

واژه‌های کلیدی: شاخص‌های پوشش گیاهی، داده‌های محیطی، دورسنجی، مدل‌سازی و هم‌افزایی.

مقدمه

کربن آلی خاک^۱ نقش مهمی در ویژگی‌های شیمیایی، فیزیکی و زیستی خاک ایفا می‌کند و روی برخی از عملکردهای مهم خاک مانند کیفیت خاک، نفوذ آب و دسترسی به عناصر غذایی تأثیر می‌گذارد که بر پایداری کشاورزی تأثیر قابل‌ملاحظه‌ای دارد [۱۲]. کربن آلی خاک به‌عنوان یک مخزن پتانسیلی برای کربن اتمسفری عمل می‌کند و هم‌چنین به‌عنوان یک عامل محافظتی در برابر تهدیدات زمین مانند فشار خاک و فرسایش خاک عمل کرده است [۱۹]. علاوه بر این، میزان بالاتر کربن آلی خاک برای کنترل و تنظیم مهم‌ترین مسائل زیست‌محیطی از جمله تغییرات آب‌وهوا، جلوگیری از فرسایش خاک، افزایش نفوذ آب به خاک و غیره الزامی است. بنابراین، برآورد دقیق کربن آلی خاک می‌تواند مسیر کلیدی برای موفقیت در مدیریت کربن به‌عنوان بخش حساسی از چرخه کربن برای انتشار دی‌اکسید کربن به اتمسفر باشد [۱۴]. تحقیقات نشان داده که اندازه‌گیری میزان کربن آلی خاک به روش آزمایشگاهی، مخرب، ناپیوسته، دشوار، هزینه‌بر و زمان‌بر است [۱۰] و [۱۲]. در این راستا، داده‌های سنجش‌ازدور با پایش مستمر زمین با ماهواره‌های متنوع اطلاعات پیوسته‌ای را به‌عنوان روش جایگزین برای برآورد کربن آلی خاک در اختیار محققین و تصمیم‌گیرندگان در امور کشاورزی و محیط‌زیست قرار داده است. در حال حاضر، متداول‌ترین ماهواره‌های به کار گرفته‌شده جهت پایش تغییرپذیری کربن آلی خاک شامل لندست ۸ [۹]، سنتینل ۲ [۲۰] و مادیس [۲] هستند. باین‌حال، کربن آلی خاک با استفاده از ماهواره‌های مختلف با وضوح زمانی - مکانی متفاوت در بازه‌های ۷۰۰-۳۵۰ نانومتر (محدوده قابل‌رؤیت)، ۷۰۰-۱۱۰۰ نانومتر (محدوده نور نزدیک فروسرخ) و ۲۵۰۰-۱۱۰۰ نانومتر (محدوده مادون‌قرمز کوتاه موج) برآورد می‌شود. مشاهده‌شده است که داده‌های سنجش از منابع چندگانه به‌وضوح مکانی و زمانی بالا نیاز دارند تا دقت ویژگی‌های زمین را بهبود بخشند [۶]. به‌عنوان مثال لیاثو و همکاران [۵] از ترکیب زمانی - مکانی داده‌های تصویر لندست ۸ و سنجند مادیس برای بهبود فنولوژی، بیوماس و عملکرد ذرت و سویا استفاده کردند. علاوه بر این، لی و همکاران [۷] تصاویر لندست ۸ و سنتینل ۲ را

برآورد کربن آلی خاک با ترکیب تصاویر ماهواره‌ای لندست ۸ و مادیس و مدل‌سازی به روش یادگیری ماشین

پریناز عبدلی^۱، علی خانمیرزائی^{۱*}، سعید حمزه^۲، شکوفه رضائی^۱ و سمیه

مقیمی^۱
تاریخ دریافت ۱۴۰۳/۰۱/۲۸ تاریخ پذیرش ۱۴۰۳/۰۲/۰۳

DOI:10.22034/wmji.2024.2026606.1064

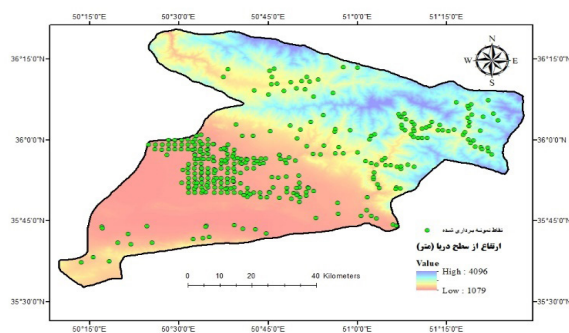
چکیده

هدف از این مطالعه، توسعه مدل‌های یادگیری ماشین جهت برآورد کربن آلی خاک با استفاده از تصاویر لندست ۸ و مادیس به‌صورت جداگانه و هم‌افزایی تصاویر فوق در خاک‌های کشاورزی در منطقه مرکزی ایران می‌باشد. در این مطالعه، نمونه‌های خاک از ۳۳۶ نقطه در خاک‌های کشاورزی در دو کاربری زراعی و باغی استان البرز جمع‌آوری شدند و کربن آلی خاک به روش اکسیداسیون تر اندازه‌گیری شد. برای تحلیل کربن آلی خاک از داده‌های سنجش‌ازدور استفاده شد که از منابع مختلف مانند تصویر لندست ۸ و سنجند مادیس به‌دست‌آمده بودند. در این مطالعه سه مدل یادگیری ماشینی شامل شبکه‌های عصبی مصنوعی (ANN)، رگرسیون بردار پشتیبان (SVR) و الگوریتم بیان ژن (GEP) استفاده شدند. این مدل‌ها به‌منظور برآورد میزان کربن آلی خاک استفاده شدند. نتایج این پژوهش نشان داد که به‌کارگیری مدل SVR در حالت استفاده هم‌زمان از داده‌های تصویر لندست ۸ و سنجند مادیس عملکرد بهتری ($R^2=0.72$ ، $RMSE=0.63$ ، $R^2=0.98$) نسبت به استفاده هر تصویر به‌صورت جداگانه داشت؛ بنابراین، این پژوهش به این نتیجه رسید که استفاده هم‌زمان (تحت عنوان هم‌افزایی) داده‌های سنجش‌ازدور از منابع مختلف می‌تواند بهبود قابل توجهی در دقت برآورد محتوای کربن آلی در خاک‌های کشاورزی در منطقه مرکزی ایران ایجاد کند. استفاده از تکنیک‌های فیوژن پیشرفته یا تکنیک‌های یادگیری عمیق برای ترکیب ویژگی‌ها در سطح بالاتر پیشنهاد می‌شود.

۱- گروه خاکشناسی، واحد کرج، دانشگاه آزاد اسلامی، کرج، ایران.

* ایمیل نویسنده مسئول: alikhanmirzai@yahoo.com

۲- گروه سنجش از دور و GIS، دانشگاه تهران، تهران، ایران



شکل ۱: منطقه مورد مطالعه (استان البرز- ایران)

جدول ۱: شاخص های طیفی استخراج شده از تصویر لندست ۸ و سنجنده مادیس

شاخص	فرمول	شاخص	فرمول
NDVI	(NIR-Red)/(NIR+Red)	GNDVI	(NIR-Green)/(NIR+Green)
RVI	(NIR)/(Red)	GRVI	(NIR)/(Green)
IPVI	(NIR)/(NIR+Red)	VARI	(Green-Red)/(Green+Red-Blue)
DVI	(NIR-Red)	NDWI	(Green-NIR)/(Green+NIR)
SAVI	1.5(NIR-Red)/(NIR+Red+0.5)	NGBDI	(Green-Red)/(Green+Blue)
GDVI	(NIR-Green)	NGRDI	(Green-Red)/(Green+Red)

کربن آلی خاک در خاک های کشاورزی تمرکز داشته باشد. هدف از این مطالعه توسعه برخی از مدل های یادگیری ماشین با استفاده از اطلاعات استخراج شده از تصویر لندست ۸ و سنجنده مادیس و نیز بررسی امکان بهبود دقت برآورد کربن آلی خاک با استفاده از هم افزایی تصویر لندست ۸ و سنجنده مادیس می باشد.

