

برآورد مقدار کربن آلی خاک با استفاده از داده‌های ابرطیفی در گستره VIS-NIR-SWIR

ابراهیم بابائیان^۱ و * وحیدرضا جلالی^۲

^۱ استادیار پژوهشکده حفاظت خاک و آبخیزداری، استادیار گروه علوم خاک، دانشگاه شهید باهنر کرمان

تاریخ دریافت: ۹۴/۵/۲۴؛ تاریخ پذیرش: ۹۴/۹/۲۲

چکیده

سابقه و هدف: مقدار کربن آلی خاک به‌عنوان یکی از مهم‌ترین ویژگی‌های خاک، اهمیت بسیاری در توسعه و مدیریت پایدار کشاورزی دارد. مقدار مواد آلی خاک که عموماً بر اساس مقدار کربن آلی آن اندازه‌گیری می‌شود، معمولاً به‌عنوان شاخص کلیدی کیفیت و عملکرد خاک قلمداد می‌گردد. وجود مواد آلی مکفی خاک، تأثیرات مثبتی بر باروری خاک، ظرفیت نگهداشت آب خاک و ترسیب کربن دارد. پژوهش‌های اخیر نشان داده است که فعالیت‌های انسانی همانند عملیات مکانیکی کاشت و فعالیت‌های توسعه‌ای اقتصادی منجر به درخطر افتادن محتوای مواد آلی خاک شده است. بنابراین پایش مستمر میزان کربن آلی خاک می‌تواند نقشی مهمی در کنترل کیفیت و عملکرد خاک داشته باشد. از آن‌جا که روش‌های معمول آزمایشگاهی اندازه‌گیری کربن آلی خاک، به‌ویژه در مقیاس‌های مکانی بزرگ، عمدتاً پرهزینه و زمان‌بر هستند به همین دلیل، روش‌های ارزیابی سریع، کم‌هزینه و با دقت میزان کربن آلی خاک، می‌تواند یک اقدام بسیار ارزشمند برای مدیریت بلندمدت خاک باشد. هدف از انجام این پژوهش مطالعه رفتار طیفی خاک به‌منظور برآورد مقدار کربن آلی خاک با استفاده از روش‌های رگرسیون حداقل مربعات جزئی (PLSR) و رگرسیون حداقل مربعات جزئی توأم با بازنمونه‌گیری (bagging-PLSR) و نیز بررسی اثر روش‌های مختلف پیش‌پردازش داده‌های طیفی بر دقت برآورد کربن آلی خاک بود.

مواد و روش‌ها: تعداد ۲۰۰ نمونه خاک در مقیاس حوزه آبخیز (داده‌های واسنجی) و ۴۰ نمونه در مقیاس مزرعه (داده‌های اعتبارسنجی) از دو عمق خاک (صفر تا ۱۰ و ۱۰ تا ۳۰ سانتی‌متر خاک) به‌صورت نمونه‌برداری مرکب در یک شعاع ۱۰ متری جمع‌آوری و پس از هوا خشک کردن از الک ۲ میلی‌متر عبور داده شد. برخی ویژگی‌های فیزیکوشیمیایی خاک‌ها در آزمایشگاه اندازه‌گیری شد. پس از آن، بازتاب طیفی خاک در گستره مرئی، مادون‌قرمز نزدیک و میانی (۲۵۰۰-۳۵۰۰ نانومتر) با استفاده از دستگاه اسپکترومتر اندازه‌گیری و ارتباط بین ۲۰۰۰ باندها با مقدار کربن آلی خاک مورد بررسی قرار گرفت.

یافته‌ها: بر اساس نتایج به‌دست آمده، بهترین روش‌های پیش‌پردازش به‌منظور واسنجی مدل PLSR شامل حذف روند به روش موجک (RPD=۱/۹۴) و SNV توأم با فیلتر میانه (RPD=۱/۹۲) بود. تعداد فاکتور مناسب مدل‌های تخمینگر در دو روش PLSR و bagging-PLSR برابر ۱۷ به‌دست آمد. نتایج اعتبارسنجی مدل‌ها به‌زای تمامی روش‌های پیش‌پردازش نشان داد، روش‌های PLSR و bagging-PLSR به‌ترتیب با گستره مقادیر RMSE برابر با

* مسئول مکاتبه: v.jalali@uk.ac.ir

۱/۱۱٪-۱/۱۶۷ و ۰/۱۴۲-۱/۰۳٪ در برآورد مقدار کربن آلی خاک همراه بودند. همچنین، به‌ازای مقادیر کربن آلی $1/2$ درصد، دقت تخمین در هر دو روش کاهش یافت.

نتیجه‌گیری: یافته‌های پژوهش نشان داد که استفاده از بازتاب طیفی خاک در محدوده مرئی، مادون‌قرمز نزدیک و مادون‌قرمز میانی (VIS، NIR و SWIR) می‌تواند مقدار کربن آلی خاک را مورد مطالعه قرار دهد. همچنین منحنی‌های طیفی خاک‌های مختلف سه مشخصه جذب در طول‌موج‌های ۱۴۱۴، ۱۹۱۳ و ۲۲۰۷ نانومتر را نشان داد که مربوط به مقدار آب موجود در شبکه کانی‌های رسی خاک و رطوبت هیگروسکوپیک خاک بوده و می‌تواند به‌عنوان مشخصه‌های منحصر‌به‌فرد هر خاک محسوب شود. این باندها اهمیت زیادی در برآورد مقدار کربن آلی خاک دارند. پیش‌پردازش داده‌های طیفی خاک و انتخاب مناسب‌ترین روش پیش‌پردازش، یکی از مهم‌ترین عوامل تأثیرگذار بر دقت روش‌های رگرسیون حداقل مربعات جزئی (PLSR) و رگرسیون حداقل مربعات معمولی با بازمنونه‌گیری (bagging-PLSR) در برآورد مقدار کربن آلی خاک بود. بر اساس نتایج به‌دست آمده، روش bagging-PLSR نسبت به روش PLSR دقت بالاتری در برآورد مقدار کربن آلی خاک نشان داد.

واژه‌های کلیدی: بازتاب طیفی خاک، پیش‌پردازش طیفی، رگرسیون حداقل مربعات جزئی، کربن آلی خاک

مقدمه

پایش‌های زیست‌محیطی و حرکت به‌سوی کشاورزی دقیق و پایدار نیاز به اطلاع از وضعیت خاک با دقت مناسب و هزینه حداقل است. از طرفی نیاز روزافزون به مدیریت پایدار اراضی، اهمیت توجه به کیفیت خاک و شاخص‌های مربوطه را بیش از پیش می‌طلبد. تاکنون مطالعات زیادی در مورد ارائه شاخص‌های کیفیت خاک انجام شده است که در بسیاری موارد اتفاق‌نظر کافی در مورد چگونگی تفسیر این شاخص‌ها حاصل نشده است (۵). مقدار کربن آلی خاک به‌عنوان یکی از مهم‌ترین شاخص‌های مرتبط با کیفیت خاک مطرح و در بسیاری از روابط کیفی خاک مدنظر قرار گرفته است (۱). همچنین برخی معتقدند، کربن آلی خاک به‌عنوان تنها شاخص مهم مطالعه کیفیت خاک مطرح است (۳۴). در بسیاری موارد کاهش مقدار کربن آلی خاک به‌عنوان مناسب‌ترین معیار ارزیابی شروع تخریب اراضی مطرح است (۲۰). اطلاع از مقدار کربن آلی خاک به‌ویژه در مقیاس‌های مکانی بزرگ نیاز به

نمونه‌برداری‌های زیاد خاک و اندازه‌گیری‌های متعدد آزمایشگاهی داشته که خود هزینه و زمان زیادی را می‌طلبد.

تاکنون مطالعات مختلفی در مورد استفاده از رفتار طیفی خاک به‌منظور برآورد ویژگی‌های مختلف خاک انجام شده است (۱۹، ۷). در بیش‌تر این مطالعات از داده‌های چندطیفی استفاده شده است. استفاده از داده‌های ابرطیفی که حاوی چند صد دامنه طیفی است، می‌تواند نتایج دقیق‌تری در مطالعات ویژگی‌های خاک ارائه نماید. پژوهش‌ها نشان داده که در گستره مرئی و مادون‌قرمز نزدیک (VIS-NIR)، طول‌موج ۴۰۰ تا ۳۰۰۰ نانومتر) و مادون‌قرمز میانی (MIR، طول‌موج ۳۰۰۰ تا ۳۰۰۰۰ نانومتر) می‌توان مقدار ماده آلی خاک را مورد مطالعه قرار داد (۱۲). همچنین در برخی دیگر از مطالعات، گستره ۴۰۰ تا ۲۵۰۰ نانومتر برای مطالعه مقدار ماده آلی خاک توصیه شده است (۳، ۳۳). در پژوهش‌های مختلف، گستره گوناگونی از دقت برای برآورد مقدار ماده آلی خاک ارائه شده است. به‌عنوان مثال، با استفاده از روش‌هایی

به منظور برآورد مقدار کربن آلی خاک با استفاده از روش‌های رگرسیون حداقل مربعات جزئی (PLSR) و رگرسیون حداقل مربعات جزئی توأم با بازنمونه‌گیری (bagging-PLSR) و (۲) مطالعه اثر روش‌های مختلف پیش‌پردازش داده‌های طیفی بر دقت مدل‌های تخمینگر مقدار کربن آلی خاک بود.

