

Digital Soil Mapping by Machine Learning Techniques

Masoumeh Mahmoudi¹, Hamid Reza Momtaz², Moslem Servati³, Hassan Mohammadi⁴

1-Department of Soil Science, Faculty of Agriculture, Urmia University, Urmia, Iran.

E-mail: masume.shno@gmail.com

2-Corresponding Author, Department of Soil Science, Faculty of Agriculture, Urmia University, Iran.

E-mail: hamidmomtaz@gmail.com

3-Department of Agricultural Extension and Landscape Architecture, Shahid Bakeri High Education Center of Miandoab, Urmia University, Iran. E-mail: m.sarvati@urmia.ac.ir

4-Department of Soil Science, Faculty of Agriculture, Shahrekord University, Iran.

E-mail: hassan.mohammadi1369@gmail.com

Received: May 15, 2024

Revised: June 15, 2024

Accepted: June 19, 2024

Published: December 21, 2024

Extended Abstract

Background and Objectives

The use of geospatial techniques for mapping soils is broadly covered by the term digital soil mapping (DSM). Soil maps have considerable significance as basic maps in many environmental and natural resources studies. Digital soil maps are based on the relationship between environmental variables and soil properties. With the development of computers and technology, digital and quantitative approaches have been developed. Continuous utilization of agricultural lands regardless of the land suitability caused soil destruction. Also, incompetency in custom methods, invention geographic information system (GIS), and remote sensing (RS) techniques cause erupt and use of digital soil mapping.

Methodology

The study area is approximately 5000 ha which is located in the west of Heris region of East Azerbaijan province, Iran. In the first study, the potential of different models to predict soil classes at different taxonomic levels was investigated. According to semi-detailed soil, survey and using stratified random sampling method, 50 pedons and 50 augers with an approximate distance of 1000 m were excavated, described and soil samples were taken from different genetic horizons. Based on the pedon descriptions and soil analytical data, pedons were classified up to the family level. Different machine learning techniques, namely boosted regression tree (BRT), random forest (RF), artificial neural networks (ANNs), and multinomial logistic regression (MLR) were used to test the predictive power for mapping the soil classes. After preparing the soil properties maps and checking their accuracy, these maps were used along with auxiliary parameters for estimating soil classes using an artificial neural network model in the R software. Finally, the accuracy and uncertainty of the model were evaluated by overall accuracy and confusion index, respectively.

Results

Results showed that the different models had the same ability for prediction of the soil classes across all taxonomic levels but a considerable decreasing trend was observed for their accuracy at subgroup and family levels. The terrain attributes were the most important auxiliary information to predict the soil classes up to the family level. The main goal of the second study was to predict soil surface properties (pH, electrical conductivity, gypsum, organic carbon, calcium carbonate equivalent, coarse fragments, and particle size distribution) using ANNs, BRT, generalized linear model (GLM), and multiple linear regression (MLR).

Among the studied models, GLM showed the highest performance to predict most soil properties whereas the best model is not necessarily able to make an accurate estimation. Also, the terrain attributes were the most important environmental covariates to predict the soil classes in all taxonomic levels, but they could not display the soil variation entirely. This shows that the unexplained variations are controlled by unobserved variations in the environment, which can be due to the management over time. Results suggested that the DSM approaches have not enough prediction accuracy for the soil classes at lower taxonomic levels that focus on the soil properties affecting land use and management. Results showed that the entry of more details in the soil classification at the lower levels of the Soil Taxonomy system while increasing the number of classes, leads to decreasing the overall accuracy and increasing uncertainty. It is noticeable that the ANNs model has a good accuracy up to the great group level through the acceptable level of overall accuracy (i.e., 75 %), hence it has a high degree of uncertainty. Therefore, the accuracy of the model could not be effective in its selection through the modeling process; however, paying attention to its uncertainty is also very important along with the model error.

Conclusion

Terrain attributes were the main predictors among different studied auxiliary information. The accuracy of the estimations with more observations is recommended to give a better understanding about the performance of DSM approach over low-relief areas. Further studies may still be required to distinguish new environmental covariates and introduce new tools to capture the complex nature of soils. Accordingly, we suggest using the other methods of soft computing for modeling in plain areas or low relief regions. Finally, the use of DSM methods is increasing over time and will eventually be considered as distinct and novel techniques.

Keywords: Artificial Neural Network, Heris, Linear regression, Prediction models, Soil classes

Data Availability Statement

Data is available on reasonable request from the authors.

Acknowledgements

This paper is published as a part of a Master's thesis supported by the Vice Chancellor for Research and Technology of the Urmia University, Iran. The authors are thankful to the Urmia University for financial supports.

Conflict of interest

The authors declare no conflict of interest.

Ethical considerations

The authors avoided data fabrication, falsification, plagiarism, and misconduct.

Cite this article: Mahmoudi, M., Momtaz H.M., Servati, M. & Mohammadi, H. (2024). Digital soil mapping by machine learning techniques. *Journal of Soil and Plant Science*, 34(4), 1–14.

<https://doi.org/10.22034/sps.2024.19179>

This work is licensed under a Creative Commons Attribution-NonCommercial 4.0 International License.



Copyright © 2024 The Authors.
Publisher: The University of Tabriz



مقاله پژوهشی

نقشه‌برداری رقومی خاک با استفاده از تکنیک‌های یادگیری ماشین

معصومه محمودی^۱، حمیدرضا ممتاز^۲✉، مسلم ثروتی^۳، حسن محمدی^۴

۱- گروه علوم و مهندسی خاک، دانشکده کشاورزی، دانشگاه ارومیه، ارومیه، ایران. رایانامه: masume.shno@gmail.com

۲- نویسنده مسئول، گروه علوم و مهندسی خاک، دانشکده کشاورزی، دانشگاه ارومیه، ارومیه، ایران. رایانامه: amidmomtaz@gmail.com

۳- گروه ترویج کشاورزی و فضای سبز، مرکز آموزش عالی شهید باکری میاندوآب، دانشگاه ارومیه، ارومیه، ایران.

رایانامه: m.sarvati@urmia.ac.ir

۴- گروه علوم و مهندسی خاک، دانشکده کشاورزی، دانشگاه شهرکرد، شهرکرد، ایران. رایانامه: hassan.mohammadi1369@gmail.com