مواد و روش ها

منطقه مورد مطالعه

مساحت استان البرز ۵۱۲۵ کیلومتر مربع است. این منطقه در عرض جغرافیایی شمالی و طول جغرافیایی شرقی به ترتیب در محدوده $35^{\circ} 32' 68''$ تا $36^{\circ} 21' 75''$ و $50^{\circ} 7' 84''$ تا $51^{\circ} 28' 73''$ قرار دارد و ارتفاع در این منطقه از ۱۰۷۹ تا ۴۰۹۶ متر از سطح دریا متغیر است (شکل ۱). دمای سالانه متوسط ۱۴/۱ درجه سانتی گراد با حداقل ۲۰- درجه سانتی گراد و حداکثر ۴۲ درجه سانتی گراد در بازه زمانی سال های ۱۳۸۰ تا ۱۴۰۱ است. میانگین بارش سالانه ۲۵۱ میلی متر در سال با حداقل ۸۹/۳ میلی متر و حداکثر ۳۷۴/۴ میلی متر در سال است. منطقه مورد مطالعه عمدتاً شامل اراضی کشاورزی قابل کشت (۷۷۴/۴ کیلومتر مربع، ۱۵/۱ درصد)، اراضی کشاورزی باغی (۱۲۹/۷ کیلومتر مربع، ۲/۵ درصد)، مرتع ها (۳۶۱۶/۲ کیلومتر مربع، ۶۹/۸ درصد)، جنگل (۱/۱ کیلومتر مربع، ۰/۰۲ درصد)، شهری (۲۱۷/۷ کیلومتر مربع، ۴/۲ درصد)، آب (۳/۶ کیلومتر مربع، ۰/۱ درصد)، زمین های بایر (۴۰۶/۳ کیلومتر مربع، ۷/۸ درصد) و زمین های آبی (۲۸/۴ کیلومتر مربع، ۰/۵ درصد) است.

ترکیب کردند تا سری تصاویر با وضوح ۱۰ متری را برای افزایش دقت نظارت سطح زمین فراهم شود. تعداد زیادی از مطالعات بر برآورد میزان کربن آلی خاک با استفاده از رگرسیون خطی چندگانه تمرکز کرده اند [۱۸]. دلایل اصلی استفاده گسترده از رگرسیون خطی چندگانه تفسیر مستقیم و قابل فهم، کارایی و سادگی آن است [۱۶]. با این حال، رابطه بین محتوای کربن آلی خاک (به عنوان متغیر خروجی) و داده های سنجنده لندست ۸ (به عنوان متغیر ورودی) به طور عمده غیرخطی است. برای پاسخ به این مسئله، گزارش شده است که مدل های یادگیری ماشین از جمله شبکه های عصبی مصنوعی^۱، ماشین بردار پشتیبان^۲ و غیره، برآوردهای دقیق تری نسبت به برآورد و نقشه برداری میزان کربن آلی خاک با استفاده از رگرسیون خطی چندگانه ارائه می دهند [۴]. یک چالش اصلی انتخاب بهترین فن مدل سازی برای برآورد محتوای کربن آلی خاک است. بسیاری از مطالعات با استفاده از داده هایی که از انواع مختلف استفاده می شوند، انجام شده اند. به همین دلیل، مدل های به دست آمده در مطالعات قبلی بیش تر قادر به برآورد محتوای کربن آلی خاک در انواع مختلف کاربری اراضی، به خصوص در خاک های کشاورزی، نیستند. برآورد کربن آلی خاک در خاک های کشاورزی به ویژه اهمیت زیادی دارد زیرا از دست دادن کربن آلی خاک سریع تر از یافتن کربن آلی خاک است [۱۷]. با این حال پژوهش های کمتری برای برآورد میزان

1. Artificial Neural Networks
2. Support Vector Machine

داده‌های سنجنش ازدور از تصاویر لندست ۸ (OLI) و سنجنده مادیس (Terra MOD09A1.006) در مردادماه ۱۳۹۸ استخراج شده‌اند. تصویر لندست ۸ با دقت زمانی ۱۶ روز و دقت مکانی ۳۰ متر و سنجنده مادیس با دقت زمانی ۲-۱ روز و دقت مکانی ۵۰۰ متر از سامانه گوگل ارث انجین دریافت و شاخص‌های طیفی و بازتاب طیفی باندها از تصویر لندست ۸ و سنجنده مادیس استخراج شدند. شاخص‌های طیفی مورد استفاده در این مطالعه، در جدول ۱ ارائه شده‌اند.

به دلیل محدودیت‌های فیزیکی در طراحی سنسورها، دستیابی به یک تصویر با کیفیت طیفی و مکانی بالا (به‌طور هم‌زمان) امکان‌پذیر نیست. از این رو محققان به دنبال راه‌کارهایی برای بهبود کیفیت تصاویر ماهواره‌ای بوده‌اند؛ به‌گونه‌ای که تصویر به‌دست‌آمده هم دقت مکانی بالا و هم دقت طیفی مناسبی داشته باشد. یکی از انواع این روش‌ها فن ادغام تصاویر^۱ و دیگری هم‌افزایی تصاویر (استفاده هم‌زمان از دیتاهای دو تصویر) است.

بررسی خاک، تجزیه و تحلیل نمونه‌ها و آزمایش‌های خاک

در این مطالعه، نمونه‌های خاک از ۳۳۶ نقطه انتخاب شده به‌صورت تصادفی از خاک سطحی (عمق صفر تا ۳۰ سانتی‌متر) گرفته شدند. نمونه‌های خاک در آزمایشگاه هوا خشک و سپس خاک خشک‌شده با استفاده از الک دو میلی‌متری غربال شدند. میزان کربن آلی خاک با استفاده از روش اکسیداسیون تر [۱۳] تعیین شد.

فن‌های مدل‌سازی: در بخش حاضر، سه روش یادگیری ماشین از جمله شبکه‌های عصبی مصنوعی پرسپترون چندلایه (ANNs)، رگرسیون بردار پشتیبان (SVR) و برنامه‌ریزی بیان ژن (GEP) برای برآورد میزان کربن آلی خاک استفاده شد.

از شبکه عصبی مصنوعی با ساختار پرسپترون چندلایه، شبکه‌های پس‌انتشار، شبکه‌های پیش‌خور تعمیم‌یافته و غیره می‌باشد که در تحقیق حاضر، از میان آن‌ها شبکه‌های عصبی با ساختار پرسپترون چندلایه با الگوریتم آموزشی پس‌انتشار به دلیل محبوبیت و کاربرد ساختار استفاده شد. هم‌چنین، از الگوریتم بیان ژن که یک روش تکاملی مبتنی بر ژنتیک است استفاده شد که برای بهینه‌سازی و کشف مدل‌های پیچیده در داده‌ها به کار می‌رود. این الگوریتم از کروموزوم‌های خطی برای نمایش فرمول‌ها یا مدل‌ها استفاده می‌کند که می‌توانند به ساختارهای درخت مانند تبدیل شوند. از فرآیندهای تکاملی مانند جهش، کراس‌اور (تلاقی ژنتیکی) و تکثیر برای ایجاد نسل‌های جدیدی از راه‌حل‌های بهبودیافته استفاده می‌کند. افزون بر این، در تحقیق حاضر رگرسیون بردار پشتیبان که یک فن یادگیری ماشین مبتنی بر نظریه رگرسیون است برای پیش‌بینی و تحلیل داده‌های پیوسته استفاده شد. این روش از مفهوم حاشیه‌ها و بردارهای پشتیبان برای ایجاد یک مدل پیش‌بینی‌کننده استفاده می‌کند که هدف آن کمینه کردن خطای پیش‌بینی است.

متغیرهای ورودی برای تولید مدل‌های مختلف

متغیرهای ورودی (شاخص‌های طیفی و بازتاب طیفی باندها) در نرم‌افزار ArcGIS استخراج شدند. حالت‌های زیر با متغیرهای ورودی مختلف برای استخراج مدل‌های شبکه عصبی مصنوعی پرسپترون چندلایه، الگوریتم بیان ژن و ماشین بردار پشتیبان در هر کلاس استفاده از زمین در نظر گرفته شدند: حالت اول: شاخص‌های طیفی و بازتاب طیفی باندها از تصویر لندست ۸ و سنجنده مادیس به‌صورت جداگانه استخراج شده‌اند. حالت دوم: استفاده هم‌زمان (هم‌افزایی) شاخص‌های طیفی و بازتاب طیفی استخراج شده از تصویر لندست ۸ و سنجنده مادیس.

