



رقومی سازی کربن آلی خاک با شبکه عصبی مصنوعی و رگرسیون خطی چندمتغیره در استان کردستان

حمید محمودزاده^۱، * حمیدرضا متین فر^۲ و روح الله تقی زاده - مهرجردی^۳

^۱دانش آموخته دکتری گروه علوم خاک، دانشگاه لرستان، ^۲دانشیار گروه علوم خاک، دانشگاه لرستان، ^۳دانشیار گروه علوم خاک، دانشگاه لرستان،

^۲دانشیار گروه علوم کشاورزی و منابع طبیعی، دانشگاه اردکان

تاریخ دریافت: ۱۳۹۹/۰۳/۲۴؛ تاریخ پذیرش: ۱۳۹۹/۰۹/۳۰

چکیده

سابقه و هدف: کربن آلی از طریق نگهداشت بخش قابل توجهی از فرم‌های آلی قابل معدنی شدن آن در خاک، نقشی حیاتی در کنترل اقلیم و پایداری محیط زیست دارد. هم‌چنین کربن آلی تأثیر کلیدی بر خصوصیات فیزیکوشیمیایی و بیولوژیکی خاک دارد؛ به نحوی که از آن به عنوان شاخص سلامت خاک نام برده می‌شود. به همین جهت، بررسی توزیع مکانی کربن آلی خاک، جهت شناسایی مناطق با پتانسیل ترسیب کربن، از الزامات برنامه‌ریزی مدیریت خاک و سیاست‌گذاری کنترل اقلیم از طریق فعالیت‌های کشاورزی می‌باشد. روش‌های مرسوم برآورد کربن آلی خاک، پرهزینه و زمان‌بر بوده و قابلیت تکرار و تعمیم به نقاط مشابه را ندارد. در سال‌های اخیر با پیشرفت فناوری و نیاز روز افزون بشر برای دستیابی به اطلاعات زودپایافت و صرفه‌جویی در هزینه، از طریق داده‌کاوی و به کمک تصاویر ماهواره‌ای و متغیرهای کمکی توپوگرافی، رقومی سازی ویژگی‌های خاک از جمله کربن آلی امکان‌پذیر شده است. نقشه‌برداری رقومی خاک در واقع توسعه یک مدل عددی یا آماری از رابطه بین متغیرهای محیطی و خصوصیات خاک است که برای داده‌های جغرافیایی زیادی به منظور تولید نقشه رقومی به کار می‌رود. سه هدف اصلی نقشه‌برداری رقومی خاک عبارت از: (۱) استنباط رابطه بین متغیرهای محیطی و خصوصیات خاک، (۲) تولید و ارائه داده‌هایی که پیوستگی خاک-زمین‌نما را بهتر نمایش می‌دهند و (۳) به کارگیری صریح دانش کارشناس در طراحی مدل می‌باشند. هم‌چنین نقشه‌برداری رقومی با ایجاد بینشی در مورد فرآیندهای خاکساز، باعث پیشرفت بالقوه پدولوژی و جغرافیای خاک می‌شود.

مواد و روش‌ها: این مطالعه در شهرستان کامیاران استان کردستان و به منظور پیش‌بینی کربن آلی خاک انجام شده است. در این پژوهش تعداد ۱۱۰ نمونه خاک بصورت تصادفی از کاربری‌های مختلف اراضی مورد آنالیز قرار گرفت. برای پیش‌بینی بهتر توزیع مکانی کربن آلی خاک در منطقه مورد مطالعه از ۱۰۱ متغیر کمکی استخراج شده از مدل رقومی ارتفاع، تصاویر ماهواره‌ای و متغیرهای اقلیمی نیز استفاده شده است. پیش‌بینی مقدار کربن آلی خاک با دو مدل رگرسیون خطی چندمتغیره و شبکه عصبی مصنوعی در محیط نرم‌افزار جامپ مدل‌سازی شد.

یافته‌ها: نتایج نشان داد که مقدار کربن آلی خاک در بخش‌های غربی و شمال‌غربی منطقه مورد مطالعه، بیش‌ترین مقدار است که شامل مناطق با پوشش جنگلی و مرتعی است. متغیرهای کمکی سطح پایه شبکه کانال (۰.۴۰٪)، باند ۴ سنجنده

* مسئول مکاتبه: matinfar.h@lu.ac.ir

OLI (۰/۲۳)، مقدار آب برگ (۰/۲۰)، زبری زمین (۰/۱۹)، فاصله عمودی تا شبکه کانال (۰/۱۸)، شیب حوزه (۰/۱۸)، شاخص تفاضل نرمال شده پوشش گیاهی (۰/۱۷)، سطح حوزه (۰/۱۶)، جهت شیب (۰/۱۶)، ارتفاع (۰/۱۶)، باند ۳ (۰/۱۵)، شاخص جذب انعکاسی (۰/۱۴)، باند ۱ (۰/۱۴)، باران (۰/۱۳)، باند ۵ (۰/۱۳)، دمای هوا (۰/۱۲)، شاخص پوشش گیاهی (۰/۱۱)، شاخص خیسی توپوگرافی (۰/۱۰) و شاخص تفاضل پوشش گیاهی (۰/۱۰) بیشترین اثر را روی مدل‌سازی کربن آلی خاک در مدل شبکه عصبی مصنوعی داشته‌اند. نتایج اعتبار سنجی مدل‌سازی نشان داد که شبکه عصبی مصنوعی عملکرد بهتری ($R^2=0/97$) نسبت به رگرسیون خطی چند متغیره ($R^2=0/59$) برای پیش‌بینی کربن آلی خاک در منطقه مورد مطالعه داشته است.

نتیجه‌گیری: نتایج این پژوهش نشان داد که پراکنش کربن آلی بیش‌تر تحت‌تأثیر عوامل توپوگرافی و اقلیم می‌باشد. در مناطقی که به هر دلیل امکان نمونه‌برداری در کل منطقه وجود ندارد، می‌توان از طریق متغیرهای کمکی مانند پارامترهای توپوگرافی، اقلیمی و پوشش گیاهی و با روش‌های نوین داده‌کاوی برای تخمین کربن آلی خاک بهره گرفت.

واژه‌های کلیدی: پهنه‌بندی کربن آلی خاک، داده کاوی، نقشه‌برداری دیجیتال خاک

مقدمه

کربن آلی خاک، شاخص مهمی در ارزیابی حاصلخیزی خاک و مدیریت اراضی است که سایر خصوصیات خاک را تحت‌تأثیر قرار می‌دهد (۴۳). ذخیره کربن آلی خاک در مقیاس جهانی (۱۵۵۰ گیگاتن) به‌ترتیب بیش از ۲ و ۲/۸ برابر کربن اتمسفر (۷۶۰ گیگاتن) و کربن ذخیره‌شده در گیاهان (۵۶۰ گیگاتن) است (۲۱). از این‌رو، اندک تغییری در کربن آلی خاک می‌تواند مقدار دی‌اکسید کربن اتمسفر را به‌شدت تحت‌تأثیر قرار داده و به‌تبع آن روی اقلیم تأثیرگذار باشد (۶). در پژوهش‌های مختلف به‌دلیل قابلیت بالای کربن آلی خاک در کنترل اقلیم، امنیت غذایی و پایداری محیط‌زیست، از آن به‌عنوان شاخص کیفیت و سلامت خاک نام‌برده شده است (۴۸). بررسی توزیع مکانی مقدار کربن آلی خاک، اساسی‌ترین قدم در سیاست‌گذاری اقلیم و مدیریت خاک، افزایش تولید غذای بشر (۴۱) و تهیه داده‌های استاندارد برای مدل‌سازی اکوسیستم (۲۳) است. هم‌چنین به‌علت افزایش جمعیت و ضرورت تأمین

نیاز غذایی و امنیت خوراک جهانی نیاز به دستیابی اطلاعات ویژگی‌های خاک رو به افزایش است. از این‌رو اقدامات انجام‌شده برای پایش و ارزیابی کربن آلی خاک، به‌طور فزاینده‌ای در حال گسترش است. از سوی دیگر نمونه‌برداری مستقیم همراه با اندازه‌گیری‌های آزمایشگاهی یا صحرایی، معمولاً پرهزینه و وقت‌گیر بوده و در نتیجه، برای اهداف کاربردی دارای محدودیت است. بنابراین کاربرد روش‌هایی برای پایش و بررسی روند تغییرات کربن آلی خاک و سایر ویژگی‌های خاک در سطح وسیع نیاز است. هم‌چنین رقوم‌سازی و بررسی تغییر مکانی کربن آلی خاک با استفاده از روش‌های آماری پیشرفته، ضرورتی انکارناپذیر است.

یکی از روش‌های پاسخ به این نیاز روزافزون بشری، استفاده از نقشه‌برداری رقومی خاک است که در آن خصوصیات خاک (مانند کربن آلی)، بر اساس ارتباط با متغیرهای محیطی (از جمله پارامترهای اقلیمی، سنجش‌ازدور و توپوگرافی) به‌صورت رقومی نقشه‌برداری می‌گردند (۱۲ و ۲۷). مک‌برانتی و

الگوریتم‌های مختلفی از ماشین یادگیری برای تهیه نقشه رقومی کربن آلی خاک استفاده شده است که از آن جمله می‌توان به شبکه‌های عصبی مصنوعی^۳ (۲۳ و ۲۵)، ماشین بردار پشتیبان^۴ (۴۵)، درختان رگرسیون تقویت شده^۵ (۲۶) و جنگل تصادفی (۱۲، ۴۴ و ۴۹) اشاره نمود. مطالعات محدودی در ارتباط با تغییرات مکانی کربن آلی خاک در ایران، به‌ویژه در غرب ایران و استان کردستان انجام شده است (۸ و ۳۱). شهرستان کامیاران در جنوب استان کردستان واقع شده و به‌عنوان یکی از قطب کشاورزی استان، به‌ویژه از نظر تولیدات باغی، پتانسیل بالایی داشته و به تبع آن مورد تهدید تخریب خاک نیز می‌باشد. از آنجایی که کشاورزی در شهرستان کامیاران بیشتر به‌صورت سنتی بوده و خاک‌های آن در معرض تخریب قابل توجه می‌باشد بنابراین پایش کربن آلی خاک به‌عنوان عاملی تأثیرگذار در کیفیت خاک ضرورتی انکارناپذیر می‌باشد. به‌طورکلی در استان کردستان و بالاخص در شهرستان کامیاران مطالعات مدونی در زمینه رقومی سازی کربن آلی خاک انجام نشده است؛ بنابراین هدف مطالعه حاضر (۱) بررسی توزیع مکانی کربن آلی خاک در شهرستان کامیاران، (۲) بررسی کارایی مدل‌های رگرسیون چندمتغیره و شبکه عصبی مصنوعی در پیش‌بینی کربن آلی خاک و (۳) ارائه نقشه‌های با مقیاس مناسب و رقومی جهت برنامه‌ریزی صحیح بهره‌برداری کشاورزی، منابع طبیعی و محیط زیست است.

