



دانشگاه علوم کشاورزی و منابع طبیعی گزن

نشریه پژوهش‌های علوم و فناوری چوب و جنگل

جلد بیست و هفتم، شماره سوم، ۱۳۹۹

۱۹-۳۶

<http://jwfst.gau.ac.ir>

DOI: 10.22069/jwfst.2020.17946.1870

کارایی شبکه عصبی مصنوعی بهینه در مدل‌سازی کربن آلی خاک مبتنی بر داده‌های میدانی و تصاویر Sentinel-2 در ارسباران

محسن لطفی^۱، * سیدیوسف عرفانی‌فرد^۲، فرشاد امیراصلانی^۲ و علی کشاورزی^۳

^۱دانشجوی کارشناسی‌ارشد گروه سنجش از دور و GIS، دانشکده جغرافیا، دانشگاه تهران، تهران، ایران،

^۲دانشیار گروه سنجش از دور و GIS، دانشکده جغرافیا، دانشگاه تهران، تهران، ایران،

^۳استادیار گروه علوم و مهندسی خاک، پردیس کشاورزی و منابع طبیعی، دانشگاه تهران، کرج، ایران

تاریخ دریافت: ۱۳۹۹/۰۲/۰۸؛ تاریخ پذیرش: ۱۳۹۹/۰۳/۱۵

چکیده

سابقه و هدف: خاک بزرگ‌ترین منبع ذخیره کربن موجود در بوم‌سازگان‌های زمینی هستند که بیش‌ترین سهم از کل ذخایر جهانی کربن زمینی را در خود جای دادند. نقشه‌برداری دقیق اطلاعات توزیع مکانی ذخیره کربن آلی خاک (SOC) یک پیش‌نیاز کلیدی جهت مدیریت منابع خاک و حفاظت از محیط‌زیست است. توسعه سریع علم سنجش از دور و استفاده از تصاویر ماهواره‌ای امکان نظارت بر ذخیره SOC در مقیاس بزرگ را فراهم می‌کند. امکان برآورد SOC یکی از موضوعات پیش روی پژوهشگران بوده است که در برخی موارد از شبکه عصبی مصنوعی برای این موضوع استفاده شده است هر چند تعیین مقادیر بهینه مؤلفه‌های مؤثر در آن دشوار است. در برخی مطالعات از الگوریتم ژنتیک برای بهینه‌سازی وزن‌های اولیه شبکه عصبی و بهبود پیش‌بینی متغیرهای خروجی استفاده شده است. اگرچه کارایی این روش در برآورد SOC با داده‌های سنجش از دور کم‌تر مورد بررسی قرار گرفته است. در این پژوهش اثر الگوریتم ژنتیک بر بهبود عملکرد شبکه عصبی مصنوعی در پیش‌بینی SOC با استفاده از تصاویر ماهواره Sentinel-2 در ناحیه رویشی ارسباران مورد بررسی قرار گرفته است.

مواد و روش‌ها: برای این منظور نمونه‌برداری از خاک با استفاده از روش نمونه‌برداری تصادفی طبقه‌بندی شده بر اساس انواع کاربری اراضی در ۴۶ نقطه و در عمق صفر تا ۱۵ سانتی‌متر انجام شد. SOC با استفاده از روش والکلی بلاک اندازه‌گیری شد. برای برازش مدل بین کربن آلی اندازه‌گیری‌شده در آزمایشگاه با ۹ شاخص طیفی و ۳ باند تصویر ماهواره‌ای که به‌طور مستقیم وارد مدل‌سازی شدند، از دو روش شبکه عصبی مصنوعی و شبکه عصبی مصنوعی بهینه‌شده با الگوریتم ژنتیک استفاده گردید. برای ارزیابی کارایی مدل‌ها از روش اعتبارسنجی متقابل

* مسئول مکاتبه: erfanifard@ut.ac.ir

(Cross Validation) استفاده شد. در نهایت مدل‌های به‌دست آمده با شاخص‌های آماری جذر میانگین مربعات خطا (RMSE)، نسبت عملکرد به انحراف (RPD)، ضریب همبستگی اسپیرمن (r)، ضریب تبیین (R^2) و هم‌چنین آزمون تی جفتی مورد ارزیابی قرار گرفتند.

یافته‌ها: نتایج نشان داد SOC برآوردی با شبکه عصبی مصنوعی بهینه‌شده با الگوریتم ژنتیک ($RMSE=1/07$ درصد، $RPD=1/89$ درصد، $R^2=0/76$) دقت بیش‌تری نسبت به نتایج شبکه عصبی مصنوعی ($RMSE=1/51$ درصد، $RPD=1/34$ درصد، $R^2=0/58$) داشت. هم‌چنین بهبود ضریب همبستگی اسپیرمن برای SOC واقعی و برآورد شده با شبکه عصبی مصنوعی بهینه‌شده ($r=0/87$) و SOC برآورد شده با شبکه عصبی مصنوعی ($r=0/76$) مشاهده گردید. SOC واقعی با SOC برآوردی با شبکه عصبی مصنوعی بهینه‌شده اختلاف معنی‌داری نداشت ($p\text{-value}=0/21$) ولی با SOC برآوردی با شبکه عصبی مصنوعی متفاوت بود ($p\text{-value}=0/02$). علاوه بر این، نتایج نشان داد که شاخص TSAVI بیش‌ترین ضریب همبستگی اسپیرمن ($0/565$) و شاخص BI2 کم‌ترین ضریب همبستگی اسپیرمن ($0/196$) را با کربن آلی خاک دارند.

نتیجه‌گیری: به‌طورکلی، نتیجه‌گیری شد که استفاده از الگوریتم ژنتیک در انتخاب مؤلفه‌های بهینه شبکه عصبی مصنوعی منجر به بهبود عملکرد این روش مدل‌سازی در برآورد نقطه‌ای SOC با استفاده از تصاویر ماهواره Sentinel-2 در منطقه مورد مطالعه شده است. هم‌چنین با توجه به نتایج به‌دست آمده، کارایی تصاویر ماهواره Sentinel-2 در برآورد SOC در منطقه مورد مطالعه تأیید شد.

واژه‌های کلیدی: الگوریتم ژنتیک، تصاویر ماهواره Sentinel-2، شبکه عصبی مصنوعی، کربن آلی خاک

مقدمه

مکانی بسیار زیاد SOC، تنوع گونه‌های خاک که باعث اریبی برآوردها می‌شود و تغییر کاربری پوشش گیاهی و زمین هستند (۱۵). این اختلاف تنوع ذخیره SOC در کاربری‌های مختلف توسط بسیاری از پژوهشگران مطالعه شده است. برای مثال دنگیز و همکاران (۲۰۱۵) در مطالعه‌ای با هدف تعیین تأثیر نوع خاک و پوشش کاربری اراضی در ذخیره SOC در یک حوزه آبخیز در ترکیه پرداختند و تأیید کردند از بین چهار کاربری مختلف جنگل، مرتع، باغ و زمین‌های زیر کشت؛ جنگل بیش‌ترین SOC ($6/30$ درصد) را داشت (۱۱). هم‌چنین آمانئول و همکاران

خاک می‌تواند دو تا سه برابر اتمسفر در خود کربن جای می‌دهد و از طریق ترسیب کربن نقش بسیار مهمی در مدیریت تغییرات اقلیم ایفا می‌کند. اگرچه کربن موجود در خاک می‌تواند موجب افزایش حاصلخیزی خاک گردد. هم‌چنین کربن آلی خاک^۱ که سهم اندکی در لایه‌های سطحی خاک دارد، برای رشد گیاهان دارای اهمیت است (۳۴). اسوارن^۲ یک سری موانع را برای برآورد دقیق محتوای کربن خاک در مقیاس جهانی گزارش داد که شامل تغییرپذیری