تحلیل آماری

برای استخراج مدل‌های برآورد یادگیری ماشین، داده‌های نمونه به‌طور تصادفی به دو زیرمجموعه تقسیم شدند. ۷۰ درصد از داده‌های نقطه‌ای برای استخراج مدل‌های یادگیری ماشین استفاده شدند و ۳۰ درصد دیگر از داده‌ها برای آزمون مدل‌ها در نظر گرفته شدند. عملکرد مدل‌های استخراج شده با استفاده از شاخص‌های آماری مانند خطای میانگین مربعات مطلق (RMSE)، ضریب تعیین (R^2)، نسبت انحراف عملکرد (RPD) و نسبت $R^2/$ RMSE ارزیابی شد.

$$R^2 = \frac{[\sum_{i=1}^n (Z_i - \bar{Z}_i)(Z_i^* - \bar{Z}_i^*)]^2}{\sum_{i=1}^n (Z_i - \bar{Z}_i)^2 \sum_{i=1}^n (Z_i^* - \bar{Z}_i^*)^2} \quad (1)$$

$$RMSE = \sqrt{\left(\frac{1}{n}\right) \sum_{i=1}^n (Z_i^* - Z_i)^2} \quad (2)$$

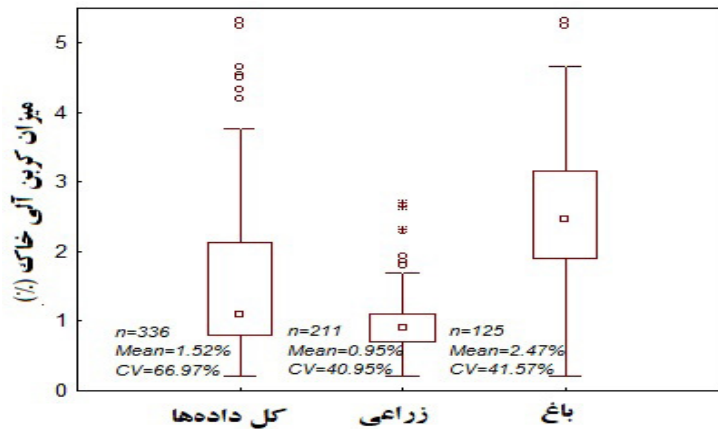
$$RPD = \frac{SD}{RMSE} \quad (3)$$

در اینجا، Z_i به معنای کربن آلی خاک برآورد شده، Z_i^* نشان‌دهنده کربن آلی خاک مشاهده‌شده است، کربن آلی خاک برآورد شده نشان‌دهنده میانگین کربن آلی خاک مشاهده‌شده، n نشان‌دهنده تعداد داده‌های کربن آلی خاک و Sd نشان‌دهنده انحراف معیار است.

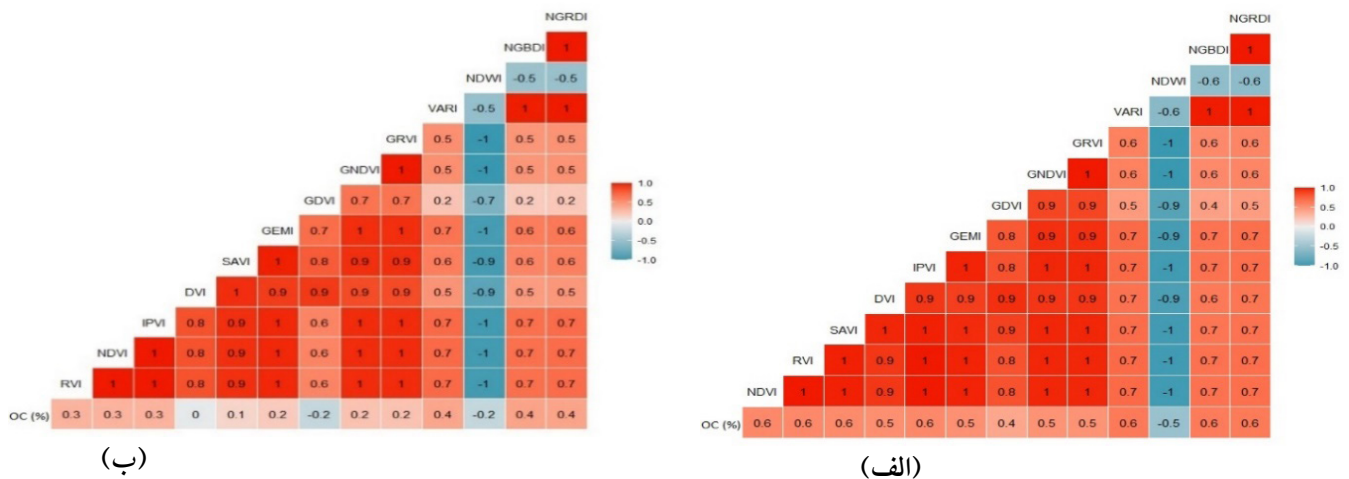
نتایج و بحث

ویژگی‌های خاک در منطقه مورد مطالعه

کربن آلی خاک اندازه‌گیری شده در خاک‌های کشاورزی در این منطقه مقادیر گسترده‌ای از ۰/۲۱ تا ۵/۳۳ درصد برای کل مجموعه داده، ۰/۲۱-۲/۷۰ درصد برای خاک‌های کشاورزی زراعی و ۰/۲۱-۵/۳۳ درصد برای خاک‌های کشاورزی باغی داشتند (شکل ۲). محتوای کربن آلی خاک اندازه‌گیری شده در مطالعه حاضر (شکل ۲) قابل‌مقایسه با مقادیری است که توسط میرزایی و همکاران [۱۰] در محدوده ۰/۱۷-۲/۶۲ درصد، اکبری و همکاران [۱] در محدوده ۰/۳۱۲-۱/۹۰۶ درصد و نوار و موازن [۱۲] در محدوده ۰/۹۴-۳/۱



شکل ۲: آمار توصیفی کربن آلی نمونه‌های مورد مطالعه به تفکیک کاربری اراضی.



شکل ۳: همبستگی بین مقدار شاخص‌های پوشش گیاهی حاصل از تصویر لندست ۸ (الف) و مادیس (ب) با میزان کربن آلی خاک

حضور ارتباط بین محتوای کربن آلی خاک و شاخص‌های طیفی سنجش‌ازدور (شکل ۳) نشان می‌دهد که نیاز به تجزیه و تحلیل دقیق برای به دست آوردن یک مدل یادگیری ماشین مناسب وجود دارد.

توسعه مدل‌های یادگیری ماشین با استفاده از اطلاعات سنجنده مادیس

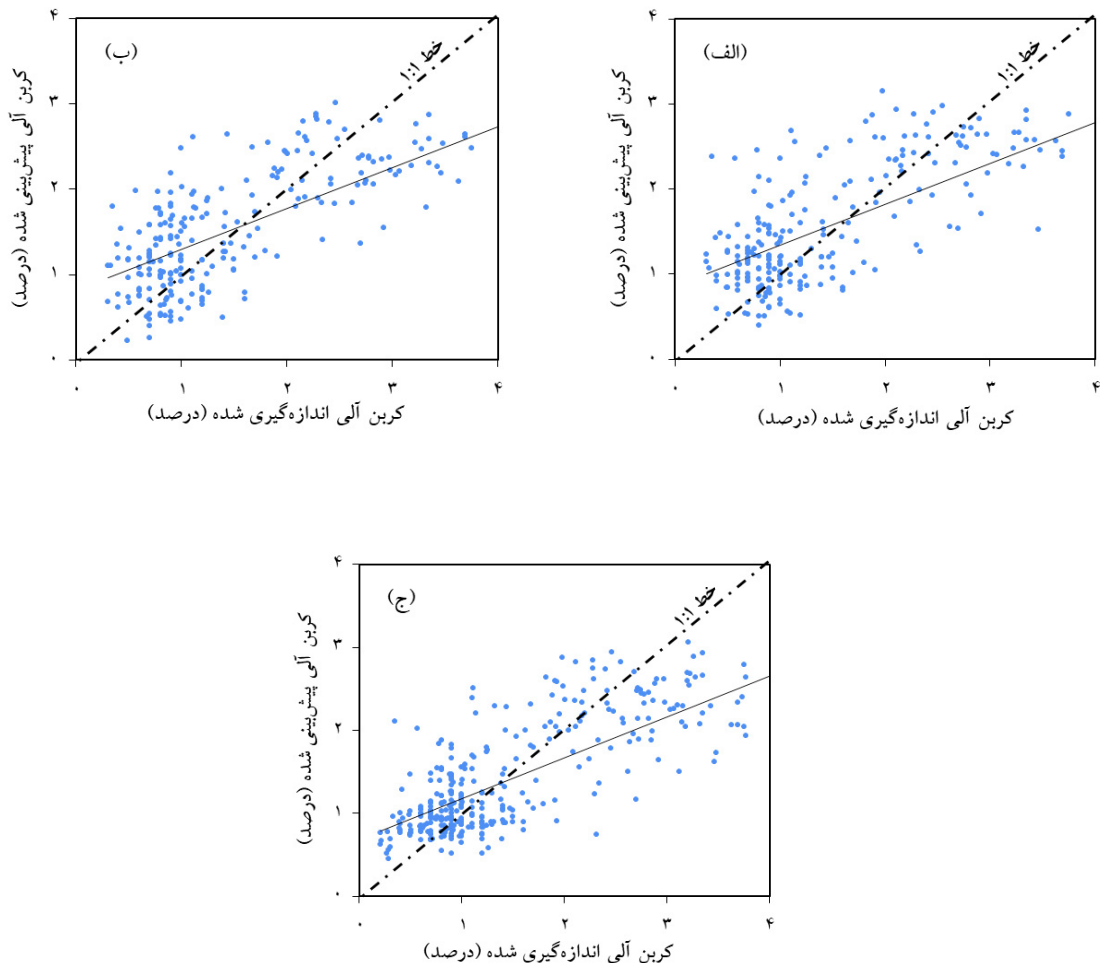
در مطالعه حاضر، کربن آلی خاک با استفاده از اطلاعات سنجنده مادیس و برخی مدل یادگیری ماشین برآورد شد. نتایج عملکرد مدل‌های یادگیری ماشین در جدول ۲ ارائه شده است. با توجه به نتایج جدول ۲، مقدار R^2 به کارگیری شبکه عصبی برای داده‌های واسنجی و صحت‌سنجی برابر با ۰/۴۹ و ۰/۵۱، میزان RMSE برای واسنجی و صحت‌سنجی به ترتیب برابر با ۰/۷۳ و ۰/۶۸، RPD برای واسنجی و صحت‌سنجی به ترتیب ۱/۴۰ و ۱/۵۰ و نسبت $R^2/$ RMSE برای هر دو حالت برابر با ۰/۶۷ و ۰/۷۵ می‌باشد. نسبت $R^2/$ RMSE یک معیار است که از ترکیب دو معیار ضریب تبیین و ریشه میانگین مربعات خطا به دست می‌آید. نسبت R^2 به RMSE می‌تواند به عنوان یک معیار جامع برای ارزیابی کیفیت مدل استفاده