مواد و روش‌ها

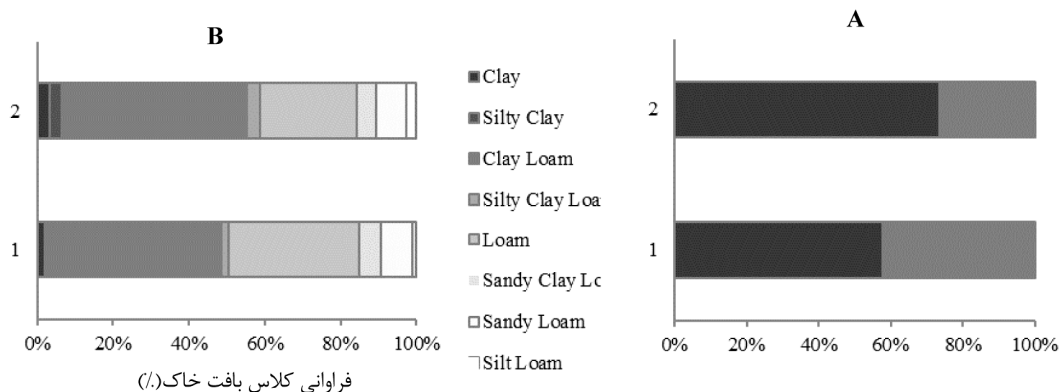
منطقه مطالعاتی بخشی از حوزه آبخیز سفیدرود و زیرحوضه زنجان‌رود است که دارای وسعت ۲۵۳ کیلومتر مربع با اقلیم مدیترانه‌ای و بیشینه و کمینه دمای سالانه ۲۵ و ۳۰- درجه سلسیوس است. توزیع بارش در منطقه یکنواخت و متوسط سالانه آن ۳۲۷ میلی‌متر است. رژیم‌های رطوبتی و حرارتی منطقه شامل زیریک و مزیک است. تیپ غالب اراضی منطقه شامل فلات و تراس‌های فوقانی (۱۳٪)، تپه‌های کم‌ارتفاع (۳۰٪) و دشت‌های دامنه‌ای (۵۷٪) و کاربری فعلی اراضی شامل کشاورزی آبی (۳/۲٪)، دیم (۷۱٪)، مرتع (۲۵٪) و اراضی متفرقه (۸٪) است. از نظر بافت خاک، خاک غالب منطقه لوم رسی و لوم و خاک‌ها در زیرگروه بزرگ Typic Calcixerepts قرار دارد (شکل ۱).

تعداد ۲۰۰ نمونه خاک در مقیاس حوضه آبخیز (داده‌های واسنجی) و ۴۰ نمونه در مقیاس مزرعه (داده‌های اعتبارسنجی) از دو عمق خاک (صفر تا ۱۰ و ۱۰ تا ۳۰ سانتی‌متر) به صورت نمونه‌برداری مرکب در یک شعاع ۱۰ متری جمع‌آوری (شکل ۲) و پس از هواخشک کردن از الک ۲ میلی‌متر عبور داده شد. برخی ویژگی‌های فیزیکوشیمیایی خاک‌ها شامل توزیع اندازه ذرات خاک (بافت خاک) به روش هیدرومتری (۱۸)، جرم ویژه ظاهری به روش کلوخه (۲۷)، جرم ویژه حقیقی به روش پیکنومتر (۱۶)، ضریب هدایت الکتریکی (EC) و واکنش خاک (pH) به روش عصاره اشباع (۲۱) و مقدار کربن آلی به روش اکسیداسیون تر در آزمایشگاه اندازه‌گیری شد (۳۹) (جدول ۱).

همچون محاسبه شاخص‌های طیفی و روش‌های رگرسیونی، برآورد مقدار ماده آلی خاک با مقادیر R^2 بین ۰/۴۶ تا ۰/۸۹ گزارش شده است (۳۶). از آنجا که بازتاب طیفی خاک در اثر برهم‌کنش پیچیده هر یک از اجزای خاک حاصل می‌شود، لازم است به‌ازای طول موجی مشخص و با استفاده از روش‌های مناسب آماری مقدار کربن آلی خاک را تعیین نمود.

تاکنون روش‌های متعددی برای آنالیز داده‌های طیفی خاک و ارتباط آن با ویژگی‌های خاک ارائه شده است. این روش‌ها عمدتاً شامل روش‌های برداشت پیوسته^۱ (CR)، رگرسیون مؤلفه‌های اصلی^۲ (PCR)، رگرسیون خطی چندگانه^۳ (MLR) و حداقل مربعات جزئی^۴ (PLS) است. روش PLS یکی از معمول‌ترین روش‌های آماری چندمتغیره است که از داده‌های طیفی در تمامی طول‌موج‌ها استفاده نموده و با گنجاندن آن‌ها در تعداد کم‌تری متغیر (فاکتورهای تخمینگر) و ایجاد یک مدل خطی، ویژگی خاک موردنظر را محاسبه می‌کند. نوع روش‌های مورد استفاده در پیش‌پردازش داده‌های طیفی و ترکیب مناسب آن‌ها، اثر قابل‌توجهی بر دقت نتایج مدل خواهد داشت. در برخی مطالعات، روش PLSR در برآورد برخی ویژگی‌های خاک دقت قابل‌قبولی نشان داده است (۴۰، ۱۴). پژوهش‌ها نشان داده تلفیق روش رگرسیون حداقل مربعات معمولی با بازنمونه‌گیری (bagging-PLSR) می‌تواند منجر به افزایش دقت مقادیر تخمینی شود (۶). علاوه بر این، از جمله مزایای روش bagging-PLSR می‌توان به بهبود نتایج مدل‌های بیش برآورد و ارائه عدم اطمینان مقادیر تخمینی، اشاره نمود (۳۸). بر همین اساس هدف از انجام این پژوهش؛ (۱) مطالعه رفتار طیفی خاک

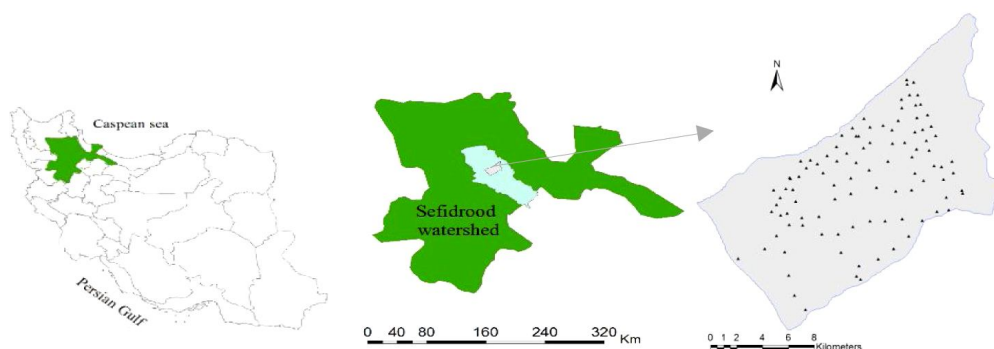
- 1- Continuum removal
- 2- Principal component regression
- 3- Multiple linear regression
- 4- Partial least square



Frequency of Soil texture classes (%)

شکل ۱- توزیع فراوانی کلاس‌های بافت خاک در عمق صفر تا ۱۰ (لایه ۱) و ۱۰-۳۰ (لایه ۲) سانتی‌متر در دو مقیاس مزرعه (الف) و حوزه آبخیز (ب).

Figure 1. Frequency of Soil texture classes in depth 0-10 cm (layer 1) and 10-30 cm (layer 2) at field scale (A) and watershed scale (B).



شکل ۲- نقشه موقعیت نقاط نمونه‌برداری خاک‌های منطقه.

Figure 2. Location map of the soils sampling points.

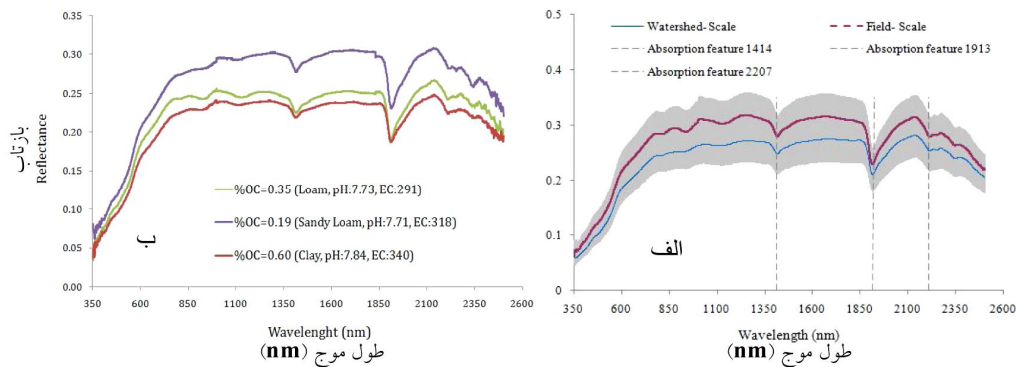
جدول ۱- آمار توصیفی برخی از ویژگی‌های مبنایی خاک مورد استفاده در فرآیند واسنجی و اعتبارسنجی مدل.

Table 1. Descriptive statistics of the basic characteristics of the soil used in the model calibration and validation.