1- Soil organic carbon: SOC

2- Eswaran

خصوصیات خاک مورد تأیید قرار گرفته است (۱۳). مطالعات پیشین تأیید کرده که شاخص‌های پوشش گیاهی به‌دست‌آمده از سنجش‌ازدور قابلیت لازم را برای ایجاد مدل‌هایی برای شبیه‌سازی پویایی کربن دارا هستند. در این راستا تکنیک‌های آماری متنوعی در نقشه‌برداری رقومی SOC استفاده شده است که عبارت‌اند از رگرسیون خطی چندگانه، رگرسیون حداقل مربعات جزئی، مدل‌های خطی تعمیم‌یافته، مدل‌های ترکیبی خطی، رگرسیون وزن‌دار جغرافیایی^۳ و رگرسیون کریجینگ (۲۴، ۲۵، ۲۶، ۳۰، ۴۱ و ۴۳). برخی پژوهش‌های ذکرشده نتیجه گرفتند که روابط بین محتوای کربن آلی خاک و عوامل محیطی معمولاً غیرخطی هستند (۳۰). از همین‌رو اخیراً چندین مطالعه از روش‌های جدیدی در یادگیری ماشین استفاده کرده‌اند، مانند شبکه عصبی مصنوعی^۴، ماشین‌های بردار پشتیبان^۵، درخت رگرسیون تقویت‌شده^۶ و جنگل تصادفی^۷ تا نقشه SOC را تهیه کنند (۱، ۱۸، ۲۸، ۳۸ و ۳۹). برای مثال چن و همکاران (۲۰۰۰) رابطه بین SOC سطحی خاک و مقادیر بازتاب باندهای آبی، سبز، قرمز تصاویر ماهواره لندست ۵ را بر اساس داده‌های ۲۸ نقطه نمونه‌برداری از منطقه مورد مطالعه واقع در ایالت جورجیا تجزیه و تحلیل کردند و محتویات SOC را با استفاده از معادله لگاریتمی خطی تخمین زدند (۷). ور و همکاران (۲۰۱۵) به توسعه و ارزیابی و مقایسه عملکرد مدل‌های رگرسیون بردار پشتیبان^۵، شبکه عصبی مصنوعی و جنگل تصادفی در پیش‌بینی و ارزیابی

(۲۰۱۸) بر روی SOC ذخیره‌شده در چهار نوع کاربری متفاوت جنگل طبیعی، جنگل دست کاشت، مراتع و زمین‌های کشاورزی در حوزه آبخیز رود نیل در اتیوپی به مطالعه پرداختند که جنگل بیش‌ترین SOC (۴/۵۸ درصد) را داشت (۲). ساسانی‌فر و همکاران (۲۰۱۸) در ناحیه رویشی ارسباران به مطالعه ویژگی‌های خاک مانند کربن آلی خاک، هدایت الکتریکی، رس، اسیدیته، شن، نیتروژن کل با ۴۴ نمونه خاک پرداختند و گزارش کردند که این مشخصه‌ها از توزیع یکنواختی برخوردار نیستند (۳۵). در حالت کلی، برآورد ویژگی‌های کمی و کیفی خاک یک رویه دشوار است. از این‌رو، نیاز به مجموعه داده‌هایی وجود دارد که بتوانند اطلاعات قابل‌قبولی را در مورد ویژگی‌های خاک از جمله SOC ارائه دهند تا در مدیریت پایدار بوم‌سازگان‌ها به کار روند.

از آن‌جاکه روش‌های مرسوم میدانی برای نظارت بر SOC وقت‌گیر و پرهزینه است، پژوهشگران اجرای رویکردهای جایگزین را که می‌توانند در شرایط مختلف و بر روی انواع خاک مورد استفاده قرار گیرند، بررسی کردند. یکی از این رویکردهای جایگزین، تکنیک‌های سنجش‌ازدور است که به‌عنوان روشی سریع، مقرون‌به‌صرفه، غیرمخرب، با دید گسترده برای برآورد ویژگی‌های مختلف خاک، از جمله SOC معرفی شده است. امروزه سامانه اطلاعات مکانی^۱ (GIS) و سنجش‌ازدور اساس نقشه‌برداری رقومی خاک^۲ (DSM) را تشکیل می‌دهند (۲۹ و ۴۰). هم‌چنین تکنیک‌های سنجش‌ازدور به‌عنوان روشی کارآمد برای ارزیابی و نقشه‌برداری از برخی

3- Geographically weighted regression
4- Artificial neural networks
5- Support vector machines
6- Boosted regression trees
7- Random forests
8- Support vector regression

1- Geographic information system
2- Digital soil mapping

SOC در جنگل‌های مائو در کنیا پرداختند که ضریب تبیین شبکه عصبی ۰/۶ شد (۳۹). کایستالدی و همکاران (۲۰۱۶) قابلیت تصاویر ماهواره‌ای Landsat-8 و Sentinel-2 و سنجنده Hyperion را برای پیش‌بینی SOC و بافت خاک با استفاده از کتابخانه طیفی آماری پوشش سطح زمین اروپا^۱ (LUCAS) و مجموعه داده‌های دیگر از قسمت مرکزی و جنوبی ایتالیا (PONMAC) بررسی کردند که دقت RPD^۲ در برآورد SOC، ۱/۳۵ به دست آمد (۵). قلی‌زاده و همکاران (۲۰۱۸) تصاویر ماهواره Sentinel-2، تصاویر ابرطیفی هوابرد (سنجنده‌های CASI و SASI) و طیف‌سنجی آزمایشگاهی را برای پایش و نقشه‌برداری بافت خاک و SOC در ۴ منطقه مورد مطالعه مختلف در جمهوری چک به کار بردند. آن‌ها از ۱۸ شاخص طیفی به‌عنوان متغیر مستقل برای آموزش مدل پیش‌بینی استفاده کردند و میانگین RPD به دست آمده ۱/۶۵ بود (۱۸). روش‌های یادگیری ماشین بر کاستی روش‌های آماری پارامتری و ناپارامتری، مانند همبستگی مکانی، غیرخطی بودن و برآزش غلبه می‌کنند و این باعث بهبود دقت پیش‌بینی روش‌های مذکور می‌شود. اگرچه روش‌های یادگیری ماشین از جمله الگوریتم پرکاربرد شبکه عصبی مصنوعی دارای مؤلفه‌هایی است که تعیین مقدار بهینه آن‌ها تأثیر زیادی بر کارایی الگوریتم دارد (۸).

مطالعات گذشته نشان می‌دهند که بهینه‌سازی شبکه عصبی مصنوعی توسط الگوریتم ژنتیک (GA) یک راه‌حل مؤثر با زمان محاسباتی کم‌تر نسبت به شبکه عصبی مصنوعی است (۱۷)؛ زیرا نواقص ناشی از طراحی شبکه‌های عصبی مصنوعی نسبت به

1- Land use cover area frame statistical survey
2- Ratio of standard deviation to RMSE of cross validation

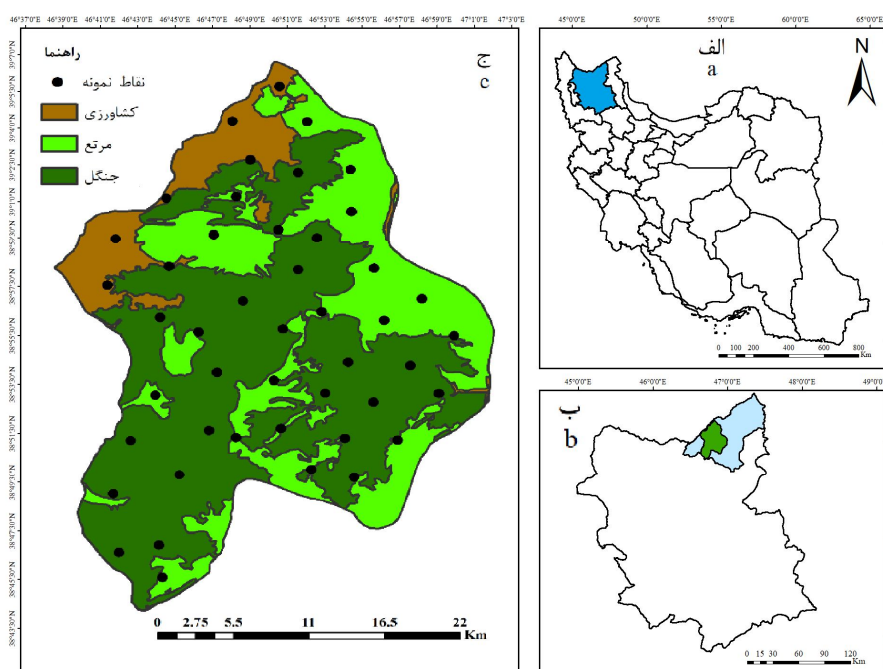
استفاده از شبکه عصبی مصنوعی به‌تنهایی از بین می‌رود. GA جزء معتبرترین الگوریتم‌های بهینه‌سازی برای جستجوی وزن‌های شبکه عصبی به حساب می‌آید، درحالی‌که الگوریتم‌های جستجوی محلی، مانند گرادیان نزولی، به دلیل افتادن در دام بهینه محلی (انتخاب یک نقطه تصادفی در فضا و پیدا کردن بیش‌ترین امتیاز محلی در نزدیکی آن منطقه) مناسب نیستند (۸). هارفام و همکاران (۲۰۰۴) نشان دادند هنگامی که یک GA برای جستجوی وزنه‌ای بهینه یا نزدیک بهینه استفاده می‌شود، افتادن در دام محلی برداشته می‌شود اما هیچ تضمینی برای همگرایی به نقطه بهینه سراسری (بهترین نقطه در فضا) وجود ندارد (۲۰). در چندین پژوهش از GA برای بهینه‌سازی وزن‌های شبکه عصبی مصنوعی استفاده کرده‌اند. برای مثال لیو و همکاران (۲۰۰۴) به بررسی دقت طبقه‌بندی پوشش زمین با استفاده از تصاویر چند طیفی پرداختند. آن‌ها نشان دادند که شبکه عصبی مصنوعی پرسپترون بهینه‌شده با الگوریتم ژنتیک عملکرد بهتری نسبت به شبکه عصبی مصنوعی پرسپترون بهینه‌شده با گرادیان نزولی داشت (۲۷). فقیه (۲۰۱۱) در مطالعه‌ای کارایی شبکه عصبی مصنوعی در برآورد مکانی بارش ماهانه را در استان کردستان مورد بررسی قرار داد. برای طراحی ساختار مدل در هر ایستگاه، با تغییر مؤلفه‌های قابل تنظیم شبکه عصبی مصنوعی مختلف ساخت و سپس شبکه عصبی را با الگوریتم ژنتیک بهینه‌سازی کرد (۱۶). نتایج نشان داد که بهینه کردن شبکه عصبی مصنوعی با الگوریتم ژنتیک برای پیدا کردن وزن‌های بهینه دقت بالایی را نسبت به خود شبکه عصبی مصنوعی دارد. پیشینه پژوهش بررسی شده نشان داد تاکنون در مورد پیش‌بینی SOC با شبکه عصبی مصنوعی روی داده سنجش‌ازدور، از الگوریتم ژنتیک برای بهینه کردن وزن‌های شبکه عصبی مصنوعی استفاده نشده است.

