

## Comparison of Different Data Mining Methods in Predicting Soil Organic Carbon Storage in Some Lands of Behbahan City

SAHEB KHORDEHBIN<sup>1\*</sup>, SAEID HOJATIA<sup>1</sup>, AHMAD LANDI<sup>1</sup>, IMAN AHMADIANFAR<sup>2</sup>

1. Department of Soil Science and Engineering, Faculty of Agriculture, Shahid Chamran University of Ahvaz, Ahvaz, Iran.

2. Department of civil Engineering, Faculty of Engineering, Behbahan Khatam Alanbia University of technology, Behbahan, Iran.

(Received: Sep. 14, 2019- Revised: Jan. 8, 2020- Accepted: Jan. 12, 2020)

### ABSTRACT

Soil organic carbon is an important factor in determining the global carbon cycle and global climate regulation. Soil is also the input/output source of carbon to the atmosphere which is depended on the land use. For this purpose, the objective of this study was to compare different methods of data mining in predicting soil organic carbon storage in irrigated, mixed cultivation (irrigated and rainfed), pasture and palm trees lands in some parts of Behbahan city in southwestern of Iran. Soil sampling from depths of 0-30 and 30-60 cm was carried out using conditional Latin hypercube square method. Organic carbon content of the soil samples was determined by Walky-Black method. Bulk density of the soils was determined using paraffin method. The auxiliary parameters used in this study included territory components, OLI sensor image data from landsat 8 and land use map. The results showed that the SAVI, NDVI, NDSI, salinity, carbonate, gypsum and clay indices have the highest correlation with the soil organic carbon stock values. The results also showed that the random forest (RF) ( $R^2= 0.983$ ,  $RMSE=2.32$ ) was the best model to predict soil organic carbon storage followed by artificial neural network model ( $R^2= 0.887$ ,  $RMSE= 4.257$ ) and Support Vector Regression Machine model (SVR) ( $R^2 = 0.707$ ,  $RMSE=7.344$ ).

**Keywords:** Artificial Neural Network, Carbon Store, Random Forest, Support Vector Machine.

## مقایسه روش‌های مختلف داده‌کاوی در پیش‌بینی ذخیره کربن آلی خاک در برخی اراضی شهرستان بهبهان

صاحب‌خورده بین<sup>۱\*</sup>، سعید حجتی<sup>۱</sup>، احمد لندی<sup>۱</sup>، ایمان احمدیان فر<sup>۲</sup>

۱. گروه علوم و مهندسی خاک، دانشکده کشاورزی، دانشگاه شهید چمران اهواز، اهواز، ایران.

۲. گروه عمران، دانشکده فنی و مهندسی، دانشگاه صنعتی خاتم الانبیاء بهبهان، بهبهان، ایران.

(تاریخ دریافت: ۱۳۹۸/۶/۲۳ - تاریخ بازنگری: ۱۳۹۸/۱۰/۱۸ - تاریخ تصویب: ۱۳۹۸/۱۰/۲۲)

### چکیده

ذخیره کربن نقش مهمی در تعیین چرخه جهانی کربن و تنظیم اقلیم جهانی ایفا می‌کند و از طرفی خاک منبع ورودی یا خروجی کربن به اتمسفر است، که به نوع کاربری اراضی وابسته است. از این رو، این مطالعه با هدف مقایسه روش‌های مختلف داده‌کاوی در پیش‌بینی ذخیره کربن آلی خاک در کاربری‌های کشاورزی آبی، کشت مخلوط (آبی-دیم)، مرتع و نخلستان موجود در بخشی از اراضی شهرستان بهبهان انجام گرفت. نمونه‌برداری از خاک به روش هایپرکیوب انجام و پس از تعیین موقعیت مکانی نقاط نمونه‌برداری، حفر پروفیل و برداشت نمونه‌های خاکی از عمق ۰-۳۰ و ۶۰-۳۰ سانتی‌متر انجام گرفت. کربن آلی نمونه‌های خاکی به روش والکی‌بلک و وزن مخصوص ظاهری آن‌ها به روش پارافین مذاب تعیین شد و سپس ذخیره کربن آلی خاک در نقاط نمونه‌برداری محاسبه گردید. پارامترهای کمکی مورد استفاده شامل اجزای سرزمین، داده‌های تصویر سنجنده OLI لندست ۸ و نقشه کاربری اراضی بود. نتایج نشان داد که شاخص‌های SAVI، NDVI، شوری، کربنات، گچ و رس بیشترین همبستگی را با مقادیر ذخیره کربن دارند. نتایج همچنین نشان داد که در همه کاربری‌ها، مدل جنگل تصادفی (RF) با بالاترین ضریب تبیین ( $R^2=0/966$ ) و کمترین مجذور میانگین مربعات خطا ( $RMSE=2/032$ ) بیشترین کارایی را در پیش‌بینی ذخیره کربن آلی خاک دارد و پس از آن مدل شبکه عصبی مصنوعی ( $RMSE=4/257$  و  $R^2=0/788$ ) و در نهایت مدل ماشین بردار پشتیبان (SVR) ( $RMSE=7/344$  و  $R^2=0/499$ ) قرار دارد.

واژه‌های کلیدی: بهبهان، جنگل تصادفی، شبکه عصبی مصنوعی، ذخیره کربن، ماشین بردار پشتیبان.

### مقدمه

ذخیره کربن نقش مهمی در تعیین چرخه جهانی کربن و تنظیم اقلیم جهانی ایفا می‌کند (Wang, 2019; Chen *et al.*, 2019) از طرفی افزایش ذخیره کربن آلی خاک از طریق مدیریت اراضی، به عنوان یک اقدام موثر در کاهش تغییرات اقلیمی پیشنهاد شده است (Chen *et al.*, 2019). ذخیره کربن آلی خاک در یک مکان با عوامل مختلف محیطی از قبیل اقلیم، پوشش گیاهی، پستی و بلندی، مواد مادری، بافت خاک و کاربری اراضی کنترل می‌شود (Jenny, 1994).

خاک به‌عنوان منبع ورودی یا خروجی کربن به اتمسفر است، که به نوع کاربری اراضی وابسته است (Lal, 2004) و داده‌های کاربری اراضی، اطلاعاتی که مربوط به تأثیرات انسان بر ذخیره کربن می‌باشد، نیز ارائه می‌نماید (Mishra, 2009). تلاش‌های زیادی برای تخمین ذخیره کربن آلی خاک در سطح جهانی و منطقه‌ای صورت گرفته است، که این برآوردها اغلب بر اساس

خصوصیات خاک و متغیرهای محیطی حاصل از مجموعه‌ای از داده‌های منطقه‌ای حاصل می‌شود. نمونه‌برداری مستقیم از خاک و اندازه‌گیری آزمایشگاهی به‌منظور اندازه‌گیری و ارزیابی کربن آلی خاک علاوه بر پرهزینه و وقت‌گیر بودن، دارای محدودیت‌هایی می‌باشد. از این رو، روش‌هایی نیاز است، که بتوان به‌سادگی سطح وسیعی از مناطق را مورد پایش<sup>۱</sup> قرار داده و روند تغییرات کربن آلی خاک را مورد بررسی قرار داد. یکی از این روش‌ها، استفاده از نقشه‌برداری رقومی خاک براساس ارتباط بین خصوصیات خاک و متغیرهای محیطی می‌باشد (McBratney *et al.*, 2003). در واقع نقشه‌برداری رقومی خاک یکی از ابزارهای کمی‌سازی رابطه بین مقادیر خصوصیات خاک و متغیرهای کمکی می‌باشد (Adhikari *et al.*, 2014).

تخمین‌های متفاوتی از ذخیره کربن در مقیاس جهانی و منطقه‌ای توسط محققین مختلف گزارش شده است که ناشی از ناهمگونی در خصوصیات خاک، پوشش گیاهی و توپوگرافی

\* نویسنده مسئول: research.beh88@gmail.com

پایین‌تر خاک (10 تا ۵۰ سانتی‌متری) تغییرات کربن آلی با کلاس‌های بافت خاک توجیه می‌شوند. Ließ *et al.* (2012) از روش جنگل تصادفی (RF) و درخت تصادفی<sup>۳</sup> (RT) برای پیش‌بینی توزیع بافت خاک استفاده نمودند و نتایج آن‌ها نشان داد، که برای پیش‌بینی مکانی بافت خاک، روش RF از دقت بالاتری نسبت به روش RT برخوردار است. Aitkenhead and Coull (2016) در اسکاتلند، برای پیش‌بینی مقدار ماده آلی و وزن مخصوص ظاهری در عمق‌های مختلف خاک از شبکه‌های عصبی مصنوعی استفاده نمودند و از متغیرهای اقلیمی، توپوگرافی، زمین‌شناسی، پوشش اراضی و نوع خاک به‌عنوان ورودی مدل استفاده نمودند. آن‌ها گزارش دادند، که ۷۹/۸ درصد تغییرات کربن و ۷۷/۹ درصد تغییرات وزن مخصوص ظاهری را می‌توان با این متغیرها پیش‌بینی کرد. تحقیقات زیادی در خصوص ذخیره کربن خاک در ایران و جهان صورت گرفته است، که در بیشتر مطالعات نقش متغیرهای محیطی در این خصوص کمتر لحاظ شده است. از این رو، در این پژوهش میزان ذخیره کربن آلی خاک در کاربری‌های موجود در بخشی از اراضی شهرستان بهبهان با مدل شبکه‌های عصبی مصنوعی، جنگل تصادفی و ماشین‌بردار پشتیبان بر اساس رابطه بین میزان ذخیره کربن با عوامل محیطی مورد بررسی قرار گرفته و توانایی مدل‌ها در پیش‌بینی ذخیره کربن مقایسه می‌شود.

