

برآورد خصوصیات مختلف خاک با استفاده از داده‌های سریع و ارزان حسگر رنگ

فاطمه چشم براه^۱، علی اصغر ذوالفقاری^{۲*}، روح الله تقی زاده مهرجردی^۳

(تاریخ دریافت: ۱۴۰۰/۰۸/۳۰ تاریخ پذیرش: ۱۴۰۱/۰۴/۲۰)

چکیده

معمولاً رنگ خاک با بسیاری از خصوصیات فیزیکی و شیمیایی در ارتباط بوده و اندازه‌گیری آن ساده، سریع و کم هزینه است. حسگر رنگ نیکس پرو (NixTMpro) از جمله وسایلی است که می‌تواند رنگ خاک را با استفاده از سیستم‌های مختلف با دقت زیادی تعیین کند. این حسگر در مقایسه با روش مرسوم دفترچه مانسل برای تعیین رنگ خاک، نسبت به شرایط محیطی و ذهنی کاربر حساسیت کمتری دارد و استفاده از آن بسیار آسان است. بنابراین، در این مطالعه از حسگر رنگ نیکس پرو برای برآورد سریع و ارزان خصوصیات مختلف خاک استفاده شد. به همین منظور، ۱۵۰ نمونه خاک از منطقه مطالعاتی نیمه‌خشک استان قزوین جمع‌آوری و مقادیر مربوط به هر ویژگی خاک در آزمایشگاه اندازه‌گیری شد. سپس با استفاده از این حسگر رنگ، طیف‌های مربوط به هر سیستم رنگ برای هر نمونه خاک ثبت شدند. برای این منظور دو روش مورد استفاده قرار گرفت. در روش اول با بررسی همبستگی بین متغیرهای سیستم رنگ سعی شد یک سیستم رنگ استاندارد که دارای بالاترین ضریب همبستگی با همه ویژگی‌های خاک است معرفی شود. در روش دوم همه متغیرهای سیستم‌های مختلف رنگ با روش حذف ویژگی بازگشتی مورد بررسی قرار گرفت، که این روش با انتخاب مهم‌ترین ویژگی‌ها، ویژگی‌هایی که دارای بالاترین دقت هستند، انتخاب می‌کند. با توجه به نتایج هر دو روش و به جهت معرفی یک سیستم رنگ استاندارد، در این مطالعه از سیستم رنگ $CIE\ L^*a^*b$ برای برآورد خصوصیات خاک استفاده شد، زیرا این سیستم نسبت به سیستم‌های دیگر رنگ همبستگی بیشتری را با خصوصیات مختلف خاک نشان داد. سپس با استفاده از الگوریتم جنگل تصادفی (Random Forest) مقادیر مربوط به هر ویژگی خاک برآورد شد. خصوصیات خاک شامل مقادیر شن، سیلت، رس، میزان شوری، مقدار کربنات کلسیم ($CaCO_3$)، مقدار ماده آلی و جرم مخصوص ظاهری خاک بودند. با توجه به نتایج حاصل از برآورد مدل جنگل تصادفی، میزان ریشه دوم میانگین مربعات خطا (RMSE)، میانگین خطا (ME)، ارزیابی فاصله چارکی (PRIQ) و مقدار ضریب همبستگی (r) برای هر ویژگی خاک تعیین شد. برای ذرات شن، سیلت و رس مقدار RMSE به ترتیب برابر با ۱۰/۰۷، ۶/۲۸، ۷/۲۶ درصد و میزان ضریب همبستگی به ترتیب برابر با ۰/۷۰، ۰/۴۹، ۰/۷۷ بود. میزان آماره PRIQ برای ذرات شن (۲/۰۹)، رس (۲/۳۷) و کربنات کلسیم (۱/۷۸) در سطح مناسب و قابل قبول می‌باشد. مقادیر RMSE برای ماده آلی، کربنات کلسیم و جرم مخصوص ظاهری به ترتیب ۰/۲۵، ۰/۱۱ درصد و میزان ضریب همبستگی به ترتیب ۰/۵۵، ۰/۷۸، ۰/۷۰ بود. با توجه به این نتایج می‌توان گفت که طیف‌های حاصل از حسگر رنگ نیکس پرو می‌تواند برای پیش‌بینی سریع خصوصیات خاک مفید باشد.

واژه‌های کلیدی: رنگ خاک، حسگر رنگ نیکس پرو، سیستم طیفی، مدل جنگل تصادفی

چشم براه ف.، ذوالفقاری ا.ع.، تقی زاده مهرجردی ر. ۱۴۰۲. برآورد خصوصیات مختلف خاک با استفاده از داده‌های سریع و ارزان حسگر رنگ. تحقیقات کاربردی خاک. جلد ۱۱، شماره ۲. صفحه: ۲۸-۴۵.

۱-دانش آموخته دکتری رشته بیابان‌دایی، گروه بیابان‌دایی، دانشکده کویرشناسی، دانشگاه سمنان

۲-دانشیار گروه بیابان‌دایی، دانشکده کویرشناسی، دانشگاه سمنان

۳- استادیار گروه مهندسی طبیعت، دانشکده کشاورزی و منابع طبیعی، دانشگاه اردکان و فوق‌دکتر در گروه علوم زمین، دانشگاه توبینگن، آلمان
*پست الکترونیک نویسنده مسئول: azolfaghari@semnan.ac.ir

مقدمه

آگاهی از خصوصیات مختلف خاک مانند بافت، ساختمان، ماده آلی، شوری، رطوبت و غیره اهمیت بسزایی در مطالعات خاکشناسی و مدیریت منابع خاک دارند. ابزارها و روش‌های مرسوم بررسی خاک، اغلب پرهزینه و متکی بر روش‌های پیچیده و طولانی هستند (Swetha & Chakraborty, 2021). اندازه‌گیری ذرات خاک به‌طور کلاسیک و با روش‌های رسوب‌سنجی با استفاده از قرائت‌های هیدرومتر انجام می‌شود. برای اندازه‌گیری ماده آلی، جرم مخصوص ظاهری و سایر خصوصیات نیز روش‌هایی استفاده می‌شوند که علاوه بر اینکه هزینه‌بر هستند، نیاز به صرف زمان طولانی و زحمت زیادی دارند. بنابراین یک روش سریع‌تر و ارزان‌تر برای به دست آوردن اطلاعات موردنیاز برای به دست آوردن اطلاعات سریع از خصوصیات خاک ارزشمند خواهد بود.

رنگ خاک و همبستگی آن با دیگر عوامل خاک مانند رطوبت خاک و پوشش گیاهی همیشه موردتوجه محققان و مهندسان کشاورزی بوده است (Anderson, 2005). رنگ خاک به‌عنوان شاخصی از فرآیندهای پدوژنتیک و عوامل تشکیل‌دهنده خاک شواهد مهمی از شرایط تشکیل خاک را نشان می‌دهد (Rezende, 2021). در بسیاری از سامانه‌های طبقه‌بندی خاک رنگ یا بازتاب خاک در محدوده طیف مرئی از ویژگی‌های تفکیک‌کننده بوده و یکی از اجزاء مهمی است که در شناسایی افق‌های سطحی و زیرسطحی مؤثر واقع می‌شود (Baumgardner et al., 1986). رنگ خاک به‌عنوان عاملی کلیدی در سیستم‌های مختلف رده‌بندی خاک مورد استفاده قرار گرفته است (FAO & I, 2015; Rezende, 2021, Staff, 2014). با توجه به اینکه رنگ خاک عمیقاً با مواد مادری خاک، کانی‌شناسی و شرایط زهکشی خاک مرتبط است، به‌عنوان یک معیار تشخیصی رایج برای افق‌های خاک در سیستم‌های مختلف طبقه‌بندی خاک در نظر گرفته شده است. به همین ترتیب علاوه بر مواد آلی، محتوای اکسیدها و بسیاری دیگر از اجزای مهم خاک مانند کربنات کلسیم (CaCO_3) که منعکس‌کننده شرایط گذشته و فعلی خاک می‌باشند با رنگ خاک در ارتباط هستند (Resende et al., 2014). بنابراین، رنگ خاک می‌تواند اطلاعات ارزشمندی در مورد خواص خاک از جمله مقدار ماده آلی خاک و درصد فراوانی ذرات

اولیه خاک ارائه دهد (Ibáñez-Asensio et al., 2013). رنگ خاک یکی از ویژگی‌های اصلی خاک است که می‌تواند به سرعت و به‌طور خودکار ارزیابی شود و رابطه قوی با ویژگی‌های خاک دارد (Aitkenhead et al., 2013). با این حال تشخیص رنگ خاک، کار ساده‌ای نبوده و تشخیص آن به‌عنوان یک ویژگی مهم با مشکل مواجه شده است (Kirillova et al., 2018b). امروزه سیستم رنگ مانسل^۱ مرجع اصلی توصیف رنگ خاک می‌باشد که در آن از سه مشخصه هیو، کروما و ولیو برای طبقه‌بندی خاک استفاده می‌شود (Dos Santos et al., 2018). بنابراین طبقه‌بندی سنتی رنگ خاک با توجه به درک محقق به رنگ‌های استاندارد موجود در مانسل انجام می‌شود. بنابراین، تشخیص رنگ خاک بستگی به نگاه محقق، شدت تأثیر عوامل محیطی مانند نور خورشید، ارتفاع، رطوبت خاک یا ایجاد سایه بر روی نمونه‌ها دارد. به همین دلیل، هیچ تضمینی برای تشخیص دقیق آن وجود ندارد، زیرا دقت هر محقق ممکن است متفاوت باشد، موضوعی که در بسیاری از مطالعات به آن اشاره شده است (Kirillova et al., 2018a; Kirillova et al., 2015; Marqués-Mateu et al., 2018; Mouazen et al., 2007).

