



سپش از دور

GIS ایران



سنجش از دور و GIS ایران  
Iranian Remote Sensing & GIS

سال دهم، شماره دوم، تابستان ۱۳۹۷  
Vol.10, No. 2, Summer 2018

۱۵-۲۲

## برآورد مقدار ماده آلی خاک، با استفاده از داده‌های طیفی و مدل‌های آماری رگرسیون

### حداقل مربعات جزئی و رگرسیون مؤلفه اصلی

حمیدرضا متین‌فر<sup>۱\*</sup>، حمید محمودزاده<sup>۲</sup>، آذر فاریابی<sup>۳</sup>

۱. دانشیار گروه علوم و مهندسی خاک، دانشکده کشاورزی، دانشگاه لرستان
۲. دانشجوی دکتری مدیریت منابع خاک، دانشکده کشاورزی، دانشگاه لرستان
۳. دانشجوی دکتری مدیریت منابع خاک، دانشکده کشاورزی، دانشگاه لرستان

تاریخ پذیرش مقاله: ۱۳۹۶/۱۲/۲۱

تاریخ دریافت مقاله: ۱۳۹۵/۷/۲۱

#### چکیده

ماده آلی خاک از مهم‌ترین ویژگی‌های فیزیکوشیمیایی خاک است که در تعیین کیفیت و مدیریت آن نقش بسزایی دارد. مقدار کربن آلی خاک دچار تغییرات زمانی و مکانی بالایی می‌شود و بنابراین، تعیین آن در آزمایشگاه هزینه‌بر، مشکل و وقت‌گیر است. طیف‌سنجی مرئی-مادون قرمز نزدیک، از نظر کاهش زمان و هزینه، روش توجیه‌پذیری است که برای بررسی کربن آلی خاک پیشنهاد شده است. هدف این پژوهش بررسی کربن آلی خاک از طریق طیف‌سنجی مرئی-مادون قرمز نزدیک و برآورد آن با استفاده از مدل‌های آماری، PCA، PLSR و PCR است. برای این منظور، چهل نمونه خاک از عمق صفر تا ۳۰ سانتی‌متری، به روش سیستماتیک تصادفی، بر مبنای مطالعات پیشین و تعیین طبقات متفاوت خاک‌های منطقه، برداشت شد. تجزیه شیمیایی خاک‌ها طبق روش‌های استاندارد صورت گرفت. بازتاب طیفی نمونه‌های خاک در محدوده طیفی ۳۵۰ تا ۲۵۰۰ نانومتر اندازه‌گیری شد و پس از اعمال روش‌های پیش‌پردازش فیلتر ساویتزکی-گلای، به کمک تحلیل مؤلفه‌های اصلی (PCA)، رگرسیون حداقل مربعات جزئی (PLSR) و رگرسیون مؤلفه اصلی (PCR)، کربن آلی خاک برآورد شد. نتایج این پژوهش نشان داد که فیلتر ساویتزکی-گلای قوی‌ترین روش پیش‌پردازش داده‌های طیفی بوده است. نتایج مدل‌سازی به روش PLSR نشان داد مقادیر  $R^2$ ، RMSE و RPD در مرحله اعتبارسنجی، برای پیش‌بینی ماده آلی، به ترتیب ۰/۸۵، ۰/۱۴ و ۲/۷۸ بوده در حالی که نتایج مدل‌سازی PCR، برای پارامترهای آماری یادشده، به ترتیب ۰/۷۸، ۰/۱۹ و ۲/۰۵ است و دقت بیشتر روش آماری PLSR را، در مقایسه با روش PCR، برای مدل‌سازی برآورد کربن آلی خاک، می‌رساند. بنابراین، به نظر می‌رسد مدل PLSR، برای پیش‌بینی سریع کربن آلی خاک‌های مناطق خشک و نیمه‌خشک، کارایی و دقت بیشتری داشته باشد.

کلیدواژه‌ها: کربن آلی خاک، طیف‌سنجی، مدل PLSR، مدل PCR، مدل PCA.

\* نویسنده عهده‌دار مکاتبات: دانشکده کشاورزی، گروه علوم مهندسی خاک دانشگاه لرستان، تلفن: ۰۹۱۲۳۴۵۰۹۶۹.

## ۱- مقدمه

ماده آلی خاک عنصری کلیدی در ارتباط با ویژگی فیزیکی، شیمیایی، بیولوژیکی، حاصلخیزی (Cambou et al., 2015)، کیفیت خاک و محیط زیست (Nieder and Benbi, 2008) و توابع خاک است (McBratney et al., 2014; Nieder and Benbi, 2008). به عبارت دیگر، ماده آلی خاک یکی از مهم ترین ویژگی های فیزیکی شیمیایی خاک است که در تعیین رفتار و مدیریت خاک تأثیر بسزایی دارد و اطلاع از ویژگی های فیزیکی خاک و تغییر پذیری مکانی آنها، برای دستیابی به کشاورزی دقیق، لازم و ضروری است (Blackmer and White, 1998; Mabit and Bernard, 2009).

یکی از مهم ترین راهکارهای پیمان کیوتو<sup>۱</sup>، ترسیب کربن اتمسفر در خاک و نگهداشت کربن آلی خاک است (Conforti et al., 2015; Cambou et al., 2015). به علت نقش خاک در چرخه جهانی کربن و نقش پراهمیت پایش ماده آلی خاک، برای برآورد کیفیت و مدیریت خاک، تقاضا برای دریافت اطلاعات کربن آلی خاک در حال افزایش است اما دستیابی به این اطلاعات با روش های متداول قبلی خسته کننده، سخت و زمان بر است (Cambou et al., 2015). محاسبه کربن آلی خاک به چند علت مشکل است: نخست اینکه افزایش کربن خاک، با توجه به تغییرات عملیات کشاورزی، بسیار اندک است و دوم، ذخیره کربن در خاک، حتی در چند متر، به طور گسترده ای متغیر است (Viscarrá Cohen et al., 2005; Christy, 2008; Zornoza et al., 2008; Rossel et al., 2006). بنابراین جمع آوری گسترده اطلاعات مکانی خاک نیازمند روش های جدیدی است که، با کمترین هزینه و در کمترین زمان ممکن، اطلاعاتی درمورد چندین ویژگی خاک ثبت کند (Summers et al., 2009; Summers et al., 2008). طیف سنجی مرئی- مادون قرمز نزدیک، از نظر زمان و هزینه، روشی توجیه پذیر

است که برای بررسی کربن آلی خاک مناسب به نظر می رسد (Cambou et al., 2015); به طوری که در بیست سال گذشته، تمایل بسیاری برای کاربرد طیف مرئی- مادون قرمز نزدیک در برآورد ویژگی خاک و به ویژه، مقدار کربن آلی خاک وجود داشته است (et al., 2007; Stevens Stenberg et al., 2010). طیف سنجی مرئی- مادون قرمز نزدیک روشی سریع، ارزان، دارای حداقل آماده سازی نمونه، بدون ضرر و تخریب برای محیط زیست، دارای توان برآورد چندین ویژگی خاک (Shepherd and Walsh, 2002; Stenberg et al., 2010; Viscarrá Rossel and Behrens, 2010; Reeves III et al., 2002; McBratney et al., 2006; Guerrero et al., 2010) و در مزرعه و آزمایشگاه (Cambou et al., 2015)، کاربردی است.

تحقیقات نشان داده است که، در گستره مرئی و مادون قرمز نزدیک (VIS-NIR)، طول موج ۴۰۰ تا ۳۰۰۰ نانومتر) و مادون قرمز میانی (MIR)، طول موج ۳۰۰۰ تا ۳۰۰۰۰ نانومتر)، می توان مقدار ماده آلی خاک را مطالعه کرد (Clark, 1999).

از آنجاکه بازتاب طیفی خاک بر اثر برهمکنش پیچیده هریک از اجزای خاک حاصل می شود، لازم است به ازای طول موجی مشخص و با استفاده از روش های آماری مناسب، مقدار کربن آلی خاک را تعیین کرد (Shepherd and Walsh, 2002; Conforti et al., 1995; Ben-Dor and Banin, 2015). برای آنالیز طیف بازتابی خاک، معمولاً از تکنیک های کالیبراسیون چندمتغیره (Viscarrá Rossel Stenberg et al., 2010; and Behrens, 2010)، مانند رگرسیون خطی چندگانه (MLR)<sup>۲</sup>، رگرسیون مؤلفه های اصلی (PCR)<sup>۳</sup>،

۱. Kyoto Protocol: پیمانی بین المللی است؛ به منظور کاهش انتشار گازهای گلخانه ای، که عامل اصلی گرم شدن زمین در دهه های اخیر محسوب می شوند.

2. Multiple Linear Regression  
3. Principal Components Regression

این پژوهش بررسی کربن آلی خاک، از طریق طیف‌سنجی مرئی- مادون قرمز نزدیک و به‌کارگیری مدل‌های آماری PCR، PLSR و PCA برای پیش‌بینی ماده آلی خاک است.