مواد و روش‌ها

محدوده مورد مطالعه: منطقه مورد مطالعه در این پژوهش، شهرستان کامیاران با وسعتی بالغ بر ۲۰۳۵ کیلومتر مربع در غرب ایران و جنوب کردستان و بین

همکاران (۲۰۰۳) چهارچوب پیش‌بینی‌کننده برای تخمین تغییرپذیری خاک در مناطق با داده‌های محدود خاک را ارائه کردند (۲۷) که از طریق ترکیب اندازه‌گیری‌های مستقیم خصوصیات خاک با یکسری متغیرهای کمکی زودیافت استخراج شده از نقشه‌های مدل رقومی ارتفاع، متغیرهای سنجش‌ازدور و متغیرهای اقلیمی عمل می‌کند (۱ و ۴۳). با توسعه روش‌های سنجش‌ازدور در طول یک دهه گذشته، دستیابی به بخش عظیمی از متغیرهای کمکی در مقیاس منطقه‌ای و جهانی با هزینه حداقل و وقت اندک میسر گردید. بنابراین، متغیرهای سنجش‌ازدور منبع مهمی برای پیش‌بینی مکانی کربن آلی خاک محسوب می‌شوند (۳۰، ۴۴ و ۴۸).

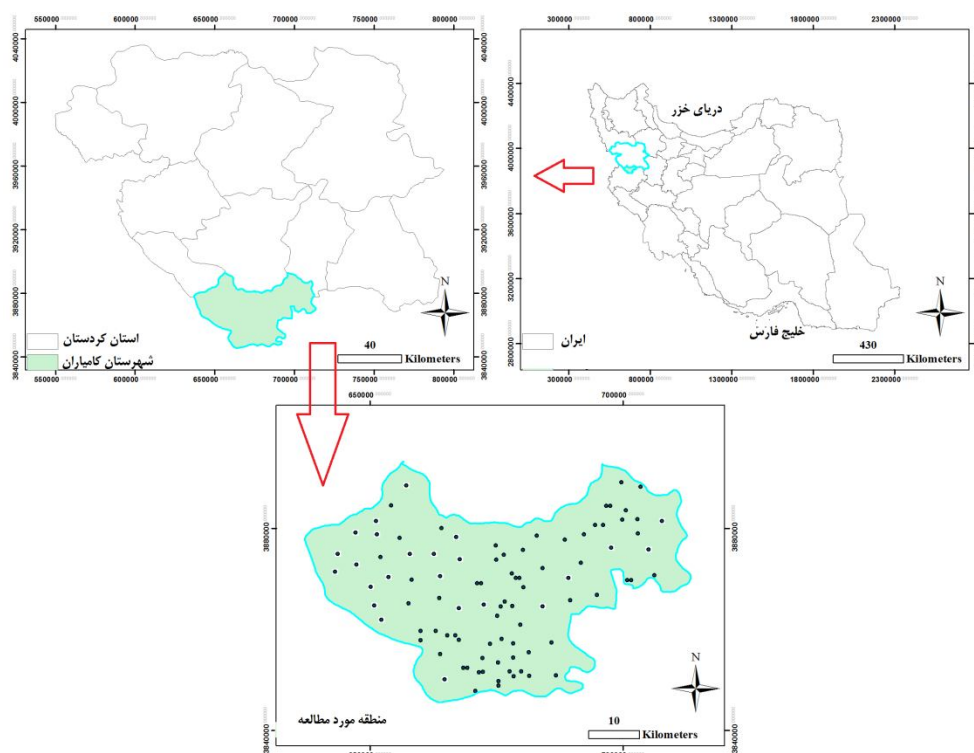
اخیراً در مطالعات زیادی رابطه بین داده‌های آزمایشگاهی و متغیرهای کمکی در دسترس از طریق ماشین یادگیری بررسی شده است که این الگوریتم‌ها محدودیت‌های روش‌های آماری پارامتریک و غیرپارامتریک را ندارند (۴۸) و توانایی قابل‌قبولی در رقومی سازی کربن آلی خاک را دارند (۴۱). طیف وسیعی از الگوریتم‌های ماشین یادگیری برای پیش‌بینی کربن آلی خاک در مقیاس منطقه‌ای و ملی استفاده شده است (۱۳). به‌عنوان مثال، وانگ و همکاران (۲۰۱۷) در لیائونینگ چین توزیع مکانی کربن آلی خاک را در پنج عمق با استفاده از مدل رگرسیون درختی^۱ بررسی کرده‌اند (۴۶) و ویشمایر و همکاران (۲۰۱۲) با روش جنگل تصادفی^۲ ذخیره کربن آلی خاک را در مونتگولیا چین برآورد کرده است (۴۹). با توجه به تنوع اقلیمی، ویژگی‌های خاک و مدیریت اراضی، روش‌های مختلفی برای پیش‌بینی ویژگی‌های خاک قابل استفاده است (۱۶).

3- Artificial neural networks
4- Support vector machine
5- Boosting Regression Tree

1- Classification and Regression Tree
2- Random Forest

طول جغرافیایی ۶۶ درجه و ۳۰ دقیقه و ۴۷ ثانیه تا ۴۷ درجه و ۲۰ دقیقه و ۳ ثانیه شرقی و عرض جغرافیایی ۳۴ درجه و ۴۴ دقیقه و ۳۵ ثانیه تا ۳۵ درجه و ۹ دقیقه و ۵۳ ثانیه شمالی قرار دارد (شکل ۱). محدوده مطالعاتی از شمال به شهرستان سنندج، از شمال شرق به قروه، از شرق به سنقر، از جنوب به کرمانشاه، از جنوب غرب به جوانرود، از غرب به پاوه و از شمال غرب به سروآباد محدود می‌شود. کامیاران جنوبی‌ترین شهرستان استان کردستان است و حصارى از کوهستان‌های شاهو و بیستون اطراف آن قرار دارند. میانگین بارندگی و دمای سالانه این منطقه به ترتیب ۵۲۰ میلی‌متر و ۱۳/۵ درجه سانتی‌گراد است (۳۷). سطح اراضی زیر کشت این شهرستان ۳۶۵۰۹ هکتار است که ۳۲۲۳۰ هکتار زراعی (۱۲۱۸۷ هکتار کشت آبی و ۲۰۰۴۳ هکتار کشت دیم) و ۴۲۷۹ هکتار باغی

۳۸۲۱ هکتار باغ آبی و ۴۵۸ هکتار باغ دیمی) است که حدود ۳۴ درصد مساحت کل شهرستان را در بر می‌گیرد (۱۸). اراضی ملی و منابع طبیعی ۶۶ درصد مساحت شهرستان را تشکیل می‌دهد که حدود ۱۱ درصد آن (۲۸،۰۰۰ هکتار) جنگل است (۱۳). کاربری‌های عمده این منطقه شامل جنگل درختان بلوط، باغ‌های درختان میوه (توت‌فرنگی، گردو، انگور)، زراعت آبی (گندم، ذرت، یونجه)، دیم (گندم، جو و نخود) و مرتع است (۱۸) (شکل ۲). رژیم رطوبتی خاک منطقه باتوجه به اطلاعات موجود منحنی‌های هم‌باران عمدتاً در دامنه رژیم زیریک و رژیم حرارتی با توجه منحنی‌های هم‌دما و متوسط سالانه دما در دامنه رژیم حرارتی مزیک قرار می‌گیرد (۲۲).



شکل ۱- موقعیت منطقه مورد مطالعه در ایران و کردستان به همراه نقاط نمونه‌برداری.

Figure 1. Location of the study area in Iran and Kurdistan along with sampling points.

همچنین سایر نسبت‌های باندی برای به‌دست آوردن مواد مادری و عوامل خاک (به‌عنوان مثال کانی‌های رس، اکسید آهن و شاخص شوری) در منطقه مطالعاتی محاسبه گردید.

توصیف مشخصات سرزمین با استفاده از مدل رقومی ارتفاع (وضوح قدرت تفکیک ۳۰ متر) دریافتی از وب‌سایت <http://srtm.csi.cgiar.org/> استخراج گردید. مدل رقومی ارتفاع برای اصلاح پستی‌وبلندی، پیش‌پردازش گردید. تعداد ۴۷ شاخص زمین با استفاده از نرم‌افزار SAGA GIS از مدل رقومی ارتفاع به‌دست آمد (۷) که خلاصه آن در جدول ۱ آورده شده است. متغیرهای اقلیمی از سازمان هواشناسی استان کردستان دریافت گردید و میانگین ۳۰ ساله دمای هوا، بارندگی و دمای خاک محاسبه گردید. سپس برای مطابقت آن‌ها با سایر نقشه‌های رستری، اطلاعات اقلیمی کل منطقه با درون‌یابی توسط ArcGIS نسخه ۱۰٫۲ به روش نزدیک‌ترین همسایه به‌دست آمد.

تمام متغیرهای کمکی استخراج‌شده از هر سه پایگاه داده برای تهیه شبکه با قدرت تفکیک مکانی ۹۰ متر (استاندارد جهانی برای رقومی‌سازی نقشه ویژگی‌های خاک) ترکیب و به‌عنوان پارامترهای پیش‌بینی‌کننده برای توسعه مدل‌های مکانی کمی اعمال گردید. در نهایت مقدار کمی صفات تمام نقشه‌های رستری در نقاط مشخص نمونه‌برداری با استفاده از نرم‌افزار SAGA GIS به‌دست آمد که ورودی اصلی مدل‌های مکانی است. مقادیر بعضی از متغیرهای پیش‌بینی‌کننده مرتبط با مقدار کربن آلی خاک در جدول ۱ آورده شده است.

برای تهیه و تجزیه و تحلیل نقشه کاربری و پوشش اراضی، ابتدا نقشه‌های توپوگرافی و تصاویر ماهواره‌ای منطقه تهیه گردید. پس از ژئورفرنس و موزاینک کردن نقشه توپوگرافی، سایر نقشه‌های شیب، جهت شیب و ارتفاع منطقه در محیط نرم‌افزار

نمونه‌برداری و متغیرهای کمکی مورد‌استفاده و تهیه نقشه کاربری اراضی: تعداد ۱۱۰ نمونه خاک از عمق صفر تا ۳۰ سانتی‌متری از نقاط مختلف اراضی شهرستان کامیاران به‌صورت طبقه‌ای تصادفی طبقه‌بندی شده برداشت شد تا نمونه در همه کاربری‌ها شانس انتخاب را داشته باشند. نمونه‌های خاک ابتدا هوا خشک گردید؛ سپس از الک دو میلی‌متری عبور داده شد و مقدار کربن آلی آن‌ها از روش تیتراسیون تر (۳۳) تعیین گردید.

بر اساس عوامل خاک‌سازی تحت عنوان معادله اسکورپین^۱ (۲۷) و مطالعات قبلی (۳، ۳۸، ۳۹ و ۴۱)، تعداد ۱۰۱ متغیر کمکی (پیش‌بینی‌کننده) برای توضیح و پیش‌بینی توزیع مکانی کربن آلی خاک استفاده گردید. متغیرهای کمکی از سه پایگاه تصاویرسنجنده OLI لندست هشت شهریور سال ۱۳۹۷، مدل رقومی ارتفاع و نقشه‌های اقلیمی به‌دست آمد. چهار تصویر از سنجنده OLI لندست هشت با قدرت تفکیک مکانی ۳۰ متر از پایگاه داده <http://earthexplorer.usgs.gov/> دریافت شد. تبدیل درجه روشنایی به رادیانس روی تمام باندها اعمال شد و با استفاده از روش جسم تاریک، تصحیح اتمسفری انجام گردید (۴۲). تعداد ۵۱ شاخص گیاهی از نسبت باندهای طیفی (۳۰) با نرم‌افزار پردازش تصاویر ENVI نسخه ۵٫۳ و نرم‌افزار آماری جامپ^۲ نسخه ۱۳٫۲ از تصاویر ماهواره‌ای استخراج گردید.

