



## بررسی ویژگی‌های جذب کربن آلی خاک با روش طیف‌سنجی آزمایشگاهی در مناطق مستعد تولید ریزگرد استان خوزستان

منصور چترنور<sup>۱</sup>، \* احمد لندی<sup>۲</sup>، احمد فرخیان فیروزی<sup>۳</sup>، علی اکبر نوروزی<sup>۴</sup> و حسینعلی بهرامی<sup>۵</sup>

<sup>۱</sup> دانشجوی دکتری گروه علوم و مهندسی خاک، دانشگاه شهید چمران اهواز، استاد گروه علوم و مهندسی خاک، دانشگاه شهید چمران اهواز و عضو مرکز پژوهشی منطقه‌ای ریزگردها، دانشگاه شهید چمران اهواز، <sup>۳</sup> دانشیار گروه علوم و مهندسی خاک، دانشگاه شهید چمران اهواز، <sup>۴</sup> دانشیار پژوهشی، پژوهشکده حفاظت خاک و آبخیزداری، <sup>۵</sup> دانشیار گروه خاکشناسی، دانشگاه تربیت مدرس

تاریخ دریافت: ۱۳۹۸/۰۳/۲۳؛ تاریخ پذیرش: ۱۳۹۸/۰۷/۲۸

### چکیده

**سابقه و هدف:** در سال‌های اخیر نواحی گسترده‌ای از استان خوزستان به دلیل عدم پوشش سطحی و مقاومت کم خاک در برابر باد فرساینده، مستعد تولید ریزگرد هستند. در بین ویژگی‌های خاک، ماده آلی با اتصال ذرات خاک نقش مهمی در مقاومت به فرسایش بادی و تولید ریزگرد دارد. با توجه به سطح گسترده کانون‌های ریزگرد استان خوزستان، استفاده از روش‌های سنتی تجزیه و تحلیل خاک پر هزینه و زمان‌بر است. روش طیف‌سنجی به دلیل مزیت سرعت عمل و سهولت جابجایی، می‌تواند هزینه و زمان اندازه‌گیری ویژگی‌های خاک را کاهش دهد. بر این اساس هدف از این پژوهش بررسی رفتار طیفی کربن آلی خاک در مناطق مرکز و جنوب استان خوزستان با استفاده از دو مدل رگرسیونی چند متغیره ماشین بردار پشتیبان (SVR) و شبکه عصبی (PLS-ANN) و تعیین طول موج کلیدی کربن آلی خاک در این مناطق است.

**مواد و روش‌ها:** در این پژوهش منطقه مطالعاتی به شبکه‌های ۲ در ۲ کیلومتری تقسیم گردید و نمونه‌برداری به روش سیستماتیک- تصادفی انجام شد. مقدار کربن آلی نمونه‌های خاک در آزمایشگاه اندازه‌گیری گردید. طیف بازتابی نمونه‌ها با استفاده از دستگاه Fildspec3 در اتاقک تاریک تعیین شد و اندازه‌گیری طیفی با سه نوع آشکارساز در محدوده مرئی تا مادون قرمز نزدیک (۲۵۰۰-۳۵۰۰ نانومتر) صورت گرفت. به منظور حذف نویز در طیف بازتابی، طیف اصلی با ۴ روش مشتق اول همراه با فیلتر ساویتزکی- گولای (FD-SG)، روش مشتق دوم به همراه فیلتر ساویتزکی- گولای (SD-SG)، واریانس نرمال استاندارد (SNV) و حذف پیوستار (CR) پیش‌پردازش شد. در ادامه عملکرد مدل‌های SVR و PLS-ANN در طیف اصلی و روش‌های پیش‌پردازش مورد مقایسه قرار گرفت.

**یافته‌ها:** نتایج نشان داد که مدل PLS-ANN دقت بیش‌تری نسبت به مدل SVR در برآورد کربن آلی خاک داشت. در مدل SVR پیش‌پردازش حذف پیوستار (CR) بهترین عملکرد (CR)  $(R^2_{CAL}=0/06$  و  $RPD_{CAL}=1/82$ )، و طیف اصلی (ROW)  $(R^2_{CAL}=0/84$  و  $RPD_{CAL}=1/66$ ) و طیف اصلی (R)  $(R^2_{CAL}=0/74$  و  $RPD_{CAL}=1/66$ ) کم‌ترین عملکرد را داشتند.

\* مسئول مکاتبه: landi@scu.ac.ir

در مدل PLS-ANN بهترین عملکرد مرتبط به روش مشتق دوم (SD-SG)  $RPD_{CAL}=2/34$  و  $RMSE_{CAL}=0/05$ ، کمترین عملکرد در روش مشتق اول (FD-SG)  $RPD_{CAL}=1/86$  و  $RMSE_{CAL}=0/1$  و  $R^2_{CAL}=0/92$  مشاهده شد.  $R^2_{CAL}=0/80$

**نتیجه‌گیری:** در این پژوهش روش‌های پیش‌پردازش سبب بهبود دقت کلی مدل‌های SVR و PLS-ANN نسبت به طیف اصلی شدند. با توجه به عملکرد روش مشتق دوم در مدل PLS-ANN که بهترین دقت را در برآورد کربن آلی خاک داشت، طول موج‌های ۸۰۰، ۱۸۰۰ و ۲۰۰۰ نانومتر به‌عنوان طول موج کلیدی کربن آلی خاک در برای مناطق مستعد تولید ریزگرد شناسایی شد.

**واژه‌های کلیدی:** پیش‌پردازش، حذف پیوستار، ماشین بردار پشتیبان، مدل شبکه عصبی

### مقدمه

به‌ویژه حضور یون آهن که بر بازتاب خاک اثر دارند (۲۵ و ۴۰) از عوامل دیگر می‌توان به اختصاصی بودن محدوده طیفی اشاره کرد که ممکن است این محدوده جذبی مختص ماده آلی باشد یا این که سایر ویژگی‌های خاک مانند اکسیدهای آهن و کانی‌های رسی نیز در آن اثر داشته باشند (۱). بنابراین لازم است که با استفاده از روش‌هایی اثر پارازیت و سایر عوامل مانند عوامل محیطی را جداسازی و به حداقل رساند. مسأله دیگر استفاده از مدل رگرسیونی چندمتغیره مناسب برای تعیین رابطه بین بازتاب خاک با ویژگی خاک موردنظر است. قابل‌ذکر است که به‌دلیل تفاوت در موقعیت طول موج، تعداد باند و هم‌چنین تفاوت در منطقه مطالعاتی نیاز است که دقت انواعی از مدل‌ها در محیط‌های مختلف آزمایش شود تا بهترین مدل در هر شرایط تعیین گردد. از روش‌های رگرسیون رایج استفاده‌شده در پردازش طیف می‌توان به رگرسیون مؤلفه‌های اصلی (PCR)، رگرسیون حداقل مربعات جزئی (PLSR)، رگرسیون چندمتغیره اسپیلاین<sup>۱</sup>، شبکه‌های عصبی مصنوعی (ANN) و ماشین‌های بردار پشتیبان (SVR) اشاره کرد (۱۱ و ۲۰).

در استان خوزستان نواحی متعددی وجود دارد که مستعد تولید ریزگرد است و در این اراضی ماده آلی خاک از مهم‌ترین پارامترهای خاک است که نقش مهمی در خصوصیات فیزیکی و شیمیایی خاک، باروری و کیفیت زمین‌های زراعی (۳۴)، اثر بر نگهداشت آب و انواع عناصر، تسریع در نفوذ سطحی آب در هنگام بارش و کاهش رواناب، هم‌چنین از عوامل مهم گردآوری ذرات و تشکیل خاکدانه محسوب می‌شود (۳۲). به‌منظور اندازه‌گیری ماده آلی خاک از روش‌های رایج آزمایشگاهی نسبتاً پیچیده، وقت‌گیر و گران‌قیمت استفاده می‌شود. روش طیف‌سنجی به‌دلیل مزیت سرعت عمل، سهولت جابجایی، می‌تواند هزینه و زمان اندازه‌گیری را کاهش دهد و جایگزین تجزیه و تحلیل آزمایشگاهی رایج در تعیین ویژگی‌های خاک از جمله ماده آلی باشد (۳۰). در طیف‌سنجی خاک‌ها مشاهده‌شده که ماده آلی خاک در ناحیه مادون‌قرمز نزدیک، مادون‌قرمز میانی و مرئی ویژگی‌های جذبی دارد (۱۴ و ۴۰).

استفاده از روش طیف‌سنجی در اندازه‌گیری ویژگی‌های خاک محدودیت‌هایی دارد؛ از جمله عواملی مثل رطوبت، اندازه ذرات و ترکیبات معدنی و

حافظه (MBL) استفاده کردند و نتیجه گرفتند که روش MBL با پیش‌پردازش مشتق جزئی با ضریب  $0/25$  دقت بهتری ( $RPD=2/23$  و  $RMSE=4/36$ ) داشته است ( $R^2=0/82$ ) در برآورد ماده آلی داشت (۱۷).