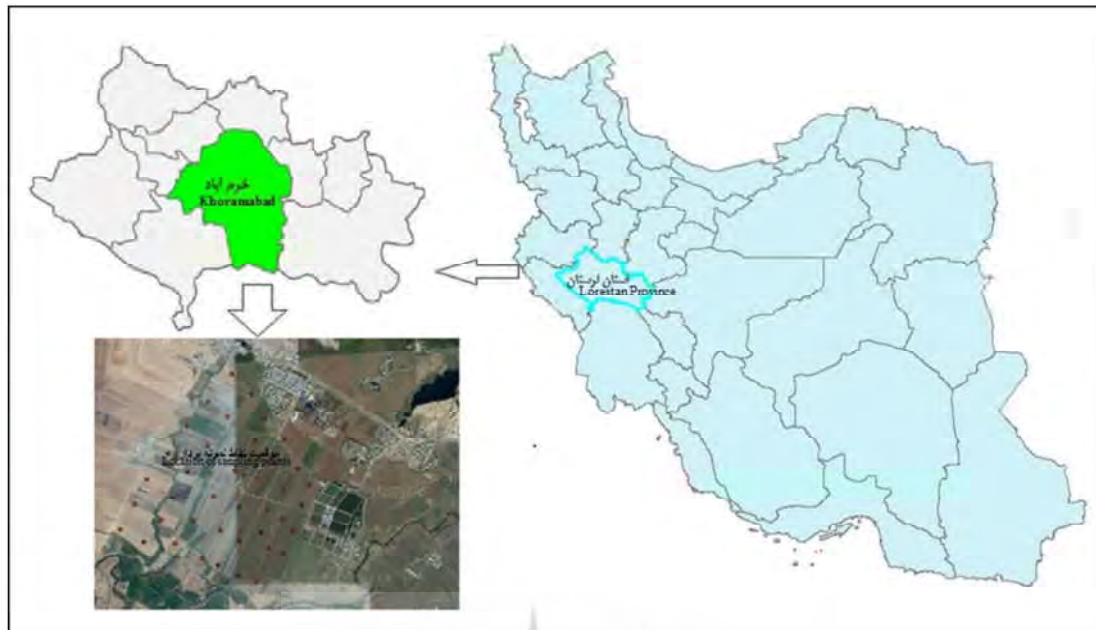
## ۲- مواد و روش‌ها

منطقه مورد مطالعه در این پژوهش بخشی از دشت خرم‌آباد، واقع در جنوب شهر خرم‌آباد، با ارتفاع حدود ۱۱۰۰ متر از سطح دریا و میانگین بارندگی سالانه ۵۱۶ میلی‌متر، است. این منطقه، با مساحت ۲۰۶ هکتار، بین ۱۱° ۱۶' ۴۸" تا ۱۷° ۱۷' ۴۸" طول شرقی و ۲۶° ۳۳' ۵۲" تا ۲۷° ۳۳' ۴۵" عرض شمالی واقع است. ناحیه مورد نظر تابستانی گرم و خشک و زمستانی سرد و مرطوب دارد. مطابق با نقشه رژیم‌های رطوبتی و حرارتی خاک‌های ایران، رژیم رطوبتی خاک‌های منطقه زیرک<sup>۵</sup> محسوب می‌شود که مؤید اقلیم مدیترانه‌ای (نیمه‌خشک) در این ناحیه است. رژیم حرارتی آن نیز ترمیک<sup>۶</sup> است (شکل ۱).

محل نمونه‌برداری براساس روش شبکه‌بندی (سیستماتیک- تصادفی) انتخاب شد. چهل نمونه خاک، از عمق ۰-۳۰ سانتی‌متر، نمونه‌برداری و پس از هواخشک کردن، از الک ۲ میلی‌متری عبور داده شد. برخی ویژگی‌های فیزیکی خاک‌ها اندازه‌گیری شد؛ شامل توزیع اندازه ذرات خاک (بافت خاک) به روش هیدرومتری (IPCC, 2007)، هدایت الکتریکی (EC) و واکنش خاک (pH) در عصاره گل اشباع (Soon and Hendershot, 2009)، آهک (Nelson and Walkly and (Sommers, 1982) و مقدار کربن آلی (Black, 1934) به روش‌های معمول آزمایشگاهی (جدول ۱).

رگرسیون حداقل مربعات جزئی (PLSR)<sup>۱</sup>، شبکه عصبی مصنوعی (ANN)<sup>۲</sup> و روش‌های حذف پیوسته (CR)<sup>۳</sup>، استفاده می‌شود (Shepherd and Walsh, 2007; Viscarra Rossel and Behrens, 2010; Viscarra Rossel and Behrens, 2010; Farifteh et al., 2007; Conforti et al., 2013; Aichi et al., 2009). در بین روش‌های آماری چندمتغیره آنالیز طیف‌سنجی مرئی- مادون قرمز نزدیک، PLSR روش متداول‌تر و کامل‌تری است (Viscarra Rossel and Behrens, 2010). روش آماری PLSR روش رگرسیونی خطی مفیدی، برای کاهش ماتریس پیچیده طیف‌ها به مؤلفه‌های متعامد<sup>۴</sup> کمتری شمرده می‌شود؛ به عبارتی دیگر، روشی مناسب برای مطالعه ویژگی‌های خاک است که داده‌های طیفی را تمامی طول موج‌ها به کار می‌گیرد و با گنجاندن آنها در متغیرهای کمتری (فاکتورهای تخمینگر) و ایجاد مدلی خطی، ویژگی خاکی مورد نظر را محاسبه می‌کند (Nocita et Blackmer et al., 1998; Reeves III et al., 2002; al., 2013; بیشترین مطالعات طیف‌سنجی مرئی- مادون قرمز نزدیک در مورد ارزیابی ماده آلی خاک بوده است. نوع روش‌های مورد استفاده در پیش‌پردازش داده‌های طیفی و ترکیب مناسب آنها اثر چشمگیری در دقت نتایج مدل خواهد داشت (Khayamim et al., 2015). روش‌های طیف‌سنجی بازتابی بر مبنای حساسیت ترکیبات آلی و معدنی خاک به بازتاب امواج مرئی و مادون قرمز نزدیک استوار شده است و از این ویژگی برای مطالعات کشاورزی و زیست‌محیطی خاک‌ها استفاده می‌شود (Khayamim et al, 2015). با وجود پژوهش‌های مهم در زمینه طیف‌سنجی بازتابی مرئی- مادون قرمز نزدیک خاک‌ها، طیف‌سنجی ویژگی‌های کیفی خاک‌های آهکی ایران به ندرت انجام شده است. با توجه به گستردگی خاک‌های آهکی در کشور و لزوم مدیریت بهتر این خاک‌ها و تولید بیشتر در آنها، ضرورت دسترسی به داده‌های خاک، از راه‌های ارزان‌تر، سریع‌تر و ایمن‌تر، تحقیق و بررسی در این مورد را ضروری می‌نماید. هدف

1. Partial least Squared Regression
2. Artificial Neural Networks
3. Continuum Removal
4. orthogonal
5. xeric
6. thermic



شکل ۱. موقعیت منطقه مورد مطالعه و نقاط نمونه برداری شده

مؤثری در بهبود کالیبراسیون دارد ( Ben-Dor and Stenberg et al., 2010; Banin, 1995). روش‌های پیش پردازش با هدف حذف و یا به حداقل رساندن تأثیرات یاد شده انجام شد. این تأثیرات شامل تغییر خط زمین، شبیدار و صاف کردن منحنی‌ها می‌شود. به دلیل امکان وجود نویز زیاد در طول موج‌های کمتر از ۴۰۰ و بیشتر از ۲۴۰۰ نانومتر، این محدوده از تمامی طیف‌ها حذف و در مدل‌سازی، به کار نرفت. سپس، فیلتر ساویتزکی-گلای<sup>۲</sup> (Savitzky and Golay, 1964) با اندازه پنجره ۱۷ پیکسل و چند جمله‌ای درجه ۲، برای از بین بردن نویزهای احتمالی دستگاه اندازه‌گیری طیفی، روی ماتریس بازتاب‌های طیفی اجرا شد. سپس به منظور کاربردی کردن نتایج حاصل از این مطالعه، محدوده‌های ۹۵۰-۱۴۰۰-۱۹۰۰ نانومتر، شامل باندهای جذب بخار آب اتمسفر (۱۴۰۰ و ۱۹۰۰ نانومتر)، از باندهای طیفی همه نمونه‌ها حذف شدند زیرا، در این محدوده‌های طیفی، تقریباً هیچ انرژی‌ای به سنجنده‌های ماهواره‌ای نمی‌رسد و اغلب سنجنده‌ها در

به منظور اندازه‌گیری بازتاب طیفی نمونه‌های خاک، دستگاه اسپکترورادیومتر زمینی<sup>۱</sup> به کار رفت. منحنی‌های طیفی نمونه‌ها در دامنه مرئی، مادون قرمز نزدیک و مادون قرمز میانی (۲۵۰۰-۳۵۰۰ نانومتر) در تاریخانه اندازه‌گیری شد. اندازه‌گیری‌های طیفی به کمک پیستول با گستره دید ۸ درجه و منبع نور ۱۵۰ وات صورت گرفت. برای کاهش خطای (افزایش نسبت سیگنال به نویز) حاصل از شرایط محیطی (رطوبت، دما، غبار)، در حین طیف‌سنجی، دستگاه اسپکترورادیومتر، در هر بار اسکن، از هر نمونه خاک ۲۵ طیف در سه تکرار برداشت کرد و برای به حداکثر رساندن نسبت سیگنال به نویز، از آنها میانگین گرفته شد و سپس مقادیر انعکاس به جذب تبدیل شد (Khayamim et al., 2015). به دلیل اینکه اندازه‌گیری‌های طیفی (به ویژه در صحرا) تحت تأثیر عوامل ناخواسته مانند رطوبت خاک، زبری سطح خاک، بقایای آلی درشت، آلودگی پروب به علت گردوغبار، تغییر فاصله سنسور و خاک و تأثیرات پرتوافکنی جوی قرار می‌گیرند؛ پیش پردازش داده‌های طیفی نقش

1. FieldSpec®3, ASD, FR, USA  
2. Savitzky-Golay

پیش پردازش طیفها و مدل سازی با استفاده از نرم افزار Unscrambler X 10.3 صورت گرفت.

## ۲-۱- تحلیل مؤلفه های اصلی

در روش تحلیل مؤلفه اصلی، متغیرهای موجود در فضای چندبعدی همبسته ای به مجموعه ای از مؤلفه های غیرهمبسته خلاصه می شوند که هر یک از آنها ترکیبی خطی از متغیرهای اصلی است. مؤلفه های غیرهمبسته حاصل مؤلفه های اصلی (PC) <sup>۴</sup> نامیده می شوند که از بردارهای ویژه ماتریس کوواریانس یا ماتریس همبستگی متغیرهای اصلی به دست می آیند.

از مشکلات پردازش داده های طیفی می توان حجم وسیع داده ها در طول موج های متفاوت، اثر هم پوشانی عوامل متفاوت در هر طول موج و اطلاعات تکراری در طول موج های مجاور را نام برد. زمانی که تعداد متغیرهای پیش بینی کننده بیشتر از نمونه های اندازه گیری شده باشد، درجه هم پوشانی زیاد و منجر به بیش برآزش می شود. در مورد چنین داده هایی، پیش از برآزش مدل، تحلیلی چندمتغیره انجام می شود تا متغیرهای مهم تأثیرگذار تعیین شوند. در این فرایند، طول موج های مهم، که با متغیرهای وابسته مرتبط اند، مشخص می شوند (عباسی، ۱۳۸۸). برای تحلیل رگرسیونی، از روش تحلیل مؤلفه های اصلی PCA استفاده می شود. اولین مؤلفه بیشترین اطلاعات را به همراه دارد و از دیدگاه آماری، بیشترین واریانس را توضیح می دهد. مؤلفه دوم بیشترین اطلاعات باقی مانده را، که مؤلفه اول توضیح نداده است، بیان می کند. با انتخاب چند مؤلفه اصلی اول، سایر مؤلفه ها از محاسبات بعدی حذف می شوند (Esbensen, 2006).