شرقی با مساحت حدوداً ۸۰ هزار هکتار قرار گرفته است (شکل ۱). این منطقه به دلیل برخورداری از بوم‌سازگان‌های طبیعی باارزش و تنوع گونه‌های جانوری و گیاهی در فهرست ذخیره‌گاه‌های زیست‌کره یونسکو قرار گرفته است. در منطقه حفاظت‌شده ارسباران حداقل ارتفاع ۲۷۰ و حداکثر ارتفاع ۲۸۷۰ متر از سطح دریا می‌باشد (۲۱). متوسط بارندگی سالیانه ۴۰۸/۱ میلی‌متر، میانگین دمای سالیانه ۱۲/۳ که حداقل دما ۲/۳- در ماه بهمن و حداکثر دما ۲۴ درجه سانتی‌گراد در ماه مرداد در یک دوره ۱۶ ساله (۱۳۷۹-۱۳۹۵) می‌باشد. طبق روش آمبرژه اقلیم منطقه جزء اقلیم سرد و مرطوب محسوب می‌شود (۳۳). عمده پوشش گیاهی در منطقه جنگل و بعد از آن مرتع و در قسمت‌های حاشیه رودخانه ارس اندکی هم زمین‌های کشاورزی قابل‌مشاهده و تفکیک می‌باشد. در مناطق جنگلی خاک منطقه بیش‌تر از نوع انتی‌سول، اینسپتی‌سول، مالی‌سول و آلفی‌سول است که از نظر ازت و ماده آلی مناسب بوده و پوشیده از هوموس نسبتاً نازکی هست. بافت خاک در بیش‌تر نقاط شنی و شنی لومی و عمدتاً سبک بوده که قابلیت نفوذپذیری آن خوب با زهکشی طبیعی و مناسب است. خاک‌های این منطقه فاقد محدودیت شوری و قلیایی بوده و میزان املاح آن ناچیز است (۳۲، ۳۳ و ۴۳). به‌طورکلی اسیدیته خاک بین ۷-۸ می‌باشد که گویای خاک‌های جنگلی است (۳۲).

با توجه به پژوهش‌های پیشین، ضرورت دارد کارایی روش مدل‌سازی شبکه عصبی مصنوعی و یافتن راهکاری برای بهینه‌سازی خودکار مؤلفه‌های آن برای برآورد SOC روی تصاویر ماهواره‌ای بیش از پیش مورد ارزیابی قرار گیرد. فرضیه اصلی مطالعه حاضر این است که با استفاده از الگوریتم ژنتیک می‌توان مؤلفه‌های شبکه عصبی مصنوعی را به نحوی تعیین نمود که امکان برآورد SOC با استفاده از تصاویر ماهواره‌ای با دقت و صحت بیش‌تر فراهم شود؛ بنابراین مطالعه حاضر با هدف بررسی کارایی شبکه عصبی مصنوعی بهینه‌شده با الگوریتم ژنتیک در برآورد SOC با استفاده از تصاویر ماهواره Sentinel-2 انجام شد. لازم به ذکر است که هدف اصلی این مطالعه تهیه نقشه مکانی SOC نیست بلکه به دنبال افزایش کارایی شبکه عصبی مصنوعی در برآورد نقطه‌ای SOC با استفاده از اطلاعات طیفی تصاویر ماهواره Sentinel-2 است. بررسی انواع شاخص‌های طیفی به‌منظور برآورد SOC نیز یکی دیگر از اهداف این پژوهش در نظر گرفته شده است. با توجه به امکان دسترسی رایگان به تصاویر ماهواره‌ای مورد استفاده، دستاوردهای این پژوهش می‌تواند در پایش تغییرات SOC در این ناحیه رویشی مورد استفاده قرار گیرد.

مواد و روش‌ها

منطقه مورد مطالعه: ناحیه رویشی ارسباران در شمال‌غرب کشور و شمال آذربایجان شرقی قرار دارد. ناحیه رویشی ارسباران محدوده بین ۴۰' ۳۸° تا ۰۹' ۳۹° عرض شمالی و ۴۲' ۴۶° تا ۰۳' ۴۷° طول



شکل ۱- ناحیه رویشی ارسباران (ج) در استان آذربایجان شرقی (ب) در شمال شرق ایران (الف). دایره‌های تیره‌رنگ روی نقشه ناحیه رویشی موقعیت مکانی نمونه‌های خاک برداشت‌شده در این پژوهش را نشان می‌دهند.

Figure 1. Arasbaran biosphere reserve (c) in East Azarbaijan province (b) in the North East of Iran (a). The dark circles on the biosphere map show location of the soil samples taken in this study.

انجام شد تا از همه طبقات شیب و جهت نمونه گرفته شود. با مراجعه به عرصه و انجام پیمایش صحرائی، موقعیت مکانی کلیه نقاط نمونه‌برداری تعیین شد. در محل هر نقطه با استفاده از بیلچه یک نمونه از خاک سطحی صفر تا ۱۵ سانتی‌متر به وزن ۵۰۰ گرم تهیه و پس از نصب برچسب مربوطه، به آزمایشگاه خاکشناسی منتقل گردید. نمونه‌ها در دمای آزمایشگاه هوا خشک شدند و سپس از الک ۲ میلی‌متری عبور داده شدند. مقدار کربن آلی آن‌ها نیز به روش تیتراسیون والکلی بلاک اندازه‌گیری شد (۳۸) که از طریق رابطه ۱ محاسبه می‌شود. هم‌چنین مشخصات آماری SOC نمونه‌های برداشت شده از منطقه مورد مطالعه از جمله میانگین، حدافل، حداکثر، انحراف معیار^۴ (SD) و ضریب تغییرات^۵ (CV) به‌دست آمد.

