

برآورد ترسیب کربن و تنفس خاک جنگل با استفاده از مدل‌های مبتنی بر یادگیری ماشین در جنگل‌های شرق استان مازندران

سیدمحمد حاجتی^{۱*}، محیا تفضلی^۲، مریم اسدیان^۳ و علی بالویی^۴

۱- استاد، گروه علوم و مهندسی جنگل، دانشکده منابع طبیعی، دانشگاه علوم کشاورزی و منابع طبیعی ساری، ساری، ایران. (s_m_hodjati@yahoo.com)

۲- دانش‌آموخته دکتری جنگلداری، دانشکده منابع طبیعی، دانشگاه علوم کشاورزی و منابع طبیعی ساری، ساری، ایران. (mahya_tafazoli@yahoo.com)

۳- دانشجوی دکتری، دانشکده منابع طبیعی، دانشگاه علوم کشاورزی و منابع طبیعی ساری، ساری، ایران. (maryam.asadiyan23@gmail.com)

۴- دانشجوی کارشناسی ارشد جنگلداری، دانشکده منابع طبیعی، دانشگاه علوم کشاورزی و منابع طبیعی ساری، ساری، ایران. (ali.baluce@gmail.com)

تاریخ دریافت: ۱۴۰۱/۰۵/۱۰ تاریخ پذیرش: ۱۴۰۱/۰۶/۰۵

چکیده

در این پژوهش مدل‌های مبتنی بر یادگیری ماشین (رگرسیون-خطی، k-نزدیک‌ترین همسایه، ماشین بردار پشتیبان، جنگل تصادفی) برای برآورد ترسیب کربن و تنفس خاک در جنگل‌های شرق استان مازندران ارزیابی شدند. پس از مشخص شدن نقاط نمونه‌برداری، در هر یک از قطعات نمونه، قطر و ارتفاع درختان اندازه‌گیری و زی‌توده روی-زمینی درختان با استفاده از مدل‌های آلومتریک جنگل هیرکانی محاسبه شد. نمونه خاک از عمق صفر تا ۲۰ سانتی‌متر تهیه و تنفس خاک با دستگاه CO₂-port اندازه‌گیری شد. تنفس خاک با استفاده از متغیرهای وزن مخصوص ظاهری، درصد رطوبت، درصد اجزای بافت، دمای خاک، نیتروژن کل، فسفر و پتاسیم قابل جذب، درصد کربن و زی‌توده درختان برآورد شد. ترسیب کربن خاک با کمک متغیرهای دما و رطوبت خاک و زی‌توده درختان برآورد شد. مدل جنگل تصادفی (RMSE=۱۰/۴۷ و R²=۰/۸۲) و ماشین بردار پشتیبان (RMSE=۰/۷۷ و R²=۰/۹۰) به ترتیب بالاترین عملکرد در برآورد ترسیب کربن و تنفس خاک داشت. متغیر رطوبت خاک در برآورد ترسیب کربن (مدل جنگل تصادفی) و تنفس خاک (مدل ماشین بردار پشتیبان) دارای بالاترین اهمیت نسبی بود. با توجه به نتایج به دست آمده می‌توان با داشتن زی‌توده روی-زمین درختان و ویژگی‌های اولیه خاک، مقدار ترسیب کربن و تنفس خاک را در جنگل با دقت مناسب برآورد کرد.

واژه‌های کلیدی: تغییر اقلیم، جنگل تصادفی، ماشین بردار پشتیبان، ویژگی‌های خاک جنگل

مقدمه

مقدار تنفس خاک اثر دارد می‌توان به نوع پوشش گیاهی، رطوبت خاک، دمای خاک و مقدار عناصر غذایی در خاک اشاره کرد (Zhu et al., 2020). همچنین لازم به ذکر است که تنفس خاک یکی از مهم‌ترین ویژگی‌های خاک است که وضعیت فعالیت زیستی را منعکس می‌کند. بنابراین با توجه به اهمیت تنفس خاک در چرخه جهانی کربن و گرمایش جهانی، می‌توان بیان کرد که برآورد و تخمین مقدار انتشار CO_2 از خاک جنگل برای درک بازخورد بین تغییرات اقلیم و اکوسیستم زمینی و جنگلی ضروری است (Tafazoli et al., 2021; Adjuik and Davis, 2022). با این حال با توجه به اینکه اندازه‌گیری مستقیم انتشار CO_2 از خاک نیاز به تجهیزات ویژه و کارشناس باتجربه دارد، می‌تواند پرهزینه و زمان‌بر باشد. همچنین با توجه به شرایط بسیار متغیر خاک، تعیین مقدار انتشار CO_2 می‌تواند با چالش همراه باشد (Adjuik and Davis, 2022). یک راهکار مناسب برای اندازه‌گیری مستقیم مقدار انتشار گازهای گلخانه‌ای خاک، استفاده از مدل‌های مبتنی بر یادگیری ماشین است (Liakos et al., 2018). الگوریتم‌های یادگیری ماشین از تشخیص الگو برای توصیف روابط بین متغیرهای ورودی و متغیرهای خروجی با ویژگی‌های یادگیری روابط یک مجموعه داده معین استفاده می‌کنند. روش یادگیری ماشین در مدل‌سازی داده‌های با ساختار پیچیده که اغلب غیرخطی هستند به سرعت در حال رشد است. مدل‌های یادگیری ماشین قدرت بیشتری برای حل و ایجاد روابط پیچیده دارند (روابط غیرخطی، غیریکنواخت، چندوجهی مشترک با کاربردهای چشم‌انداز و محیط‌زیست) زیرا محدود به مفروضات سنتی در مورد ویژگی‌های داده نیستند.

با توجه به اهمیت موضوع بیان شده، پژوهش‌هایی در ارتباط با برآورد ویژگی‌های خاک و همچنین تنفس

تغییر اقلیم و افزایش گرمایش جهانی یکی از چالش‌های مهم در بحث توسعه پایدار است. از مهم‌ترین عوامل ایجاد تغییر اقلیم و پدیده گرمایش جهانی می‌توان به انتشار غیرمجاز گازهای گلخانه‌ای اشاره کرد (Osabohien et al., 2019) که مهم‌ترین این گازها شامل دی‌اکسید نیتروژن (N_2O)، دی‌اکسید کربن (CO_2) و متان (CH_4) هستند (Osabohien et al., 2019). با توجه به اینکه تجمع دی‌اکسید کربن حجم بیشتری را در اتمسفر به خود اختصاص داده است، از کلیدی‌ترین و مهم‌ترین گازهای گلخانه‌ای به‌شمار می‌رود که استفاده از سوخت‌های فسیلی و جنگل‌زدایی سبب افزایش مقدار آن در جو شده است (Hojjati et al., 2022). در این میان، جنگل‌ها که مهم‌ترین اکوسیستم‌های خشکی نیز به حساب می‌آیند، نقش مهمی در چرخه گاز دی‌اکسید کربن ایفا می‌کنند و یکی از مهم‌ترین محل ذخیره و ترسیب کربن هستند (Pan et al., 2011). خاک جنگل نیز منبع مهمی برای ذخیره کربن در بوم‌سازگان جنگل است.

در بوم‌سازگان جنگل، فرآیند تنفس خاک که حدود ۳۰ تا ۸۰ درصد از تنفس را در بوم‌سازگان جنگلی به خود اختصاص می‌دهد، نقش بسیار مهمی در چرخه گاز دی‌اکسید کربن دارد که به‌طور مستقیم بر غلظت دی‌اکسید کربن در جو و در نتیجه در پدیده گرمایش جهانی اثر خواهد داشت (Naderi et al., 2022). بنابراین، می‌توان بیان کرد که هر عاملی که منجر به تغییر در مقدار تنفس خاک جنگل شود به‌طور مستقیم روی پدیده تغییر اقلیم و گرمایش جهانی نیز اثرگذار خواهد بود. انتشار CO_2 از خاک بیشتر ناشی از تنفس ریشه‌ای و فرآیندهای تجزیه و تنفس هتروتروف (حاصل از تجزیه مواد آلی خاک) میکروارگانیسم‌های خاک است (Adjuik and Davis, 2022). مهم‌ترین عواملی که روی

کاهش تراکم دی‌اکسید کربن اتمسفری در مجامع علمی و سیاسی جهان مطرح شده است (Alazmani et al., 2021). ترسیب کربن خاک بخش مهمی از ترسیب کربن در اکوسیستم خشکی است و تأثیر شدیدی بر دی-اکسید کربن اتمسفری دارد (Hojjati et al., 2020). به-علت نقش خاک در چرخه جهانی کربن و نقش پر اهمیت پایش ماده آلی خاک برای برآورد کیفیت و مدیریت خاک، تقاضا برای دریافت اطلاعات کربن آلی خاک در حال افزایش است. با این حال، اندازه‌گیری ترسیب کربن خاک نیز در سطوح وسیع بوم‌سازگان جنگل می‌تواند پرهزینه و زمان‌بر باشد. بنابراین پژوهش در مورد کاربرد و بررسی کارایی مدل‌های مبتنی بر یادگیری ماشین برای برآورد کربن خاک و همچنین ترسیب کربن خاک نیز لازم و ضروری به‌نظر می‌رسد. در این راستا (Matinfar et al., 2021) به ارزیابی روش‌های یادگیری ماشین در نقشه‌برداری رقومی کربن آلی خاک‌های زراعی پرداخت و بیان کرد که مدل جنگل تصادفی به‌ترتیب با مقادیر ضریب تبیین و مجذور میانگین مربعات خط به‌ترتیب ۰/۷۵ و ۰/۲۵ درصد بهترین کارایی را نسبت به دیگر رویکردهای مورد استفاده (کوئیست، رگرسیون حداقل مربعات جزئی، رگرسیون خطی چندمتغیره و کریجینگ معمولی) در منطقه دشت خرم داشت.