این نواحی تصویربرداری نمی کنند. داده های طیفی معمولاً تحت تأثیر پخش نور در محیط، گردوغبار و بخار آب هوا قرار می گیرند و با نویز همراه اند. به منظور حذف این آثار که منجر به رفتار غیرخطی باندها، جابه جایی طیفی و کاهش نسبت سیگنال به نویز می شوند، از الگوریتم های گوناگون پیش پردازش استفاده شده است. انواع روش های پیش پردازش شامل فیلتر میانه، فیلتر ساویتزکی- گلائی (Ibid.)، نرمال سازی براساس میانگین و فیلتر گوسین صورت گرفته است.

پس از پیش پردازش، به کمک آنالیز مؤلفه های اصلی، اطلاعات طیفی به دست آمده تحلیل می شوند تا بتوان جهات دارای حداکثر واریانس داده های طیفی را جست و جو کرد. سپس، از رگرسیون حداقل مربعات جزئی و رگرسیون مؤلفه اصلی، برای پیش بینی کربن آلی، استفاده شد. بدین منظور، ابتدا مجموعه داده ها به صورت تصادفی به دو گروه کالیبراسیون و اعتبارسنجی تقسیم شد. تعداد نمونه ها، برای گروه کالیبراسیون، سی (حدود دوسوم داده ها) و برای اعتبارسنجی، ده نمونه (حدود یک سوم داده ها) در نظر گرفته شد. مقادیر کربن آلی، برای گروه اعتبارسنجی، پیش بینی و دقت مدل ها ارزیابی شد. به قصد ارزیابی دقت در هر یک از مراحل تخمین، آماره های متفاوتی به کار رفت که شامل ضریب تبیین ( $R^2$ )<sup>۱</sup>، ریشه میانگین مربعات خطا (RMSE)<sup>۲</sup>، نسبت انحراف دقت (RPD)<sup>۳</sup> می شود.

آماره RPD یکی از متداول ترین شاخص های ارزیابی دقت در مطالعات طیفسنجی است که از نسبت خطای استاندارد داده ها به RMSE مدل تخمینگر به دست می آید ( $RPD=SD/RMSE$ ). زمانی که مقدار RPD کمتر از ۱/۵ باشد، پیش بینی مدل ضعیف است؛ اگر بین ۱/۵ تا ۲ باشد، پیش بینی قابل قبول و در صورتی که بیشتر از ۲ باشد، پیش بینی کاملاً مناسب است (Williams, 2001; Summers et al., 2011; Summers et al., 2009). توصیف آماری متغیرها،

1. Coefficients of Determination
2. Root Mean Square Error of Prediction
3. Ratio of Prediction to Deviation
4. Principal Component

تشکیل شده است. یکی از مهم ترین دلایل استفاده از تکنیک حداقل مربعات جزئی این است که این تکنیک به فرض نرمال بودن جامعه و حجم نمونه متکی نیست. به طور کلی، دو نوع رویکرد برای بر آورد پارامترهای مدل معادلات ساختاری وجود دارد که عبارت اند از رویکرد مبتنی بر کوواریانس و رویکرد مبتنی بر واریانس. رویکرد اول در تلاش است اختلاف بین کوواریانس های نمونه و آنچه مدل نظری پیش بینی کرده است به کمترین میزان برساند. برخلاف رویکرد نخست، رویکرد حداقل مربعات جزئی حداکثر کردن واریانس متغیرهای وابسته ای است که با استفاده از متغیرهای مستقل تعریف می شود. همانند سایر مدل های معادلات ساختاری، مدل حداقل مربعات جزئی نیز دارای بخش ساختاری است که منعکس ارتباط بین متغیرهای پنهان و یک جزء اندازه گیری را نشان می دهد. این تکنیک امکان بررسی هم زمان روابط متغیرهای پنهان و متغیرهای قابل مشاهده را فراهم می آورد. این روش زمانی استفاده می شود که حجم نمونه کوچک باشد و یا توزیع متغیرها نرمال نباشد.

### ۳- نتایج و بحث

جدول ۱ برخی از ویژگی های فیزیکی و شیمیایی خاک منطقه مورد مطالعه را، در دو گروه کالیبراسون و اعتبارسنجی، نشان می دهد. نتایج نشان داد که مقدار آهک خاک های منطقه بالاست و اسیدیته خاک منطقه نیز خنثی، به سمت قلیا، است. فراوانی بافت خاک منطقه نشان داد که بیش از ۷۰٪ خاک ها دارای بافت لوم رسی است و کمتر از ۵٪ خاک ها بافت لوم رسی سیلنتی دارند (شکل ۲). جدول ۲ آنالیز آماری برخی ویژگی های خاک مورد مطالعه را نشان می دهد. میانگین مقدار ماده آلی خاک ۲/۲۷٪ است که بالابودن مقدار ماده آلی را در دشت خرم آباد، در مقایسه با میانگین کشوری (۰/۵٪)، نشان می دهد.

۲-۲- مدل رگرسیون حداقل مربعات جزئی یا PLSR  
تحلیل رگرسیون حداقل مربعات جزئی شبیه به تحلیل مؤلفه های اصلی (PCA) است؛ با این تفاوت که در روش PLSR، متغیرهای وابسته  $Y$  (ماتریس غلظت ها) نیز، هم زمان با متغیرهای مستقل  $X$  ماتریس داده های طیفی، به تعداد مؤلفه های اصلی تبدیل می شود؛ به گونه ای که بیشترین درصد تغییرات در مؤلفه های اولیه مطرح می شود. بنابراین، یکی از مزایای رگرسیون PLSR تعیین مؤثرترین و یا حساس ترین طول موج ها به نسبت تغییرات متغیر  $Y$ ، براساس تعیین مؤلفه های اصلی انتخاب شده از متغیر  $X$ ، به شمار می رود. در واقع، هدف اصلی رگرسیون ساختن مدلی خطی براساس رابطه (۱) است:

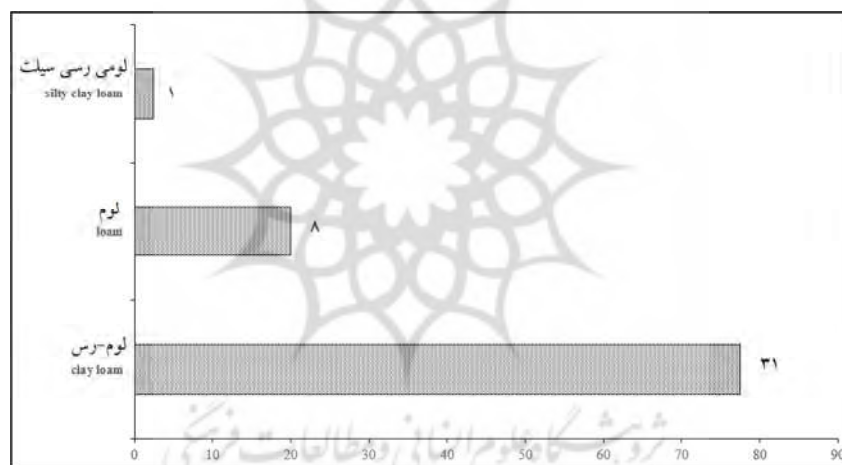
$$Y = Xb + E \quad \text{رابطه (۱)}$$

$Y$ : متغیر برآورد شده (در رگرسیون PLSR) در شکل ماتریس با ابعاد (متغیر  $Y \times 1$ ) و  $X$  متغیر مستقل پیشگویی کننده با ابعاد طول موج  $\times$  بازتابندگی است. ضمناً  $b$  ماتریسی شامل ضرایب رگرسیون و  $E$  ماتریسی از خطاهای باقی مانده است.

روش PLSR روش نسبتاً جدیدی برای ساخت معادلات رگرسیون است. این روش برای رگرسیون تک متغیری و چندمتغیری استفاده می شود؛ بنابراین، ممکن است چندین متغیر وابسته داشته باشد. برای ایجاد ارتباط بین متغیرهای وابسته و متغیرهای مستقل، PLS متغیرهای تبیینی (مستقل) جدیدی ایجاد می کند. برخلاف مدلیابی معادلات ساختاری مبتنی بر کوواریانس، کمترین مجذورات جزئی (PLS)، به جای بازتولید ماتریس کوواریانس تجربی، بر بیشترین واریانس تبیین شده متغیرهای وابسته با استفاده از متغیرهای مستقل تمرکز دارد. مشابه با تمامی مدلیابی معادلات ساختاری، مدل کمترین مجذورات جزئی از بخشی ساختاری، که ارتباط بین متغیرهای پنهان را نشان می دهد، و یک مؤلفه اندازه گیری، که شیوه ارتباط متغیرهای پنهان و نشانگرهای آنها را منعکس می کند،

جدول ۱. برخی ویژگی‌های خاک‌های به‌کاررفته در مرحله کالیبراسیون و اعتبارسنجی

داده‌های اعتبارسنجی (validation data)	داده‌های کالیبراسیون (calibration data)	پارامتر (parameter)
0.67	0.63	EC (dS.m <sup>-1</sup> ) شوری خاک
7.80	7.81	pH واکنش خاک
1.48	1.28	O.C (%) کربن آلی
34.01	34.83	T.N.V (%) مواد خنثی‌شونده
28	30	sand (%) درصد شن
41	39	silt (%) درصد سیلت
31	31	clay (%) درصد رس
لوم رس (clay loam)	لوم رس (clay loam)	بافت خاک



شکل ۲. فراوانی بافت خاک در منطقه مورد مطالعه، برحسب درصد

جدول ۲. آنالیز آماری برخی از ویژگی‌های خاک مورد مطالعه

ویژگی خاک (soil character)	انحراف معیار (standard deviation)	حداکثر (maximum)	حداقل (minimum)	میانگین (average)
T.N.V (%)	2.84	39.5	27.5	34.77
O.C (%)	0.33	2.63	1.05	1.32
pH	0.09	7.98	7.62	7.82
EC (dS.m <sup>-1</sup> )	0.22	1.71	0.44	0.65
Clay	4.51	40	20.5	30.80
Silt	5.39	54	31.5	39.26
Sand	5.72	41.5	13.5	29.95

### ۳-۱- ویژگی‌های طیفی خاک

به‌طور کلی، بازتاب طیفی خاک، در طول موج‌های پایین و بالا، ضعیف است و بنابراین وجود رطوبت، اکسیدهای آهن و آلومینیوم، و مقدار و نوع رس از اجزای اصلی خاک اند که طیف بازتابی را تحت تأثیر مستقیم قرار می‌دهند و ممکن است باعث بروز خطا در طیف‌سنجی نیز بشوند.

