical conductivity with depth using EM38 and EM31 signal data and a laterally constrained inversion model. *Soil Research*, 48(5), 434–446. <https://doi.org/10.1071/SR09149>
  80. Tudi, M., Ruan, H. D., Wang, L., Lyu, J., Sadler, R., Connell, D., Chu, C., & Phung, D. (2021). Agriculture development, pesticide application and its impact on the environment. *International Journal of Environmental Research and Public Health/International Journal of Environmental Research and Public Health*, 18(3), 1112. <https://doi.org/10.3390/ijerph18031112>.
  81. Vanmaercke, M., Panagos, P., Vanwallegghem, T., Hayas, A., Foerster, S., Borrelli, P., Rossi, M., Torri, D., Casalí, J., Borselli, L., Vigiak, O., Maerker, M., Haregeweyn, N., De Geeter, S., Zgłobicki, W., Biielders, C., Cerdà, A., Conoscenti, C., De Figueiredo, T., . . . Poesen, J. (2021). Measuring, modelling and managing gully erosion at large scales: A state of the art. *Earth-science Reviews*, 218, 103637. <https://doi.org/10.1016/j.earscirev.2021.103637>.
  82. Wadoux, A. M. C., Minasny, B., & McBratney, A. B. (2020). Machine learning for digital soil mapping: Applications, challenges and suggested solutions. *Earth-Science Reviews*, 210, 103359.
  83. Walkley, A., & Black, I. A. (1934). An Examination of the Degtjareff Method for Determining Soil Organic Matter and a Proposed Modification of the Chromic Acid Titration Method. *Soil Science*, 37(1), 29-38.

84. Yang, R. M., Liu, L. A., Zhang, X., He, R. X., Zhu, C. M., Zhang, Z. Q., & Li, J. G. (2022). The effectiveness of digital soil mapping with temporal variables in modeling soil organic carbon changes. *Geoderma*, 405, 115407.
85. Yoni, M., & Senou, I. (2024). Distribution of soil carbon according to pedological profile in the Sudanese savannah Agro-systems of Bondoukuy, western Burkina Faso. *World Journal Of Advanced Research and Reviews*, 24(1), 265–277. <https://doi.org/10.30574/wjarr.2024.24.1.2653>
86. Zare, E., Wang, J., Zhao, D., Arshad, M., & Triantafilis, J. (2021). Scope to map available water content using proximal sensed electromagnetic induction and gamma-ray spectrometry data. *Agricultural Water Management*, 247, 106705.
87. Zeraatpisheh, M., Bakhshandeh, E., Hosseini, M., & Alavi, S. M. (2020). Assessing the effects of deforestation and intensive agriculture on the soil quality through digital soil mapping. *Geoderma*, 363, 114139.
88. Zhang, G. (2010). Vertical Distribution Characterization of Electrical Conductivity of Urban Soil under Different Land Use Type. *Journal of Soil and Water Conservation*. [https://en.cnki.com.cn/Article\\_en/CJFDTOTAL-TRQS201004038.htm](https://en.cnki.com.cn/Article_en/CJFDTOTAL-TRQS201004038.htm)
89. Zhang, P., Yin, Z. Y., & Jin, Y. F. (2022). Machine learning-based modelling of soil properties for geotechnical design: review, tool development and comparison. *Archives of Computational Methods in Engineering*, 1-17.
90. Zhang, Y., Ji, W., Saurette, D. D., Easher, T. H., Li, H., Shi, Z., ... & Biswas, A. (2020). Three-dimensional digital soil mapping of multiple soil properties at a field-scale using regression kriging. *Geoderma*, 366, 114253.
91. Zhou, J., Li, E., Wei, H., Li, C., Qiao, Q., & Armaghani, D. J. (2019). Random forests and Cubist algorithms for predicting shear strengths of rockfill materials. *Applied Sciences*, 9(8), 1621. <https://doi.org/10.3390/app9081621>.