درصد در خاک‌های کشاورزی با شرایط آب‌وهوای نیمه‌خشک یافت شده‌اند. کربن آلی خاک برای کل مجموعه داده، خاک‌های کشاورزی زراعی و خاک‌های کشاورزی باغی پس از حذف داده‌های نویزی به توزیع نرمال مطابق با آزمون همبستگی کولموگراف-اسمیرنوف واکنش مثبت نشان داده نشد و سپس از همبستگی اسپیرمن استفاده گردید.

تحلیل همبستگی بین میزان کربن آلی خاک و شاخص‌های طیفی سنجش‌ازدور: در فرآیند مدل‌سازی، انتخاب صحیح متغیرهای ورودی منجر به دقت نتیجه می‌شود [۶]. استفاده از متغیرهای ورودی با همبستگی بالا باعث کاهش دقت مدل به دلیل هم خطی بودن متغیرها حاصل می‌شود [۱۵]. ضرایب همبستگی بین داده‌های کربن آلی اندازه‌گیری شده و شاخص‌های طیفی استخراج‌شده از سنجنده مادیس، تصاویر لندست ۸ بررسی شد (شکل ۳). همبستگی مثبت بین میزان کربن آلی خاک و شاخص‌های طیفی سنجش‌ازدور، به‌استثنای NDWI استخراج‌شده از تصاویر تصویر لندست ۸، NDWI و GDVI استخراج‌شده از سنجنده مادیس وجود دارد.

جدول ۲: نتایج مدل‌سازی کربن آلی خاک با داده‌های طیفی مشتق‌شده از ماهواره مادیس

آزمون (۱۰۰)				آموزش (۲۳۶)				
RPD	R ² /RMSE	RMSE	R ²	RPD	R ² /RMSE	RMSE	R ²	
۱/۵۰	۰/۷۵	۰/۶۸	۰/۵۱	۱/۴۰	۰/۶۷	۰/۷۳	۰/۴۹	شبکه عصبی
۱/۴۸	۰/۷۲	۰/۶۹	۰/۵۰	۱/۴۲	۰/۷۲	۰/۷۲	۰/۵۲	الگوریتم بیان ژن
۱/۴۴	۰/۸۳	۰/۷۱	۰/۵۲	۱/۵۰	۰/۸۴	۰/۶۸	۰/۵۷	مدل بردار پشتیبان



شکل ۴: نمودار یک‌به‌یک داده‌های مشاهده‌ای و برآوردی کربن آلی خاک برای داده‌های صحت‌سنجی (شبکه عصبی (الف)، الگوریتم بیان ژن (ب)، و مدل بردار پشتیبان (ج)).

و مدل الگوریتم بیان ژن برآورد شد. با توجه به نتایج جدول ۲، مقدار آماره R^2 برای واسنجی و صحت‌سنجی برابر با ۰/۵۲ و ۰/۵۰، مقدار RMSE برای واسنجی و صحت‌سنجی به ترتیب برابر با ۰/۷۲ و ۰/۶۹، RPD برای واسنجی و صحت‌سنجی به ترتیب ۱/۴۲ و ۱/۴۸ و نسبت R^2 /RMSE برای هر دو حالت به ترتیب برابر با ۰/۷۲ برای هر دو حالت واسنجی و صحت‌سنجی می‌باشد. از دیگر سو، کربن آلی خاک با استفاده و مدل بردار پشتیبان نیز برآورد شد. با توجه به نتایج جدول ۲، مقدار R^2 برای واسنجی و صحت‌سنجی برابر با ۰/۵۷ و ۰/۵۲، میزان RMSE برای واسنجی و صحت‌سنجی به

شود. این معیار نشان می‌دهد که چه میزان توانایی مدل در توضیح دادن تغییرات وابسته را در مقایسه با خطا دارد. مقدار بالای این نسبت نشان‌دهنده عملکرد بهتر مدل است. در مورد مقادیری که در جدول ۲ ارائه شده است، نسبت R^2 /RMSE برای حالت واسنجی ۰/۶۷ و برای حالت صحت‌سنجی ۰/۷۵ است. به‌طورکلی، نسبت R^2 به RMSE یک معیار کاربردی است که ترکیبی از دقت و قابلیت توضیح دادن مدل را مورد بررسی قرار می‌دهد. این نسبت می‌تواند به مقایسه و انتخاب بهترین مدل برای برآورد و توضیح دادن داده‌ها کمک کند. افزون بر این، کربن آلی خاک با استفاده از تصویر مادیس

ترتیب برابر با ۰/۶۸ و ۰/۷۱، RPD برای واسنجی و صحت سنجی به ترتیب ۱/۵۰ و ۱/۴۴ و نسبت R²/RMSE برای هر دو حالت برابر با ۰/۸۴ و ۰/۸۳ می‌باشد. علاوه بر این، دقت برآورد کربن آلی خاک با استفاده از اطلاعات ماهواره مادیس در اطراف خط ۱:۱ که در واقع بیان‌گر رابطه بین مقادیر پیش‌بینی و اندازه‌گیری شده متغیرها است ارائه شده است (شکل ۴).

نتایج استفاده از اطلاعات تصاویر مادیس و مدل‌های یادگیری ماشین نشان داد که نسبت R²/RMSE برای مدل‌های مذکور به ترتیب برابر با ۰/۶۷، ۰/۷۲ و ۰/۸۴ می‌باشد (جدول ۲). مدل الگوریتم بیان ژن مدل مناسب‌تری برای برآورد کربن آلی خاک با استفاده از سنجنده مادیس می‌باشد. در این راستا، نتایج حاصل از مطالعه حاضر با مطالعات ژانگ و همکاران [۲۱] مطابقت دارد. ژانگ و همکاران [۲۱] کربن آلی خاک را با استفاده از تصاویر سنتینل ۲ و مادیس و مدل‌های یادگیری ماشین را در شمال چین بررسی کردند. نتایج این تحقیق نشان داد که مدل بردار پشتیبان عملکرد بهتری در مقایسه با مدل شبکه عصبی مصنوعی در برآورد کربن آلی خاک داشت.