شکل ۳ طیف انعکاسی محدوده مرئی- مادون قرمز نزدیک به‌دست‌آمده از چهل نمونه خاک را نشان می‌دهد. شکل کلی همه طیف‌ها، در همه خاک‌ها، تقریباً مشابه است؛ به‌گونه‌ای که ویژگی‌های جذبی آب هیگروسکوپی و آب موجود در شبکه رس‌ها، در طول موج‌های ۱۴۱۴ و ۱۹۱۳ و ۲۲۰۷ نانومتر، در همه طیف‌ها مشاهده می‌شود؛ البته اختلافاتی در میزان شدت انعکاس وجود دارد. لی و همکاران<sup>۱</sup> (۲۰۰۷) گزارش دادند که طول موج‌های ۴۶۵-۴۵۰، ۹۶۵، ۱۴۰۹ و ۲۲۰۰-۱۷۷۵ نانومتر مهم‌ترین طول موج‌ها برای تشخیص کربن آلی شمرده می‌شوند. ویژگی جذبی رس در ۲۲۰۰ نانومتر مشخص می‌شود. کانی‌های خاک نور را، در ناحیه‌های مرئی- مادون قرمز نزدیک و میانی، جذب می‌کنند. پاسخ طیفی کانی‌های رسی در نتیجه ارتعاش مولکول‌های آب ساختمانی، گروه‌های هیدروکسیل، سیلیکات ساختمانی و کاتیون‌های اکتاهدرال، تتراهدرال و بین‌لایه‌ای است (Bouma, 2002). اکسیدهای آهن در محدوده مرئی و کربنات‌ها در ۲۳۰۰ نانومتر مشاهده می‌شوند. مقادیر انعکاس از طریق رابطه زیر به جذب تبدیل می‌شود و شکل ۳ نمودارهای جذب طیف مرئی- مادون قرمز شش نمونه خاک را نشان می‌دهد. در این رابطه  $A$  و  $R$ ، به‌ترتیب، مقادیر جذب و انعکاس برحسب درصد را نشان می‌دهند:

$$A = \log(1/R) \quad (2)$$

ویژگی جذبی آب در طول موج ۱۴۰۰ و ۱۹۰۰ نانومتر، به‌علت فرکانس ارتعاشات پایه مولکول آب، شامل کشش متقارن و نامتقارن پیوند O-H و خمش O-H، مشاهده می‌شود (Walvoort and Cambou, 2015). تمامی طیف‌های خاک در طول موج‌های ۱۴۰۰، ۱۹۰۰ و ۲۲۰۰ نانومتر، بیشترین میزان

جذب را دارند. جذب در باند ۱۴۰۰ نانومتر مربوط به گروه‌های عاملی O-H و C-H آلیفاتیک است؛ درحالی‌که جذب در طول موج ۱۹۰۰ نانومتر با گروه آمیدی N-H و O-H مرتبط می‌شود. در طول موج ۲۲۰۰ گروه‌های O-H فنولی، N-H آمیدی، N-H آمین و C-H آلیفاتیک عامل جذب می‌باشند (Cozzolino and Moron, 2003; Fideñcio et al., 2002).

منحنی میانگین طیفی برای سه نمونه خاک، در شرایط مقدار متفاوت کربن آلی و pH و شوری خاک، نشان داد که، میانگین طیفی نمونه اول، با وجود مقدار بالای کربن آلی خاک و امکان کاهش بازتاب طیفی، بالاتر از دو نمونه دارای مقدار کربن آلی کمتر است که علت آن ممکن است میانگین بیشتر شوری باشد (دست‌کم دوبرابر بیشتر از دو نمونه دیگر است) (شکل ۴). این نشان می‌دهد که ویژگی‌های فیزیکی و شیمیایی خاک اثر متقابل پیچیده‌ای در رفتار طیفی خاک دارند (Santra et al., 2009).

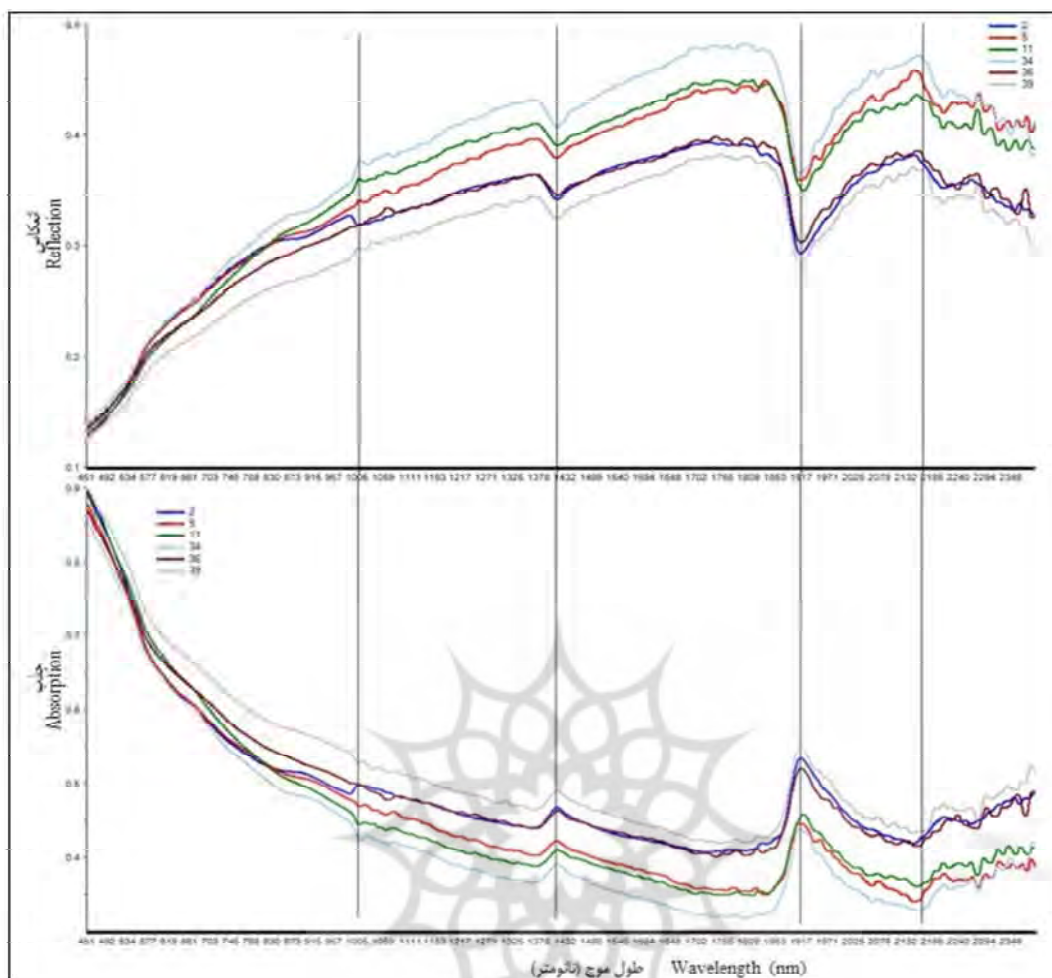
به‌طور کلی، با افزایش مقدار ماده آلی خاک، بازتاب طیفی کاهش یافت. تحقیقات نشان داده است هریک از اجزای ماده آلی خاک نیز اثر بازتاب متفاوتی در هریک از طول موج‌ها دارد (Shepherd and Walsh, 2002). در صورتی که مقدار کربن آلی خاک کم باشد باندهای مرئی در برآورد ماده آلی خطای بیشتری خواهند داشت. بنابراین انتظار می‌رود این روش‌ها در مناطق با مقدار کم مواد آلی نتایج ضعیفی داشته باشند؛ بنابراین، انتظار می‌رود نتایج این روش‌ها، در مناطق دارای سطح کربن آلی بسیار کمتر، ضعیف‌تر باشد.

### ۳-۲- تحلیل مؤلفه‌های اصلی

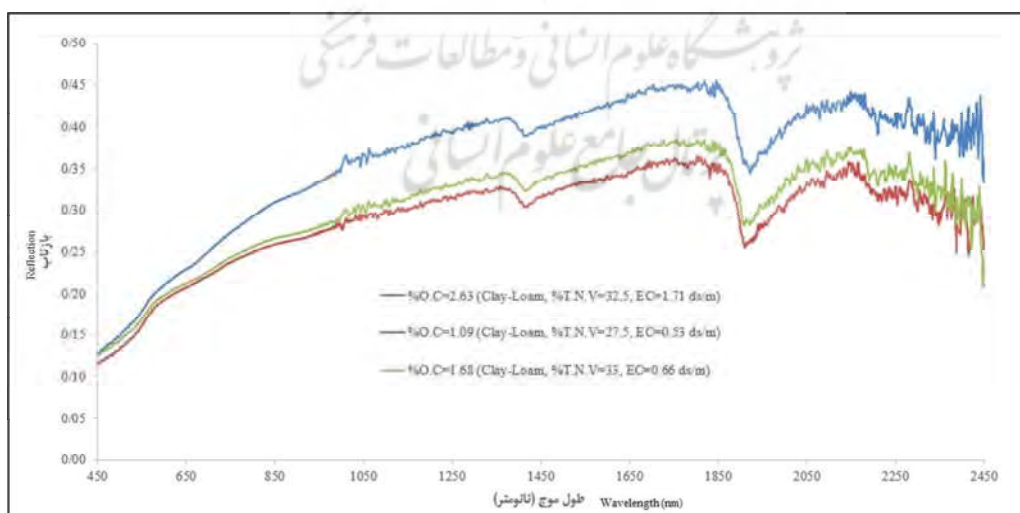
شکل ۵ نمودار امتیاز<sup>۲</sup> تحلیل مؤلفه‌های اصلی برای مقادیر جذبی طیف‌های محدوده مرئی- مادون قرمز نزدیک را نشان می‌دهد.