داده‌های میدانی و تصاویر ماهواره Sentinel-2: نمونه‌برداری از خاک‌های منطقه مطالعاتی در تابستان سال ۱۳۹۸ صورت گرفت. در مجموع تعداد ۴۶ نمونه خاک به روش تصادفی طبقه‌بندی شده^۱ و با استفاده از یک سامانه موقعیت‌یاب جهانی^۲ (GPS) دقیق جمع‌آوری شد (شکل ۱). بدین منظور ابتدا با استفاده از نقشه مدل رقومی ارتفاع^۳ (DEM) با قدرت تفکیک مکانی ۳۰ متر، نقشه شیب در چهار طبقه بر اساس درجه (۰-۱۵، ۱۵-۳۰، ۳۰-۴۵ و ۴۵-۶۰) و جهت شیب نیز در چهار طبقه (شمال، جنوب، شرق و غرب) تهیه شد. سپس با در نظر گرفتن نقشه‌های شیب، جهت شیب و کاربری اراضی (شامل جنگل، مرتع و کشاورزی) که از سایت OpenStreetMap اخذ شده بود، توزیع مکانی نمونه‌های خاک به نحوی

4- Standard deviation
5- Coefficient of variation

1- Stratified random sampling
2- Global positioning system: GPS
3- Digital elevation model: DEM

رادیومتریکی و هندسی در پیش‌پردازش محصول Level-1C توسط USGS انجام شده است. به‌منظور برآورد کربن آلی خاک نیز از مجموعه‌ای از شاخص‌های طیفی و باندها که در جدول ۱ نشان داده شده است، استفاده شد. دلیل آن این است که نواحی طیفی شاخص‌های پوشش گیاهی نسبت به تغییرات در SOC حساس هستند (۲۳). برای انتخاب این شاخص‌ها ابتدا ماتریس همبستگی بین ۴۴ متغیر که شامل باندهای اصلی ماهواره Sentinel-2 و شاخص‌های مشتق‌شده از تصاویر ماهواره Sentinel-2 بود، برقرار شد. شاخص‌های ویژه‌ای نیز با استفاده از داده‌های Sentinel-2 ارائه شدند (مانند HCSI: heavy metal Cd stress-sensitive spectral index) ولی با SOC همبستگی نداشتند. از میان ۴۴ متغیر مرتبط آن‌هایی که همبستگی بالایی با SOC و همبستگی پایینی با خود داشتند، انتخاب شد. با توجه به این‌که نمونه‌برداری از منطقه مورد مطالعه به‌صورت نقطه‌ای و توسط یک GPS دقیق در هر سه کاربری (زیر تاج درختان در جنگل، مرتع، کشاورزی) انجام شد، مکان نمونه‌ها و پیکسل تصویر ماهواره منطبق شدند. ارزش رقمی مربوط به نمونه‌ها با روی هم انداختن نقاط نمونه‌برداری و تصویر ماهواره Sentinel-2 در نرم‌افزار ArcGIS استخراج شد. با توجه به این‌که پیکسل سایز باندهای مختلف با هم برابر نبود، سایز پیکسل همه باندهای مورد استفاده با روش Resampling به ۲۰ متر تبدیل شد. با توجه به این‌که تمام بازتاب روی تصویر ماهواره‌ای در سه کاربری منطقه مطالعاتی مربوط به گیاهان بودند، شاخص‌هایی که صرفاً به خاک مربوط بودند در این مطالعه با کربن آلی همبستگی نداشتند. تنها شاخص TSAVI در این پژوهش در نظر گرفته شد که از بین شاخص‌های خاک بررسی شده، بیش‌ترین همبستگی را با کربن آلی خاک در این مطالعه داشت. با

$$\%SOC = \frac{V_1 - V_2}{S} \times N \times \frac{0.003}{0.77} \times 100 \quad (1)$$

که در آن، %SOC درصد کربن آلی خاک، N نرمالیزه فروآمونیم سولفات، V_1 و V_2 به‌ترتیب فروآمونیم سولفات مصرفی برای شاهد و نمونه بر حسب میلی‌لیتر و S وزن خاک خشک‌شده در هوای آزاد بر حسب گرم است. عدد $0.003/0.77$ نیز میلی‌اکی‌والان کربن هست، هم‌چنین در این روش ۷۷ درصد کربن آلی اکسید می‌شود و این موضوع در فرمول محاسبات نیز لحاظ شده است (۳۸).

هم‌چنین تصویر بدون ابر ماهواره Sentinel-2 برای تاریخ ۱۲ تیرماه ۱۳۹۸ از سایت^۱ USGS دانلود شد. ماهواره Sentinel-2 دارای ۱۳ باند است. توان تفکیک مکانی این باندها متغیر بوده، باندهای ۲، ۳، ۴ و ۸ دارای توان تفکیک مکانی ۱۰ متر و باندهای ۵، ۶، ۷، ۸a، ۱۱ و ۱۲ دارای توان تفکیک مکانی ۲۰ متر و باندهای ۱، ۹ و ۱۰ دارای توان تفکیک مکانی ۶۰ متر هستند، از میان ۱۳ باند این تصویر، باندهای Coastal Aerosol (1)، Water vapor (9)، SWIR Cirrus (10) که دربردارنده اطلاعات ویژه‌ای نیستند، حذف شده و از فرآیند پردازشی کنار گذاشته شدند.

پردازش و پیش‌پردازش تصاویر ماهواره‌ای: تصویر دانلود شده در قالب محصول Level-1C تولید می‌شود، بنابراین به‌صورت بازتابندگی بالای اتمسفر^۲ (TOA) است و برای این‌که محصول Level-2A به‌دست آید که بازتابندگی پایین اتمسفر^۳ (BOA) است تصویر با استفاده از پردازنده Sen2Cor (v.5.5.5) در نرم‌افزار SNAP از نظر اتمسفری تصحیح شد و محصول سطح یک به سطح دو تبدیل شد. تصحیحات

1- United States Geological Survey
2- Top of atmosphere
3- Bottom of atmosphere

استفاده از ۴۳۱ پیکسل خالص خاک و با استفاده از فضای ویژگی باند قرمز و مادون قرمز نزدیک، صفر به عنوان عرض از مبدأ خط خاک ۱/۱ به عنوان شیب خط خاک در معادله خط خاک (NIR=1.0814×RED-0.00009) تعیین گردید. با استفاده از ضرایب همبستگی اسپیرمن روابط بین کربن آلی خاک و متغیرهای ورودی استخراج شد و تأثیر بیش‌ترین و کم‌ترین متغیر بر مدل‌سازی پیش‌بینی SOC نشان داده شد.

جدول ۱- مشخصات متغیرهای ورودی.

Table 1. Specifications of variables.

منبع Reference	جزئیات Detail	فرمول Formula	شاخص Index
(36)	-	$\frac{\rho\text{NIR} - \rho\text{Red}}{\rho\text{NIR} + \rho\text{Red}}$	NDVI
(11)	-	$\frac{\rho\text{NIR}}{\rho\text{NIR} + \rho\text{Red}}$	IPVI
(24)	G=2.5, C ₁ =6, C ₂ =7.5, L=1	$G \frac{\rho\text{NIR} - \rho\text{Red}}{\rho\text{NIR} + C_1 \times \rho\text{Red} - C_2 \times \rho\text{Blue} + L}$	EVI
(39)	-	$\frac{\rho\text{Green} - \rho\text{Red}}{\rho\text{Green} + \rho\text{Red}}$	GRVI
(34)	-	$\frac{\rho\text{Red} \times \rho\text{Red}}{\rho\text{Green} \times \rho\text{Green} \times \rho\text{Green}}$	RI
(3)	a=Soil line intercept s=Soil line slope X=Adjustment factor to minimize soil noise	$\frac{s(\rho\text{NIR} - s \times \rho\text{Red} - a)}{(a \times \rho\text{NIR} + \rho\text{Red} - a \times s + X \times (1 + s \times s))}$	TSAVI
(10)	-	$B_2 \times 0.3037 + B_3 \times 0.2793 + B_4 \times 0.4743 + B_8 \times 0.5585 + B_{11} \times 0.5082 + B_{12} \times 0.1863$	TCB
(43)	-	$\frac{\rho\text{NIR} - \text{SWIR}_1}{\rho\text{NIR} + \text{SWIR}_1}$	LSWI
(15)	-	$\sqrt{\frac{(\rho\text{Red} \times \rho\text{Red}) + (\rho\text{Green} \times \rho\text{Green}) + (\rho\text{NIR} \times \rho\text{NIR})}{3}}$	BI2
-	-	-	Band4
-	-	-	Band11
-	-	-	Band8a

جلوگیری نمود. برای استانداردسازی از رابطه ۳ استفاده کرده و داده‌ها را بین (۱ و -۱) استاندارد شد.

$$X_n = \left(\frac{X - X_{\min}}{X_{\max} - X_{\min}} \right) \times 2 - 1 \quad (3)$$

که در آن، X_n متغیر نرمال شده، X متغیر نرمال نشده، X_{\max} حداکثر داده‌ها، X_{\min} حداقل داده‌ها است.