با توجه به ضرورت بهینه‌سازی مدیریت خاک جهت برآوردن تقاضای جهانی غذا از طریق تجزیه و تحلیل گسترده خاک با وضوح بالا ضروری است (Barthès et al., 2019). همچنین پرهزینه و وقت‌گیر بودن روش‌های سنتی تعیین خصوصیات خاک، استفاده از ابزارهایی مانند اسپکتروفتومتر و رنگ‌سنج‌ها به‌جای دفترچه رنگ مانسل می‌تواند مفید واقع گردد (Thompson et al., 2013). خوشبختانه دانشمندان روش‌های جدیدی از تشخیص رنگ را آزمایش و ایجاد نموده‌اند که علاوه بر مزه‌پسند بودن، دارای هزینه کمتری نیز می‌باشند (Levin et al., 2005). اخیراً، حسگر رنگی به نام نیکس پرو برای ارزیابی خاک توسعه یافته و نتایج بسیار خوبی از آن حاصل شده است. این حسگر برای شناسایی رنگ اجسام معرفی شده و قادر است مشخصات رنگ را در قالب سیستم‌های رنگی مختلف بیان کند. این حسگر به راحتی قابل حمل و قابل شارژ بوده و با استفاده از منبع نوری که در آن تابش به اجسام وجود دارد، آن را به یک جایگزین ارزشمند برای مانسل تبدیل کرده است. داده‌های اندازه‌گیری شده توسط این حسگر با استفاده از

^۱ Munsell

شیمیایی خاک و بر اساس مدل داده‌کاوی جنگل تصادفی می‌باشد.

مواد و روش‌ها

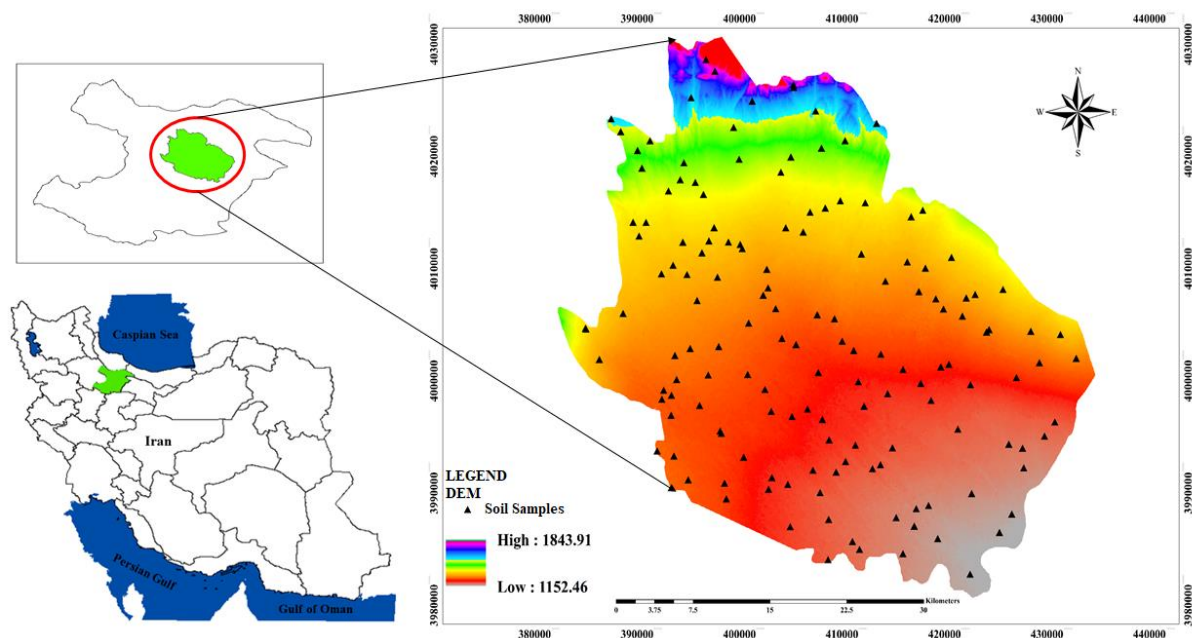
منطقه مطالعاتی

منطقه مورد مطالعه در استان قزوین در منطقه مرکزی ایران، بین عرض جغرافیایی $36^{\circ}15'$ تا $36^{\circ}43'$ عرض شمالی و $50^{\circ}02'$ و $50^{\circ}17'$ طول شرقی واقع شده است (شکل ۱). مساحت منطقه مورد مطالعه ۱۵۲۰۰۰ هکتار و اقلیم آن نیمه‌خشک می‌باشد. میانگین دمای، بارش و پتانسیل تبخیر و تعرق سالانه در منطقه مورد مطالعه به ترتیب $14/3^{\circ}\text{C}$ ، ۳۰۱ میلی‌متر و ۱۲۵۰ میلی‌متر در سال است (Nasrollahi *et al.*, 2021). این منطقه شامل کوه‌ها، تپه‌های شیب‌دار و ملایم، زمین‌های مسطح و دشت‌های سیل‌آبی برای کشاورزی می‌باشد. کاربری اراضی در منطقه مورد مطالعه مراتع دارای پوشش گیاهی غنی و فقیر (۰/۱۰) از منطقه مورد مطالعه) و زمین‌های کشاورزی می‌باشد که برای تولید گندم، جو و ذرت (۰/۹۰) از منطقه مورد مطالعه) استفاده می‌شود. بر اساس رده‌بندی WRB خاک‌های موجود در منطقه از نوع Calcisols و Cambisols هستند (Sabetizade *et al.*, 2021).

نرم‌افزاری که بر روی گوشی تلفن همراه نصب می‌شود به راحتی قابل دریافت است (Stiglitz *et al.*, 2016; Stiglitz *et al.*, 2017).

جا و همکاران (Jha *et al.*, 2021) با استفاده از حسگر رنگ نیکس پرو مقدار کل آهن خاک را برآورد نمودند. تاکنون مطالعات مختلفی در زمینه برآورد کربن آلی خاک با استفاده از حسگر رنگ نیکس پرو صورت گرفته است. به‌عنوان نمونه، سوتا و چاکرابورتی (Swetha & Chakraborty, 2021) از اسکن نیکس پرو و PXRF و رگرسیون جنگل تصادفی (RF) برای برآورد SOC خاک با ترکیبات مختلف داده‌ها استفاده کردند (Mukhopadhyay *et al.*, 2020).

در ایران مطالعه‌ای که با استفاده از این حسگر صورت گرفته است توسط رئیسی و همکاران (Raeesi *et al.*, 2019) می‌باشد که میزان ماده آلی خاک در مناطق خشک و نیمه‌خشک را مورد بررسی و ارزیابی قرار دادند. اما مطالعه دیگری که ارزیابی توانایی این حسگر را در برآورد سایر خصوصیات خاک نشان دهد، صورت نگرفته است. بنابراین، این مطالعه با هدف (۱) ارزیابی کارایی داده‌های حسگر رنگ نیکس پرو برای برآورد برخی از خصوصیات فیزیکی و



شکل ۱- موقعیت منطقه مطالعاتی و نقاط نمونه‌برداری
Figure 1. Position of Study area and sampling points

نمونه برداری خاک و تجزیه و تحلیل آزمایشگاهی

تعداد ۱۵۰ نمونه خاک از عمق ۰ تا ۳۰ سانتی متری جمع آوری شد. نقاط نمونه برداری خاک با استفاده از روش نمونه برداری مکعب لاتین (conditional Latin Hypercube Sampling) cLHS (Minasny *et al.*, 2008) مشخص شدند. این روش از متغیرهای کمکی که از تصاویر ماهواره-ای لندست ۸ و داده‌های رقومی ارتفاع به دست می‌آید برای تعیین نقاط نمونه برداری استفاده می‌کند. در اینجا از باندها و شاخص‌های مستخرج از تصاویر لندست مانند شاخص پوشش گیاهی و همچنین داده‌های رقومی ارتفاع، شیب و شاخص رطوبت برای تعیین مکان نقاط نمونه برداری استفاده شد.

در آزمایشگاه مقادیر ذرات شن (Sand)، سیلت (Silt)، رس (Clay)، جرم مخصوص ظاهری (Bulk Density)، مقدار ماده آلی (Organic Matter)، میزان شوری (EC) و مقدار کربنات کلسیم (CaCO_3) در نمونه‌های خاک جمع آوری شده، تعیین شد. جرم مخصوص ظاهری خاک با استفاده از روش کلوخه و پارافین به دست آمد. شوری خاک در عصاره گل اشباع و با استفاده از EC متر و مقدار کربنات کلسیم (آهک) با استفاده از روش کلسی متری تعیین شدند. برای تعیین مقادیر ذرات شن، رس، سیلت و مقدار ماده آلی در ابتدا، نمونه‌های خاک، هوا خشک شده و از الک ۲ میلی متری عبور داده شدند. سپس برای اندازه‌گیری ذرات بافت خاک شامل شن، سیلت و رس از متافسفات سدیم و هم زدن مکانیکی برای پراکندگی خاک رس استفاده شد (Gee & Or, 2002). سپس با استفاده از هیدرومتر ASTM غلظت ذرات شن، سلت و رس به دست آمد. ماده آلی خاک با استفاده از روش سوزاندن تر تعیین شد (Walkley & Black, 1934).

انجام تست رنگ با استفاده از حسگر رنگ نیکس پرو

حسگر رنگ نیکس پرو ابزاری کوچک، تقریباً ارزان قیمت و ساده است که اخیراً استفاده از آن جهت اسکن رنگ اشیا گسترش زیادی یافته است (شکل ۲). این حسگر رنگ، خود دارای دیود تابش نور (منبع نور LED) است و باطری قابل شارژ دارد. همچنین، پس از آنالیز رنگ اجسام، مقادیر حاصل را در سیستم‌های رنگی متفاوت مانند $\text{CIE}L^*a^*b^*$ ، RGB و XYZ و CMYK تولید و ذخیره می‌کند (Stiglitz

2016). سیستم رنگ RGB می‌تواند به‌عنوان یک مکعب با ۲۵۵ نقطه گسسته در هر طرف تنظیم شود. با توجه به درجه رنگ شامل رنگ‌های قرمز (Red)، سبز (Green) و آبی (Blue) است. سیستم رنگ XYZ از نظر کارکرد تقریباً مشابه سیستم رنگ RGB بوده و متغیرهای آن به ترتیب X نمایانگر رنگ قرمز، Y بیانگر میزان روشنایی و Z تا حدودی آبی را مشخص می‌کند. در سیستم رنگ $\text{CIE}L^*a^*b^*$ متغیر L نشان‌دهنده درجه روشنی تا تیرگی (lightness to darkness)، متغیر a^* نشان‌دهنده قرمزی تا سبزی (- redness + to greenness) و متغیر b^* نمایانگر زردی تا رنگ آبی (- yellowness + to blueness) است (Stiglitz *et al.*, 2017). سیستم رنگ CMYK، از مخفف سرواژه‌های Cyan (فیروزه‌ای)، Magenta (سرخایی)، Yellow (زرد) و Black (سیاه) تشکیل شده است. در این سیستم رنگ مقادیر برای هر کد به صورت درصد بیان می‌شود، درحالی‌که در سایر سیستم‌ها مقادیر بین ۰-۱۰۰ ارائه شدند. حسگر رنگی قدرتمند نیکس پرو اندازه‌گیری دقیق رنگ هر سطح را ارائه می‌دهند و داده‌ها را به برنامه تلفن همراه می‌فرستند.