ماده آلی خاک با خاصیت چسبندگی و گردآوری ذرات خاک، سبب جلوگیری از پراکندگی ذرات و تولید ریز گرد می‌شود، از طرفی دیگر تعیین مقدار این پارامتر در خاک‌های مناطق مستعد تولید ریز گرد استان خوزستان با توجه به گستردگی که دارند (مساحت حدود ۳۵۰ هزار هکتاری (۱۶))، با روش‌های آزمایشگاهی سخت و پرهزینه است. در بررسی منابع صورت گرفته، پژوهشی در مورد استفاده از طیف‌سنجی و مدل‌های رگرسیونی برای ارزیابی طیفی ماده آلی در این مناطق انجام نشده است. بنابراین هدف از این پژوهش ارزیابی مقدار کربن آلی خاک در مناطق مستعد تولید ریز گرد استان خوزستان با استفاده از دو مدل رگرسیونی شبکه عصبی (PLS-ANN) و ماشین بردار پشتیبان (SVR) است که عملکرد این مدل‌ها در پنج نوع طیف شامل: طیف بازتابی اصلی، روش‌های پیش‌پردازش مشتق اول صاف شده با ساویتزکی-گولای<sup>۴</sup>، مشتق دوم صاف شده با ساویتزکی-گولای، واریانس نرمال استاندارد و روش حذف پیوسته مورد مقایسه قرار گرفت. از اهداف دیگر این مطالعه می‌توان به تعیین طول موج کلیدی تحت‌تأثیر کربن آلی در منطقه مورد مطالعه اشاره کرد.

### مواد و روش‌ها

منطقه مورد مطالعه: منطقه مطالعاتی در مرکز و جنوب استان خوزستان واقع در جنوب غربی ایران، با مساحت ۱۱۰ هزار هکتار و در موقعیت جغرافیایی با مختصات طول  $31^{\circ} 19'$  تا طول  $30^{\circ} 24'$  و عرض

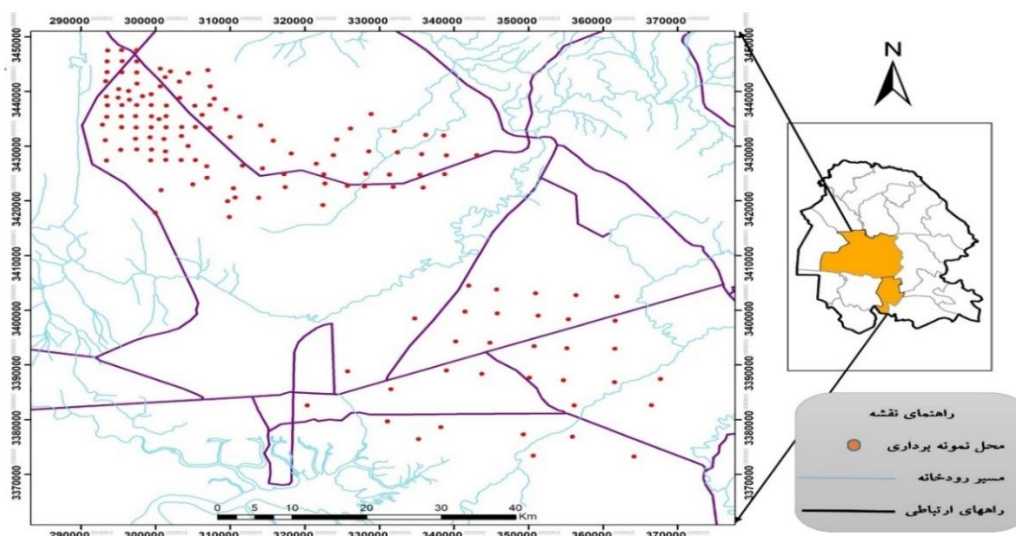
طی پژوهشی هی و همکاران (۲۰۰۹) از مدل رگرسیون گام‌به‌گام<sup>۱</sup> و روش پیش‌پردازش مشتق جزئی در برآورد ماده آلی خاک استفاده کردند و نتیجه گرفتند که مشتق با ضریب یک بهترین برآورد ( $RMSE=0/88$  و  $R^2=0/88$ ) را داشت (۱۵). تیان و همکاران (۲۰۱۳) با استفاده از مدل‌های PLSR و PLS-BPNN<sup>۲</sup> و روش‌های پیش‌پردازش مشتقات، ماده آلی خاک‌های چین را ارزیابی کرده و مشاهده نمودند که مدل PLSR در روش مشتق اول بهترین دقت برآورد ( $RMSE=3/37$  و  $R^2=0/90$ ) را داشت (۴۰). خویمی و جانشی (۲۰۱۳) در روشی مشابه از مدل‌های PLSR و LS-SVM در طیف‌سنجی خاک استفاده کردند و مشاهده کردند که مدل LS-SVM در برآورد ماده آلی، نیتروژن، فسفر و پتاسیم تبادلی در خاک دقت قابل‌قبولی را با ضریب تعیین به‌ترتیب  $0/87$ ،  $0/82$ ،  $0/76$  و  $0/73$  نشان داد (۴۹). خیامیم و همکاران (۲۰۱۵) در بررسی خاک‌های استان اصفهان با روش طیف‌سنجی، ویژگی‌های ماده آلی، کربنات کلسیم و گچ را با دقت مناسبی با ضریب تعیین ( $R^2$ )  $0/61$ ،  $0/45$  و  $0/80$  برآورد کردند (۲۱). نوار و همکاران (۲۰۱۵) از مدل‌های رگرسیونی PLSR، MARS و SVR در برآورد ویژگی‌های خاک در مصراستفاده کرده و مشاهده کردند که مدل MARS<sup>۳</sup> و روش پیش‌پردازش حذف پیوستار (CR) دقت قابل‌قبولی را در برآورد درصد رس و مواد آلی، با ضرایب تعیین به‌ترتیب  $0/85$  و  $0/90$  به‌دست آوردند (۲۷). هونگ و همکاران (۲۰۱۹) برای بهبود برآورد ماده آلی خاک با روش طیف‌سنجی از روش‌های پیش‌پردازش مشتق اول، مشتق دوم و مشتق جزئی و هم‌چنین مدل‌های حداقل مربعات جزئی (PLS) و جنگل تصادفی (RF) و یادگیری بر پایه

- 1- Stepwise regression analysis
- 2- Backpropagation neural network
- 3- Multivariate adaptive regression splines

4- Savitzky-Golay filter

است. بافت غالب منطقه سیلتی بوده و طبق آمار سازمان هواشناسی کشور تعداد روزهای گرد و غباری در یک سال بین ۸۰ تا ۱۳۳ روز متغیر است. نمونه برداری: پس از انجام مطالعات مقدماتی، محدوده مورد نظر به شبکه‌های ۲ در ۲ کیلومتری تقسیم‌بندی و نمونه برداری به صورت سیستماتیک-تصادفی انجام شد. تعداد ۱۴۲ نمونه از سطح خاک تا عمق ۵ سانتی متری جمع‌آوری شد. برای تعیین کربن آلی، ابتدا نمونه‌های خاک هوا خشک‌شده و سپس کوبیدن و عبور از الک دو میلی متری انجام شد. در آزمایشگاه مقدار کربن آلی نمونه‌های خاک با روش والکلی و بلاک (۱۹۸۲) تعیین گردید (۲۹).

جغرافیایی ۲۱' ۴۹° تا ۲۸' ۴۹° قرار دارد. متوسط درجه حرارت سالانه این منطقه ۳۱/۲ درجه سانتی‌گراد و میانگین بارش ۲۶۱ میلی‌متر است. از نظر پوشش گیاهی داری زمین لخت و فاقد پوشش گیاهی و مناطقی نیز دارای گیاهان پراکنده با تراکم بسیار کم است. همچنین کشت دیم نیز در مناطقی وجود دارد که در سال‌های اخیر به علت خشک‌سالی بسیار کم شده است. که از نظر ژئومورفولوژی در گروه دشت‌های رسوبی و زمین‌های پست و شور قرار دارد و به دلیل صعود مویینه‌ای املاح در سطح خاک تجمع یافته است. میانگین شوری این مناطق ۳۴/۸ دسی‌زیمنس بر متر است و در نقاطی شوری به ۹۶ دسی‌زیمنس بر متر رسیده و بلورهای نمک در سطح خاک نمایان شده



شکل ۱- موقعیت کانون ریز گرد و محل‌های نمونه برداری.

Figure 1. Dust center situation and soil sample location.

شد. اندازه‌گیری طیفی با سه نوع آشکارساز در محدوده (۲۵۰۰-۳۵۰۰ نانومتر) شامل محدوده مرئی تا مادون قرمز نزدیک صورت گرفت. نمونه‌های خاک در فاصله ۲۰ سانتی متری و زاویه ۴۵ درجه از لامپ هالوژنی دستگاه قرار گرفته و اندازه‌گیری طیفی به روش غیرتماسی و در ارتفاع ۳ سانتی متر از بالای

اندازه‌گیری بازتاب نمونه‌های خاک: از هر نمونه مقداری خاک به پتری‌دیشی با قطر ۸ و عمق ۲ سانتی متر منتقل شد. جهت جلوگیری از مزاحمت و نویز ناشی از نور محیطی، نمونه‌های خاک به اتاق تاریک منتقل و تعیین بازتاب خاک با استفاده از طیف‌سنج آزمایشگاهی ASD FieldSpec3 انجام

در رابطه ۲  $\lambda_i$  طول موج در هر باند،  $R'(\lambda_i)$  و  $R''(\lambda_i)$  به ترتیب مشتق اول و دوم هر برای طول موج،  $\Delta\lambda$  فاصله بین طول موج  $\lambda_i$  و  $\lambda_{i-1}$  است. در این مطالعه  $\Delta\lambda=10$  نانومتر انتخاب شده است.