## مواد و روش‌ها

### منطقه مورد مطالعه

منطقه مطالعاتی در ناحیه‌ای به وسعت حدود ۳۱۷۸ هکتار در ۲۰ کیلومتری شهرستان بهبهان به مختصات ۳۰ درجه و ۴۰ دقیقه عرض شمالی و ۵۰ درجه و ۳۰ دقیقه طول شرقی واقع شده است. منطقه دارای چهار کاربری مرتع، اراضی کشاورزی آبی، اراضی کشت مخلوط (کشاورزی آبی-دیم) و نخلستان می‌باشد. اراضی با کاربری کشاورزی آبی با ۱۱۹۴ هکتار، ۵۳/۱۷٪، اراضی با کاربری مرتع با ۱۹۹ هکتار، ۶/۲۸٪، اراضی با کاربری کشاورزی آبی-دیم با ۱۶۹۰ هکتار، ۳۷/۶۲٪ و اراضی با کاربری نخلستان با ۹۳ هکتار، ۲/۹۳٪ مساحت کل منطقه را شامل می‌شود. میانگین بارندگی سالانه منطقه، ۳۲۳ میلی‌متر، متوسط درجه حرارت سالانه، ۲۵ درجه سانتی‌گراد و میانگین تبخیر سالانه ۳۸۱۸ میلی‌متر می‌باشد. منطقه دارای رژیم رطوبتی و حرارتی به‌ترتیب یوستیک و هایپرترمیک می‌باشند (Amirian Chakan *et al.*, 2017).

می‌باشد؛ به‌طوری که میزان تخمین‌ها تا شش برابر توسط محققین مختلف گزارش شده است. به‌عنوان مثال، در مقیاس جهانی، Bohn (1976)، مقدار کل ذخیره کربن را ۳۰۰۰ پتاگرم<sup>۱</sup> برآورد کرد، در حالی که Eswaran *et al.* (1993) این مقدار را ۷۰۶ پتاگرم تخمین زده‌اند (Mingjun *et al.*, 2017). نوع کاربری اراضی به‌طور قابل ملاحظه‌ای بر مقدار و توزیع ذخیره کربن در عمق‌های مختلف تاثیرگذار است؛ به‌طوری که با افزایش عمق مقدار ذخیره کربن خاک کاهش می‌یابد (Zhang *et al.*, 2013). تغییر در نوع کاربری اراضی نیز با تغییر میزان کربن آلی ورودی به خاک، سبب تغییر در میزان ذخیره کربن خاک می‌شود (Zhang *et al.*, 2013).

مدل‌ها و روش‌های آماری مختلفی از جمله رگرسیون درختی (Akpa *et al.*, 2016; Wang *et al.*, 2017)، رگرسیون کریجینگ (Yigini and Panagos, 2016)، جنگل تصادفی<sup>۲</sup> (RF) (Akpa *et al.*, 2016)، روش‌های زمین‌آماری (Mitran *et al.*, 2018) و شبکه‌های عصبی مصنوعی (Aitkenhead and Coull, 2016) برای مدل کردن مقادیر ذخیره کربن آلی خاک توسعه و به‌کار گرفته شده است. Minasny *et al.* (2016) نقشه رقومی ذخیره کربن آلی خاک را با استفاده از مدل رگرسیون درختی، جنگل تصادفی<sup>۳</sup> و شبکه عصبی مصنوعی در دو منطقه در اندونزی تهیه کردند و ضرایب تبیین را به‌ترتیب ۰/۹۹-، ۰/۹۵-، ۰/۹۸- و ۰/۹۳- و ۰/۸۶-، ۰/۵۹- گزارش کردند. Somaratne *et al.* (2005) از مدل شبکه‌عصبی مصنوعی و رگرسیون خطی چندمتغیره برای مدل‌سازی مقدار کربن آلی خاک در شش نوع کاربری اراضی استفاده نمودند. آن‌ها از عوامل محیطی شامل ارتفاع، جهت شیب، میانگین سالانه درجه حرارت و میانگین سالانه بارندگی به‌عنوان پارامترهای ورودی شبکه برای مدل‌سازی مقدار کربن آلی خاک استفاده نموده و نشان دادند، که تخمین‌های ارائه شده توسط شبکه عصبی مصنوعی از صحت بالاتری نسبت به رگرسیون خطی چندمتغیره برای پیش‌بینی مقدار کربن آلی خاک برخوردار است. Grimm *et al.* (2008) نیز از روش درخت تصمیم برای پیش‌بینی کربن آلی خاک استفاده نمودند. نتایج آن‌ها نشان داد، که برآوردهای مقدار کربن آلی خاک با استفاده از مدل درخت تصمیم از دقت بالایی در پیش‌بینی کربن آلی خاک برخوردار است و تغییرات کربن آلی خاک در بخش بالایی خاک (۰ تا ۱۰ سانتی متری) به‌وسیله شاخص توپوگرافی ترکیبی<sup>۴</sup> (CTI) و فاکتور LS از متغیرهای محیطی توپوگرافی توجیه می‌شوند، ولی در لایه‌های

نمونه‌برداری و تجزیه‌های آزمایشگاهی

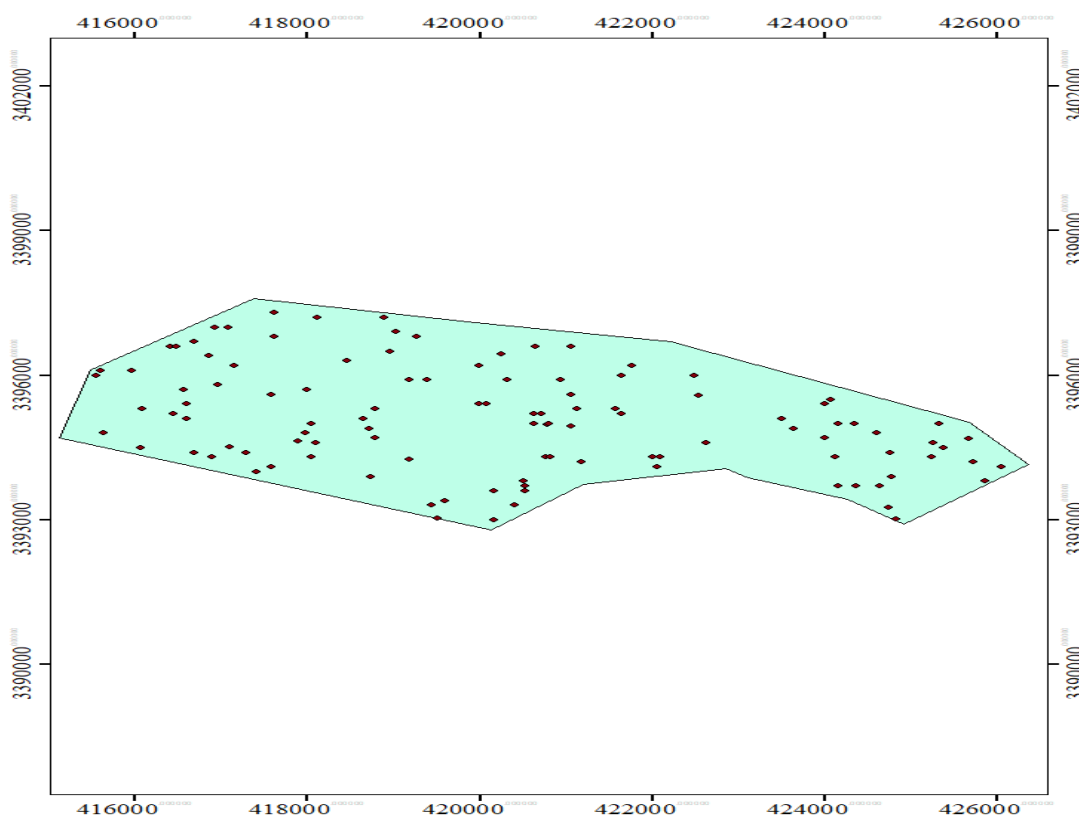
نظر کاربر تعیین گردید. در این پژوهش با روش هایپرکیوب در منطقه مورد مطالعه، موقعیت مکانی ۱۱۰ نقطه مطالعاتی مشخص گردید. در نقاط مورد نظر، نمونه‌برداری از عمق ۰-۳۰ و ۳۰-۶۰ سانتی‌متری صورت گرفت و پس از خشک کردن نمونه‌ها در هوا، کوبیدن و عبور دادن از الک ۲ میلی‌متری، کربن آلی خاک به روش والکی‌بلک اندازه‌گیری شد (Walkley and Black, 1934). جهت اندازه‌گیری وزن مخصوص ظاهری نیز یک کلوخه در هر نقطه نمونه‌برداری برداشت و وزن مخصوص ظاهری خاک به روش کلوخه و غوطه‌وری در پارافین مذاب (Blake and Hartge, 1986) محاسبه گردید. پس از محاسبه درصد کربن آلی خاک و وزن مخصوص ظاهری خاک، ذخیره کربن آلی در واحد سطح از رابطه (۱) محاسبه گردید (Gomes et al., 2019).

$$\text{SC} = 100 \times \text{OC} \times \text{Bd} \times \text{D} \times \text{ (رابطه ۱)}$$

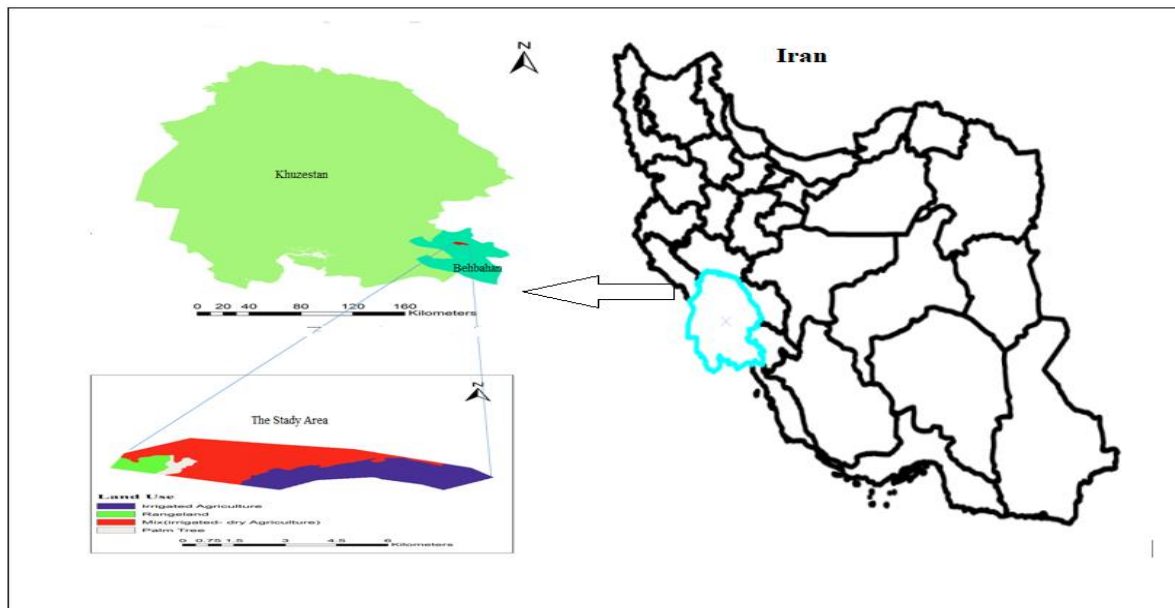
(درصد ذرات بزرگتر از ۲ میلی‌متر -۱)

که در این رابطه، SC، بیانگر ذخیره کربن آلی ( $\text{ton} \cdot \text{ha}^{-1}$ )، OC، درصد کربن آلی خاک، Bd، جرم مخصوص ظاهری خاک ( $\text{g} \cdot \text{cm}^{-3}$ ) و D عمق خاک (m) می‌باشد.