تاریخ بازنگری: ۱۴۰۳/۰۳/۲۶

تاریخ دریافت: ۱۴۰۳/۰۲/۲۶

تاریخ انتشار: ۱۴۰۳/۱۰/۰۱

تاریخ پذیرش: ۱۴۰۳/۰۳/۳۰

چکیده

کاستی‌های روش‌های کلاسیک، ابداع سامانه اطلاعات جغرافیایی و تکنیک‌های سنجش از دور، ضرورت استفاده از نقشه‌برداری رقومی خاک را دوچندان نموده است. پژوهش حاضر برای بررسی توانایی تکنیک‌های یادگیری ماشین در توصیف پراکنش خاک‌ها در منطقه‌ای با وسعت حدود ۵۰۰۰ هکتار در غرب شهرستان هریس استان آذربایجان شرقی انجام شد. در این پژوهش از داده‌های بانک خاک، شامل ویژگی‌های فیزیکی و شیمیایی ۵۰ خاکرخ و ۵۰ مته که با استفاده از روش طبقه‌بندی تصادفی، حفر و تشریح شده بودند، استفاده شد. نتایج نشان داد که برای تمامی مدل‌های مورد مطالعه (رگرسیون درختی توسعه یافته، درخت تصمیم‌گیری تصادفی و شبکه‌های عصبی مصنوعی)، با پایین رفتن سطح رده‌بندی (از رده به گروه بزرگ)، مقادیر صحت عمومی کاهش یافت. از میان مدل‌های انتخابی، مدل رگرسیون درختی تصمیم یافته بالاترین کارایی را برای تخمین اکثر ویژگی‌های مورد مطالعه داشت، اما مناسب‌ترین مدل برای تخمین ویژگی‌های خاک، به‌طور حتم نمی‌تواند تخمین درستی از آن ویژگی‌های اراضی داشته باشد. از سوی دیگر، اگرچه مدل‌های مختلف از ویژگی‌های محیطی متفاوتی برای تخمین استفاده نموده‌اند، ولی اجزای اراضی، توانایی زیادی در تخمین ویژگی‌های خاک حتی در اراضی مسطح داشته‌اند. نتیجه‌گیری جامع و قطعی در مورد روش‌های نقشه‌برداری رقومی برای تخمین ویژگی‌های خاک در مناطق مسطح دارای ابهام است. شایان ذکر است که تخمین صحیح می‌تواند متأثر از تغییرپذیری ویژگی‌های خاک، مدل تخمین، تعداد نمونه‌های صحرایی و توانایی ویژگی‌های محیطی کاربردی در بیان تغییرات سطوح مختلف رده‌بندی باشد.

واژه‌های کلیدی: رگرسیون خطی، شبکه‌های عصبی مصنوعی، کلاس‌های خاک، مدل‌های تخمینی، هریس

استناد به این مقاله: محمودی، م.، ممتاز، ح.، ثروتی، م. و محمدی، ح. (۱۴۰۳). نقشه‌برداری رقومی خاک با استفاده از تکنیک‌های

یادگیری ماشین. نشریه دانش خاک و گیاه، ۳۴(۴)، ۱-۱۴.

<https://doi.org/10.22034/sps.2024.19179>

مقدمه

جهان با چالش‌هایی مانند افزایش جمعیت، امنیت غذایی، تغییر آب و هوا، تخریب زمین، کمبود منابع آبی و پایداری زیست‌بوم مواجه است که همه این چالش‌ها بر کمیت و کیفیت محصولات کشاورزی تولید شده از خاک اثر دارند. خاک به‌عنوان یکی از اجزای زیست‌بوم، نقش اساسی در زنجیره‌های غذایی، انرژی و آب ایفا می‌کند (Zhu et al., 1997; Lal et al., 2017). همچنین، انسان برای تامین نیازهای کشاورزی خود به خاک نیاز دارد. با در نظر گرفتن چنین نقش‌های مهم، مدیریت خاک برای تولید بلند مدت، پایداری و سلامت، اجتناب‌ناپذیر است. برای این کار کارشناسان حوزه کشاورزی، نیازمند شناسایی دقیق و توزیع مکانی ویژگی‌های خاک در بسیاری از فعالیت‌های زراعی و زیست‌محیطی هستند (Wadoux et al., 2019).

یکی از روش‌های نوین ارائه شده در شناسایی خاک، نقشه برداری رقومی خاک^۱ نام دارد. نقشه برداری رقومی خاک در حال حاضر به‌طور گسترده برای تخمین کلاس‌ها و ویژگی‌های خاک کاربرد دارد (Arrouays et al., 2014). نقشه رقومی خاک یک شبکه از پیکسل‌ها می‌باشد که به‌صورت شبکه‌ای سازماندهی شده است و هر پیکسل یک موقعیت جغرافیایی حاوی داده‌های خاک است. از روش نقشه برداری رقومی خاک برای تهیه نقشه‌های اولیه خاک، به روزرسانی اطلاعات خاک، تصمیم‌گیری برای کاربری‌های آبی خاک و ارزیابی خطرهای زیست‌محیطی در منطقه می‌توان استفاده کرد (Care et al., 2007). افزایش توان محاسباتی رایانه‌ها، توسعه سیستم اطلاعات جغرافیایی (GIS) و افزایش روزافزون میزان تنوع اطلاعات جغرافیایی، باعث پیشرفت‌هایی قابل ملاحظه در تولید نقشه‌های رقومی خاک شده است. نقشه‌های رقومی خاک بر مبنای برنامه‌های کشوری در مقیاس‌های مختلف تهیه می‌شوند که عموماً توسط چارچوب فنون نقشه برداری رقومی خاک (McBratney et al., 2003) تدوین شده است. ایشان با پیروی از ایده‌های داکوچائف و ینی مدل اسکورپن را به‌عنوان روابط کمی تجربی یک مشخصه خاک و عوامل تشکیل آن توصیف کردند (Minasny and McBratney, 2007; Padarian et al., 2019). اساس روش‌های نقشه برداری رقومی مبتنی بر رابطه اسکورپن (رابطه ۱) است که در این رابطه، Sc: کلاس خاک، s: داده‌های مزرعه‌ای و آزمایشگاهی در نقاط یا در نقشه‌های خاک، c: اقلیم، o: گیاهان، جانوران و انسان‌ها، t: توپوگرافی، p: مواد مادری، a: فاکتور زمان و n: موقعیت مکانی می‌باشد. تابع f، یک برنامه یادگیری نظارت‌شده را نشان می‌دهد که برای تخمین کلاس یا ویژگی‌های خاک می‌توان به‌کار گرفت (McBratney et al., 2003).

$$Sc = (s, c, o, r, p, a, n) \quad [۱]$$

در پهنه‌بندی رقومی خاک روش‌های مختلفی از جمله: توابع انتقالی، زمین آمار، مدل‌های آماری تجربی، طبقه‌بندی فازی و تکنیک یادگیری ماشین استفاده می‌شوند (MacMillan et al., 2004). هر پژوهشگر برای تهیه نقشه رقومی خاک می‌تواند از روش‌های مختلف استفاده کند. روش‌های زیادی در نقشه برداری رقومی خاک وجود دارد که عمدتاً در گروه روش‌های یادگیری ماشین قرار دارند. یادگیری ماشینی^۲ به‌عنوان روشی برای تجزیه و تحلیل داده‌ها و زمینه‌ای از هوش مصنوعی، از تکنیک‌های داده‌کاوی برای یادگیری و ساختن یک مدل استفاده می‌کند (Khaledian and Miller, 2020). (2014) Heung et al. از روش درخت رگرسیون برای نقشه برداری مواد مادری خاک در مقیاس منطقه‌ای با استفاده از نقشه خاک موجود و نقشه ارتفاعی رقومی استفاده و مشاهده کردند که دقت نقشه تولید شده داری ضریب کاپای ۷۹/۵ درصد با اعتبارسنجی از طریق ۳۰۷ نقطه و ۸۹/۹ درصد با اعتبارسنجی از طریق نقشه قدیمی داشت. نتایج آن‌ها درخت رگرسیون را به‌عنوان یک روش ماشین یادگیری و داده‌کاوی با دقت بالا نشان داد. (2014) Adhikari et al. نیز با استفاده از روش تصمیم‌گیری درختی استفاده از ۱۷ متغیر محیطی کمکی و ۱۱۷۰ خاکرخ، نقشه کلاس‌های خاک کشور دانمارک را تهیه کردند. (2012) Rodriguez-Galiano et al. نیز درختان تصمیم‌گیری تصادفی را برای طبقه‌بندی پوشش گیاهی یک منطقه در اسپانیا استفاده کردند. نتایج آن‌ها نشان داد که دقت طبقه‌بندی با این روش ۹۲ درصد بود. (2012) Lie et al.

