

اثر مقیاس متغیرهای محیطی در نقشه برداری رقومی ویژگی‌های خاک در یک منطقه نیمه خشک غرب ایران

زیبا مقصودی^۱، حمیدرضا متین فر^{۱*}، سید روح اله موسوی^۲

۱- دانشجوی دکتری مدیریت منابع خاک - منابع خاک و ارزیابی اراضی

۱- استاد گروه علوم و مهندسی خاک دانشکده کشاورزی دانشگاه لرستان؛ matinfar.h@lu.ac.ir

۲- دکتری مدیریت منابع خاک- منابع خاک و ارزیابی اراضی

چکیده

تهیه نقشه‌های دقیق از ویژگی‌های خاک برای ارزیابی جنبه‌های متعدد مدیریت اراضی کشاورزی از جمله پتانسیل تولید و ذخایر کربن آلی ضروری است. هدف از این مطالعه بهبود دقت نقشه برداری از طریق انتخاب مقیاس بهینه برای پیش‌بینی شش ویژگی خاک شامل کربن آلی (SOC)، کربنات کلسیم معادل (CCE)، pH، شن، سیلت و رس در منطقه خرم آباد است. برای این منظور تعداد ۱۵۷ نمونه در منطقه خرم آباد از عمق ۰-۳۰ سانتی متری و براساس الگوی نمونه برداری تصادفی تهیه گردید و برای اندازه‌گیری ویژگی‌ها به آزمایشگاه منتقل شدند. متغیرهای محیطی توپوگرافی و سنجش از دور از مدل رقومی ارتفاع (DEM) و ماهواره لندست ۸ استخراج شدند. متغیرهای محیطی بهینه توسط روش باروتا در منطقه خرم آباد انتخاب شدند. مدل‌سازی ویژگی‌های خاک توسط مدل‌های یادگیری ماشین جنگل تصادفی (RF)، رگرسیون بردار پشتیبان (SVR)، کوبیست (CB) و مدل‌های ترکیبی اجرا شدند. نتایج نشان داد که در مناطق با توپوگرافی کم استفاده از مقیاس‌های میانی (اندازه سلول ۳۰ تا ۹۰ متر) منجر به پیش‌بینی ویژگی‌های خاک با دقت بالاتر شد. زیرا استفاده از اندازه‌های سلولی بالاتر نویزهایی را ایجاد می‌کنند که از دقت جلوگیری می‌کند. بنابراین، انتخاب وضوح و یا به عبارتی اندازه سلول در DSM باید متناسب با نیازهای خاص، ایجاد تعادل بین ثبت الگوهای معنادار و کارایی محاسباتی باشد.