الگوریتم ژنتیک: الگوریتم ژنتیک از تئوری تکامل بیولوژیکی که داروین پیشنهاد کرده الهام گرفته شده است (۱۹). یک کاربرد متداول الگوریتم ژنتیک، استفاده از آن به‌عنوان تابع بهینه‌کننده است. در الگوریتم ژنتیک هر پدیده از یک سری ترکیبات درست شده است که به‌صورت کروموزوم نمایش داده می‌شود. در این روش هدف انتخاب بهترین صفات در کروموزوم برای رسیدن به نقطه بهینه است. در الگوریتم ژنتیک یک مقدار تطابق با توجه به بقیه شرایط محیط به هر کدام از کروموزوم‌ها اختصاص داده می‌شود. کروموزوم‌هایی که دارای تابع تطابق با مقدار بالا هستند و قدرت بیش‌تری دارند در جهت بهینه شدن سیستم انتخاب می‌شوند و نسل بعدی را تولید می‌کنند. برای تولید نسل بعد، دو عملگر متداول الگوریتم ژنتیک که عبارت‌اند از عملگر جهش^۳ و عملگر برش^۴ به‌ترتیب با احتمال‌های P_c و P_m بر روی نسل اولیه اعمال می‌شود. سپس با محاسبه میزان تطابق کروموزوم‌های به‌دست آمده و مقایسه آن‌ها با نسل قبل، کروموزوم‌های نسل بعد انتخاب می‌شوند تا بستری برای اعمال مجدد عملیات الگوریتم ژنتیک باشند. این عملیات هم‌چنان ادامه می‌یابد تا این‌که اهداف مسأله محقق شود و سیستم

شبکه عصبی مصنوعی: نوع ساختار در شبکه عصبی مصنوعی برگرفته از مدل سلول‌های عصبی در سیستم بیولوژیک بدن انسان می‌باشد. شبکه عصبی مصنوعی از طریق لایه ورودی و خروجی با فضای بیرون در ارتباط است و تمام ارتباطات موجود در داخل شبکه بستگی به طراح مسأله دارد. تعداد نرون‌ها در لایه ورودی بستگی به داده‌هایی دارد که از فضای بیرون یک پدیده در دسترس ماست. هر نرون معرف یک الگوست که آموزش می‌دهد یا آموزش می‌بیند. این آموزش می‌تواند به روش‌های مختلفی از جمله الگوریتم ژنتیک، گرادیان نزولی^۱ و ... انجام شود. هر نرون می‌تواند ارزش خروجی Z_1 بر اساس جمع وزنی ورودی‌های X_j بر اساس رابطه ۲ محاسبه کند (۱۲).

$$Z_1 = f_1 \left(\sum_{j=1}^k W_j^1 X_j \right) \quad (2)$$

که در آن، W_j^1 جمع ضرایب وزن‌های سیناپسی، اندیس ۱ نشان‌دهنده تعداد نرون‌ها و f_1 تابع فعال‌سازی^۲ شبکه عصبی مصنوعی می‌باشد. شبکه عصبی مصنوعی مورد استفاده در این پژوهش از نوع پرسپترون می‌باشد. این شبکه از سه لایه به‌صورت کلی تشکیل شده، لایه ورودی، لایه پنهان و لایه خروجی. لایه پنهان خود می‌تواند دارای چندین لایه باشد، به‌عبارت دیگر به‌جز لایه ورودی و خروجی سایر لایه‌ها به‌عنوان لایه پنهان شناخته می‌شوند. برای طراحی شبکه عصبی، تابع سیگموئید به‌عنوان تابع فعال‌سازی لایه پنهان و تابع فعال‌سازی لایه خروجی از نوع خطی انتخاب شد. با تنظیم داده‌های ورودی در یک محدوده خاص می‌توان نرون‌ها را در محدوده مطلوب قرار داده و از اشباع زودهنگام نرون‌ها

3- Mutation
4- Crossover

1- Gradient descent
2- Activation functions

که در آن‌ها، y_t ، \hat{y}_t ، \bar{y}_t و SD به ترتیب مقادیر مشاهده شده، پیش‌بینی شده، میانگین و انحراف معیار می‌باشند. با استفاده از مقایسه ضریب تعیین که مقادیری بین صفر تا یک را به خود اختصاص می‌دهد، بدین ترتیب که هرچه به یک نزدیک‌تر باشد، مدل از دقت بالاتری در پیش‌بینی برخوردار خواهد بود. RMSE در حالت ایده‌آل باید برابر صفر باشد، هرچه به صفر نزدیک‌تر باشد دقت مدل افزایش می‌یابد. RPD نیز با استفاده از سیستم طبقه‌بندی توصیه شده توسط چانگ و همکاران (۲۰۰۱) مورد استفاده قرار گرفت. به طور کلی، در $RPD < 1$ مدل بسیار ضعیف و توصیه نمی‌شود، $1/4 < RPD < 1$ مدل ضعیف هست، $1/8 < RPD < 1/4$ یک مدل مناسب هست و از آن می‌شود برای ارزیابی و پیش‌بینی استفاده کرد، $2 < RPD < 1/8$ مدل خوب است و برای پیش‌بینی‌های کمی امکان‌پذیر است، $2/5 < PRD < 2$ مدل بسیار خوب و $PRD < 2/5$ مدل عالی هست (۷). هم‌چنین پس از نرمال کردن داده‌ها به روش Z-score، از آزمون تی جفتی برای مقایسه SOC واقعی با مقادیر برآورد شده از طریق شبکه عصبی مصنوعی و شبکه عصبی مصنوعی بهینه استفاده شد.

نتایج و بحث

نتایج آماری کربن آلی خاک در منطقه مورد مطالعه در جدول ۲ ارائه شده است. آزمون نرمالیت کولموگروف-اسمیرنوف نشان داد که متغیر کربن آلی دارای توزیع نرمال نمی‌باشد. اختلاف بیش‌ترین و کم‌ترین حد کربن آلی خاک که به ترتیب ۱۰/۵ و ۰/۶۱ است، زیاد بوده و این اختلاف تحت تأثیر دو عامل می‌تواند باشد. تغییرپذیری زیاد مکانی خصوصیات خاک (مانند اندازه ذرات خاک، چگالی ظاهری، ضخامت خاک و غلظت SOC) وجود

به جواب مطلوب دست پیدا کند. نوع تلاقی به کاررفته یک نقطه^۱ و نوع انتخاب رولت^۲ است و به منظور انجام مدل‌سازی از نرم‌افزار متلب ۲۰۱۸ استفاده گردید.

ارزیابی دقت و صحت: در این پژوهش از روش Cross validation برای ارزیابی دقت مدل‌ها استفاده شد (۴). استفاده از این روش در شبکه عصبی مصنوعی بدین شکل بود که هر شبکه با ۴۵ نمونه آموزش داده شد. سپس نمونه باقی‌مانده که در آموزش شرکت نداشت با استفاده از این شبکه شبیه‌سازی شد. پس از تکرار این مراحل به تعداد ۴۶ مرتبه، تعداد ۴۶ نمونه شبیه‌سازی شده به دست آمد. سپس نمونه‌های شبیه‌سازی شده که در واقع مقادیر SOC تخمین زده شده بودند در مقابل مقادیر اندازه‌گیری شده SOC قرار گرفته و مقدار ضریب تبیین (R^2)، خطای جذر میانگین مربعات (RMSE)، نسبت عملکرد به انحراف (RPD) محاسبه شد (رابطه‌های ۴ تا ۶). هم‌چنین به دلیل پیروی نکردن توزیع داده‌ها از توزیع نرمال از ضریب همبستگی اسپیرمن برای مقایسه بین کربن آلی اندازه‌گیری شده و برآورد شده از طریق مدل‌های موردنظر استفاده شد. برای انجام محاسبات آماری از نرم‌افزار SPSS و Excel استفاده شد.

$$RMSE = \sqrt{\frac{\sum (\hat{y}_t - y_t)^2}{n}} \quad (4)$$

$$RPD = \frac{SD}{RMSE} \quad (5)$$

$$R^2 = 1 - \frac{\sum_1^N (y_t - \hat{y}_t)^2}{\sum_1^N (y_t - \bar{y}_t)^2} \quad (6)$$

- 1- One point
2- Roulette

تسریع اکسایش مواد آلی می‌شود. این نتایج با سایر مطالعات مطابقت می‌کند. ساسانی‌فر و همکاران (۱۳۹۶) در مطالعه ویژگی‌های خاک ناحیه رویشی ارسباران همانند مطالعه حاضر در مورد تنوع مکانی کربن آلی خاک که دارای توزیع یکنواختی نیست، به نتایج مشابهی دست یافتند (۳۶). هم‌چنین بیش‌تر بودن مقدار SOC جنگل در مقایسه با سایر کاربری‌ها در ناحیه رویشی ارسباران در مطالعه حاضر در راستای دستاوردهای پژوهش‌های پیشین (۲ و ۱۱) بود.