1- Digital soil mapping (DMS)

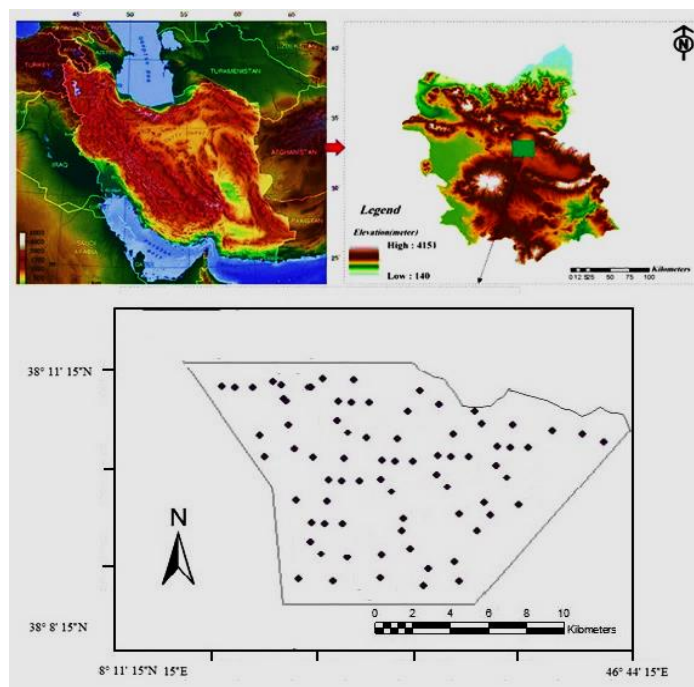
2- Machine learning

درختان تصمیم‌گیری تصادفی و درخت رگرسیون را برای تخمین توزیع بافت خاک استفاده کرده و مشاهده کردند که روش درختان تصمیم‌گیری تصادفی دارای دقت بالاتری نسبت به روش درخت رگرسیون بود. (Camera et al. (2017) با استفاده از دو مدل رگرسیون لجستیک چند متغیره و جنگل‌های تصادفی اقدام به تهیه نقشه رقومی کشور قبرس کردند که نتایج آن‌ها نشان داد که مدل جنگل‌های تصادفی نسبت به مدل رگرسیون لجستیک چند متغیره از توانایی بالاتری برای تخمین کلاس‌های خاک منطقه داشته و همچنین، این تخمین با خطای کمتری نسبت به مدل دیگر همراه بوده است. (Pahlavanrad et al. (2014) از نقشه برداری رقومی برای بروز رسانی نقشه‌های خاک استفاده کردند. در این مطالعه از روش نمونه برداری مکعب لاتین مشروط و تکنیک درختان تصمیم‌گیری تصادفی جهت تخمین نقشه خاک در سطوح گروه و زیرگروه استفاده شد. نتایج نشان داد نقشه بروز شده سری خاک ۱۳/۴ درصد دقیق‌تر از نقشه سری خاک معمولی بوده است. هدف از این تحقیق بررسی توانایی تکنیک‌های یادگیری ماشین در توصیف پراکنش خاک‌ها در غرب شهرستان هریس آذربایجان شرقی می‌باشد. همچنین، کارایی مدل‌های مختلف رقومی برای تخمین کلاس‌های خاک با استفاده از شاخص‌های ارزیابی صحت مورد بررسی قرار گرفت. تخمین ویژگی‌های خاک سطحی (صفر تا ۳۰ سانتی‌متری) نیز با استفاده از مدل‌های شبکه‌های عصبی مصنوعی، رگرسیون درختی توسعه یافته، مدل خطی تعمیم‌یافته و رگرسیون خطی چندمتغیره انجام شد.

مواد و روش‌ها

منطقه مورد مطالعه

منطقه مورد مطالعه با وسعتی حدود ۵۰۰۰ هکتار، در شمال شرق تبریز و غرب شهرستان هریس قرار دارد. این منطقه در محدوده $38^{\circ} 11' 15''$ تا $38^{\circ} 7' 30''$ عرض شمالی و $46^{\circ} 39' 30''$ تا $46^{\circ} 44' 30''$ طول شرقی واقع شده است. دارای رژیم حرارتی مزیک و رژیم رطوبتی اریدیک هم‌مرز با زیرک است (شکل ۱).



شکل ۱- موقعیت منطقه مورد مطالعه همراه با نقاط نمونه برداری.

داده‌های خاک

داده خاک می‌تواند یا از طریق مشاهدات صحرایی، تجزیه‌های آزمایشگاهی و پایگاه داده خاک به دست آید. در این پژوهش از داده‌های اخذ شده از ۵۰ خاکرخ و ۵۰ مته حفر شده (به صورت تصادفی) در منطقه مورد مطالعه استفاده شد.

اجزای سرزمین

بر پایه مدل رقومی ارتفاع با تفکیک مکانی ۳۰ متر که از وبسایت مدل رقومی ارتفاع جهانی استر تهیه شده بود، ویژگی‌های درصد شیب^۱، جهت شیب^۲، انحناي خالص^۳، انحناي نیمرخ^۴، انحناي سطحی^۵، جهت جریان^۶، تجمع جریان^۷، مستقیم^۸، مدت تابش^۹، تابش پخشیده^{۱۰}، شاخص قدرت جریان^{۱۱}، شاخص خیسی^{۱۲} و شاخص همواری دره با درجه تفکیک بالا^{۱۳} با استفاده از نرم‌افزار SAGA تعیین شدند.

متغیرهای کمکی محیطی

متغیرهای کمکی محیطی در واقع نماینده عوامل خاکسازي هستند. در این پژوهش از شاخص گیاهی تفاضلی نرمال شده (NDVI)، شاخص رس^{۱۴}، شاخص کربنات^{۱۵}، شاخص گیاهی عمودی (PVI) حاصل از تصاویر ماهواره‌ی لندست ۸ (در ۷ نوامبر ۲۰۱۸)، استفاده شد.

تعیین اندازه پیکسل پارامترهای محیطی

برای تهیه نقشه‌های رقومی نیمه تفصیلی، مطابق با اندازه پیکسلی پیشنهاد شده برای نقشه برداری در سطح نیمه تفصیلی، تمامی پارامترهای محیطی به نقشه‌های شبکه‌ای رستری با اندازه پیکسل ۵۰ متر تبدیل شدند.