با توجه به اطلاعات موجود، تاکنون پژوهشی در ارتباط با امکان به‌کارگیری مدل‌های مبتنی بر یادگیری ماشین برای برآورد ترسیب کربن و تنفس خاک در بوم-سازگان جنگل، به‌ویژه در جنگل‌های شمال کشور انجام نشده است. بنابراین هدف از اجرای این پژوهش ارزیابی مدل‌های مبتنی بر یادگیری ماشین برای برآورد ترسیب کربن و تنفس خاک با استفاده از داده‌های زی-توده روی زمینی درختان و ویژگی‌های خاک بود.

خاک با استفاده از مدل‌های مبتنی بر یادگیری ماشین انجام شده است. (Adjuik and Davis (2022) در پژوهش خود به ارزیابی مدل‌های یادگیری ماشین برای پیش‌بینی انتشار گازهای گلخانه‌ای خاک پرداختند. برای این منظور الگوریتم‌های یادگیری ماشین K نزدیک‌ترین همسایه (KNN)، ماشین بردار پشتیبان (SVR)، جنگل تصادفی (RF) و رگرسیون تقویت گرادیان (GB) استفاده شد. نتایج نشان داد که مدل رگرسیون GB عملکرد بهتری نسبت به دیگر مدل‌ها داشت. همچنین بیان کرد که استفاده از الگوریتم‌های یادگیری ماشین برای مدل‌سازی پیش‌بینی انتشار CO₂ مناسب است. (Hamrani et al. (2020) در پژوهش خود قابلیت استفاده از مدل‌های رگرسیون کلاسیک، یادگیری سطحی و یادگیری عمیق برای پیش‌بینی انتشار گازهای گلخانه‌ای خاک در زمین کشاورزی را مورد بررسی قرار دادند. مدل‌های رگرسیون کلاسیک به‌طور رضایت‌بخشی تغییرات چرخه‌ای و فصلی انتشار CO₂ را شبیه‌سازی کردند. (Ebrahimi et al. (2019) در پژوهش خود به تخمین تنفس پایه خاک در عمق ۲۰ سانتی‌متری در کاربری‌های مختلف با استفاده از شبکه عصبی مصنوعی و روش رگرسیون خطی پرداختند. به-طور کلی، روش رگرسیون خطی در مقایسه با شبکه عصبی مصنوعی دارای ضریب تبیین کمتری بود. بنابراین، نتایج نشان می‌دهد که مدل‌سازی با شبکه عصبی مصنوعی یک روش قابل اعتماد برای پیش‌بینی تنفس خاک است.

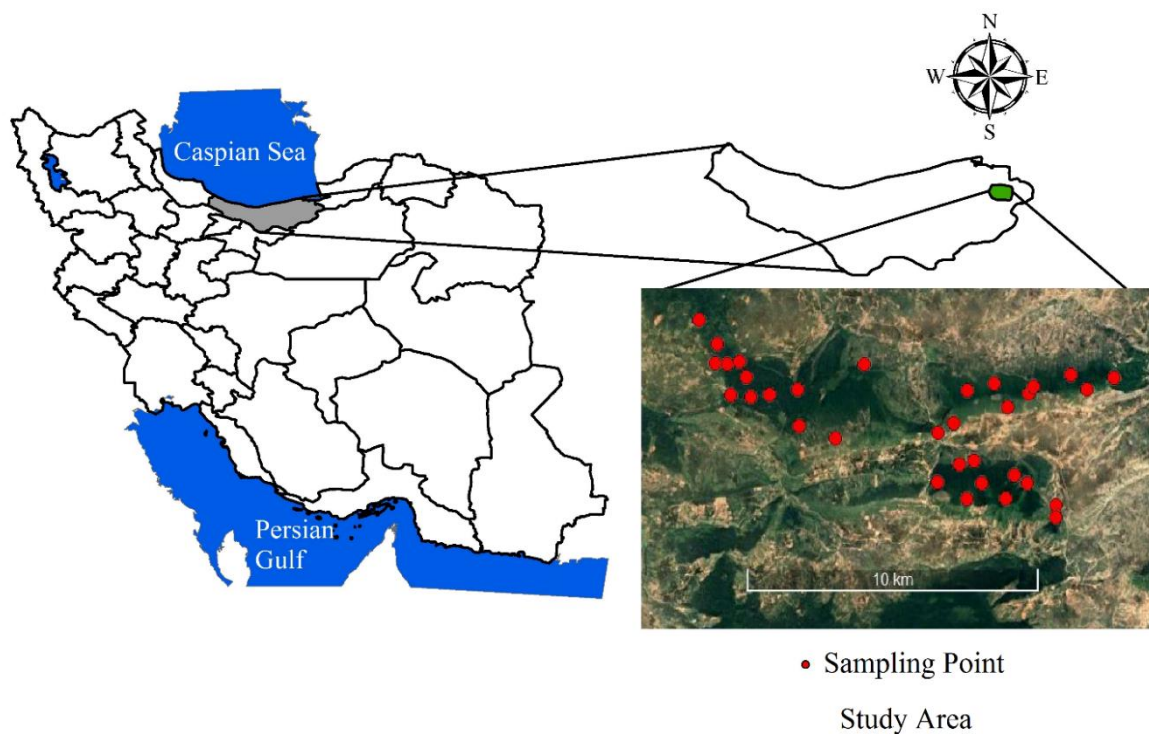
با توجه به افزایش نگرانی‌ها در زمینه گرمایش جهانی و تغییر اقلیم، ظرفیت خاک و قابلیت آن نیز در ترسیب کربن به‌صورت پایدار مورد توجه ویژه‌ای قرار گرفته است. در سال‌های اخیر توجه به ماده آلی خاک در رابطه با ترسیب کربن افزایش یافته و دستیابی به افزایش ترسیب کربن خاک به‌عنوان روش مناسبی برای

مواد و روش‌ها

منطقه مورد بررسی

این پژوهش در جنگل‌های سری شیلر، در حوزه آبخیز ۷۶، شهرستان بهشهر، استان مازندران انجام شد (شکل

۱). مساحت سری ۱۲۹۹۰ هکتار، کمترین ارتفاع ۷۵۰ متر و بیشترین ارتفاع ۱۸۵۰ متر است. متوسط بارندگی سالانه ۳۱۷ میلی‌متر و کمترین مقدار درجه حرارت ۲/۵ درجه سانتی‌گراد و بیشترین مقدار درجه حرارت ۲۱/۵ درجه سانتی‌گراد است.



شکل ۱ - موقعیت منطقه مورد بررسی

Figure 1. The location of study area

روش انجام پژوهش

با توجه به وسعت زیاد منطقه، مراکز قطعات نمونه به صورت تصادفی در کاربری جنگل تعیین شدند (Mannan et al., 2018). در هر یک از قطعات نمونه (۱۰۰۰ مترمربع) ویژگی‌های کمی درختان شامل قطر برابر سینه و ارتفاع کل تمام درختان (قطر برابر سینه بیشتر از ۷/۵ سانتی‌متر) به ترتیب با استفاده از خط‌کش دو بازو و دستگاه شیب‌سنج اندازه‌گیری و ثبت شد.

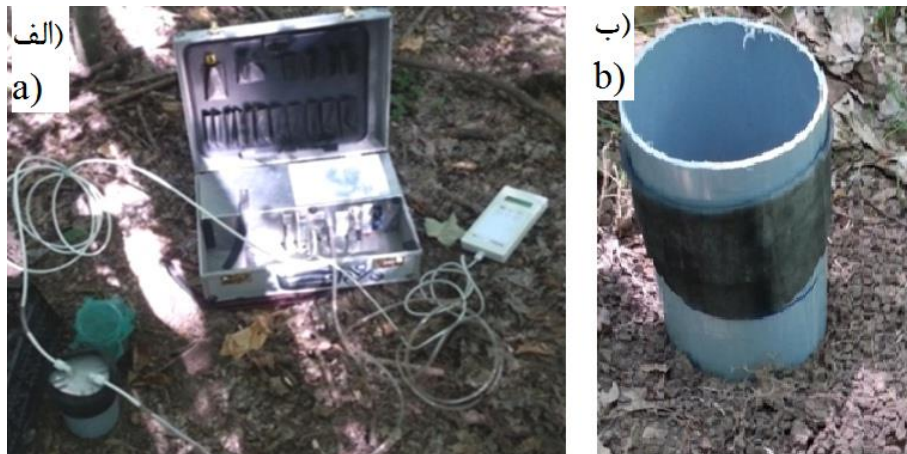
برای بررسی ویژگی‌های خاک، در مرکز و چهار گوشه هر یک از قطعات نمونه پس از کنارزدن لایه لاشبرگ از عمق صفر تا ۲۰ سانتی‌متر با استفاده از روش

استوانه فلزی (قطر دهانه هشت سانتی‌متر) نمونه خاک با رعایت فاصله یک متر از تنه درختان تهیه شد (Molina et al., 2022). سپس نمونه‌ها باهم آمیخته شده و یک نمونه برای هر قطعه نمونه استخراج شد. نمونه‌های خاک درون پلاستیک قرار داده شده و به آزمایشگاه منتقل شدند.