برای تعیین رنگ نمونه‌ها در ابتدا نمونه‌های خاک هوا خشک شدند. سپس رنگ خاک نمونه‌ها در شرایط مرطوب و خشک تعیین شد. از هر نمونه خاک به صورت جداگانه و تقریباً به قطر یک اینچ بر روی سطح قرار داده شدند و با استفاده از حسگر رنگ، رنگ نمونه‌های خشک در سیستم‌های مختلف تعیین شد (Stiglitz *et al.*, 2016). برای ایجاد شرایط مرطوب و تعیین رنگ نمونه‌ها در این شرایط، سطح هر نمونه با آب مقطر با استفاده از قطره‌چکان مرطوب شد و بعد از چند ثانیه حسگر بر روی نمونه مرطوب قرار داده شد (Stiglitz *et al.*, 2017). در هنگام آزمایش رنگ خاک، LED حسگر به‌طور کامل توسط نمونه خاک پوشانده شد تا از ورود نور خارجی جلوگیری کند. با کلیک روی گزینه اتصال به دستگاه حسگر توسط گوشی همراه هوشمند، طیف قابل مشاهده به صورت کد برای هر نمونه خاک مشخص می‌شود و در گوشی تلفن همراه نتایج مختلف از تجزیه کدهای مربوط به هر سیستم طیفی یعنی RGB، CMYK و XYZ و $\text{CIE}L^*a^*b^*$ ذخیره می‌شود که مقدار هر کد به صورت درصد است (شکل ۲). پس از آماده‌سازی

نمونه‌ها، همه آن‌ها اسکن شدند. سپس از نتایج حاصل از اسکن که در حافظه دستگاه تلفن هوشمند ذخیره شده بودند، خروجی گرفته شد. سپس همبستگی بین متغیرها با ضریب همبستگی (r) در سیستم‌های رنگ مختلف، مورد بررسی قرار می‌گیرد. به این دلیل که ممکن است بین متغیرهای مستقل در یک سیستم رنگ همبستگی بالایی وجود داشته باشد که باعث خودهمبستگی و اشتباه در مدل‌سازی شود. بنابراین در ابتدا همبستگی بین متغیرهای حاصل از داده‌های حسگر رنگ تعیین شد. به منظور تعیین بهترین متغیرها در برآورد خصوصیات خاک همبستگی بین متغیرها حاصل از حسگر در هر کدام از سیستم‌های رنگ با خصوصیات خاک به دست آمد. با تعیین میزان همبستگی بین متغیرهای هر سیستم طیفی، سیستمی که متغیرهای آن بیشترین میزان همبستگی با خصوصیات خاک را داشت به عنوان متغیر مستقل برای مدل‌سازی و برآورد خصوصیات خاک انتخاب شدند.

مدل جنگل تصادفی (RF)

در این مطالعه از روش غیرخطی جنگل تصادفی با استفاده از داده‌های حسگر رنگ که به عنوان متغیرهای مستقل برای برآورد خواص خاک به عنوان متغیرهای هدف (وابسته) شامل شن، سیلت، رس، کربنات کلسیم، شوری، جرم

مخصوص ظاهری خاک و ماده آلی استفاده شد. جنگل تصادفی (RF) یک الگوریتم ترکیبی است که از چندین درخت تصمیم (Decision Tree) برای الگوریتم خود استفاده می‌کند. این روش به تنظیمات زیادی نیاز ندارد (Kuhn & Johnson, 2013). این الگوریتم مجموعه داده‌ها را به ntree که نشان‌دهنده تعداد درختان است تقسیم می‌کنند. سپس مجموعه‌ای از درخت‌های تصمیم، باهم یک جنگل را تولید می‌کنند (Hastie et al., 2005) که این جنگل می‌تواند نتایج بهتری را نسبت به یک درخت اتخاذ نماید.

در این مطالعه، تعداد درختان ۱۰۰۰ و حداقل تعداد گره هر درخت برابر با پنج در نظر گرفته شد. به منظور استفاده از تمامی داده‌های موجود برای بررسی دقت مدل، تمامی داده‌ها به ۳ گروه تقسیم شدند سپس در هر بار اجرای مدل یک گروه (حدود ۳۰٪ داده‌ها) به عنوان داده تست و دو گروه دیگر که حدود ۷۰ درصد داده‌ها را شامل می‌شدند به عنوان داده آموزش استفاده می‌شدند. بنابراین بر اساس روش مذکور تمامی داده‌ها به عنوان داده تست مورد ارزیابی قرار می‌گیرند. محاسبات در نرم‌افزار RStudio (version 1.4.1106) و با پکیج Random Forest انجام شد (شکل ۲).



شکل ۲- نمودار نحوه عملکرد حسگر اسکن رنگ نیکس پرو برای برآورد خصوصیات خاک

Figure 2. Functional diagram of the NIX™Pro color sensor application for prediction of different soil properties

متغیرهای اضافی قابلیت تعمیم مدل را کاهش می‌دهد و همچنین ممکن است دقت کلی طبقه بندی را کاهش دهد. علاوه بر این، افزودن متغیرهای بیشتر و بیشتر به یک مدل،

انتخاب متغیر

هنگام ساخت یک مدل یادگیری ماشین همه متغیرهای مجموعه داده برای ساخت یک مدل مفید نیستند. افزودن

رابطه ۲:

$$r = \frac{\sum_{i=1}^n (x_i - \bar{x})(y_i - \bar{y})}{\sqrt{[\sum_{i=1}^n (x_i - \bar{x})]^2} \sqrt{[\sum_{i=1}^n (y_i - \bar{y})]^2}}$$

رابطه ۳:

$$PRIQ = \frac{IQ}{RMSE}$$

رابطه ۴:

$$ME = \frac{\sum_{i=1}^n (y_i - x_i)}{n}$$

RPIQ نسبت فاصله بین چارک‌های اول و سوم داده‌های اندازه‌گیری شده (Q3 - Q1) به مقدار آماره RMSE می‌باشد و می‌توان با استفاده از رابطه (۴) آن را محاسبه کرد. Q1 مقداری است که درجایی که ۲۵ درصد داده‌های اندازه‌گیری شده زیر آن قرار دارد ممکن است پیدا شود و Q3 است مقداری است که زیر ۷۵ درصد نمونه‌ها ممکن است یافت شوند (Bellon-Maurel *et al.*, 2010; Niazi *et al.*, 2015)

نتایج و بحث

خلاصه آماری داده‌های اندازه‌گیری شده

در جدول (۱) خلاصه آماری خصوصیات خاک آمده است. حداقل جرم مخصوص ظاهری ۱/۲۱ و حداکثر آن ۱/۹۵ گرم بر سانتی‌متر مکعب می‌باشد. همان‌طور که نتایج ضریب تغییرات (CV) نیز نشان می‌دهد در بین متغیرهای مورد مطالعه کمترین درصد ضریب تغییرات (CV) مربوط به جرم مخصوص ظاهری خاک (۰/۸۱۹۶٪) است. میزان شوری خاک در کمترین مقدار ۰/۴ dS/m و حداکثر آن ۷/۴ dS/m است که دارای تنوع بالایی است و بیش‌ترین درصد ضریب تغییرات (۰/۹۴/۱۸٪) در بین خصوصیات خاک را داشته است. ضریب تغییرات نشان‌دهنده دامنه تغییرپذیری یک ویژگی می‌باشد. هر چه مقدار این ضریب کمتر باشد نشان‌دهنده پراکنش یکنواخت‌تر ویژگی می‌باشد. با توجه به کلاس تغییرپذیری ضریب تغییرات بر اساس معیار ارائه‌شده توسط ویلدینگ (Wilding, 1985) اگر درصد ضریب تغییرات کمتر از ۱۵ باشد، تغییرپذیری کم، اگر این ضریب بین ۱۵

پیچیدگی کلی مدل را افزایش می‌دهد. بنابراین، انتخاب ویژگی به بخشی ضروری از ساخت مدل‌های یادگیری ماشین تبدیل می‌شود. هدف از انتخاب ویژگی در یادگیری ماشینی یافتن بهترین مجموعه از ویژگی‌ها است که به فرد امکان می‌دهد مدل‌های مفیدی از ویژگی‌های مورد مطالعه بسازد.

در این تحقیق از دو روش برای انتخاب متغیرها استفاده شد. الف) بررسی همبستگی بین سیستم‌های مختلف رنگ و انتخاب سیستم رنگی که متغیرهای آن بیشترین همبستگی را با ویژگی‌های خاک دارند. این روش این امکان را برای ما به وجود می‌آورد که براحتی بتوانیم یک سیستم رنگ استاندارد را معرفی کنیم. ب) انتخاب متغیرها با استفاده از حذف ویژگی بازگشتی (REF) در پکیج (Caret) در نرم افزار Rstudio پیاده‌سازی شد تا گروه کوچکی از متغیرهای مهم را پیدا کند تا بتواند بهترین نتایج را برای پیش‌بینی خصوصیات خاک داشته باشد. REF یک الگوریتم انتخاب معکوس است که به طور مکرر از روش رتبه بندی استفاده می‌کند تا مجموعه‌های مختلفی از مناسب‌ترین متغیرها را با استفاده از همه پیش‌بینی‌کننده‌ها ارائه دهد. این روش این امکان را به وجود می‌آورد تا بتوان متغیرهای محیطی را که در نقشه برداری رقومی خاک کاربرد دارند، ایجاد کرد. با توجه به اینکه هدف این مطالعه معرفی یک سیستم استاندارد می‌باشد بنابراین پس از مشاهده نتایج هر دو روش سعی شده است یک سیستم استاندارد که بیشترین همبستگی را با همه ویژگی‌های خاک دارد معرفی شود.