در این پژوهش از روش نمونه‌برداری هایپرکیوب<sup>۱</sup> (cLHS) که توسط Minasny and McBratney (2006) معرفی شده است برای انتخاب محل‌های نمونه‌برداری خاک استفاده گردید. مدل رقومی ارتفاع<sup>۲</sup>، داده‌های رقومی تصاویر ماهواره‌ای لندست ۸ و نقشه کاربری اراضی به‌عنوان ورودی‌های مدل هایپرکیوب استفاده گردید. مدل رقومی ارتفاعی به همراه باندهای ماهواره‌ای لندست ۸ پس از آماده‌سازی در نرم‌افزار ArcGIS نسخه ۱۰.۴ به نرم‌افزار SAGA GIS نسخه ۲.۳.۲ منتقل و پارامترهای توپوگرافی از مدل رقومی ارتفاع به شیوه ارائه شده توسط Hengl et al. (2004) استخراج گردید. پارامترهای شاخص رس، شاخص کربنات، شاخص پوشش گیاهی (NDVI)<sup>۳</sup> و شاخص پوشش گیاهی با تنظیم انعکاس زمین (SAVI) از باندهای ماهواره‌ای لندست در هر پیکسل از منطقه مورد مطالعه در نرم‌افزار JMP نسخه ۱۳ استخراج گردید. پس از استخراج عوامل محیطی، یک فایل با پسوند txt از آنها ایجاد و با استفاده از نرم‌افزار هایپرکیوب (cLHS)، مختصات نقاط نمونه‌برداری بر اساس تعداد نقاط مورد



شکل ۱- موقعیت مکانی نقاط مشخص شده به روش هایپرکیوب



شکل ۲- موقعیت منطقه مورد مطالعه

jmp نسخه ۱۳ استخراج شد (۲ تا ۶):

$$\text{Brightness Index} = \sqrt{B_4^2 + B_5^2} \quad (\text{رابطه ۲})$$

$$\text{Clay Index} = \frac{B_6}{B_7} \quad (\text{رابطه ۳})$$

$$\text{Salinity Index} = \frac{B_4 - B_5}{B_3 + B_5} \quad (\text{رابطه ۴})$$

$$\text{NDVI} = \frac{B_5 - B_4}{B_5 + B_4} \quad (\text{رابطه ۵})$$

$$\text{SAVI} = (1 + L) \frac{B_5 - B_4}{B_5 + B_{4+1}} \quad (\text{رابطه ۶})$$

#### آنالیز حساسیت<sup>۸</sup>

از آنجایی که برخی پارامترهای محیطی اگر به‌عنوان ورودی مدل انتخاب گردند، کارایی الگوریتم یادگیری را کاهش می‌دهند و در نتیجه باعث کاهش دقت مدل می‌شوند، برای انتخاب مهمترین پارامتر محیطی موثر بر ذخیره کربن از آنالیز حساسیت استفاده گردید. آنالیز حساسیت در نرم‌افزار Matlab نسخه ۲۰۱۷ انجام گرفت، که برای انجام آن، نرم‌افزار با ثابت نگه داشتن تمام پارامترها و تنها با تغییر در یکی از آن‌ها در هر مرحله اجرا گردید و میزان تغییر در پارامتر خروجی (ذخیره کربن) محاسبه شد. بر این اساس، در این مطالعه از میان ۲۴ پارامتر استخراج شده از مدل رقومی ارتفاع و شاخص‌های سنجش از دور، ۱۵ پارامتر به‌عنوان ورودی مدل انتخاب و بقیه پارامترها حذف شدند.

#### استخراج داده‌های کمکی توپوگرافی و سنجش از دور

بر اساس مدل اسکورپین، در نقشه برداری رقومی خاک علاوه بر داده‌های خاک، به داده‌های کمکی (داده‌های محیطی) نیز نیاز می‌باشد. در این مطالعه برای استخراج داده‌های کمکی علاوه بر نقشه کاربری اراضی از نقشه مدل رقومی ارتفاع و داده‌های لندست ۸ استفاده گردید. جهت تهیه نقشه مدل رقومی ارتفاع از نقشه موجود در سامانه نقشه برداری کشور با اندازه سلولی ۳۰ متر و برای تهیه داده‌های ماهواره‌ای از باندهای ماهواره لندست ۸ مربوط به ماه می سال ۲۰۱۷ استفاده گردید.

مدل رقومی ارتفاعی در نرم‌افزار ArcGIS نسخه ۱۰.۴ تهیه و سپس لایه مذکور به فرمت img تبدیل و به نرم‌افزار SAGA GIS نسخه ۲.۳.۲ منتقل و پارامترهای توپوگرافی از جمله شیب<sup>۱</sup>، جهت شیب<sup>۲</sup>، انحنای شیب<sup>۳</sup>، شیب حوزه<sup>۴</sup>، شاخص همواری دره با درجه تفکیک بالا<sup>۵</sup>، شاخص همواری قله با درجه تفکیک بالا<sup>۶</sup> و شاخص خیزی<sup>۷</sup> از آن استخراج گردید. پارامترهای شاخص روشنایی (Brightness Index)، شاخص کربنات (Carbonate Index)، شاخص رس (Clay Index)، شاخص شوری (Salinity Index)، پوشش گیاهی نرمال شده (NDVI) و شاخص پوشش گیاهی با تنظیم انعکاس زمین (SAVI)، که نشان‌دهنده ترکیبات زمین‌شناسی مواد سطحی می‌باشند، از ترکیب باندهای ماهواره لندست ۸ در نرم‌افزار

6. Multi-Resolution Ridge Flatness Index

7. Wetness Index

8. Sensitivity Analysis

1. Slope

2. Aspect

3. Curvature

4. Catchment Slope

5. Multi-Resolution Valley Bottom Flatness Index

مدل‌های به‌کار رفته جهت پیش‌بینی ذخیره کربن آلی خاک

کننده روش ماشین‌بردار پشتیبان (SVM) است.

مدل شبکه عصبی مصنوعی (Artificial Neural Network)

در این تحقیق برای ساخت شبکه عصبی مصنوعی از پرسپترون چندلایه (MLP) که مهمترین مدل شبکه عصبی مصنوعی می‌باشد (Tiryaki et al., 2014) با یک لایه مخفی با تابع فعال‌سازی سیگموئیدی در لایه مخفی و تابع فعال‌سازی خطی در لایه خروجی استفاده گردید. لایه‌های ورودی شبکه، دارای ۱۵ ورودی و لایه خروجی شامل یک نرون یعنی ذخیره کربن آلی در نظر گرفته شد. تعداد نرون‌های لایه مخفی و همچنین تعداد اپوک‌ها در فرآیند آموزش به‌صورت سعی و خطا به‌دست آمد، که بررسی‌های اولیه نشان داد، تعداد ۸ نرون در لایه مخفی و تعداد ۵۰۰ اپوک دارای بهترین ترکیب در پیش‌بینی ذخیره کربن آلی می‌باشد. ساخت شبکه عصبی مصنوعی و انجام پیش‌بینی ذخیره کربن آلی خاک در نرم‌افزار Matlab نسخه ۲۰۱۷ صورت گرفت.

جنگل تصادفی (Random Forest)

در الگوریتم جنگل تصادفی چندین درخت تصمیم وجود دارد و مجموعه‌ای از درخت‌های تصمیم با هم یک جنگل را تولید می‌کنند و هر درخت تصمیم این جنگل زیر مجموعه‌ای از داده‌ها را به خود اختصاص داده و الگوریتم یادگیری را روی آن‌ها انجام می‌دهد و هنگام پیش‌بینی، هر درخت تصمیم بر اساس الگوریتم‌های یادگرفته شده، یک پیش‌بینی نتایج را ارائه می‌دهند (Pouladi et al., 2019). جهت ساخت جنگل تصادفی و پیش‌بینی ذخیره کربن آلی خاک از کرنل پایه شعاعی نرم‌افزار Weka نسخه ۳.۹ استفاده گردید. با اجرای مدل و تفسیر میانگین مربعات خطا در داده‌های آموزش و آزمون، جایی که با افزایش تعداد درختان تصمیم، تغییری در کاهش میانگین مربعات خطا مشاهده نگردد، به‌عنوان تعداد درخت بهینه انتخاب می‌شود (Shataee et al., 2012).