در آزمایشگاه ویژگی‌های فیزیکی خاک شامل درصد رطوبت به روش وزنی، چگالی ظاهری به روش کلوخه (Piper, 2019) و بافت خاک به روش هیدرومتری (Huluka, 2014) اندازه‌گیری شدند.

(شکل ۲ الف). تنفس خاک با دستگاه CO₂ port (ساخت آلمان) اندازه‌گیری شد (شکل ۲ ب) (Hojjati and Lamersdorf, 2010). تمام اندازه‌گیری‌ها در فاصله زمانی ۸ تا ۱۱ صبح (به وقت محلی) انجام شد. دما و رطوبت حجمی خاک به‌طور هم‌زمان با تنفس خاک در کنار لوله‌های پلاستیکی اندازه‌گیری شد. دمای خاک با استفاده از دماسنج در عمق ۱۰ سانتی‌متری اندازه‌گیری شد. رطوبت خاک با استفاده از دستگاه قابل‌حمل سنجش رطوبت خاک (روش TDR) در عمق ۱۰ سانتی‌متری مورد اندازه‌گیری قرار گرفت (Tafazoli et al., 2021).

ویژگی‌های شیمیایی خاک شامل اسیدیته خاک (pH) به روش پتانسیومتری (Piper, 2019)، هدایت الکتریکی (EC) به روش هدایت سنجی (نسبت خاک به آب برابر با یک به دو و نیم بود)، کربن آلی به روش والکلی و بلاک، نیتروژن به روش کجلدال (Yeomans and Bremner, 1991)، فسفر قابل‌جذب با روش اولسن (Olsen et al., 1954)، پتاسیم با روش عصاره‌گیری با استات آمونیم (Piper, 2019) اندازه‌گیری شدند. برای اندازه‌گیری تنفس خاک لوله‌های پلاستیکی (PVC؛ قطر ۱۰ و ارتفاع ۲۰ سانتی‌متر) در عمق ۱۰ سانتی‌متری خاک نصب شد (روش اتاقک تاریک)



شکل ۲- الف) اندازه‌گیری تنفس خاک با دستگاه CO₂ - Port؛ ب) لوله‌های پلاستیکی برای اندازه‌گیری تنفس خاک

Figure 2. a) Measurement of soil respiration with CO₂-Port device; b) plastic tubes (PVC) for measuring soil respiration

$$ABG = 0.1173 \times (WD \times DBH^2 \times H)^{0.928} \quad \text{رابطه (۳)}$$

در روابط بالا، G سطح مقطع درختان، DBH قطر برابرسینه، V حجم درختان، H ارتفاع کل درختان، f ضریب کاهش قطری، AGB روی زمینی درختان و WD چگالی چوب است. پس از اندازه‌گیری مشخصه‌های درصد کربن و وزن مخصوص ظاهری، مقادیر ترسیب کربن خاک برحسب تن در هکتار از رابطه ۴ محاسبه شد (Razakamanarivo et al., 2011).

تجزیه و تحلیل داده‌ها

با استفاده از قطر درختان ابتدا سطح مقطع درختان (رابطه ۱) و در آخر حجم درختان (رابطه ۲) بر حسب مترمکعب محاسبه شد. زی‌توده روی زمینی درختان (AGB) با استفاده از مدل‌های آلومتریک برای جنگل هیرکانی ($R^2=0.3975$ ، رابطه ۳)، محاسبه شد (Sharifi et al., 2013):

$$G = \frac{\pi}{4} DBH^2 \quad \text{رابطه (۱)}$$

$$V = G \times H \times f \quad \text{رابطه (۲)}$$

GridSearchCV در بسته Scikit-learn تعیین شد (Adjuik and Davis, 2022).

برای پیش‌بینی تنفس خاک از متغیرهای مستقل شامل: وزن مخصوص ظاهری، درصد رطوبت، درصد اجزای بافت خاک (شن، سیلت و رس)، دمای خاک، نیتروژن کل، فسفر قابل جذب، پتاسیم قابل جذب، کربن آلی و زی‌توده روی زمینی درختان استفاده شد. این عوامل در سراسر جهان به‌عنوان عوامل اصلی کنترل کننده انتشار گاز به‌ویژه CO₂ شناخته شده‌اند (Ebrahimi et al., 2019). همچنین برای پیش‌بینی ترسیب کربن خاک از متغیرهای مستقل شامل: دما و رطوبت خاک و زی‌توده روی زمینی درختان استفاده شد. در ادامه اهمیت نسبی متغیرها با استفاده از شاخص درصد افزایش میانگین خطای پیش‌بینی در زبان برنامه‌نویسی پایتون با تابع permutation importance در بسته Scikit-learn محاسبه شد. در این روش یک متغیر پیش‌بینی‌کننده زمانی با اهمیت شناخته می‌شود که به هم زدن و جابه‌جایی مقادیر آن منجر به افزایش خطای پیش‌بینی شود (Adjuik and Davis, 2022).

برای مدل‌سازی، داده‌ها به دو گروه داده‌های واسنجی (۸۰ درصد) و داده‌های اعتبارسنجی (۲۰ درصد) تقسیم شدند (Adjuik and Davis, 2022). در ادامه به‌منظور ارزیابی دقت در هر یک از روش‌های مدل‌سازی مورد استفاده در برآورد مقدار ترسیب کربن و تنفس خاک از آماره‌های ضریب تبیین (R²)، اریبی مدل (Bias)، میانگین قدر مطلق خطا (Mean Absolute Error)، جذر میانگین مربعات خطا (Root Mean Square Error) استفاده شد. ضریب تعیین نشان می‌دهد که چند درصد تغییرات متغیر وابسته به‌وسیله متغیر مستقل تعیین می‌شوند یا به‌عبارت دیگر نشان‌دهنده این است که چه مقدار از تغییرات متغیر وابسته تحت تأثیر متغیر مستقل مربوطه بوده است و مابقی

$$\text{رابطه (۴)} \quad C_s = 10000 \times OC (\%) \times Bd \times E$$

که در رابطه فوق، C_s، ترسیب کربن (kg/ha)، OC، درصد کربن آلی، Bd، وزن مخصوص ظاهری خاک (g/cm³) و E عمق لایه به سانتی‌متر است. در ادامه در این پژوهش مقدار ترسیب کربن و تنفس خاک با چهار الگوریتم یادگیری ماشین شامل رگرسیون خطی چندگانه، K-نزدیک‌ترین همسایه، ماشین بردار پشتیبان و جنگل تصادفی با استفاده از زبان برنامه‌نویسی پایتون (نسخه ۳.۸.۵) و بسته Scikit-learn برنامه‌ریزی و برآورد شد. بسته Scikit-learn از بسته‌های متن‌باز و قدرتمند در زبان برنامه‌نویسی پایتون است که برای اهداف یادگیری ماشین به‌کار می‌رود. در مدل جنگل تصادفی از سه متغیر قابل تنظیم برای بهبود مدل شامل تعداد متغیرهای مورد استفاده در ساخت هر درخت، تعداد درختان جنگل تصادفی، و حداقل گره‌های انتهایی استفاده می‌شود (Estévez et al., 2022). متغیرهای قابل تنظیم برای بهبود مدل ماشین بردار پشتیبان، بستگی به نوع مجموعه‌ای توابع ریاضی (Kernel) مورد استفاده دارد. این مجموعه توابع می‌تواند انواع متفاوتی داشته باشند که به‌عنوان مثال می‌توان به توابع غیرخطی، چندجمله‌ای، تابع پایه شعاعی (Radial Basis Function یا به‌اختصار RBF) اشاره کرد. در این پژوهش از تابع RBF استفاده شد که دو متغیر مهم قابل تنظیم آن شامل جریمه و گاما است (Hamrani et al., 2020). متغیرهای قابل تنظیم برای مدل نزدیک‌ترین همسایه شامل تعداد همسایه‌ها و تابع فاصله برای تخمین داده‌های جدید است. در این پژوهش از فاصله منهتن (Manhattan) استفاده شد و بهترین تعداد همسایه با استفاده از کمینه مقدار خطا تعیین شد (Adjuik and Davis, 2022). پس از برنامه‌ریزی مدل‌ها، بهترین متغیرهای مدل‌ها با استفاده از تابع

اندازه‌گیری شده با هم برابر باشند یعنی برابر با صفر باشد (Hodson, 2022).