کاربری‌های متفاوت اعم از جنگل که بیش‌ترین SOC و زمین‌های کشاورزی که کم‌ترین SOC را در خود ذخیره می‌کنند. هم‌چنین در مطالعه حاضر میانگین SOC ذخیره شده در جنگل (۳/۹۶ درصد)، مرتع (۲/۵۳ درصد) و زمین‌های کشاورزی (۱/۶۵ درصد) به‌دست آمد. بیش‌ترین SOC ذخیره‌شده در جنگل می‌باشد که این می‌تواند به دلیل ورود بیش‌تر لاشبرگ پوشش گیاهی به خاک و کاهش تجزیه مواد آلی باشد. از طرف دیگر، کم‌ترین میزان SOC در زمین‌های زیر کشت می‌تواند ناشی از کاهش ورود لاشبرگ پوشش گیاهی به خاک و خاکورزی مکرر باشد که باعث

جدول ۲- نتایج آماری کربن آلی خاک.

Table 2. Statistical results of soil organic carbon.

کربن آلی خاک (درصد) Soil organic carbon (%)	مشخصات Characteristics
3.11	میانگین Mean
0.61	حداقل Minimum
10.5	حداکثر Maximum
2.06	انحراف معیار Standard deviation
66	ضریب تغییرات Coefficient of variation

رابطه متغیر Band8a با SOC مثبت و در سطح احتمال پنج درصد ($\alpha=0/05$) است. مقایسه ضریب همبستگی اسپیرمن SOC اندازه‌گیری شده با برآورد شده از طریق مدل‌های موردنظر نشان داد که همبستگی بین SOC اندازه‌گیری شده با برآورد شده با شبکه عصبی مصنوعی بهینه ($r=0/87$) نسبت به شبکه عصبی مصنوعی ($r=0/76$) بیش‌تر است.

با توجه به جدول ۳، کم‌ترین ارتباط بین SOC و BI2 ($r=0/196$) است و بیش‌ترین ارتباط بین SOC و TSAVI ($r=0/565$) است که رابطه در سطح احتمال یک درصد ($\alpha=0/01$) معنی‌دار است. هم‌چنین بین SOC و متغیرهای NDVI، IPVI، EVI، GRVI، RI، TSAVI، LSWI، Band 4 رابطه مستقیم و مثبت و بین SOC و Band 11 رابطه منفی در سطح احتمال یک درصد ($\alpha=0/01$) وجود دارد و

مطالعات سایر پژوهشگران در سایر موضوعات مطابقت می‌کند (۱۶ و ۲۷). در مطالعه حاضر نیز با اعمال الگوریتم ژنتیک ضریب تبیین برآورد SOC روی تصاویر ماهواره Sentinel-2 افزایش پیدا کرده است. در مطالعه حاضر که با مدل شبکه عصبی مصنوعی پرسپترون و بهینه‌سازی شبکه عصبی مصنوعی با الگوریتم ژنتیک انجام شده است، از معیارهای ارزیابی دقت متفاوت نسبت به پژوهش‌های پیشین استفاده شده است که یکی از این‌ها RPD هست و با توجه به نتایج جدول ۴ به ترتیب برای داده‌های تست مدل شبکه عصبی مصنوعی، شبکه عصبی مصنوعی بهینه‌شده با الگوریتم ژنتیک ۱/۳۴ و ۱/۸۹ می‌باشد. مدل شبکه عصبی مصنوعی در همان محدوده دقت RPD برای پیش‌بینی SOC در یافته‌های سایرین قرار دارد، ولی SOC پیش‌بینی‌شده از طریق شبکه عصبی مصنوعی بهینه از بعضی از یافته‌ها بیشتر است و دلیل آن استفاده از الگوریتم ژنتیک برای بهینه کردن وزن‌های شبکه عصبی مصنوعی است.

ورودی‌های مسأله بر اساس متغیرهای جدول ۱ برای مدل‌سازی توسط شبکه عصبی مصنوعی و شبکه عصبی مصنوعی بهینه‌شده با الگوریتم ژنتیک استفاده شدند. نتایج حاصل از مدل‌سازی شبکه عصبی مصنوعی و شبکه عصبی مصنوعی بهینه در جدول ۴ نتایج نشان داده شده است. با توجه به نتایج، R^2 و RPD با اعمال الگوریتم ژنتیک افزایش پیدا کرده است و هم‌چنین RMSE نیز کاهش پیدا کرده است. این افزایش و کاهش نشان‌دهنده تأثیر الگوریتم ژنتیک در بهینه کردن وزن‌های شبکه عصبی مصنوعی در مدل‌سازی کربن آلی خاک هست. هم‌چنین شکل ۲ درستی این امر را تأیید می‌کند. مقایسه میانگین‌ها نیز نشان داد SOC واقعی با SOC برآوردی با شبکه عصبی مصنوعی بهینه اختلاف معنی‌داری نداشت ($t=1/28$, $p\text{-value}=0/21$) ولی با SOC برآوردی با شبکه عصبی مصنوعی متفاوت بود ($t=2/47$, $p\text{-value}=0/02$). طبق نتایج این پژوهش بهینه‌سازی شبکه عصبی مصنوعی با الگوریتم ژنتیک باعث افزایش دقت و صحت مدل در پیش‌بینی می‌شود که با

جدول ۳- ضریب همبستگی اسپرمن متغیرهای ورودی با کربن آلی خاک (* همبستگی در سطح ۰/۰۵ معنی‌دار، ** همبستگی در سطح ۰/۰۱ معنی‌دار است).

Table 3. Spearman correlation coefficients of the input variables and soil organic carbon (* Correlation is significant at 0.05 level, ** Correlation is significant at 0.01 level).

متغیر Variable	ضریب همبستگی Correlation coefficient
Band8a	0.353*
Band11	-0.235**
Band4	0.478**
BI2	0.196
LSWI	0.400**
TCB	-0.215
TSAVI	0.565**
RI	0.351**
GRVI	0.530**
EVI	0.517**
IPVI	0.499**
NDVI	0.496**

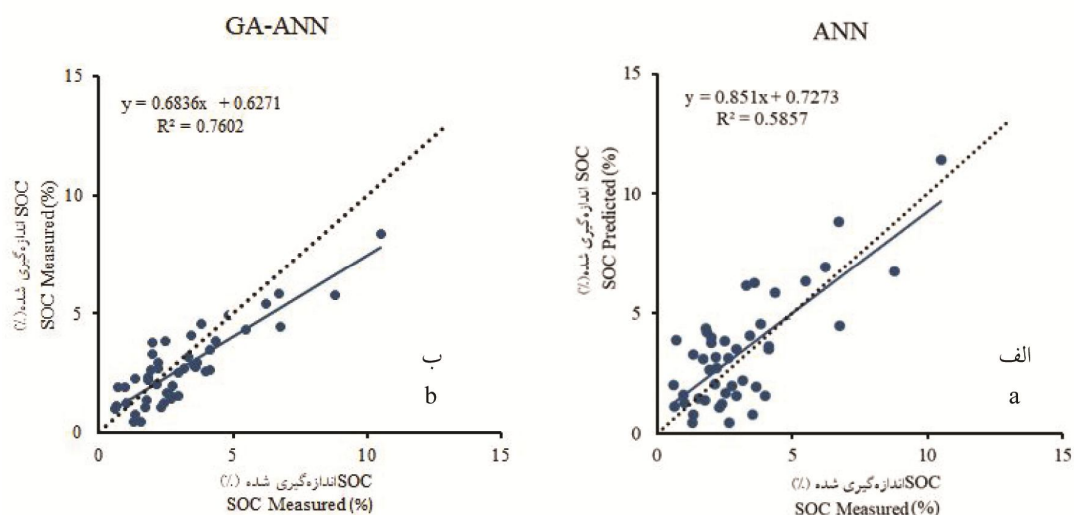
جدول ۴- نتایج حاصل از مدل‌سازی شبکه عصبی مصنوعی و شبکه عصبی مصنوعی بهینه‌شده با الگوریتم ژنتیک.

Table 4. Results of modeling of artificial neural network and artificial neural network optimized by genetic algorithm.