1. Lee et al.  
2. score

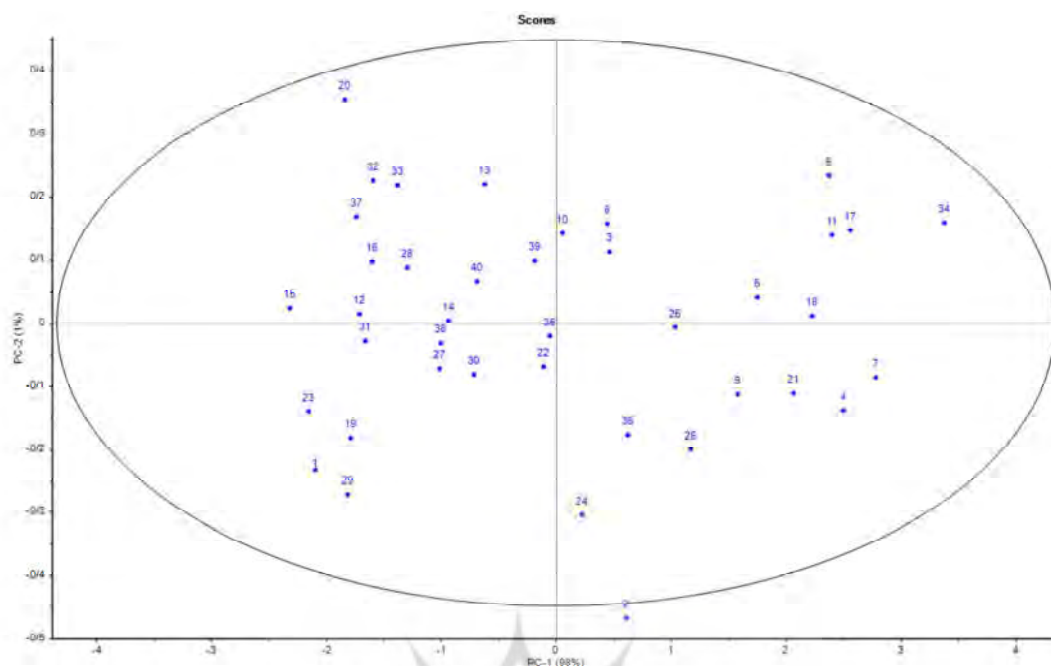




شکل ۳. نمودار انعکاس و جذب طیفی بخشی از نمونه‌های خاک منطقه مورد مطالعه



شکل ۴. منحنی میانگین طیفی برای سه نمونه خاک، با مقادیر متفاوت کربن آلی و آهک و شوری



شکل ۵. نمودار تحلیل مؤلفه‌های اصلی برای طیف‌های جذبی محدوده مرئی- مادون قرمز نزدیک

مؤلفه نخست ۹۸٪ از واریانس داده‌ها را به خود اختصاص می‌دهد و هریک از مؤلفه‌های دوم و سوم ۱٪ از تغییرات را توجیه می‌کند. با استفاده از این نمودار، می‌توان نمونه‌هایی را، که دارای طیفی متفاوت از بقیه‌اند، شناسایی کرد و چنانچه پس از بررسی‌های لازم، داده پرت تشخیص داده شوند، می‌توان از مجموعه داده‌ها حذفشان کرد. برطبق این نمودار، هیچ خوشه‌بندی خاصی در نمونه‌های خاک، براساس ویژگی‌های طیفی‌شان، مشاهده نمی‌شود. خاک‌های دارای بازتاب طیفی متفاوت با سایر خاک‌ها تأثیر چشم‌گیری در دقت مدل رگرسیونی خواهند داشت (Khayamim et al., 2015).

۳-۳- پیش‌بینی کربن آلی خاک  
مدل رگرسیون حداقل مربعات جزئی (PLSR) کربن آلی خاک، در شکل‌های ۶ و ۷، مقادیر اندازه‌گیری شده در برابر مقادیر پیش‌بینی شده ماده آلی را نشان می‌دهد. برای ارزیابی دقت مدل، از سه پارامتر  $R^2$ ، RMSE و

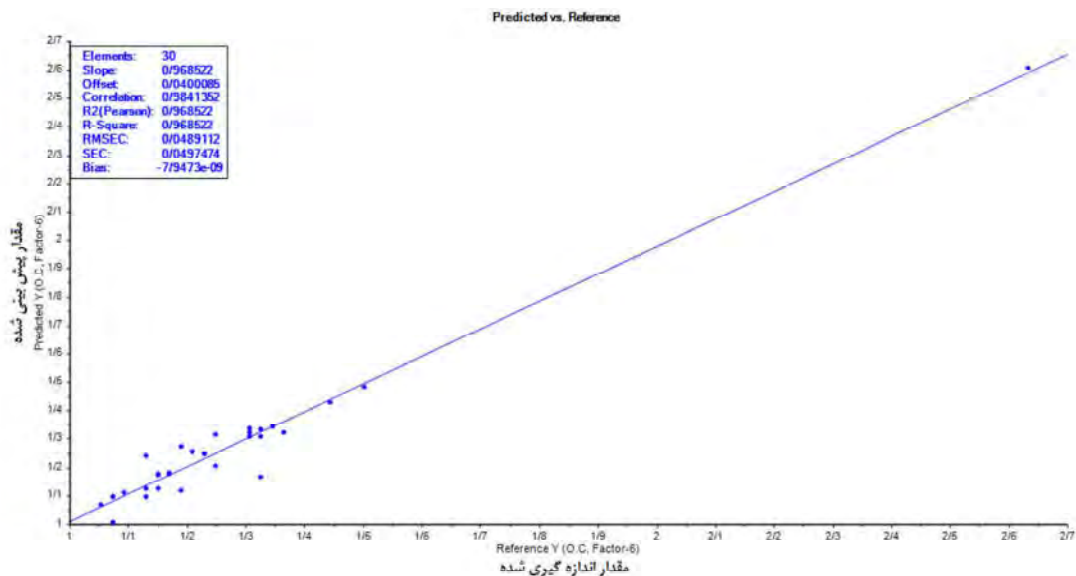
شکل ۵. نمودار تحلیل مؤلفه‌های اصلی برای طیف‌های جذبی محدوده مرئی- مادون قرمز نزدیک

مؤلفه نخست ۹۸٪ از واریانس داده‌ها را به خود اختصاص می‌دهد و هریک از مؤلفه‌های دوم و سوم ۱٪ از تغییرات را توجیه می‌کند. با استفاده از این نمودار، می‌توان نمونه‌هایی را، که دارای طیفی متفاوت از بقیه‌اند، شناسایی کرد و چنانچه پس از بررسی‌های لازم، داده پرت تشخیص داده شوند، می‌توان از مجموعه داده‌ها حذفشان کرد. برطبق این نمودار، هیچ خوشه‌بندی خاصی در نمونه‌های خاک، براساس ویژگی‌های طیفی‌شان، مشاهده نمی‌شود. خاک‌های دارای بازتاب طیفی متفاوت با سایر خاک‌ها تأثیر چشم‌گیری در دقت مدل رگرسیونی خواهند داشت (Khayamim et al., 2015).

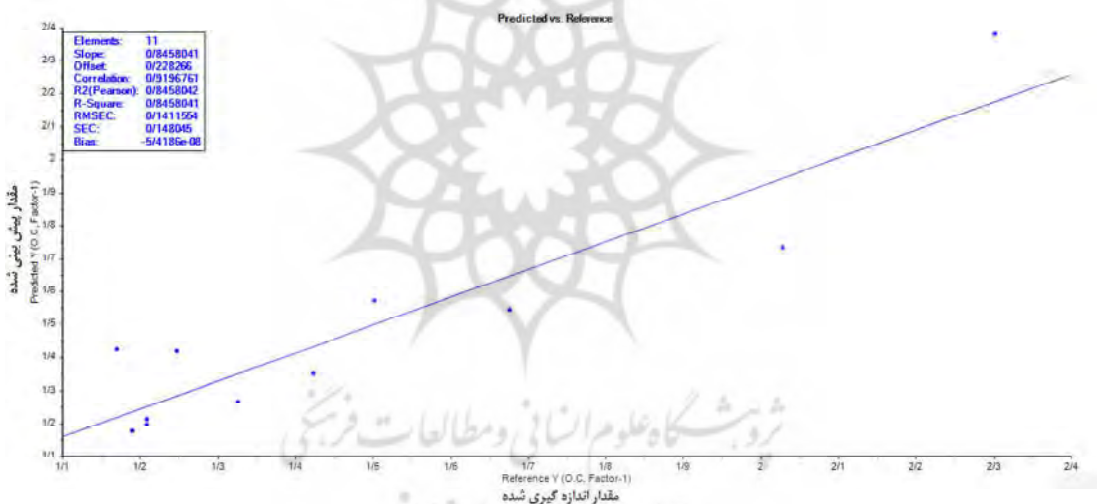
### ۳-۳- پیش‌بینی کربن آلی خاک

مدل رگرسیون حداقل مربعات جزئی (PLSR) کربن آلی خاک، در شکل‌های ۶ و ۷، مقادیر اندازه‌گیری شده در برابر مقادیر پیش‌بینی شده ماده آلی را نشان می‌دهد. برای ارزیابی دقت مدل، از سه پارامتر  $R^2$ ، RMSE و

1. Saeys et al.



شکل ۶. مقادیر اندازه گیری شده در مقابل مقادیر پیش بینی شده ماده آلی در داده های کالیبراسیون (مدل PLSR)



شکل ۷. مقادیر اندازه گیری شده در مقابل مقادیر پیش بینی شده ماده آلی در داده های اعتبارسنجی (مدل PLSR)