نتایج

مشخصات کمی درختان و ویژگی‌های خاک در منطقه موردبررسی

نتایج مربوط به مشخصات کمی درختان در منطقه موردبررسی در جدول ۱ گزارش شده است. در این پژوهش زی توده روی زمینی درختان رو به‌عنوان متغیر مستقل در برآورد ترسیب کربن و تنفس خاک مورد استفاده قرار گرفت. همچنین ویژگی‌های فیزیکی و شیمیایی خاک در منطقه موردبررسی نیز در جدول ۲ گزارش شده است.

تغییرات متغیر وابسته مربوط به دیگر عوامل هستند. مقدار آن بین صفر و یک است و هرچه مقدار آن نزدیک به یک نزدیک‌تر باشد نشان می‌دهد که مدل داده‌ها را خوب برازش کرده است. مقدار اریبی نشان‌دهنده میانگین تفاوت بین مقدار واقعی و مقدار برآورد شده است و هرچه مقدار آن به صفر نزدیک‌تر باشد، مدل برازش بهتری دارد. آماره MAE میانگین قدر مطلق تفاضل بین مقدار پیش‌بینی و واقعی را محاسبه می‌کند. MAE برابر صفر نشان‌دهنده تطبیق کامل مقادیر اندازه‌گیری شده و برآورد شده است. مقدار RMSE نشان‌دهنده این است مقادیر پیش‌بینی شده تا چه حد اندازه‌گیری را بیشتر یا کمتر تخمین زده‌اند. بهترین حالت برای RMSE این است که مقادیر پیش‌بینی شده و

جدول ۱- مشخصات کمی درختان (میانگین \pm اشتباه معیار) در منطقه موردبررسی

Table 1. Quantitative characteristics of trees (mean \pm standard error) in the study area

میانگین \pm اشتباه معیار Mean \pm standard error	مشخصات کمی درختان Quantitative Characteristics of trees
185.31 \pm 16.32	تعداد در هکتار Number per hectare
28.89 \pm 1.42	قطر برابر سینه (سانتی‌متر) dbh (cm)
16.3 \pm 0.84	ارتفاع کل (متر) Total height
19.28 \pm 2.35	سطح مقطع (مترمربع در هکتار) Basal area (m ² /ha)
36.53 \pm 3.9	حجم (مترمکعب در هکتار) Volume (m ³ /ha)
235.91 \pm 30.57	زی توده روی زمینی (کیلوگرم در هکتار) Above ground biomass (kg/ha)

جدول ۲- ویژگی‌های خاک (میانگین \pm اشتباه معیار) در منطقه مورد بررسیTable 2. Soil properties (mean \pm standard error) in the study area

میانگین \pm اشتباه معیار	ویژگی‌های خاک	میانگین \pm اشتباه معیار	ویژگی‌های خاک
Mean \pm SE	Soil Properties	Mean \pm SE	Soil Properties
0.19 \pm 0.01	نیترژن کل (درصد) Total nitrogen (%)	34.05 \pm 1.54	درصد رطوبت Moisture (%)
4.72 \pm 0.21	درصد کربن آلی Organic carbon (%)	21.47 \pm 0.23	دما (درجه سانتی‌گراد) Temperature ($^{\circ}$ C)
5.76 \pm 0.55	فسفر قابل جذب (میلی‌گرم بر کیلوگرم) Available phosphorus (mg/kg)	1.66 \pm 0.04	چگالی ظاهری Bulk density (g/m ³)
672.67 \pm 23.78	پتاسیم قابل جذب (میلی‌گرم بر کیلوگرم) Available potassium (mg/kg)	59.32 \pm 1.60	درصد شن Sand (%)
5.87 \pm 0.75	غلظت آمونیم (میلی‌گرم بر کیلوگرم) Ammonium concentration (mg/kg)	27.11 \pm 1.52	درصد سیلت Silt (%)
33.66 \pm 1.87	غلظت نیترات (میلی‌گرم بر کیلوگرم) Nitrate concentration (mg/kg)	13.56 \pm 0.56	درصد رس Clay (%)
76.12 \pm 3.15	ترسیب کربن خاک (کیلوگرم در هکتار) Soil carbon sequestration (kg/ha)	6.47 \pm 0.07	اسیدیته pH
57.48 \pm 0.88	تنفس خاک (مول کربن در ساعت در مترمربع) Soil respiration (mol C.h ⁻¹ .m ⁻²)	0.51 \pm 0.014	هدایت الکتریکی (دسی‌زیمنس بر متر) EC (ds/m)

برآورد ترسیب کربن و تنفس خاک

معیارهای ارزیابی مدل‌های مختلف پیش‌بینی تنفس خاک در جدول ۴ گزارش شده است. بر این اساس، مدل یادگیری رگرسیون خطی پایین‌ترین عملکرد (RMSE= ۲/۴۸ و R²= ۰/۶۵) را داشت. این درحالی است که مدل ماشین بردار پشتیبان بهترین عملکرد را داشت (RMSE= ۰/۷۷ و R²= ۰/۹۰). همچنین بعد از مدل ماشین بردار پشتیبان، مدل جنگل تصادفی بهترین عملکرد را داشت (RMSE= ۲/۳۴ و R²= ۰/۹۳). همبستگی مقدار اندازه‌گیری شده و برآورد شده تنفس خاک برای مدل‌های مورد بررسی در شکل ۵ نشان داده شده است. همچنین پراکنش مقادیر برآورد شده تنفس خاک در مقابل باقی‌مانده‌ها (اختلاف بین مقدار برآورد شده و واقعی) در شکل ۶ نشان داده شده است.

جدول ۳ معیارهای ارزیابی برای مدل‌های مختلف پیش‌بینی ترسیب کربن را ارائه می‌کند. بر اساس نتایج به‌دست‌آمده، رگرسیون خطی پایین‌ترین عملکرد (RMSE= ۱۲/۰۷ و R²= ۰/۳۶) را داشت. در حالی که مدل جنگل تصادفی بالاترین عملکرد (RMSE= ۱۰/۴۷ و R²= ۰/۸۲) را برای برآورد ترسیب کربن خاک به‌همراه داشت. پس از مدل جنگل تصادفی، مدل ماشین بردار پشتیبان عملکرد مناسب‌تری (RMSE= ۱/۱۶ و R²= ۰/۵۸) را برای ترسیب کربن خاک به‌همراه داشت. همبستگی مقدار اندازه‌گیری شده و برآورد شده ترسیب کربن خاک برای مدل‌های مورد بررسی در شکل ۳ نشان داده شده است. همچنین پراکنش مقادیر برآورد شده ترسیب کربن خاک در مقابل باقی‌مانده‌ها (اختلاف بین مقدار برآورد شده و واقعی) در شکل ۴ نشان داده شده است.

جدول ۳- معیارهای ارزیابی مدل‌های یادگیری ماشین برای پیش‌بینی ترسیب کربن خاک

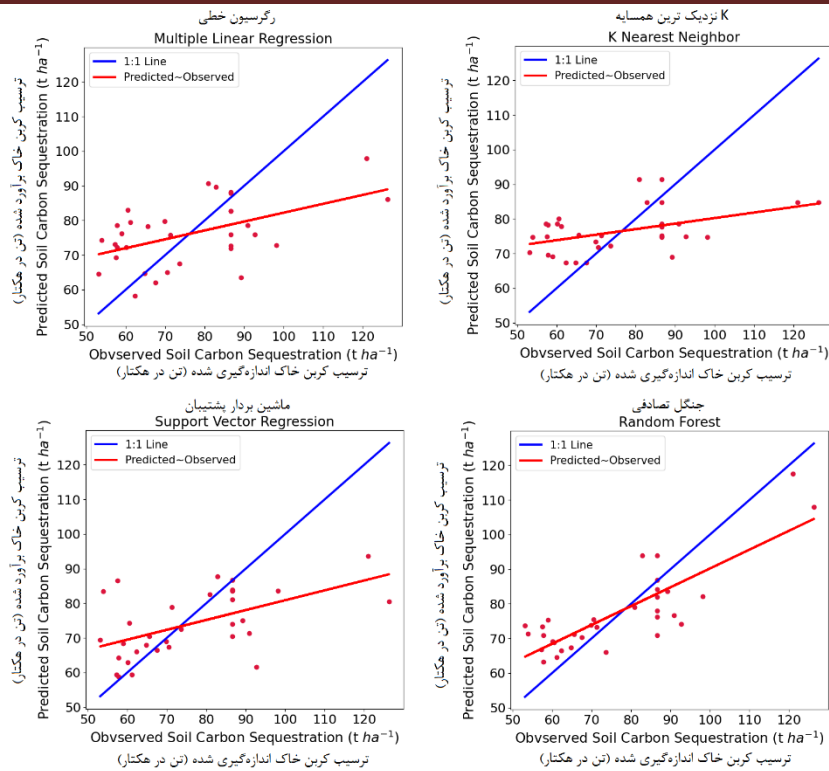
Table 3. Evaluation of machine learning models for predicting soil carbon sequestration