با استفاده از اطلاعات لندست ۸

نتایج برآورد کربن آلی خاک با استفاده از اطلاعات تصاویر لندست ۸ با برخی مدل‌های یادگیری ماشین در جدول ۳ ارائه شده است. با توجه به نتایج جدول ۳، به‌کارگیری مدل شبکه عصبی مصنوعی نشان داد که مقدار R² برای واسنجی و صحت‌سنجی برابر با ۰/۵۵ و ۰/۵۲، میزان RMSE برای واسنجی و صحت‌سنجی به ترتیب برابر با ۰/۷۰ و ۰/۶۷، RPD برای واسنجی و صحت‌سنجی به ترتیب ۱/۴۶ و ۱/۵۲، و نسبت R²/RMSE برای هر دو حالت برابر با ۰/۷۹ و ۰/۷۸ می‌باشد. هم‌چنین، نتایج به توسعه مدل الگوریتم بیان ژن در جدول ۳ ارائه شده است. با توجه به نتایج جدول ۳، مقدار R² برای واسنجی و صحت‌سنجی برابر با ۰/۴۸ و ۰/۴۹، میزان RMSE برای واسنجی و صحت‌سنجی به ترتیب ۱/۳۸ و ۱/۴۸، و نسبت R²/RMSE برای هر دو حالت برابر با ۰/۶۵ و ۰/۷۱ می‌باشد. هم‌چنین، کربن آلی خاک با استفاده از تصویر لندست ۸ و مدل بردار پشتیبان برآورد شد. با توجه به نتایج جدول ۳، مقدار R² برای واسنجی و

صحت‌سنجی برابر با ۰/۵۵ و ۰/۴۹، میزان RMSE برای واسنجی و صحت‌سنجی به ترتیب برابر با ۰/۶۹ و ۰/۷۳، RPD برای واسنجی و صحت‌سنجی به ترتیب ۱/۴۸ و ۱/۴۰، و نسبت R²/RMSE در دو حالت برابر با ۰/۸۰ و ۰/۶۷ می‌باشد. علاوه بر این، دقت برآورد کربن آلی خاک برای مدل‌های توسعه داده شده با استفاده از اطلاعات ماهواره لندست ۸ در اطراف خط ۱:۱ ارائه شده است (شکل ۵).

تصویر لندست ۸ ثابت و مدل‌های شبکه عصبی پرسپترون چندلایه، الگوریتم بیان ژن و ماشین بردار پشتیبان با یکدیگر مقایسه شدند. نسبت R²/RMSE برای مدل‌های مذکور به ترتیب برابر با ۰/۷۲، ۰/۶۵ و ۰/۸۰ می‌باشد. مدل الگوریتم بیان ژن مدل مناسب‌تری برای برآورد کربن آلی خاک با استفاده از تصویر لندست ۸ می‌باشد. به‌هرحال نتایج مطالعات انجام شده در برآورد کربن آلی خاک به دلیل شرایط محیطی حاکم متفاوت بوده است. برای مثال، محمودآبادی و همکاران [۸] سه مدل عددی رگرسیون خطی چندگانه، شبکه عصبی مصنوعی و برنامه‌ریزی بیان ژن به‌منظور استفاده در معادله اسکورپن برای تخمین ویژگی‌های فیزیکی و شیمیایی مورد مقایسه قرار دادند. نتایج نشان داد که بر اساس مقدار آماره‌های ارزیابی در فاز آزمون، روش شبکه عصبی پرسپترون چندلایه دارای دقت بیش‌تری در برآورد مقادیر رس، شن، سیلت، کربن آلی، نیتروژن، کربنات کلسیم، ظرفیت آب اشباع و چگالی خاک در مقایسه با دیگر مدل‌های بررسی شده برخوردار بود. روش شبکه عصبی پرسپترون چندلایه برای تخمین ویژگی‌های خاک رضایت‌بخش بود به‌گونه‌ای که این مدل توانست ۶۰ تا ۸۵ درصد تغییرپذیری خصوصیات موردبررسی را نشان دهد. هم‌چنین، میرزائی و همکاران [۱۰] با استفاده از اطلاعات ماهواره لندست نشان دادند که شبکه عصبی مصنوعی عملکرد قابل قبولی در تخمین ماده آلی خاک دارد. هم‌چنین، نتایج تحقیقات گومزو همکاران [۳] در کشور برزیل نیز جهت بررسی میزان ذخیره کربن آلی خاک با مدل‌های جنگل تصادفی، کیوبیست، مدل خطی تعمیم‌یافته و ماشین بردار پشتیبان انجام گرفته بود، نشان داد که روش جنگل تصادفی در تمام عمق‌های استاندارد (۰-۵، ۵-۱۵، ۱۵-۳۰، ۳۰-۶۰، ۶۰-۱۰۰) بهترین عملکرد را در برآورد ذخیره کربن آلی خاک دارد. افزون بر این، مینانسی و همکاران [۱۱] با استفاده از سه مدل شبکه عصبی مصنوعی، جنگل تصادفی و رگرسیون درختی نقشه رقومی کربن آلی خاک را در دو منطقه در اندونزی تهیه کردند. نتایج این تحقیق

جدول ۳: نتایج مدل‌سازی کربن آلی خاک با داده‌های طیفی مشتق شده از ماهواره لندست ۸

آزمون (۱۰۰)				آموزش (۲۳۶)				
RPD	R ² /RMSE	RMSE	R ²	RPD	R ² /RMSE	RMSE	R ²	
۱/۵۲	۰/۷۸	۰/۶۷	۰/۵۲	۱/۴۶	۰/۷۹	۰/۷۰	۰/۵۵	شبکه عصبی
۱/۴۸	۰/۷۱	۰/۶۹	۰/۴۹	۱/۳۸	۰/۶۵	۰/۷۴	۰/۴۸	الگوریتم بیان ژن
۱/۴۰	۰/۶۷	۰/۷۳	۰/۴۹	۱/۴۸	۰/۸۰	۰/۶۹	۰/۵۵	مدل بردار پشتیبان

نشان داد که عملکرد مدل شبکه عصبی مصنوعی با استفاده از تصویر لندست ۸ بهتر بود.