ماشین بردار پشتیبان (SVR)

SVM، یک مدل یادگیری نظارت شده است و از پرکاربردترین الگوریتم‌های ماشین‌بردار پشتیبان می‌باشد. الگوریتم ماشین‌بردار پشتیبان یک روش دسته‌بندی است، که توسط Vapnik et al. (1963) ابداع شد و در سال ۱۹۹۵ برای حالت غیرخطی تعمیم داده شد (Cortes and Vapnik, 1995). در این تحقیق از نرم‌افزار Weka نسخه ۳.۹ جهت اجرای الگوریتم استفاده گردید. لازم به توضیح است، که الگوریتم SOM در نرم‌افزار Weka پیاده‌سازی

ارزیابی کارایی مدل‌ها در برآورد ذخیره کربن

از آماره‌های مختلفی می‌توان برای مقایسه کارایی مدل‌های مختلف در برآورد ذخیره کربن استفاده کرد. برای این منظور ابتدا داده‌ها به دو دسته آموزش (۷۰ درصد داده‌های نمونه‌برداری شده) و آزمون (۳۰ درصد داده‌های نمونه‌برداری شده) به‌طور تصادفی تقسیم گردید و سپس از آماره‌های ضریب همبستگی (R)، ضریب تبیین (R-Square)، مجذور میانگین مربعات خطا (RMSE)، قدر مطلق مربعات خطا (MAE)، درصد میانگین خطا (MAPE)، خطای مطلق نسبی (RAE) و مجذور نسبی خطا (RSE) جهت بررسی مقایسه دقت مدل‌های مختلف در پیش‌بینی ذخیره کربن استفاده گردید (روابط ۷ تا ۱۲) و (جدول ۳).

$$R^2 = \frac{\sum_{i=1}^n (T_i - Y_i)^2}{\sum_{i=1}^n \bar{Y} - Y_i} \quad \text{(رابطه ۷)}$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (T_i - Y_i)^2}{n}} \quad \text{(رابطه ۸)}$$

$$MAE = \frac{1}{n} \sum_{i=1}^n |T_i - Y_i| \quad \text{(رابطه ۹)}$$

$$SRE = \frac{1}{n} \sum_{i=1}^n \left( \frac{Y_i - T_i}{Y_i} \right)^2 \quad \text{(رابطه ۱۰)}$$

$$RAE = \frac{\sum_{i=1}^n |Y_i - T_i|}{\sum_{i=1}^n |\bar{T} - T_i|} \quad \text{(رابطه ۱۱)}$$

$$RSE = \frac{\sum_{i=1}^n (Y_i - T_i)}{\sum_{i=1}^n (\bar{T} - T_i)} \quad \text{(رابطه ۱۲)}$$

در روابط فوق،  $T_i$  مقدار واقعی (هدف)،  $Y_i$  مقدار پیش‌بینی شده،  $\bar{Y}$  میانگین مقادیر پیش‌بینی شده و  $\bar{T}$  میانگین مقادیر واقعی می‌باشد.

تجزیه و تحلیل آماری

ابتدا نرمال بودن داده‌ها با استفاده از آزمون Kolmogorov-Smirnov بررسی شد، که نتایج این آزمون حاکی از نرمال بودن داده‌ها بود. همچنین همگن بودن داده‌ها با استفاده از آزمون لون بررسی و مورد تایید قرار گرفت. به‌منظور مقایسه میزان ذخیره کربن خاک در کاربری‌های گوناگون از آنالیز واریانس یک طرفه (ANOVA) و جهت مقایسه میانگین ذخیره کربن در کاربری‌های مختلف از آزمون دانکن استفاده گردید.

نتایج و بحث

نتایج تحلیل آماری

نتایج توصیف آماری ذخیره کربن در جدول (۱) آورده شده است. آزمون‌های آماری در نرم‌افزار SPSS نسخه شماره ۲۴ انجام گرفت. برای بررسی ارتباط بین عوامل محیطی با ذخیره کربن، ضریب همبستگی بین آنها در نرم‌افزار Matlab انجام گرفت، که نتایج آن در جدول شماره (۲) آورده شده است. نتایج این جدول نشان

گردید و مقادیر پیش‌بینی شده برای داده‌های تست و آموزشی این مدل‌ها تعیین گردید. رابطه بین مقادیر واقعی و پیش‌بینی شده دو مدل جنگل تصادفی و ماشین‌بردار پشتیبان در شکل‌های (۴) و (۵) نشان داده شده است.

خصوصیات فیزیکی و شیمیایی خاک در نقاط نمونه‌برداری، که شامل درصد ذرات خاک (سیلت، رس و شن)، درصد آهک، هدایت الکتریکی، مقدار ماده آلی، واکنش خاک و درصد رطوبت اشباعی می‌باشد برای عمق ۰-۳۰ سانتی متر و ۳۰-۶۰ سانتی متر تحلیل شدند. آمار توصیفی این خصوصیات و ضریب همبستگی بین پارامترهای خاک و مقادیر ذخیره کربن در جدول (۴) و (۵) نشان داده شده است. نتایج نشان داد، که درصد آهک ( $p < 0.05$ ) و هدایت الکتریکی ( $p < 0.01$ ) در عمق ۰-۳۰ و ۳۰-۶۰ و مقدار کربن آلی خاک ( $p < 0.01$ ) در هر دو عمق ۰-۳۰ و ۳۰-۶۰ با مقدار ذخیره کربن ارتباط معنی‌داری دارند.

#### مقایسه عملکرد روش‌های مختلف پیش‌بینی ذخیره کربن آلی خاک

مقایسه آماره‌های ضریب تبیین (R-Square)، مجذور میانگین مربعات خطا (RMSE) در مدل‌های مختلف (جدول ۳) نشان می‌دهد، که برای هر دو عمق ۰-۳۰ و ۳۰-۶۰ سانتی‌متر، مدل جنگل تصادفی نسبت به دو مدل شبکه عصبی مصنوعی و ماشین‌بردار پشتیبان دارای ضریب تبیین بیشتر (۰/۹۶۶) باشد. ( $R^2 = 0.929$ ) و مجذور میانگین مربعات خطا کمتری می‌باشد. ( $RMSE = 2/538, RMSE = 2/32$ ) که بیانگر کارایی بیشتر مدل جنگل تصادفی در پیش‌بینی ذخیره کربن آلی خاک می‌باشد. پس از آن به ترتیب مدل شبکه عصبی مصنوعی ( $R^2 = 0.788, R^2 = 4/257$ ) و مدل ماشین‌بردار پشتیبان (SVR) ( $R^2 = 0.499$ ) قرار دارند. مقایسه پارامترهای قدر مطلق مربعات خطا (MAE)، درصد میانگین خطا (MAPE) و خطای مطلق نسبی (RAE) (جدول ۳) همچنین نشان می‌دهد، که مقادیر این آماره‌ها در مدل جنگل تصادفی نسبت به دیگر مدل‌ها کمتر می‌باشد، که نشان‌دهنده مقدار خطای کمتر بین مقادیر مشاهده‌ای و پیش‌بینی شده ذخیره کربن آلی خاک در این مدل می‌باشد.

#### ذخیره کربن در کاربری‌های مختلف

نتایج جدول (۱) نشان می‌دهد که میانگین ذخیره کربن در عمق ۰-۳۰ سانتی‌متری خاک نسبت به عمق ۳۰-۶۰ سانتی‌متری بیشتر می‌باشد و این مقدار برای عمق ۰-۳۰ سانتی‌متری بین ( $ton/ha^{-1} = 77/84 - 23/09$ ) برای کاربری مرتع و کشاورزی آبی-دیم متغیر می‌باشد. به‌طوری که کاربری مرتع دارای کم‌ترین

می‌دهد، که این همبستگی در عمق ۰-۳۰ در مقایسه با عمق ۳۰-۶۰ بیشتر بوده ولی در هر دو عمق مذکور شاخص‌های SAVI, EVI و شاخص رس بیشترین همبستگی مثبت و شاخص‌های شوری خاک، کربنات و گچ، بیشترین همبستگی منفی را با ذخیره کربن دارند.

از آنجایی که واریانس ذخیره کربن در کاربری‌های مختلف با همدیگر اختلاف معنی‌داری نداشتند، شرط همگنی واریانس‌ها در کاربری‌های مختلف مورد تایید قرار گرفت. جهت مقایسه میانگین ذخیره کربن در بین کاربری‌های مختلف از تحلیل واریانس یک طرفه و آزمون دانکن استفاده گردید که سطح معنی‌داری (p-value) آزمون ANOVA نشان داد، که ذخیره کربن در کاربری‌های مختلف با همدیگر اختلاف معنی‌داری داشته که وجود این اختلاف بین کاربری مرتع و کاربری نخلستان بیشتر بود. نتایج همچنین نشان داد که متوسط ذخیره کربن آلی خاک در کاربری نخلستان بیشتر از دیگر کاربری‌ها بوده و پس از آن به ترتیب کاربری کشت مخلوط (کشاورزی آبی-دیم)، کشاورزی آبی و مرتع قرار دارند.

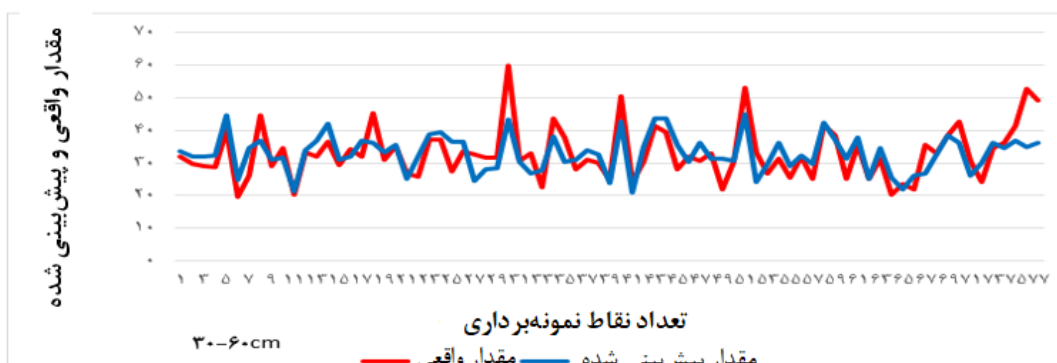
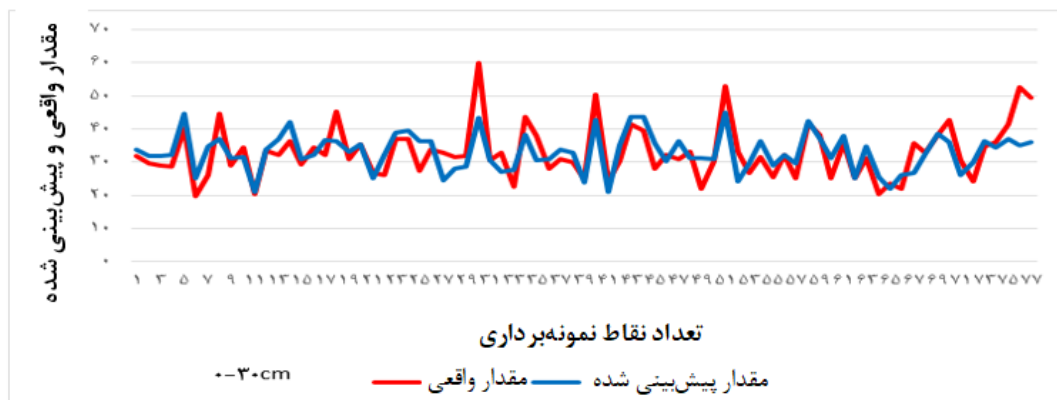
جهت مقایسه عملکرد و کارایی مدل‌های مختلف در پیش‌بینی ذخیره کربن، آماره‌ها در نرم‌افزار Matlab محاسبه گردید که نتایج آن در جدول شماره (۳) آورده شده است. نمودار رگرسیونی مقادیر پیش‌بینی شده و واقعی ذخیره کربن در مدل‌های مختلف در نرم‌افزار Excel انجام شد و نتایج آن در شکل (۶) نشان داده شده است.