جذر میانگین	مربعیات خطا (مول کربن در ساعت در مترمربع)	درصد میانگین قدر مطلق خطا Mean absolute percentage error	میانگین قدر مطلق خطا (مول کربن در ساعت در مترمربع) Mean absolute error (mol C.h ⁻¹ .m ⁻²)	اریبی (مول کربن در ساعت در مترمربع) Bias (mol C.h ⁻¹ .m ⁻²)	ضریب تبیین (R ²) Coefficient of determination (R ²)	
22.59	12.07	16.14	11.23	-1.83	0.36	رگرسیون خطی Linear regression
16.02	9.21	13.05	8.61	-5.13	0.20	K نزدیک‌ترین همسایه K nearest neighbor
2.03	1.16	10.5	0.93	-0.10	0.58	ماشین بردار پشتیبان Support vector regression
18.22	10.47	14.11	8.85	-6.13	0.82	جنگل تصادفی Random forest

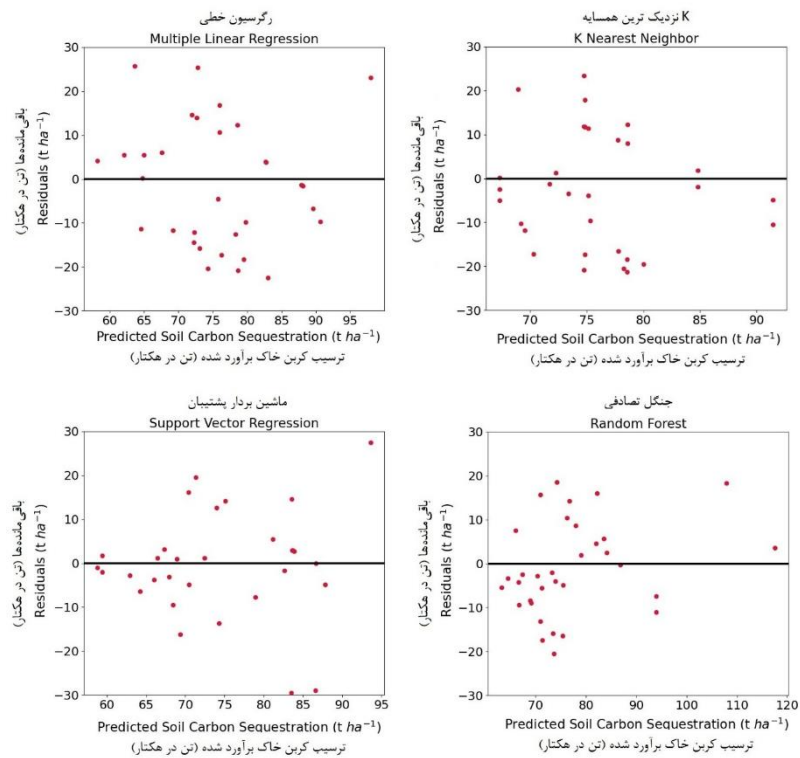
جدول ۴- معیارهای ارزیابی مدل‌های یادگیری ماشین برای پیش‌بینی انتشار گاز CO₂

Table 4. Evaluation of machine learning models for predicting CO₂ flux

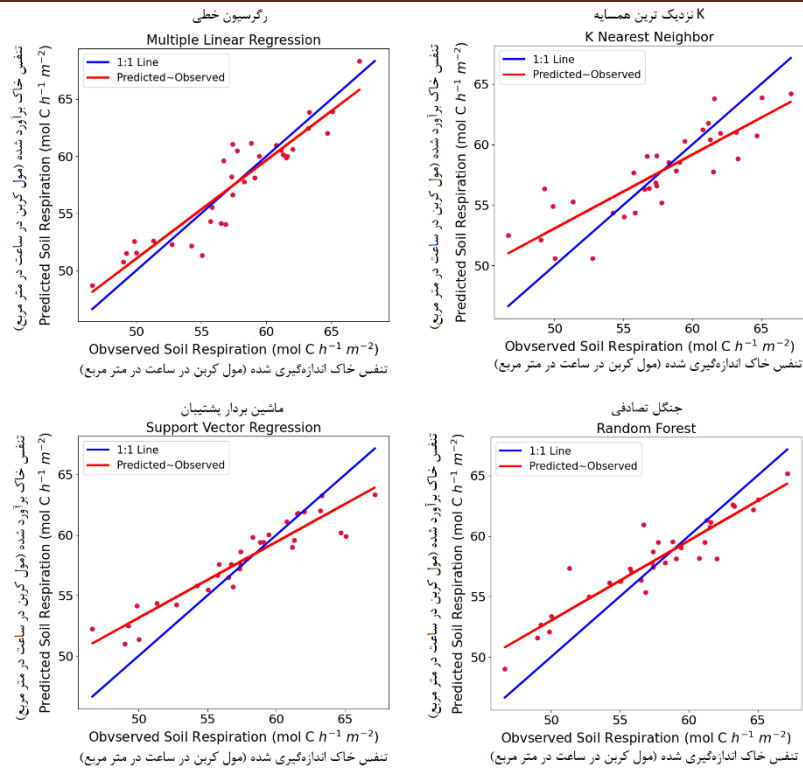
جذر میانگین	مربعیات خطا (مول کربن در ساعت در مترمربع)	درصد میانگین قدر مطلق خطا Mean absolute percentage error	میانگین قدر مطلق خطا (مول کربن در ساعت در مترمربع) Mean absolute error (mol C.h ⁻¹ .m ⁻²)	اریبی (مول کربن در ساعت در مترمربع) Bias (mol C.h ⁻¹ .m ⁻²)	ضریب تبیین (R ²) Coefficient of determination (R ²)	
4.31	2.48	2.39	4.24	-2.39	0.65	رگرسیون خطی Linear regression
3.74	2.15	1.72	3.02	0.35	0.69	K نزدیک‌ترین همسایه K nearest neighbor
1.35	0.77	0.65	1.65	-0.11	0.90	ماشین بردار پشتیبان Support vector regression
4.08	2.34	2.02	3.50	0.14	0.93	جنگل تصادفی Random forest



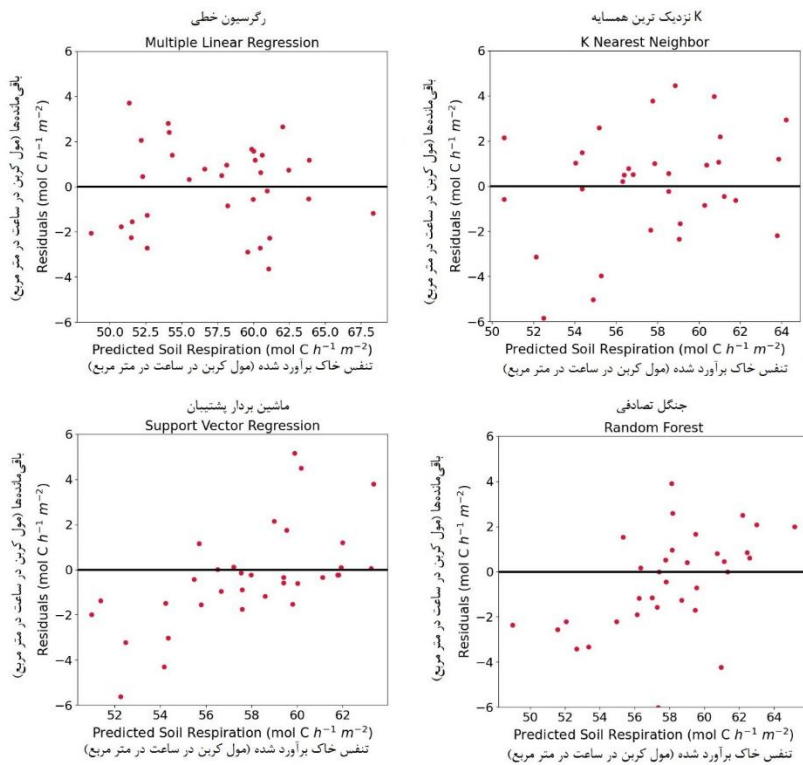
شکل ۳- همبستگی بین مقادیر اندازه‌گیری شده و برآورد شده در مدل‌های مورد استفاده برای برآورد توسیع کربن خاک
Figure 3. Scatter plots of the measured and estimated values of the models used to estimate soil carbon sequestration



شکل ۴- پراکنش مقادیر باقی‌مانده در مقابل مقادیر برآورد شده در مدل‌های مورد استفاده برای برآورد توسیع کربن خاک
Figure 4. Scatter plots of the residuals and estimated values of the models used to estimate soil carbon sequestration