ارزیابی کارایی مدل جنگل تصادفی

جهت ارزیابی کارایی مدل جنگل تصادفی در برآورد خصوصیات خاک، از آماره‌های ضریب همبستگی (r)، میانگین خطا (ME)، ارزیابی فاصله چارکی (PRIQ) و ریشه دوم میانگین مربعات خطا (RMSE) استفاده شد که نحوه محاسبات هر کدام از آن‌ها در زیر آمده است. در این روابط X_i مقادیر مشاهداتی، Y_i مقادیر برآورد شده و n تعداد داده‌ها می‌باشد.

رابطه ۱:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (x_i - \bar{x}_i)^2}{n}}$$

تغییرپذیری کم، سیلت در کلاس تغییرپذیری متوسط (۱۶/۳۴) و سایر خصوصیات خاک در کلاس با تغییرپذیری زیاد قرار گرفتند.

تا ۳۵ درصد باشد، تغییرپذیری متوسط و اگر ضریب تغییرات بیشتر از ۳۵٪ باشد، تغییرپذیری ویژگی زیاد است. بر همین اساس جرم مخصوص ظاهری خاک در کلاس

جدول ۱- خلاصه آماری داده‌های اندازه‌گیری شده از ویژگی‌های خاک

Table 1. Summary statistics of soil samples in study area

Name	Min.	Max.	Mean	Median	1st Qu	3rd Qu	CV	SD
Sand (%)	4.66	64.48	28.34	28.94	16.23	37.29	50.05	14.18
Silt (%)	26.54	65.6	43.91	43.82	39.18	48.7	16.34	7.17
Clay (%)	6	56	27.76	27	18.75	36.00	40.71	11.29
Organic matter (%)	0.27	4.42	1.61	1.61	1.14	2.01	42.86	0.69
Bulk Density (gr cm ³)	1.21	1.95	1.55	1.57	1.48	1.64	8.96	0.14
EC (dS m ⁻¹)	0.40	7.4	0.99	0.66	0.55	0.79	94.18	0.94
CaCO ₃ (%)	3.18	17.73	9.80	10.02	6.32	12.73	37.22	3.65

مورد بررسی قرار گرفتند. باین‌حال، در این موارد بهتر است متغیرهای جدیدی از طریق ترکیب متغیرهای مورد مطالعه ایجاد شوند که نه تنها رابطه قوی با متغیرهای اولیه داشته باشند بلکه کمترین میزان همبستگی را با یکدیگر ایجاد کنند.

جدول (۴) و (۵) ضرایب همبستگی بین متغیرهای مختلف حسگر رنگ با خصوصیات مختلف خاک را به ترتیب در دو حالت مرطوب و خشک نشان می‌دهد. نتایج ضرایب همبستگی بین متغیرهای مختلف رنگ و ویژگی‌های خاک متفاوت می‌باشد. مقادیر همبستگی بین متغیرهای به‌دست‌آمده از حسگر رنگ در نمونه‌های مرطوب بیشتر از نمونه‌های خشک بود. به همین دلیل از نتایج نمونه‌های مرطوب برای برآورد مقادیر در مدل‌سازی استفاده شد. در مطالعه‌ای که توسط موخوپادهایای و چاکرابورتی (Mukhopadhyay *et al.*, 2020) برای برآورد کربن آلی خاک با استفاده از حسگر رنگی نیکس پرو صورت گرفت، با بررسی همبستگی بین حالت خشک و مرطوب نتایج حاصل از حالت خشک برای برآورد کربن آلی قوی‌تر از حالت مرطوب بود. از داده‌های حسگر رنگ برای طبقه‌بندی رنگ خاک در مناطق استوایی نیز استفاده شده است (Mancini *et al.*, 2020). پس این حسگر قابلیت استفاده در شرایط آب و هوایی مختلف را دارد.

بررسی همبستگی متغیرهای مختلف سیستم‌های رنگ با ویژگی‌های خاک

الف) بررسی همبستگی بین سیستم‌های مختلف رنگ در برآورد خصوصیات خاک

در اینجا همبستگی بین متغیرهای حسگر رنگ به‌عنوان داده‌های ورودی برای مدل‌سازی مورد بررسی قرار گرفتند. در جداول (۲) و (۳) نتایج همبستگی بین متغیرهای حسگر رنگ در سیستم‌های مختلف در دو حالت خشک و مرطوب آمده است. استفاده هم‌زمان از این سیستم‌های مختلف در هنگام مدل‌سازی باعث ایجاد خطای مدل در برآورد خصوصیات مختلف خاک می‌شود. نتایج جدول (۴) نشان می‌دهد که همبستگی بین متغیر X و R که هر دو قرمز هستند و بین متغیرهای Y و G که سبز هستند در سیستم‌های رنگی RGB و XYZ (۰/۹۹) بود. همبستگی متغیر a* در سیستم رنگ CIEL*a*b* با متغیر X مقدار ۰/۶۷ بود و با متغیر R مقدار ۰/۷۲ بود. بیشترین تفاوت بین کدهای رنگی در سیستم رنگ CIEL*a*b* وجود داشت. برای برآورد خصوصیات مختلف خاک با استفاده از داده‌های حسگر رنگ، در نظر گرفتن همه متغیرهای مشتق شده از سیستم‌های رنگی مختلف، ضروری نیست. در اینجا متغیرهای سیستم رنگ CIEL*a*b* انتخاب شدند و

جدول ۲- ضرایب همبستگی بین متغیرهای حسگر رنگ NIXTMPro (حالت مرطوب)Table 2. Correlation between the NIXTMPro color sensor variables (in moist samples soil)

Parameters	L	a*	b*	X	Y	Z	R	G	B	C	M	Y	K
L	1	0.61	0.84	0.99	0.99	0.97	0.99	0.99	0.96	-0.89	-0.89	-0.28	-0.99
a*	0.61	1	0.8	0.67	0.62	0.49	0.72	0.56	0.45	-0.88	-0.21	0.17	-0.61
b*	0.84	0.8	1	0.85	0.83	0.69	0.89	0.81	0.65	-0.95	-0.6	0.25	-0.83
X	0.99	0.67	0.85	1	0.99	0.96	0.99	0.98	0.94	-0.91	-0.85	-0.28	-0.98
Y	0.99	0.62	0.83	0.99	1	0.97	0.99	0.99	0.96	-0.89	-0.88	-0.3	-0.99
Z	0.97	0.49	0.69	0.96	0.97	1	0.94	0.98	0.99	-0.79	-0.92	-0.5	-0.97
R	0.99	0.72	0.89	0.99	0.99	0.94	1	0.98	0.92	-0.95	-0.82	-0.2	-0.98
G	0.99	0.56	0.81	0.98	0.99	0.98	0.98	1	0.97	-0.86	-0.91	-0.32	-0.99
B	0.96	0.45	0.65	0.94	0.96	0.99	0.92	0.97	1	-0.75	-0.92	-0.52	-0.96
C	-0.89	-0.88	-0.95	-0.91	-0.89	-0.79	-0.95	-0.86	-0.75	1	0.62	-0.03	0.88
M	-0.89	-0.21	-0.6	-0.85	-0.88	-0.92	-0.82	-0.91	-0.92	0.62	1	0.42	0.89
Y	-0.28	0.17	0.25	-0.28	-0.3	-0.5	-0.2	-0.32	-0.52	-0.03	0.42	1	0.3
K	-0.99	-0.61	-0.83	-0.98	-0.99	-0.97	-0.98	-0.99	-0.96	0.88	0.89	0.3	1

جدول ۳- ضرایب همبستگی بین متغیرهای حسگر رنگ نیکس پرو (حالت خشک)

Table 3. Correlation between the NIXTMPro color sensor variables (in dry samples soil)

Parameters	L	a*	b*	X	Y	Z	R	G	B	C	M	Y	K
L	1	0.11	0.59	0.99	0.99	0.97	0.99	0.99	0.98	-0.95	-0.97	-0.9	-0.98
a*	0.11	1	0.46	0.14	0.1	0.02	0.23	0.06	0.02	-0.36	0.09	0.04	-0.12
b*	0.59	0.46	1	0.59	0.57	0.43	0.66	0.57	0.43	-0.72	-0.52	-0.19	-0.58
X	0.99	0.14	0.59	1	0.99	0.98	0.99	0.99	0.97	-0.97	-0.97	-0.9	-0.95
Y	0.99	0.1	0.57	0.99	1	0.99	0.98	0.99	0.98	-0.96	-0.98	-0.91	-0.95
Z	0.97	0.02	0.43	0.98	0.99	1	0.95	0.98	0.99	-0.91	-0.97	-0.96	-0.93
R	0.99	0.23	0.66	0.99	0.98	0.95	1	0.98	0.96	-0.98	-0.95	-0.86	-0.97
G	0.99	0.06	0.57	0.99	0.99	0.98	0.98	1	0.98	-0.94	-0.98	-0.91	-0.98
B	0.98	0.02	0.43	0.97	0.98	0.99	0.96	0.98	1	-0.9	-0.97	-0.96	-0.96
C	-0.95	-0.36	-0.72	-0.97	-0.96	-0.91	-0.98	-0.94	-0.9	1	0.89	0.8	0.91
M	-0.97	0.09	-0.52	-0.97	-0.98	-0.97	-0.95	-0.98	-0.97	0.89	1	0.9	0.94
Y	-0.9	0.04	-0.19	-0.9	-0.91	-0.96	-0.86	-0.91	-0.96	0.8	0.9	1	0.86
K	-0.98	-0.12	-0.58	-0.95	-0.95	-0.93	-0.97	-0.98	-0.96	0.91	0.94	0.86	1

متغیرها از این نتایج به دست می‌آید، برای مطالعه استفاده شود.

مقدار همبستگی متغیرهای a* و b* در این سیستم رنگ با مقدار EC به ترتیب برابر با ۰/۴۰ و ۰/۵۲، با کربنات کلسیم (CaCO₃) به ترتیب برابر ۰/۵۹ و ۰/۶۵ و همچنین برای جرم مخصوص ظاهری خاک مقدار ضریب همبستگی با متغیرهای a* و b* به ترتیب برابر با ۰/۲۱ و ۰/۲۵ می‌باشد. مقدار ضریب همبستگی در سیستم رنگ XYZ با خصوصیات EC و CaCO₃ به ترتیب با میانگین ۰/۴۴ و ۰/۵۹ می‌باشد اما با سایر خصوصیات خاک همبستگی پایینی دارد.