۱- اجزای فاکتورهای مدل SCORPAN به شرح ذیل به‌دست آمد؛ S: ویژگی کربن آلی خاک در نقاط نمونه‌گیری، C: ویژگی‌های اقلیمی که از نقشه‌های اقلیمی به‌دست آمده است، O: شامل پوشش زمین و پوشش طبیعی است که از تصاویر ماهواره‌ای سنجنش از دور حاصل شده است، R: توپوگرافی که شامل شاخص‌های زمین است که از مدل رقومی ارتفاع به‌دست آمده است، P: مواد مادری که شامل ویژگی‌های لیتولوژی است که از تصاویر سنجنش از دور استخراج شده است، A: زمان، N: موقعیت فضایی یا جغرافیایی

2- JMP

R^2 : ضریب تعیین رگرسیونی^۱ (ضریب تعیین نشانگر درصدی از تغییرات متغیر وابسته است که به وسیله متغیر مستقل تبیین می‌شود و با استفاده از رابطه زیر محاسبه می‌شود).

RMSE: خطای جذر میانگین مربعات^۲ (تفاوت میان مقدار پیش‌بینی شده توسط مدل و مقدار واقعی می‌باشد که ابزاری مناسب برای مقایسه خطاهای پیش‌بینی توسط یک مجموعه داده‌است).

$$R^2 = \left(\frac{\sum_{i=1}^n (O_i - \bar{O})(P_i - \bar{P})}{\sqrt{\sum_{i=1}^n (O_i - \bar{O})^2 \sum_{i=1}^n (P_i - \bar{P})^2}} \right)^2$$

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (P_i - P_o)^2}$$

Pi: کربن آلی پیش‌بینی شده

Oi: کربن آلی اندازه‌گیری شده در نقاط

n: تعداد نمونه

\bar{P} : متوسط کربن آلی پیش‌بینی شده

\bar{O} : متوسط کربن آلی اندازه‌گیری شده

ضریب تبیین بالاتر و جذر میانگین مربعات خطای کم‌تر، نشانگر دقت بالاتر مدل می‌باشد.

انتخاب ویژگی‌ها: دشوارترین و مهم‌ترین مرحله در الگوریتم یادگیری ماشین، انتخاب ویژگی^۳ برای کاهش بعد و افزایش کارکرد الگوریتم‌های یادگیری است. اطلاعات نامناسب و زائد ممکن است دقت پیش‌بینی در الگوریتم‌های یادگیری ماشین رایج را کاهش دهد (۱۳ و ۲۹). انتخاب ویژگی فرآیند کاهش تعداد متغیرهای ورودی هنگام ایجاد یک مدل پیش‌بینی است. مطلوب است تعداد متغیرهای ورودی را کاهش داد تا هم هزینه محاسباتی مدل‌سازی کاهش و هم عملکرد مدل بهبود یابد. روش‌های انتخاب

GIS ۱۰,۲ تولید گردید. پس از تصحیح و پردازش تصاویر ماهواره‌ای در محیط نرم‌افزاری ENVI 5.3 انجام شد. سپس با استفاده از طبقه‌بندی نظارت شده و الگوریتم حداکثر احتمال، نقشه طبقه‌بندی شده کاربری اراضی تولید شد.

مدل‌سازی کربن آلی خاک و ارزیابی دقت مدل: در این مطالعه برای برآورد کربن آلی خاک از شبکه عصبی مصنوعی و رگرسیون خطی چندمتغیره استفاده شد. مدل شبکه عصبی مصنوعی پرسپترون چندلایه با سه لایه مخفی و یک تابع فعال‌سازی سیگموئیدی در لایه مخفی و تابع فعال‌سازی خطی در لایه خروجی بوده و تعداد نوروں ۲ تا ۱۰ متغیر که بهترین تعداد نوروں با آزمون و خطا تعیین گردید. در این مطالعه از نرم‌افزار جامپ برای مدل‌سازی پیش‌بینی کربن آلی خاک استفاده گردید.

برای بررسی ارزیابی نقشه کربن آلی خاک، اعتبارسنجی به روش تقسیم داده‌های موجود به دو دسته آزمون و آموزش بوده است. در این مرحله، تعداد ۸۲ داده برای آموزش و ۲۸ داده برای آزمون در نظر گرفته شد. خروجی شبکه عصبی مصنوعی که بعد از آموزش برآورد گردید مقدار کربن آلی است. به‌منظور ارزیابی صحت شبکه، مقدار برآورد شده و مقدار واقعی هر نمونه در یک نمودار رسم شد و ضریب همبستگی بین داده‌های واقعی و تخمین‌زده شده مورد بررسی قرار گرفت. شبکه عصبی مصنوعی در مرحله آموزش بین متغیرهای ورودی (متغیرهای کمکی استخراج‌شده از مدل رقومی ارتفاع و تصاویر ماهواره‌ای) و متغیر خروجی (درصد کربن آلی خاک) روابط ریاضی منطقی موجود را پیدا کرده و به‌عنوان یک مدل شبکه برای تخمین داده‌هایی که در مرحله آموزش استفاده نشده است، استفاده می‌گردد و بدین‌صورت مقدار خروجی مدل پیش‌بینی می‌گردد. برای ارزیابی نتایج عملکرد مدل از آماره‌های به شرح زیر استفاده شد.

- 1- Regression coefficients of determination
- 2- Root mean of squared error
- 3- Feature selection

زیاد است که این اختلاف نشانگر تأثیرپذیری بالای افق سطحی خاک‌ها در منطقه مورد مطالعه از اعمال مدیریت‌های گوناگون، ویژگی‌های محیطی و تخریب اراضی می‌باشد. بیش‌ترین مقدار کربن آلی خاک در مناطق جنگلی (۲/۷۱ درصد) و کم‌ترین مقدار آن در اراضی زراعی (۰/۴۶ درصد) مشاهده گردید. وجود توپوگرافی متغیر و هم‌چنین کاربری‌های متنوع اراضی در منطقه مورد مطالعه (شکل ۲) باعث تغییر در توزیع رطوبت خاک و پوشش گیاهی شده که آن‌هم به نوبه خود تغییرات کربن آلی خاک منطقه را به دنبال خواهد داشت. هیستوگرام مقدار کربن آلی خاک دارای توزیع نرمال بوده و ضریب تغییرات آن (۴۲/۶۱ درصد) در دامنه تغییرپذیری بالا قرار گرفته است. ویلدینگ (۱۹۸۵) مقدار ضریب تغییرات کربن آلی خاک در سه کلاس با تغییرپذیری بالا ($CV > ۳۵\%$)، متوسط ($۱۵\% < CV < ۳۵\%$) و پایین ($CV < ۱۵\%$) تقسیم کرده است (۵۱). نتایج نشان داد که مقدار کربن آلی خاک در منطقه مورد مطالعه، توزیع مکانی ناهمگن دارد که ممکن است ناشی از تفاوت در کاربری اراضی، الگوی کشت و عملیات مدیریتی متفاوت و موقعیت توپوگرافی منطقه باشد. مشابه این نتایج هم توسط فلاحت‌کار و همکاران (۲۰۱۶) گزارش شده است (۱۰). هان و همکاران (۲۰۱۰) نشان دادند که ویژگی‌های خاک (مانند کربن آلی و پ-هاش) مستقل نیستند، بلکه متغیرهای با توزیع تصادفی بوده که با فرآیندهای خاک‌سازی و اقلیمی وابسته هستند (۱۴).

ویژگی مبتنی بر آماری شامل ارزیابی رابطه بین هر متغیر ورودی و متغیرهای هدف با استفاده از آمار و انتخاب آن دسته از متغیرهای ورودی است که بیش‌ترین ارتباط را با متغیرهای هدف دارند. این روش‌ها می‌توانند سریع و مؤثر باشند. می‌توان از تکنیک‌های مختلفی برای رتبه‌بندی اهمیت متغیرها، از جمله انتخاب ویژگی مبتنی بر همبستگی^۱، تجزیه و تحلیل مؤلفه‌های اصلی^۲، (۳۴)، آنالیز فاکتور^۳ و آنالیز حساسیت^۴ استفاده کرد. در این پژوهش با استفاده از الگوریتم CfsSubsetEval موجود در بسته نرم‌افزاری WEKA از روش انتخاب ویژگی مبتنی بر همبستگی استفاده گردید (۱۳). انتخاب ویژگی مبتنی بر همبستگی یک الگوریتم کاملاً خودکار است و نیازی به آستانه‌های از پیش تعیین شده یا تعداد ویژگی‌ها ندارد. این الگوریتم داده‌های متغیرها را با توجه به یک تابع ارزیابی اکتشافی مبتنی بر همبستگی، مرتب می‌کند و متغیرهای مربوط به همبستگی‌های بالا را نگه می‌دارد (۴۳). داده‌های نامناسب، با همبستگی کم نشان داده شده است. انتخاب ویژگی مبتنی بر همبستگی معمولاً بیش از نیمی از ویژگی‌ها را از بین می‌برد. در این پژوهش، این الگوریتم تعداد متغیرها را از ۱۰۱ به ۴۳ لایه کاهش داده است.

نتایج و بحث

در جدول ۱ خلاصه خصوصیات آماری کربن آلی خاک در عمق ۰ تا ۳۰ سانتی‌متر نمونه‌های تجزیه شده و برخی از متغیرهای کمکی نشان داده شده است. مقدار کربن آلی خاک در دامنه ۰/۴۶ تا ۲/۷۱ درصد و میانگین آن ۱/۲۸ درصد است. تفاوت بین حداقل و حداکثر مقدار کربن آلی خاک (۲/۲۵ درصد)

- 1- Correlation based feature selection (CFS)
- 2- Principal component analysis (PCA)
- 3- Factor analysis
- 4- Sensitivity analysis

جدول ۱- خلاصه پارامترهای آماری مربوط به کربن آلی خاک نمونه‌های تجزیه شده و برخی از متغیرهای کمکی.

Table 1. Summary of statistical parameters related to analyzed soil organic carbon samples and some auxiliary parameters.