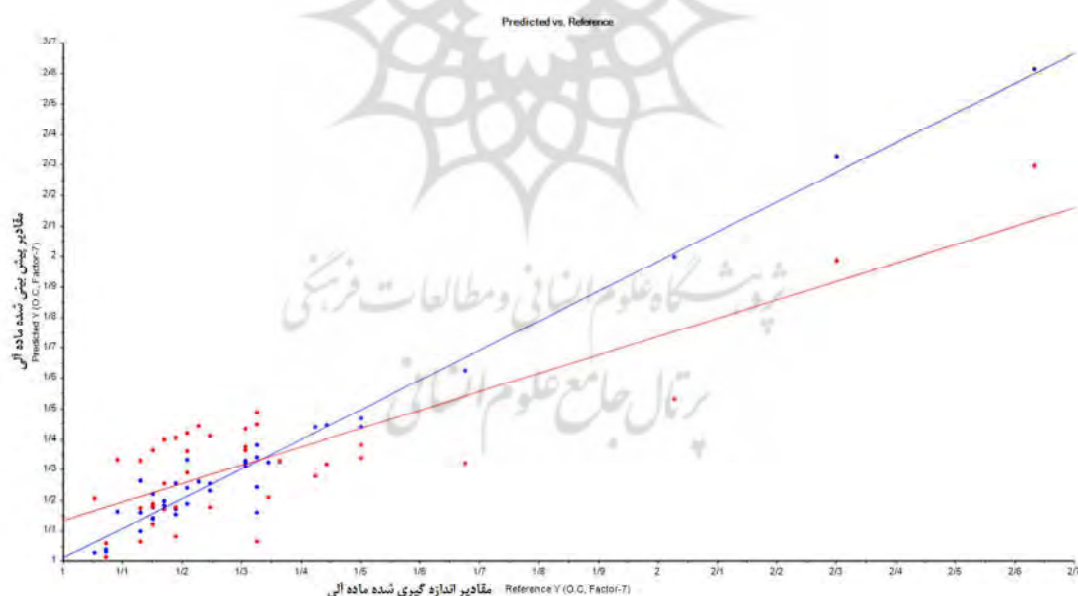
بیشترین مقدار  $R^2$  و RPD را داشته است. مقدار RPD به دست آمده در این پژوهش را می توان با مقادیری که دیگر محققان به دست آورده اند مقایسه کرد. استونس و همکاران<sup>۲</sup> (۲۰۰۷) مقدار RPD را، برای طیفسنجی مزرعه، ۲/۲۹ تا ۳/۰۴ و برای طیفسنجی تصویر، ۱/۴۶ گزارش کرده اند. لی و همکاران (۲۰۰۷) هم مقدار RPD را، برای برآورد کربن آلی با استفاده از مدل

دقت مدل، با وجود مقدار RPD کمتر از ۲، ضعیف است؛ وقتی بین ۲ و ۲/۵ باشد، پیش بینی تا حدودی مطلوب است و برای مقادیر بین ۲/۵ و ۳، پیش بینی مدل خوب برآورد می شود (Zornoza et al., 2008). نتایج طیفسنجی زورنوزا و همکاران<sup>۱</sup> (۲۰۰۸) نشان داد که مدل PLSR کربن آلی خاک را به خوبی پیش بینی کرده و مقادیر  $R^2$ ، RMSE و RPD به ترتیب ۰/۹۸، ۶/۲۵ و ۵/۷۵ بوده است. براساس همین پژوهش، پیش بینی کربن آلی خاک با مدل PLSR

1. Zornoza et al.
2. Stevens et al.

گرفتند که در تخمین مقدار کربن آلی خاک، بین مقادیر اندازه‌گیری شده و پیش‌بینی شده، تطابق بالایی وجود دارد و  $R^2$ ، RMSE و RPD، به ترتیب، ۰/۹۱ و ۰/۳۶ و ۳/۴ بوده است. تنوع نتایج تخمین مواد آلی خاک دلایل متفاوتی دارد: مواد آلی، در محدوده مادون قرمز نزدیک، ویژگی جذبی خاصی دارند اما این ویژگی‌ها در بیشتر موارد، ضعیف است و در مواردی که کانی‌شناسی خاک بسیار متنوع باشد یا دارای شن معتابهی باشد، به‌علت پخش نور در حضور کوارتز، تخمین ماده آلی ضعیف می‌شود (Stenberg et al., 2010). اختلاف مقادیر واقعی و پیش‌بینی شده، با استفاده از مدل، درباره تخمین مقادیر کربن آلی چشمگیر است و کم‌برآوردی و بیش‌برآوردی در داده‌های پیش‌بینی شده مشاهده می‌شود (شکل ۸). شایان ذکر است که تمامی تفسیرهای صورت‌گرفته براساس نتایج آنالیزهای آزمایشگاهی است.

PLSR، ۲/۸۷ به‌دست آورده‌اند. سامرز و همکاران<sup>۱</sup> (۲۰۰۹)، با بررسی سبده نمونه خاک در منطقه جنوب استرالیا، روش طیف‌سنجی مرئی-مادون قرمز نزدیک را در برآورد کربن آلی، مقدار رس، اکسیدهای آهن و کربنات‌های خاک ارزیابی کردند. آنها نتیجه گرفتند که برای درصد رس، کربن آلی، اکسید آهن و کربنات‌های خاک، به ترتیب،  $R^2$  برابر است با ۰/۶۶، ۰/۵۷، ۰/۶۱ و ۰/۶۹ و مقادیر RMSE نیز، به ترتیب، ۳/۱۳، ۰/۳۵، ۰/۲۳ و ۲/۹ درصد گزارش شده است. خیامیم و همکاران<sup>۲</sup> (۲۰۱۵) نتیجه گرفتند که مقدار  $R^2$  در گروه اعتبارسنجی برای ماده آلی، کربنات‌ها و گچ، به ترتیب، ۰/۶۱، ۰/۴۵، ۰/۸ و مقدار RMSE نیز، به ترتیب، ۰/۳، ۱۲/۵۱ و ۳/۴۳ درصد است. این نتایج نشان داد که پیش‌بینی مدل، برای گچ، کاملاً مناسب و برای ماده آلی، قابل قبول است؛ درحالی‌که برای کربنات‌های خاک، ضعیف است. آچی و همکاران<sup>۳</sup> (۲۰۰۹) نتیجه

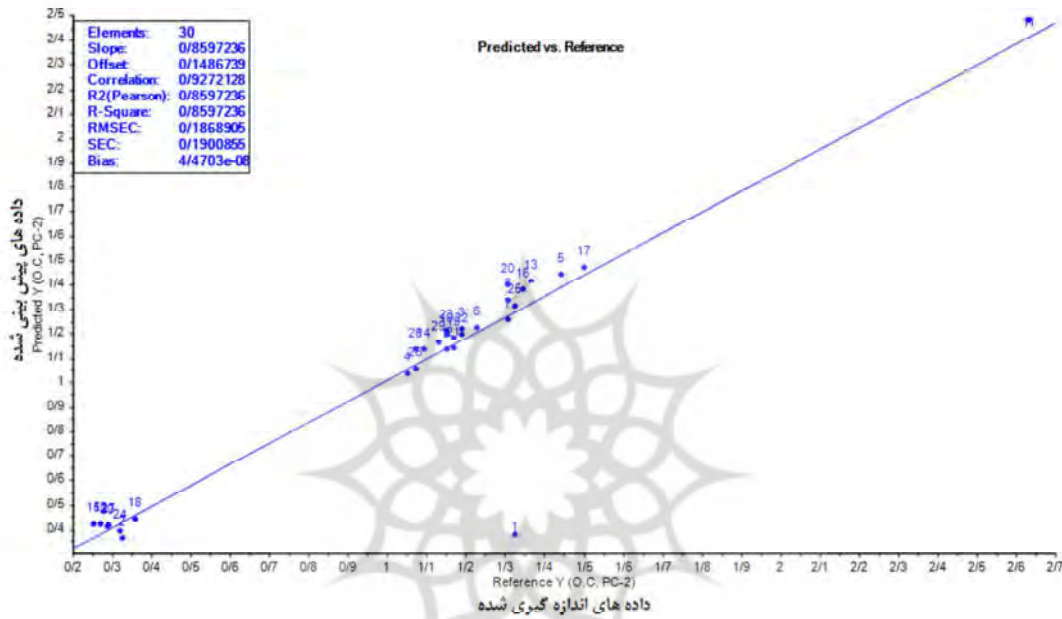


شکل ۸. مقادیر اصلی و پیش‌بینی شده کربن آلی خاک در مدل PLSR (رنگ آبی مربوط به داده‌های اندازه‌گیری شده و رنگ قرمز مربوط به داده‌های پیش‌بینی شده است)

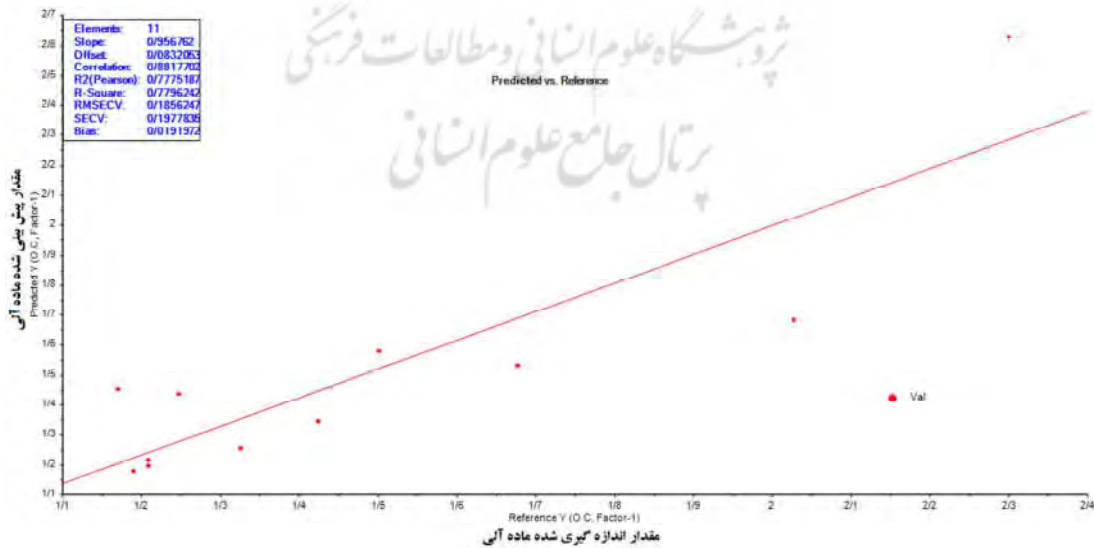
1. Summers et al.
2. Khayamim et al.
3. Aichi et al.