واریانس نرمال استاندارد (SNV) نوعی تغییر جهت دار در منحنی طیف است و داده‌ها را در جهتی تغییر داده و سبب تمرکز آن‌ها در ستون یا دسته‌ای از داده‌ها می‌شود (۳). روش حذف پیوستار برای جدا کردن و تحلیل عارضه‌ها در طیف بازتابی استفاده می‌شود. در این روش انواع پدیده‌های جذبی موجود در طیف مورد استفاده قرار می‌گیرد (۶). بنابراین حذف پیوستار اغلب بر روی همه پدیده‌های جذبی طیف اجرا شده و برای مطالعه روابط بین عوارض جذبی و انواع ترکیبات مورد استفاده قرار می‌گیرد (۱۸). حذف پیوستار اثرات نامطلوب عوامل محیطی و جذب‌های اتمسفری را به حداقل می‌رساند (۲۲). با انجام این روش نویز ناشی از عوامل نامشخص حذف و یا به حداقل می‌رسد. در این روش واریانس مطلق حذف و جذب حداکثر بارزتر می‌شود (۹).

**روش مدل‌سازی:** مدل شبکه عصبی پرسپترون چندلایه (MLP) سیستمی محاسباتی با تعدادی از عناصر پیوسته و پاسخ‌های پویای ورودی به خروجی است (۴). مدل شبکه‌های عصبی معمولاً در لایه‌ها سازمان‌دهی می‌شوند. لایه‌ها از تعدادی گره‌های متصل به هم تشکیل شده که دارای یک تابع فعال است. الگوها از طریق لایه ورودی به شبکه ارائه می‌شود. فرایند مدل‌سازی از طریق یک یا چند لایه پنهان انجام می‌شود و پردازش واقعی از طریق سیستمی از اتصالات وزنی صورت می‌گیرد. لایه‌های پنهان به یک لایه خروجی پیوند می‌شود که پاسخ خروجی است (۲۶). در این مدل از الگوریتم پس‌انتشار که یک الگوریتم بسیار رایج در آموزش شبکه MLP است، استفاده می‌شود (۱۰).

نمونه‌ها انجام شد. برای واسنجی نوری دستگاه طیف‌سنج، از یک صفحه سفید با ضریب انعکاس معلوم و برابر یک استفاده شد. برای هر نمونه ۱۰ طیف بازتابی اندازه‌گیری شده و از طیف‌های به دست آمده در هر اسکن، میانگین‌گیری در محیط نرم‌افزار ViewSpec انجام شد و به‌عنوان طیف نمونه خاک در کتابخانه طیفی ذخیره گردید.

**پیش‌پردازش:** روش‌های پیش‌پردازش سبب حذف پارازیت ناشی از اثرات پایه‌ای، قله‌ها و باندهای همپوشانی شده و هم‌چنین بهبود قدرت تفکیک و حساسیت طیف و حذف اثرات ناشی از فاکتورهای محیطی می‌شود (۴۱ و ۵۰). در ابتدا دو بخش نویزدار در ابتدا و انتهای طیف که به ترتیب در محدوده ۳۵۰-۴۵۰ و ۲۵۰۰-۲۴۵۰ نانومتر است و هم‌چنین دو وقفه حاصل از تغییر آشکارساز<sup>۱</sup> در محدوده ۹۰۰ و ۱۷۰۰ نانومتر حذف شد (۳۵). در ادامه پیش‌پردازش‌هایی شامل: فیلتر ساویتزکی- گولای به صورت تابع چندجمله‌ای درجه دوم و نیز تعداد ۲۳ نقطه هموارساز) به همراه مشتق اول (FD-SG)، فیلتر ساویتزکی- گولای به همراه مشتق دوم (SD-SG)، روش نرمال استاندارد بر اساس میانگین به همراه فیلتر ساویتزکی- گولای و حذف پیوستار (CR) بر روی طیف اصلی انجام شد (۳).

در روش مشتق‌گیری نقاط عطف و بازتاب‌های حداکثر وضوح بیش‌تری را نشان می‌دهد. رابطه مشتق اول و دوم به صورت زیر است (۹):

$$R'(\lambda_i) = [R(\lambda_i) - R(\lambda_{i-1})] / 2\Delta\lambda \quad (1)$$

$$R''(\lambda_i) = [R'(\lambda_i) - R'(\lambda_{i-1})] / 2\Delta\lambda \\ = [R(\lambda_{i+1}) - 2R(\lambda_i) + R(\lambda_{i-1})] / \Delta\lambda^2 \quad (2)$$

$$|\xi|_{\varepsilon} = \begin{cases} 0 & \text{if } |\xi| \leq \varepsilon \\ \xi - \varepsilon & \text{otherwise} \end{cases} \quad (6)$$

$$\text{Minimize: } \frac{1}{2} (w \cdot w) + C \sum_{i=1}^N (\xi_i^+ + \xi_i^-)$$

$$\text{constraints: } \begin{cases} wx_i + b - y_i \leq \varepsilon + \xi_i^+ \\ y_i - wx_i + b \leq \varepsilon + \xi_i^- \\ \xi_i^+ \geq 0, \xi_i^- \geq 0 \end{cases} \quad (7)$$

$i = 1, 2, 3, \dots, N$

که در آن‌ها،  $w$  وزن بردار  $x$  مقدار بایاس،  $\xi_i^+$  و  $\xi_i^-$  متغیرهای کمبود و  $C$  ثابت گنجایش که خطای مربوط به فاصله از  $\varepsilon$  را کنترل می‌کند.

**معیارهای ارزیابی:** برای ارزیابی دقت مدل‌ها از سه شاخص آماری ریشه میانگین مربعات خطا (RMSE)، ضریب تعیین ( $R^2$ ) و نسبت عملکرد به انحراف (RPD) که معادلات آن‌ها در ادامه آورده شده، استفاده گردید (۵ و ۱۹).

$$R^2 = \frac{\sum_{i=1}^n (y_i - \bar{y})(y_{ij} - \bar{y}_{ij})}{\sqrt{\sum_{i=1}^n (y_i - \bar{y})^2 \sum_{i=1}^n (y_{ij} - \bar{y}_{ij})^2}} \quad (8)$$

$$\text{RMSE} = \sqrt{\frac{\sum_{i=1}^n \sum_{j=1}^m (\hat{y}_{ij} - y_i)^2}{mn}} \quad (9)$$

$$\text{RPD} = \frac{\text{SD}}{\text{RMSE}} \quad (10)$$

$$\text{SD} = \sqrt{\frac{\sum_{i=1}^n (y_i - \bar{y})^2}{n-1}} \quad (11)$$

$$\Delta w_{ij}(n) = -\varepsilon \times \frac{\partial E}{\partial w_{ij}} + \alpha \times \Delta w_{ij}(n-1) \quad (3)$$

که در آن، در این روابط  $\Delta w_{ij}(n)$  و  $\Delta w_{ij}(n-1)$  به ترتیب افزایش وزن بین گره‌های  $i$  ام و  $j$  ام،  $E$  خطا، و  $\alpha$  ضریب مومتم و ضریب یادگیری (مقدار آن‌ها بین ۰ تا ۱ است. خروجی هر نرون در رابطه زیر تعریف می‌شود:

$$y = f(u) = f\left(\sum_{i=1}^n p_i W_{j,i} + b_j\right) \quad (4)$$

که در آن،  $W_{j,i}$  مقدار وزن اتصال بین نرون  $j$  ام لایه مذکور با نرون  $i$  ام لایه قبل،  $b_j$  وزن مربوط به بایاس Bias برای نرون  $j$  ام،  $P_i$  مقدار خروجی از نرون  $i$  ام لایه قبل،  $a$  مقدار خروجی از نرون  $j$  ام،  $f$  تابع آستانه نرون  $j$  ام است.

**ماشین بردار پشتیبان<sup>۱</sup> (SVR):** این روش بر اساس نظریه مفاهیم آماری<sup>۲</sup> استوار است و جواب کلی را با کمینه کردن خطای ساختاری تعیین می‌کند (۴۳).

مجموعه توابع به منظور پیش‌بینی تعیین می‌شود و  $F(x)$  به گونه‌ای تعیین می‌شود که  $x$  بیش‌ترین حاشیه را از مقادیر آموزشی  $y$  داشته باشد. در این روش منحنی با ضخامت  $\varepsilon$ ، با کم‌ترین خطای برآورد به داده‌ها برازش داده می‌شود (۳۹). مقدار انحراف از رابطه ۶ تعیین می‌شود و با وارد کردن متغیرهای کمبود و اصل کمینه‌سازی و بهینه‌سازی خطای ساختاری<sup>۳</sup>، رابطه ۷ برقرار می‌شود.

$$f(x) = w \cdot x + b \quad (5)$$

- 1- Support Vector Regression
- 2- Statistical learning theory
- 3- Structural Risk Minimization

**نتایج و بحث**

مشخصات آماری کربن آلی اندازه‌گیری شده در نمونه‌های خاک و همچنین همبستگی بین کربن آلی با سایر ویژگی‌های خاک به ترتیب در جدول‌های ۱ و ۲ ارائه شده است. بر اساس این نتایج، مقدار کربن آلی با شوری خاک همبستگی معنی‌دار (۰/۴۶) و با درصد رس، همبستگی معنی‌دار (۰/۴۱-) در سطح ۵ درصد و با ویژگی‌های درصد شن و سیلت و مقدار گچ همبستگی معنی‌داری نداشت.