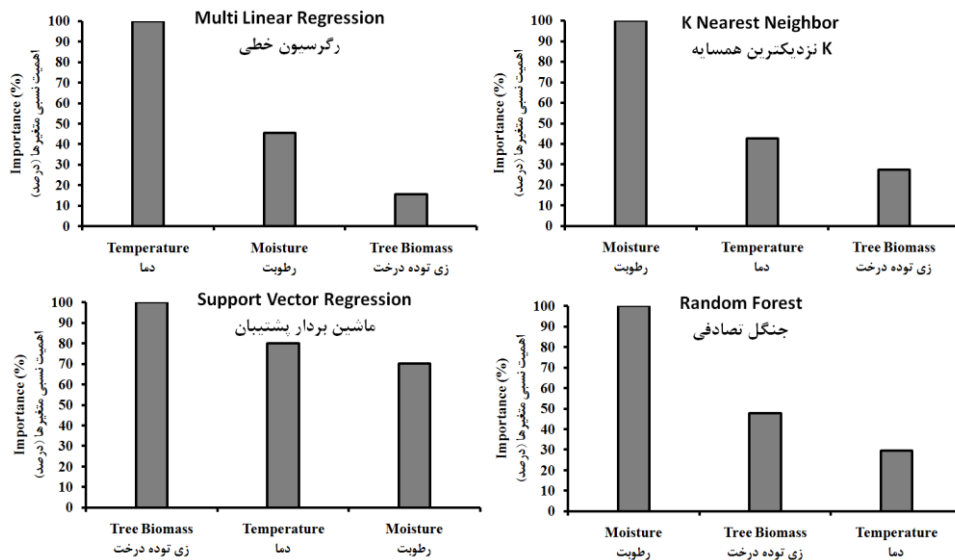
شکل ۵- همبستگی بین مقادیر اندازه‌گیری شده و برآورد شده در مدل‌های مورد استفاده برای برآورد تنفس خاک
 Figure 5. Scatter plots of the measured and estimated values in the models used to estimate soil respiration



شکل ۶- پراکنش مقادیر باقی‌مانده در مقابل مقادیر برآورد شده در مدل‌های مورد استفاده برای برآورد تنفس خاک
 Figure 6. Scatter plots of the residuals and estimated values of the models used to estimate soil respiration

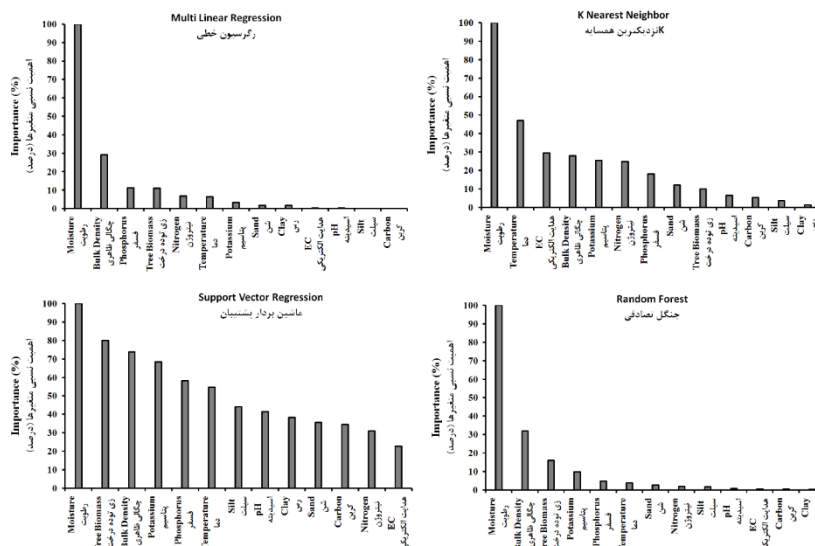
تصادفی رطوبت خاک مهم‌ترین عامل بود. در مدل ماشین بردار پشتیبان، متغیر زی توده روی زمینی درختان مهم‌ترین عامل بود. در برآورد تنفس خاک روند مشابهی دیده شد و درصد رطوبت خاک به‌عنوان مهم‌ترین متغیر در تمام مدل‌های مورد استفاده شناسایی شد.

اهمیت نسبی متغیرهای مورد استفاده در فرآیند مدل‌سازی ترسیب کربن و تنفس خاک به‌ترتیب در شکل‌های ۷ و ۸ نشان داده شده است. در برآورد ترسیب کربن خاک با استفاده از مدل رگرسیون خطی، دمای خاک به‌عنوان مهم‌ترین عامل شناسایی شد در حالی که در مدل نزدیک‌ترین همسایه و جنگل تصادفی



شکل ۷- اهمیت نسبی متغیرهای مورد بررسی در پیش‌بینی ترسیب کربن آلی خاک با استفاده از مدل‌های یادگیری ماشین

Figure 7. The relative importance of the studied variables in predicting soil organic carbon sequestration using the machine learning models



شکل ۸- اهمیت نسبی متغیرهای مورد بررسی در پیش‌بینی تنفس خاک با استفاده از مدل‌های یادگیری ماشین

Figure 8. The relative importance of the studied variables in predicting soil respiration using the machine learning models

بحث

مدل‌های یادگیری ماشین مانند جنگل تصادفی از روابط عملکردی بین پیش‌بینی‌کننده‌ها و متغیر وابسته استفاده می‌کنند که در مرحله آموزش الگوریتم اجرا می‌شوند. در همین راستا (Ebrahimi et al., 2019) در پژوهش خود به برآورد تنفس خاک در کاربری‌های مختلف با استفاده از شبکه عصبی مصنوعی (ANN) و مدل‌های رگرسیون خطی پرداختند و بیان کردند که رگرسیون خطی مقدار R^2 کمتری نسبت به شبکه عصبی مصنوعی داشت. به‌طور کلی نتایج پژوهش مذکور نشان داد که مدل ANN می‌تواند تنفس خاک را با استفاده از اندازه‌گیری ویژگی‌های فیزیکی و شیمیایی و زیستی خاک به‌عنوان متغیر ورودی پیش‌بینی و برآورد کند. همچنین (Tavares et al., 2018) نیز به ارزیابی مدل جنگل تصادفی برای برآورد انتشار گاز دی‌اکسید کربن پرداختند و بیان کردند که مدل مذکور از دقت مناسبی برخوردار است.

به‌طور کلی، نتایج این پژوهش نشان داد که بعد از مدل‌های ماشین بردار پشتیبان و جنگل تصادفی، مدل K نزدیک‌ترین همسایه عملکرد به نسبت قابل قبولی (ترسیب کربن: $RMSE=9/21$ و $R^2=0/2$ ؛ تنفس خاک $RMSE=2/15$ و $R^2=0/69$) به همراه داشت. با این وجود عملکرد بهتر و مناسب‌تر مدل K نزدیک‌ترین همسایه ($R^2=0/80$) در برآورد تنفس خاک در دیگر پژوهش‌ها نیز گزارش شده است. در این پژوهش مقدار K با دقت تنظیم و انتخاب شد زیرا مقدار کم می‌تواند منجر به برآزش بیش از حد داده‌ها شود و یک مقدار K بزرگ‌تر می‌تواند منجر به عدم تناسب داده شود (Adjuik and Davis, 2022). برای این منظور محدوده‌ای از K از ۱ تا ۵۰ مورد بررسی قرار گرفت و $K=8$ مقدار بهینه به‌دست آمد.

با توجه به اطلاعات موجود، بیشتر پژوهش‌های در خصوص تخمین کربن آلی خاک با استفاده از مدل‌های

باوجود اینکه در سال‌های اخیر، مدل‌های یادگیری ماشین به‌طور گسترده برای تخمین تنفس خاک در مقیاس منطقه‌ای و جهانی استفاده شده است (Zhu et al., 2020). با این حال، تاکنون پژوهشی در مورد عملکرد مدل‌های مختلف یادگیری ماشین را در تخمین ترسیب کربن و تنفس خاک در جنگل‌های با ارزش شمال کشور ارزیابی نکرده است. در این پژوهش، عملکرد ۴ مدل رگرسیون خطی، K نزدیک‌ترین همسایه، ماشین بردار پشتیبان و جنگل تصادفی برای تخمین تنفس خاک در جنگل‌های شمال کشور ارزیابی شد. نتایج نشان داد که برای ترسیب کربن خاک مدل جنگل تصادفی ($RMSE=10/47$ و $R^2=0/82$) و برای تنفس خاک مدل ماشین بردار پشتیبان ($RMSE=0/77$ و $R^2=0/90$) می‌توانند تخمین مناسبی را به‌همراه داشته باشند. این را می‌توان با این واقعیت توضیح داد که تنفس خاک وابستگی قوی به متغیرهای محیطی انتخاب شده دارد و هر ۴ مدل یادگیری ماشین دارای قابلیت‌های کافی برای یادگیری این وابستگی‌های اساسی برای تخمین تنفس خاک هستند، اگرچه اصول یادگیری هر مدل تفاوت‌هایی دارند (Zhu et al., 2020). نتایج مشابه این پژوهش در خصوص عملکرد بهتر مدل ماشین بردار پشتیبان در پژوهش (Adjuik and Davis, 2022) نیز گزارش شد. همچنین (Hamrani et al., 2020) مقدار R^2 را ۰/۹۲ برای مدل SVR و ۰/۹۶ برای مدل جنگل تصادفی در پیش‌بینی تنفس خاک گزارش کرد. به‌طور کلی می‌توان بیان کرد که نتایج این پژوهش در مقایسه با دیگر پژوهش‌ها نیز عملکرد قابل قبولی ارائه کرد. در مجموع نتایج این پژوهش نشان داد که مدل رگرسیون خطی عملکرد پایین‌تری نسبت به دیگر مدل‌های یادگیری داشت. برخلاف مدل‌های رگرسیون ساده که به درک سطح فرآیند از تغییرپذیری بستگی دارد،