کلمات کلیدی: نقشه برداری رقومی، متغیرهای محیطی، مقیاس بهینه، مدل‌سازی مکانی

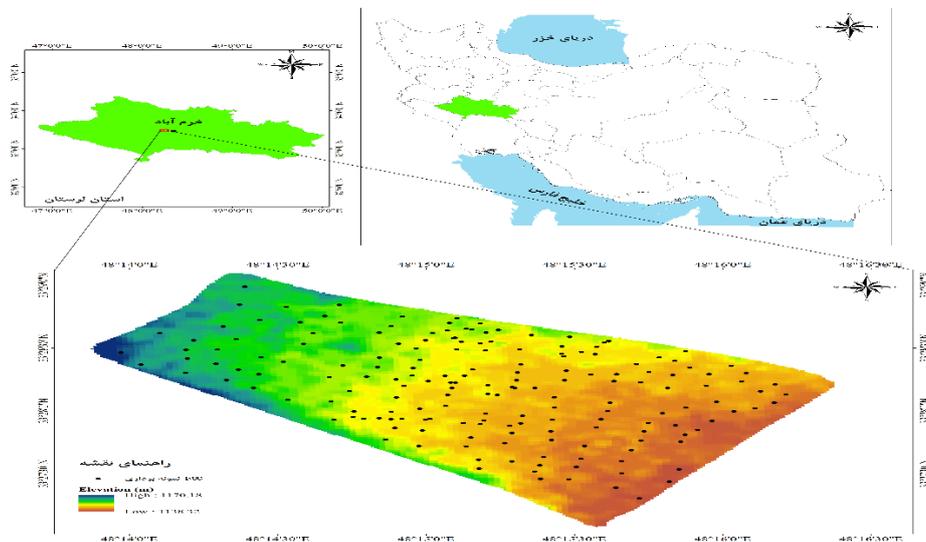
مقدمه

یکی از عوامل مؤثر بر پیش‌بینی خصوصیات و یا کلاس‌های خاک، مقیاس داده‌ها و متغیرهای پیش‌بینی‌کننده است. با این حال، به دلیل پیچیدگی‌های زیادی که تاکنون در جنبه‌های مختلف فرآیند نقشه برداری وجود داشته، این موضوع چندان مورد توجه قرار نگرفته است. در شرایط فعلی، نیاز است که به‌طور عمیق‌تر مورد بررسی و توجه قرار گیرد (Malone et al., 2017). وضوح داده‌ها یکی از مهمترین ویژگی‌هایی است که هنگام انتخاب داده باید در نظر گرفته شود. بسیاری از منابع داده با وضوح بالا در حال حاضر در دسترس هستند، اما ممکن است مشکل مورد نظر را برطرف نکنند (Hengl et al., 2014). هنگام انتخاب مناسب‌ترین وضوح، مقیاس ویژگی‌ها یا ویژگی‌های فیزیکی روی زمین‌نما باید در نظر گرفته شود. بنابراین، انواع تفکیک پذیری - مکانی، طیفی، زمانی و رادیومتری - باید در نظر گرفته شوند (Hengl et al., 2014). وضوح مکانی برای همه منابع داده اعمال می‌شود و برابر با اندازه سلول شبکه است. موفقیت DSM به‌طور قابل توجهی به کیفیت متغیرهای کمکی محیطی ورودی بستگی دارد (Zeraatpisheh et al., 2020) و در مواقعی اغلب بر انتخاب الگوریتم مدل‌سازی اولویت دارد (keshin and Grunwald, 2018). یک شاخص مهم برای کیفیت این متغیرهای کمکی محیطی ورودی، وضوح است. در DSM، متغیرهای کمکی محیطی در قالب پیکسل نشان داده می‌شوند و وضوح آن‌ها اغلب بر اساس عرض پیکسل DEM تعیین می‌شود. بسیاری از مطالعات اهمیت تحلیل چند مقیاسی زمین در DSM را برجسته کرده‌اند (Cavazzi et al., 2013)، (Behrens et al., 2018). Cavazzi و همکاران (۲۰۱۳)، در پژوهش خود با عنوان آیا مدل رقومی ارتفاع با وضوح مکانی پایین (ریز مقیاس) در نقشه برداری رقومی خاک مناسب‌ترین انتخاب می‌باشد؟ گزارش نمودند که در مناطق با تنوع توپوگرافی عمدتاً اندازه وضوح مکانی بسیار ریز تا خیلی درشت، در مناطق هموار و همگن اندازه سلولی خیلی درشت و در مناطق مختلط مقادیر حدواسط مقیاس در نقشه برداری خاک مناسب می‌باشد. میلر و همکاران (Miller et al., 2018)، در مطالعه‌ای تأثیر انتخاب پیش‌بینی‌کننده‌های

چند مقیاسی برای مدل‌سازی ویژگی‌های خاک مورد بررسی قرار دادند. در این تحقیق مجموعه‌ای از ۴۱۲ پیش‌بینی کننده بالقوه جمع‌آوری گردید که شامل ارتفاع، مشتقات سطح زمین، شاخص‌های هیدرولوژیکی و همچنین شاخص‌های سنجش از دور و نزدیک بود. نتایج نشان داد که عملکرد مدل با در نظر گرفتن متغیرهای پیش‌بینی کننده در مقیاس‌های چندگانه بهتر عمل می‌کند بر همین اساس، هدف این مطالعه شامل تعیین تأثیرگذارترین متغیرهای کمکی پیش‌بینی کننده ویژگی‌های خاک در مقیاس‌های مکانی مختلف، بررسی دقت الگوریتم یادگیری ماشین در تهیه نقشه ویژگی‌های خاک در مقیاس‌های مکانی مختلف می‌باشد.

مواد و روش‌ها

منطقه مورد مطالعه در استان لرستان حدفاصل طول جغرافیایی ۴۱۲۶۴۶ و ۶۷۳۵۷۵ متر و عرض جغرافیایی ۳۶۱۰۳۰۲ و ۳۸۰۳۲۲۱ متر در سیستم مختصات متریک و منطقه ۳۹ شمالی در خرم‌آباد واقع شده است (شکل ۱). مساحت دشت خرم‌آباد ۶۶۹ هکتار و متوسط تغییرات ارتفاع ۱۱۳۷ تا ۱۱۶۸ متر از سطح دریا می‌باشد. رده‌بندی خاک‌ها براساس سامانه رده‌بندی آمریکایی (۲۰۱۴) و نقشه خاک با مقیاس ۱/۵۰۰۰۰ موجود در دو رده اینسپتیسول و انتیسول قرار دارند. نمونه‌برداری از خاک‌ها شامل انتخاب ۱۵۷ نمونه در منطقه خرم‌آباد از عمق ۰-۳۰ سانتی‌متر و براساس الگوی نمونه‌برداری تصادفی و نظر کارشناس انجام گرفت. ویژگی‌های فیزیکی و شیمیایی منتخب در این مطالعه به ترتیب توزیع اندازه ذرات خاک (شن، سیلت و رس) به روش هیدرومتری (Gee and Bauder, 1986)، کربنات کلسیم معادل به روش تیتراسیون برگشتی (Nelson et al., 1980)، واکنش خاک (pH) در عصاره گل اشباع (McLean, 1982) و مقدار کربن آلی خاک به روش اکسیداسیون تر (Jackson, 1969) اندازه‌گیری شدند.