انحراف استاندارد Std Dev	ضریب تغییرات CV (%)	چولگی Kurtosis	کشیدگی Skewness	میانگین Mean	حداکثر Max	حداقل Min	خصوصیت Properties
0.55	42.61	1.14	0.76	1.28	2.71	0.46	Soil organic carbon
220.60	20.52	-0.23	-0.77	220.60	1474	606	Band 1 (0.45-0.52 μ)
436.76	20.95	-0.42	-0.47	2084.22	2879.46	1091.85	Band 2 (0.63-0.69 μ)
639.82	20.34	-0.11	-0.27	3145.27	4499.51	1490.61	Band 4 (0.77-0.90 μ)
593.54	19.86	-0.89	2.12	2987.70	4004.92	671.31	Band 5 (1.55-1.75 μ)
139.05	8.74	0.06	0.003	1590.45	1879.56	1214.83	Chanel network base level ^۱
0.001	148.45	2.48	7.56	0.0008	0.006	0	Vector Terrain Ruggedness ^۲
0.15	11.70	1.10	3.06	1.29	1.88	0.87	Leaf Water Content ^۳
117.38	163.35	2.20	5.02	71.86	534.005	0	Vertical Distance to Channel Network ^۴
0.10	53.36	0.61	-0.08	0.20	0.48	0.02	Catchment Slope ^۵
0.09	46.21	5.02	4.90	0.20	0.60	0.10	Normalized Difference Vegetation Index ^۶
180.51	10.86	0.77	0.45	1662.28	2218.60	1379.34	Digital elevation model ^۷
1.55	46.08	-0.40	-0.54	3.37	6.21	0.09	Aspect ^۸
0.13	20.36	2.62	9.47	0.63	1.28	0.49	Reflectance Absorption Index ^۹
0.05	0.36	0.45	-1.12	12.73	12.82	12.67	Annual air temperature average ^{۱۰}
36.23	9.53	-0.42	-1.09	520	850.5	307.15	Annual rainfall average ^{۱۱}
0.05	-44.18	-0.36	2.62	-0.12	0.07	-0.3	Vegetation Index ^{۱۲}
551.44	51.97	1.75	3.30	1061.05	3235.61	379.69	Difference Vegetation Index ^{۱۳}
361.18	20.01	-1.15	6.17	1289.19	2152.46	-393.46	Canopy Index ^{۱۴}

۱- سطح پایه شبکه کانال (Chanel network base level)

۲- زبری زمین (Vector Terrain Ruggedness)

۳- مقدار آب برگ (Leaf Water Content)

۴- فاصله عمودی تا شبکه کانال (Vertical Distance to Channel Network)

۵- شیب حوزه (Catchment Slope)

۶- شاخص تفاضلی پوشش گیاهی نرمال شده (Normalized Difference Vegetation Index)

۷- مدل رقومی ارتفاع (Digital elevation model)

۸- جهت شیب (Aspect)

۹- شاخص جذب انعکاسی (Reflectance Absorption Index)

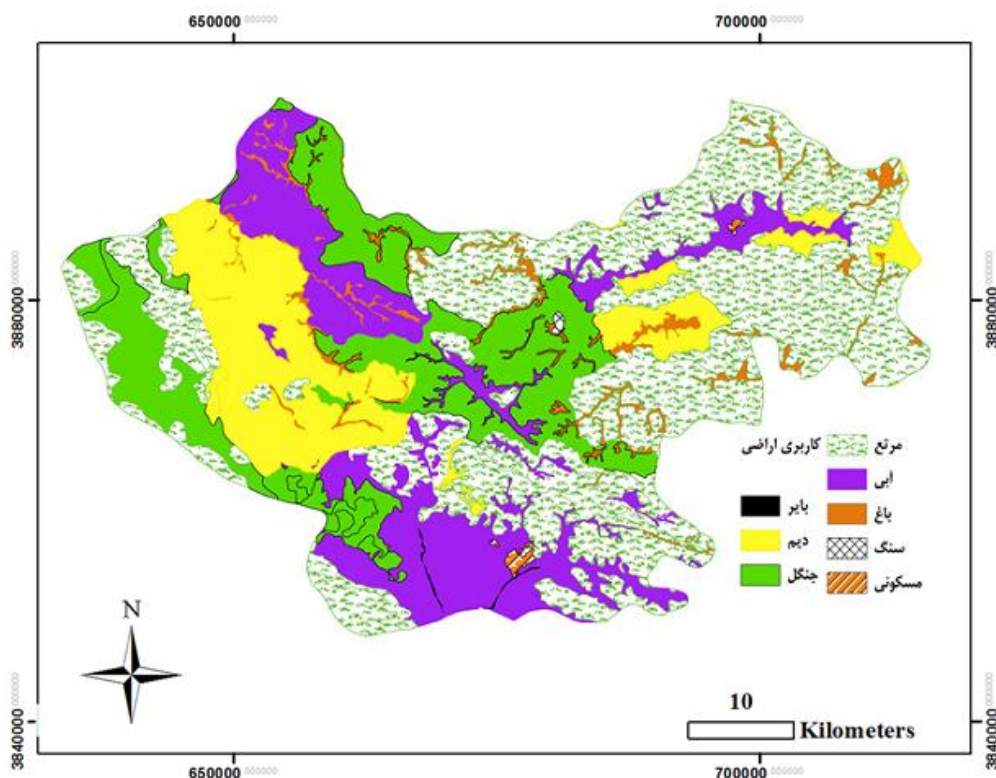
۱۰- دمای هوای سالانه (Annual air temperature)

۱۱- میانگین سالانه بارندگی (Annual rainfall average)

۱۲- شاخص پوشش گیاهی (Vegetation Index)

۱۳- شاخص تفاضلی پوشش گیاهی (Difference Vegetation Index)

۱۴- شاخص تاج پوشش (Canopy Index)



شکل ۲- کاربری‌های اراضی موجود در منطقه مورد مطالعه.

Figure 2. Land uses in the study area.

تصحیح شده (۱۰٪) و شاخص تاج پوشش (۱۰٪) به ترتیب بیشترین اثر را روی مدل‌سازی کربن آلی خاک در مدل شبکه عصبی مصنوعی داشته‌اند (شکل ۳). هم‌چنین در مدل‌سازی کربن آلی خاک توسط رگرسیون خطی چندمتغیره پارامترهای ناهمواری زمین بردار^۳ بیشترین نقش را در پیش‌بینی کربن آلی خاک در منطقه مورد مطالعه داشته است (شکل ۴)، به عبارت دیگر، ناهمواری زمین از طریق پارامترهای اقلیمی، زمینی و مدیریتی مقدار کربن آلی خاک را تحت‌تأثیر قرار داده است. بررسی نتایج انتخاب ویژگی نشان داد که متغیرهای کمکی استخراج شده از مدل رقومی ارتفاع، تصاویر ماهواره‌ای و اقلیم در پیش‌بینی کربن آلی خاک نقش داشته‌اند. به عبارتی می‌توان استنباط کرد که در منطقه مورد مطالعه، توپوگرافی، پوشش گیاهی و اقلیم از عوامل اصلی خاک‌سازی هستند و

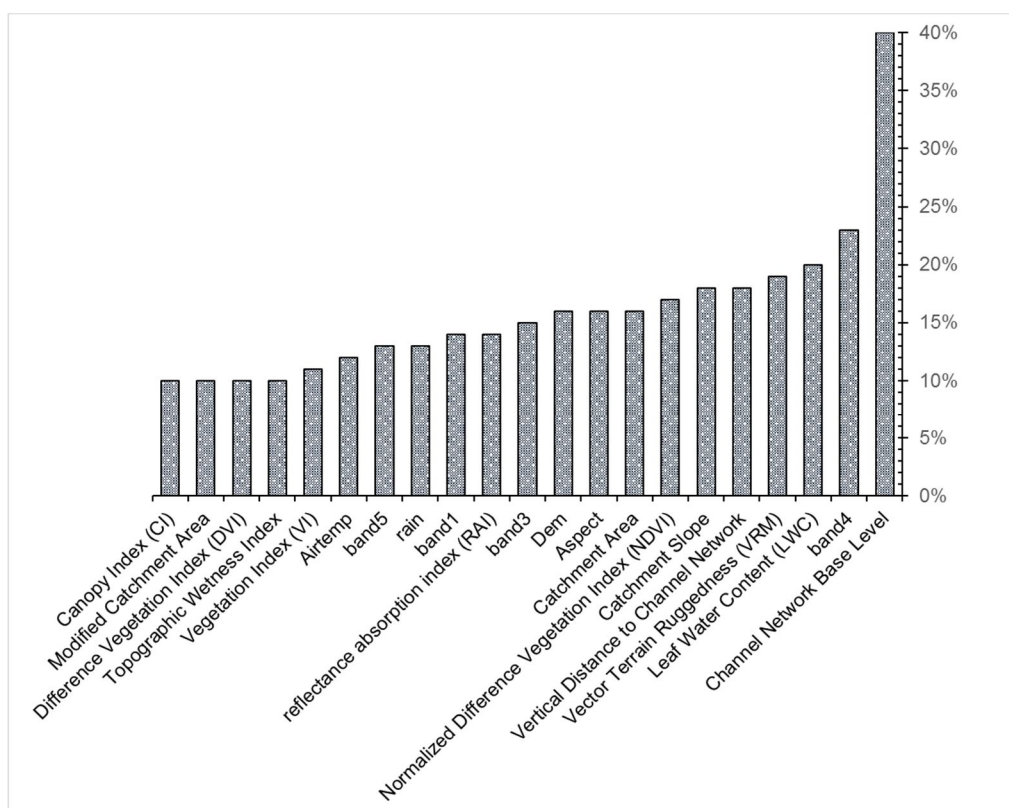
انتخاب ویژگی: در مدل‌سازی با استفاده از متغیرهای کمکی همواره برخی از پارامترها بیشترین نقش را در پیش‌بینی داشته است که این پارامترها توسط آنالیز انتخاب ویژگی قابل تشخیص می‌باشد. آنالیز انتخاب ویژگی نشان داد که متغیرهای کمکی سطح پایه شبکه کانال^۱ (۴۰٪)، باند ۴ (۲۳٪)، مقدار آب برگ^۲ (۲۰٪)، زبری زمین (۱۹٪)، فاصله عمودی تا شبکه کانال (۱۸٪)، شیب حوزه (۱۸٪)، شاخص تفاضل نرمال شده پوشش گیاهی (۱۷٪)، سطح حوزه (۱۶٪)، جهت شیب (۱۶٪)، ارتفاع (۱۶٪)، باند ۳ (۱۵٪)، شاخص جذب انعکاسی (۱۴٪)، باند ۱ (۱۴٪)، باران (۱۳٪)، باند ۵ (۱۳٪)، دمای هوا (۱۲٪)، شاخص پوشش گیاهی (۱۱٪)، شاخص خیسی توپوگرافی (۱۰٪)، شاخص تفاضل پوشش گیاهی (۱۰٪)، سطح حوزه

3- Vector Terrain Rugedness

1- Channel network base level
2- Leaf water content

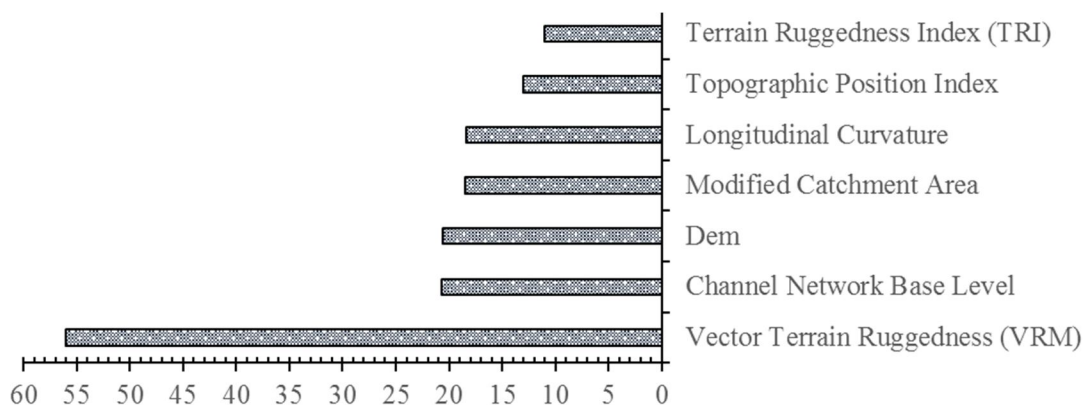
پارامترهای سرزمین (از جمله شیب، شاخص خیسی) و پوشش گیاهی (شاخص NDVI) را در اولویت تشخیص داده‌اند (۴۱). وانگ و همکاران (۲۰۱۸) در بررسی پراکنش کربن آلی خاک در مراتع استرالیا نتیجه‌ای مشابه پژوهش حاضر گزارش داده‌اند؛ به‌نحوی که باندهای ۱، ۲، ۳ و بارندگی بیش‌ترین نقش را در پیش‌بینی کربن آلی خاک داشته‌اند (۴۷). هم‌چنین در پژوهشی دیگر در چین که توسط همان پژوهشگران انجام شد، ارتفاع، شیب، شاخص خیسی و شاخص NDVI مهم‌ترین پارامترهای کمکی در پیش‌بینی کربن آلی خاک تا عمق ۱۰۰ سانتی‌متری خاک گزارش شده است (۴۷).