نقشه برداری رقومی و مدل سازی

این مرحله از مهم‌ترین مراحل نقشه برداری رقومی می‌باشد، چرا که در این مرحله ارتباط بین داده‌های خاک و داده‌های کمکی محیطی برقرار می‌گردد و در نهایت بعد از این مرحله می‌توان تخمین رده‌های خاک را انجام داد. برای پهنه‌بندی کلاس‌های خاک تا سطح گروه بزرگ، داده‌های کلاس خاک (گروه بزرگ) کدگذاری شده و به عنوان متغیر وابسته در نظر گرفته شد. در مرحله بعد این لایه‌ها ابتدا به نقاط نمونه برداری شده و سپس به کل منطقه مورد مطالعه تعمیم داده شد و در نهایت تمامی این لایه‌های اطلاعاتی شامل داده‌های خاک و داده‌های کمکی ذخیره و مدل سازی شد.

روش‌های مدل سازی در تخمین ویژگی‌های خاک

برای تخمین ویژگی‌های خاک، رگرسیون درختی توسعه یافته، مدل خطی تعمیم یافته، مدل‌های شبکه عصبی مصنوعی، و رگرسیون خطی چندمتغیره مورد بررسی قرار گرفتند. برای انتخاب مناسب‌ترین مجموعه از ویژگی‌های محیطی، در رابطه با مدل خطی تعمیم یافته و رگرسیون خطی چندمتغیره، معنادار بودن اهمیت متغیرهای محیطی (Aksoy et al., 2012)، میزان همبستگی بین ویژگی‌های خاک با پارامترهای محیطی (Padarian et al., 2012) و استفاده از بسته نرم‌افزاری leaps و تابع Regsubsets (Lumley, 2009) مد نظر قرار گرفتند. اهمیت نسبی پارامترهای محیطی

- 1- Slope
- 2- Aspect
- 3- Curvature
- 4- Profile curvature
- 5- Plan curvature
- 6- Flow direction
- 7- Flow accumulation
- 8- Direct radiation
- 9- Duration radiation
- 10- Diffuse radiation
- 11- Stream power index
- 12- Wetness index
- 13- Multi-resolution valley bottom flatness index
- 14- Clay index
- 15- Carbonate index

(Friedman, 2001) معیار انتخاب مهم‌ترین پارامترهای ورودی برای مدل‌های غیرخطی شبکه‌های عصبی مصنوعی و رگرسیون درختی توسعه یافته بود. مدل‌سازی با استفاده از بسته‌های نرم‌افزاری مختلف gbm و nnet در نرم‌افزار R انجام شد. در نهایت، نقشه‌هایی با اندازه پیکسلی ۵۰ متر تهیه گردید.

ارزیابی صحت مدل‌ها در تخمین ویژگی‌های خاک

مدل‌های مورد مطالعه با ۸۰ درصد داده‌ها (۸۰ خاکرخ) آموزش داده شدند و اعتبارسنجی با استفاده از ۲۰ درصد داده‌ها (۲۰ خاکرخ) انجام شد. تفاوت بین مقادیر پیش‌بینی شده و مقادیر واقعی مشاهدات با استفاده از ضریب تبیین (R^2) و خطای جذر میانگین مربعات^۱ (RMSE) تعیین شدند.

روش‌های مدل‌سازی در تخمین کلاس‌های خاک

برای این کار مدل‌های رگرسیون درختی توسعه یافته، درختان تصمیم‌گیری تصادفی و شبکه‌های عصبی مصنوعی برای تخمین کلاس‌های خاک استفاده شد. ابتدا کلاس‌های خاک در سطوح مختلف رده‌بندی آمریکایی (رده، زیررده، گروه بزرگ) همراه با پارامترهای محیطی تعریف شدند. سپس، بر اساس ارتباط کلاس‌های خاک با پارامترهای محیطی، تخمین انجام شد. مدل‌سازی با استفاده از بسته نرم‌افزاری Caret موجود در نرم‌افزار R انجام شد. برای تمامی مدل‌ها، با استفاده از شاخص اثر نسبی سهم هر متغیر در تخمین تعیین و ویژگی‌های مؤثر برای مدل‌سازی انتخاب شدند.

ارزیابی صحت مدل‌ها در تخمین کلاس‌های خاک

مدل‌های مورد مطالعه با ۸۰ درصد داده‌ها (۸۰ پروفیل) توسعه یافتند و اعتبارسنجی با استفاده از ۲۰ درصد داده‌ها (۲۰ خاکرخ) انجام شد. ارزیابی صحت تخمین با استفاده از ماتریس خطا (Congalton, 1991) و شاخص‌های صحت عمومی و کاپای تعدیل‌شده انجام شد (Byrt et al., 1993).