شکل ۱- منطقه مورد مطالعه

برای مدل‌سازی نقشه ویژگی‌های خاک منطقه از متغیرهای ژئومورفومتری و سنجش از دور استخراج شده از مدل رقومی ارتفاع ایران و همچنین تصاویر ماهواره لندست ۸ و ۹ (با قدرت تفکیک مکانی ۳۰ متر) توسط نرم افزار SAGA GIS استفاده گردید. در بررسی اثر مقیاس متغیرهای محیطی با مقیاس‌های ۳۰ تا ۲۵۰ متر مورد ارزیابی قرار گرفتند. برای پیش‌بینی مکانی ویژگی‌های خاک منتخب شامل شن، سیلت، رس، pH، CCE و SOC از چهار الگوریتم یادگیری ماشین جنگل تصادفی (RF)، کوبیست (CB)، رگرسیون بردار پشتیبان (SVR) و مدل‌های ترکیبی استفاده شد. به‌طور کلی مجموعه داده‌ها به صورت تصادفی ۸۰ درصد برای آموزش و ۲۰ درصد برای اعتبار سنجی استفاده شد. برای این منظور از بسته تخصصی "Caret" و "randomForest" و "svmRadial" در محیط نرم افزار RStudio نسخه 4.1.1717 استفاده شد.

انتخاب متغیرهای محیطی و ارزیابی مدل‌ها

در این مطالعه از الگوریتم باروتا (Kursa and Witold, 2010) با استفاده از مدل جنگل تصادفی (RF) در محیط نرم افزاری R و بسته تخصصی Boruta و RandomForest پیاده سازی شد تا متغیرهای کمکی مورد نیاز برای پیش‌بینی ویژگی‌های خاک مورد نظر انتخاب گردند. صحت مدل‌ها با استفاده از شاخص‌های آماری ضریب تعیین^۱ (R^2)، میانگین ریشه خطای مربع^۲ (RMSE) و میانگین خطای مطلق^۳ (MAE) و ضریب همبستگی تطابق^۴ (CCC) برآورد شد.

نتایج و بحث:

خلاصه آماری ویژگی‌های خاک

نتایج خلاصه آماری ویژگی‌های اندازه‌گیری شده خاک دشت خرم آباد در «جدول ۱» نشان داده شده است. کمترین و بیشترین میزان رس در نمونه‌های خاک به ترتیب ۱۸/۵۶ و ۶۸/۵۶ درصد و میانگین آن در تمام نمونه‌ها ۳۷/۵۸ درصد است. کمترین میزان سیلت و شن به ترتیب ۸/۷۲ و ۱۰ و بیشترین آن‌ها برابر با ۴۴/۷۲ و ۵۴/۷۲ درصد اندازه‌گیری شد. همچنین میانگین سیلت و شن به ترتیب برابر با ۲۹/۳۵ و ۱/۳۳ درصد است. ضریب تغییرات^۵ (CV) به عنوان معیاری برای تفسیر تغییرپذیری داده‌ها براساس طبقه‌بندی ارائه شده توسط ویلدینگ (Wilding, 1985) استفاده شد. براین اساس ذرات تشکیل دهنده خاک شامل رس، سیلت و شن دارای تغییرپذیری متوسط (بین ۱۵ تا ۳۵ درصد) است. کربنات کلسیم معادل (CCE) به عنوان عامل محدود کننده رشد گیاه در منطقه از خیلی کم (۰/۶ درصد) تا زیاد (۲۸ درصد) با میانگین ۷/۶۹ درصد متغیر است. همچنین مقدار متوسط کربن آلی خاک (SOC) ۰/۵۸ درصد می‌باشد مقادیر pH از ۷ تا ۸/۷۲ (خنثی تا قلیایی) متغیر بود. مقدار میانگین pH نشان می‌دهد که خاک منطقه مورد مطالعه از نظر اسیدیته در وضعیت نسبتاً قلیایی است. همچنین مقادیر ضریب تغییرات نشان‌دهنده تغییرپذیری زیاد CCE و SOC (به ترتیب ۶۶/۱۲ و ۶۶/۹۴ درصد) و تغییرپذیری کم pH خاک است