نقش مؤثری در توزیع مکانی کربن آلی خاک دارند. اسکندری و همکاران (۲۰۱۸) در بررسی توزیع مکانی کربن آلی خاک در شهرستان مریوان (شهرستان هم‌جوار منطقه پژوهش حاضر) نتیجه‌ای مشابه را گزارش کرده‌اند (۹)، هم‌چنین دای و همکاران (۲۰۱۴) در مدل‌سازی توزیع مکانی کربن آلی خاک توسط شبکه عصبی مصنوعی گزارش دادند که پارامترهای سرزمین (مانند ارتفاع)، اقلیم (مانند بارندگی و درجه حرارت) و تصاویر ماهواره‌ای (مانند شاخص پوشش گیاهی) بیش‌ترین نقش را ایفا کرده‌اند (۸). تقی‌زاده-مهرجردی و همکاران (۲۰۱۵) نیز در پهنه‌بندی کلاس خاک و پراکنش کربن آلی خاک در شهرستان بانه توسط روش‌های مختلف داده‌کاوی،



شکل ۳- انتخاب ویژگی مدل‌سازی کربن آلی خاک با شبکه عصبی مصنوعی (از راست به چپ: سطح پایه شبکه کانال، باند ۴، مقدار آب برگ، ناهمواری‌های زمین بردار، فاصله عمودی تا شبکه کانال، شیب حوزه، شاخص تفاضلی نرمال شده پوشش گیاهی، سطح حوزه، جهت شیب، مدل رقومی ارتفاع، باند ۳، شاخص جذب انعکاسی، باند ۱، بارندگی، باند ۵، دمای هوا، شاخص پوشش گیاهی، شاخص خیسی توپوگرافی، شاخص تفاضلی پوشش گیاهی، مساحت اصلاح‌شده حوزه، شاخص تاج پوشش گیاهی).

Figure 3. Feature selection of soil organic carbon modeling with artificial neural network.



شکل ۴- انتخاب ویژگی مدل‌سازی کربن آلی خاک با رگرسیون خطی چندمتغیره (از پایین به بالا: ناهمواری‌های زمین بردار، سطح پایه شبکه کانال، مدل رقومی ارتفاع، مساحت اصلاح‌شده حوزه، انحنای طولی، شاخص موقعیت توپوگرافی، شاخص ناهمواری زمین).

Figure 4. Feature selection of soil organic carbon modeling with multivariate linear regression.

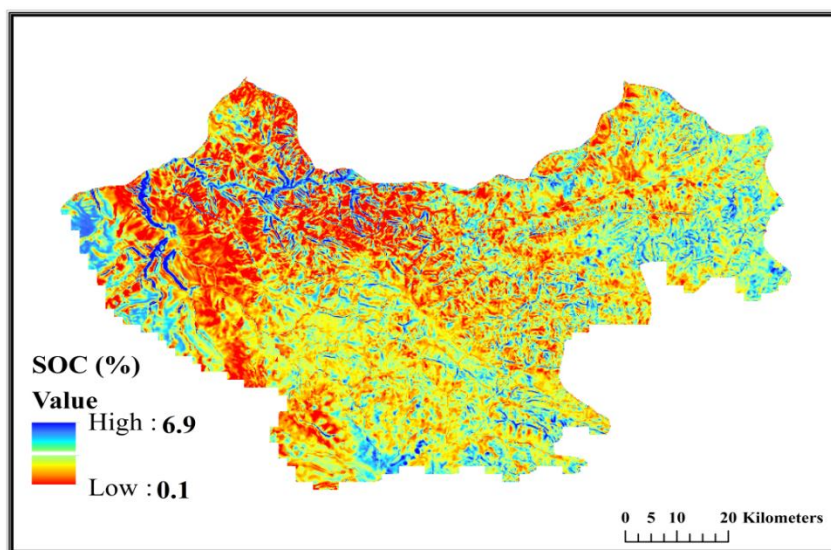
نیز نشان داده است که خاک‌های تحت پوشش جنگل دارای مقادیر بالاتر کربن آلی (۱۸۴ تن کربن بر هکتار) هستند (۳ و ۱۷). با افزایش ارتفاع به علت کاهش دما و افزایش نزولات جوی، کربن آلی خاک بیشتر از ارتفاعات پایین‌تر است که بیشتر کاربری‌های زراعی هستند (۹). چلیک (۲۰۰۵) هم در بررسی تأثیر تغییر کاربری بر میزان ماد آلی خاک در سه کاربری جنگل، مرتع و اراضی زراعی نشان داد که مقدار کربن آلی در کاربری زراعت نسبت به جنگل و مرتع دارای مقدار پایین‌تری است (۴). پارامترهای توپوگرافی از طریق عوامل فرسایش و رسوب، حفظ پوشش گیاهی، عمق توسعه ریشه، زهکشی خاک، که همگی روی مقدار تجزیه کربن آلی خاک و نیز پوشش گیاهی اثر دارند، توزیع مکانی و تشکیل خاک را تحت تأثیر قرار می‌دهد (۹).

بونفاتی و همکاران (۲۰۱۶) در مطالعه‌ای که برای نقشه‌برداری رقومی کربن آلی خاک در مناطق زراعی برزیل انجام داد پارامترهای حاصل از مدل رقومی را مهم‌ترین متغیرهای کمکی معرفی کرده است (۳) که با نتایج مطالعه چارترین و همکاران (۲۰۱۷) مشابه بوده است (۵). بر اساس آزمون انتخاب ویژگی، در میان داده‌های مربوط به تصاویر ماهواره‌ای در طول موج

توزیع مکانی کربن آلی خاک: شکل‌های ۵ و ۶ پیش‌بینی کربن آلی خاک را با مدل‌های شبکه عصبی مصنوعی و رگرسیون خطی چندمتغیره در منطقه مورد مطالعه نشان می‌دهد. مطابق با نقشه رقومی کربن آلی خاک در روش شبکه عصبی مصنوعی بیش‌ترین مقادیر کربن آلی خاک در بخش‌های غربی قرار دارد که عمدتاً دارای پوشش جنگلی، مرتع و باغ بوده و دارای فیزیوگرافی کوه و تپه هستند. مناطق تحت پوشش زراعت که در دشت‌ها قرار گرفته‌اند دارای کربن آلی خاک کم‌تری می‌باشند که ممکن است به علت کشت متراکم ذرت و گندم و همچنین دیگر محصولات جالیزی باشد؛ در کشت متراکم میزان عملیات زراعی بالاست و در نتیجه ورود بقایای آلی به خاک کاهش و همچنین کربن آلی ذخیره شده در خاک در معرض تجزیه میکروبی قرار خواهد گرفت. شکل ۷ میانگین کربن آلی خاک را در کاربری‌های زراعت دیم (۱/۳۰ درصد)، زراعت آبی (۱/۳۰ درصد)، جنگل (۱/۵۰ درصد)، مرتع (۱/۶۰ درصد) و باغ (۱/۵۰ درصد) نشان می‌دهد که از نظر آماری اختلاف معنی‌داری ندارند. میزان کربن آلی خاک در کاربری زراعت کم‌تر از بقیه کاربری‌ها است. مطالعات نقشه‌برداری رقومی کربن آلی خاک در جنوب برزیل

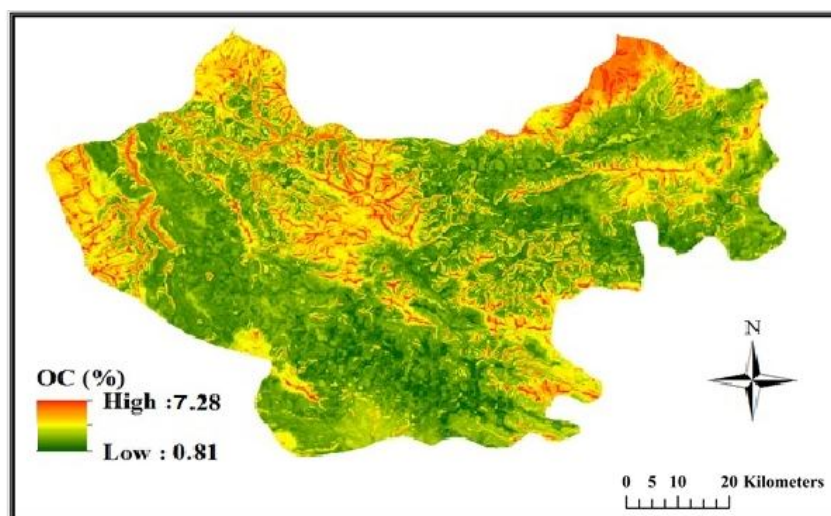
بازتاب بیشتر طیف الکترومغناطیس، بیشترین مقدار باند ۳ را دارد. مالون و همکاران (۲۰۰۹) ارتفاع بالای شبکه آبراه‌ای، شاخص قدرت جریان و باندهای ۳ و ۴ را مهم‌ترین متغیرهای کمکی در پیش‌بینی توزیع کربن آلی خاک و مقدار شاخص NDVI را در کاربری جنگل بیش از کاربری زراعت گزارش داده‌اند (۲۵).