داده‌های اعتبارسنجی مدل model validation Data			داده‌های واسنجی مدل model calibration Data			ویژگی خاک soil characteristics
حداقل Min	حداکثر Max	انحراف معیار SD	حداقل Min	حداکثر Max	انحراف معیار SD	
16	48	0.07	13	63	0.1	2000-50 μm
27	50	0.05	15	45	0.06	$<2 \mu\text{m}$
1.3	1.6	0.12	1.2	1.7	0.1	$\rho_b (\text{g cm}^{-3})$
2.6	2.8	0.05	2.6	2.9	0.04	$\rho_p (\text{g cm}^{-3})$
7.2	7.6	0.07	7.0	7.9	0.19	pH
0.44	2.1	0.33	0.24	7.3	0.52	EC [dS m^{-1}]
0.05	1.8	0.23	0.06	1.9	0.26	OC (%)

نانومتر) در تاریکخانه اندازه‌گیری شد. اندازه‌گیری‌های طیفی به کمک پیستول با گستره دید ۸ درجه (فاصله پیستول تا نمونه برابر ۲۸ سانتی‌متر، قطر دید ۴ سانتی‌متر) و منبع نور ۱۵۰ وات (با زاویه ۲۵ درجه از محور عمود بر نمونه) انجام شد.

به منظور اندازه‌گیری بازتاب طیفی نمونه‌های خاک، از دستگاه اسپکترورادایومتر زمینی^۱ استفاده شد. نمونه‌های هواخشک شده خاک از الک ۲ میلی‌متر عبور داده شد و منحنی‌های طیفی آن‌ها در دامنه مرئی، مادون قرمز نزدیک و مادون قرمز میانی (۲۵۰۰-۳۵۰۰

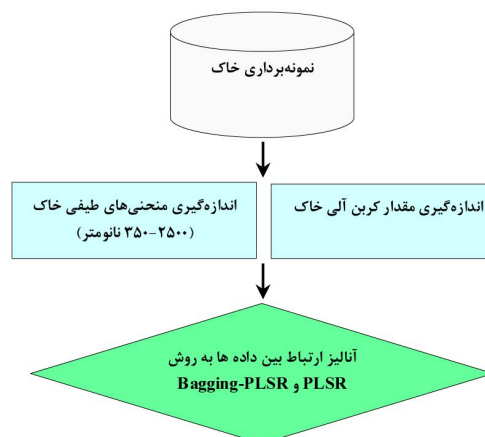


شکل ۳- میانگین و انحراف معیار منحنی‌های بازتاب طیفی خاک در دو مقیاس حوزه آبخیز و مزرعه (الف)، منحنی‌های طیفی سه خاک با مقادیر مختلف کربن آلی، بافت، pH و EC (ب). خطوط قائم باندهای جذبی به دلیل وجود آب هیگروسکوپی و آب موجود در شبکه رس‌ها در طول موج‌های ۱۴۱۴، ۱۹۱۳ و ۲۲۰۷ نانومتر را نشان می‌دهد.

Figure 3. The mean and SD of spectral reflectance curves of soil in the watershed and field scale (a), Spectral curves of soils with different amounts of organic carbon, texture, pH and EC (b).

آبخیز و ۱۰۰۰۰ منحنی طیفی در مقیاس مزرعه ثبت و اقدام به ایجاد یک کتابخانه طیفی از خاک‌های منطقه شد (شکل ۲). شکل ۴، روندنمای (فلوچارت) کلی روش انجام پژوهش را نشان می‌دهد.

به منظور کاهش خطا (افزایش نسبت سیگنال به نویز) حاصل از شرایط محیطی (رطوبت، دما، غبار) در حین طیف‌سنجی، از هر نمونه خاک ۲۵ اسکن تهیه و مجموعاً ۵۰۰۰ منحنی طیفی در مقیاس حوزه



شکل ۴- روندنمای کلی روش انجام تحقیق.

Figure 4. The general flowchart of research method.

1- FieldSpec®3, ASD, FR, USA

که در آن، t بردار رتبه‌ها^۶ (رابطه ۲)، p بارگذاری‌های طیفی^۷، q بارگذاری‌های متغیرهای خاک، E و f به ترتیب خطای باقی‌مانده^۸ متغیرهای تخمینگر و پاسخ حاصل از اثر تخمینی فاکتور a ام ($a=1,2,\dots,A$) است.

$$t_a = X_{a-1} \cdot w_a \quad (2)$$

در رابطه ۲، w_a ضریب وزنی مقیاس‌سازی شده ($w_a = cX'_{a-1}y_{a-1}$) و c فاکتور مقیاس است. ضرایب مدل رگرسیون (b و b_0) به کمک روابط ۳ و ۴ محاسبه و در مدل تخمینگر (رابطه ۵) بر اساس تعداد فاکتورهای مناسب اتخاذ شده، برای برآورد متغیر مورد نظر (\hat{y}_i) مورد استفاده قرار می‌گیرد. در نهایت با استفاده از رابطه ۶، مقادیر تخمینی متغیر مورد نظر ($\hat{y}_{bag}(y)$) محاسبه می‌شود. در رابطه زیر x_i باند طیفی است.

$$b = W(P'W)^{-1} \cdot q \quad (3)$$

$$b_0 = \bar{y} - \bar{x}'b \quad (4)$$

$$\hat{y}_i = b_0 + x'_i \cdot b \quad (5)$$

روش رگرسیون حداقل مربعات جزئی توأم با بازنمونه‌گیری (bagging-PLSR): روش بازنمونه‌گیری مورد استفاده در این روش، روش نمونه‌گیر bootstrap می‌باشد. در این روش به‌طور تصادفی اقدام به نمونه‌برداری از داده‌ها شده و سپس جایگزینی مجدد داده‌ها از طریق همان داده

در ادامه با استفاده از برنامه‌های RS3 و IDL-ENVI میانگین منحنی‌های طیفی هر یک از نمونه‌های خاک محاسبه و در تجزیه و تحلیل‌ها مورد استفاده قرار گرفت. به‌منظور واسنجی میزان بازتابش ثبت‌شده از نمونه‌های خاک، از یک صفحه سفید مینا به ابعاد ۰/۱۵×۰/۱۵ متر مربع و در هر ۳ دقیقه و منطبق با شرایط نمونه‌های خاک استفاده شد.

به‌دلیل وجود پراکندگی^۱ زیاد در طول‌موج‌های ۳۵۰-۴۰۰ و ۲۵۰۰-۲۶۰۰، این محدوده از تمامی طیف‌ها حذف و در مدل‌سازی مورد استفاده قرار نگرفت. داده‌های طیفی معمولاً متأثر از اثرهای پخش نور در محیط، گردوغبار و بخار آب هوا قرار گرفته و با پراکندگی همراه است. به‌منظور حذف این اثرات که منجر به رفتار غیرخطی باندها، جابجایی طیفی و کاهش نسبت سیگنال به پراکندگی می‌شود، از الگوریتم‌های مختلف پیش‌پردازش موجود در برنامه 3.1 ParLeS استفاده شد. الگوریتم‌های پیش‌پردازش مورد مطالعه شامل روش‌های تصحیح (نرمال‌سازی به روش‌های MSC، SNV، حذف روند به روش موجک) و صاف و یکنواخت‌سازی^۲ (فیلترهای میان‌ه، موجک و savitzky-golay) طیف‌ها بود (۳۶).

روش رگرسیون حداقل مربعات جزئی^۳ (PLSR): در این روش، متغیر پاسخ^۴ (y) به‌واسطه متغیرهای پنهان به ماتریس متغیرهای تخمینگر^۵ ارتباط داده می‌شود:

$$\begin{aligned} X &= t_1 p'_1 + \dots + t_a p'_a + E_a \\ y &= t_1 q_1 + \dots + t_a q_a + f_a \end{aligned} \quad (1)$$

- 1- Noise
- 2- Smoothing
- 3- Partial Least Square Regression
- 4- Response
- 5- Predictor

- 6- Scores
- 7- Spectral Loadings
- 8- Residual

$$SD = \sqrt{\frac{\sum_{i=1}^N (y_i - \bar{y})^2}{N-1}} \quad (10)$$

$$MBE = \frac{1}{N} \sum_{i=1}^N (y'_i - y_i) \quad (11)$$

که در آنها، y'_i مقادیر تخمینی، y_i مقادیر مشاهداتی، \bar{y} میانگین مشاهدات، N تعداد داده‌ها، m تعداد فاکتورهای مدل و SD انحراف معیار داده‌ها است.

انتخاب تعداد فاکتورهای مناسب (متغیرهای تخمینگر) در مدل PLSR به کمک آماره RMSE (۳۶) انجام شد. آماره RPD یکی از متداول‌ترین شاخص‌های ارزیابی دقت در مطالعات طیف‌سنجی است که از نسبت خطای استاندارد داده‌ها به RMSE مدل تخمینگر به دست می‌آید. بر اساس مطالعات انجام‌شده، مقادیر $RPD > 2$ بیانگر دقت بسیار بالا، $1/4 < RPD < 2$ بیانگر دقت خوب تا متوسط و مقادیر $RPD < 1/4$ بیانگر دقت ضعیف فرآیند تخمین است (۲۶، ۱۰). شکل ۵، روندنمای نحوه پیش‌پردازش داده‌های ابرطیفی در برآورد مقدار کربن آلی خاک را به شکل شماتیک نشان می‌دهد.