مقایسه مقادیر ضریب تبیین، میانگین مربعات خطا و قدر مطلق مربعات خطا و دیگر پارامترهای محاسبه شده در جدول (۳) نشان داد، که در همه کاربری‌ها برای عمق ۰-۳۰ سانتی‌متر مدل جنگل تصادفی (RF) بیشترین کارایی ( $R^2 = 0.983$ ) را در پیش‌بینی ذخیره کربن آلی خاک داشته و پس از آن مدل شبکه عصبی مصنوعی ( $R^2 = 0.887$ ) و در نهایت مدل ماشین‌بردار پشتیبان (SVR) ( $R^2 = 0.707$ ) قرار دارد. همچنین آماره‌های جدول (۳) نشان می‌دهد، که برای عمق ۳۰-۶۰ سانتی‌متر نیز مدل جنگل تصادفی نسبت به بقیه مدل‌ها کارایی بیشتری در پیش‌بینی ذخیره کربن دارد.

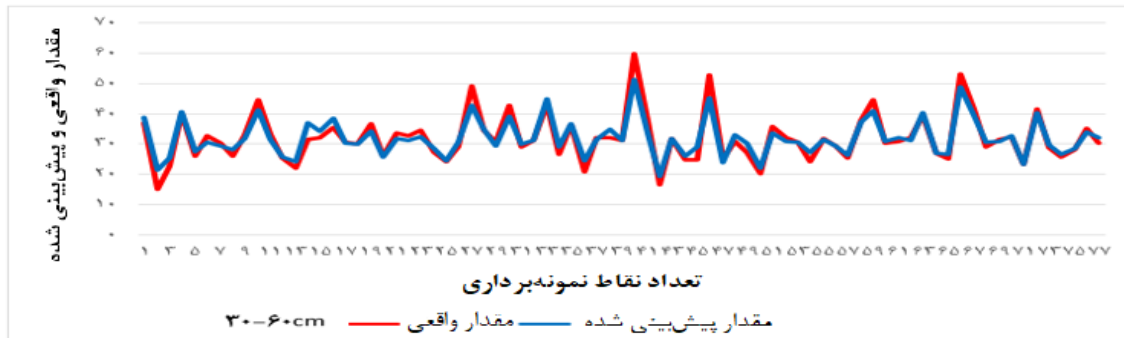
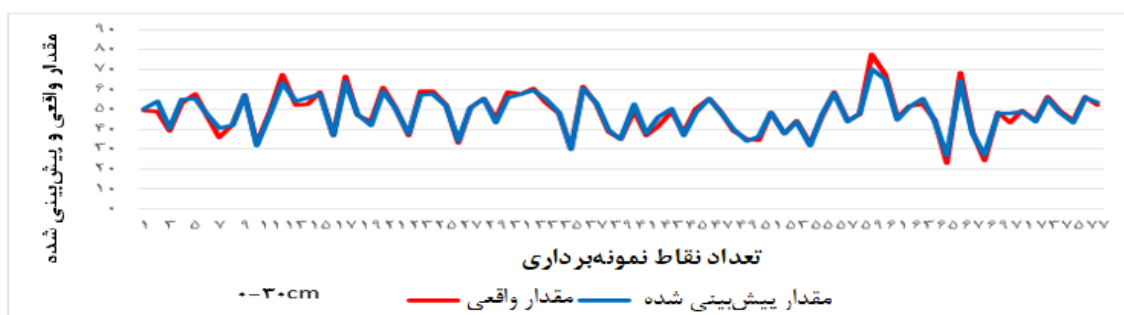
پس از آماده‌سازی فایل‌ها و انجام پیش‌پردازش بر روی داده‌ها و تقسیم داده‌ها به دو بخش آموزش و آزمون، مدل شبکه عصبی در نرم‌افزار Matlab اجرا گردید. سپس برای داده‌های آموزشی و آزمون مقادیر برآورد شده تعیین گردید. رابطه بین مقادیر واقعی و مقادیر پیش‌بینی شده برای داده‌های مذکور در مدل شبکه عصبی مصنوعی در شکل (۳) نشان داده شده است. همچنین مدل جنگل تصادفی (RF) ماشین‌بردار پشتیبان (SVR) در نرم‌افزار وکا (Weka) اجرا

ذخیره کربن پیش‌بینی شده و واقعی مدل‌ها در شکل (۶) نشان داده شده است. نتایج نشان داد، که در سه مدل جنگل تصادفی، شبکه عصبی مصنوعی و ماشین بردار پشتیبان، میزان خطا بین مقدار واقعی و پیش‌بینی ذخیره کربن در مدل جنگل تصادفی از دیگر مدل‌ها کمتر می‌باشد.

ذخیره کربن و کاربری کشاورزی آبی-دیم دارای بیش‌ترین ذخیره کربن می‌باشد، ولی متوسط ذخیره کربن در هر دو عمق ۰-۳۰ و ۳۰-۶۰ سانتی‌متری در کاربری نخلستان از بقیه کاربری‌ها بیشتر می‌باشد. برای عمق ۰-۳۰ سانتی‌متری خاک، حداکثر و حداقل میزان ذخیره کربن به ترتیب ۵۹/۸ و ۳۰/۳ ( $tonha^{-1}$ ) برای کاربری نخلستان و کاربری مرتع مشاهده گردید. مقایسه مقادیر

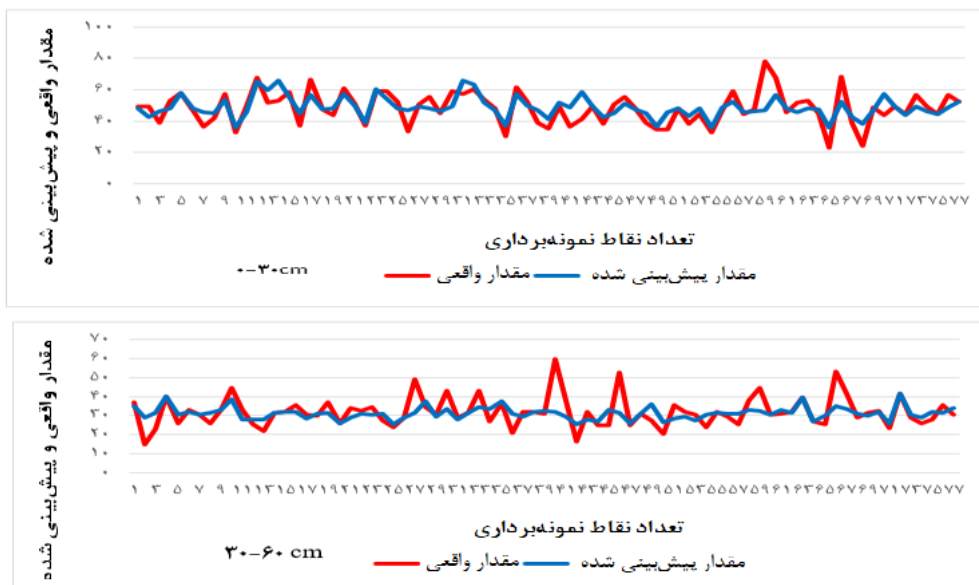


شکل ۳- نمودار مقادیر واقعی و پیش‌بینی شده ذخیره کربن داده‌های آموزشی در مدل شبکه عصبی مصنوعی



شکل ۴- نمودار بین مقادیر واقعی و پیش‌بینی شده داده‌های آموزشی ذخیره کربن در مدل جنگل تصادفی





شکل ۵- نمودار بین مقادیر واقعی و پیش‌بینی شده داده‌های آموزشی ذخیره کربن در مدل ماشین‌بردار پشتیبان

جدول ۱- توصیف آماری داده‌های ذخیره کربن

عمق نمونه‌برداری (cm)	کاربری اراضی	تعداد	میانگین	انحراف استاندارد	خطای استاندارد	کمینه	بیشینه
۰-۳۰	کشاورزی آبی	۴۵	۴۷/۹۱۷	۸/۲۰۶۷	۱/۲۲۳۴	۳۳/۱	۶۷/۶
	کشاورزی آبی - دیم	۵۵	۴۷/۶۹۲	۱۰/۸۹۷۶	۱/۴۶۹۴	۲۵/۴	۷۷/۸
	نخلستان	۳	۵۹/۰۱۸	۶/۶۷۸۶	۳/۸۵۵۹	۵۳/۵	۶۶/۴
	مرتع	۷	۳۰/۷۹۶	۸/۹۳۳۲	۳/۳۷۶۳	۲۳/۱	۴۹
۳۰-۶۰	کشاورزی آبی	۴۵	۳۱/۱۰۵	۶/۶۹۳۶	۱	۱۵/۱	۵۲/۶
	کشاورزی آبی - دیم	۵۵	۳۳/۹۱۶	۷/۵۲۳۴	۱/۰۱۴۵	۱۶/۵	۵۳
	نخلستان	۳	۵۱/۶۲۴	۷/۴۴۳۳	۴/۲۹۷۴	۴۵/۲	۵۹/۸
	مرتع	۷	۲۴/۲۴۴	۳/۶۸۰۵	۱/۳۹۱۱	۱۹/۷	۳۰/۳

\*ذخیره کربن بر حسب  $(tonha^{-1})$  می‌باشد.