نتایج و بحث

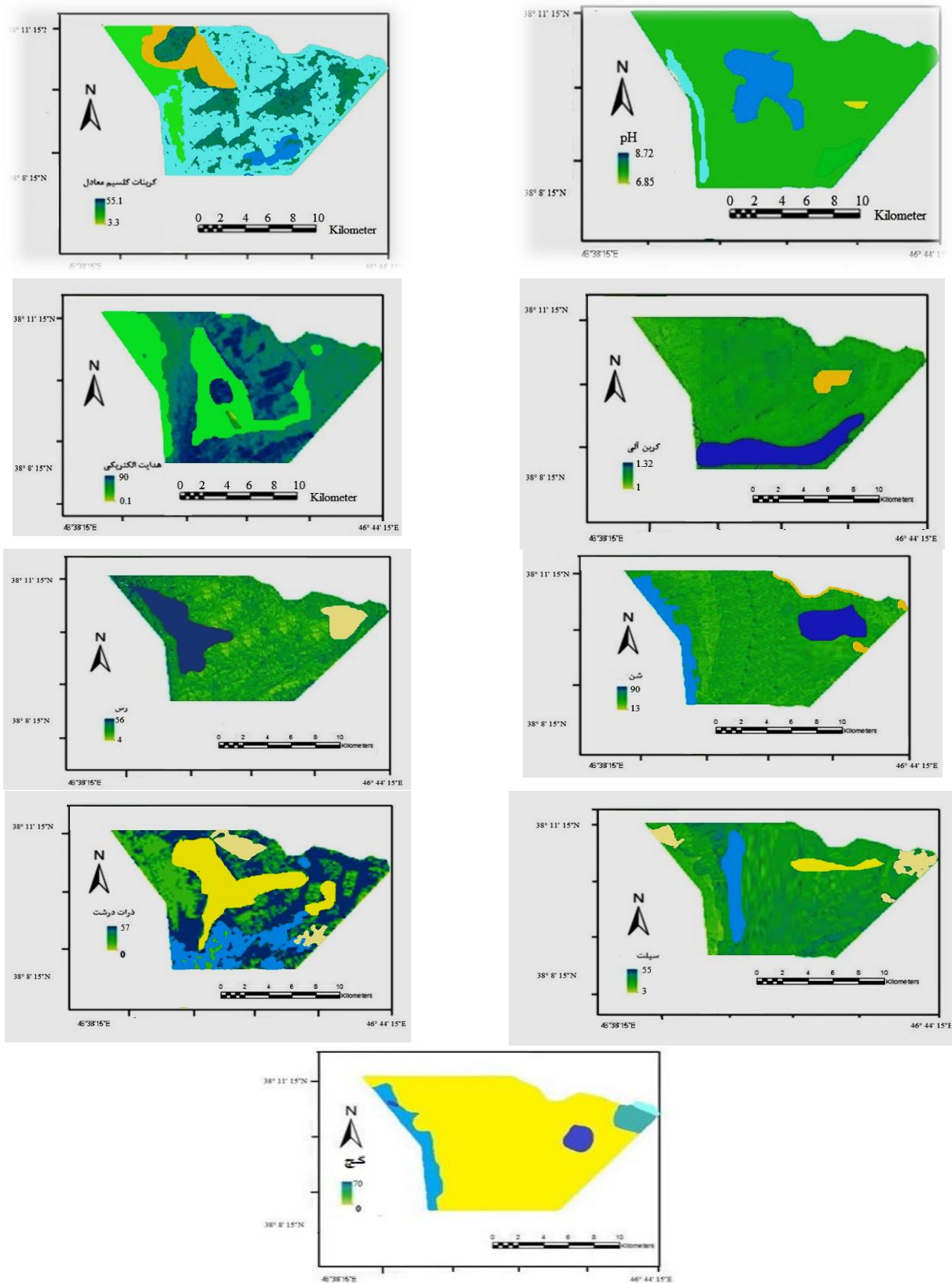
تخمین ویژگی‌های خاک

از میان مدل‌های مورد مطالعه (رگرسیون درختی توسعه یافته، خطی تعمیم یافته، رگرسیون خطی و شبکه‌های عصبی مصنوعی)، خطی تعمیم یافته بالاترین توانایی را در تخمین pH، گچ، قابلیت هدایت الکتریکی، رس، سیلت، شن و کربنات کلسیم معادل داشت. مناسب‌ترین مدل برای تخمین میزان کربن آلی و ذرات درشت، به ترتیب مدل‌های رگرسیون درختی توسعه یافته و شبکه‌های عصبی مصنوعی بودند (جدول ۱).

جدول ۱- نتایج ارزیابی مدل‌های مختلف برای تخمین ویژگی‌های خاک.

مدل	معیار ارزیابی	pH	قابلیت هدایت الکتریکی	کربن آلی	کربنات کلسیم معادل	رس	سیلت	شن	گچ	ذرات درشت
رگرسیون درختی توسعه یافته	RMSE	۱۰	۱۴	۱۸	۱۵	۲۶	۲۵	۲۳	۲۲	۲۴
	R^2	۰/۹	۰/۷۲	۰/۸۳	۰/۷۷	۰/۴۴	۰/۶۳	۰/۶۶	۰/۷۷	۰/۵۵
مدل خطی تعمیم یافته	RMSE	۹	۱۱	۱۹	۱۳	۲۲	۲۱	۲۰	۱۹	۲۱
	R^2	۰/۹۳	۰/۸۸	۰/۸	۰/۸۶	۰/۷۵	۰/۷۷	۰/۷۹	۰/۸۲	۰/۷۶
رگرسیون خطی	RMSE	۱۴	۱۷	۲۲	۱۹	۲۶	۲۴	۲۳	۲۲	۲۳
	R^2	۰/۷۲	۰/۶۷	۰/۷	۰/۶۸	۰/۵۷	۰/۶۵	۰/۶۵	۰/۷۶	۰/۵۹
شبکه‌های عصبی مصنوعی	RMSE	۱۰	۱۴	۲۰	۱۴	۲۴	۲۴	۲۲	۲۲	۱۹
	R^2	۰/۹	۰/۷۲	۰/۷۹	۰/۷۳	۰/۵۵	۰/۶۵	۰/۶۹	۰/۷۱	۰/۸۱

^۱-Root mean square error



شکل ۲- نقشه های پراکنش مکانی برای تمام ویژگی های خاکی مؤثر بر تخمین کلاس ها.

شاخص‌های سنجش از دور از قبیل شاخص کربنات و شاخص گچ می‌توانند برای بررسی تغییرات مواد مادری مؤثر باشند. برای مدل‌سازی کربن آلی خاک با استفاده از مدل رگرسیون درختی توسعه یافته، مقادیر بهینه برای پارامترهای پیچیدگی درخت، مقدار یادگیری، سهم داده‌های آموزشی و تعداد درختان به ترتیب ۲/۹، ۰/۰۰۱، ۰/۴۸ و ۱۸۸۵ بودند. همچنین، برخی پارامترهای محیطی مانند تابش پخشیده، ارتفاع و شاخص کربنات، بیشترین اثر را در پیش‌بینی کربن آلی خاک داشتند.

از طرفی نتایج نشان داد که بین احتمال تخمین کربن آلی خاک با تابش پخشیده، ارتفاع و شاخص کربنات، روند معکوس وجود داشت. نتایج نشان داد که مدل شبکه عصبی مصنوعی، شاخص همواری دره با درجه‌ی تفکیک بالا، شاخص گیاهی عمودی و نقشه ژئومورفولوژی تا سطح شکل سرزمین را برای تخمین ذرات درشت انتخاب نمود. لذا، با این شاخص‌ها و ویژگی‌های خاک می‌توان به ترسیم نقشه‌های پیوسته ویژگی‌های خاک موفق شد. شکل ۲ نقشه تخمینی ویژگی‌های اراضی را در منطقه مورد مطالعه نشان می‌دهد. در بین ویژگی‌های اراضی ۶ ویژگی شن، سیلت، رس، کربنات کلسیم، گچ، کربن آلی، pH خاک و هدایت الکتریکی بیشترین اثر را در تخمین کلاس‌های خاک داشتند.

نتایج نشان داد که بیشترین و کمترین صحت تخمین، به ترتیب برای pH با استفاده از مدل رگرسیون خطی تعمیم‌یافته و ذرات درشت با استفاده از مدل شبکه‌های عصبی مصنوعی به دست آمد. مقادیر RMSE کمتر از ۴۰، حاکی از تخمین صحیح و مقادیر بزرگ‌تر از ۷۰ به معنای آن است که پیش‌بینی دارای عدم قطعیت زیاد می‌باشد (Minasny and McBratney, 2016). بر این اساس، مدل رگرسیون خطی تعمیم‌یافته، تخمین صحیحی برای پیش‌بینی pH خاک، قابلیت هدایت الکتریکی، درصد رس، سیلت، شن، گچ و کربنات کلسیم معادل داشت و این در حالی است که مناسب‌ترین مدل برای تخمین ذرات درشت (شبکه‌های عصبی مصنوعی) نتوانست تخمین صحیحی را برای این ویژگی انجام دهد (جدول ۲). بنابراین، نتایج حاکی از آن است که مناسب‌ترین مدل برای تخمین ویژگی‌های خاک الزاماً نمی‌تواند تخمین صحیحی از آن ویژگی داشته باشد و این نکته‌ای است که در مسایل مدل‌سازی نباید از خاطر دور نگاه داشته شود.

جدول ۲- مقادیر جذر میانگین مربعات خطای نسبی با استفاده از مناسب‌ترین مدل برای ویژگی‌های خاک.