نمونه‌برداری شده صورت می‌گیرد. اگر سری داده‌های واسنجی به صورت $Z_N = (z_1, z_2, \dots, z_n)$ در نظر گرفته شود که در آن $z_n = (x_n, y_n)$ و N تعداد داده‌ها باشد و نمونه‌برداری از این سری B مرتبه تکرار شود ($b=1, 2, \dots, B$)، تعداد B سری داده راه‌انداز^۱ شده (Z_N^{*b}) تشکیل خواهد شد. هر یک از سری‌های حاصل حدود ۶۳٪ داده‌ها را پوشش داده و منجر به ایجاد سری‌هایی هم‌تا و منطبق با تعداد داده‌های اولیه می‌شود. سپس مناسب‌ترین مدل PLSR به هر یک از این سری‌های جدید برازش داده شده و مقادیر پیش‌بینی ($\hat{f}^{*b}(y)$) به دست می‌آید. در نهایت با استفاده از رابطه ۶، مقادیر تخمینی متغیر مورد نظر محاسبه می‌شود.

$$\hat{f}_{bag}(y) = \frac{1}{B} \sum_{b=1}^B \hat{f}^{*b}(y) \quad (6)$$

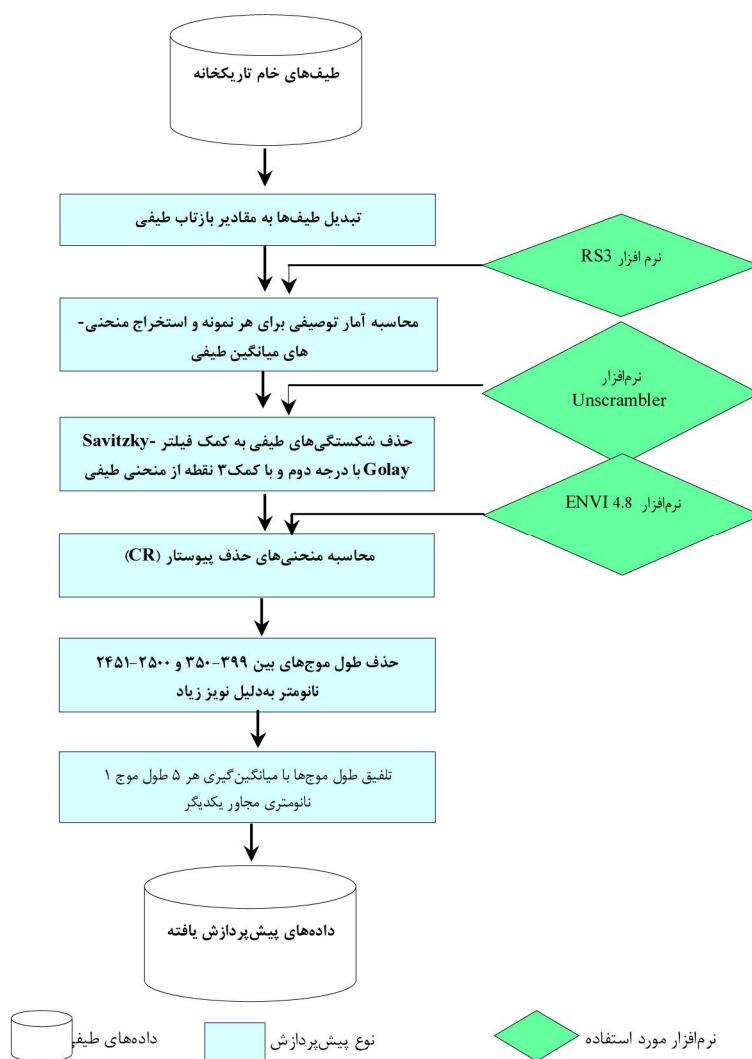
دقت آزمایی: به منظور ارزیابی دقت در هر یک از مراحل فرآیند تخمین، از آماره‌های مختلفی استفاده شد. این آماره‌ها شامل ضریب تبیین (R^2)، جذر میانگین مربعات خطا^۲ (RMSE)، نسبت انحراف دقت^۳ (RPD) و میانگین انحراف خطا^۴ (MBE) است.

$$R^2 = 1 - \frac{\sum_{i=1}^N (y'_i - y_i)^2}{\sum_{i=1}^N (y'_i - \bar{y})^2} \quad (7)$$

$$RMSE = \sqrt{\frac{1}{N} \sum_{i=1}^N (y_i - y'_i)^2} \quad (8)$$

$$RPD = \frac{SD}{RMSE} \times 100\% \quad (9)$$

- 1- Bootstrap
- 2- Root Mean Square Error Cross Validation
- 3- Ratio Performance Deviation
- 4- Mean Bias Error



شکل ۵- روندنمای نحوه پیش‌پردازش داده‌های ابرطیفی در برآورد مقدار کربن آلی.

Figure 5. The flowchart of pre-processing hyperspectral data to estimate the amount of soil organic carbon.

رس، دامنه جرم ویژه ظاهری و حقیقی خاک نیز وسیع و میانگین آن‌ها برابر $1/45$ و $2/75$ گرم بر سانتی مترمکعب به دست آمد. این تغییرات عمدتاً به دلیل وجود تنوع در تیپ اراضی و شیب منطقه است. گستره مقدار کربن آلی خاک در مقیاس مزرعه برابر $(\pm 0/26) 1/8 - 0/05$ درصد و در مقیاس حوضه برابر $(\pm 0/23) 1/9 - 0/06$ درصد اندازه‌گیری شد. بنابراین گستره مناسبی برای مدل‌سازی وجود دارد. مقادیر کربن آلی خاک از نظر وضعیت نرمال بودن به

نتایج و بحث

ویژگی‌های فیزیکی و بازتاب طیفی خاک: توزیع کلاس‌های بافت خاک در دو مقیاس مزرعه و حوضه آبخیز در شکل ۱ نشان داده شده است. تعداد هشت کلاس بافت در مقیاس حوضه و دو کلاس در مقیاس مزرعه شناسایی شد. در جدول ۱ نتایج آمار توصیفی برخی ویژگی‌های فیزیکوشیمیایی خاک به همراه مقدار کربن آلی خاک ارائه شده است. بر اساس جدول ۱، علاوه بر تغییرات زیاد درصد شن و

روش کلموگروف- اسمیرنف در سطح پنج درصد مورد آزمون قرار گرفت که نتایج نشان داد توزیع داده‌ها نرمال است. مقدار کربن آلی خاک در مقیاس حوزه آبخیز به منظور آزمون مدل (داده‌های واسنجی) و مقادیر در مقیاس مزرعه به منظور دقت‌آزمایی مدل (داده‌های اعتبارسنجی) مورد استفاده قرار گرفت.

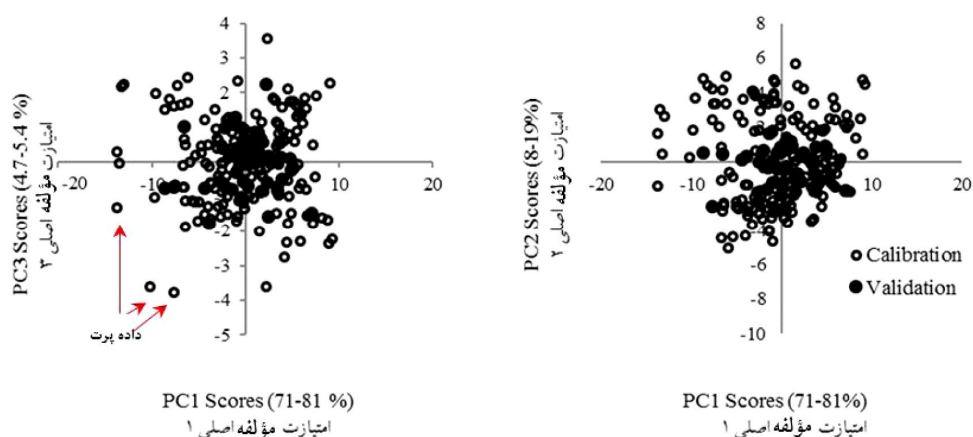
در شکل ۳ میانگین و گستره تغییرات منحنی‌های بازتاب طیفی خاک در دامنه مرئی (VIS)، مادون قرمز نزدیک (NIR) و مادون قرمز میانی (SWIR) در دو مقیاس مزرعه و حوضه ارائه شده است. منحنی میانگین طیفی در مقیاس مزرعه بالاتر از منحنی متوسط طیفی در مقیاس حوزه آبخیز قرار گرفته است. این می‌تواند به دلیل میانگین پایین اسیدیته خاک، ماده آلی و هدایت الکتریکی داده‌های مزرعه نسبت به حوزه آبخیز باشد که باعث افزایش بازتاب در دامنه NIR و SWIR شده است (۲۷). این نتایج نشان می‌دهد ویژگی‌های فیزیکی و شیمیایی خاک اثر متقابل پیچیده‌ای بر رفتار طیفی خاک دارند. به‌طور کلی با افزایش مقدار ماده آلی خاک، بازتاب طیفی کاهش می‌یابد. پژوهش‌ها نشان داده هر یک از اجزای ماده آلی خاک نیز اثر بازتاب متفاوتی در هر یک از طول موج‌ها دارد (۳۱). بر اساس شکل ۳، با افزایش طول موج از دامنه مرئی، مقدار بازتاب طیفی خاک به تدریج افزایش یافته و حداکثر به مرز ۳۵ درصد می‌رسد. از طرفی، در طول موج‌های بین ۱۳۷۰ تا ۱۵۰۰، ۱۸۶۰ تا ۲۰۵۰ و ۲۲۰۰ تا ۲۲۳۰ نانومتر مقادیر بازتاب طیفی خاک‌ها به شدت کاهش می‌یابد. این طول موج‌ها، طول موج‌های جذبی نام داشته و مربوط به جذب آب بر روی سطوح رس‌ها (طول موج ۱۴۱۴)، پیوندهای H-O-H ملکول‌های آب حبس شده در شبکه رس‌ها (طول موج ۱۹۱۳) و پیوندهای

OH با فلزات آلومینیم، آهن و سیلیس (طول موج ۲۲۰۷) است (۱۳). از طرفی هرچه مقدار رس خاک بیش‌تر باشد، عمق باندهای جذبی (فرورفتگی‌ها) بیش‌تر خواهد بود. زیرا خاک‌های رسی مقدار رطوبت بیش‌تری را در خود نگه می‌دارند. وجود چنین باندهای جذبی توسط دیگر پژوهشگران نیز گزارش شده است (۴، ۲۷). مطالعات نشان داده طول موج‌های ۱۴۰۰، ۱۹۰۰ و ۲۲۰۰ نانومتر باندهایی هستند که به بسیاری از ویژگی‌های خاک از جمله مقدار کربن آلی خاک حساس است (۲) (شکل ۳). بنابراین در فرآیند تخمین ماده آلی خاک طول موج‌های ذکر شده نقش عمده را ایفا می‌کنند.