در تحقیق حاضر نیز در مورد ماده آلی خاک، با بررسی معنی‌دار بودن داده‌ها در سطح ۰/۰۵ درصد، در حالت خشک نتیجه بهتری به دست آمد. مقدار ضریب همبستگی بین ماده آلی خاک و سیستم رنگ CIEL*a*b* در حالت خشک بین ۰/۱۳- (۰/۲۷) درصد قرار گرفت (جدول ۵). بنابراین، با مشاهده نتایج مطالعات مختلفی که بر روی کربن آلی در حالت خشک و مرطوب صورت گرفته، نیاز است که قبل از ورود داده‌ها به مدل‌سازی، همبستگی متغیرهای هر سیستم رنگی در حالت خشک و مرطوب مقایسه شود و بهترین حالتی که با توجه به معنادار بودن و همبستگی

جدول ۴- ضرایب همبستگی بین خصوصیات خاک و متغیرهای حسگر رنگ (حالت مرطوب)

Table 4. Correlation between the different soil properties and color sensor variables (in moist soil)

* indicates significance at the 0.05 level

Name	Sand	Silt	Clay	EC	Organic matter	CaCO ₃	Bulk Density
L	-0.19*	0.21*	0.11*	0.44*	-0.08*	0.58*	-0.15*
a*	-0.52*	0.21*	0.51*	0.40*	0.05 ^{ns}	0.59*	-0.22*
b*	-0.43*	0.22*	0.39*	0.52*	-0.04 ^{ns}	0.66*	-0.26*
X	-0.24*	0.23*	0.15*	0.48*	-0.07*	0.58*	-0.15*
Y	-0.21*	0.22*	0.12*	0.47*	-0.07*	0.57*	-0.14*
Z	-0.10*	0.19*	-0.01 ^{ns}	0.39*	-0.08*	0.48*	-0.08*
R	-0.27*	0.22*	0.20*	0.47*	-0.06*	0.62*	-0.17*
G	-0.15*	0.20*	0.069*	0.43*	-0.09*	0.56*	-0.13*
B	-0.06*	0.18*	-0.04 ^{ns}	0.35*	-0.08*	0.47*	-0.07*
C	0.41*	-0.24*	-0.37*	-0.49*	0.02 ^{ns}	-0.65*	0.21*
M	-0.03 ^{ns}	-0.14*	0.12*	-0.36*	0.12*	-0.39*	0.05 ^{ns}
Y	-0.30*	0.01 ^{ns}	0.37*	0.08*	0.03 ^{ns}	0.08*	-0.13*
K	0.18*	-0.19*	-0.10*	-0.42*	0.08*	-0.58*	0.15*

*معنی داری در سطح ۰/۰۵ درصد، ns عدم معنی داری در سطح ۰/۰۵ درصد

جدول ۵- ضرایب همبستگی بین خصوصیات خاک و متغیرهای حسگر رنگ (حالت خشک)

Table 5. Correlation between the different soil properties and color sensor variables (in dry soil)

* indicates significance at the 0.05 level

Name	Sand	Silt	Clay	EC	Organic matter	CaCO ₃	Bulk Density
L	-0.23*	0.20*	0.16*	0.32*	-0.13*	0.43*	-0.05 ^{ns}
a*	-0.05 ^{ns}	-0.04 ^{ns}	0.08*	0.08*	-0.19*	0.13*	-0.11*
b*	-0.18*	0.17*	0.11*	0.25*	-0.27*	0.12*	-0.10*
X	-0.23*	0.22*	0.15*	0.33*	-0.14*	0.42*	-0.05 ^{ns}
Y	-0.23*	0.23*	0.15*	0.32*	-0.13*	0.41*	-0.04 ^{ns}
Z	-0.28*	0.21*	0.14*	0.31*	-0.10*	0.42*	-0.02 ^{ns}
R	-0.23*	0.20*	0.17*	0.33*	-0.16*	0.43*	-0.06 ^{ns}
G	-0.23*	0.21*	0.16*	0.32*	-0.12*	0.42*	-0.04 ^{ns}
B	-0.22*	0.19*	0.15*	0.3*	-0.08*	0.45*	-0.03 ^{ns}
C	0.23*	-0.21*	-0.15*	-0.34*	0.20*	-0.39*	0.07*
M	0.23*	-0.23*	-0.14*	-0.32*	0.10*	-0.39*	0.03 ^{ns}
Y	0.19*	-0.16*	-0.14*	-0.27*	0.03*	-0.45*	0.006 ^{ns}
K	0.22*	-0.16*	-0.17*	-0.31*	0.19*	-0.47*	0.05 ^{ns}

*معنی داری در سطح ۰/۰۵ درصد، ns عدم معنی داری در سطح ۰/۰۵ درصد

مختلف خاک از خود نشان داد. در این مطالعه سیستم رنگ CIEL*a*b* مورد بررسی قرار گرفت. نتایج حاصل از تجزیه و تحلیل این سیستم در مورد خصوصیات مختلف خاک شامل ذرات خاک (شن، سیلت، رس)، جرم ظاهری خاک، مقدار ماده آلی خاک، میزان شوری و کربنات کلسیم نشان می‌دهد که این سیستم رنگ می‌تواند برای برآورد این خصوصیات مورد استفاده قرار گرفت. در مطالعه دیگری که توسط جا و همکاران (Jha et al. 2021) بر روی مقدار آهن (Fe) خاک با استفاده از حسگر رنگ نیکس پرو انجام شد، از سه سیستم رنگ شامل CIEL*a*b*، RGB و CMYK

همان‌طور که نتایج جدول (۴) نشان می‌دهد در سیستم رنگی CMYK دامنه تغییرات ضریب همبستگی از 0.01 تا 0.45 متفاوت می‌باشد. به‌طور کلی سیستم رنگ CMYK با خصوصیاتمانند جرم مخصوص ظاهری، مقادیر شن، سیلت و رس همبستگی پایینی را نشان دادند. به‌طوری‌که متغیر M با جرم مخصوص ظاهری و میزان شن رابطه معناداری نداشتند. همچنین متغیر Y با میزان سیلت و ماده آلی معنادار نبودند. در مجموع، از بین چهار سیستم رنگی حسگر، سیستم رنگ CIEL*a*b* بیشترین میزان همبستگی را نسبت به سایر سیستم‌های رنگ با خصوصیات

طبقه‌بندی را دارند، انتخاب شوند. در این مطالعه، از مدل طبقه‌بندی الگوریتم جنگل تصادفی برای پیاده‌سازی الگوریتم بازگشتی استفاده شد. سپس با استفاده از معیار اهمیت ویژگی، اهمیت هر متغیر محاسبه شد. دقت طبقه‌بندی با استفاده از روش cross 10-Fold validation محاسبه و مهم‌ترین ویژگی‌ها که بالاترین دقت را مدل‌سازی دارند، انتخاب شدند (جدول ۶). نتایج این جدول مهم‌ترین ویژگی‌های سیستم‌های مختلف رنگ را در ارتباط با هر خصوصیت نشان می‌دهد. دو متغیر A و B از بین سه متغیر سیستم رنگ $CIEL^*a^*b^*$ در تمام خصوصیات خاک مهم‌ترین ویژگی بودند. همچنین دو متغیر از بین چهار متغیر سیستم رنگ CMYK مهم‌ترین ویژگی‌ها در بین متغیرهای مختلف سیستم رنگ بودند. در سیستم رنگ RGB با سه متغیر، تنها متغیر B با ویژگی‌های خاک مرتبط بود.

استفاده شد. سیستم رنگ CMYK نسبت به دو سیستم دیگر نتایج بهتری از برآورد میزان آهن ارائه داد. اکثر مطالعات انجام‌شده با حسگر رنگ برای برآورد ماده آلی خاک می‌باشد (Mikhailova *et al.*, 2017a; Mikhailova *et al.*, 2017b; Stiglitz *et al.*, 2018).

ب) انتخاب ویژگی با استفاده از حذف ویژگی به روش بازگشتی (REF)

در این روش ابتدا طبقه‌بندی با استفاده از همه متغیرهای حسگر انجام شده و برای هر متغیر یک مقدار اهمیت محاسبه می‌شود. سپس با استفاده از مهم‌ترین متغیرها طبقه‌بندی انجام شده و دقت آن محاسبه می‌شود. این کار برای هر تعداد ویژگی مختلف از مهم‌ترین ویژگی‌ها تکرار می‌شود تا در نهایت ویژگی‌هایی که بیشترین دقت در

جدول ۶- مهم‌ترین متغیرهای هر سیستم رنگ با استفاده از روش حذف بازگشتی در ارتباط با ویژگی‌های خاک
Table 6. The importance variables of color systems using REF in relation to soil properties

Color Systems	CIEL* <i>a</i> * <i>b</i> *			XYZ			RGB			CMYK				
	Parameters	L	A	B	X	Y	Z	R	G	B	C	M	Y	K
Sand		*	*					*			*	*		
Silt		*	*					*			*	*		
Clay		*	*					*			*	*		
Organic matter		*	*					*			*	*		
Bulk Density		*	*					*			*	*		
EC		*	*					*			*	*		
CaCO ₃		*	*					*			*	*		

می‌توان به سه دسته ضعیف ($\leq 1/6$)، قابل قبول ($1/6 - 2$) و عالی (≥ 2) طبقه‌بندی کرد (Bellon-Maurel *et al.*, 2010; Niazi *et al.*, 2015). میزان این آماره در مورد مقادیر کربنات کلسیم ($3/11$)، رس ($2/26$) و شن ($2/25$) عالی می‌باشد.

میانگین خطا (ME) معیاری از وجود اریب در برآورد مدل است و هرچه به صفر نزدیک‌تر باشد، دقت مدل بیشتر است. مقادیر منفی و مثبت آماره ME نشان می‌دهد که مقدار برآورد به ترتیب کمتر و بیشتر از اندازه‌گیری شده می‌باشد. مقدار این آماره در شن برابر با $0/27$ می‌باشد که نشان دهنده بیش برآورد مدل می‌باشد. در مورد سیلت برابر با $0/11$ است که نشان‌دهنده کم برآورد بودن مدل در مورد ذرات سیلت است. مقادیر این آماره در مقدار ماده آلی، شوری و جرم مخصوص ظاهری بسیار ناچیز و نزدیک به صفر است که نشان‌دهنده دقت بالای مدل در برآورد این خصوصیات می‌باشد.