ویژگی‌های خاک	مدل مناسب	RMSE (%)	ویژگی‌های خاک	مدل مناسب	RMSE (%)
pH	رگرسیون خطی تعمیم‌یافته	۱۳	ذرات درشت	شبکه‌های عصبی مصنوعی	۷۵
قابلیت هدایت الکتریکی	رگرسیون خطی تعمیم‌یافته	۲۹	شن	رگرسیون خطی تعمیم‌یافته	۱۹
کربن آلی	رگرسیون درختی تعمیم یافته	۳۸	سیلت	رگرسیون خطی تعمیم‌یافته	۱۹
کربنات کلسیم معادل	رگرسیون خطی تعمیم‌یافته	۲۵	رس	رگرسیون خطی تعمیم‌یافته	۲۱
گچ	رگرسیون خطی تعمیم‌یافته	۲۶			

نتایج حاکی از آن است که مدل رگرسیون خطی تعمیم‌یافته با انتخاب تجمع جریان، شاخص گیاهی عمودی، شاخص کربنات و شاخص خیزی توانسته است مناسب‌ترین تخمین برای میزان کربنات کلسیم معادل خاک را داشته

باشد. توپوگرافی، به‌عنوان یکی از فاکتورهای خاکسازي بر شرایط هیدرولوژیکی، الگوهای جریان و میزان آب موجود در خاک اثرگذار است و تمامی این عوامل می‌توانند میزان آبتشویی کربناتهای خاک را کنترل نمایند. Alijani and Sarmadian (2014) با مطالعه خاک‌های منطقه کوهین قزوین گزارش کردند که مقادیر کربنات کلسیم معادل خاک با شاخص خیسی و انحناي نیمرخ ارتباط معناداری داشتند. همچنین، با توجه به اینکه مواد مادری منطقه به‌طور عمده شامل رسوباتی از جنس آهک بودند، شاخص کربنات می‌تواند در صحت تخمین میزان کربنات کلسیم معادل اثرگذار باشد. گزارش شده است که شاخص‌های سنجش از دور کربنات و گچ می‌توانند برای بررسی تغییرات مواد مادری مؤثر باشند (Boettinger et al., 2008; Mosleh et al., 2016).

تخمین کلاس‌های خاک

مقادیر کاپای بزرگ‌تر از ۰/۸، ۰/۸ تا ۰/۴ و کمتر از ۰/۴، به‌ترتیب نشان‌دهنده تطابق شدید، متوسط و کم بین مقادیر تخمینی و اندازه‌گیری شده می‌باشد (Congalton and Green, 1998). بنابراین، با توجه به مقادیر شاخص کاپا و کاپای تعدیل‌شده، بین کلاس‌های خاک در سطح رده، تحت‌رده و گروه بزرگ به‌ترتیب مطابقت شدید، متوسط تا شدید و متوسط وجود دارد و این در حالی است که این تطابق در سطح زیرگروه و خانواده ضعیف می‌باشد (جدول ۳). برای هر دو مدل مورد مطالعه، با افزایش سطح رده‌بندی، مقادیر صحت عمومی کاهش یافت و این کاهش در سطوح زیرگروه و خانواده چشمگیرتر بود. نتیجه مشابهی به‌وسیله سایر پژوهشگران نیز گزارش شده است (Pahlavanrad et al., 2014; Taghizadeh Mehrjardi et al., 2015).

جدول ۳- ارزیابی صحت مدل‌های مختلف برای تخمین سطوح مختلف رده‌بندی.

خانواده	زیرگروه	گروه بزرگ	زیررده	رده	معیار ارزیابی	مدل
۰/۱۴	۰/۴۷	۰/۶۴	۰/۸۵	۰/۹۸	صحت عمومی	درخت تصمیم‌گیری تصادفی
-۰/۱	۰/۲	۰/۱۸	-۰/۰۹	۰/۹۸	کاپا کاپای	
-۰/۹۹	۰	۰	۰/۵۹	۰/۹۸	تعدیل‌یافته	
۰/۰۹	۰/۲۹	۰/۶۶	۰/۸۸	۰/۹۸	صحت عمومی	رگرسیون درختی توسعه یافته
۰/۰۸	۰/۰۳	۰/۲۵	-۰/۰۴	۰/۹۸	کاپا کاپای	
۰/۰۶	-۰/۰۲	۰/۳۷	۰/۷۱	۰/۹۸	تعدیل‌یافته	
۰/۰۳	۰/۲۶	۰/۴۸	۰/۸۹	۰/۹۳	صحت عمومی	شبکه‌های عصبی مصنوعی
-۰/۸۸	۰	۰	-۰/۰۲	۰	کاپا	
-۰/۸۹	-۰/۰۷	۰/۱۳	۰/۷۵	۰/۹۱	کاپای تعدیل‌یافته	

نتایج حاکی از آن است که صحت تخمین در سطوح پایین‌تر رده‌بندی، کاهش می‌یابد و این کاهش در سطوح زیرگروه و فامیل قابل توجه می‌باشد. در سطوح پایین‌تر رده‌بندی خاک، علاوه بر آنکه ویژگی‌های بیشتری از خاک مد نظر قرار می‌گیرند، معمولاً تعداد کلاس‌های خاک در هر سطح، افزایش می‌یابد و از طرفی تعداد مشاهدات در هر کلاس خاک کم می‌شود. بنابراین، حضور کم برخی خاک‌ها سبب می‌شود که فرایند آموزش مدل به‌خوبی انجام نشود و شاید به همین دلیل است که در سطوح پایین‌تر رده‌بندی، مقادیر صحت عمومی کاهش می‌یابد و قابلیت اعتماد به نتایج کمتر می‌شود.

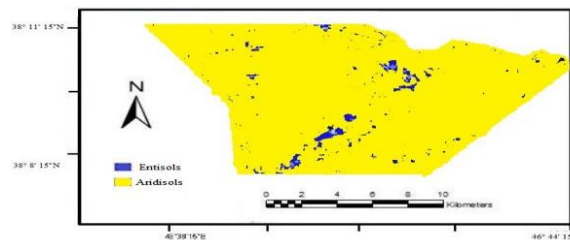
که نتایج مشابهی به وسیله سایر پژوهشگران نیز گزارش شده است (Jafari et al., 2013; Bagheri Bodaghabadi et al., 2015).

نتایج این پژوهش نشان داد که اجزای اراضی (جدول ۴) بیشترین اثر را در تخمین کلاس‌های خاک تا سطح گروه بزرگ داشتند و اثر آن‌ها در سطوح مختلف متفاوت بود. در سطوح پایین رده‌بندی (زیر گروه و خانواده) همه ویژگی‌های محیطی اثر یکسانی در تخمین کلاس‌های خاک داشتند که مؤید این مطلب باشد که در مناطق مسطح نیز توپوگرافی از مهم‌ترین فاکتورهای خاکساختی است که چنین نتایجی به وسیله سایر پژوهشگران نیز گزارش شده است (Mosleh et al., 2016).