جدول ۲. ضریب همبستگی بین عوامل محیطی با ذخیره کربن

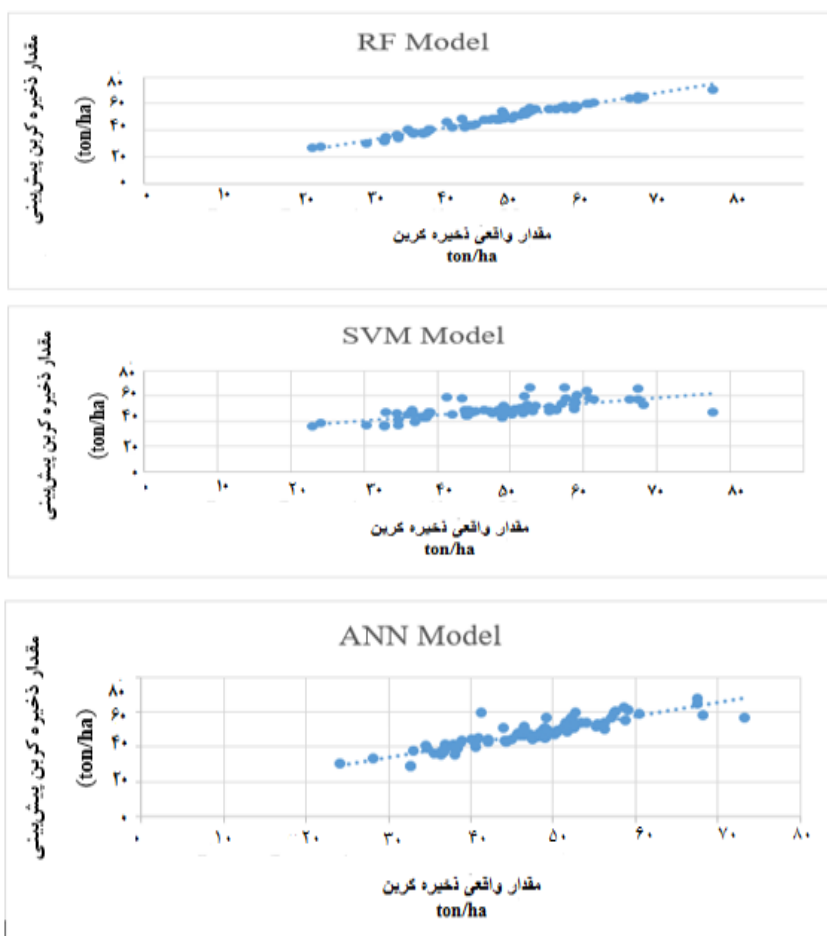
عوامل محیطی	ضریب همبستگی (عمق ۰-۳۰ سانتی‌متری)	ضریب همبستگی (عمق ۳۰-۶۰ سانتی‌متری)
EVI	۰/۶۵	۰/۵۳
SAVI	۰/۶۹	۰/۵۶
NDSI	-۰/۶۹	-۰/۵۶
Carbonate Index	-۰/۶۳	-۰/۵۲
Clay Index	۰/۵۲	۰/۳۵
NDVI	۰/۶۹	۰/۵۷
Gypsum Index	-۰/۷۱	-۰/۵۳
Aspect	۰/۲۱	-
B1	-۰/۶۹	-۰/۴۳
B2	-۰/۶۳	-۰/۴۵
B3	-۰/۶۴	۰/۳۱
B4	-۰/۶۷	-۰/۵۳
B5	۰/۳۵	-۰/۴۶
B6	-۰/۶۹	-۰/۵۰
b7	-۰/۶۳	-۰/۴۲

b1 تا b7، باندهای ۱ تا ۷ ماهواره لندست ۸، EVI، شاخص پوشش گیاهی بازسازی شده، NDSI، شاخص شوری خاک، SAVI، شاخص پوشش گیاهی با تنظیم انعکاس زمین و NDVI، شاخص پوشش گیاهی نرمال شده می‌باشد.

جدول ۳- آماره‌های مقایسه عملکرد روش‌های مختلف در پیش‌بینی ذخیره کربن

آماره‌های مقایسه عملکرد							عمق (cm)	داده	مدل	
MAPE	MSRE	RAE	RSE	MAE	RMSE	R-Square	r	(Depth)	(Data)	(Model)
۳/۱۵۹۱۹	۰/۰۰۱۹۶	۰/۱۸۳	۰/۰۳۹	۱/۴۷	۲/۰۳۲	۰/۹۶۶	۰/۹۸۳	۰-۳۰	Train	RF
۶/۹۳۳۳۸	۰/۰۱۰۹۵	۰/۳۷۳	۰/۲۲۰	۳/۰۳۷	۴/۹۶۲	۰/۷۸۴	۰/۸۸۵	۰-۳۰	Test	RF
۱۱/۹۲۲۲	۰/۰۲۹	۰/۶۶۱	۰/۵۰۶	۵/۳۲۷	۷/۳۴۴	۰/۴۹۹	۰/۷۰۷	۰-۳۰	Train	SVR
۱۳/۹۱۵۲	۰/۰۳۴۲	۰/۶۶۷	۰/۳۷۸	۵/۴۱۷	۶/۵۰۳	۰/۷۶۱	۰/۸۷۲	۰-۳۰	Test	SVR
۶/۲۸۸۶	۰/۰۰۷	۰/۴۵۹	۰/۲۷۰	۲/۹۵۱	۴/۲۵۷	۰/۷۸۸	۰/۸۸۷	۰-۳۰	Train	ANN
۱۰/۳۲۳	۰/۰۲۰	۰/۵۹۴	۰/۴۹۶	۴/۷۵۸	۶/۸۹۰	۰/۷۳۷	۰/۸۵۸۲۵	۰-۳۰	Test	ANN
۵/۹۱۰۵	۰/۰۰۷۲	۰/۳۲۵	۰/۱۰۴	۱/۸۳۴	۲/۵۳۸	۰/۹۲۹	۰/۹۶۴	۳۰-۶۰	Train	RF
۱۵/۸۱۴۸	۰/۰۳۷۹	۰/۷۸۲	۰/۵۷۳	۵/۰۶۷	۶/۰۴۷	۰/۴۴۱	۰/۶۶۴	۳۰-۶۰	Test	RF
۱۴/۱۸۳۴	۰/۰۴۵	۰/۷۹۰	۰/۷۴۵	۴/۴۵۵	۶/۷۸۹۹	۰/۲۷۶	۰/۵۲۵	۳۰-۶۰	Train	SVR
۱۵/۰۶۶۶	۰/۰۳۵	۰/۷۸۲	۰/۶۳۴	۵/۰۶۳	۶/۳۶۳	۰/۳۳۹	۰/۶۳۱	۳۰-۶۰	Test	SVR
۱۳/۱۶۴	۰/۰۲۶	۰/۷۳۳	۰/۴۹۲	۴/۳۴۰	۵/۵۵۸	۰/۵۰۸	۰/۷۱۳	۳۰-۶۰	Train	ANN
۱۷/۷۰۶	۰/۰۴۷	۱/۱۶۰	۱/۲۶۳	۵/۴۰۹	۶/۵۴۱	۰/۳۵۵	۰/۵۹۶	۳۰-۶۰	Test	ANN

r، ضریب همبستگی، R-Square، ضریب تبیین، RMSE، مجذور میانگین مربعات خطا، MAE، قدرمطلق مربعات خطا، RSE، مجذور نسبی خطا، RAE، خطای مطلق نسبی، MAPE، درصد میانگین خطا، MSER، میانگین مربع نسبی خطا



شکل ۶- نمودار بین مقدار واقعی و پیش‌بینی شده ذخیره کربن در مدل‌های مختل

۱- Mean Absolute Error  
 ۲- Relative Squared Error  
 ۳- Relative Absolute Error  
 ۴- Mean Squared Relative Error

جدول ۴- توصیف آماری خصوصیات خاک

عمق	خصوصیات خاک	واحد	حداقل	حداکثر	میانگین	انحراف معیار	چولگی	کشیدگی
۰-۳۰ (cm)	SOC	gCkg-1	۴/۷	۱۷/۱۳	۹/۸	۰/۲۵۲	۰/۲۹۸	-۰/۱۸۹
	Caco3	%	۳۹/۷	۴۹/۳۷	۴۸/۳۹	۱/۴۱	-۴/۴۷	۲۱/۵۸
	Ec	Ms/cm	۰/۲۹	۷/۶	۲/۱	۱/۳۳	۱/۵۷	۲/۹۳
	pH	Log[H <sup>+</sup> ]	۶/۷	۷/۸۲	۷/۳۴	۰/۲۶	-۰/۳۶	-۰/۶۸
	clay	%	۱۰	۶۱/۵	۳۲/۸	۱۱/۱	-۰/۰۱۳	-۰/۶۸
	Silt	%	۸	۵۲	۳۹/۵	۸/۴۱	۱/۹۶	۴/۰۷
	sand	%	۸/۵	۶۶	۲۷/۶	۱۲/۲۸	۰/۷۸	۰/۰۷۸
۳۰-۶۰ (cm)	SP	%	۲۳/۷	۵۱	۴۰/۴	۴/۰۶	-۰/۶۴	۲/۱
	SOC	gCkg-1	۳,۱۹	۱۱,۱	۶,۷	۰/۱۷	۰/۶۶	۰/۱۹
	Caco3	%	۳۳	۴۹/۴۲	۴۸/۲۸	۱/۸۹	-۵/۹۶	۴۵/۴
	Ec	Ms/cm	۰/۲۹	۷	۱/۹۴	۱/۱۴	۱/۶۳	۳/۸۸
	pH	Log[H <sup>+</sup> ]	۶/۹	۷/۹۲	۷/۴۵	۰/۲۳	-۱/۷۳	۳/۴۱
	clay	%	۸	۶۲	۳۵/۶	۱۱/۸	-۰/۲۹	-۰/۱۰۴
	Silt	%	۶	۵۴/۵	۳۸/۱۱	۹	-۱/۷۳	۳/۴۱
	sand	%	۱۰	۸۶	۲۶/۲۴	۱۳/۷۶	۱/۶۳	۳/۳۹
	SP	%	۱۷/۵	۵۴	۴۰/۱۱	۵/۰۳	-۰/۶۲	۴/۲۵