برآورد خصوصیات خاک با استفاده از روش حذف ویژگی بازگشتی (REF) در مدل جنگل تصادفی بر اساس متغیرهای مختلف سیستم‌های رنگ

پس از تعیین مهم‌ترین متغیرهای هر سیستم، اقدام به برآورد مقادیر مختلف خصوصیات خاک شد. نتایج اعتبارسنجی حاصل از مدل جنگل تصادفی در جدول (۷) نشان داده شده است. میزان خطای RMSE در جرم مخصوص ظاهری دارای کمترین مقدار ($0/11$) و در برآورد مقدار شن دارای بیش‌ترین مقدار ($9/33$) است. میزان ضریب همبستگی و آماره خطای RMSE در کربنات کلسیم به ترتیب برابر با $0/80$ و $2/06$ است که بیش‌ترین میزان ضریب همبستگی در بین خصوصیات خاک می‌باشد. در مورد مقدار شوری خاک ضریب همبستگی $0/77$ و مقدار خطا برابر با $0/58$ می‌باشد که نتیجه قابل قبولی در برآورد حاصل از متغیرهای مختلف سیستم رنگ نیکس پرو می‌باشد. مناسب بودن یک مدل را می‌توان با توجه به آماره PRIQ نشان داد. برای تجزیه و تحلیل خاک، مقادیر PRIQ را

جدول ۷- نتایج اعتبار سنجی مدل جنگل تصادفی در برآورد خصوصیات مختلف خاک با حسگر رنگ تصادفی بر اساس متغیرهای مختلف سیستم های رنگ

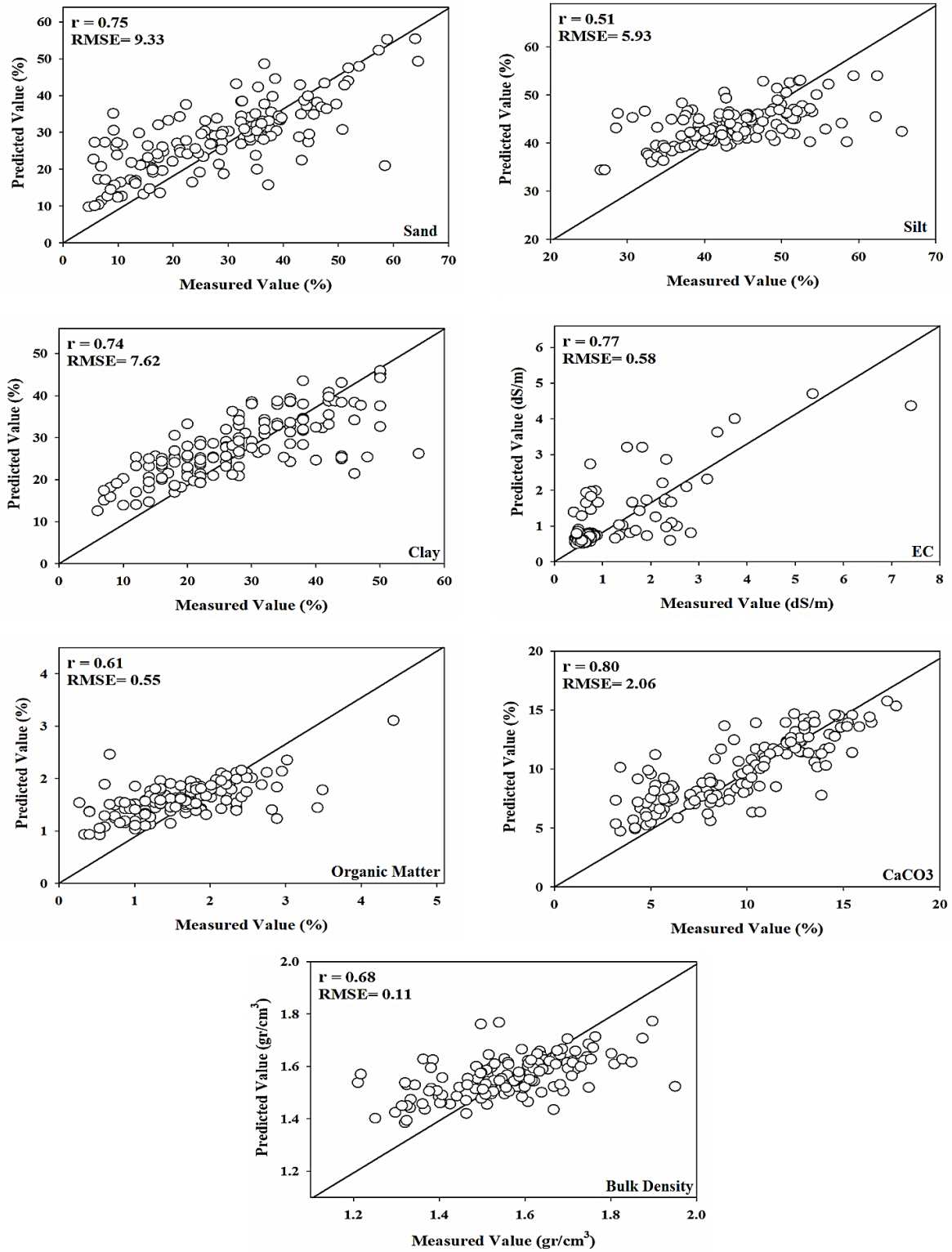
Table 7. Validation results of the Random Forest Model in predicating different soil properties using color sensor based on different variables of color systems

Name	Sand	Silt	Clay	EC	Organic matter	CaCO ₃	Bulk Density
RMSE(%)	9.33	5.93	7.62	0.58	0.61	2.06	0.11
r	0.75	0.51	0.74	0.77	0.55	0.80	0.68
PRIQ	2.25	1.60	2.26	0.41	1.42	3.11	1.45
ME	0.27	-0.11	0.16	-0.006	-0.007	0.15	0.001

برای برآورد ویژگی های خاک باشد اما از چهار متغیر این سیستم، دو متغیر آن با ویژگی های خاک ارتباط داشت.

بنابراین پس از بررسی همه متغیرهای سیستم های مختلف و بررسی نتایج همبستگی متغیرها در دو حالت خشک و مرطوب، سیستم رنگی که متغیرهای بیشتری از آن، در مدلسازی نقش داشتند، علاوه بر آن در بررسی حالت مرطوب و خشک نمونه های خاک، نتایج قابل قبول تری داشتند، به عنوان سیستم رنگ استاندارد انتخاب شد. در این تحقیق، سیستم رنگ $CIEL^*a^*b^*$ به عنوان سیستم استاندارد قابل قبول معرفی شد.

نمودار مربوط به هر کدام از خصوصیات در شکل (۳) نشان داده شده است. همانطور که گفته شد این نتایج مربوط به یک سیستم رنگ استاندارد نمی باشد. در واقع در اینجا تمام متغیرهای سیستم های مختلف رنگ مورد بررسی قرار گرفته است. هدف از این مطالعه معرفی یک سیستم رنگ استاندارد است که علاوه بر برآورد نتایج قابل قبول، بتواند تعداد متغیرهای ورودی برای مدلسازی را کاهش داده و بتواند در زمان صرفه جویی کند و برآوردی سریع و آسان ارائه دهد. همانطور که نتایج ضریب همبستگی نشان داد از سیستم رنگ $CIEL^*a^*b^*$ که دارای سه متغیر است، دو متغیر آن بیشترین همبستگی را با ویژگی های خاک داشت. سیستم رنگ $CMYK$ هم می تواند سیستم رنگ مناسبی



شکل ۳- نمودار مقادیر اندازه‌گیری شده در مقابل کل مقادیر برآورد شده برای خصوصیات مختلف خاک با استفاده از روش REF در مدل جنگل تصادفی بر اساس متغیرهای مختلف سیستم های رنگ

Figure 3. Scatter Plots of measured values versus predicted values for different soil properties by the random forest model based on different variables of color systems

قرار گرفت. این آماره می‌تواند معیار طبیعی‌تری برای نشان دادن میانگین خطا باشد (Willmott & Matsuura, 2005). مقدار ضریب r برای ذرات شن، سیلت و رس به ترتیب برابر با ۰/۷۱، ۰/۴۹ و ۰/۷۷ می‌باشد. همچنین برای مقدار ماده آلی برابر با ۰/۵۵ است که این نتیجه برای مدل RF نتیجه قابل قبولی است. در مطالعه ای که توسط وِثا و چاکرابورتی (Swetha & Chakraborty, 2021) روی ترکیب بافت خاک با حسگر رنگ نیکس پرو انجام شد، از رگرسیون جنگل تصادفی (RF) برای پیش‌بینی SOC خاک با ترکیب‌های مختلف داده‌های بافت خاک استفاده شد. این مدل نیز در مطالعه آن‌ها توانست نتایج رضایت‌بخشی از پیش‌بینی کربن آلی خاک با دقت اعتبار $r=0.70$ نشان دهد. میانگین خطای ME برای جرم مخصوص ظاهری خاک برابر با 0.004 که حداقل مقدار اریب (میانگین خطا) می‌باشد و بسیار نزدیک به مقدار ایده آل ME یعنی صفر است. در مورد میزان سیلت، رس و آهک ME منفی می‌باشد که نشان می‌دهد مقدار برآورد شده کمتر از میزان اندازه‌گیری شده است.