اهمیت را داشت. همچنین در مدل جنگل تصادفی که بالاترین دقت را در تخمین ترسیب کربن خاک داشت، متغیر رطوبت خاک مهم‌ترین عامل شناسایی شد. در ارتباط با اهمیت نسبی متغیرها در تخمین تنفس خاک، تقریباً روند مشابهی مشاهده شد و در تمامی مدل‌ها رطوبت خاک به‌عنوان مهم‌ترین متغیر شناسایی شد. همچنین در مدل‌های ماشین بردار پشتیبان و جنگل تصادفی مشاهده می‌شود که زی‌توده روی زمینی به- ترتیب در اولویت دوم و سوم اهمیت شناسایی شده است. با وجود همبستگی مثبت دمای خاک با تنفس خاک در پژوهش این، ولی دمای خاک به‌عنوان مهم‌ترین عامل شناسایی نشد. نتایج این پژوهش در خصوص اهمیت بالای رطوبت خاک در دیگر پژوهش‌ها نیز گزارش شد (Ni et al., 2019).

در اینجا باید تأکید کرد که در این پژوهش هدف ارزیابی مدل‌های یادگیری ماشین برای پیش‌بینی ترسیب کربن و تنفس خاک با ویژگی‌هایی بود که اندازه‌گیری آن آسان‌تر از ترسیب کربن و تنفس خاک است و داده‌های آن ممکن است در مناطق دیگر نیز موجود باشد. بنابراین با توجه به نتایج ذکر شده در بالا، استفاده از مدل‌های یادگیری ماشین حاوی متغیرهای آسان برای اندازه‌گیری می‌تواند نتایج مناسبی را برای برآورد ترسیب کربن و تنفس خاک فراهم کند. این روش تنها با اندازه‌گیری مشخص‌های آسان برای اندازه‌گیری در عرصه امکان ارزیابی یک مشخصه بسیار دشوار قابل ارزیابی مانند ترسیب کربن و تنفس خاک را فراهم می‌کند. برآورد مقدار تنفس خاک بر اساس روش‌های آزمایشگاهی نیز دقیق هستند. با این وجود، برای تعیین و مقایسه معتبر بین خاک‌ها در مناطق مختلف، اندازه‌گیری تنفس معمولاً باید در شرایط طبیعی و در عرصه انجام شود.

نتیجه‌گیری کلی

یادگیری ماشین در اراضی کشاورزی انجام شده و پژوهش‌های کمی در ارتباط با بوم‌سازگان جنگل گزارش شده است. بر این اساس (Zhang et al., 2017) در خصوص پیش‌بینی کربن آلی خاک به دو روش جنگل تصادفی و رگرسیون خطی برتری روش جنگل تصادفی را نسبت به روش رگرسیون خطی نشان دادند و بیان داشتند که برتری مدل جنگل تصادفی نسبت به مدل‌های خطی در پیش‌بینی کربن آلی خاک مربوط به روابط غیرخطی است که بین کربن آلی خاک و متغیرهای پیش‌بینی کننده وجود دارد. Zeraatpisheh et al. (2019) برای برآورد ویژگی‌های کربن آلی، رس و آهک در استفاده از مدل جنگل تصادفی به‌عنوان مدل مناسب‌تر در مقایسه با مدل کویست، رگرسیون درختی و رگرسیون خطی چندمتغیره گزارش کردند.

برای تحلیل اثرهای متغیرهای مختلف بر عملکرد مدل‌های یادگیری ماشین در تخمین ترسیب کربن و تنفس خاک، اهمیت نسبی متغیرها ارزیابی شد (Were et al., 2015). با توجه به اینکه تفسیر مدل‌های یادگیری ماشین نسبت به مدل‌های سنتی (مانند رگرسیون خطی) پیچیده‌تر و سخت‌تر است، بنابراین به‌طور معمول از شاخص اهمیت نسبی متغیرها استفاده می‌شود (Adjuik and Davis, 2022).

پژوهش در مورد عوامل مؤثر بر مقدار کربن آلی و تنفس خاک در هر منطقه می‌تواند در پیش‌بینی تغییرات آن تأثیر به‌سزایی داشته باشد. بررسی مقدار کربن آلی خاک و عوامل مؤثر بر آن لازمه مدیریت مناسب چرخه کربن در مقیاس‌های بزرگ است. در ارتباط با تخمین ترسیب کربن خاک با استفاده از مدل‌های یادگیری ماشین، روند مشابهی برای اهمیت متغیرها مشاهده نشد. در رگرسیون خطی مقدار دما و رطوبت خاک به‌عنوان مهم‌ترین متغیرها شناسایی شد، درحالی که در مدل ماشین بردار پشتیبان متغیر زی‌توده روی زمینی بالاترین

درختان) و همچنین زی توده روی زمینی در بیشتر نقاط موجود است و برای دیگر مناطق نیز با استفاده از روش‌های سنجش از دور قابل محاسبه است، می‌توان با داشتن دما و رطوبت خاک، مقدار ترسیب کربن خاک را نیز اندازه‌گیری کرد. همچنین با توجه به اینکه اندازه-گیری تنفس خاک نیاز به تجهیزات ویژه مناسب و همچنین به‌کارگیری کارشناس باتجربه دارد، بنابراین می‌توان با استفاده از مدل‌های یادگیری ماشین و همچنین اندازه‌گیری ویژگی‌های اولیه خاک، مقدار تنفس خاک را در بوم‌سازگان جنگل با دقت قابل قبول اندازه‌گیری کرد. با توجه به کمبود اطلاعات، انجام پژوهش‌های آینده در ارتباط با ارزیابی دیگر الگوریتم‌های یادگیری ماشین مانند شبکه‌های عصبی مصنوعی و XGBoost در برآورد ترسیب کربن و تنفس خاک لازم و ضروری به نظر می‌رسد.

دسترسی به اطلاعات تکمیلی

اطلاعات تکمیلی و اضافی در خصوص روش‌های مورد استفاده در این پژوهش از طریق مکاتبه با نویسنده مسئول در اختیار علاقه‌مندان قرار خواهد گرفت.

References

- Adjuik, T. A.; Davis, S. C., Machine Learning Approach to Simulate Soil CO₂ Fluxes under Cropping Systems. *Agronomy* **2022**, *12* (1), 197.
- Alazmani, M.; Hojati, S. M.; Waez-Mousavi, S. M.; Tafazoli, M., Effect of alder plantation age on soil carbon sequestration. *Forest Research and Development* **2021**, *7* (2), 279-291. (In persian)
- Ebrahimi, M.; Sarikhani, M. R.; Sinigani, A. A. S.; Ahmadi, A.; Keesstra, S., Estimating the soil respiration under different land uses using artificial neural network and linear regression models. *Catena* **2019**, *174*, 371-382.
- Estévez, V.; Beucher, A.; Mattbäck, S.; Boman, A.; Auri, J.; Björk, K.-M.; Österholm, P.,

همان‌طور که داده‌های بیشتری از طریق اندازه‌گیری‌های مستقیم در دسترس عموم قرار می‌گیرد، ابزارهای ترکیبی بیشتری برای تفسیر حجم عظیم اطلاعات در پایگاه‌های داده مورد نیاز خواهد بود. با دردسترس بودن داده‌های بیشتر، نیاز به روش‌هایی برای استفاده از این داده‌ها برای درک اینکه چگونه ویژگی‌های خاک بر انتشار گازهای گلخانه‌ای تأثیر می‌گذارد، لازم و ضروری است. ظهور الگوریتم‌های مدل‌سازی یادگیری ماشین در چند دهه گذشته فرصت‌هایی را برای استفاده از این پایگاه‌های داده موجود برای تحلیل و شبیه‌سازی انتشار گازهای گلخانه‌ای در بوم‌سازگان مختلف فراهم کرده است. در این پژوهش، چهار مدل یادگیری ماشین (رگرسیون-خطی، k-نزدیک‌ترین همسایه، ماشین بردار پشتیبان، جنگل تصادفی) از نظر تخمین ترسیب کربن و تنفس خاک ارزیابی و مقایسه شد. نتایج نشان داد که هر چهار مدل در جنگل‌های شمال کشور به‌خوبی ترسیب کربن و تنفس خاک را برآورد کردند، به‌طوری که مدل جنگل تصادفی برای ترسیب کربن و مدل ماشین بردار پشتیبان برای تنفس خاک بالاترین عملکرد را داشت. به‌طور کلی، با توجه به اینکه داده‌های مربوط به آماربرداری جنگل‌های شمال کشور (شامل قطر، ارتفاع و نوع گونه

Machine learning techniques for acid sulfate soil mapping in southeastern Finland. *Geoderma* **2022**, *406*, 115446.