که در آن‌ها،  $y_{ij}$  مقدار برآورد شده در  $j$  تکرار در  $i$  نمونه،  $y_i$  مقدار اندازه‌گیری شده در نمونه  $i$ ،  $n$  تعداد نمونه‌ها در مراحل واسنجی و اعتبارسنجی،  $m$  تعداد تکرار و  $\bar{y}$  میانگین مقادیر اندازه‌گیری شده است. در رابطه ۱۰ SD انحراف معیار نمونه است و در رابطه ۱۱ محاسبه می‌شود. آماره RPD عملکرد مدل را ارزیابی می‌کند. مقادیر کم‌تر از ۱/۴، ۲ تا ۱/۴ و بیش‌تر از ۲ به ترتیب نشان‌دهنده عملکرد ضعیف، قابل‌قبول و عالی در مدل‌سازی است (۵).

جدول ۱- توصیف آماری کربن آلی اندازه‌گیری شده نمونه‌های خاک.

**Table 1. Statistical description of soil organic matter measured in soil samples.**

آماره							
بیشینه Max (%)	میانه Med (%)	کمینه Min (%)	میانگین MEAN (%)	ضریب تغییرات CV	انحراف معیار ST.DEV	چولگی SKEW	کشیدگی KURT
1.26	0.56	0.059	0.62	0.38	0.23	0.17	0.20

جدول ۲- همبستگی بین کربن آلی با سایر ویژگی‌های اندازه‌گیری شده خاک.

**Table 2. Correlation between soil organic carbons with other properties.**

کربن آلی Organic carbon (%)	گچ (mg/ 100gr) Gypsum	درصد رس Clay (%)	درصد شن Sand (%)	درصد سیلت Silt (%)	شوری EC (DS.M <sup>-1</sup> )	ویژگی خاک Soil properties
1	0.37	-0.41*	0.35	0.18	0.46*	کربن آلی Organic carbon (%)
	1	0.42*	-0.54*	0.43*	-0.17	گچ (mg/ 100gr) Gypsum
		1	-0.76**	-0.40*	0.55*	درصد رس Clay (%)
			1	-0.38*	-0.51*	درصد شن Sand (%)
				1	-0.23	درصد سیلت Silt (%)
					1	شوری EC (DS.M <sup>-1</sup> )



بهترین عملکرد و ضعیف‌ترین عملکرد مربوط به طیف اصلی بود. در گروه واسنجی روش‌های پیش‌پردازش دقت قابل‌قبولی ( $RPD_{CAL} < 2$ ) را نشان دادند، در گروه اعتبارسنجی طیف اصلی عملکرد ضعیفی ( $RPD_{VAL} < 1/4$ ) داشت اما چهار روش پیش‌پردازش دقت قابل‌قبولی ( $RPD_{VAL} > 1/4$ ) را در برآورد کربن آلی نشان دادند. بر اساس نتایج حاصل از عملکرد مدل‌های SVM PLS-ANN در برآورد کربن آلی خاک، مدل PLS-ANN عملکرد بهتری را نسبت به مدل SVM در هر دو سری واسنجی و اعتبارسنجی داشته است.

قابل‌ذکر است که در مرحله واسنجی در تمام روش‌های پیش‌پردازش به جز روش حذف پیوستار، عملکرد مدل PLS-ANN نسبت به مدل SVM بهتر بوده که این نتیجه با توجه به رفتار پیچیده خاک و رابطه غیرخطی بین طیف با ویژگی‌های خاک، قابل انتظار است و با پژوهش تیان و همکاران (۲۰۱۳) (۴۰) مطابقت دارد. مقایسه دو مدل PLS-ANN و SVR، در نمودار هم‌راستایی داده‌های اندازه‌گیری و برآورد شده، نشان داد که شیب مدل PLS-ANN هم‌راستایی بیشتری با خط (۱:۱) دارد و توزیع داده‌ها در اطراف این خط نشان‌دهنده برتری نسبی این مدل است. در مقایسه مراحل واسنجی و اعتبارسنجی در مدل PLS-ANN نشان داد که داده‌های مرحله واسنجی توزیع بهتری در اطراف خط (۱:۱) دارند، بر این اساس نمودارهای مرحله اعتبارسنجی در شکل ۳ آورده شده است. در این نمودارها روش مشتق دوم و حذف پیوستار توزیع بهتری در اطراف خط ۱:۱ دارد و در طیف اصلی، داده‌ها پراکندگی بیشتری نسبت به خط ۱:۱ نشان دادند.

نتایج برآورد آماری کربن آلی خاک با استفاده از مدل‌های SVR و PLS-ANN در ۵ دسته طیفی در جدول ۳ ارائه شده است. بر اساس نتایج در مدل PLS-ANN طیف اصلی و چهار روش پیش‌پردازش دارای دقت عالی ( $RPD_{CAL} = 1/4$ ) در مرحله واسنجی و دقت قابل‌قبولی ( $1/4 < RPD_{VAL} < 2$ ) در مرحله اعتبارسنجی بوده‌اند. که در این بین در مرحله واسنجی بهترین عملکرد مربوط به روش مشتق دوم (SD-SG) ( $RPD_{CAL} = 2/34$  و  $RMSE_{CAL} = 0/05$ )،  $R^2_{CAL} = 0/92$  بوده و کم‌ترین عملکرد در روش مشتق اول (FD-SG) ( $RPD_{CAL} = 1/86$  و  $RMSE_{CAL} = 0/1$ )،  $R^2_{CAL} = 0/80$  مشاهده شد. در مرحله اعتبارسنجی کم‌ترین عملکرد برآورد در طیف اصلی ( $RPD_{VAL} = 1/44$ ) و بیش‌ترین عملکرد نیز در روش مشتق دوم ( $RPD_{VAL} = 1/72$ ) و  $R^2_{VAL} = 0/70$ ،  $RMSE_{VAL} = 0/15$  و  $R^2_{VAL} = 0/74$ ،  $RMSE_{VAL} = 0/11$  مشاهده شد. در گروه واسنجی سه روش مشتق دوم، واریانس نرمال استاندارد و حذف پیوسته سبب افزایش دقت برآورد مدل شدند ولی در مشتق اول، دقت برآورد مدل نسبت به طیف اصلی کاهش نشان داد. در گروه اعتبارسنجی چهار روش پیش‌پردازش دقت برآورد بیش‌تری نسبت به طیف اصلی داشتند که نشان می‌دهد پیش‌پردازش سبب حذف نویز احتمالی و افزایش دقت برآورد مدل شده است.

در مدل SVR و مرحله واسنجی، روش پیش‌پردازش حذف پیوستار (CR) بهترین عملکرد (CR) ( $RPD_{CAL} = 1/82$  و  $RMSE_{CAL} = 0/06$ ،  $R^2_{CAL} = 0/84$ ) و طیف اصلی (ROW) ( $RPD_{CAL} = 1/66$ ) و کم‌ترین عملکرد ( $R^2_{CAL} = 0/74$ ،  $RMSE_{CAL} = 0/14$ ) را داشتند. در مرحله اعتبارسنجی روش مشتق دوم



جدول ۳- نتایج معیارهای ارزیابی مراحل واسنجی و اعتبارسنجی در مدل‌های SVR و PLS-ANN.

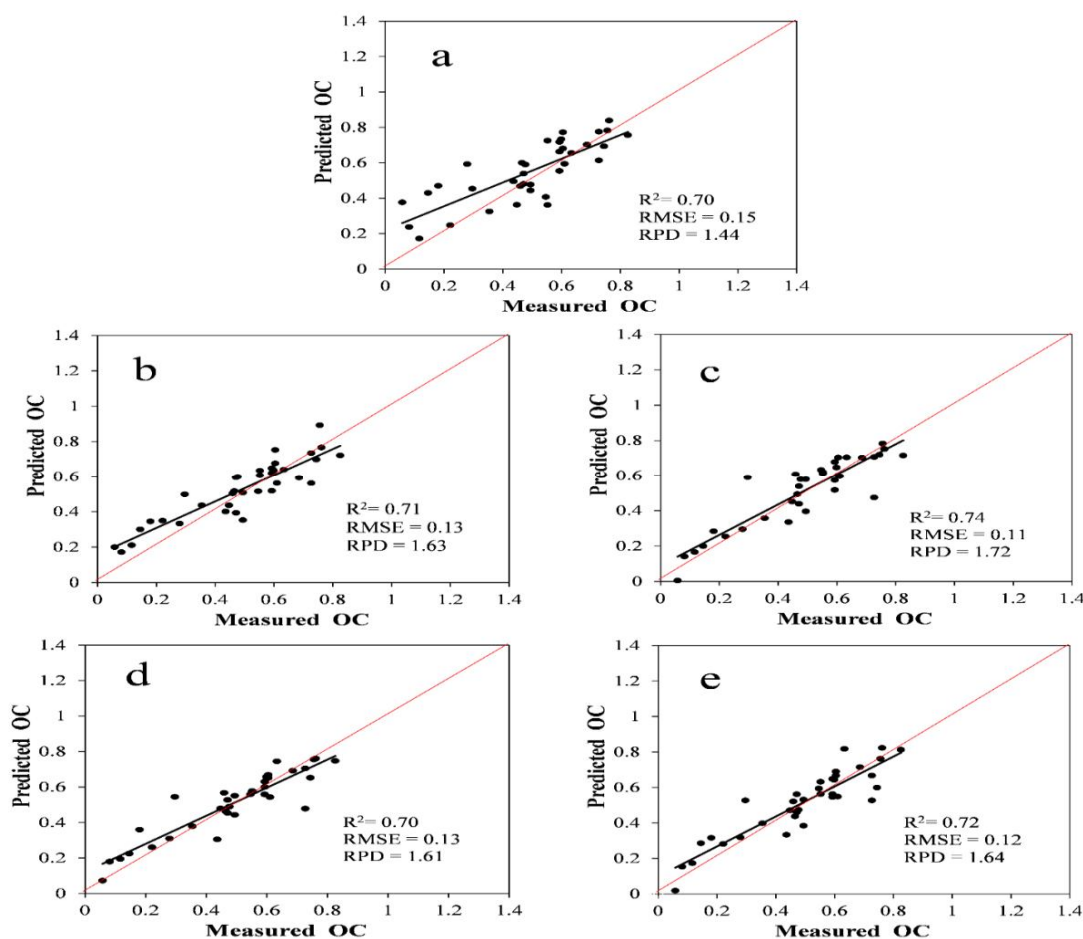
**Table 3. The results of evaluation criteria of calibration and validation steps in PLS-ANN and SVR models.**