تجزیه مؤلفه‌های اصلی^۱ (PCA): پس از انجام پیش‌پردازش‌های طیفی، تجزیه و تحلیل مؤلفه‌های اصلی بر روی داده‌های طیفی انجام شد. در شکل ۶ نمودار پراکنش دوبعدی داده‌های واسنجی و اعتبارسنجی در قالب سه مؤلفه اصلی (PC) شامل PC1، PC2 و PC3 نشان داده شده است.

مشاهده می‌شود که تغییرات واریانس بازتاب طیفی خاک در دو مقیاس حوزه آبخیز و مزرعه تقریباً مشابه و نزدیک به یکدیگر است. همچنین بیش از ۹۴ درصد تغییرات واریانس طیفی داده‌ها می‌تواند در قالب PC1، PC2 و PC3 ارائه شود. این نشان می‌دهد باندهای طیفی فاقد همبستگی خطی نسبت به یکدیگر هستند. داده‌های طیفی که بیش از دو برابر انحراف معیار از میانگین طیفی فاصله داشتند به‌عنوان داده پرت شناسایی و از مجموعه داده‌های واسنجی حذف شدند (۸، ۱۷).

1- Principal component analysis



شکل ۶- نمودار پراکنش دوبعدی داده‌های طیفی مربوط به واسنجی و اعتبارسنجی در قالب سه مؤلفه اصلی PC1، PC2 و PC3 و موقعیت داده‌های پرت.

Figure 6. Two-dimensional spectral plot of the calibration and validation of the three main components of PC1, PC2 and PC3 and the position of outliers.

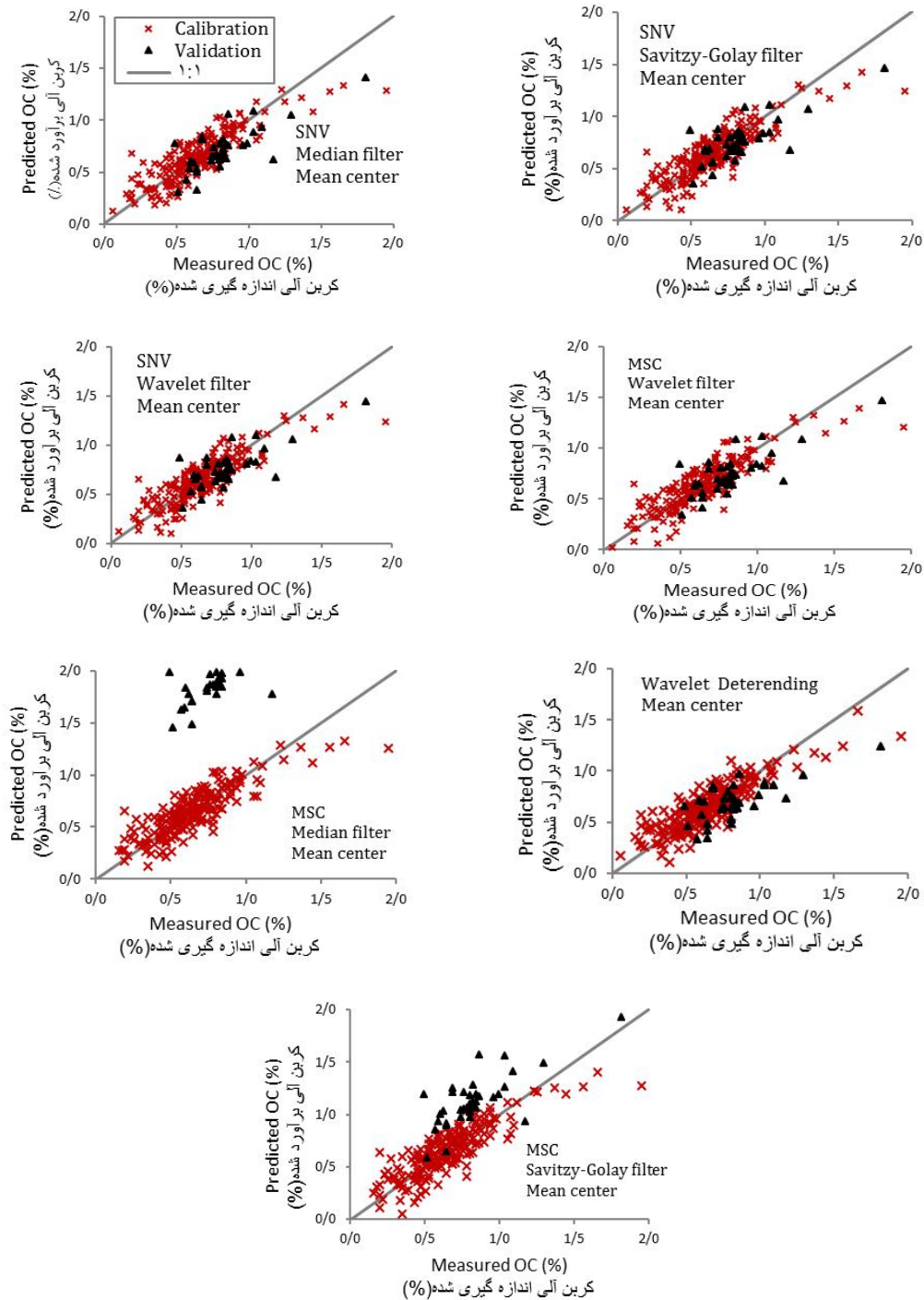
توأم با فیلترهای Savitzky-Golay و موجک (به‌منظور افزایش نسبت سیگنال به نویز)، پایین‌ترین دقت برای آزمون (RMSE=۰/۱۴۱٪، RPD=۱/۸۸) و اعتبارسنجی (RMSE=۰/۳۴۸٪، RPD=۰/۶۸) مدل PLSR به‌دست آمد. در مدل bagging-PLSR، روش SNV توأم با فیلتر موجک نتایج دقیق‌تر (RMSE=۰/۱۴۷٪، $R^2=۰/۷۰$) و روش SNV توأم با فیلتر Savitzky-Golay با عدم اطمینان بیشتری (RMSE=۰/۱۵۷٪، $R^2=۰/۶۸$) همراه بود (جدول ۲). بر اساس شکل ۷، روش‌های مختلف پیش‌پردازش اثر متفاوتی بر دقت مقادیر آزمون و اعتبارسنجی مدل PLSR نشان داده‌اند. به‌طورکلی، روش‌های حذف روند به روش موجک، SNV و MSC (به‌منظور کاهش اثرات پخش نور و حذف جابجایی‌های طیفی) و فیلترهای میانه، موجک و Savitzky-Golay (به‌منظور کاهش پراکندگی‌های طیفی)، به‌ترتیب بیش‌ترین کارایی را نشان دادند (جدول ۲). پژوهش‌ها نشان داده، استفاده از روش SNV توأم با فیلتر میانه نتایج دقیق‌تری در مقایسه با سایر روش‌های پیش‌پردازش ارائه می‌دهد (۲۳). در روش SNV حذف

اثر رویکردهای مختلف پیش‌پردازش بر دقت مدل: از آن‌جا که در پژوهش‌ها مناسب‌ترین روش پیش‌پردازش داده‌های طیفی به‌منظور برآورد مقدار کربن آلی خاک گزارش نشده است، پیش از واسنجی مدل اثر روش‌های مختلف پیش‌پردازش و تلفیق آن‌ها بر روی نتایج آزمون مدل مورد بررسی قرار گرفت (شکل ۷). با انجام هر یک از روش‌های مختلف پیش‌پردازش، مشخصه‌های جذب و انعکاس متفاوتی بر روی منحنی‌های طیفی حاصل شد. نتایج نشان داد انتخاب روش‌های مختلف پیش‌پردازش و انتخاب ترکیبی آن‌ها منجر به تغییرات قابل‌توجه در دقت مدل و اعتبارسنجی نتایج می‌شود. در روش PLSR، بهترین روش پیش‌پردازش به‌منظور حذف اثرات پخش نور محیط^۱ بر طیف‌ها و جابجایی‌های طیفی، روش حذف روند به روش موجک^۲ بود. در این روش مقدار RMSE و RPD حاصل از آزمون مدل به‌ترتیب برابر ۰/۱۳۶٪ و ۱/۹۴٪ و اعتبارسنجی مدل برابر ۰/۲۰۹٪ و ۱/۱۲٪ به‌دست آمد. در روش MSC^۳

- 1- Light scatter and baseline correction
- 2- Wavelet detrending
- 3- Multiplicative signal correction

در حالی که روش MSC برای یک طیف که میانگینی از تمامی طیف‌ها است، مورد استفاده قرار می‌گیرد.