۰/۶۰۰-۰/۵۲۵ میکرومتر و شاخص NDVI مهم‌ترین متغیرهای کمکی پیش‌بینی‌کننده کربن آلی خاک هستند (۹). بازتاب طیف الکترومغناطیسی به دلیل بالا بودن مقدار پوشش گیاهی کمتر خواهد بود. به عبارتی طول موج مادون قرمز توسط پوشش گیاهی جذب می‌شود. به همین دلیل مقادیر درجه روشنایی در باند ۳ (قرمز) کمتر است. در کاربری زراعت به دلیل پوشش گیاهی کمتر نسبت به کاربری جنگل و مرتع، جذب کمتر و



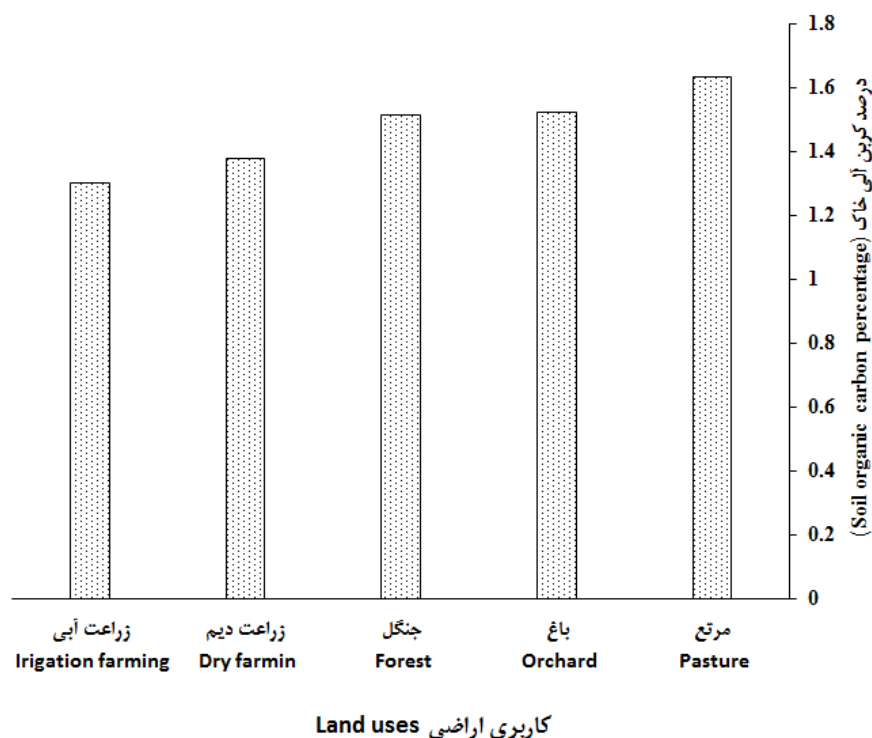
شکل ۵- توزیع مکانی کربن آلی پیش‌بینی شده خاک توسط شبکه عصبی مصنوعی.

Figure 5. Spatial distribution of predicted soil organic carbon created by artificial neural network.



شکل ۶- توزیع مکانی کربن آلی پیش‌بینی شده خاک توسط رگرسیون خطی چندمتغیره.

Figure 6. Spatial distribution of predicted soil organic carbon created by Multivariate linear regression.



شکل ۷- مقدار کربن آلی پیش‌بینی شده خاک به تفکیک کاربری‌های موجود در منطقه مورد مطالعه.
Figure 7. The amount of predicted soil organic carbon in the study area land uses.

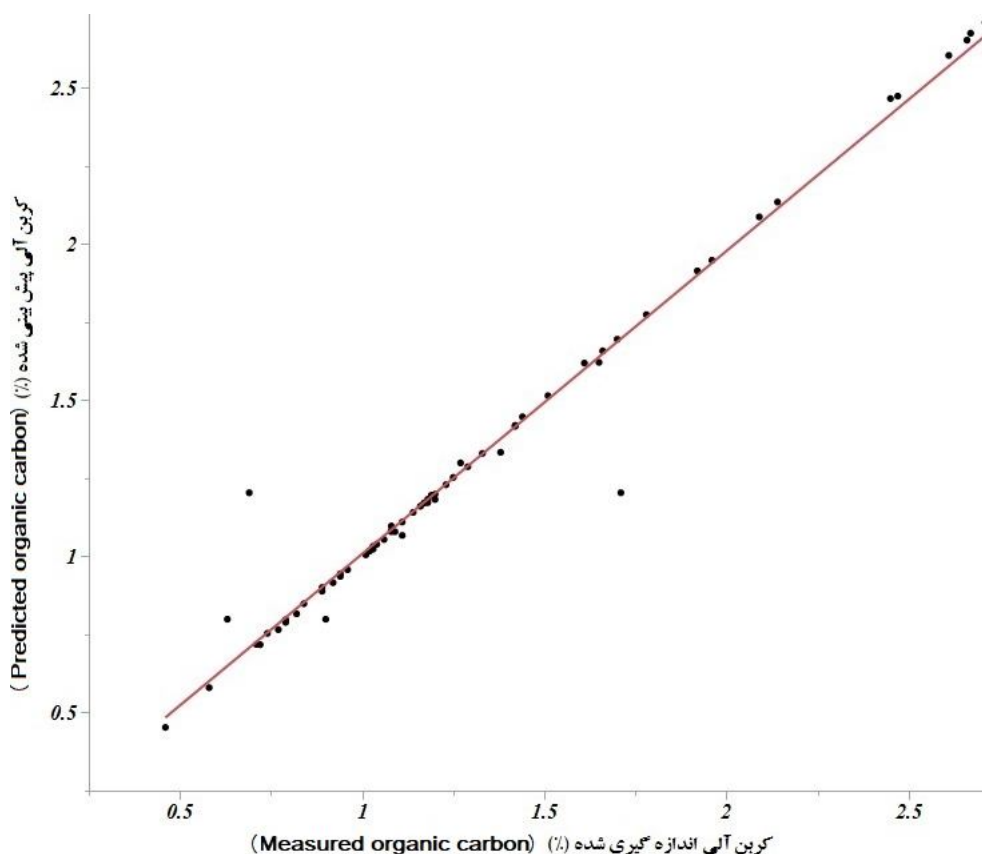
میانگین مربعات خطا $0/09$ و ضریب تبیین $0/97$ است. با توجه به وجود پارامترهای گوناگون اقلیم، توپوگرافی و پوشش گیاهی در مدل‌سازی مقدار 3 درصد از تغییرات کربن آلی خاک که توسط مدل شبکه عصبی پیش‌بینی نشده است می‌تواند مربوط به عوامل مدیریتی تأثیرگذار بر تغییرات کربن آلی خاک باشد. مدل رگرسیونی دارای ریشه میانگین مربع خطا $0/59$ و ضریب تبیین $0/60$ است، به عبارتی دیگر مدل رگرسیون خطی توانسته تنها 60 درصد از تغییرات کربن آلی سطح زمین را توجیه نماید و مدل مذکور قادر به توجیه 40 درصد تغییرات نبوده است که می‌تواند به خاطر وجود روابط غیرخطی بین پدیده‌ها باشد، برخلاف مدل رگرسیونی، شبکه عصبی مصنوعی قادر به پیش‌بینی روابط غیرخطی بین پدیده‌ها بوده و بالا بودن ضریب تبیین آن می‌تواند مربوط به این موضوع باشد، مشابه این نتیجه توسط

مقایسه مدل‌های تخمین کربن آلی خاک: شکل‌های 8 و 9 دیاگرام‌های پراکنش داده‌های تست برای مدل‌های شبکه عصبی مصنوعی و رگرسیون خطی چندمتغیره را نشان می‌دهد. همان‌طور که ملاحظه می‌شود بهترین خط برازش داده شده دارای زاویه نزدیک به 45 درجه می‌باشد که نشان از دقت بالای برآورد توسط مدل شبکه عصبی مصنوعی می‌باشد. مشابه این نتیجه در مطالعه سرمیدان و همکاران نیز تأیید گردیده است (۳۵).

نتایج ارزیابی مدل‌های شبکه عصبی مصنوعی و رگرسیون خطی در تخمین کربن آلی خاک در منطقه مورد مطالعه در جدول 2 آورده شده است و برای شناسایی دقیق‌ترین روش، اساس قضاوت کم‌ترین RMSE و بیش‌ترین R^2 در نظر گرفته می‌شود. طبق نتایج به دست آمده بهترین مدل برای پیش‌بینی کربن آلی خاک روش شبکه عصبی مصنوعی با ریشه

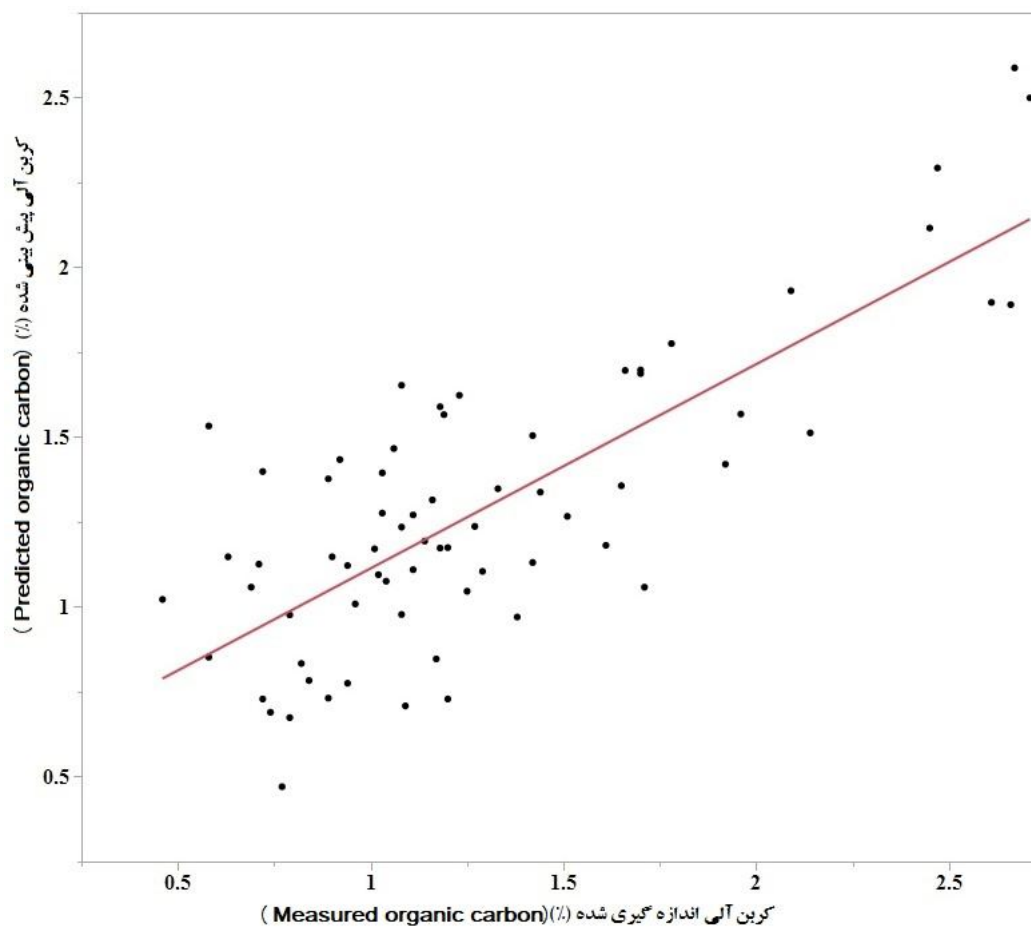
(۲۰۱۶) و مقیمی و همکاران (۲۰۱۵) است (۱۱، ۱۵ و ۲۸). به‌طورکلی با توجه به معیارهای ارزیابی می‌توان بیان داشت که نقشه‌های حاصل از شبکه عصبی مصنوعی نسبت به روش رگرسیون خطی چندمتغیره، مناسب‌ترین الگوی پراکنش کربن آلی خاک را در اختیار بهره‌برداران منطقه مورد مطالعه قرار خواهد داد. نتایج این پژوهش با نتایج لی و همکاران (۲۰۱۳)، لاهوتی و همکاران (۲۰۱۸) و سرمدیان و همکاران (۱۳۸۹) مطابقت دارد (۲۰، ۲۳ و ۳۵).