معیارهای ارزیابی						روش‌های مدل‌سازی و پیش‌پردازش	
اعتبارسنجی			واسنجی			پیش‌پردازش	ویژگی خاک
RPD	RMSE	R <sup>2</sup>	RPD	RMSE	R <sup>2</sup>		
1.44	0.15	0.70	1.92	0.09	0.81	Row	PLSR-ANN
1.63	0.13	0.71	1.86	0.1	0.80	FD-SG	
1.72	0.11	0.74	2.34	0.05	0.92	SD-SG	
1.61	0.13	0.70	2.12	0.1	0.82	SNV	
1.64	0.12	0.72	2.10	0.09	0.85	CR	
1.38	0.17	0.52	1.66	0.14	0.74	Row	SVR
1.48	0.14	0.66	1.75	0.07	0.78	FD-SG	
1.66	0.12	0.69	1.78	0.08	0.80	SD-SG	
1.54	0.15	0.68	1.72	0.11	0.71	SNV	
1.70	0.07	0.75	1.82	0.06	0.84	CR	

روش CR با نشان دادن خصوصیات جذبی در طیف، دقت عملکرد مدل را افزایش می‌دهد (۷). که در پژوهش حاضر مدل SVR در روش حذف پیوستار بهترین عملکرد را نسبت به سایر طیف‌ها نشان داده است. در پژوهش‌هایی از این روش حذف پیوستار و سایر روش‌ها سبب حذف نویز و بهبود عملکرد مدل‌سازی طیفی شده‌اند، از جمله: پژوهش ونگ و همکاران (۲۰۰۸) که با روش حذف پیوستار در طول موج‌های ۲۰۵۲ و ۲۲۰۳ نانومتر و با روش پیش‌پردازش مشتق دوم طول موج‌هایی را در محدوده ۴۴۰، ۱۳۹۰، ۱۴۳۰، ۱۷۴۰، ۱۸۷۰، ۱۹۰۰، ۲۰۱۰ و ۲۲۷۰ نانومتر برای شوری خاک معرفی کردند (۴۶). ناوار و همکاران (۲۰۱۶) با روش‌های پیش‌پردازش حذف پیوستار، مشتق اول و مشتق دوم دقت قابل‌قبولی ( $R^2 > 0.50$ ) را برای کربن آلی به دست آوردند (۲۸). خو و همکاران (۲۰۱۶) با استفاده از روش پیش‌پردازش مشتق دوم خصوصیات جذبی را برای شوری خاک در طول موج‌های ۴۴۰، ۱۳۹۰، ۱۴۳۰، ۱۷۴۰، ۱۸۷۰، ۱۹۰۰، ۲۰۱۰ و ۲۲۷۰ نانومتر نشان دادند (۴۸).

در پژوهش حاضر روش مشتق اول در مدل PLS-ANN دقت برآورد کم‌تری نسبت به طیف اصلی نشان داد که ممکن است به دلیل کاهش شیب پایه‌ای و خطی‌سازی طیف، تولید نویز کرده و دقت برآورد را کاهش دهد (۴۴). در این پژوهش مشتق‌گیری مرتبه اول و دوم در مدل SVR و مشتق مرتبه دوم در مدل PLS-ANN سبب کاهش تغییرات و افزایش دقت مدل‌های شد. که نشان می‌دهد که این دو روش در حذف نویز مؤثر بوده‌اند (۳).

هنگامی که رابطه خطی بین غلظت خصوصیات خاک و طیف بازتابی غالب باشد مدل PLSR عملکرد بهتری دارد (۳۳). در صورتی که رابطه غیرخطی حاکم باشد دقت مدل‌های غیرخطی در برآورد ویژگی‌های خاک بیشتر است که در پژوهش حاضر و پژوهش ونگ و همکاران (۲۰۱۸) (۴۴) مدل شبکه عصبی عملکرد کلی بهتری را در برآورد کربن آلی خاک نشان داده‌اند. نتایج پژوهش حاضر با مطالعات دوتو و همکاران (۲۰۱۸) (۱۱) و استواری و همکاران (۲۰۱۸) (۳۱) مطابقت دارد.



شکل ۲- مدل‌سازی کربن آلی خاک با استفاده از داده‌های اعتبارسنجی (validation) برازش‌یافته بر مدل PLS-ANN: a: طیف اصلی (ROW)، b: مشتق اول طیف نرم‌شده با سائویتزکی- گولای (FD-SG)، c: مشتق دوم طیف نرم‌شده با سائویتزکی- گولای (SD-SG)، d: روش واریانس استاندارد نرمال (SNV)، e: حذف پیوستار (CR).

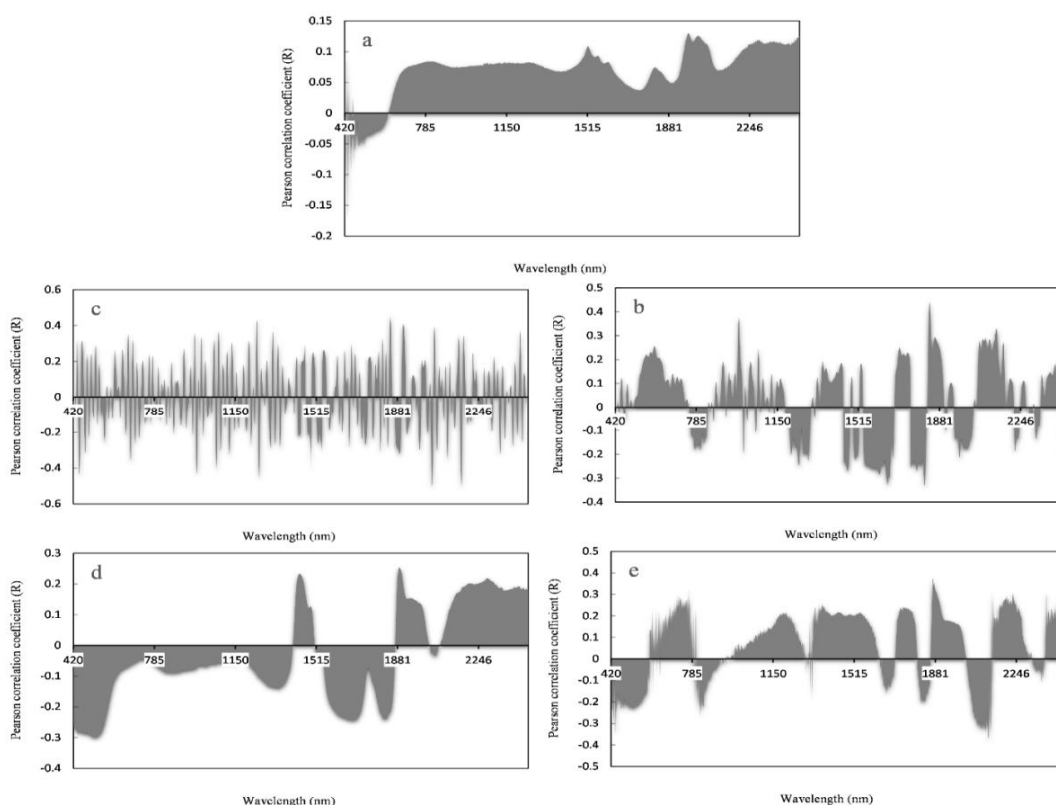
Figure 2. Modeling of Soil organic matter fitted on PLS-ANN model, using validation data. a: main spectra (ROW), b: the first derivative with the Savitzky-Golay filter (FD-SG), c: the second derivative with the Savitzky-Golay filter (SD-SG), d: the standard normal variant (SNV) and e: the continuum removal method (CR).