برآورد خصوصیات خاک با استفاده از مدل جنگل تصادفی

بر اساس سیستم رنگ $CIEL^*a^*b^*$

مقادیر میانگین مربعات خطا (RMSE)، ضریب همبستگی (r)، نسبت فاصله چارکی (PRIQ) و میانگین خطا (ME) برای هر ویژگی در جدول (۸) نشان داده شده است. بیشترین و کمترین مقدار ضریب همبستگی به ترتیب برای مقادیر کربنات کلسیم ۰/۷۹ و محتوای سیلت ۰/۴۹ می‌باشد. مقدار این ضریب برای ماده آلی خاک برابر با ۰/۵۵ و میزان خطای RMSE برابر با ۰/۵۷ می‌باشد. مقدار این آماره برای محتوای شن ۲/۰۹ و برای محتوای رس ۲/۶۷ می‌باشد که نشان‌دهنده پایایی قوی مدل RF برای برآورد این مقادیر می‌باشد. مقدار آماره PRIQ برای کربنات کلسیم برابر با ۱/۷۸ می‌باشد که مقداری قابل قبول است. همچنین مقدار شوری منطقه در اکثر نقاط کمتر از 2 dS m^{-1} است که ممکن است رنگ خاک آن را به خوبی نشان ندهد به همین دلیل مقدار آماره PRIQ در مورد EC خاک برابر با ۰/۴۷ می‌باشد. همچنین، میزان خطای مدل با استفاده از آماره ME در مورد خصوصیات مختلف خاک مورد بررسی

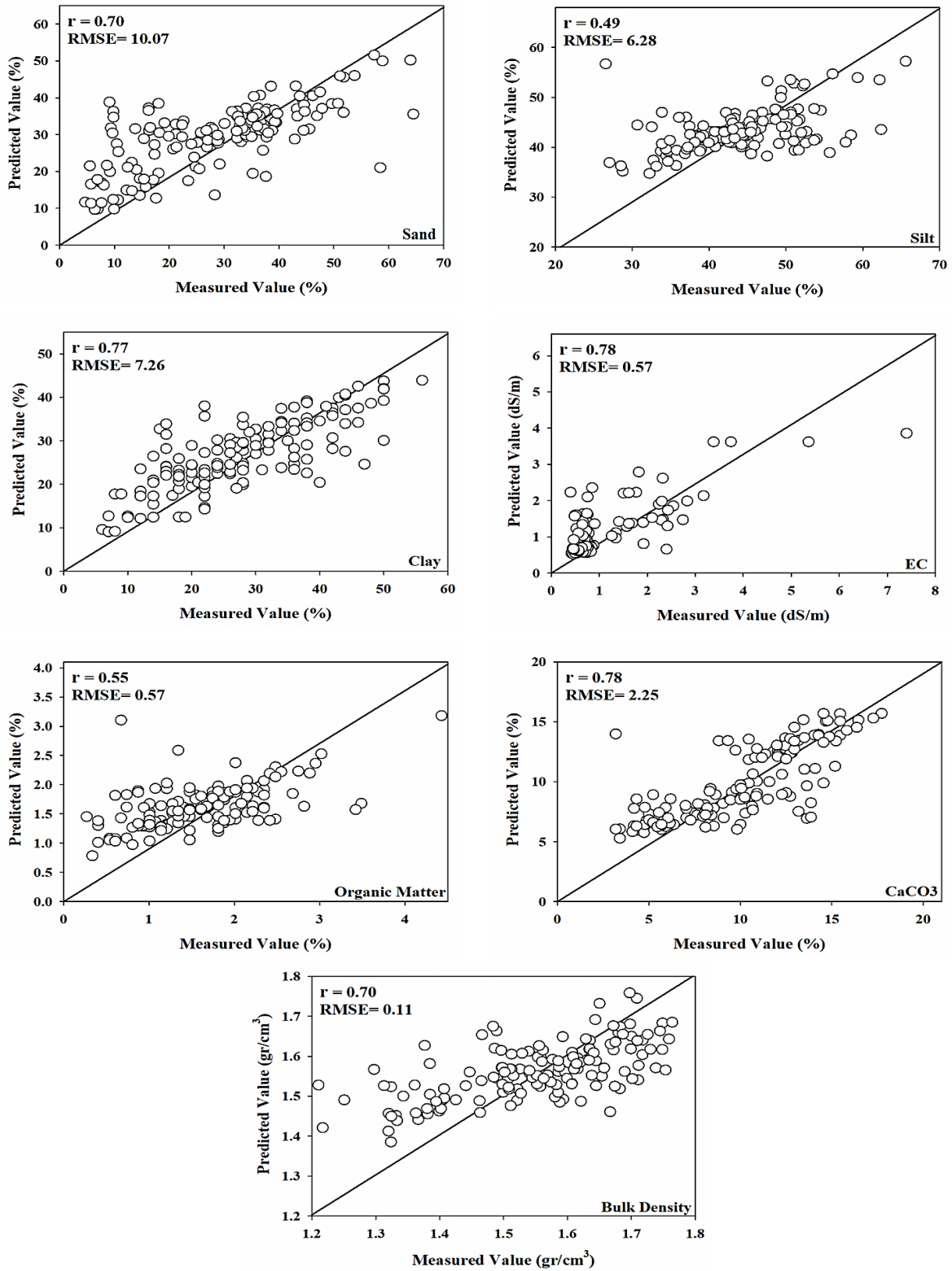
جدول ۸- نتایج اعتبار سنجی مدل جنگل تصادفی در برآورد خصوصیات مختلف خاک با حسگر رنگ بر اساس سیستم رنگ $CIEL^*a^*b^*$

Table 8. Validation results of the Random Forest Model in predicating different soil properties using color sensor based on $CIEL^*a^*b^*$ color system

Name	Sand	Silt	Clay	EC	Organic matter	CaCO ₃	Bulk Density
RMSE(%)	10.07	6.28	7.26	0.59	0.57	0.02	0.11
r	0.71	0.49	0.77	0.78	0.55	0.79	0.7
PRIQ	2.09	1.52	2.37	0.41	1.51	1.78	1.46
ME	1.08	-0.61	-0.81	0.05	0.02	-0.07	0.004

تصادفی مقادیر ذرات شن، سیلت، رس، ماده آلی، جرم مخصوص ظاهری، مقداری شوری و کربنات کلسیم خاک برآورد شدند. نمودار مربوط به هر یک از این خصوصیات در شکل (۴) نشان داده شده است.

در این مطالعه برای برآورد خصوصیات خاک از مدل جنگل تصادفی استفاده شد. بعد از اینکه سیستم $CIEL^*a^*b^*$ که در بررسی میزان همبستگی متغیرهای آن با خصوصیات خاک معنی‌دار بود، به عنوان سیستم مناسب برای برآورد خصوصیات خاک انتخاب شد، با استفاده از مدل جنگل



شکل ۴- نمودار مقادیر اندازه گیری شده در مقابل کل مقادیر برآورد شده برای خصوصیات مختلف خاک با استفاده از مدل جنگل تصادفی

بر اساس سیستم رنگ *a*b* CIEL*

Figure 4. Scatter Plots of measured values versus predicted values for different soil properties by the random forest model based on CIEL*a*b* color system

نتیجه‌گیری کلی

در این مطالعه قابلیت حسگر رنگ نیکس پرو برای برآورد سریع خصوصیات مختلف خاک مورد بررسی قرار گرفت. با توجه به بررسی میزان همبستگی بین نتایج حاصل از بررسی نمونه‌ها در دو حالت خشک و مرطوب، برای مدل‌سازی دقیق‌تر بهتر است که قبل از ورود داده‌ها به مدل، داده‌هایی انتخاب شوند که همبستگی بالاتری با خصوصیات خاک دارند. با بررسی متغیرهای مختلف سیستم‌های رنگ، با اعمال فیلتر ضریب همبستگی در مدل جنگل تصادفی، از متغیرهای چهار سیستم رنگ مختلف، سیستم رنگ $CIEL^*a^*b^*$ سه متغیره، با همبستگی بالای ۰/۹۵ درصد با ویژگی‌های مختلف خاک و عملکرد بهتر این سیستم در بررسی حالت خشک و مرطوب نمونه‌ها، به عنوان سیستم استاندارد معرفی شد. در اینجا میزان همبستگی مقدار ماده آلی در حالت خشک در سیستم رنگ $CIEL^*a^*b^*$ بین ۰/۰۴ تا ۰/۰۸ بود. در حالی که این مقدار در حالت مرطوب بین ۰/۱۳ تا ۰/۲۷ بود. در مورد سایر خصوصیات خاک، میزان همبستگی در حالت مرطوب بالاتر بود. نتایج نشان داد که داده‌های حسگر رنگ می‌تواند در جمع‌آوری اطلاعات سریع و ارزان خاک مفید باشند. علاوه بر این، دستگاه حسگر به راحتی قابل حمل است و می‌توان به سادگی در مزرعه از آن استفاده کرد. برنامه این حسگر به راحتی قابل نصب بر روی گوشی تلفن همراه است که می‌تواند طیف‌های رنگی

حاصل از نمونه برداری خاک را جمع‌آوری کند و با دقت قابل قبولی مورد ارزیابی قرار دهد. در این تحقیق، از مدل غیرخطی جنگل تصادفی (RF) برای برآورد خصوصیات خاک با استفاده از داده‌های حسگر رنگ نیکس پرو با نمونه‌های خاکی که از مناطق نیمه‌خشک ایران جمع‌آوری شده بود، استفاده شد. ارزیابی مدل جنگل تصادفی در برآورد این خصوصیات نشان داد که این مدل غیرخطی می‌تواند برآورد قابل قبولی در برآورد این خصوصیات نشان دهد. استفاده از این حسگر، سریع، قابل تکرار و قابل اعتماد است زیرا می‌توان با استفاده از آن بدون نیاز به تجهیزات دقیق و با کاهش نیاز به تجزیه و تحلیل آزمایشگاهی در شرایط کمبود منابع به عنوان ابزاری مفید و کارآمد مورد استفاده قرار گیرد.

منابع مالی

این مطالعه توسط صندوق حمایت از پژوهشگران و فناوران (INSF) با شماره ۹۷۰۱۲۵۵۷ پشتیبانی شد. نویسندگان از حمایت‌های مالی ارائه شده توسط INSF سپاسگزار هستند.