جدول ۱- خلاصه آماری ویژگی‌های اندازه‌گیری شده خاک منطقه خرم آباد

ویژگی‌ها	کمترین	بیشترین	میانگین	SD	CV%
رس	۱۸/۵۶	۶۸/۵۶	۳۷/۵۸	۷/۷۱	۲۰/۵۱
سیلت	۸/۷۲	۴۴/۷۲	۲۹/۳۵	۶/۱۲	۲۰/۸۶
شن	۱۰	۵۴/۷۲	۳۳/۱	۶/۶	۱۹/۹۴
کربنات کلسیم معادل (CCE)	۰/۶	۲۷	۷/۶۹	۵/۰۹	۶۶/۱۲
کربن آلی خاک (SOC)	۰/۰۳	۲/۳۴	۰/۵۸	۰/۳۹	۶۶/۹۴
اسیدیته (pH)	۷	۸/۷۲	۷/۸۱	۰/۲	۲/۵۷

انتخاب متغیرهای محیطی کمکی

پراکنش مکانی برخی از متغیرهای محیطی مورد استفاده برای پیش‌بینی توزیع مکانی ویژگی‌های خاک مورد نظر در منطقه خرم آباد در «شکل ۲» نشان داده شده است.

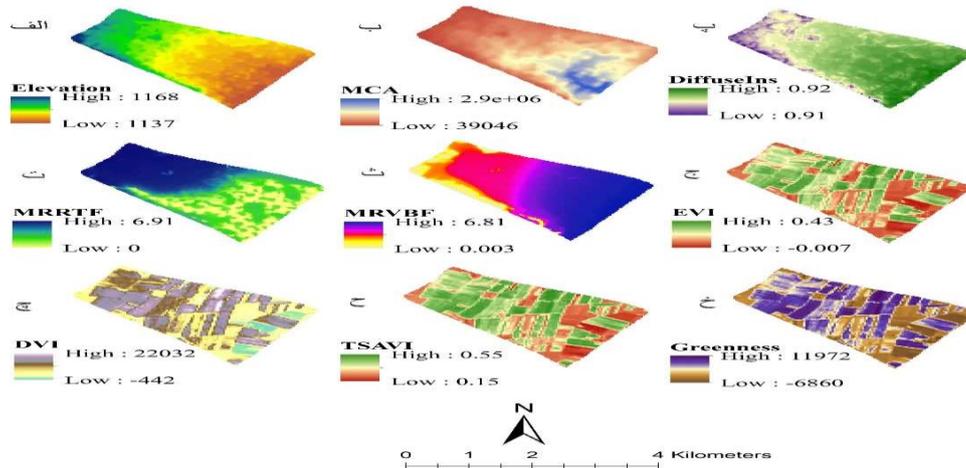
¹ Coefficient Determination

² Root Mean Square Error

³ Mean Absolute Error

⁴ Concordance correlation coefficient

⁵ Coefficient of variation



شکل ۲- پراکنش مکانی متغیرهای کمکی الف: ارتفاع، ب: مساحت حوضه آبریز اصلاح شده، پ: تابش پخشیده، ت: شاخص همواری قله با درجه تفکیک بالا، ث: شاخص همواری دره با درجه تفکیک بالا، ج: شاخص پوشش گیاهی تقویت شده، چ: شاخص تفاوت پوشش گیاهی، ح: شاخص تبدیل یافته پوشش گیاهی تعدیل کننده اثر خاک، خ: سبزیگی

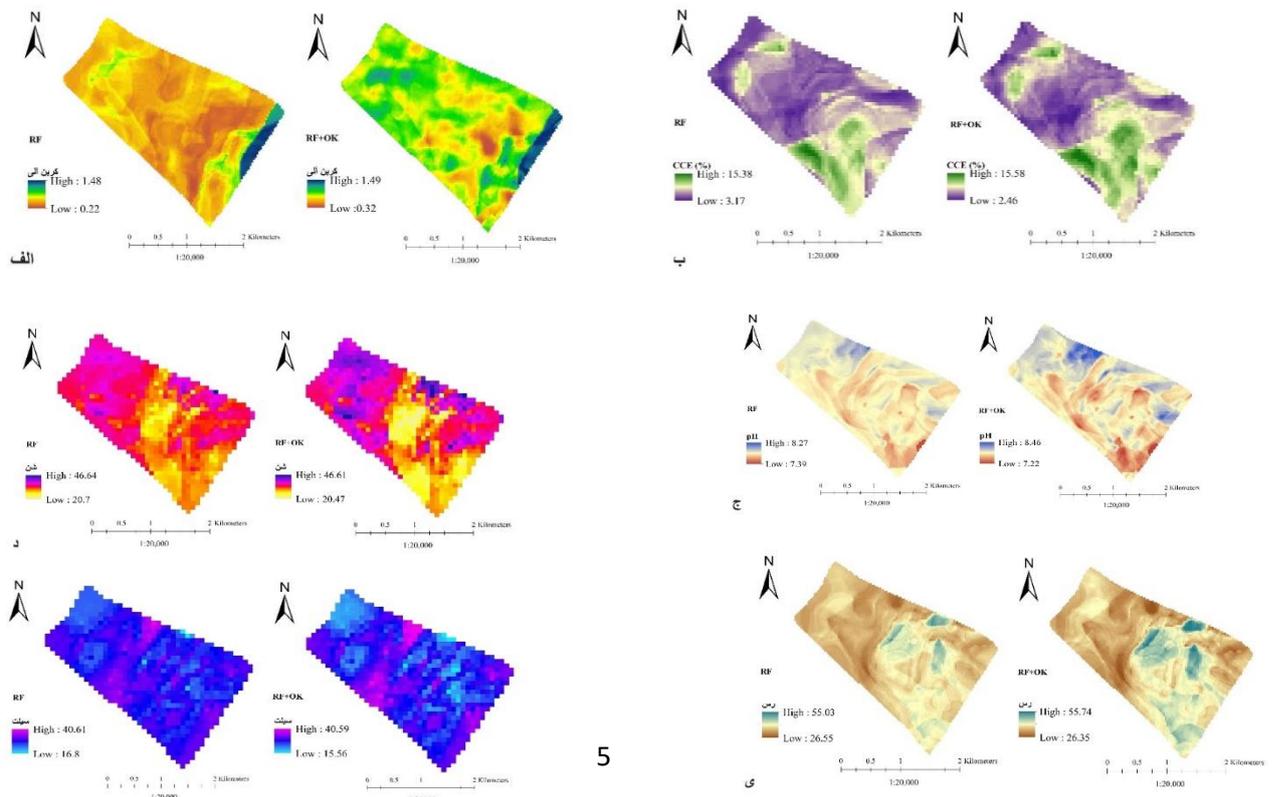
ارزیابی مدل سازی ویژگی های خاک در اندازه های سلولی مختلف