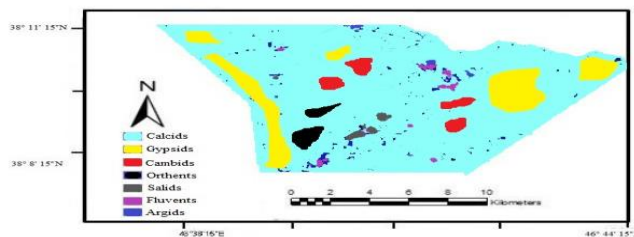
شکل‌های ۳، ۴، ۵ به ترتیب نقشه طبقه‌بندی سطوح مختلف رده‌بندی در سطح رده، زیررده و گروه بزرگ را نشان می‌دهند. نتایج مؤید این مطلب است که دقت نقشه برای سطح رده ۹۸ درصد بود و ۹۵ درصد اراضی را ارییدی‌سول‌ها و ۵ درصد را انتی‌سول‌ها به خود اختصاص دادند. دقت نقشه زیررده‌ها ۸۵ درصد بود و بیشترین درصد را زیررده کلسیدز و کمترین مقدار را زیررده فلوننتز به خود اختصاص دادند. در سطح گروه بزرگ نیز دقت نقشه‌ها ۶۴ درصد بود و بیشترین پراکنش را هاپلوکلسیدز و کمترین پراکنش را تورفلوننتزها به خود اختصاص دادند. در سطوح پایین رده‌بندی دقت نقشه کم بود و لذا به علت پایین بودن اعتبار برای آن نقشه تهیه نشد.

جدول ۴- مهم‌ترین پارامترهای محیطی جهت تخمین کلاس‌های خاک.

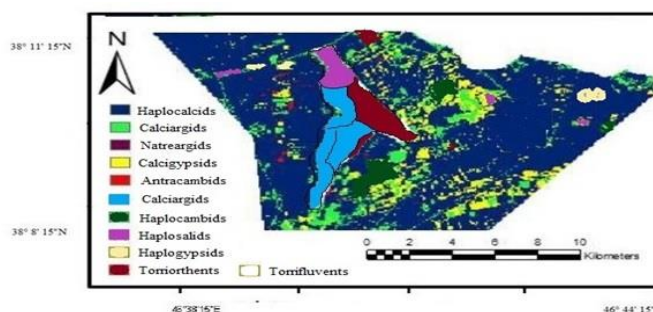
مدل	سطوح رده‌بندی	مهمترین پارامترهای محیطی
درخت تصمیم‌گیری تصادفی	رده	جهت شیب، تابش مستقیم، انحناى سطحی
	زیررده گروه بزرگ	جهت شیب، ارتفاع، تابش پخشیده، شاخص خیسی ارتفاع، تابش پخشیده
رگرسیون درختی توسعه یافته	رده	جهت شیب، انحناى سطحی
	زیررده گروه بزرگ	جهت شیب، ارتفاع، شاخص خیسی ارتفاع، تابش پخشیده، شاخص گیاهی عمودی
شبکه‌های عصبی مصنوعی	رده	شاخص گیاهی تفاضلی، انحناى سطحی
	زیررده گروه بزرگ	شیب، شاخص کربنات شاخص گیاهی عمودی، تجمع جریان



شکل ۳- نقشه رقومى خاک‌های منطقه در سطح رده.



شکل ۴- نقشه رقومی خاک‌های منطقه در سطح زیررده.



شکل ۵- نقشه رقومی خاک‌های منطقه در سطح گروه بزرگ.

نتیجه‌گیری کلی

نتایج این تحقیق نشان داد که در تمامی سطوح رده‌بندی، مدل‌های انتخابی، توانایی یکسانی در پیش‌بینی کلاس‌های خاک دارند. همچنین، نتایج حاکی از آن است که صحت پیش‌بینی در سطوح پایین‌تر رده‌بندی، کاهش می‌یابد و این کاهش در سطوح زیرگروه و فامیل قابل توجه می‌باشد. در سطوح پایین‌تر رده‌بندی خاک، علاوه بر آن‌که ویژگی‌های بیشتری از خاک مد نظر قرار می‌گیرند، معمولاً تعداد کلاس‌های خاک در هر سطح، افزایش می‌یابد و از طرفی تعداد مشاهدات در هر کلاس خاک کم می‌شود. بنابراین، حضور کم برخی خاک‌ها سبب می‌شود که فرایند آموزش مدل به خوبی انجام نشود و شاید به همین دلیل است که در سطوح پایین‌تر رده‌بندی، مقادیر صحت عمومی کاهش می‌یابد و قابلیت اعتماد به نتایج کمتر می‌شود. پس می‌توان بیان نمود که زمانی که تعداد مشاهدات زیاد، تعداد کلاس‌های خاک کم و توزیع مشاهدات بین کلاس‌های خاک تقریباً یکسان باشد می‌توان انتظار داشت که کلاس‌های پیش‌بینی شده از صحت بالایی برخوردار باشند. همچنین، نتایج نشان داد که در منطقه مورد مطالعه، روش نقشه‌برداری رقومی برای تخمین کلاس‌های خاک در سطح زیرگروه و خانواده، دارای صحت قابل قبولی نیست. صحت بالای نقشه‌های رقومی خاک در برابر میزان هزینه‌ای که برای انجام آن‌ها می‌شود؛ زمانی می‌تواند ارزشمند باشد که بتواند پیچیدگی‌های خاک را در سطوح پایین رده‌بندی خاک با دقت قابل قبولی بیان نماید. از میان مدل‌های انتخابی، مدل رگرسیون درختی تعمیم یافته بالاترین کارایی را برای تخمین اکثر ویژگی‌های مورد مطالعه دارد اما بهترین مدل برای تخمین ویژگی‌های خاک، الزاماً نمی‌تواند تخمین صحیحی از آن ویژگی داشته باشد. از سوی دیگر، اگرچه مدل‌های مختلف، پارامترهای محیطی متفاوتی را برای پیش‌بینی استفاده نموده‌اند ولی اجزای سرزمین، توانایی زیادی در پیش‌بینی ویژگی‌های خاک حتی در مناطق مسطح داشته‌اند. با این وجود، نتیجه‌گیری قطعی در رابطه با اینکه آیا روش‌های رقومی برای پیش‌بینی ویژگی‌های خاک در مناطق مسطح که سایر فاکتورهای خاکساختی نیز دارای تغییرات چندانی نمی‌باشند؛ مناسب هستند یا نه، بسیار

مشکل است. به نظر می‌رسد که تخمین صحیح ویژگی‌های خاک می‌تواند متأثر از تغییرپذیری ویژگی‌های خاک، نوع مدل، تعداد مشاهدات و توانایی پارامترهای محیطی استفاده شده در بیان تغییرات باشد. از آنجا که مدیریت اراضی می‌تواند اثر زیادی بر برخی ویژگی‌های فیزیکی و شیمیایی خاک‌ها داشته باشد احتمال دارد برخی تغییرات مکانی ویژگی‌های خاک که توسط مدل‌های مزبور قابل تشخیص نبوده است به فاکتور مدیریت باز گردد.