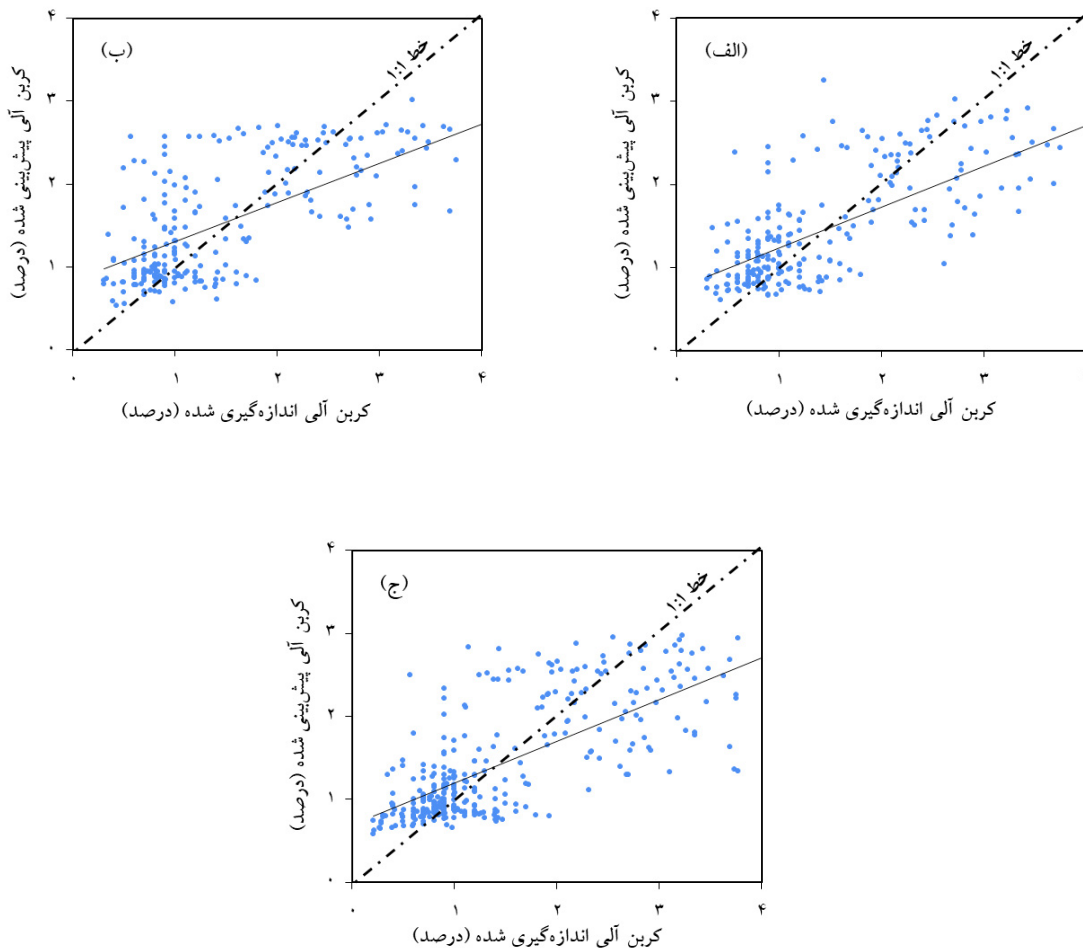
هم‌افزایی تصویر لندست ۸ و سنجنده مادیس

نتایج برآورد کربن آلی خاک با استفاده از هم‌افزایی تصاویر لندست ۸ و مادیس در جدول ۴ ارائه شده است. با توجه به نتایج جدول ۴، مقدار R^2 برای واسنجی و صحت‌سنجی برابر با ۰/۵۸ و ۰/۵۳، میزان RMSE برای واسنجی و صحت‌سنجی به ترتیب برابر با ۰/۶۶ و ۰/۶۷، RPD برای واسنجی و صحت‌سنجی به ترتیب ۱/۵۴ و ۱/۵۵، و نسبت R^2 /RMSE در دو حالت برابر با ۰/۸۸ و ۰/۷۹ می‌باشد. افزون بر این، نتایج برآورد کربن آلی خاک با مدل الگوریتم بیان ژن نشان داد که مقدار R^2 برای واسنجی و صحت‌سنجی برابر با ۰/۵۲ و ۰/۴۵، میزان RMSE برای واسنجی و صحت‌سنجی به ترتیب برابر با ۰/۷۱ و ۰/۷۳، RPD برای واسنجی و صحت‌سنجی به ترتیب ۱/۴۴ و ۱/۳۷، و نسبت R^2 /RMSE برای هر دو حالت برابر با ۰/۷۳

و ۰/۶۲ می‌باشد (جدول ۴). با توجه به نتایج جدول ۴، دقت مدل بردار پشتیبان با استفاده از اطلاعات هم‌افزایی تصاویر ماهواره‌ای برابر با R^2 برای واسنجی و صحت‌سنجی برابر با ۰/۶۲ و ۰/۵۴، میزان RMSE برای واسنجی و صحت‌سنجی به ترتیب برابر با ۰/۶۳ و ۰/۶۹، RPD برای واسنجی و صحت‌سنجی به ترتیب ۱/۶۲ و ۱/۴۵، و نسبت R^2 /RMSE برای هر دو حالت برابر ۰/۹۸ و ۰/۷۸ می‌باشد. علاوه بر این، دقت برآورد کربن آلی خاک برای مدل‌های توسعه داده شده با استفاده از هم‌افزایی تصاویر ماهواره‌ای در اطراف خط ۱:۱ ارائه شده است (شکل ۶).

مقایسه نتایج سه روش فوق

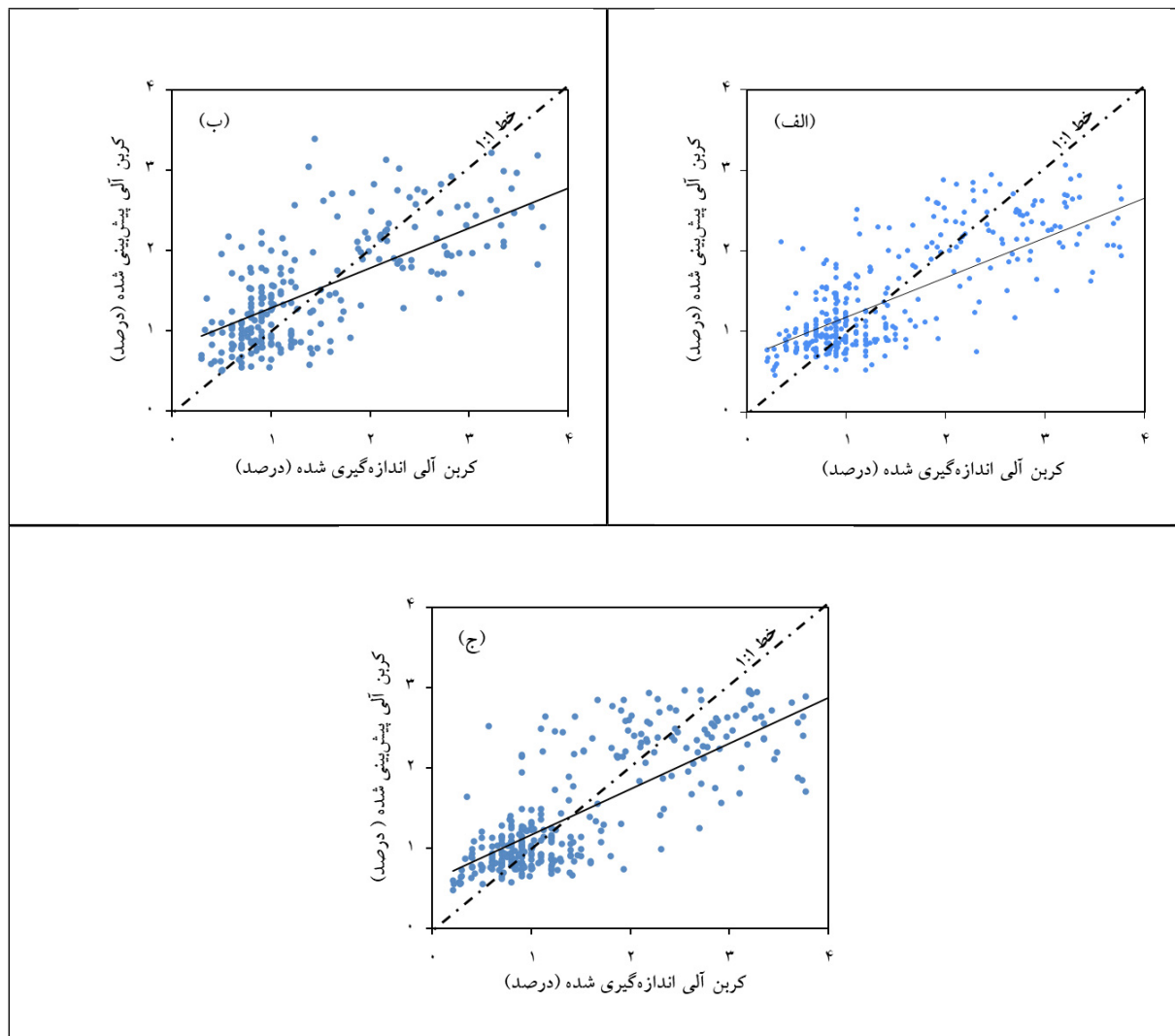
بر اساس نتایج به‌دست‌آمده در بخش‌های پیشین، میزان نسبت R^2 /RMSE برای سه مدل پیشرفته هوش مصنوعی (شبکه عصبی پرسپترون چندلایه، الگوریتم بیان ژن و رگرسیون بردار پشتیبان) در استفاده از داده‌های سنجنده مادیس به ترتیب ۰/۶۷، ۰/۷۲ و ۰/۸۴



شکل ۵: نمودار یک‌به‌یک داده‌های مشاهده‌ای و برآوردی کربن آلی خاک برای داده‌های صحت‌سنجی (شبکه عصبی (الف)، الگوریتم بیان ژن (ب)، و مدل بردار پشتیبان (ج)).

جدول ۴: نتایج مدل‌سازی کربن آلی خاک با داده‌های طیفی مشتق شده از هم‌افزایی ماهواره مادیس و لندست ۸

آزمون (۱۰۰)				آموزش (۲۳۶)				
RPD	R ² /RMSE	RMSE	R ²	RPD	R ² /RMSE	RMSE	R ²	
۱/۵۵	۰/۸۸	۰/۶۶	۰/۵۸	۱/۵۴	۰/۸۸	۰/۶۶	۰/۵۸	شبکه عصبی
۱/۳۷	۰/۶۲	۰/۷۳	۰/۴۵	۱/۴۴	۰/۷۳	۰/۷۱	۰/۵۲	الگوریتم بیان ژن
۱/۴۵	۰/۷۸	۰/۶۹	۰/۵۴	۱/۶۲	۰/۹۸	۰/۶۳	۰/۶۲	مدل بردار پشتیبان



شکل ۶: نمودار یک‌به‌یک داده‌های مشاهده‌ای و برآوردی کربن آلی خاک برای داده‌های صحت‌سنجی (شبکه عصبی (الف)، الگوریتم بیان ژن (ب)، و مدل بردار پشتیبان (ج)).