به‌منظور پهنه‌بندی کربن آلی خاک، روش‌های درون‌یابی ارزیابی و روش وزنی معکوس فاصله، با RMSE کمتری، برای پهنه‌بندی انتخاب شد (شکل ۱۱). نتایج نشان می‌دهد که بیش از ۵۰٪ منطقه دارای کربن آلی کمتر از ۲٪ است.

۳-۴- مدل رگرسیون مؤلفه اصلی (PCR)  
مقدار کربن آلی خاک با استفاده از رگرسیون مؤلفه اصلی نیز به‌دست آمد. مقادیر  $R^2$ ، RMSE و RPD در مرحله کالیبراسیون (شکل ۹) و اعتبارسنجی (شکل ۱۰)، برای پیش‌بینی ماده آلی، به‌ترتیب، ۰/۸۶، ۰/۱۹، ۱/۴۷ و ۰/۷۸، ۰/۱۹ و ۲/۰۵ است.



شکل ۹. مقادیر اندازه‌گیری شده در مقابل مقادیر پیش‌بینی شده ماده آلی، در داده‌های کالیبراسیون (مدل PCR)



شکل ۱۰. مقادیر اندازه‌گیری شده در مقابل مقادیر پیش‌بینی شده ماده آلی، در داده‌های اعتبارسنجی (مدل PCR)

جدول ۳. مقایسه روش‌های متفاوت پیش‌پردازش

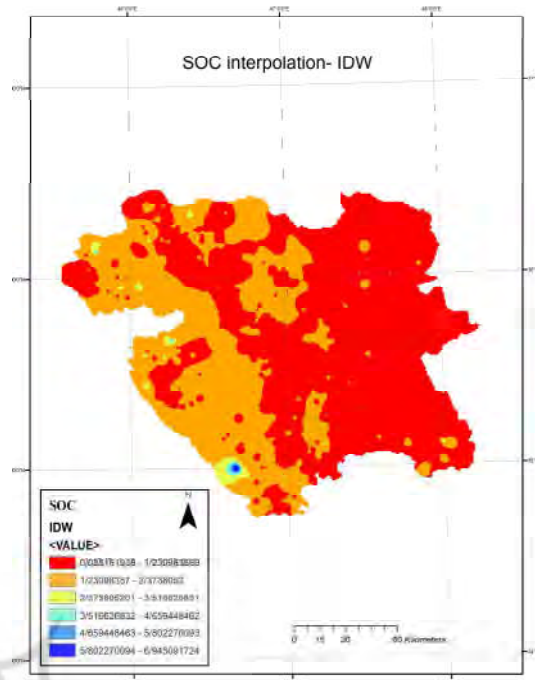
RMSE	R	روش پیش‌پردازش method of preprocessing
0.086	0.83	بدون پیش‌پردازش (without preprocessing)
0.089	0.92	نرمال‌سازی براساس میانگین (moving average)
0.052	0.97	فیلتر ساویتزکی-گلای (Savitzky-Golay filter)
0.089	0.92	فیلتر میانه (median filter)
0.088	0.92	فیلتر گوسین (Gaussian filter)

ماده آلی، به ترتیب، ۰/۸۵ و ۰/۱۴ و ۲/۷۸ است در حالی که نتایج مدل‌سازی PCR برای پارامترهای آماری بالا، به ترتیب، ۰/۷۸ و ۰/۱۹ و ۲/۰۵ بوده است و نشان از دقت بالاتر PLSR، در مقایسه با PCR، برای مدل‌سازی برآورد کربن آلی خاک دارد. از این رو پیش‌بینی می‌شود، در شرایط اقلیمی و خاکی مناطق خشک و نیمه‌خشک، مدل PLSR کارایی بیشتری برای پیش‌بینی سریع کربن آلی خاک داشته باشد. پیش‌بینی مدل‌ها برای درصد کربن آلی کاملاً مناسب بود؛ روش طیف‌سنجی مرئی-مادون قرمز نزدیک قابلیت اندازه‌گیری هم‌زمان چند ویژگی خاک را دارد و دقت مدل‌سازی نیز، تا حد بسیاری، قابل قبول است؛ بنابراین پیشنهاد می‌شود، در مطالعات بعدی، این روش برای برآورد پارامترهای زودیافت خاک، به صورت هم‌زمان به کار رود.

#### ۵- منابع

عباسی، م.، ۱۳۸۸، بررسی مشخصه‌های طیفی برگ‌گونه‌های راش، ممرز، توسکا، بلوط و انجیلی با استفاده از طیف‌سنج زمینی، پایان‌نامه دکتری، جنگلداری، دانشکده منابع طبیعی دانشگاه تهران، ص ۱۱۲.

ملاح نوکنده، س.، همایی، م.، نوروزی، ع.ا.، ۱۹۹۳، بررسی امکان پذیری برآورد مواد آلی خاک با استفاده از تصاویر ابرطیفی هایپرسپکترومتری مناطق ایوانکی و ارومیه، نشریه علمی- پژوهشی مهندسی و مدیریت آبخیز. جلد ششم، شماره ۳، صص. ۲۰۰-۱۹۰.



شکل ۱۱. نقشه درون‌یابی کربن آلی خاک منطقه مطالعاتی

#### ۳-۵- تأثیر روش‌های متفاوت پیش‌پردازش در پیش‌بینی مدل

در مورد ماده آلی، بهترین نتایج مدل‌سازی در روش فیلتر ساویتزکی-گلای و ضعیف‌ترین نتایج در حالت به کار نبردن پیش‌پردازش است. نتایج پیش‌بینی کربن آلی خاک، در این پژوهش، طبق روش پیش‌پردازش فیلتر ساویتزکی-گلای حاصل شده که دقت برآورد مدل‌های پیش‌بینی کربن آلی را افزایش داده است. نتایج خیامیم و همکاران (۲۰۱۵) نیز نشان داد که پیش‌پردازش ساویتزکی-گلای برای برآورد کربن آلی خاک با مدل PLSR، دارای  $R^2$  و RMSE، به ترتیب، برابر با ۰/۶۸ و ۰/۲۶ بوده و مؤید کارایی این روش پیش‌پردازش در افزایش دقت مدل است.

#### ۴- نتیجه‌گیری

نتایج این پژوهش نشان داد که فیلتر ساویتزکی-گلای قوی‌ترین روش پیش‌پردازش داده‌های طیفی بوده است. طبق نتایج مدل‌سازی به روش PLSR، مقادیر  $R^2$  و RMSE و RPD در مرحله اعتبارسنجی برای پیش‌بینی

- Aichi, H., Fouad, Y., Walter, C., Viscarra Rossel, R.A., Chabaane, Z.L. & Sanaa, M., 2009, **Regional Predictions of Soil Organic Carbon Content from Spectral Reflectance Measurements**, *Biosys. Eng.*, 104, PP. 442–446.
- Ben-Dor, E. & Banin, A., 1995, **Near-Infrared Analysis as a Rapid Method to Simultaneously Evaluate Several Soil Properties**, *Soil Sci. Soc. Am. J.*, 59, PP. 364–372.
- Blackmer, A.M. & White, S.E., 1998, **Using Precision Farming Technology to Improve Management of Soil and Fertilizer Nitrogen**, *Australian Journal of Agricultural Research*, 49, PP. 555–564.
- Bouma, J., 2002, **Land Quality Indicators of Sustainable Land Management across Scales**, *Agriculture, Ecosystems and Environment*, 88 (2), PP. 129–136.
- Brunet, D., Barthes, B.G., Chotte, J.L. & 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, PP. 106–117.
- Cambou, A., Cardinael, R., Kouakoua, E., Villeneuve, M., Durand, C. & Barthès, B.G., 2015, **Prediction of Soil Organic Carbon Stock Using Visible and Near Infrared Reflectance Spectroscopy (VNIRS) in the Field**, *Geoderma*, 261, PP. 151–159.
- Chang, C.W., Laird, D.A., Mausbach, M.J. & Hurburgh, Jr. C.R., 2001, **Near-Infrared Reflectance Spectroscopy-Principal Components Regression Analyses of Soil Properties**, *Soil Science Society of America Journal*, 65 (2), PP. 480–490.
- Christy, C.D., 2008, **Real-Time Measurement of Soil Attributes Using on-the-go near Infrared Reflectance Spectroscopy**, *Computers and Electronics in Agriculture*, 61, PP. 10–19.
- Clark, R.N., 1999, **Spectroscopy of Rocks and Minerals, and Principles of Spectroscopy**, In: Rencz, A.N. (Ed.), *Remote Sensing for Earth Sciences, Manual of Remote Sensing*, John Wiley and Sons Inc., Toronto, PP. 3–58.
- Cohen, M.J., Prenger, J.P. & DeBusk, W.F., 2005, **Visible-Near Infrared Reflectance Spectroscopy for Rapid, Non-Destructive Assessment of Wetland Soil Quality**, *Journal of Environmental Quality*, 34, PP. 1422–1434.
- Conforti, M., Castrignanò, A., Robustelli, G., Scarciglia, F., Stelluti, M., Buttafuoco, G., 2015, **Laboratory-Based Vis-NIR Spectroscopy and Partial Least Square Regression with Spatially Correlated Errors for Predicting Spatial Variation of Soil Organic Matter Content**, *Catena*, 124, PP. 60–67.
- Conforti, M., Froio, R., Matteucci, G., Caloiero, T., Buttafuoco, G., 2013, **Potentiality of Laboratory Visible and near Infrared Spectroscopy for Determining Clay Content in Forest Soils: A Case Study from High Forest Beech (Fagus Sylvatica) in Calabria (southern Italy)**, *EQA Int. J. Environ. Qual.*, 11, PP. 49–64.
- Cozzolino, D. & Moron, A., 2003, **The Potential of Near-Infrared Reflectance Spectroscopy to Analyse Soil Chemical and Physical Characteristics**, *Journal of Agricultural Sciences*, 140, PP. 65–71.
- Esbensen, K.H., 2006, **Multivariate Data Analysis**, CAMO Software AS. 5<sup>th</sup> Edition, 589 Pages.
- Farifteh, J., Van Der Meer, F., Atzberger, C. & Carranza, E.J.M., 2007, **Quantitative Analysis of Salt-Affected Soil Reflectance Spectra: A Comparison of Two Adaptive Methods (PLSR and ANN)**, *Remote Sens. Environ.*, 110, PP. 59–78.
- Fideñcio, P.H., Poppi, R.J., Andrade, J.C. & Cantarella, H., 2002, **Determination of Organic Matter in Soil Using Near-Infrared Spectroscopy and Partial Least Squares Regression**, *Communications in Soil Science and Plant Analysis*, 33, PP. 1607–1615.
- Hartemink, A., & McSweeney, K. (Eds). (2014), **Soil Carbon**. Springer Science & Business Medien op.3-16.