اثرات پخش نور و ذرات ریز موجود در محیط در مورد هر یک از طیف‌ها انجام می‌شود و سپس هر طیف به مقداری با میانگین صفر و واریانس یک نرمال می‌شود.



شکل ۷- نمودار اثر روش‌های مختلف پیش‌پردازش داده‌های بازتاب طیفی بر دقت نتایج آزمون و اعتبارسنجی روش PLSR.

Figure 7. The effect of different preprocessing methods of spectral reflectance data on the accuracy of the test results and validation of PLSR method.

RMSE بین ۰/۱۶۷ تا ۱/۱۱ درصد و RPD بین ۰/۲۱ تا ۱/۴۱ در حال تغییر است. همچنین بر اساس نتایج به دست آمده از داده‌های اعتبارسنجی، روش PLSR به‌ازای تمامی روش‌های پیش‌پردازش به‌جز روش MSC توأم با فیلترهای میان‌ه و Savitzky-Golay، حالت کم‌برآورد (۰/۱۳۲٪- تا ۰/۶۵=MBE) را نشان می‌دهد (جدول ۲). بنابراین روش PLSR مقدار ماده آلی خاک را در دامنه‌های پایین با دقت بیش‌تری برآورد می‌کند. پژوهش‌های به‌عمل آمده نشان داده روش MSC نسبت به روش SNV از کارایی کم‌تری برخوردار است (۱۵). گستره مقادیر شاخص RPD به‌ازای داده‌های واسنجی در روش PLSR بین ۱/۸۸-۱/۹۴ و به‌ازای داده‌های اعتبارسنجی بین ۱/۴۱-۰/۲۱ محاسبه شد. بنابراین بر اساس مقدار این شاخص می‌توان گفت مدل تخمینگر از درجه دقت خوب تا متوسط برخوردار است. در پژوهشی کامبول و همکاران (۲۰۱۲) نشان دادند روش PLSR قادر است مقدار کربن آلی خاک را با دقت مناسب (RMSE=۰/۳۲٪) برآورد کند (۹). همچنین چانگ و لایرد (۲۰۰۱) مقدار کربن آلی خاک را به کمک روش PLSR و با RMSE برابر ۰/۶۵ درصد برآورد نمودند (۱۰). بنابراین، در پژوهش حاضر مقدار کربن آلی خاک با دقت بسیار بالاتری نسبت به مطالعات پیشین برآورد شده است.

بر اساس مطالعات، هرچه تعداد فاکتورها و مقدار RMSE آزمون مدل کم‌تر باشد، آن مدل PLSR از مطلوبیت بالاتری برخوردار است (۳۶). بر اساس نوع روش پیش‌پردازش، مناسب‌ترین مدل PLSR برای برآورد مقدار کربن آلی خاک به‌ترتیب دارای ۱۷، ۱۹، ۱۶ و ۱۵ فاکتور است. بنابراین افزایش تعداد فاکتورهای تخمینگر مدل لزوماً منجر به افزایش دقت تخمین نمی‌شود (۳۲)، بلکه نوع و ترکیب مناسب روش‌های پیش‌پردازش طیفی اهمیت بیش‌تری در افزایش دقت مدل PLSR دارد.

در برخی پژوهش‌های انجام شده، مقدار RMSE تخمین کربن آلی به روش PLSR برابر ۰/۳۶ درصد به‌دست آمده (۳۵) که در مقایسه با نتایج پژوهش حاضر با خطای بیش‌تری همراه است. همچنین در برخی مطالعات، تنها گستره‌های ۵۰۰ تا ۹۰۰ و ۱۳۵۰ تا ۱۴۶۰ نانومتر مورد استفاده قرار گرفته است. نتایج این مطالعات نشان داده مدل PLSR در برآورد مقدار کربن آلی خاک به‌ازای ۱۲ فاکتور (RMSE=۳/۳۷، $R^2=0/90$) نسبت به هشت فاکتور (RMSE=۵/۱۲، $R^2=0/76$) دقت بالاتری دارد (۳۴).

بر اساس شکل ۷، به‌ازای مقدار کربن آلی کم‌تر از ۱/۲ درصد، داده‌های آزمون و اعتبارسنجی از توزیع نسبتاً مطلوبی در اطراف خط ۱:۱ برخوردار هستند و با عبور مقدار کربن آلی خاک از این مرز، حالت کم‌برآوردی برای مدل PLSR ایجاد می‌شود. نتایج اعتبارسنجی روش PLSR نشان داد که پراکنش نقاط در اطراف خط ۱:۱ نیز بسته به نوع روش پیش‌پردازش متفاوت است، به‌طوری‌که مقادیر

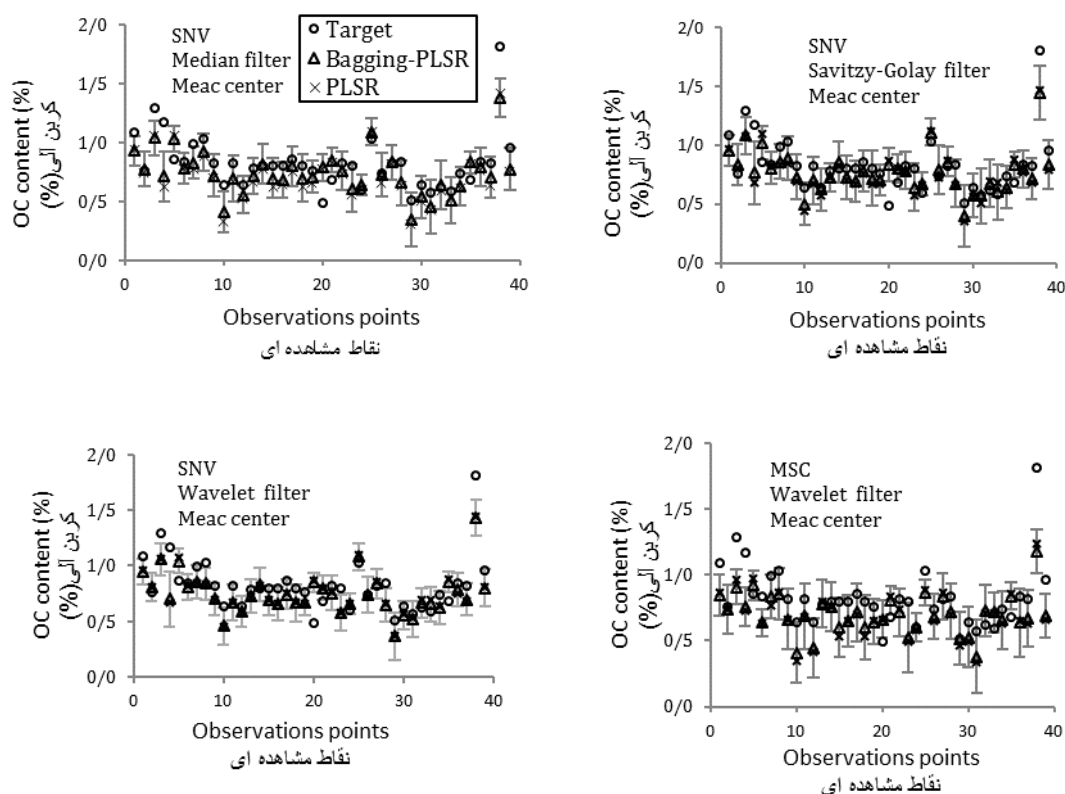
جدول ۲- نتایج ارزیابی روش‌های مختلف پیش‌پردازش بر دقت روش‌های PLSR و bagging-PLSR در برآورد مقدار کربن آلی خاک.

Table 2. Results of the preprocessing accuracy on bagging-PLSR and PLSR method to estimate the amount of soil organic carbon.

RPD	SDE	MBE (%)	RMSE (%)	R ²	تعداد فاکتور No. F	انحراف استاندارد SD	میانگین Mean (%)	تعداد N	روش‌های مختلف پیش‌پردازش Preprocessing Methods
SNV, Median filter, Mean center									
1.92	0.138		0.138	0.73	19	0.26	0.64	200	PLSR _{Cal.}
1.27	0.151	-0.109	0.185	0.60		0.26	0.82	40	PLSR _{Val.}
	0.151		0.151	0.70	19	0.26	0.64	200	Bagging-PLSR _{Cal.}
1.45	0.144	-0.079	0.163	0.63		0.19	0.74	40	Bagging-PLSR _{Val.}
SNV, Savitzky-Golay filter, Mean center									
1.90	0.140		0.140	0.72	16	0.26	0.64	200	PLSR _{Cal.}
1.41	0.156	-0.065	0.167	0.57		0.2	0.76	40	PLSR _{Val.}
	0.156		0.157	0.68	16	0.26	0.64	200	Bagging-PLSR _{Cal.}
1.66	0.135	-0.047	0.142	0.67		0.18	0.78	40	Bagging-PLSR _{Val.}
SNV, Wavelet filter, Mean center									
1.90	0.140		0.139	0.72	16	0.26	0.64	200	PLSR _{Cal.}
1.40	0.156	-0.068	0.168	0.56		0.19	0.76	40	PLSR _{Val.}
	0.147		0.147	0.70	16	0.26	0.64	200	Bagging-PLSR _{Cal.}
1.44	0.149	-0.071	0.164	0.60		0.19	0.75	40	Bagging-PLSR _{Val.}
MSC, Wavelet filter, Mean center									
1.88	0.141		0.141	0.72	15	0.26	0.64	200	PLSR _{Cal.}
1.36	0.153	-0.083	0.172	0.58		0.20	0.74	40	PLSR _{Val.}
	0.153		0.154	0.69	15	0.26	0.64	200	Bagging-PLSR _{Cal.}
1.58	0.141	-0.055	0.149	0.64		0.20	0.77	40	Bagging-PLSR _{Val.}
Wavelet Deterending									
1.94	0.137		0.136	0.73	17	0.26	0.64	200	PLSR _{Cal.}
1.12	0.165	-0.132	0.209	0.51		0.18	0.69	40	PLSR _{Val.}
	0.153		0.155	0.70	17	0.26	0.64	200	Bagging-PLSR _{Cal.}
1.17	0.161	-0.121	0.200	0.54		0.15	0.70	40	Bagging-PLSR _{Val.}
MSC, Median filter, Mean center									
1.90	0.140		0.139	0.72	17	0.26	0.64	200	PLSR _{Cal.}
0.21	0.164	1.1	1.11	0.57		0.23	1.93	40	PLSR _{Val.}
	0.147		0.147	0.70	17	0.26	0.64	200	Bagging-PLSR _{Cal.}
0.23	0.142	1.02	1.03	0.64		0.21	1.85	40	Bagging-PLSR _{Val.}
MSC, Savitzky-Golay filter, Mean center									
1.88	0.141		0.141	0.72	16	0.26	0.64	200	PLSR _{Cal.}
0.68	0.173	0.302	0.348	0.54		0.24	1.13	40	PLSR _{Val.}
	0.148		0.148	0.7	16	0.26	0.64	200	Bagging-PLSR _{Cal.}
1.07	-0.151	0.162	0.220	0.61		0.22	0.99	40	Bagging-PLSR _{Val.}