بر اساس پژوهش ناوار و همکاران (۲۰۱۶) (۲۸) و کونگ و موازن (۲۰۱۱) (۲۳) دامنه تغییرات غلظت ویژگی‌های خاک عامل مهمی در دقت برآورد مدل رگرسیونی است و با افزایش دامنه تغییرات دقت برآورد افزایش می‌یابد. در این پژوهش دقت گروه اعتبارسنجی نسبت به واسنجی کم‌تر بود زیرا ۳۰ درصد از کل نمونه‌ها در این گروه مورد استفاده قرار گرفت.

در پژوهش حاضر محدوده غلظت کربن آلی خاک ۰/۲۶-۰/۵۹ است و ضریب تغییرات به مقدار ۰/۳۸ است و نشان می‌دهد که داده‌های اندازه‌گیری شده از گستردگی مناسبی برخوردار بوده و سبب بهبود دقت برآورد کربن آلی در هر دو مدل شده‌اند. بر اساس پژوهش وایلینگ (۱۹۸۵) (۴۷) گستردگی داده‌ها با آماره ضریب تغییرات (CV) در محدوده ۰/۱۵-۰/۳۵ به‌عنوان گستردگی متوسط در نظر گرفته شده است و

(۲۰۱۵) (۲۷) طول موج‌هایی در محدوده ۶۰۰ نانومتر و ۱۹۰۰ نانومتر برای ماده آلی خاک گزارش کرده و هم‌چنین مشاهده کردند که طول موج ۶۰۰ نانومتر همبستگی قوی را با رس خاک نیز نشان داد. سایر ویژگی‌های خاک بر محدوده تحت‌تأثیر کربن آلی در بازتاب خاک‌ها اثرگذار هستند و در صورتی که ویژگی‌های خاک همبستگی معنی‌داری را باهم داشته باشند، این همبستگی بر روی طول موج کلیدی در این محدوده‌ها تأثیر دارد. این نتیجه برای ویژگی‌هایی مانند عناصر محلول که نشانه بارز طیفی ندارند، بیش‌تر نمایان است (۳۰). در طیف بازتابی خاک یک طول موج خاص با توجه به نوع خاک و طیف به‌دست آمده، می‌تواند با انواعی از ویژگی‌های خاک همبستگی داشته باشد. در پژوهش حاضر طول موج ۸۰۰، ۱۸۰۰ و ۲۰۰۰ نانومتر با کربن آلی خاک همبستگی داشته که در پژوهش‌های قبلی مشاهده شد که ویژگی جذب در طول موج ۲۲۰۰ نانومتر می‌تواند در اثر جذب اکسیدهای آلومینیوم باشد و ویژگی جذب کوچک در نزدیکی ۲۲۸۰ نانومتر ممکن است به‌علت اکسید آهن ایجاد شود. محدوده جذبی در ۱۱۶۰ نانومتر می‌تواند در اثر کانی کوارتز و قله نسبتاً ضعیف در نزدیکی ۱۸۰۰ نانومتر ممکن است نشانه کانی کائولینیت باشد (۲۴). طول‌موج‌های دامنه ۷۰۰-۴۰۰ نانومتر مربوط به کروموفورها و هیومیک اسید است (۳۸). که در روش مشتق دوم این طول موج همبستگی بالایی را با کربن آلی نشان داد (۳۳ و ۳۴).

محدوده‌های طیفی و طول موج کلیدی: در شکل ۳ همبستگی بین کربن آلی و طیف خاک بر اساس مدل PLS-ANN ارائه شده است. در طیف اصلی همبستگی در محدوده ۱۵۲۰-۱۵۰۰، ۲۰۳۰-۱۹۸۰ مشاهده شد. در روش پیش‌پردازش مشتق اول همبستگی‌های مثبت و منفی قوی در محدوده ۱۲۴۰، ۱۸۰۰ و ۴۴۰، ۲۰۳۰، ۱۴۵۰ و ۲۲۰۰ نانومتر، در روش مشتق دوم همبستگی در ۸۰۰، ۱۸۰۰، ۲۰۰۰، ۱۶۰۰ و ۱۸۰۰ نانومتر، در روش واریانس استاندارد نرمال، همبستگی در محدوده ۱۵۰۰، ۱۸۸۰، ۲۳۰۰ و ۵۵۰ نانومتر و در روش حذف پیوستار همبستگی در ۷۰۰، ۱۸۹۰ و ۴۵۰ نانومتر مشاهده گردید. با مراجعه به نتایج مدل‌سازی در جدول ۳ و بررسی روش‌های طیفی دارای دقت بیش‌تر در برآورد کربن آلی، مشاهده شد که مدل PLS-ANN و روش مشتق دوم بهترین برآورد کربن آلی خاک را نشان دادند. بنابراین بر اساس نتایج، همبستگی طیفی در محدوده ۸۰۰، ۱۸۰۰، ۲۰۰۰ نانومتر را می‌توان محدوده یا طول موج کلیدی را برای کربن آلی خاک معرفی کرد. در پژوهش‌های قبلی طول موج‌هایی در انواع دامنه‌های طیفی برای ماده آلی و سایر ویژگی‌های خاک تعیین شده است. طی پژوهشی راسل و همکاران (۲۰۰۶) (۳۷) طول موج‌هایی در محدوده طول موج‌های ۴۱۰، ۵۷۰ و ۶۶۰ نانومتر، وانگ و همکاران (۴۵) طول موج‌هایی در ۴۴۰، ۵۶۰، ۶۲۵، ۷۴۰ و ۱۳۳۶ نانومتر، نوکیتا و همکاران (۲۰۱۴) (۳۰) در محدوده طول موج‌های ۵۵۰-۶۸۰ نانومتر و ناوار و همکاران



شکل ۳- نمودار همبستگی بین کربن آلی با بازتاب خاک در انواع طیف‌ها شامل a: طیف اصلی (ROW)، b: مشتق اول همراه با فیلتر ساویتزی- گولای (FD-SG)، c: مشتق دوم همراه با فیلتر ساویتزی- گولای (SD-SG)، d: روش واریانس نرمال استاندارد (SNV)، e: حذف پیوستار (CR).  
**Figure 3. Diagram of Correlation between soil organic carbon with reflection in different spectra including: a: main spectra (ROW), b: the Savitzky-Golay filter (SG), c: the first derivative with the Savitzky-Golay filter (FD-SG), d: the second derivative with the Savitzky-Golay filter (SD-SG), e: the standard normal variant (SNV) and f: the continuum removal method (CR).**

مقدار جذب نور مواد آلی خاک، در محدوده مرئی بیش‌تر از مادون‌قرمز است (۲). بنابراین در پژوهش‌هایی از رنگ خاک در تشخیص ماده آلی استفاده شده است. راسل و همکاران (۲۰۰۸) (۳۶) و فیسترو و همکاران (۲۰۰۲) (۱۳) از محدوده مرئی برای تعیین ماده آلی خاک استفاده کرده‌اند. یودلهوون و همکاران (۲۰۰۳) (۴۲) گزارش دادند که در صورت وجود ترکیباتی مانند گچ و آهک در خاک، سبب روشن شدن رنگ نمونه خاک و همچنین تولید خطا در برآورد ماده آلی خاک می‌شود.

با توجه به شور بودن منطقه مطالعاتی در استان خوزستان که در برخی نقاط تا شوری شدید ۹۰ دسی‌زیمنس بر متر هم رسیده است. در بازتاب این خاک‌ها می‌توان طول موج‌هایی را به انواع کانی‌های نمکی و همچنین به شوری خاک نسبت داد. در پژوهش‌هایی مشاهده شده که ویژگی‌های جذبی در طول موج‌های ۱۰۰۰، ۱۲۰۰، ۱۴۰۰ و ۱۹۰۰ مربوط به نمک‌های هیدراته بیسکوفیت<sup>۱</sup> و اپسومیت<sup>۲</sup> و طول موج‌های ۱۴۰۰ و ۱۹۰۰ ویژگی‌های جذب نمک‌های هالیت<sup>۳</sup> و دندارنیت<sup>۴</sup> (۸ و ۱۲) است.

- 1- Bischofite
- 2- Epsomite
- 3- Halite
- 4- Thendarnite

روش پیش‌پردازش با بیش‌ترین دقت، بیش‌ترین فاکتور معنادار را دارد بر این اساس می‌توان نتیجه گرفت که در این روش نیز پیش‌پردازش مشتق دوم بیش‌ترین فاکتور و بیش‌ترین دقت برآورد را داشته است. از این فاکتورها به‌عنوان ورودی مدل شبکه عصبی استفاده گردید.

**معادلات رگرسیونی در مدل PLSR:** در جدول ۴ روابط رگرسیونی حاصل از مدل PLSR در طیف اصلی و پیش‌پردازش‌های مشتق اول، مشتق دوم، واریانس نرمال استاندارد و روش حذف پیوستار برای برآورد کربن آلی خاک ارائه شده است. فاکتورهای اولیه حاوی اطلاعات اصلی هستند. درحالی‌که فاکتورهای بعدی نویزها را هم شامل می‌شود (۳).

جدول ۴- رابطه رگرسیونی تعیین‌شده با استفاده از مدل PLSR برای کربن آلی خاک در انواع روش‌های پیش‌پردازش.