References

- Aitkenhead M.J., Coull M., Towers W., Hudson G. and Black, H.I.J. 2013. Prediction of soil characteristics and colour using data from the National Soils Inventory of Scotland. *Geoderma*, 200: 99-107.
- Anderson, S. 2005. *Soils: Genesis and Geomorphology*. Cambridge University Press, Cambridge, 13 817p.
- Barthès B.G., Kouakoua E., Clairotte M., Lallemand J., Chapuis-Lardy L., Rabenarivo M. and Rousset S. 2019. Performance comparison between a miniaturized and a conventional near infrared reflectance (NIR) spectrometer for characterizing soil carbon and nitrogen. *Geoderma* 338: 422-429.
- Baumgardner M.F., Silva L.F., Biehl L.L., Stoner Baumgardner M.F., Silva L.F., Biehl L.L. and Stoner E.R. 1986. Reflectance properties of soils. *Advances in agronomy*, 38: 1-44.
- Bellon-Maurel V., Fernandez-Ahumada E., Palagos B., Roger J.M. and McBratney A. 2010. Critical review of chemometric indicators commonly used for assessing the quality of the prediction of soil attributes by NIR spectroscopy. *TrAC Trends in Analytical Chemistry*, 29(9): 1073-1081.
- Dos Santos J.C.B., Le Pera E., de Souza Júnior V.S., de Oliveira C.S., Juilleret J., Corrêa M.M. and de Azevedo A.C. 2018. Porosity and genesis of clay in gneiss saprolites: the relevance of saprolithology to whole regolith pedology. *Geoderma*, 319: 1-13.
- Gee G.W. and Or D. 2002. 2.4 Particle-size analysis. *Methods of soil analysis: Part 4 physical methods, Soil Science Society of America Book Series*, pp. 255-293
- Hastie T., Tibshirani R., Friedman J.H. and Friedman, J.H. 2009. *The elements of statistical learning: data mining inference, and prediction*. New York, springer, 758p.
- Ibáñez-Asensio S., Marques-Mateu A., Moreno-Ramón H. and Balasch, S. 2013. Statistical relationships between soil colour and soil attributes in semiarid areas. *Biosystems Engineering*, 116(2): 120-129.
- Jha G., Sihi D., Dari B., Kaur H., Nocco M.A., Ulery A. and Lombard K. 2021. Rapid and inexpensive assessment of soil total iron using Nix Pro color sensor. *Agricultural & Environmental Letters*, 6(3): e20050.
- Kirillova N.P., Sileva T.M., Ul'yanova T.Y., Smirnova I.E., Ul'yanova A.S. and Burova E.K. 2018. Color diagnostics of soil horizons (by the example of soils from Moscow region). *Eurasian Soil Science*, 51(11): 1348-1356.
- Kirillova N.P., Vodyanitskii Y.N. and Sileva, T.M. 2015. Conversion of soil color parameters from the Munsell system to the CIE-L* a* b* system. *Eurasian soil science* 48(5): 468-475.
- Kirillova N.P., Grauer-Gray J., Hartemink A.E., Sileva T.M., Artemyeva Z.S. and Burova E.K. 2018. New perspectives to use Munsell color charts with electronic devices. *Computers and Electronics in Agriculture* 155: 378-385.
- Kuhn M. and Johnson K. 2013. *Applied predictive modeling*. New York, Springer, 600p.
- Levin N., Ben-Dor E. and Singer, A. 2005. A digital camera as a tool to measure colour indices and related properties of sandy soils in semi-arid environments. *International Journal of Remote Sensing*, 26(24): 5475-5492.
- Marqués-Mateu Á., Moreno-Ramón H., Balasch S. and Ibáñez-Asensio, S. 2018. Quantifying the uncertainty of soil colour measurements with Munsell charts using a modified attribute agreement analysis. *Catena*, 171: 44-53.
- Mikhailova E.A., Stiglitz R.Y., Post C.J., Schlautman M.A., Sharp J.L. and Gerard P.D. 2017. Predicting soil organic carbon and total nitrogen in the Russian Chernozem from depth and wireless color sensor measurements. *Eurasian Soil Science* 50(12): 1414-1419.
- Mikhailova E., Stiglitz R., Post C., Schlautman M.A., Sharp J. and Gerard, P. 2017. Developing Predictive Soil Organic C and N Models for Glaciated Soils Using Quantitative Color Sensor Measurements. *ASA, CSSA and SSSA International Annual (2017)*, (Abs.)
- Minasny B. and McBratney, A.B. 2008. Regression rules as a tool for predicting soil properties from infrared reflectance spectroscopy. *Chemometrics and intelligent laboratory systems* 94(1): 72-79.
- Mouazen A.M., Karoui R., Deckers J., De Baerdemaeker J. and Ramon H. 2007. Potential of visible and near-infrared spectroscopy to derive colour groups utilising the Munsell soil colour charts. *Biosystems Engineering*, 97(2): 131-143.

- Mukhopadhyay S. and Chakraborty S. 2020. Use of diffuse reflectance spectroscopy and Nix pro color sensor in combination for rapid prediction of soil organic carbon. *Computers and Electronics in Agriculture*, 176: 105630.
- Nasrollahi M., Zolfaghari A.A. and Yazdani M.R. 2021. Spatial and temporal properties of reference evapotranspiration and its related climatic parameters in the main agricultural regions of Iran. *Pure and Applied Geophysics*, 178(10): 4159-4179.
- Niazi N.K., Singh B. and Minasny B. 2015. Mid-infrared spectroscopy and partial least-squares regression to estimate soil arsenic at a highly variable arsenic-contaminated site. *International Journal of Environmental Science and Technology*, 12(6): 1965-1974.
- Raeesi M., Zolfaghari A.A., Yazdani M.R., Gorji M. and Sabetizade M. 2019. Prediction of soil organic matter using an inexpensive colour sensor in arid and semiarid areas of Iran. *Soil Research*, 57(3):276-286.
- Resende M., Curi N., Rezende S.D., Corrêa G.F. and Ker J.C. 2014. Pedologia base para distinção de ambientes. rev. ampl. *Lavras: Editora UFLA*
- Rezende É.A., 2021. Estudo da influência da Zona de Cisalhamento de Três Corações na ocorrência de voçorocamentos. *Revista de Geografia-PPGEO-UFJF*, 11(1): 120-135.
- Sabetizade M., Gorji M., Roudier P., Zolfaghari A.A. and Keshavarzi, A. 2021. Combination of MIR spectroscopy and environmental covariates to predict soil organic carbon in a semi-arid region. *Catena*, 196: 104844.
- Staff S.S. 2014. Keys to soil taxonomy. 13th Ed. United States Department of Agriculture: Washington, DC, USA, 436p.
- Stiglitz R., Mikhailova E., Post C., Schlautman M. and Sharp J. 2016. Evaluation of an inexpensive sensor to measure soil color. *Computers and Electronics in Agriculture*, 121: 141-148.
- Stiglitz R., Mikhailova E., Post C., Schlautman M. and Sharp J. 2017. Using an inexpensive color sensor for rapid assessment of soil organic carbon. *Geoderma*, 286: 98-103.
- Stiglitz R.Y., Mikhailova E.A., Sharp J.L., Post C.J., Schlautman M.A., Gerard P.D. and Cope M.P. 2018. Predicting soil organic carbon and total nitrogen at the farm scale using quantitative color sensor measurements. *Agronomy*, 8(10): 212.
- Swetha R.K. and Chakraborty, S. 2021. Combination of soil texture with Nix color sensor can improve soil organic carbon prediction. *Geoderma*, 382: 114775.
- Thompson J.A., Pollio A.R. and Turk, P.J. 2013. Comparison of Munsell soil color charts and the GLOBE soil color book. *Soil Science Society of America Journal*, 77(6): 2089-2093.
- Walkley A. and 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.
- Wilding L.P. 1985. Spatial variability: its documentation, accomodation and implication to soil surveys. In *Soil spatial variability, Las Vegas NV, 30 November-1 December, 1984*: 166-194
- Willmott C.J. and Matsuura K. 2005. Advantages of the mean absolute error (MAE) over the root mean square error (RMSE) in assessing average model performance. *Climate research*, 30(1): 79-82.

Estimation of Different Soil Properties Using Fast and Inexpensive Color Sensor Data

Fatemeh Cheshmberah¹, Ali. A Zolfaghari^{2*}, Ruhollah Taghizadeh-Mehrjardi³

(Received: November, 2021 Accepted: July, 2022)

Abstract

Soil color is one of the obvious characteristics of soil that usually has much to do with other soil properties. The Nix™Pro color sensor can detect different amounts of soil color and allow you to check different soil properties. This sensor is less sensitive to the user's environmental and mental conditions than the conventional method of Mansell manual to determine the color of the soil, and it is effortless to use. Therefore, in this study, Nix™Pro color sensor was used for fast and cheap estimation of different soil properties. For this purpose, 150 soil samples were collected from the semi-arid study area of Qazvin province. Values related to each soil characteristic were measured in the laboratory. Then, using this color sensor, the spectra of each color system were recorded for each soil sample. Two methods were used for this purpose. In the first method, by examining the correlation between the variables of the color system, an attempt was made to introduce a standard color system with the highest correlation coefficient with all soil properties. In the second method, all variables of different color systems were examined by the method of recursive feature elimination, which selects the features that have the highest accuracy by choosing the most important features. According to the results of both methods and to introduce a standard color system, in this study, CIEL*a*b color system was used to estimate soil properties. Because this system showed the highest degree of correlation with different soil properties. Then, using the random forest algorithm, the values related to each soil feature were estimated. Soil properties include sand, silt, clay, salinity, calcium carbonate (CaCO₃), organic matter and soil bulk density. According to the results of random forest forecast, the amount of root mean square error (RMSE), mean error (ME), ratio of performance to interquartile distance (PRIQ), and the value of correlation coefficient (r) for each soil feature were determined. For sand, silt and clay particles, the amount of RMSE was 10.07, 6.28, 7.26%, and the correlation was 0.70, 0.49, and 0.77, respectively. PRIQ statistics for sand particles (2.09), clay (2.37), and calcium carbonate (1.78) are at an excellent and acceptable level. The RMSE values of organic matter, calcium carbonate and soil bulk density were 0.57, 0.02, 0.11%, and the correlation coefficient were 0.55, 0.58, and 0.70, respectively. Based on these results, it can be said that the spectra obtained from the Nix™ Pro color sensor can be helpful in the rapid prediction of soil properties.

Keywords: soil color, Nix™ Pro color sensor, spectral system, random forest model

Cheshmberah F., Zolfaghari A.A. and Taghizadeh-Mehrjardi R. 2023. Estimation of Different Soil Properties Using Fast and Inexpensive Color Sensor Data. *Applied Soil Research*, 11(2): 28-45.

1. PhD in Combating Desertification, Semnan University

2. Associate Professor, Faculty of Desert Studies, Semnan University

3. Assistance Professor Faculty of Agriculture and Natural Resources Ardakan University and Department of Geosciences, Soil Science and Geomorphology, University of Tübingen, Tübingen, Germany

* Corresponding Author Email: azolfaghari@semnan.ac.ir