سرمدیان و همکاران (۱۳۸۹) و امینی و همکاران (۲۰۰۵) نیز گزارش شده است (۲ و ۳۵). مدل‌های رگرسیونی قادر به پیش‌بینی این روابط نیستند (۳۶). با توجه به شکل ۸ با افزایش ارتفاع، مقدار کربن آلی افزایش یافته است. لی و همکاران (۲۰۱۳) در مطالعه خود مدل‌سازی کربن آلی خاک با شبکه عصبی را نسبت به رگرسیون خطی چندمتغیره واقع‌بینانه‌تر گزارش کردند (۲۳) که این نتیجه مشابه گزارش‌های کوانگ و همکاران (۲۰۱۵)، قلیپور و همکاران



شکل ۸- دیاگرام پراکنش برای مقادیر کربن آلی اندازه‌گیری شده و پیش‌بینی شده با مدل شبکه عصبی مصنوعی.

Figure 8. Distribution diagram for organic carbon values measured and predicted by artificial neural network model.



شکل ۹- دیاگرام پراکنش برای مقادیر کربن آلی اندازه گیری شده و پیش بینی شده با مدل رگرسیون خطی چندمتغیره.

Figure 9. Distribution diagram for organic carbon values measured and predicted by multivariate linear regression model.

جدول ۲- اعتبارسنجی مدل سازی.

Table 2. Modeling validation.

RMSE	R square	نسبت f	میانگین مربعات Average of squares	مجموع مربعات Sum of squares	درجه آزادی Degrees of freedom	منبع تغییرات Source of changes	نوع مدل Model type
0.09	0.97	2344.60	19.21	19.21	1	مدل Model	شبکه عصبی مصنوعی ANN
		Prob > f	0.008	0.55	67	خطا Error	
0.59	0.60	101.09	7.38	7.38	1	مدل Model	رگرسیون خطی چندمتغیره Multivariate linear regression
		Prob > f	0.07	4.89	67	خطا Error	

نتیجه گیری

زمان، هزینه و قابلیت دسترسی تهیه نمونه از همه قسمت‌های منطقه مورد مطالعه را محدود می‌نماید، از این‌رو استفاده از متغیرهای کمکی بخشی از راه‌حل افزایش دقت تخمین ویژگی‌های خاک در این نقاط است. علی‌رغم موفقیت‌های حاصل از نقشه‌برداری رقومی خاک، هنوز زمینه‌های ناشناخته‌ای پیش‌روی علوم خاک وجود دارد تا بتواند الگوی رقومی پذیرفته شده جهانی را تدوین کند؛ به‌نحوی که نیازها و فشارهای جامعه جهانی را پاسخ دهد و جامعه را به سمت مدیریت پایدار محیط‌زیست هدایت کند.

مدل رقومی ارتفاع، تصاویر ماهواره‌ای و اقلیم و مشتقات آن‌ها برای پیش‌بینی کربن آلی خاک، به‌عنوان متغیر کمکی جهت پیش‌بینی کربن آلی خاک، مورد استفاده قرار گرفتند. این متغیرهای کمکی، تنها به‌عنوان پیش‌بینی‌کننده تجربی مطرح نیستند، بلکه آن‌ها دلایل بیوفیزیکی قابل‌توجهی به‌عنوان محرک‌های توزیع کربن خاک را ارائه می‌دهند. متغیرهای کمکی می‌تواند عرضه مواد آلی به خاک و پتانسیل تجزیه آن (اثر دما و رطوبت) را منعکس کند. ویژگی‌های مختلف اقلیمی منطقه و پارامترهای زمینی مشتق شده از DEM (هم به لحاظ منطقی و ریاضی، و هم از نظر تجربی)، دارای رابطه نزدیکی با کربن آلی خاک هستند. دو متغیر آب و هوایی یعنی دما و بارش (رطوبت) بیش‌ترین تأثیر را در توزیع کربن آلی خاک دارند. هم‌چنین مهم‌ترین ویژگی‌های DEM که در توزیع مکانی کربن آلی خاک دارای تأثیر بیش‌تر می‌باشند، عبارت از سطح پایه شبکه کانال، ناهمواری زمین بردار و ارتفاع هستند. در پژوهش حاضر متغیرهای کمکی مربوط به ویژگی‌های زمین اثر مهم‌تری نسبت به متغیرهای سنجش‌ازدور در پیش‌بینی کربن آلی خاک داشته است. رابطه بین متغیرهای کمکی و کربن آلی خاک به شرایط محیطی، قدرت

تفکیک مکانی تصویر و وسعت منطقه مورد مطالعه بستگی دارد. ارتباط کربن آلی خاک با ویژگی‌های زمینی به قدرت تفکیک مکانی مورد نظر بستگی دارد. برای تفکیک مکانی کوچک‌تر از ۱۰۰ متر، ویژگی‌های زمینی محلی (ارتفاع، شیب، جهت، انحنای پیش‌بینی‌کننده‌های خوبی هستند، که در کنار عامل فرسایش نیروی محرکه توزیع کربن آلی خاک هستند. در تفکیک مکانی بیش از ۱۰۰ متر، ویژگی‌های زمینی محلی کمتر مهم‌اند و موقعیت در چشم انداز (لنداسکیپ) مهم‌تر به نظر می‌رسد.

نتایج این مطالعه نشان داد در مناطقی که به هر دلیل امکان نمونه‌برداری در کل منطقه وجود ندارد، می‌توان از طریق متغیرهای کمکی مانند پارامترهای توپوگرافی، اقلیمی و پوشش گیاهی و با روش‌های نوین داده‌کاوی برای تخمین کربن آلی خاک بهره گرفت. مدل شبکه عصبی مصنوعی نسبت به رگرسیون خطی چندمتغیره دارای دقت بالاتری در پیش‌بینی کربن آلی است. با توجه به دقت قابل‌قبول روش‌های نقشه‌برداری خاک، امکان ارزیابی عدم قطعیت نقشه‌های تولیدی، امکان تعیین متغیرهای محیطی مهم در تغییرات خاک و هم‌چنین رشد سریع نقشه‌برداری رقومی خاک و تغییر از فاز تحقیق به فاز مدیریتی در سطح لزوم توجه بیش‌تر ارگان‌ها و متولیان اجرایی به پروژه نقشه‌برداری رقومی خاک، به‌ویژه کربن آلی خاک، در کشور را می‌طلبد. استراتژی‌های پایدار مدیریت محیط‌زیست و کاهش تغییرات آب و هوایی به درک درستی از توزیع مکانی ذخیره کربن آلی خاک نیاز دارند. بنابراین، نقشه ذخیره کربن آلی حاصل از این پژوهش می‌تواند در تشخیص مناطق دارای پتانسیل ترسیب کربن و مدیریت حاصلخیزی خاک مناسب باشد. این نوع مطالعات برای ارزیابی تأثیر تغییر کاربری اراضی بر تغییر کربن آلی خاک نیز مفید خواهد بود.

که نشان‌دهنده این است که خطاهای اندک موجود در پژوهش حاضر می‌تواند مربوط به توپوگرافی و پوشش گیاهی باشد (۲۴).

علی‌رغم انتظار ما، خطای پیش‌بینی کربن آلی خاک در این پژوهش، در منطقه با کم‌ترین تراکم نمونه بیش از منطقه با بالاترین تراکم نمونه نمی‌باشد

منابع

- Adhikari, K., Minasny, B., Greve, B.G., and Greve, M.H. 2014. Constructing a soil class map of Denmark based on the FAO legend using digital techniques. *Geoderma*. 214-215: 101-113.
- Amini, M., Afyuni, M., Fathianpour, N., Khademi, H., and Fluchler, H. 2005. Continuous soil pollution mapping using fuzzy logic and spatial interpolation. *Geoderma*. 124: 223-233.
- Bonfatti, B.R., Hartemink, A.E., Giasson, E., Tornquist, C.G., and Adhikari, K. 2016. Digital mapping of soil carbon in a viticultural region of Southern Brazil. *Geoderma*. 261: 204-221.
- Celik, I. 2005. Land-use effects on organic matter and physical properties of soil in a southern Mediterranean highland of Turkey. *Soil and Tillage Research*. 83:2. 270-277.
- Chartin, C., Stevens, A., Goidts, E., Krüger, I., Carnol, M., and van Wesemael, B. 2017. Mapping Soil Organic Carbon stocks and estimating uncertainties at the regional scale following a legacy sampling strategy (Southern Belgium, Wallonia). *Geoderma Regional*. 9: 73-86.
- Chen, S., Arrouays, D., Angers, D.A., Chenu, C., Barré, P., Martin, M.P., Saby, N.P., and Walter, C. 2019. National estimation of soil organic carbon storage potential for arable soils: A data-driven approach coupled with carbon-landscape zones. *Science of the Total Environment*. 666: 355-367.
- Conrad, O., Bechtel, B., Bock, M., Dietrich, H., Fischer, E., Gerlitz, L., Wehberg, J., Wichmann, V., and Böhner, J. 2015. System for automated geoscientific analyses (SAGA) v.2.1.4. *Geoscientific Model Development*. 8: 7. 1991-2007.
- Dai, F., Zhou, Q., Lv, Z., Wang, X., and Liu, G. 2014. Spatial prediction of soil organic matter content integrating artificial neural network and ordinary kriging in Tibetan Plateau. *Ecological Indicators*, 45: 184-194.
- Eskandari, Sh., Nabiollahi, K., and Taghizadeh-Mehrjardi, R. 2018. Digital Mapping of Soil Organic Carbon (Case Study: Marivan, Kurdistan Province). *Journal of Water and Soil*. 2: 4. 737-750 (In Persian)
- Falahatkar, S., Hosseini, S.M., Ayoubi, Sh., and Salmanmahiny, A. 2016. Predicting soil organic carbon density using auxiliary environmental variables in northern Iran. *Archives of Agronomy and Soil Science*. 62: 3. 375-393.
- Gholipour, S., Kadkhodaei, A., Makkipour, M., and Abadi-Chalaksaraee, A.R. 2016. Comparison of artificial neural network, ΔLogR and cluster analysis for the assessment of organic carbon in hydrocarbon-bearing formations. *Geoscience*. 25: 98. 147-158 (In Persian)
- 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.
- Hall, M., Frank, E., Holmes, G., Pfahringer, B., Reutemann, P., and Witten, I.H. 2009. The WEKA data mining software: an update. *ACM SIGKDD explorations newsletter*. 11: 1. 10-18.
- Han, F., Hu, W., Zheng, J., Du, F., and Zhang, F. 2010. Estimating soil organic carbon storage and distribution in a catchment of Loess Plateau, China. *Geoderma*. 154: 3-4. 261-266.