ارزیابی مدل سازی ویژگی های خاک در اندازه های سلولی مختلف شامل ۳۰، ۶۰، ۹۰، ۱۵۰ و ۲۵۰ متر انجام شد. براساس داده های اعتبار سنجی مدل RF در منطقه خرم آباد برای سه ویژگی SOC، pH و رس افزایش اندازه سلولی از ۳۰ متر به ۲۵۰ متر منجر به کاهش R^2 و افزایش RMSE شد. به طوریکه بیشترین صحت برای SOC با R^2 و RMSE برابر با ۰/۸ و ۰/۱۴، برای pH با R^2 و RMSE برابر با ۰/۷۵ و ۰/۰۹ و برای رس با R^2 و RMSE برابر با ۰/۷۹ و ۱۰/۷۹ در اندازه سلولی ۳۰ متر اندازه گیری شد. در حالیکه کربنات کلسیم در اندازه سلولی ۶۰ متر با R^2 و RMSE برابر با ۰/۷۹ و ۱/۹۵ نسبت به سایر اندازه سلولی بیشترین صحت را نشان داد. همچنین پیش بینی مکانی شن و سیلت در اندازه سلولی ۹۰ متر با R^2 و RMSE به ترتیب ۰/۷۷ و ۲/۴۵ برای شن و ۰/۷۷ و ۲/۴۲ برای سیلت با بیشترین صحت محاسبه شد. با توجه به McBratney و Rossel (۲۰۰۸) پیش بینی ویژگی های خاک منطقه خرم آباد توسط مدل RF با صحت خوب (۰/۶۱ تا ۰/۸۱) انجام شد. مقادیر CCC برای هر شش ویژگی SOC، CCE، pH، شن، سیلت و رس به ترتیب برابر با ۰/۷۹، ۰/۸۲، ۰/۷۶، ۰/۸۲، ۰/۷۵ و ۰/۸ محاسبه شد. مقادیر CCC محاسبه شده نیز نشان دهنده صحت اندازه گیری خوب توسط مدل RF است. مقادیر نسبتاً پایین MAE نشان می دهد که تمام نقشه های ویژگی های خاک تهیه شده معمولاً نزدیک به مقادیر اندازه گیری شده هستند. همچنین نتایج اعتبار سنجی مدل های ترکیبی نشان داد مدل ترکیبی RF-OK منجر به تهیه نقشه یا R^2 برابر با ۰/۸۸، ۰/۸۶، ۰/۸۲، ۰/۸۴، ۰/۸۳ و ۰/۸۶ به ترتیب برای CCE با اندازه سلولی ۶۰ متر، pH با اندازه سلولی ۳۰ متر، رس با اندازه سلولی ۳۰ متر، شن با اندازه سلولی ۹۰ متر، سیلت با اندازه سلولی ۹۰ متر و SOC با اندازه سلولی ۳۰ متر شد. به طور کلی استفاده از مدل ترکیبی با انتخاب بهترین اندازه سلولی منجر به افزایش ۷، ۹، ۶، ۶ و ۲ درصدی در صحت پیش بینی SOC، CCE، pH، شن، رس و سیلت شد.