جدول ۵- ضریب همبستگی بین ذخیره کربن و برخی خصوصیات خاک

خصوصیات خاک	عمق (cm)	آهک %	کربن آلی %	هدایت الکتریکی (Ms/cm)	رطوبت اشباعی	واکنش خاک	وزن مخصوص ظاهری (g/cm <sup>3</sup> )	رس %	سیلت %	شن %
ضریب همبستگی	۰-۳۰	-۰/۲*	۰/۹۳۳**	-۰/۱۶۹	-۰/۰۰۳	۰/۰۰۵	-۰/۰۰۶	۰/۱۰۸	-۰/۰۱۰	۰/۱۰۸
	۳۰-۶۰	-۰/۱۰۱	۰/۹۱**	-۰/۳۹۷**	-۰/۰۹۸	-۰/۰۸۶	۰/۱۰۳	-۰/۰۸۱	۰/۱۱۸	-۰/۰۰۷

## بحث و نتیجه‌گیری

پوشش گیاهی به‌خصوص در مناطق خشک و نیمه‌خشک دارند. مقایسه نتایج ذخیره کربن در کاربری‌های مختلف منطقه مورد مطالعه نشان داد که الگوی میزان ذخیره کربن در کاربری مرتع و کشاورزی با الگوی گزارش شده توسط دیگر محققین متفاوت می‌باشد. به‌عنوان مثال *Padilla et al.* (2010) با مطالعه بر روی مراتع نوادا، تغییر کاربری اراضی مرتعی را مهمترین عامل کاهش میزان ذخیره کربن مراتع منطقه دانسته‌اند. *and Tiessen Stewart* (1983) نیز بیان داشتند که تغییر کاربری اراضی مرتعی باعث کاهش ماده آلی و در نتیجه کاهش ذخیره کربن خاک می‌شود. از آنجا که میزان ماده آلی خاک در هر منطقه توسط اقلیم (دما و رطوبت) و عوامل بیولوژیکی (از جمله برگرداندن بقایای گیاهی) و ویژگی‌های فیزیکی و شیمیایی خاک کنترل می‌شود، در مناطق خشک، به‌دلیل کمبود برگشت بقایای گیاهی و دمای بالا، میزان کربن آلی خاک، بیشتر تحت تاثیر ویژگی‌های خاک قرار می‌گیرند و چون کربن آلی خاک سریع‌تر و بیشتر از هر ویژگی دیگر خاک، تحت تاثیر اقدامات مدیریتی قرار می‌گیرد،

آگاهی از وضعیت ذخیره کربن آلی خاک و تاثیر کاربری‌های مختلف در ذخیره کربن آلی خاک، در راستای کاهش تغییرات اقلیمی ناشی از گازهای گلخانه‌ای امری لازم و ضروری می‌باشد. از این رو، این مطالعه جهت پیش‌بینی میزان ذخیره کربن آلی خاک در کاربری‌های مختلف با مدل‌های شبکه عصبی مصنوعی، جنگل تصادفی و ماشین‌بردار پشتیبان و مقایسه کارایی و دقت این مدل‌ها انجام گرفت. مقادیر ذخیره کربن آلی خاک در کاربری‌های مختلف نشان داد، که متوسط ذخیره کربن آلی خاک در کاربری نخلستان بیشتر از بقیه کاربری‌ها بوده و پس از آن به ترتیب کاربری کشت مخلوط (کشاورزی آبی-دیم)، کشاورزی آبی و مرتع قرار دارد. دلیل بیشتر بودن مقدار ذخیره کربن در کاربری نخلستان را می‌توان به پوشش گیاهی بیشتر و در نتیجه وجود بقایای گیاهی بیشتر و عدم آتش‌سوزی بقایای گیاهی و چرای توسط دام نسبت داد زیرا بقایای گیاهی با کاهش تبخیر از سطح خاک، افزایش رطوبت برای رشد گیاهان، اثر مثبتی بر

یکسان در نظر گرفته شده تا امکان مقایسه آن‌ها فراهم گردد. انتخاب بهترین روش برآورد ذخیره کربن آلی، بر اساس بیشترین مقدار  $R^2$  و کمترین مقدار RMS انجام شد. میزان مقادیر  $R^2$  در مدل‌های مختلف مذکور به ترتیب  $(R^2=0/966)$ ،  $(R^2=0/788)$  و  $(R^2=0/499)$  به دست آمد که نشان می‌دهد در مدل جنگل تصادفی، از ۱۰۰ درصد تغییرات مربوط به ذخیره کربن، ۹۶/۶ درصد تغییرات مربوط به تغییرات عواملی می‌باشد که در انتخاب ورودی‌ها به مدل لحاظ شده است و تنها ۳/۴ درصد تغییرات کربن مربوط به عواملی می‌باشد که در ورودی به مدل لحاظ نشده است. از آنجا که کربن آلی خاک یکی از مهم‌ترین خصوصیات خاک است، تلاش‌های زیادی صورت گرفته است تا دقت تخمین‌های ماده آلی خاک و در نتیجه میزان ذخیره کربن، با استفاده از تکنیک‌های داده‌کاوی بهبود یابد. این مقاله نیز به بررسی تغییرات ذخیره کربن خاک در دو عمق ۰-۳۰ و ۳۰-۶۰ سانتی‌متری خاک، با استفاده از عوامل محیطی کمکی با سه مدل جنگل تصادفی، شبکه عصبی مصنوعی و ماشین‌بردار پشتیبان در کاربری‌های موجود در بخشی از شهرستان بهبهان پرداخته شده است. نتایج نشان داد، که شاخص‌های SAVI، NDVI، شاخص شوری، شاخص کربنات، شاخص گچ و رس بیشترین همبستگی را با مقادیر ذخیره کربن دارند. Minasny *et al.* (2013) نیز از شاخص‌های SAVI و NDVI جهت پیش‌بینی ماده آلی خاک استفاده کرده و همبستگی خوبی را بین این شاخص‌ها و مقادیر پیش‌بینی شده اعلام نمودند. بررسی دقت و کارایی سه مدل مختلف شبکه‌های عصبی مصنوعی، جنگل تصادفی و ماشین‌بردار پشتیبان در پیش‌بینی میزان ذخیره کربن نشان داد، که در همه کاربری‌ها، مدل جنگل تصادفی نتیجه بهتری دارد. نتایج تحقیقات Gomes *et al.* (2019) در کشور برزیل نیز جهت بررسی میزان ذخیره کربن آلی خاک با مدل‌های جنگل تصادفی، کیوبیست، مدل خطی تعمیم یافته و ماشین‌بردار 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 پیش‌بینی ذخیره کربن آلی خاک دارد. نتایج تحقیقات Kuang *et al.* (2015)، نیز کارایی خوب شبکه‌های عصبی را در برآورد کربن آلی خاک نشان داد. نتایج تحقیقات Zhang *et al.* (2017)، در خصوص پیش‌بینی کربن آلی خاک به دو روش جنگل تصادفی و رگرسیون خطی نیز برتری روش جنگل تصادفی را در پیش‌بینی کربن آلی را نسبت به روش رگرسیون نشان داد و اعلام نمودند، که برتری مدل جنگل تصادفی نسبت به مدل رگرسیون خطی در پیش‌بینی کربن آلی خاک مربوط به روابط غیرخطی است، که بین کربن آلی خاک و متغیرهای

هر گونه تغییر در مدیریت اراضی، می‌تواند تاثیر چشم‌گیری بر کیفیت خاک در کاربری‌های مختلف داشته باشد. در منطقه مورد مطالعه نیز، مدیریت بقایای گیاهی نقش مهمی در ذخیره کربن آلی خاک داشته، به طوری که، به دلیل کاهش بارش باران سال‌های اخیر در منطقه مورد مطالعه، چرای بیش از حد در مراتع صورت گرفته، که سبب کاهش بازگشت بقایای گیاهی به خاک شده است. کاهش بازگشت بقایای گیاهی، سبب کاهش کربن ورودی به خاک و در نتیجه کاهش ذخیره کربن آلی کاربری مراتع منطقه شده است. Luo and Zhou (2006) نیز در مطالعات خود چرای بیش از حد مراتع را عاملی برای افزایش تصاعد کربن آلی خاک و در نتیجه کاهش میزان ذخیره کربن آلی خاک عنوان کردند. از طرفی در منطقه مورد مطالعه، بیشتر اراضی کشاورزی در یک دوره تناوب، مورد کشت یونجه قرار می‌گیرند، که علاوه بر استفاده از حاصلخیزکننده‌های آلی، به دلیل استفاده از کود حیوانی در هنگام آماده‌سازی اراضی، افزایش میزان کربن آلی این اراضی را سبب شده است.

نتایج تحقیق همچنین نشان داد، که اختلاف میزان ذخیره کربن بین عمق ۰-۳۰ و ۳۰-۶۰ در کاربری کشاورزی نسبت به کاربری نخلستان و مرتع کمتر می‌باشد، که یک دلیل اصلی برای این امر، انجام عملیات خاک‌ورزی می‌باشد که سبب مخلوط شدن لایه‌های زیرین خاک با درصد کربن آلی کمتر با خاک سطحی حاوی کربن آلی بیشتر شده و در نتیجه موجب کاهش کربن آلی خاک سطحی در قیاس با حالت اولیه می‌شود.