- Hamrani, A.; Akbarzadeh, A.; Madramootoo, C. A., Machine learning for predicting greenhouse gas emissions from agricultural soils. *Science of The Total Environment* **2020**, *741*, 140338.

Hodson, T.O., Root-mean-square error (RMSE) or mean absolute error (MAE): when to use them or not. *Geoscientific Model Development* **2022**, *15* (14), 5481-5487.

- Hojjati, S. M.; Hashemi, S. A.; Hosseyni, S. M.; Asadiyan, M.; Tafazoli, M., The Effect of plantation with native and exotic species on soil CO₂ emissions (The case study: Darabkola forest). *Journal of Plant Ecosystem Conservation* **2020**, *8* (16), 95-110. (In persian)

- Hojjati, S. M.; Tafazoli, M.; Imani, M.; Alazmani, M.; Fallah, A.; Pourmajidian, M. R., Variation in Carbon Sequestration and Soil Properties in Relation to Stand Age in Maple and Alder Plantations. *Journal of Sustainable Forestry* **2022**, 1-15.
- Hojjati, S. M.; Lamersdorf, N. P., Effect of canopy composition on soil CO₂ emission in a mixed sprucebeech forest at Solling, Central Germany. *Journal of Forestry Research* **2010**, 21 (4), 461-464.
- Huluk, G.; Miller, R., Particle size determination by hydrometer method. *Southern Cooperative Series Bulletin* **2014**, 419, 180-184.
- Liakos, K. G.; Busato, P.; Moshou, D.; Pearson, S.; Bochtis, D., Machine learning in agriculture: A review. *Sensors* **2018**, 18 (8), 2674.
- Mannan, A.; Feng, Z.; Ahmad, A.; Liu, J.; Saeed, S.; Mukete, B., Carbon dynamic shifts with land use change in Margallah Hills National Park, Islamabad (Pakistan) from 1990 to 2017. *Applied Ecology and Environmental Research* **2018**, 16 (3), 3197-3214.
- Matinfar, H. R.; Maghsodi, Z.; Mousavi, S. R.; Rahmani, A., Evaluation and Prediction of Topsoil organic carbon using Machine learning and hybrid models at a Field-scale. *Catena* **2021**, 202, 105258.
- Molina, A. J.; Bautista, I.; Lull, C.; del Campo, A.; González-Sanchis, M.; Lidón, A., Effects of Thinning Intensity on Forest Floor and Soil Biochemical Properties in an Aleppo Pine Plantation after 13 Years: Quantity but Also Quality Matters. *Forests* **2022**, 13 (2), 255.
- Naderi, M.; Kialashaki, A.; Veisi, R.; Sheykheslami, A.; Tafazoli, M., Effect of Site on Soil Properties and Carbon Sequestration in Populus deltoids Stand in Sari. *Ecology of Iranian Forest* **2021**, 9 (18), 187-195. (In persian)
- Ni, X.; Liao, S.; Wu, F.; Groffman, P. M., Short-term precipitation pulses stimulate soil CO₂ emission but do not alter CH₄ and N₂O fluxes in a northern hardwood forest. *Soil Biology and Biochemistry* **2019**, 130, 8-11.
- Olsen, S. R., *Estimation of available phosphorus in soils by extraction with sodium bicarbonate*. US Department of Agriculture: 1954.
- Osabohien, R.; Matthew, O.; Aderounmu, U.; Olawande, T., Greenhouse gas emissions and crop production in West Africa: Examining the mitigating potential of social protection. *International Journal of Energy Economics and Policy* **2019**, 9 (1), 57.
- Pan, Y.; Birdsey, R. A.; Fang, J.; Houghton, R.; Kauppi, P. E.; Kurz, W. A.; Phillips, O. L.; Shvidenko, A.; Lewis, S. L.; Canadell, J. G., A large and persistent carbon sink in the world's forests. *science* **2011**, 333 (6045), 988-993.
- Piper, C. S., *Soil and plant analysis*. Scientific Publishers: 2019.
- Razakamanarivo, R. H.; Grinand, C.; Razafindrakoto, M. A.; Bernoux, M.; Albrecht, A., Mapping organic carbon stocks in eucalyptus plantations of the central highlands of Madagascar: A multiple regression approach. *Geoderma* **2011**, 162 (3-4), 335-346.
- Sharifi, A.; Amini, J.; Pourshakouri, F., Allometric model development for Above-Ground Biomass estimation in Hyrcanian forests of Iran. *World Applied Sciences Journal* **2013**, 28 (9), 1322-1330.
- Tafazoli, M.; Hojjati, S. M.; Jalilvand, H.; Lamersdorf, N.; Tafazoli, M., Effect of nitrogen addition on soil CO₂ efflux and fine root biomass in maple monocultures of the hyrcanian region. *Annals of Forest Science* **2021**, 78 (2), 1-11.
- Tavares, R. L. M.; Oliveira, S. R. d. M.; Barros, F. M. M. d.; Farhate, C. V. V.; Souza, Z. M. d.; Scala Junior, N. L., Prediction of soil CO₂ flux in sugarcane management systems using the random forest approach. *Scientia Agricola* **2018**, 75, 281-287.
- Were, K.; Bui, D. T.; Dick, Ø. B.; Singh, B. R., A comparative assessment of support vector regression, artificial neural networks, and random forests for predicting and mapping soil organic carbon stocks across an Afromontane landscape. *Ecological Indicators* **2015**, 52, 394-403.
- Yeomans, J.; Bremner, J., Carbon and nitrogen analysis of soils by automated combustion techniques. *Communications in Soil Science and Plant Analysis* **1991**, 22 (9-10), 843-850.
- Zeraatpisheh, M.; Ayoubi, S.; Jafari, A.; Tajik, S.; Finke, P., Digital mapping of soil properties using multiple machine learning in a semi-arid region, central Iran. *Geoderma* **2019**, 338, 445-452.
- Zhang, H.; Wu, P.; Yin, A.; Yang, X.; Zhang, M.; Gao, C., Prediction of soil organic carbon in an intensively managed reclamation zone of eastern China: A comparison of multiple linear regressions and the random forest

model. *Science of the Total Environment* **2017**, 592, 704-713.
Zhu, X.; He, H.; Ma, M.; Ren, X.; Zhang, L.;
Zhang, F.; Li, Y.; Shi, P.; Chen, S.; Wang, Y.,

Estimating ecosystem respiration in the
grasslands of northern China using machine
learning: Model evaluation and comparison.
Sustainability **2020**, 12 (5), 2099.

Estimation of carbon sequestration and forest soil respiration using machine learning models in Eastern Forests of Mazandaran Province

S. M. Hojjati*¹, M. Tafazoli², M. Asadiyan³ and A. Baluee⁴

1- Professor, Department of Forest science and engineering, Sari Agricultural Sciences and Natural Resources University, Sari, I. R. Iran. (s_m_hodjati@yahoo.com)

2- PhD in Forest Soil Science, Department of Forest science and engineering, Sari Agricultural Sciences and Natural Resources University, Sari, I. R. Iran. (mahya_tafazoli@yahoo.com)

3- PhD Student in Forestry, Department of Forest science and engineering, Sari Agricultural Sciences and Natural Resources University, Sari, I.R. Iran. (maryam.asadiyan23@gmail.com)

4- MSc Student in Forestry, Department of Forest science and engineering, Sari Agricultural Sciences and Natural Resources University, Sari, I. R. Iran. (ali.baluee@gmail.com)

Received: 01.08.2022 Accepted: 27.08.2022

Abstract

In this study, machine learning based models (linear regression, k-nearest neighbor, support vector regression, random forest) were evaluated to estimate carbon sequestration and soil respiration in the forests of eastern Mazandaran province. After identifying the sampling points, in each plot, the diameter and height of the trees were measured and the above-ground biomass of the trees was calculated using allometric models of Hyrcanian forest. Soil samples were taken from 0-20 cm depth and soil respiration was measured by CO₂-port. Soil respiration was estimated using bulk density, moisture content, texture, soil temperature, total nitrogen, available phosphorus and potassium, organic carbon and above-ground trees biomass. Soil carbon sequestration was estimated using soil temperature and moisture and above-ground tree biomass. Random forest model (RMSE = 10.47 and R² = 0.82) and support vector machine (RMSE = 0.77 and R² = 0.90) had the highest performance in estimating carbon sequestration and soil respiration, respectively. Soil moisture had the highest relative importance in estimating carbon sequestration (random forest model) and soil respiration (support vector machine model). According to the obtained results, it is possible to estimate the amount of carbon sequestration and soil respiration in the forest with appropriate accuracy by having the above-ground biomass of trees and basic soil characteristics.

Keywords: Climate change, Random Forest, Support vector regression, Forest soil properties.

* Corresponding author

Tel: +989117434910