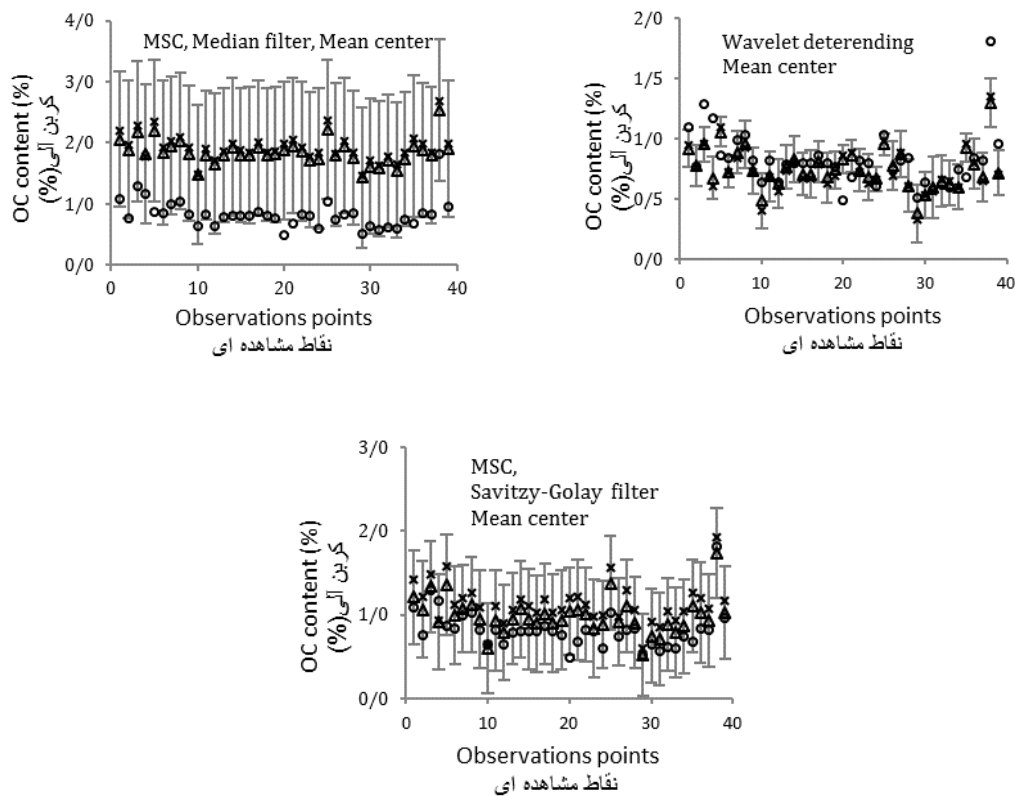
در شکل ۸ نتایج دقت آزمایشی روش‌های PLSR و bagging-PLSR در برآورد مقدار کربن آلی خاک با استفاده از داده‌های اعتبارسنجی در مقیاس مزرعه ارائه شده است. بر اساس این نتایج، روش bagging-PLSR دقت بالاتری نسبت به روش PLSR در تمامی روش‌های پیش‌پردازش در برآورد کربن آلی نشان داد (جدول ۲). روش bagging-PLSR نیز مشابه روش PLSR، به‌ازای تمامی روش‌های پیش‌پردازش حالت کم‌برآورد (مقدار MBE بین تا $-0/047$ تا $-0/121$) و به‌ازای روش MSC توأم با فیلترهای میان‌ه و Savitzky-Golay، حالت بیش‌برآورد ($MBE=0/162-1/02$ ٪) را نشان داد.

مطالعات نشان داده استفاده از روش bagging-PLSR در بیش‌تر موارد منجر به افزایش دقت مقادیر تخمینی شده و از بیش‌برآوردی مدل جلوگیری می‌کند (۶، ۳۷). بر اساس جدول ۲، تعداد بهینه فاکتورهای تخمینگر روش bagging-PLSR نیز مشابه با روش PLSR به‌دست آمده است. در این روش مقدار RMSE مرحله آزمون و اعتبارسنجی مدل به‌ازای روش‌های مختلف پیش‌پردازش به‌ترتیب بین $0/146$ تا $0/157$ و $0/142$ تا $1/03$ درصد به‌دست آمد. گستره مقادیر شاخص RPD در روش bagging-PLSR به‌ازای داده‌های آزمون بین $0/23$ تا $1/66$ محاسبه شد (جدول ۲).



شکل ۸- نتایج اعتبارسنجی روش‌های PLSR و bagging-PLSR به‌ازای روش‌های مختلف پیش‌پردازش (Target: مقادیر مطلوب، خطوط عمودی ۹۵٪ فاصله اطمینان مقادیر تخمینی روش bagging-PLSR را نشان می‌دهد).

Figure 8. The results of bagging-PLSR and PLSR validation method for different preprocessing methods.



ادامه شکل ۸-

Continue Figure 8.

به دست آمده، روش bagging-PLSR نسبت به روش PLSR دقت بالاتری در برآورد مقدار کربن آلی خاک نشان داد. از طرفی، دقت تخمین دو روش بسته به گستره مقدار کربن آلی خاک متفاوت است. به عبارتی، به ازای مقادیر کربن آلی $< 1/2$ درصد، دقت مدل‌ها تا حدودی کاهش می‌یابد. یکی از مزایای استفاده از داده‌های طیفی آن است که با افزایش مقیاس مکانی، امکان استفاده از باندهای تصاویر ماهواره‌ای در برآورد ویژگی‌های مورد نظر وجود خواهد داشت. به طور کلی روش‌های پیش‌پردازش و مدل‌های مورد استفاده در این پژوهش می‌تواند به عنوان مطالعات مبنای پژوهش‌های آینده مورد استفاده قرار گیرد. همچنین لازم است این مدل‌ها در مورد سایر خاک‌ها و مناطق دیگر کشور نیز مورد آزمون و ارزیابی قرار گیرند.

نتیجه‌گیری

نتایج این مطالعه نشان داد با استفاده از بازتاب طیفی خاک در محدوده VIS، NIR و SWIR می‌توان مقدار کربن آلی خاک را مورد مطالعه قرار داد. منحنی‌های طیفی خاک‌های مختلف سه مشخصه جذب در طول موج‌های ۱۴۱۴، ۱۹۱۳ و ۲۲۰۷ را نشان داد که مربوط به مقدار آب موجود در شبکه کانی‌های رسی خاک و رطوبت هیگروسکوپی خاک بوده و می‌تواند به عنوان مشخصه‌های منحصربه‌فرد هر خاک محسوب شود. این باندها اهمیت زیادی در برآورد مقدار کربن آلی خاک دارند. پیش‌پردازش داده‌های طیفی خاک و انتخاب مناسب‌ترین روش پیش‌پردازش، یکی از مهم‌ترین عوامل تأثیرگذار بر دقت روش‌های PLSR و bagging-PLSR در برآورد مقدار کربن آلی خاک بود. بر اساس نتایج

راهنمایی در پردازش‌های طیفی سپاسگزاری
می‌نماید.