15. 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.
16. Kumar, S., and Lal, R. 2011. Mapping the organic carbon stocks of surface soils using local spatial interpolator. *Journal of Environmental Monitoring*, 13: 11. 3128-3135.
17. Kumar, S., Lal, R., and Liu, D. 2012. A geographically weighted regression kriging approach for mapping soil organic carbon stock. *Geoderma*. 189: 627-634.
18. Kurdistan Province Management and Planning Organization, Deputy of Statistics and Information. 2018. Statistical yearbook of Kurdistan province in 2017. Country Planning and Budget Organization. (In Persian)
19. Lahooti, P., Emadi, S.M., Bahmanyar, M.A., and Ghajar-Sepanlou, M. 2019. Soil Organic Carbon Mapping By Geostatistics and Artificial Neural Network Methods (Kohgiluyeh and Boyer-Ahmad Province). *Journal of Water and Soil*. 32: 6. 1135-1148. (In Persian)
20. Lal, R. 2004. Carbon sequestration in soils of central Asia. *Land Degradation and Development*. 15: 6. 563-572.
21. Li, Q.Q., Yue, T.X., Wang, C.Q., Zhang, W.J., Yu, Y., Li, B., Yang, J., and Bai, G.C. 2013. Spatially distributed modeling of soil organic matter across China: An application of artificial neural network approach. *Catena*. 104: 210-218.
22. Mahab Gostar Zagros Consulting Engineers. Detailed-executive watershed management studies of Khamesan watershed. Volume Five: Land Suitability Assessment. 2014. Forests, Rangelands and Watershed Management Organization, General Department of Natural Resources of Kurdistan Province. Detailed watershed management study report. (In Persian)
23. Mahmoudzadeh, H., Matinfar, H.R., Taghizadeh-Mehrjardi, R., and Kerry, R., 2020. Spatial prediction of soil organic carbon using machine learning techniques in western Iran. *Geoderma Regional*. 21:e00260. doi.org/10.1016/j.geodrs.2020.e00260.
24. Malone, B.P., McBratney, A., Minasny, B., and Laslett, G. 2009. Mapping continuous depth functions of soil carbon storage and available water capacity. *Geoderma*. 154: 1-2. 138-152.
25. Martin, M., Wattenbach, M., Smith, P., Meersmans, J., Jolivet, C., Boulonne, L., and Arrouays, D. 2010. Spatial distribution of soil organic carbon stocks in France. *Biogeosciences*. 8: 1053-1065.
26. McBratney, A.B., Santos, M.M., and Minasny, B. 2003. On digital soil mapping. *Geoderma*. 117: 1-2. 3-52.
27. Moghimi, S., Parvizi, Y., Mahdian, M.H., and Masih-Abadi M.H. 2015. Comparative application of multiple linear regression and artificial neural networks for simulating the effects of topographic factors on organic carbon changes in soil. *Journal of Watershed Engineering and Management*, 6: 4. 312-322. (In Persian)
28. Mollazade, K., Omid, M., and Arefi, A. 2012. Comparing data mining classifiers for grading raisins based on visual features. *Computers and electronics in agriculture*. 84: 124-131.
29. Mulder, V., De Bruin, S., Schaeppman, M., and Mayr, T. 2011. The use of remote sensing in soil and terrain mapping a review. *Geoderma*. 162: 1-2. 1-19.
30. Nabiollahi, K., Eskandari, Sh., Taghizadeh-Mehrjardi, R., Kerry, R., and Triantafalis, J. 2019. Assessing soil organic carbon stocks under land-use change scenarios using random forest models. *Carbon Management*. 10: 63-77.
31. Narjary, B., Meena, M.D., Kumar, S., Kamra, S.K., Sharma, D.K., and Triantafalis, J. 2019. Digital mapping of soil salinity at various depths using an EM38. *Soil Use and Management*. 35: 2. 232-244.

32. Nelson, D.W., and Sommers, L.E. 1975. A rapid and accurate procedure for estimation of organic carbon in soil. *Proceedings of the Indiana Academy of Science*. 84: 456-462.
33. Omid, M., Mahmoudi, A., and Omid, M.H. 2010. Development of pistachio sorting system using principal component analysis (PCA) assisted artificial neural network (ANN) of impact acoustics. *Expert Systems with Applications*. 37: 10. 7205-7212.
34. Pilevar, Shahri, A.R., Ayoubi, Sh., and Khademi, H. 2011. Regression (MLR) Models to Predict Soil Organic Carbon Using Digital Terrain Analysis (Case Study: Zargham Abad Semirrom, Isfahan Province). *Journal of Water and Soil*. 24: 6. 1151-1163 (In Persian)
35. Sarmadian, F., Taghizadeh-Mehrjardi, R., Mohammad-Asgari, H., and Akbarzadeh, Ali. 2010. A Comparison of Neuro-Fuzzy, Artificial Neural Network and Multivariate Regression for Prediction of some Soil Properties (Case Study: Golestan Province). *Iranian Soil and Water Research*. 41: 2. 211-220. (In Persian)
36. Sarvati, M.R., and Ebrahimi, A. 2016. Geographical, climatic and climatic characteristics of Kamyaran region. *The First International Conference on Natural Hazards and Environmental Crises in Iran, Solutions and Challenges*. Ardabil, <https://civilica.com/doc/548900/certificate/print/>.
37. Schillaci, C., Acutis, M., Lombardo, L., Lipani, A., Fantappie, M., Märker, M., and Saia, S. 2017. Spatio-temporal topsoil organic carbon mapping of a semi-arid Mediterranean region: The role of land use, soil texture, topographic indices and the influence of remote sensing data to modelling. *Science of the Total Environment*. 601: 821-832.
38. Shelukindo, H.B., Semu, E., Singh, B., and Munishi, P. 2014. Predictor variables for soil organic carbon contents in the Miombo woodlands ecosystem of Kitonga forest reserve, Tanzania. *International Journal of Agricultural Sciences*. 4: 7. 222-231.
39. Taghizadeh-Mehrjardi, R., Minasny, B., Sarmadian, F., and Malone, B. 2014. Digital mapping of soil salinity in Ardakan region, central Iran. *Geoderma*. 213: 15-28.
40. Taghizadeh-Mehrjardi, R., Toomanian, N., Khavaninzadeh, A., Jafari, A., and Triantafilis, J. 2016. Predicting and mapping of soil particle-size fractions with adaptive neuro-fuzzy inference and ant colony optimization in central Iran. *European Journal of Soil Science*. 67: 6. 707-725.
41. Taghizadeh-Mehrjardi, R., Neupane, R., Sood, K., and Kumar, S. 2017. Artificial bee colony feature selection algorithm combined with machine learning algorithms to predict vertical and lateral distribution of soil organic matter in South Dakota, USA. *Carbon Management*, 8: 3. 277-291.
42. Taghizadeh-Mehrjardi, R., Nabiollahi, K., Minasny, B., and Triantafilis, J. 2015. Comparing data mining classifiers to predict spatial distribution of USDA-family soil groups in Baneh region, Iran. *Geoderma*. 253: 67-77.
43. Vågen, T.G., and Winowiecki, L.A. 2013. Mapping of soil organic carbon stocks for spatially explicit assessments of climate change mitigation potential. *Environmental Research Letters*. 8: 1. 015011.
44. Viscarra-Rossel, R., and Behrens, T. 2010. Using data mining to model and interpret soil diffuse reflectance spectra. *Geoderma*. 158: 1-2. 46-54.
45. Wang, S., Zhuang, Q., Wang, Q., Jin, X., and Han, C. 2017. Mapping stocks of soil organic carbon and soil total nitrogen in Liaoning Province of China. *Geoderma*. 305: 250-263.
46. Wang, B., Waters, C., Orgill, S., Cowie, A., Clark, A., Liu, D.L., Simpson, M., McGowen, I., and Sides, T. 2018. Estimating soil organic carbon stocks using different modelling techniques in the semi-arid rangelands of eastern Australia. *Ecological Indicators*. 88: 425-438.

47. Were, K., Bui, D.T., Dick, Ø.B., and Singh, B.R. 2015. 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*. 52: 394-403.
48. Wiesmeier, M., Spörlein, P., Geuß, U., Hangen, E., Haug, S., Reischl, A., Schilling, B., von Lützow, M., and Kögel-Knabner, I. 2012. Soil organic carbon stocks in southeast Germany (Bavaria) as affected by land use, soil type and sampling depth. *Global Change Biology*. 18: 7. 2233-2245.
49. Wilding, L. 1985. Spatial variability: its documentation, accommodation and implication to soil surveys, in *Proceedings Soil spatial variability. Workshop*, Pp: 166-194.
50. Yang, R.M., Zhang, G.L. Liu, F., Lu, Y.Y., Yang, F., Yang, F., Yang, M., Zhao, Y.G., and Li, D.C. 2016. Comparison of boosted regression tree and random forest models for mapping topsoil organic carbon concentration in an alpine ecosystem. *Ecological Indicators*. 60: 870-878.



Digitization of soil organic carbon with artificial neural network and multivariate linear regression in Kurdistan province

H. Mahmoudzadeh¹, *H.R. Matinfar² and R. Taghizadeh-Mehrjerdi³

¹Ph.D. Graduate, Dept. of Soil Science, Lorestan University,

²Associate Prof., Dept. of Soil Science, Lorestan University,

³Associate Prof., Dept. of Agricultural Science and Natural Resources, Ardakan University

Received: 06.13.2020; Accepted: 12.20.2020

Abstract

Background and Objectives: Organic carbon plays a vital role in climate control and environmental sustainability by retaining a significant portion of its mineralizable organic forms in the soil. Organic carbon also has a key effect on physicochemical and biological properties of soil; in a way, it is called as an indicator of soil health. Therefore, the study of spatial distribution of soil organic carbon to identify areas with carbon sequestration potential is one of the requirements of soil management planning and climate control policy through agricultural activities. Conventional methods for estimating soil organic carbon are costly and time consuming and cannot be replicated and generalized to similar points. In recent years, with the advancement of technology and the growing human need for access to accessible information and cost savings, through data mining and with the help of satellite images and auxiliary topographic variables, digitization of soil properties, including Organic carbon is made possible. Digital soil mapping is the development of a numerical or statistical model of the relationship between environmental variables and soil properties that is used for large geographic data to produce a digital map. The three main goals of digital soil mapping are: 1) inferring the relationship between environmental variables and soil properties, 2) producing and presenting data that better illustrate soil-geolocation, and 3) explicitly applying expert knowledge in design. They are models. Digital mapping also provides potential advances in soil pedology and geography by providing insights into landfilling processes.

Materials and Methods: This study was conducted in Kamyaran city of Kurdistan province in order to predict soil organic carbon. In this study, 110 soil samples were randomly analyzed from different land uses. To better predict the spatial distribution of soil organic carbon in the study area, 101 auxiliary variables extracted from digital elevation model, satellite images and climatic variables have been used. Prediction of soil organic carbon was modeled with two models of multivariate linear regression and artificial neural network in Jump software environment.

Results: The results showed that the amount of soil organic carbon in the western and northwestern parts of the study area is the highest, which includes areas with forest and rangeland cover. Auxiliary variables: base level of canal network (40%), band 4 (23%), leaf water content (20%), ground roughness (19%), vertical distance to canal network (18%), slope (18%), Normalized vegetation differentiation index (17%), area level (16%), slope direction (16%), height (16%), band 3 (15%), reflective absorption index (14%), band 1 (14%), Rain (13%), band 5 (13%), air temperature (12%), vegetation index (11%), topographic wetness index (10%) and vegetation difference index (10%) had the greatest effect On the modeling of soil organic carbon in the artificial neural network model. The results of modeling validation

* Corresponding Author; Email: matinfar.h@lu.ac.ir

showed that the artificial neural network performed better ($R^2 = 0.97$) than the multivariate linear regression ($R^2 = 0.59$) prediction of soil organic carbon in the study area.

Conclusion: The results of this study showed that the distribution of organic carbon is more influenced by topographic and climatic factors. In areas where sampling is not possible in the whole area for any reason, it can be used through auxiliary variables such as topographic, climatic and vegetation parameters and with modern data mining methods to estimate soil organic carbon.

Keywords: Data Mining, Digital Soil Mapping, Soil Organic Carbon