میزان ذخیره کربن با پارامترهای فیزیکی و شیمیایی خاک نیز مورد بررسی قرار گرفت. نتایج این تحقیق نشان داد که میزان ذخیره کربن با درصد آهک و هدایت الکتریکی همبستگی منفی و با میزان کربن آلی خاک همبستگی مثبت دارد. یک دلیل برای کاهش میزان ذخیره کربن خاک با افزایش میزان هدایت الکتریکی و آهک خاک، تاثیر منفی این عوامل بر روی پوشش گیاهی و در نتیجه کاهش ورود لاشبرگ به خاک می‌باشد، که سبب کاهش ماده آلی و در نتیجه ذخیره کربن خاک می‌شوند. نتایج تحقیقات Chahouki *et al.* (2011) نیز این موضوع را تایید می‌کند. نتایج تحقیقات Brahim *et al.* (2011) که در مراتع نیمه‌خشک کشور تونس، جهت بررسی ارتباط ویژگی‌های خاک با میزان کربن آلی دو نوع خاک رسی و خاک شنی انجام گرفت، نتایج نشان داد که میزان ذخیره کربن آلی خاک تحت تاثیر پارامترهای فیزیکی و شیمیایی قرار می‌گیرد.

از آنجا که هدف این مطالعه، مقایسه بررسی کارایی و دقت سه مدل مختلف جنگل تصادفی، شبکه عصبی مصنوعی و ماشین‌بردار پشتیبان می‌باشد، سعی شد تا ورودی‌های مدل‌ها

پیش‌بینی کننده‌ها وجود دارد. نتایج تحقیقات Dharumarajan *et al.* (2017) که در منطقه گرمسیری نیمه‌خشک جنوب هندوستان در خصوص چگونگی توزیع مکانی خصوصیات خاک، با استفاده از مدل جنگل تصادفی انجام گرفته بود نیز نشان داد، که مدل جنگل تصادفی جهت بررسی توزیع مکانی خصوصیات خاک نتایج خوبی را ارائه می‌دهد.

نتایج مقایسه کارآیی و عملکرد مدل‌ها (جدول ۳)، همچنین نشان داد، که مقادیر MAE در تمامی مدل‌ها مثبت می‌باشد، که نشان می‌دهد که مقادیر واقعی از مقادیر پیش‌بینی میزان ذخیره کربن بیشتر بوده ولی تفاوت این مقادیر برای مدل جنگل تصادفی نسبت به دیگر مدل‌ها کمتر می‌باشد.

### سپاسگزاری

از گروه خاک‌شناسی دانشگاه شهید چمران اهواز جهت در اختیار قراردادن امکانات آزمایشگاه و همچنین از حمایت مالی این دانشگاه تشکر و قدردانی می‌گردد.

### REFERENCES

- Adhikari, K., Hartemink, A.E., Minasny, B., Kheir, R.B., Greve, M.B., and Greve, M.H.J.P.O. (2014). Digital mapping of soil organic carbon contents and stocks in Denmark. *PLoS One*, 9(8), 105519.
- Aitkenhead, M.J., and Coull, M.C. (2016). Mapping soil carbon stocks across Scotland using a neural network model. *Geoderma*, 262, 187-198.
- Akpa, S.I.C., Odeh, I.O.A., Bishop, T.F.A., Hartemink, A.E., and Amapu, I.Y. (2016). Total soil organic carbon and carbon sequestration potential in Nigeria. *Geoderma*, 271, 202-215.
- Amirian Chakan, A., Taghizadeh-Mehrjardi, R., Kerry, R., Kumar, S., Khordehbin, S., and Khanghah, S.Y. (2017). Spatial 3D distribution of soil organic carbon under different land use types. *Environmental Monitoring and Assessment*, 189(3), 131.
- Blake, G.R., and Hartge, K.H. (1986). Bulk density 1. *Methods of soil analysis: part 1—physical and mineralogical methods*, (methodsofsoilan1), 363-375.
- Brahim, N., Blavet, D., Gallali, T., and Bernoux, M. (2011). Application of structural equation modeling for assessing relationships between organic carbon and soil properties in semiarid Mediterranean region. *International Journal of Environmental Science and Technology*, 8(2), 305-320.
- Bohn, H. L. (1976). Estimate of organic carbon in world soils 1. *Soil Science Society of America Journal*, 40(3), 468-470.
- Chahouki, M. A. Z., Ahvazi, L. K., and Azarnivand, H. (2011). Environmental factors affecting distribution of vegetation communities in Iranian rangelands. *Vegetos*, 24(1), 1-15.
- Chen, S., Arrouays, D., Angers, D. A., Martin, M. P., and Walter, C. (2019). Soil carbon stocks under different land uses and the applicability of the soil carbon saturation concept. *Soil and Tillage Research*, 188, 53-58.
- Cortes, C., and Vapnik, V. (1995). Support-vector networks. *Machine learning*, 20(3), 273-297.
- Dharumarajan, S., Hegde, R., and Singh, S. K. (2017). Spatial prediction of major soil properties using Random Forest techniques-A case study in semi-arid tropics of South India. *Geoderma Regional*, 10, 154-162.
- Eswaran, H., Van Den Berg, E., and Reich, P. (1993). Organic carbon in soils of the world. *Soil Science Society of America journal*, 57(1), 192-194.
- 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.
- Grimm, R., Behrens, T., Märker, M., and Elsenbeer, H. (2008). Soil organic carbon concentrations and stocks on Barro Colorado Island—Digital soil mapping using Random Forests analysis. *Geoderma*, 146(1-2), 102-113.
- Hengl, T., Heuvelink, G.B.M., and Stein, A. (2004). A generic framework for spatial prediction of soil variables based on regression-kriging. *Geoderma*, 120(1-2), 75-93.
- Jenny, H. (1994). *Factors of soil formation: a system of quantitative pedology*. Courier Corporation.
- Kuang, B., Tekin, Y., and Mouazen, A. M. (2015). Comparison between artificial neural network and partial least squares for on-line visible and near infrared spectroscopy measurement of soil organic carbon, pH and clay content. *Soil and Tillage Research*, 146, 243-252.
- Lal, R. (2004). Soil carbon sequestration to mitigate climate change. *Geoderma*, 123(1-2), 1-22.
- Ließ, M., Glaser, B., and Huwe, B. (2012). Uncertainty in the spatial prediction of soil texture: comparison of regression tree and Random Forest models. *Geoderma*, 170, 70-79.
- Luo, Y. and Zhou, X., 2006. *Soil Respiration and the Environment*. 320pp.
- McBratney, A. B., Santos, M. M., and Minasny, B. (2003). On digital soil mapping. *Geoderma*, 117(1-2), 3-52.
- Mingjun, T., Lixiong, Z., Wenfa, X., Zhiling, H., Zhixiang, Z., Zhaogui, Y., and Pengcheng, W.J.S.O.T.T.E. (2017). Spatial variability of soil

- organic carbon in Three Gorges Reservoir area, China. *Science of the Total Environment*, 599, 1308-1316.
- Mishra, U. (2009). Predicting storage and dynamics of soil organic carbon at a regional scale (Doctoral dissertation, The Ohio State University).
- Mitran, T., Mishra, U., Lal, R., Ravisankar, T., and Sreenivas, K. (2018). Spatial distribution of soil carbon stocks in a semi-arid region of India. *Geoderma Regional*, 15, e00192.
- Minasny, B., and McBratney, A. B. (2006). A conditioned Latin hypercube method for sampling in the presence of ancillary information. *Computers and Geosciences*, 32(9), 1378-1388.
- Minasny, B., McBratney, A. B., Malone, B. P., and Wheeler, I. (2013). Digital mapping of soil carbon. In *Advances in Agronomy* (Vol. 118, pp. 1-47). Academic Press.
- 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.
- Padilla, F. M., Vidal, B., Sánchez, J., and Pugnaire, F. I. (2010). Land-use changes and carbon sequestration through the twentieth century in a Mediterranean mountain ecosystem: implications for land management. *Journal of Environmental Management*, 91(12), 2688-2695.
- Pouladi, N., Møller, A. B., Tabatabai, S., and Greve, M. H. (2019). Mapping soil organic matter contents at field level with Cubist, Random Forest and kriging. *Geoderma*, 342, 85-92.
- Shataee, S., Kalbi, S., Fallah, A., and Pelz, D. (2012). Forest attribute imputation using machine-learning methods and ASTER data: comparison of k-NN, SVR and random forest regression algorithms. *International journal of remote sensing*, 33(19), 6254-6280.
- Somarathne, S., Seneviratne, G., and Coomaraswamy, U. (2005). Prediction of soil organic carbon across different land-use patterns. *Soil Science Society of America Journal*, 69(5), 1580-1589.
- Tiessen, H. J. W. B., and Stewart, J. W. B. (1983). Particle-size fractions and their use in studies of soil organic matter: II. Cultivation effects on organic matter composition in size fractions 1. *Soil Science Society of America Journal*, 47(3), 509-514.
- Tiryaki, S., and Aydın, A. (2014). An artificial neural network model for predicting compression strength of heat treated woods and comparison with a multiple linear regression model. *Construction and Building Materials*, 62, 102-108.
- Vapnik, V.N., Lerner A. Ya. (1963) "Recognition of Patterns with help of Generalized Portraits", *Avtomat. i Telemekh.*, 24(6), 774-780
- 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.
- Wang, Y. F., Liu, L., and Shanguan, Z. P. (2017). Carbon storage and carbon sequestration potential under the Grain for Green Program in Henan Province, China. *Ecological Engineering*, 100, 147-156.
- Wang, Z. (2019). Estimating of terrestrial carbon storage and its internal carbon exchange under equilibrium state. *Ecological Modelling*, 401, 94-110.
- Yigini, Y., and Panagos, P. (2016). Assessment of soil organic carbon stocks under future climate and land cover changes in Europe. *Science of the Total Environment*, 557, 838-850.
- Zhang, C., Liu, G., Xue, S., and Sun, C. (2013). Soil organic carbon and total nitrogen storage as affected by land use in a small watershed of the Loess Plateau, China. *European Journal of Soil Biology*, 54, 16-24.
- Zhang, H., Wu, P., Yin, A., Yang, X., Zhang, M., and Gao, C. (2017). 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*, 592, 704-713.