پهنه بندی مکانی ویژگی های خاک منطقه خرم آباد

پهنه بندی مکانی ویژگی های خاک پیش بینی شده توسط مدل های RF و RF-OK در «شکل ۴» ارائه شده است. با توجه به «شکل ۴ الف» SOC پیش بینی شده توسط مدل ترکیبی RF-OK از ۰/۳۲ تا ۱/۴۹ متغیر است. در حالیکه مدل RF این مقدار را بین ۰/۲۲ تا ۱/۴۸ پیش بینی کرده است. بیشترین و کمترین مقدار شن به ترتیب ۲۰/۴۷ تا ۴۶/۶۱ درصد و بیشترین و کمترین مقدار رس به ترتیب ۲۶/۳۵ تا ۵۵/۷۴ درصد پیش بینی شده است. همچنین تغییرات سیلت از ۱۵/۵۶ تا ۴۰/۵۹ درصد توسط مدل RF-OK پیش بینی شده که بیشترین مقادیر آن در بخش مرکزی تجمع یافته و سایر نواحی از میزان کمتری از سیلت برخوردارند. کمترین و بیشترین مقدار CCE برابر با ۲/۴۶ تا ۱۵/۵۸ درصد توسط مدل RF-OK پیش بینی شد. مطابق با تغییرات شن و رس، بیشترین مقدار CCE در جنوب غرب و بخش هایی از شمال منطقه مورد مطالعه تجمع یافته است (شکل

۴ ب). بخش‌های مرکزی و جنوب منطقه مورد مطالعه pH نزدیک به خنثی داشته (۷/۲۲) و برای شمال غرب و بخشی از جنوب غرب، pH بالاتری (۸/۴۶) پیش‌بینی شده است (شکل ۴ ج). به طور کلی عملکردهای اعتبار سنجی بین ۰/۶۱ تا ۰/۸۱ برای مدل‌های یادگیری منفرد قابل مقایسه و سازگار با مطالعات قبلی است که از مدل‌های یادگیری ماشین استفاده کرده‌اند (Khanal et al., 2018)، (Emadi et al., 2020)، (Saurette et al., 2022). در این مطالعه ۹۰ درصد از مساحت منطقه مورد مطالعه دارای شیب کمتر از ۵ درجه می‌باشد و تغییرات توپوگرافی در آن یکنواخت است. با این حال به نظر می‌رسد اندازه سلولی کمتر (در این مطالعه کمتر از ۹۰ متر) منجر به پیش‌بینی ویژگی‌های خاک با صحت بیشتر شده است. مطالعات مختلفی در مورد تاثیر اندازه سلول بر عملکرد مدل‌های یادگیری ماشین انجام شده است که نشان می‌دهد استفاده از مقیاس‌های میانی دقت پیش‌بینی را بهبود می‌بخشد. به عنوان مثال، Behrens و همکاران (۲۰۱۸) به این نتیجه رسیدند که مقیاس‌های میانی (۲۰-۹۰ متر) ممکن است با مقیاس‌های فرآیندهای پدوژنتیکی مهم انطباق بیشتری داشته باشند و دقت پیش‌بینی را بهبود بخشند. این نتیجه با تحقیقات قبلی مطابقت دارد و بر اهمیت اطلاعات زمین در مقیاس ریز برای نقشه برداری دقیق خاک تأکید دارد (Grunwald et al., 2011)، (Flynn et al., 2022)، (Safanelli et al., 2022). استفاده از متغیرهای محیطی کمکی مقیاس‌بندی شده سبب شد تا پیش‌بینی ویژگی‌های خاک در منطقه خرم‌آباد با مقادیر R^2 بالا (خوب) انجام شود. اندازه‌گیری دقیق CCE و SOC که دارای تغییرپذیری زیاد می‌باشند دشوار است اما با استفاده از متغیرهای محیطی با اندازه سلولی کمتر پیش‌بینی این دو ویژگی با صحت بیشتر انجام شد. اگرچه ویژگی‌های pH، شن، سیلت و رس دارای تغییرپذیری کمتر در منطقه خرم‌آباد می‌باشند با این حال استفاده از مقیاس‌های میانی منجر به پیش‌بینی با صحت بالا برای این ویژگی‌ها نیز شد. نتایج اهمیت متغیرهای کمکی مقیاس‌بندی شده در پیش‌بینی ویژگی‌های خاک در منطقه مطالعاتی بیان‌گر تاثیر بیشتر متغیرهای توپوگرافی می‌باشد. پهلوان راد و اکبری مقدم (۲۰۱۸) نیز تاکید داشتند که متغیرهای استخراج شده از مدل رقومی ارتفاع مانند عمق دره و فاصله از جریان آب متغیرهای مهمی در پیش‌بینی ماده آلی هستند. نتایج پژوهش حاضر نشان داد که که زیرمجموعه‌ای از متغیرهای کمکی مقیاس‌بندی شده برای نقشه‌برداری ویژگی‌های خاک منجر به عملکرد بهتر مدل پیش‌بینی کننده خواهد شد. در این مطالعه مدل RF در اندازه‌های سلولی مختلف نسبت به دو مدل دیگر عملکرد بالاتری داشت. موسوی و همکاران (۲۰۲۲) نیز در مطالعات خود تایید نمودند که مدل RF یک الگوریتم یادگیری قدرتمند برای پیش‌بینی ویژگی‌ها و شاخص حاصلخیزی خاک است. متین‌فر و همکاران (۲۰۲۱) بیان کردند که مدل ترکیبی RF-OK دقت پیش‌بینی کربن آلی خاک را در مقایسه با مدل منفرد ۹ و ۱۴ درصد افزایش داد.