سنجده مادیس و تصاویر لندست ۸، مؤثرترین روش برای برآورد کربن آلی در خاک است. به‌طورکلی، بهترین مدل توسعه‌یافته در تحقیق حاضر یعنی مدل ماشین بردار پشتیبان با استفاده از اطلاعات هم‌افزایی تصاویر ماهواره‌های لندست ۸ و مادیس توانستند ۶۲ درصد از تغییرات کربن آلی خاک را نشان دهند. این دلایل متفاوتی می‌تواند داشته باشد. از جمله دلایل مهم می‌توان به کاربری موردمطالعه در این تحقیق اشاره کرد. کاربری موردمطالعه در تحقیق حاضر اراضی کشاورزی با نسبت ۸۵/۸ درصد اراضی زراعی و ۱۴/۲ درصد اراضی باغی اشاره کرد. در خاک‌های کشاورزی به‌خصوص

به دست آمد (جدول ۵). هنگام استفاده از تصاویر لندست ۸ این نسبت‌ها برای همان سه مدل به ترتیب به ۰/۷۹، ۰/۶۵ و ۰/۸۰ تغییر یافتند. اما با بهره‌گیری از ترکیب دو داده‌ی مذکور (هم‌افزایی داده‌ها)، نسبت R²/RMSE برای سه مدل به ترتیب به ۰/۷۳، ۰/۸۸ و ۰/۹۸ افزایش یافت (جدول ۵). این افزایش قابل‌توجه نشان می‌دهد که استفاده هم‌زمان از سنجده مادیس و تصاویر لندست ۸ در مقایسه با استفاده مجزا از هرکدام، به‌طور معناداری دقت و کارایی تخمین کربن آلی خاک را بهبود می‌بخشد. بنابراین، می‌توان نتیجه گرفت که استفاده از مدل رگرسیون بردار پشتیبان، همراه با هم‌افزایی داده‌های

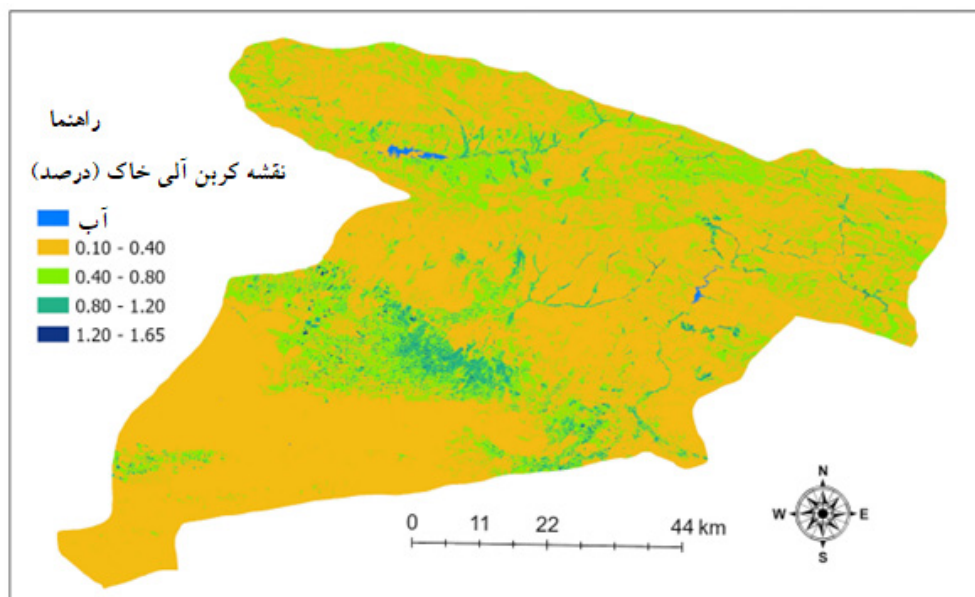
دارد. در منطقه مرکزی، میزان کربن آلی بین ۰/۵ تا ۱/۲ درصد متغیر است، که نشان‌دهنده تأثیر تراکم بالای باغات بر ساختار و عملکرد خاک در ذخیره کربن است. بخش‌های شمالی منطقه نیز شاهد تغییراتی در میزان کربن آلی خاک (۰/۴ تا ۰/۸ درصد) هستند که این موضوع به وجود جنگل‌های کم‌پوش و مراتع ارتباط داده می‌شود، افزایش نسبی کربن آلی در این مناطق تأکیدی بر نقش حیاتی اکوسیستم‌های طبیعی در حفظ تعادل کربنی دارد. در مقابل، بخش‌های جنوبی و به‌ویژه جنوب غربی، که با کاهش ماده آلی خاک مواجه هستند، به دلیل پوشش گیاهی کم‌تر و فرسایش زمین‌های کشاورزی دیم مطرح می‌شوند، این وضعیت بر چالش‌های مرتبط با حفاظت از خاک و جلوگیری از تخریب آن در مناطق کشاورزی تأکید دارد، که این امر اهمیت استراتژی‌های مدیریتی مؤثر در حفظ سلامت خاک و تقویت ذخیره کربن را نمایان می‌سازد.

اراضی زراعی هر ساله عملیات متنوع زراعی انجام می‌شود. یکی از مهم‌ترین عملیات زراعی، خاک‌ورزی است که موجب بهم‌خوردن خاک سطحی می‌گردد. به‌هرحال، به هم خوردن تأثیر زیادی در دقت مدل‌های یادگیری ماشین دارد [۱۰ و ۲۱]. هم‌چنین، این امر می‌تواند ناشی از فعالیت‌های کشاورزی، سوزاندن بقایای گیاهی و تغییر کاربری توسط انسان‌ها و یا ناهنجاری‌های محیطی مانند فرسایش خاک باشد.

توزیع کربن آلی خاک در مناطق مختلف منطقه مورد مطالعه نشان‌دهنده تفاوت‌های قابل‌توجهی است که به‌وضوح در شکل ۷ قابل‌مشاهده است. تحلیل‌ها بر اساس نقشه کربن آلی خاک، بیانگر آن است که میزان کربن آلی در نواحی نزدیک به رودخانه‌ها که عمدتاً به کاربری‌های باغی و زراعت آبی اختصاص دارند، بالاتر است. این افزایش به خصوصیات بیژئوشیمیایی خاک که تحت تأثیر مستقیم کاربری زمین و نوع پوشش گیاهی قرار دارد، اشاره

جدول ۵: مقایسه دقت برآورد کربن آلی خاک با استفاده از مدل‌های مختلف

کربن آلی خاک				داده‌های طیفی مورد استفاده	فن مدل‌سازی
رتبه	R ² /RMSE	RMSE	R ²		
۸	۰/۶۷	۰/۷۳	۰/۴۹	مادیس	شبکه عصبی پرسپترون چندلایه
۵	۰/۷۹	۰/۷۰	۰/۵۵	لندست ۸	
۲	۰/۸۸	۰/۶۶	۰/۵۸	هم‌افزایی لندست ۸ و مادیس	
۷	۰/۷۲	۰/۷۲	۰/۵۲	مادیس	رژن
۹	۰/۶۵	۰/۷۴	۰/۴۸	لندست ۸	
۶	۰/۷۳	۰/۷۱	۰/۵۲	هم‌افزایی لندست ۸ و مادیس	ماشین بردار الگوریتم بیان پشتیبان
۳	۰/۸۴	۰/۶۸	۰/۵۷	مادیس	
۴	۰/۸۰	۰/۶۹	۰/۵۵	لندست ۸	
۱	۰/۹۸	۰/۶۳	۰/۶۲	هم‌افزایی لندست ۸ و مادیس	



شکل ۷: نقشه کربن آلی خاک منطقه مورد مطالعه

in Pearl River Basin using input variable selection and machine learning techniques. *Science of the Total Environment*. 731: 139099.

7- Li, Y., Shi, Q., He, L., Cai, R., Meng, L., Li, J., and Plaza, A. 2022. Fusing Sentinel-2 and Landsat-8 surface reflectance data via pixel-wise local normalization. *IEEE Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. 15: 7359-7374.

8- Mahmoudabadi, A., Karimi, A. R., Haqnia, G. and Sepehr, A. 2017. Performance evaluation of multivariate regression, artificial neural network and Biangen programming in estimating some soil properties. *Journal of water and soil protection research*. 24 (2): 23-44 (In Persian).