R ²	RPD (%)	RMSE (%)	نرمال‌سازی Normalization	تعداد لایه پنهان Number of hidden layers	تعداد نورون Number of neurons	نوع مدل Model type
0.58	1.34	1.51	[-1,1]	2	22	ANN(MLP)
0.76	1.89	1.07	[-1,1]	2	34	GA-ANN

احتمالاً علت آن استفاده از تصاویر Sentinel-2 می‌باشد که نسبت به تصاویر ماهواره لندست ۸ توان تفکیک طیفی و مکانی بالاتری دارند (۴۰). با توجه به این‌که بخشی از ناحیه رویشی ارسباران را تاج پوشش جنگلی فراگرفته، SOC به صورت غیرمستقیم برآورد شده است که این روش با مطالعات چن و همکاران (۲۰۰۰) و ور و همکاران (۲۰۱۵) مطابقت دارد (۷ و ۴۰).

مقدار RPD حاصل از شبکه عصبی مصنوعی با نتایج کایستالدی و همکاران (۲۰۱۶) و قلی‌زاده و همکاران (۲۰۱۸) در یک محدوده قرار داشت ولی از شبکه عصبی مصنوعی بهینه‌کم‌تر بود (۵ و ۱۸). علت این موضوع ممکن است به عملکرد بهتر شبکه عصبی مصنوعی بهینه به دلیل انتخاب پارامترها با الگوریتم ژنتیک بازگردد. هم‌چنین ضریب تبیین پژوهش حاضر بیش‌تر از نتایج مطالعه ور و همکاران (۲۰۱۵) شد که



شکل ۲- نمودار مقادیر اندازه‌گیری‌شده و پیش‌بینی‌شده کربن آلی خاک در دو مدل شبکه عصبی مصنوعی پرسپترون (الف) و شبکه عصبی مصنوعی بهینه‌شده با الگوریتم ژنتیک (ب).

Figure 2. Diagram of measured and predicted amounts of soil organic carbon in two models of perceptron artificial neural network (a), and artificial neural network optimized by genetic algorithm (b).

نتیجه‌گیری کلی

به‌طور کلی این مقاله با هدف ارزیابی استفاده از الگوریتم ژنتیک برای بهینه کردن شبکه عصبی مصنوعی در مدل‌سازی و پیش‌بینی SOC در ناحیه رویشی ارسباران با تصاویر Sentinel-2 انجام شد. با توجه به نتایج ارزیابی مدل به‌دست‌آمده می‌توان نتیجه گرفت به‌کار بردن الگوریتم ژنتیک در بهینه کردن شبکه عصبی برای هدف موردنظر مؤثر بوده و نسبت به شبکه عصبی از دقت و صحت بیشتری برخوردار است. ماهواره Sentinel-2 که دارای توان تفکیک طیفی و مکانی مناسب برای این موضوع است و همچنین دوره بازگشت ۵ روزه را دارد که این امکان را فراهم می‌کند که به‌طور مداوم بتوان SOC را پایش کرد. از بین متغیرهای مورد استفاده برای مدل‌سازی پیش‌بینی کربن آلی خاک TSAVI

بیش‌ترین همبستگی را با SOC داشت. همچنین استفاده از شاخص‌های مناسب به‌دلیل این‌که مناطق طیفی شاخص‌ها نسبت به تغییرات در SOC حساس هستند، می‌تواند مفید باشد. لازم به ذکر است که این مطالعه به دنبال تهیه نقشه مکانی SOC در ارسباران نبود بلکه هدف اصلی بهبود عملکرد شبکه عصبی مصنوعی در برآورد نقطه‌ای SOC با اطلاعات طیفی تصاویر Sentinel-2 بود. با توجه به این‌که مدل‌های یادگیری ماشین و الگوریتم‌های تکاملی اخیراً پیشرفت بهتری در زمینه مدل‌سازی و پیش‌بینی کرده‌اند، پیشنهاد می‌شود از الگوریتم‌های تکاملی به‌ویژه الگوریتم ژنتیک برای بهینه کردن روش‌های یادگیری ماشین در برآورد SOC با استفاده از داده‌های سنجنش‌ازدور استفاده شود.

منابع

- Adhikari, K., and Hartemink, A.E. 2015. Digital mapping of topsoil carbon content and changes in the Driftless area of Wisconsin, USA. *Soil Science Society of America J.* 79: 1. 155-164.
- Amanuel, W., Yimer, F., and Karlton, E. 2018. Soil organic carbon variation in relation to land use changes: the case of Birr watershed, upper Blue Nile River Basin, Ethiopia. *J. of Ecology and Environment.* 42: 1. 16-27.
- Baret, F., and Guyot, G. 1991. Potentials and limits of vegetation indices for LAI and APAR assessment. *Remote Sensing of Environment.* 35: 2. 161-173.
- Browne, M.W. 2000. Cross-validation methods. *J of mathematical psychology.* 44: 1. 108-32.
- Castaldi, F., Palombo, A., Santini, F., Pascucci, S., Pignatti, S., and Casa, R. 2016. Evaluation of the potential of the current and forthcoming multispectral and hyperspectral imagers to estimate soil texture and organic carbon. *Remote Sensing of Environment.* 179: 54-65.
- Chang, C.W., Laird, D.A., Mausbach, M.J., and Hurburgh, J. 2001. Near-infrared reflectance spectroscopy - principal components regression analyses of soil properties. *Soil Science Society of America J.* 65: 2. 480-490.
- Chen, F., Kissel, D.E., West, L.T., and Adkins, W. 2000. Field-scale mapping of surface soil organic carbon using remotely sensed imagery. *Soil Science Society of America J.* 64: 2. 746-753.
- Chiroma, H., Noor, A.S.M., Abdulkareem, S., Abubakar, A.I., Hermawan, A., Qin, H., Hamza, M.F., and Herawan, T. 2017. Neural networks optimization through genetic algorithm searches: A review. *J. of Applied Mathematics and Information Sciences.* 11: 6. 1543-1564.
- Cohen, W.B., and Spies, T.A. 1992. Estimating structural attributes of Douglas-fir/western hemlock forest stands from Landsat and SPOT imagery. *Remote Sensing of Environment.* 41: 1. 1-17.

10. Crippen, R.E. 1990. Calculating the vegetation index faster. *Remote Sensing of Environment*. 34: 1. 71-73.
11. Dengiz, O., Sağlam, M., and Türkmen, F. 2015. Effects of soil types and land use land cover on soil organic carbon density at Madendere watershed. *Eurasian J. of Soil Science*. 4: 2. 82-87.
12. Dreyfus, G. 2005. *Neural networks: Methodology and applications*. Springer-Verlag, Berlin, Germany. 322p.
13. Elachi, C., and Zyl, J. 2006. *Introduction to the physics and techniques of remote sensing*. John Wiley and Sons, New Jersey, U.S.A. 513p.
14. Escadafal, R. 1989. Remote sensing of arid soil surface color with Landsat thematic mapper. *Advances in Space Research*. 9: 1. 159-163.
15. Eswaran, H., Van Den Berg, E., and Reich, P. 1993. Organic carbon in soils of the World. *Soil Science Society of America J.* 57: 1. 192-194.
16. Fagih, H. 2011. Evaluation of artificial neural network application and optimization using genetic algorithm in estimation of monthly precipitation data (case study: Kurdistan region). *J. of Water and Soil Science*. 14: 51. 27-44. (In Persian)
17. Furtuna, R., Curteanu, S., and Leon, F. 2011. An elitist non-dominated sorting genetic algorithm enhanced with a neural network applied to the multi-objective optimization of a polysiloxane synthesis process. *Engineering Applications of Artificial Intelligence*. 24: 5. 772-785.
18. Gholizadeh, A., Žižala, D., Saberioon, M., and Borůvka, L. 2018. Soil organic carbon and texture retrieving and mapping using proximal, airborne and Sentinel-2 spectral imaging. *Remote Sensing of Environment*. 218: 89-103.
19. Goldberg, D.E., and Holland, J.H. 1988. Genetic algorithms and machine learning. *Machine Learning*. 3: 2. 95-99.
20. Harpham, C., Dawson, C.W., and Brown, M.R. 2004. A review of genetic algorithms applied to training radial basis function networks. *Neural Computing and Applications*. 13: 3. 193-201.
21. Hejazi, A. 2009. An analysis on the phytogeomorphological potential of Arasbaran biosphere storage. *J. of Geography and Planning*. 13: 33-39. (In Persian)
22. Huete, A., Didan, K., Miura, T., Rodriguez, E.P., Gao, X., and Ferreira, L.G. 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment*. 83: 2. 195-213.
23. Jin, X., Song, K., Du, J., Liu, H., and Wen, Z. 2017. Comparison of different satellite bands and vegetation indices for estimation of soil organic matter based on simulated spectral configuration. *Agricultural and Forest Meteorology*. 245: 57-71.
24. Karunaratne, S.B., Bishop, T.F.A., Baldock, J.A., and Odeh, I.O.A. 2014. Catchment scale mapping of measurable soil organic carbon fractions. *Geoderma*. 220: 14-23.
25. Kumar, S., Lal, R., and Liu, D. 2012. A geographically weighted regression kriging approach for mapping soil organic carbon stock. *Geoderma*. 190: 627-634.
26. Kumar, S., Lal, R., Liu, D., and Rafiq, R. 2013. Estimating the spatial distribution of organic carbon density for the soils of Ohio, USA. *J. of Geographical Sciences*. 23: 2. 280-296.
27. Liu, Z., Liu, A., Wang, C., and Niu, Z. 2004. Evolving neural network using real coded genetic algorithm (GA) for multispectral image classification. *Future Generation Computer Systems*. 20: 7. 1119-1129.
28. Martin, M.P., Wattenbach, M., Smith, P., Meersmans, J., Jolivet, C., Bouillon, L., and Arrouays, D. 2011. Spatial distribution of soil organic carbon stocks in France. *Biogeosciences*. 8: 5. 1053-1065.
29. McBratney, A.B., Mendonça Santos, M.L., and Minasny, B. 2003. On digital soil mapping. *Geoderma*. 117: 2. 3-52.