شکل ۴- پهنه‌بندی مکانی ویژگی‌های خاک شامل الف: SOC، ب: CCE، ج: pH، د: شن، ه: سیلت، ی: رس در منطقه خرم‌آباد توسط دو

مدل RF-OK و RF

نتیجه‌گیری

مطالعه حاضر با هدف شناسایی مقیاس بهینه جهت بهینه‌سازی دقت نقشه‌برداری رقومی خاک برای ویژگی‌های SOC، CCE، pH، شن، سیلت و رس خاک در منطقه خرم‌آباد انجام شد. نتایج به دست آمده نشان داد به طور کلی مدل RF در پیش‌بینی ویژگی‌های خاک دارای عملکرد بهتری نسبت به مدل‌های SVR و CB است. مدل RF برای پیش‌بینی تمام ویژگی‌های خاک مذکور منجر به پیش‌بینی با بیشترین دقت شد. اندازه سلول میانی ۳۰ تا ۹۰ متر بهینه‌ترین اندازه سلول برای هر شش ویژگی خاک در منطقه مورد مطالعه شناسایی شد. استفاده از مدل RF در منطقه خرم‌آباد برای ویژگی‌های SOC، pH و رس با اندازه سلول ۳۰ متر و ویژگی‌های شن و سیلت با اندازه سلولی ۹۰ متر و CCE با اندازه سلولی ۶۰ متر منجر به پیش‌بینی با صحت خوب شد. از طرفی بهترین اندازه سلول برای پیش‌بینی CCE، pH، سیلت (با دقت پیش‌بینی خوب) و شن (با دقت پیش‌بینی متوسط) اندازه سلولی ۳۰ متر توسط مدل RF و برای SOC و رس اندازه سلولی به ترتیب ۶۰ و ۹۰ متر بهینه‌ترین اندازه سلولی توسط مدل‌های SVR و CB (با دقت پیش‌بینی خوب) شناسایی شد.

فهرست منابع

1. Behrens, T., Schmidt, K., MacMillan, R. A., Viscarra Rossel, R. A. (2018). Multi-scale digital soil mapping with deep learning. *Scientific reports*, 8(1): 15244.
2. Cavazzi, S., Corstanje, R., Mayr, T., Hannam, J., Fealy, R. (2013). Are fine resolution digital elevation models always the best choice in digital soil mapping?. *Geoderma*, 195: 111-121.
3. Emadi, M., Taghizadeh-Mehrjardi, R., Cherati, A., Danesh, M., Mosavi, A., Scholten, T. (2020). Predicting and mapping of soil organic carbon using machine learning algorithms in Northern Iran. *Remote Sensing*, 12(14): 2234
4. Flynn, T., Rozanov, A., Ellis, F., de Clercq, W., & Clarke, C. (2022). Farm-scale digital soil mapping of soil classes in South Africa. *South African Journal of Plant and Soil*, 39(3), 175-186.
5. Grunwald, S., Thompson, J. A., Boettinger, J. L. (2011). Digital soil mapping and modeling at continental scales: finding solutions for global issues. *Soil Science Society of America Journal*, 75(4): 1201-1213.
6. Gee, G. W., Bauder, J. W. (1986). Particle-size analysis 1. *Methods of soil analysis: part 1—physical and mineralogical methods, (methodsofsoil1)*. 383-411.
7. Gomes, L. C., Faria, R. M., de Souza, E., Veloso, G. V., Schaefer, C. E. G., Fernandes Filho, E. I. (2019). Modelling and mapping soil organic carbon stocks in Brazil. *Geoderma*, 340: 337-350.
8. Hengl, T., De Jesus, J. M., MacMillan, R. A., Batjes, N. H., Heuvelink, G. B., Ribeiro, E., Gonzalez, M. R. (2014). SoilGrids1km—global soil information based on automated mapping. *PloS one*, 9(8): 105992.
9. Jackson, Marion LeRoy. (1969). *Soil chemical analysis-advanced course*. 895-pp.
10. Kursa, Miron B., Witold R. (2010). Feature selection with the Boruta package. *Journal of statistical software*, 36: 1-13.
11. Keskin, H., Grunwald, S. (2018). Regression kriging as a workhorse in the digital soil mapper's toolbox. *Geoderma* 326: 22-41.
12. Khanal, S., Fulton, J., Klopfenstein, A., Douridas, N., Shearer, S. (2018). Integration of high resolution remotely sensed data and machine learning techniques for spatial prediction of soil properties and corn yield. *Computers and electronics in agriculture*, 153: 213-225.
13. Matinfar, H. R., Maghsodi, Z., Mousavi, S. R., Rahmani, A. (2021). Evaluation and Prediction of Topsoil organic carbon using Machine learning and hybrid models at a Field-scale. *Catena*, 202: 105258.
14. Malone, B. P., Styc, Q., Minasny, B., McBratney, A. B. (2017). Digital soil mapping of soil carbon at the farm scale: A spatial downscaling approach in consideration of measured and uncertain data. *Geoderma*, 290, 91-99.
15. Miller, B. A., Koszinski, S., Wehrhan, M., Sommer, M. (2015). Impact of multi-scale predictor selection for modeling soil properties. *Geoderma*, 239: 97-106.
16. McLean, E. O. (1982). Soil pH and lime requirement. *Methods of soil analysis: Part 2 Chemical and microbiological properties* 9, 199-224.
17. Nelson, D. W., Sommers, L. E. (1980). Total nitrogen analysis of soil and plant tissues. *Journal of the Association of Official Analytical Chemists*, 63(4): 770-778.