منابع مورد استفاده

References

- Adhikari, K., Minasny, B., Greve, M.B. & Greve, M.H. (2014). Constructing a soil class map of Denmark based on the FAO legend using digital techniques. *Geoderma*, 214, 101–113. <https://doi.org/10.1016/j.geoderma.2013.09.023>
- Aksoy, E., Panagos, P. & Montanarella, L. (2012). Spatial prediction of soil organic carbon of Crete using geostatistics. Pp. 149–159. In: Minasny B, et al. (eds.) *Digital soil assessments and beyond*. CRC Press, London.
- Alijani, Z. & Sarmadian, F. (2014). The role of topography in changing of soil carbonate content. *Indian Journal of Science and Research*, 6, 263–271.
- Arrouays, D., McKenzie, N., Hempel, J., de Forges, A.R. & McBratney, A.B. (2014). *Global soil map: basis of the global spatial soil information system*. CRC press.
- Bagheri Bodaghabadi, M., Martinez-Casasnovas, J.A., Salehi, M.H., Mohammadi, J., Esfandiarpour Borujeni, I., Toomanian, N. & Gandomkar, A. (2015). Digital soil mapping using artificial neural networks and terrain-related attributes. *Pedosphere*, 25, 580–591. [https://doi.org/10.1016/S1002-0160\(15\)30038-2](https://doi.org/10.1016/S1002-0160(15)30038-2)
- Boettinger, J.L., Ramsey, R.D., Bodily, J.M., Cole, N.J., Kienast Brown, S., Nield, S.J., Saunders, A.M. & Stum, A.K. (2008). Landsat spectral data for digital soil mapping. Pp. 193–203. In: Hartemink A.E. et al. (eds.) *Digital soil mapping with limited data*. Springer, Australia.
- Byrt, T., Bishop, J. & Carling, J.B. (1993). Bias, prevalence and kappa. *Journal of Clinical Epidemiology*, 46, 423–429. [https://doi.org/10.1016/0895-4356\(93\)90018-v](https://doi.org/10.1016/0895-4356(93)90018-v)
- Camera, C., Zomeni, Z., Noller, J.S., Zissimos, A.M., Christoforou, I.C. & Bruggeman, A. (2017). A high-resolution map of soil types and physical properties for Cyprus: A digital soil mapping optimization. *Geoderma*, 285, 35–49. <https://doi.org/10.1016/j.geoderma.2016.09.019>
- Carré, F., McBratney, A.B., Mayr, T. & Montanarella, L. (2007). Digital soil assessments: Beyond DSM. *Geoderma*, 142(1-2), 69–79. <https://dx.doi.org/10.1016/j.geoderma.2007.08.015>
- Congalton, R. (1991). A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment*, 37, 35–46. [https://doi.org/10.1016/0034-4257\(91\)90048-B](https://doi.org/10.1016/0034-4257(91)90048-B)
- Congalton, R.G. and Green, K. (1998). *Assessing the accuracy of remotely sensed data: principles and practices*. CRC Press.
- Friedman, J.H. (2001). Greedy function approximation: a gradient boosting machine. *The Annals of Statistics*, 29, 1189–1232. DOI: [10.1214/aos/1013203451](https://doi.org/10.1214/aos/1013203451)
- Heung, B., Bulmer, C.E. & Schmidt M.G. (2014). Predictive soil parent material mapping at a regional-scale: a random forest approach. *Geoderma*, 214, 141–154. <https://doi.org/10.1016/j.geoderma.2013.09.016>
- Jafari, A., Ayoubi, S., Khademi, H., Finke, P.A. & Toomanian, N. (2013). Selection of a taxonomic level for soil mapping using diversity and map purity indices: A case study from an Iranian arid region. *Geomorphology*, 201, 86–97. <https://doi.org/10.1016/j.geomorph.2013.06.010>
- Khaledian, Y & Miller, BA, (2020). Selecting appropriate machine learning methods for digital soil mapping. *Applied Mathematical Modelling*, 81, 401–418. <https://doi.org/10.1016/j.apm.2019.12.016>
- Lal, R., Mohtar, R.H., Assi, A.T., Ray, R., Baybil, H. & Jahn, M. (2017). Soil as a basic nexus tool: soils at the center of the food–energy–water nexus. *Current Sustainable/Renewable Energy Reports*, 4(3), 117–129. <https://doi.org/10.1007/s40518-017-0082-4>

- Lie, M., Glaser, B. & Huwe, B. (2012). Uncertainty in the spatial prediction of soil texture: comparison of regression tree and Random Forest models. *Geoderma*, 170, 70–79. <https://doi.org/10.1016/j.geoderma.2011.10.010>
- Lumley, T. (2009). Regression subset selection. <http://CRAN.com>.
- MacMillan, R.A., Jones, R.K. & McNabb, D.H. (2004). Defining a hierarchy of spatial entities for environmental analysis and modeling using digital elevation models (DEMs). *Computers, Environment and Urban Systems*, 28(3), 175–200. [https://doi.org/10.1016/S0198-9715\(03\)00019-X](https://doi.org/10.1016/S0198-9715(03)00019-X)
- McBratney, A.B., Santos, M.M. & Minasny, B. (2003). On digital soil mapping. *Geoderma*, 117(1-2), 3–52. [https://doi.org/10.1016/S0016-7061\(03\)00223-4](https://doi.org/10.1016/S0016-7061(03)00223-4)
- Minasny, B. & McBratney, A.B. (2007). Spatial prediction of soil properties using EBLUP with the Matérn covariance function. *Geoderma*, 140, 324–336.
- 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>
- Mosleh, Z., Salehi, M.H., Jafari, A., Esfandiarpour Borujeni, I. & Mehnatkesh, A. (2016). The effectiveness of digital soil mapping to predict soil properties over low-relief areas. *Environmental Monitoring Assessment*, 31, 188–195. <https://doi.org/10.1007/s10661-016-5204-8>
- Padarian, J., Minasny, B. & McBratney, A.B. (2019). Using deep learning for digital soil mapping. *Soil Discussion*, 5(1), 79–89. <http://dx.doi.org/10.5194/soil-5-79-2019>
- Padarian, J., Perez-Quezada, J. & Seguel, S. (2012). Modeling the distribution of organic carbon in the soils of Chile. Pp. 329–333. In: Minasny B, et al. (eds.) *Digital soil assessments and beyond*. CRC Press, London, UK. <http://dx.doi.org/10.1201/b12728-65>
- Pahlavanrad, M.R., Toomanian, N., Khormali, F., Brungard, C.W., Komaki, S.B. & Bogaert, P. (2014). Updating soil survey maps using random forest and conditioned Latin hypercube sampling in the loess derived soils of northern Iran. *Geoderma*, 232, 97–106. <https://doi.org/10.1016/j.geoderma.2014.04.036>
- Rodriguez-Galiano, V.F., Ghimire, B., Rogan, J., Chica-Olmo, M. & Rigol-Sanchez, J.P. (2012). An assessment of the effectiveness of a random forest classifier for land-cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, 67, 93–104. <https://doi.org/10.1016/j.isprsjprs.2011.11.002>
- Taghizadeh Mehrjardi, R., Nabiollahi, K., Minasny, B. & Triantafilis, J. (2015). Comparing data mining classifiers to predict spatial distribution of USDA-family soil groups in Baneh region, Iran. *Geoderma*, 253–254, 67–77. <https://doi.org/10.1016/j.geoderma.2015.04.008>
- Wadoux, A.M.C. (2019). Using deep learning for multivariate mapping of soil with quantified uncertainty. *Geoderma*, 351, 59–70. <https://doi.org/10.1016/j.geoderma.2019.05.012>
- Zhu, A.X., Band L., Vertessy, R. & Dutton, B. (1997). Derivation of soil properties using a soil land inference model (SoLIM). *Soil Science Society of America Journal*, 61(2), 523–533. <https://doi.org/10.2136/sssaj1997.03615995006100020022x>