30. Meersmans, J., De Ridder, F., Canters, F., De Baets, S., and Van Molle, M. 2008. A multiple regression approach to assess the spatial distribution of Soil Organic Carbon (SOC) at the regional scale (Flanders, Belgium). *Geoderma*. 143: 2. 1-13.
31. Pouget, M., Madeira, J., Lefloch, E., and Kamal, S. 1990. Caracteristiques spectrales des surfaces sableuses de la region cotiere nord-ouest de l'Egypte: application aux donnees satellitaires SPOT. *J. De teledetection*. 12: 27-39.
32. Rasuly, A., Naghdifar, R., and Rasoli, M. 2010. Detecting of Arasbaran forest changes applying image processing procedures and GIS techniques. *Procedia Environmental Sciences*. 2: 454-464. (In Persian)
33. Rezaei, H., Jsfarzadeh, A.A., Alijanpour, A., Shahbazi, F., and Valizadeh Kamran, K. 2016. Genetically evolution of Arasbaran forests soils along altitudinal transects of Kaleybar Chai Sofla Sub-Basin. *Iranian J. of Water and Soil Science*. 26: 1. 151-166. (In Persian)
34. Rouse, J., Haas, J.R., Schell, J., and Deering, D. 1974. Monitoring vegetation systems in the great plains with ERTS. *Proceedings of the 3rd ERTS Symposium*. 1: 309-317.
35. Rumpel, C., Amiraslani, F., Koutika, L.S., Smith, P., Whitehead, D., and Wollenberg, E. 2018. Put more carbon in soils to meet Paris climate pledges. *Nature*. 564: 32-34.
36. Sasanifar, S., Alijanpor, A., Banjshafi, A., Eshagirad, J., and Molai, M. 2018. The impact of conservation-based management on the physical and chemical properties of soil in Arasbaran forests. *Iranian J. of Forest and Poplar Research*. 26: 1. 104-117. (In Persian)
37. Tucker, C.J. 1979. Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*. 8: 2. 127-150.
38. Walkley, A.J., and Black, I. 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: 29-38.
39. Wang, C., Cui, Y., Ma, Z., Guo, Y., Wang, Q., Xiu, Y., Xiao, R., and Zhang, M. 2019. Simulating spatial variation of soil carbon content in the Yellow River Delta: Comparative analysis of two artificial neural network models. *Wetlands*. 13: 29-38.
40. 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.
41. Xiao, X., Zhang, Q., Braswell, B., Urbanski, S., Boles, S., Wofsy, S., Moore, B., and Ojima, D. 2004. Modeling gross primary production of temperate deciduous broadleaf forest using satellite images and climate data. *Remote Sensing of Environment*. 91: 256-270.
42. Yang, Y., Fang, J., Tang, Y., Ji, C., Zheng, C., He, J., and Zhu, B. 2008. Storage, patterns and controls of soil organic carbon in the Tibetan grasslands. *Global Change Biology*. 14: 7. 1592-1599.
43. Zebardast, L., Jafari, H., Badehyan, Z., and Asheghmoala, M. 2011. Assessment of the trend of changes in land cover of Arasbaran protected area using satellite images of 2002, 2006 and 2008. *Environmental Research J*. 1: 1. 23-33. (In Persian)
44. Zhang, Y., Guo, L., Chen, Y., Shi, T., Luo, M., Ju, Q., Zhang, H., and Wang, S. 2019. Prediction of soil organic carbon based on Landsat 8 monthly NDVI data for the Jiangnan plain in Hubei province, China. *Remote Sensing*. 11: 14. 1683.



Efficiency of optimized artificial neural network in soil organic carbon modeling based on *in-situ* measurements and Sentinel-2 images in Arasbaran

M. Lotfi¹, *S.Y. Erfanifard², F. Amir Aslani² and A. Keshavarzi³

¹M.Sc. Student, Dept. of Remote Sensing and GIS, Faculty of Geography, University of Tehran, Tehran, Iran,

²Associate Prof., Dept. of Remote Sensing and GIS, Faculty of Geography, University of Tehran, Tehran, Iran,

³Assistant Prof., Dept. of Soil Sciences and Engineering, College of Agriculture and Natural Resources,
University of Tehran, Karaj, Iran

Received: 04.27.2020; Accepted: 06.04.2020

Abstract

Background and Objectives: Soils are the largest carbon pool in terrestrial ecosystems, which account for the greatest amount of the global total terrestrial carbon stocks. Accurate mapping of Soil Organic Carbon (SOC) spatial distribution is a key assumption for soil resource management and environmental protection. The rapid development of remote sensing and the application of satellite images provide an excellent opportunity to monitor large-scale SOC storage. Estimating SOC is one of the research topics that artificial neural networks are applied for this purpose in some studies, although parameter optimization is difficult. In previous studies, genetic algorithms have been used to optimize the artificial neural network initial weights and improve the prediction of the output variables. However, the effectiveness of this method in estimating the SOC by remote sensing has been less studied. In this study, the effect of the genetic algorithm on artificial neural network training to predict SOC on Sentinel-2 satellite images in the Arasbaran vegetation zone was investigated.

Materials and Methods: For this purpose, soil sampling was performed using the stratified sampling method at 46 points from depths varying between 0 and 15 cm. SOC was measured by the Walkley-Black titration method. To fit the model between the measured organic carbon in the laboratory, 9 spectral indices and three bands of the satellite image, and two methods were used namely, artificial neural network and artificial neural network optimized by genetic algorithm. Cross-validation was used to evaluate the models efficiently. Finally, the precision of the obtained models was evaluated with statistical indices of Root Mean Square Error (RMSE), Ratio of Performance to Deviation (RPD), Spearman's correlation coefficient (r), coefficient of determination (R^2), and paired sample t-test.

Results: The results showed that the precision of SOC estimated by artificial neural network optimized by a genetic algorithm (RMSE = 1.07%, RPD = 1.89%, $R^2 = 0.76$) was higher than artificial neural network results (RMSE = 1.51%, RPD = 1.34%, $R^2 = 0.58$). Also, the Spearman correlation coefficient for SOC estimated with optimized artificial neural network ($r = 0.87$) was higher compared to estimated SOC with an artificial neural network ($r = 0.76$). Observed SOC was not significantly different from SOC estimated by an optimized artificial neural network (p-value=0.21) while it was different from estimated SOC by an artificial neural network (p-value=0.02). Besides, the results showed that the TSAVI index had the highest Spearman correlation coefficient (0.565), and the BI2 index had the lowest Spearman correlation coefficient (0.196) with soil organic carbon.

*Corresponding author: erfanifard@ut.ac.ir

Conclusion: Generally, it was concluded that the use of a genetic algorithm in the selection of artificial neural network parameters improved the performance of this modeling method in estimating soil organic carbon on Sentinel-2 satellite images in the study area. Also, the performance of Sentinel-2 satellite images in estimating soil organic carbon in the study area was validated.

Keywords: Artificial neural networks, Genetic algorithm, Sentinel-2 satellite images, Soil organic carbon