**Table 4. The regression equation determined using the PLSR model for soil organic matter in a variety of preprocessing methods.**

رابطه رگرسیونی	طیف
$Y = 0.569 + 15.19X_1 + 42.56X_2 + 49.62X_3$	طیف اصلی
$Y = 0.0163 + 0.0009X_1 - 0.011X_2 + 0.002X_3 - 0.06X_4 - 0.05X_5 - 0.05X_6$	مشتق اول
$Y = 0.564 + 1743.98X_1 + 7508.8X_2 + 5748.12X_3 + 7607.87X_4 + 6044.17X_5 + 2534.26X_6 + 5322.6X_7 + 3715.64X_8$	مشتق دوم
$Y = 0.32 + 0.0125X_1 + 0.19X_2 - 0.124X_3 - 0.75X_4 + 0.7X_5 - 0.11X_6$	واریانس نرمال استاندارد
$Y = 0.21 + 0.07X_1 - 0.19X_2 - 0.079X_3 - 0.641X_4 - 0.12X_5 + 0.20X_6$	حذف پیوستار

نسبتاً بهتری نسبت به سایر روش‌ها داشت. در مدل PLS-ANN روش مشتق دوم و در مدل SVR روش حذف پیوستار بهترین عملکرد را نشان دادند. استفاده از روش رگرسیون PLS و تعیین مؤلفه‌های معنادار در هر دسته طیفی سبب کاهش حجم محاسبات و افزایش دقت سرعت پردازش اطلاعات شد. در این پژوهش در هر روش طیفی محدوددهایی به‌عنوان طول موج با همبستگی بالا با کربن آلی به‌دست آمد. با مراجعه به نتایج برآورد کربن آلی خاک در مدل PLS-ANN و تعیین روش پیش‌پردازش با بهترین عملکرد، محدوده ۸۰۰، ۱۸۰۰ و ۲۰۰۰ نانومتر به‌عنوان محدوده تحت‌تأثیر و طول موج کلیدی برای کربن آلی در مناطق مستعد تولید گرد و غبار معرفی گردید. با

### نتیجه‌گیری کلی

این مطالعه در مرکز و جنوب استان خوزستان در مناطقی با پتانسیل تولید ریزگرد انجام گرفت. بر این اساس عملکرد دو مدل SVR و PLS-ANN در ۵ دسته طیفی شامل: طیف اصلی (ROW)، پیش‌پردازش‌های SD-SG، FD-SG، SNV و CR در برآورد کربن آلی خاک مورد مقایسه قرار گرفت. یافته‌های پژوهش عبارتند از:

بر اساس شاخص‌های آماری  $R^2$ ، RMSE و RPD، مدل PLS-ANN عملکرد بهتری نسبت به مدل SVR در برآورد کربن آلی نشان داد. در برآورد کربن آلی خاک، نوع عملکرد دو مدل نسبت به انواع روش‌های پیش‌پردازش تفاوت بیش‌تری را نشان دادند و در هر مدل یک روش پیش‌پردازش عملکرد

است از طول موج کلیدی به همراه روش سنجش از دور برای تهیه نقشه دقیق‌تر کربن آلی خاک این مناطق استفاده شده است.

ترکیب محدوده تحت‌تأثیر کربن آلی در روش طیف‌سنجی با روش‌های دورسنجی، می‌توان دقت نقشه‌های برآورد کربن آلی را افزایش داد که در مطالعه‌ای که به موازات این پژوهش در حال انجام

## منابع

1. Bartholomeus, H., Schaepman, M., Kooistra, L., Stevens, A., Hoogmoed, W., and Spaargaren, O. 2008. Spectral reflectance based indices for soil organic carbon quantification. *Geoderma*, 145: 1-2. 28-36.
2. Baumgardner, M.F., Silva, L.F., Biehl, L.L., and Stoner, E.R. 1986. Reflectance properties of soils *Advances in Agronomy*, 38: 1-44.
3. CAMO, A. 1998. The Unscrambler User Manual. CAMO ASA Norway. Pp: 72-82.
4. Caudill, M. 1987. Neural networks primer, part I. *AI expert*, 2: 12. 46-52.
5. Chang, C.W., Laird, D.A., Mausbach, M.J., and Hurburgh, C.R. 2001. Near-infrared reflectance spectroscopy-principal components regression analyses of soil properties. *Soil Sci. Soc. Amer. J.* 65: 2. 480-490.
6. Clark, R.N. 1999. Spectroscopy of rocks and minerals, and principles of spectroscopy. *Manual of Remote Sensing*, 3: 3-58.
7. Clark, R.N., and Roush, T.L. 1984. Reflectance spectroscopy: Quantitative analysis techniques for remote sensing applications. *J. Geophys. Res. Solid Earth*. 89: B7. 6329-6340.
8. Crowley, J.K. 1991. Visible and near-infrared (0.4-2.5  $\mu\text{m}$ ) reflectance spectra of playa evaporite minerals. *J. Geophys. Res. Solid Earth*. 96: B10. 16231-16240.
9. Curran, P.J., Dungan, J.L., and Peterson, D.L. 2001. Estimating the foliar biochemical concentration of leaves with reflectance spectrometry: testing the Kokaly and Clark methodologies. *Remote Sensing of Environment*, 76: 3. 349-359.
10. Demuth, H., and Beale, M. 1998. *Neural network toolbox: For use with MATLAB*, Natick, MA: The Math Works. Inc, 14-79.
11. Dotto, A.C., Dalmolin, R.S.D., ten Caten, A., and Grunwald, S. 2018. A systematic study on the application of scatter-corrective and spectral-derivative preprocessing for multivariate prediction of soil organic carbon by Vis-NIR spectra. *Geoderma*, 314: 262-274.
12. Farifteh, J., Van der Meer, F., Van der Meijde, M., and Atzberger, C. 2008. Spectral characteristics of salt-affected soils: A laboratory experiment. *Geoderma*, 145: 3-4. 196-206.
13. Fyströ, G. 2002. The prediction of C and N content and their potential mineralisation in heterogeneous soil samples using Vis-NIR spectroscopy and comparative methods. *Plant and soil*, 246: 2. 139-149.
14. Gaffey, S., McFadden, L., Nash, D., and Pieters, C. 1993. Ultraviolet, visible and near-infrared reflectance spectroscopy: Laboratory spectra of geologic materials. *Remote geochemical analysis: Elemental and Mineralogical Composition*, 151: 43-77.
15. He, T., Wang, J., Lin, Z., and Cheng, Y. 2009. Spectral features of soil organic matter. *Geo-spatial Information Science*, 12: 1. 33-40.
16. Heidarian, P., Azhdari, A., Joudaki, M., Khatooni, J.D., and Firoozjaei, S.F. 2018. Integrating Remote Sensing, GIS, and Sedimentology Techniques for Identifying Dust Storm Sources: A Case Study in Khuzestan, Iran. *J. Ind. Soc. Rem. Sens.* 46: 7. 1113-1124.
17. Hong, Y., Chen, S., Liu, Y., Zhang, Y., Yu, L., Chen, Y., Liu, Y., Cheng, H., and Liu, Y. 2019. Combination of fractional order derivative and memory-based learning algorithm to improve the estimation accuracy of soil organic matter by visible and near-infrared spectroscopy. *Catena*, 174: 104-116.



18. Huang, Z., Turner, B.J., Dury, S.J., Wallis, I.R., and Foley, W.J. 2004. Estimating foliage nitrogen concentration from HYMAP data using continuum removal analysis. *Remote Sensing of Environment*, 93: 1-2. 18-29.
19. Ji, W., Adamchuk, V.I., Biswas, A., Dhawale, N.M., Sudarsan, B., Zhang, Y., Rossel, R.A.V., and Shi, Z. 2016. Assessment of soil properties in situ using a prototype portable MIR spectrometer in two agricultural fields. *Biosystems Engineering*, 152: 14-27.
20. Ji, W., Shi, Z., Huang, J., and Li, S. 2016. Correction: In Situ Measurement of Some Soil Properties in Paddy Soil Using Visible and Near-Infrared Spectroscopy. *PloS one*, 11: 7. P. e0159785.
21. Khayamim, F., Khademi, H., Stenberg, B., and Wetterlind, J. 2015. Capability of vis-NIR Spectroscopy to Predict Selected Chemical Soil Properties in Isfahan Province. *JWSS-Isfahan University of Technology*, 19: 72. 81-92.
22. Kokaly, R.F. 2011. PRISM: Processing routines in IDL for spectroscopic measurements (installation manual and user's guide, version 1.0): NO: 193. P. 432.
23. Kuang, B., and Mouazen, A. 2011. Calibration of visible and near infrared spectroscopy for soil analysis at the field scale on three European farms. *Europ. J. Soil Sci.* 62: 4. 629-636.
24. Le Guillou, F., Wetterlind, W., Rossel, R.V., Hicks, W., Grundy, M., and Tuomi, S. 2015. How does grinding affect the mid-infrared spectra of soil and their multivariate calibrations to texture and organic carbon? *Soil Research*, 53: 8. 913-921.
25. Lobell, D.B., and Asner, G.P. 2002. Moisture effects on soil reflectance. *Soil Sci. Soc. Amer. J.* 66: 3. 722-727.
26. Mohamed, E., Saleh, A., Belal, A., and Gad, A.A. 2018. Application of near-infrared reflectance for quantitative assessment of soil properties. *The Egypt. J. Rem. Sens. Space Sci.* 21: 1. 1-14.
27. Nawar, S., Buddenbaum, H., and Hill, J. 2015. Estimation of soil salinity using three quantitative methods based on visible and near-infrared reflectance spectroscopy: a case study from Egypt. *Arabi. J. Geosci.* 8: 7. 5127-5140.
28. Nawar, S., Buddenbaum, H., Hill, J., Kozak, J., and Mouazen, A.M. 2016. Estimating the soil clay content and organic matter by means of different calibration methods of vis-NIR diffuse reflectance spectroscopy. *Soil and Tillage Research*, 155: 510-522.
29. Nelson, D.W., and Sommers, L.E. 1982. Total carbon, organic carbon and organic matter: P 539-579. In: A.L. Page, R.H. Miller and D.R. Keeney. *Methods of soil analysis. Part 2 Chemical and Microbiological Properties*, Madison, WI.
30. Nocita, M., Stevens, A., Toth, G., Panagos, P., van Wesemael, B., and Montanarella, L. 2014. Prediction of soil organic carbon content by diffuse reflectance spectroscopy using a local partial least square regression approach. *Soil Biology and Biochemistry*, 68: 337-347.
31. Ostovari, Y., Ghorbani-Dashtaki, S., Bahrami, H.A., Abbasi, M., Dematte, J.A.M., Arthur, E., and Panagos, P. 2018. Towards prediction of soil erodibility, SOM and CaCO<sub>3</sub> using laboratory Vis-NIR spectra: A case study in a semi-arid region of Iran. *Geoderma*, 314: 102-112.
32. Powlson, D., Brookes, P., Whitmore, A., Goulding, K., and Hopkins, D. 2011. Soil organic matters. *Europ. J. Soil Sci.* 62: 1-62.
33. Rossel, R.V., and Behrens, T. 2010. Using data mining to model and interpret soil diffuse reflectance spectra. *Geoderma*, 158: 1-2. 46-54.
34. Rossel, R.V., Behrens, T., Ben-Dor, E., Brown, D., Demattê, J., Shepherd, K.D., Shi, Z., Stenberg, B., Stevens, A., and Adamchuk, V. 2016. A global spectral library to characterize the world's soil. *Earth-Science Reviews*, 155: 198-230.