- Guerrero, C., Viscarra Rossel, R.A. & Mouazen, A. M., 2010, **Diffuse Reflectance Spectroscopy in Soil Science and Land Resource Assessment**, Geoderma, 158, PP. 1–2.
- IPCC (**International Panel on Climate Change**), 2007. Climate change 2007: synthesis report. In: Writing Team, Core, Pachauri, R.K., Reisinger, A. (Eds.), Contribution of Working Groups I, II and III to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change. IPCC, Geneva.
- Khayamim F, Khademi H, Stenberg B, Wetterlind J. **Capability of vis-NIR Spectroscopy to Predict Selected Chemical Soil Properties in Isfahan Province. JWSS - Isfahan University of Technology**. 2015; 19 (72):81–92
- Lee, K.S., Lee, D.H., Sudduth, K.A., Chung, S.O. & Drummond, S.T., 2007, **Wavelength Identification for Reflectance Estimation of Surface and Subsurface Soil Properties**, ASABE Annual International Meeting, St. Joseph, Michigan.
- Mabit, L. & Bernard, C., 2009, **Spatial Distribution and Content of Soil Organic Matter in an Agricultural Field in Eastern CANADA, as Estimated from Geostatistical Tools**, Earth Surf. Proc. Land, 35, PP. 278–283.
- McBratney, A.B., Stockmann, U., Angers, D., Minasny, B. & Field, D., 2014, **Challenges for Soil Organic Carbon Research**, Insoil Carbon (pp.3-16).
- McBratney, A.B., Minasny, B. & Viscarra Rossel, R.A., 2006, **Spectral Soil Analysis and Inference Systems: A Powerful Combination for Solving the Soil Data Crisis**, Geoderma, 136, PP. 272–278.
- Nelson, D.W. & Sommers, L.E., 1982, **Total Carbon, Organic Carbon and Organic Matter**, In: **Methods of Soil Analysis**, A.L. Page et al. (Eds.), Part 2, 2nd ed., Agron, Monogr, 9, ASA and SSSA Madison, WI., PP. 101–129.
- Nieder, R. & Benbi, D.K., 2008, **Carbon and Nitrogen in the Terrestrial Environment**, Springer.
- Nocita, M., Stevens, A., Noon, C. & vanWesemael, B., 2013, **Prediction of Soil Organic Carbon for Different Levels of Soil Moisture Using Vis-NIR Spectroscopy**, Geoderma, 199, PP. 37–42.
- Reeves III, J.B., McCarty, G.W. & Mimmo, T., 2002, **The Potential of Diffuse Reflectance Spectroscopy for the Determination of Carbon Inventories in Soils**, Environ. Pollut. 116, PP. S277–S284.
- Saeyns, W., Mouazen, A.M. & Ramon, H., 2005, **Potential for Onsite and Online Analysis of Pig Manure Using Visible and Near Infrared Reflectance Spectroscopy**, Biosystems Engineering, 91, PP. 393–402.
- Santra, P., Sahoo, R.N., Das, B.S., Samal, R.N., Pattanaik, A.K. & Gupta, V.K., 2009, **Estimation of Soil Hydraulic Properties Using Proximal Spectral Reflectance in Visible, Near-Infrared, and Shortwave-Infrared (VIS–NIR–SWIR) Region**, Geoderma, 152, PP. 338–349.
- Sarkhot, D.V., Grunwald, S., Ge, Y. & Morgan, C.L.S., 2011, **Comparison and Detection of Total and Available Soil Carbon Fractions Using Visible/Near Infrared Diffuse Reflectance Spectroscopy**, Geoderma, 164, PP. 23–32.
- Savitzky, A. & Golay, M.J., 1964, **Smoothing and Differentiation of Data by Simplified Least Squares Procedures**, Anal. Chem., 36 (8), PP. 1627–1639.
- Shepherd, K.D. & Walsh, M.G., 2007, **Infrared Spectroscopy Enabling an Evidence Based Diagnostic Surveillance Approach to Agricultural and Environmental Management in Developing Countries**, Journal of Near Infrared Spectroscopy, 15 (1), PP. 1–19.
- Shepherd, K.D. & Walsh, M.G., 2002, **Development of Reflectance Spectral Libraries for Characterization of Soil Properties**, Soil Sci. Soc. Am. J., 66, PP. 988–998.
- Stenberg, B., Viscarra Rossel, R.A., Mouazen, A.M. & Wetterlind, J., 2010, **Visible and near infrared spectroscopy in soil science**, Adv. Agron., 107, PP. 163–215.



- Soon, Y.K. & Hendershot, W.H., 2009, In Carter, M.R., Gregorich, E.G. (eds.), **Soil Sampling and Methods of Analysis**, part III: Soil Chemical Analyses, CRC Press Taylor & Francis Group, USA, PP: 173–178.
- Stevens, A., Wesemael, B.V., Bartholomeus, H.M., Rossilon, D., Tychon, B. & Ben-Dor, E., 2007, **CARBIS Final Report: Detecting Soil Carbon and Its Spatial Variability by Imaging Spectroscopy**, Belspo, Louvain-la-Neuve, P. 14.
- Summers, D., Lewis, M., Ostendorf, B. & Chittleborough, D., 2011, **Visible Near-Infrared Reflectance Spectroscopy As a Predictive Indicator of Soil Properties**, *Ecol. Indic.*, 11, PP. 123–131.
- Summers, D., Lewis, M., Ostendorf, B. Chittleborough, D. J., & In, S.S.C. 2009, **Spectral discrimination of Soil Properties**. Discriminating and mapping soil variability with hyper spectral reflectance data, 34.
- Viscarra Rossel, R.A. & Behrens, T., 2010, **Using Data Mining to Model and Interpret Soil Diffuse Reflectance Spectra**, *Geoderma*, 158, PP. 46–54.
- Viscarra Rossel, R.A., Walvoort, D.J.J., McBratney, A.B., Janik, L.J. & Skjemstad, J.O., 2006, **Visible, Near Infrared, Mid Infrared or Combined Diffuse Reflectance Spectroscopy for Simultaneous Assessment of Various Soil Properties**, *Geoderma*, 131, PP. 59–75.
- Walkly, A. & Black, I.A., 1934, **An Examination of Digestion Method for Determining Soil Organic Matter and a Proposed Modification of the Chromic Acid Titration**, *Soil Science Society of American Journal*, 37, PP. 29–38.
- Walvoort, D.J.J. & McBratney, A.B., 2001, **Diffuse Reflectance Spectrometry As a Proximal Sensing Tool for Precision Agriculture**, In: Grenier, G., Blackmore, S. (Eds.), *ECPA 2001, Third European Conference on Precision Agriculture*, Vol. 1., Agro Montpellier, PP. 503–507.
- Williams, P.C., 2001, **Implementation of Near-Infrared Technology**, PP. 145–169, In: Williams, P., Norris, K. (Eds.), **Near-infrared Technology in the Agricultural and Food Industries**, American Association of Cereal Chemists Inc., St. Paul, MN.
- Zornoza, R., Guerrero, C., Mataix-Solera, J., Scow, K.M., Arcenegui, V. & Mataix-Beneyto, J., 2008, **Near Infrared Spectroscopy for Determination of Various Physical, Chemical and Biochemical Properties in Mediterranean Soils**, *Soil Biology & Biochemistry*, 40, PP. 1923–1930.



## Estimation Soil Organic Matter (SOM) Content Using Visible and Near Infrared Spectral data, PLSR and PCR Statistical Models

Matinfar, H.<sup>1\*</sup>, Mahmoodzadeh, H.<sup>2</sup>, Fariabi, A.<sup>3</sup>

1. Associate Prof. of Soil Science, Remote Sensing and GIS

2. Ph.D. Student of Soil Science, Lorestan University, Soil Science, Genesis and Classification

3. Ph.D. Student of Soil Science Department, Lorestan University, Soil Science, Genesis and Classification

### Abstract

Soil organic matter is one of the most important Physical and chemical properties of soil that it is critical in determining the quality and management of soils. Quantify of soil organic carbon due to the high spatial variability and changes over time is difficult. Near-infrared-visible spectroscopy is a feasible method to reduce the time and cost to check the organic carbon. The aim of this study was to evaluate soil organic carbon through near-infrared and visible spectroscopy with the statistical models PLSR and PCR. For this purpose, 40 soil samples from depths of 0 to 30 cm were collected by systematic random method based on previous studies and determination of different classes of soils in the region. Chemical analysis of soils was performed according to standard methods. Spectral reflectance of soil samples in the range of 350 to 2500 nm was measured then after applying the preprocessing methods such as Savitzky and Golay filter, Soil organic carbon were calculated by principal component analysis (PCA), regression partial least squares (PLSR) and principal component regression (PCR) models. The results of this study showed that the Savitzky and Golay filter was the strongest preprocessing method for spectral data. Coefficients of determination ( $R^2$ ), root mean square error of Prediction (RMSE) and ratio of prediction to deviation (RPD) in the calibration and validation to predict organic matter, respectively, 0.97, 0.05, 5.09 and 0.85, 0.14, 2.78 respectively. Therefore, for dry and semi-arid soils of the PLSR model, it is more efficient to predict the organic carbon of the soil. The results showed that the PLSR model has better performance than the PCR model in soil organic carbon estimation.

**Keywords:** organic matter, spectroscopy, PLSR, PCR, principal component analysis.