18. Pahlavan-Rad, M. R., Akbarimoghaddam, A. (2018). Spatial variability of soil texture fractions and pH in a flood plain (case study from eastern Iran). *Catena*, 160: 275-281.
19. Rossel, R. V., McBratney, A. B. (2008). Diffuse reflectance spectroscopy as a tool for digital soil mapping. In *Digital soil mapping with limited data*. Springer Netherlands, 165-172 pp.
20. Saurette, D. D., Berg, A. A., Laamrani, A., Heck, R. J., Gillespie, A. W., Voroney, P., Biswas, A. (2022). Effects of sample size and covariate resolution on field-scale predictive digital mapping of soil carbon. *Geoderma*, 425: 116054.
21. Safanelli, J. L., Demattê, J. A. M., Santos, N. V. D., Rosas, J. T. F., Silvero, N. E. Q., Bonfatti, B. R., & Mendes, W. D. S. (2021). Fine-scale soil mapping with Earth Observation data: a multiple geographic level comparison. *Revista Brasileira de Ciência do Solo*, 45: 0210080.
22. Wilding, L. P. (1985). Spatial variability: its documentation, accommodation and implication to soil surveys. 166-189.
23. Zeraatpisheh, M., Jafari, A., Bodaghabadi, M. B., Ayoubi, S., Taghizadeh-Mehrzardi, R., Toomanian, N., Xu, M. (2020). Conventional and digital soil mapping in Iran: Past, present, and future. *Catena*, 188: 104424.
24. Zeraatpisheh, M., Jafari, A., Bodaghabadi, M. B., Ayoubi, S., Taghizadeh-Mehrzardi, R., Toomanian, N., & Xu, M. (2020). Conventional and digital soil mapping in Iran: Past, present, and future. *Catena*, 188:104424.

The effect of the scale of environmental variables in digital mapping of soil properties in a semi-arid region of western Iran

Ziba Maghsodi¹, Hamid Reza Matinfar*¹, Seyed Roohollah Mousavi²

*¹ Department of Soil Science, Faculty of Agriculture, Lorestan University, Lorestan, Khorramabad, Iran

² Department of Soil Science and Engineering, Faculty of Agriculture, University of Tehran, Karaj, Iran

Abstract

Preparing accurate maps of soil properties is essential for assessing various aspects of agricultural land management, including production potential and organic carbon reserves. The aim of this study is to improve mapping accuracy by selecting the optimal scale for predicting six soil properties, including Soil Organic Carbon (SOC), Calcium Carbonate Equivalent (CCE), pH, sand, silt, and clay in the Khorramabad region. For this purpose, 157 samples were collected from a depth of 0-30 cm in the Khorramabad area based on a random sampling pattern and were sent to the laboratory for analysis. Environmental variables such as topography and remote sensing data were extracted from a Digital Elevation Model (DEM) and Landsat 8 satellite images. The optimal environmental variables were selected using the Boruta method in the Khorramabad area. Soil properties modeling was performed using machine learning models, including Random Forest (RF), Support Vector Regression (SVR), Cubist (CB), and ensemble models. The results showed that in areas with low topography, the use of intermediate scales (cell sizes of 30 to 90 meters) resulted in higher prediction accuracy for soil properties. This is because larger cell sizes introduce noise that hinders accuracy. Therefore, the choice of resolution, or in other words, cell size in Digital Soil Mapping (DSM), should be tailored to specific needs, balancing meaningful pattern capture with computational efficiency.

Keyword: Digital Mapping, Environmental Variables, Optimal Scale, Spatial Modeling