35. Rossel, R.V., Cattle, S.R., Ortega, A., and Fouad, Y. 2009. In situ measurements of soil colour, mineral composition and clay content by vis-NIR spectroscopy. *Geoderma*, 150: 3-4. 253-266.
36. Rossel, R.V., Fouad, Y., and Walter, C. 2008. Using a digital camera to measure soil organic carbon and iron contents. *biosystems engineering*, 100: 2. 149-159.
37. Rossel, R.V., Walvoort, D., McBratney, A., Janik, L.J., and Skjemstad, J. 2006. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma*, 131: 1. 59-75.
38. Shi, T., Chen, Y., Liu, H., Wang, J., and Wu, G. 2014. Soil organic carbon content estimation with laboratory-based visible-near-infrared reflectance spectroscopy: Feature selection. *Applied Spectroscopy*, 68: 8. 831-837.
39. Smola, A.J., and Schölkopf, B. 2004. A tutorial on support vector regression. *Statistics and Computing*, 14: 3. 199-222.
40. Tian, Y., Zhang, J., Yao, X., Cao, W. and Zhu, Y. 2013. Laboratory assessment of three quantitative methods for estimating the organic matter content of soils in China based on visible/near-infrared reflectance spectra. *Geoderma*, 202: 161-170.
41. Tong, P., Du, Y., Zheng, K., Wu, T., and Wang, J. 2015. Improvement of NIR model by fractional order Savitzky-Golay derivation (FOSGD) coupled with wavelength selection. *Chemometrics and Intelligent Laboratory Systems*, 143: 40-48.
42. Udelhoven, T., Emmerling, C., and Jarmer, T. 2003. Quantitative analysis of soil chemical properties with diffuse reflectance spectrometry and partial least-square regression: A feasibility study. *Plant and soil*, 251: 2. 319-329.
43. Vapnik, V., and Vapnik, V. 1998. *Statistical learning theory* Wiley. New York. Pp: 156-160.
44. Wang, J., Ding, J., Abulimiti, A., and Cai, L. 2018. Quantitative estimation of soil salinity by means of different modeling methods and visible-near infrared (VIS-NIR) spectroscopy, Ebinur Lake Wetland, Northwest China. *PeerJ*, 6, p.e4703.
45. Wang, J., He, T., Lv, C., Chen, Y., and Jian, W. 2010. Mapping soil organic matter based on land degradation spectral response units using Hyperion images. *Inter. J. Appl. Earth Observ. Geoinfor.* 12: 171-180.
46. Weng, Y., Gong, P., and Zhu, Z. 2008. Soil salt content estimation in the Yellow River delta with satellite hyperspectral data. *Can. J. Rem. Sens.* 34: 3. 259-270.
47. Wilding, L. 1985. Spatial variability: its documentation, accommodation and implication to soil surveys. Paper presented at the Soil spatial variability. Las Vegas NV, 30 November-1 December 1984 (pp. 166-194).
48. Xu, C., Zeng, W., Huang, J., Wu, J., and van Leeuwen, W. 2016. Prediction of soil moisture content and soil salt concentration from hyperspectral laboratory and field data. *Remote Sensing*, 8: 1. p. 42.
49. Xuemei, L., and Jianshe, L. 2013. Measurement of soil properties using visible and short wave-near infrared spectroscopy and multivariate calibration. *Measurement*, 46: 10. 3808-3814.
50. Zheng, K.Y., Zhang, X., Tong, P.J., Yao, Y., and Du, Y.P. 2015. Pretreating near infrared spectra with fractional order Savitzky-Golay differentiation (FOSGD). *Chinese Chemical Letters*, 26: 3. 293-296.



## Investigation of absorbance characteristics of soil organic carbon using laboratory spectroscopy in dust sensitive areas of Khuzestan province, Iran

M. Chatrenour<sup>1</sup>, \*A. Landi<sup>2</sup>, A. Farrokhian Firouzi<sup>3</sup>, A.A. Noroozi<sup>4</sup>  
and H.A. Bahrami<sup>5</sup>

<sup>1</sup>Ph.D. Student, Dept. of Soil Science and Engineering, Shahid Chamran University of Ahvaz,  
<sup>2</sup>Professor, Dept. of Soil Science and Engineering, Shahid Chamran University of Ahvaz and Dust  
Research Center, Shahid Chamran University of Ahvaz, <sup>3</sup>Associate Prof., Dept. of Soil Science and  
Engineering, Shahid Chamran University of Ahvaz, <sup>4</sup>Associate Prof., Soil Conservation and Watershed  
Management Research Institute, <sup>5</sup>Associate Prof., Dept. of Soil Science, Tarbiat Modares University  
Received: 06.13.2019; Accepted: 10.20.2019

### Abstract

**Background and Objectives:** In recent years, due to the lack of surface coating and low soil resistance to wind erosion, the large area of Khuzestan province is sensitive to dust production. Among the soil characteristics, the organic matter by collecting soil particles, has important role in soil resistance to wind erosion and dust production. Since these areas are so wide, the use of traditional methods of soil analysis is really costly and time consuming. The spectroscopy approach, due to the advantage of speed and easy movement, can reduce the cost and time of measurement. The aim of this study is to investigate the spectral behavior of soil organic carbon in central and southern regions of Khuzestan province by using two multivariate regression, Support Vector Regression (SVR) and neural network (PLS-ANN) and key wavelength determination of soil organic matter in these areas.

**Materials and Methods:** In this research, the study area was divided into 2 km square grids and systematic and random sampling Methods were performed. The soil organic matter in samples was measured in the laboratory. The Reflectance spectra of soil samples were determined using FieldSpec setup in dark room. And spectral measurements were carried out with three types of detectors in range of visible to near infrared (3500-2500 nm). To eliminate the noise in normal reflectance spectra, the main spectra were preprocessed by four methods, including the first derivative with the Savitzky-Golay filter (FD-SG), the second derivative with the Savitzky-Golay filter (SD-SG), the standard normal variant method (SNV) and the continuum removed method (CR). Next, the performance of SVR and PLS-ANN models in main spectra preprocessed method were compared.

**Results:** The results showed that the PLS-ANN model had better accuracy compared to SVR model in estimating organic carbon. In SVR models, the continuum removal method (CR) had the best performance ( $R^2_{CAL}=0.84$ ,  $RMSE_{CAL}=0.06$  and  $RPD_{CAL}=1.82$ ) and the main Spectra had the worst performance ( $R^2_{CAL}=0.74$ ,  $RMSE_{CAL}=0.14$  and  $RPD_{CAL}=1.66$ ). In PLS-ANN models, the best performance belonged to the second derivative (SD-SG), ( $R^2_{CAL}=0.92$ ,  $RMSE_{CAL}=0.05$  and  $RPD_{CAL}=2.34$ ) and the worst performance was related to the first derivative (FD-SG), ( $R^2_{CAL}=0.80$ ,  $RMSE_{CAL}=0.1$  and  $RPD_{CAL}=1.86$ ).

**Conclusion:** In this study, the preprocessing methods improved the overall accuracy of SVR and PLS-ANN models compared to the main spectrum. According to the second derivative method, in PLS-ANN which had the best accuracy in estimating soil organic carbon, the Wavelength ranges around 800, 1800 and 2000 nm were identified as the key wavelength of the organic carbon in sensitive centers to dust production.

**Keywords:** Continuum Removal Method, Neural Network, Preprocessing, Support Vector Machine

\* Corresponding Author; Email: landi@scu.ac.ir