9- Mirchooli, F., Kiani-Harchegani, M., Darvishan, A. K., Falahatkar, S., and Sadeghi, S. H. 2020. Spatial distribution dependency of soil organic carbon content to important environmental variables. *Ecological Indicators*. 116: 106473.

10- Mirzaee, S., Ghorbani-Dashtaki, S., Mohammadi, J., Asadi, H., and Asadzadeh, F. 2016. Spatial variability of soil organic matter using remote sensing data. *Catena*. 145: 118-127.

11- Minasny, B., Setiawan, B. I., Arif, C., Saptomo, S. K., and Chadirin, Y. 2016. Digital mapping for cost-effective and accurate prediction of the depth and carbon stocks in Indonesian peatlands. *Geoderma*. 272: 20-31.

12- Nawar, S., and Mouazen, A. M. 2019. On-line vis-NIR spectroscopy prediction of soil organic carbon using machine learning. *Soil and Tillage Research*. 190: 120-127.

13- Nelson, D. W., and Sommers, L. E. 1983. Total carbon, organic carbon, and organic matter. *Methods of soil analysis: Part 2 chemical and microbiological properties*. 9: 539-579.

14- Paustian, K., Larson, E., Kent, J., Marx, E., and Swan, A. 2019. Soil C sequestration as a biological negative emission strategy. *Frontiers in Climate*. 1: 482133.

15- Peng, J., Biswas, A., Jiang, Q., Zhao, R., Hu, J., Hu, B., and Shi, Z. 2019. Estimating soil salinity from remote sensing and terrain data in southern Xinjiang Province, China. *Geoderma*. 337: 1309-1319.

16- Seifi, M., Ahmadi, A., Neyshabouri, M. R., Taghizadeh-Mehrjardi, R., and Bahrami, H. A. 2020. Remote and Vis-NIR spectra sensing potential for soil salinization estimation in the eastern coast of Urmia hyper saline lake, Iran. *Remote Sensing Applications: Society and Environment*. 20: 100398.

17- Soussana, J. F., Loiseau, P., Vuichard, N., Ceschia, E., Balesdent, J., Chevallier, T., and Arrouays, D. 2004. Carbon cycling and sequestration opportunities in temperate grasslands. *Soil use*

نتیجه‌گیری کلی

در این مطالعه، از تصاویر ماهواره‌ای لندست ۸ و داده‌های سنجنده مادیس برای برآورد میزان کربن آلی موجود در خاک استفاده شده است. به منظور افزایش دقت برآوردها، فن هم‌افزایی داده‌های حاصل از تصاویر لندست ۸ و سنجنده مادیس به کار گرفته شد. هم‌چنین، کارایی سه مدل پیشرفته یادگیری ماشینی (شبکه عصبی پرسپترون چندلایه، الگوریتم بیان ژن و ماشین بردار پشتیبان) در برآورد کربن آلی خاک مورد ارزیابی قرار گرفت. عملکرد این مدل‌ها با استفاده از داده‌های طیفی و بر اساس معیارهای آماری R^2 و RMSE مورد ارزیابی و مقایسه قرار گرفت. نتایج نشان داد که استفاده هم‌زمان از تصاویر لندست ۸ و داده‌های سنجنده مادیس (هم‌افزایی داده‌ها) نسبت به استفاده جداگانه از هر یک از آن‌ها، منجر به دقت بیش‌تر با ضریب تبیین بالاتر و خطای پایین‌تر می‌شود. این مطالعه هم‌چنین نشان داد که مدل ماشین بردار پشتیبان در رتبه اول و سپس مدل شبکه عصبی مصنوعی پرسپترون چندلایه، در ترکیب داده‌های مختلف ماهواره‌ای برای دستیابی به نتایج دقیق‌تر بسیار مؤثر هستند. بنابراین، استفاده از هم‌افزایی تصاویر لندست ۸ و سنجنده مادیس می‌تواند به عنوان روشی قابل اعتماد و مؤثر در برآورد کربن آلی خاک در مناطق مورد مطالعه استفاده شود.

منابع

1- Akbari, M., Goudarzi, I., Tahmoures, M., Elveny, M., and Bakhshayeshi, I. 2021. Predicting soil organic carbon by integrating Landsat 8 OLI, GIS and data mining techniques in semi-arid region. *Earth Science Informatics*. 14(4): 2113-2122 (In Persian).

2- Chen, D., Chang, N., Xiao, J., Zhou, Q., and Wu, W. 2019. Mapping dynamics of soil organic matter in croplands with MODIS data and machine learning algorithms. *Science of the Total Environment*. 669: 844-855.

3- Gomes, L. C., Faria, R. M., de Souza, E., Veloso, G. V., Schaefer, C. E. G., and Fernandes Filho, E. I. 2019. Modelling and mapping soil organic carbon stocks in Brazil. *Geoderma*. 340: 337-350.

4- Gray, J., Karunaratne, S., Bishop, T., Wilson, B., and Veeragathipillai, M. 2019. Driving factors of soil organic carbon fractions over New South Wales, Australia. *Geoderma*. 353: 213-226.

5- Liao, C., Wang, J., Dong, T., Shang, J., Liu, J., and Song, Y. 2019. Using spatio-temporal fusion of Landsat-8 and MODIS data to derive phenology, biomass and yield estimates for corn and soybean. *Science of the total environment*. 650: 1707-1721.

6- Li, W., Fang, H., Qin, G., Tan, X., Huang, Z., Zeng, F., Du, H. and Li, S. 2020. Concentration estimation of dissolved oxygen

- 20- Zepp, S., Heiden, U., Bachmann, M., Wiesmeier, M., Steininger, M., and Van Wesemael, B. 2021. Estimation of soil organic carbon contents in croplands of Bavaria from SCMaP soil reflectance composites. *Remote Sensing*. 13(16): 3141.
- 21- Zhang, M., Zhang, M., Yang, H., Jin, Y., Zhang, X., and Liu, H. 2021. Mapping regional soil organic matter based on sentinel-2a and modis imagery using machine learning algorithms and google earth engine. *Remote Sensing*. 13(15): 2934.
- 18- Wang, B., Gray, J.M., Waters, C.M., Anwar, M.R., Orgill, S.E., Cowie, A.L., Feng, P. and Li Liu, D. 2022. Modelling and mapping soil organic carbon stocks under future climate change in south-eastern Australia. *Geoderma*. 405: 115442.
- 19- Wenzel, W. W., Duboc, O., Golestanifard, A., Holzinger, C., Mayr, K., Reiter, J., and Schiefer, A. 2022. Soil and land use factors control organic carbon status and accumulation in agricultural soils of Lower Austria. *Geoderma*. 409: 115595.



Abstract

Estimation of Soil Organic Carbon Using Data Synergy of Landsat 8 Satellite Imagery and MODIS Sensor Through Machine Learning Models

P. Abdoli¹, A. Khanmirzaei^{*2}, S. Hamzeh³, Sh. Rezaei⁴ and S. Moghimi⁵

Received: 2024/04/16 Accepted: 2024/05/22

The main objective of this study was to develop machine learning models for predicting soil organic carbon (SOC) content in agricultural soils enriched with calcium-rich materials in the central region of Iran. Soil samples were collected from 336 points in agricultural lands with two land uses, namely cropland and orchard, from Alborz Province. The soil organic carbon was measured using the Walkley-Black oxidation method. Remote sensing data, obtained from various sources such as Landsat 8 imagery and MODIS sensor, were utilized for soil organic carbon analysis. Three machine learning models, including Artificial Neural Networks (ANN), Support Vector Regression (SVR), and Gene Expression Programming (GEP), were employed in this study to estimate the soil organic carbon content. The results of this research demonstrated positive outcomes, indicating that the simultaneous use of Landsat 8 imagery and MODIS sensor data through the SVR model yielded better performance ($R^2 = 0.62$, $RMSE = 0.63$, $R^2/RMSE = 0.98$) compared to the use of each image separately. Therefore, this study concluded that the simultaneous synergy of remote sensing data from different sources (referred to as data synergy) can significantly improve the accuracy of predicting soil organic carbon content in calcium-enriched agricultural soils in the central region of Iran. It is recommended to use advanced fusion techniques and deep learning methods for combining features at a higher level.

Keywords: Environmental data, Modelling, and combination, Remote sensing, Vegetation indices.

1. Department of Soil Science, Karaj Branch, Islamic Azad University, Karaj, Iran
Corresponding author, Email: alikhanmirzai@yahoo.com

2. Remote Sensing Department, University of Tehran, Tehran, Iran