سپاسگزاری

بدین‌وسیله نویسندگان از آقای دکتر ویسکارا
راسلاستاد دانشگاه سیدنی استرالیا به خاطر کمک و

منابع

1. Arshad, M.A., and Martin, S. 2002. Identifying critical limits for soil quality indicators in agro-ecosystems. *Agriculture, Ecosystems and Environment*. 88: 2. 153-160.
2. Ben-Dor, E., and Banin, A. 1995. Near infrared analysis as a rapid method to simultaneously evaluate several soil properties. *Soil Sci. Soc. Amer. J.* 59: 364-372.
3. Ben-Dor, E. 2002. Quantitative remote sensing of soil properties. *Advances in Agronomy*. Academic Press, Pp: 173-243.
4. Ben-Dor, E., Irons, J.R., and Epema, G.F. 1999. Soil reflectance, P 111-188. In: A.N. Rencz (Ed.), *Remote Sensing for Earth Sciences. Manual of Remote Sensing*. John Wiley and Sons, Inc., Toronto.
5. Bouma, J. 2002. Land quality indicators of sustainable land management across scales. *Agriculture, Ecosystems and Environment*. 88: 2. 129-136.
6. Breiman, L. 1996. Bagging predictors. *Machine Learning*. 24: 2. 123-140.
7. Brown, D.J., Shepherd, K.D., Walsh, M.G., Dewayne Mays, M., and Reinsch, T.G. 2006. Soil characterization with VNIR diffuse reflectance spectroscopy. *Geoderma*. 132: 273-290.
8. Brunet, D., Barthes, B.G., Chotte, J.L., and Feller, C. 2007. Determination of carbon and nitrogen contents in alfisols, oxisols and ultisols from Africa and Brazil using NIRS analysis: Effects of sample grinding and set heterogeneity. *Geoderma*. 139: 106-117.
9. Cambule, A.H., Rossiter, D.G., Stoorvogel, J.J., and Smaling, E.M.A. 2012. Building a near infrared spectral library for soil organic carbon estimation in the Limpopo National Park, Mozambique. *Geoderma*. 183-184: 41-48.
10. 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: 480-490.
11. Clark, R.N. 1999. Spectroscopy of rocks and minerals, and principles of spectroscopy, P 3-58. In: A.N. Rencz (Ed.), *Remote Sensing for Earth Sciences. Manual of Remote Sensing*. John Wiley and Sons, Inc., Toronto.
12. Clark, R.N., King, T.V.V., Klejwa, M., Swayze, G.A., and Vergo, N. 1990. High spectral resolution reflectance spectroscopy of minerals. *J. Geophysic. Res.* 95: 12653-12680.
13. Cozzolino, D., and Moron, A. 2003. The potential of near-infrared reflectance spectroscopy to analyse soil chemical and physical characteristics. *J. Agric. Sci.* 140: 65-71.
14. Esbensen, K. 1994. *Multivariate Data Analysis in Practice*. CAMO Software AS, Oslo, 598p.
15. Flint, A.L., and Flint, L.E. 2002. Particle density, P 229-240. In: J.H. Dane and G.C. Topp (Eds.), *Methods of Soil Analysis, Part 4-Physical Methods*. Soil Science Society of America, Inc., Madison, WI, USA.
16. Fyströ, G. 2002. The prediction of C and N content and their potential mineralization in heterogeneous soil samples using VIS-NIR spectroscopy and comparative methods. *Plant Soil*. 246: 139-149.
17. Gee, G.W., and Bauder, J.W. 1986. Particle-size analysis, P 383-412. In: A. Klute (Ed.), *Methods of Soil Analysis Part I, Physical and Mineralogical Methods*. Soil Science Society of America, Madison, WI, USA.
18. Geladi, P., and Kowalski, B.R. 1986. Partial least-squares regression: a tutorial. *Analytica Chimica Acta*. 185: 1-17.
19. Gisladdottir, G., and Stocking, M. 2005. Land degradation control and its global environmental benefits. *Land Degradation and Development*. 16: 99-112.
20. Gupta, P.K. 2000. *Soil, Plant, Water and Fertilizer Analysis*. Agrobios, New Delhi, India.

21. Temporal and spatial variability of soil organic matter and total nitrogen in an agricultural ecosystem as affected by farming practices. *Geoderma*. 139: 3-4. 336-345.
22. Kissling, M., Hegetschweiler, K.T., and Rusterholz, H.P. 2009. Short-term and long-term effects of human trampling on above-ground vegetation, soil density, soil organic matter and soil microbial processes in suburban beech forests. *Appl. Soil Ecol.* 42: 3. 303-314.
23. Munson, S.A., and Carey, A.E. 2004. Organic matter sources and transport in an agriculturally dominated temperate watershed. *Appl. Geochem.* 19: 1111-1121.
24. Page, A.L., Miller, R.H., and Keeney, D.R. 1982. *Methods of Soil Analysis, Part II, Physical Properties*. ASA, SSSA, Madison, WI.
25. Prescott, C.E., and Maynard, D.G. 2000. Humus in northern forests: friend or foe? *For. Ecol. Manag.* 133: 23-36.
26. Seely, B., Welham, C., and Juan, A. 2010. Blanco Towards the application of soil organic matter as an indicator of forest ecosystem productivity: deriving thresholds, developing monitoring systems, and evaluating practices. *Ecol. Indic.* 5: 10. 999-1008.
27. Sharma, S., Mohanty, B.P., and Zhu, J. 2006. Including topography and vegetation attributes for developing pedotransfer functions. *Soil Sci. Soc. Amer. J.* 70: 1430-1440.
28. Shepherd, K.D., and Walsh, M.G. 2002. Development of reflectance spectral libraries for characterization of soil properties. *Soil Sci. Soc. Amer. J.* 66: 3. 988-998.
29. Shepherd, K.D., and Walsh, M.G. 2007. Infrared spectroscopy-enabling an evidence based diagnostic surveillance approach to agricultural and environmental management in developing countries. *J. Near Inf. Spect.* 15: 1. 1-19.
30. Shukla, M.K., Lal, R., and Ebinger, M. 2006. Determining soil quality indicators by factor analysis. *Soil and Tillage Research.* 87: 2. 194-204.
31. Six, J., and Paustian, K. 2014. Aggregate-associated soil organic matter as an ecosystem property and a measurement tool. *Soil Biol. Biochem.* 68: 4-9.
32. 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*. 202-203: 161-170.
33. Viscarra Rossel, R.A. 2008. ParLeS software for chemometric analysis of spectroscopic data. *Chemometrics and Intelligent Laboratory Systems.* 90: 72-83.
34. Viscarra Rossel, R.A., Walter, C., and Fouad, Y. 2003. Assessment of two reflectance techniques for the quantification of field soil organic carbon, P 697-703. In: J. Stafford and A. Werner (Eds.), *Precision Agriculture. Fourth European Conference on Precision Agriculture*. Wageningen Academic Publishers, Berlin.
35. Walkley, A.J., and Black, I.A. 1934. An examination of Degtjareff method for determining soil organic matter and a proposed modification of the chromic acid titration method. *Soil Science.* 37: 29-38.
36. Walvoort, D.J.J., and McBratney, A.B. 2001. Diffuse reflectance spectrometry as a proximal sensing tool for precision agriculture, P 503-507. In: G. Grenier and S. Blackmore (Eds.), *ECPA 2001, Third European Conference on Precision Agriculture*, vol. 1. Agro Montpellier.



Estimating Soil Organic Carbon Using Hyperspectral Data in Visible, Near-infrared and Shortwave-infrared (VIS-NIR-SWIR) Range

E. Babaeian¹ and *V.R. Jalali²

¹Assistant Prof., Soil Conservation and Watershed Management Research Institute,

²Assistant Prof., Dept. of Soil Science, Shahid Bahonar University of Kerman

Received: 08/15/2015; Accepted: 12/13/2015

Abstract

Background and Objectives: Soil organic carbon is a main soil property and particularly important for development and sustainable management of agricultural systems. Soil Organic Matter content, which is typically measured in the form of soil organic carbon SOC content, is commonly regarded as a key indicator of soil quality and utilization, The presence of SOM has been proved to be beneficial for soil productivity, water holding capacity and carbon sequestration. Earlier studies showed that SOM is vulnerable to anthropogenic activities such as farming practices and other economic development activities. Conventional laboratory analyses for measuring soil organic carbon, especially in large scale, are expensive and time consuming. For this reason, fast and accurately assess the amount of organic carbon can be a very valuable measure for long-term management of soil. The objectives of this study were: i) studying of proximal spectral reflectance of soils for estimating soil organic carbon by PLSR and bagging-PLSR methods and ii) investigating the effects of different preprocessing methods on performance of estimated soil organic carbon.

Materials and Methods: A total of 200 composite soil samples on watershed scale (calibration data) and 40 soil samples on farm scale (validation data) from two different depths (0-10 and 10-30 cm) collected within a radius of 10 meters and after air drying, they were passed through 2 mm sieve. Some physicochemical characteristics of soils were measured in the laboratory. Consequently, proximal spectral reflectance of the soil samples within the VIS-NIR and SWIR (400-2500 nm) range was measured using a handheld spectroradiometer apparatus and correlation between 2000 bands and soil organic carbon were determined.

Results: Results indicated that the best preprocessing methods to calibrate PLSR model were wavelet detrending (RPD=1.94) and SNV with median filter (RPD=1.92). The best PLSR and bagging-PLSR model for the estimation was obtained with 17 factors. Bagging-PLSR method has high performance (RMSE=0.142-1.03 %) than PLSR method (RMSE=0.167-1.11 %) for estimating soil organic carbon. In both methods, the accuracy was decreased while soil organic carbon was bigger than 1.2 percent.

Conclusion: Using the soil spectral reflectance in the range of VIS, NIR and SWIR can examine the soil organic carbon. The spectral curves of different soils showed three absorbance properties at wave lengths 1414, 1913 and 2207 that was the amount of water in clay network and soil hygroscopic water, so they can be considered as a unique characteristics for each soil. These spectral bands are very important to estimate the amount of soil organic carbon. Soil spectral data pre-processing and selection of the most suitable pre-processing method was one of the most important factors affecting the accuracy of bagging-PLSR and PLSR method to estimate the amount of soil organic carbon. Based on the results, the method of bagging-PLSR showed higher accuracy than the PLSR method to estimate the amount of soil organic carbon.

Keywords: Soil spectral reflectance, Spectral preprocessing, PLSR; Soil organic carbon

* Corresponding Authors; Email: v.jalali@uk.ac.ir