

مقایسه روش‌های آماری حداقل مربعات جزئی و رگرسیون مولفه اصلی در برآورد مقادیر گچ، کربنات‌ها و ماده آلی خاک با استفاده از طیف‌سنجی مرئی و مادون قرمز نزدیک

- الهام مهرابی گوهری^۱، حمیدرضا متین فر^{۲*}، اعظم جعفری^۳، روح الله تقی زاده مهرجردی^۴، فاطمه خیامیم^۵
۱. دانشجوی دکتری، گروه علوم خاک، دانشکده کشاورزی، دانشگاه لرستان و مربی گروه کشاورزی، دانشگاه پیام نور
 ۲. دانشیار، گروه علوم خاک، دانشکده کشاورزی، دانشگاه لرستان
 ۳. استادیار، گروه علوم خاک، دانشکده کشاورزی، دانشگاه باهنر
 ۴. استادیار، گروه علوم خاک، دانشکده کشاورزی و منابع طبیعی، دانشگاه اردکان
 ۵. دانش آموخته دکتری، گروه خاکشناسی، دانشکده کشاورزی، دانشگاه صنعتی اصفهان

تاریخ پذیرش مقاله: ۱۳۹۸/۰۷/۱۶

تاریخ دریافت مقاله: ۱۳۹۷/۱۰/۲

چکیده

داده‌های طیفی محدوده مرئی - مادون قرمز که با حداقل هزینه و صرف وقت تهیه می‌شوند، کاربرد وسیعی در برآورد خصوصیات فیزیکی و شیمیایی خاک دارند. مطالعه حاضر، با هدف بررسی توانایی این روش در برآورد مقدار ماده آلی، کربنات‌ها و درصد گچ سطح خاک صورت پذیرفته است. بر اساس تکنیک هایپرکیوب، محل ۱۱۵ پروبیل شناسایی و سپس نمونه‌برداری از افق‌های خاک انجام گرفت و مقدار ماده آلی، درصد گچ و آهک خاک نیز با روش‌های استاندارد اندازه‌گیری شد. آنالیز طیفی خاک افق‌های سطحی مورد نظر با استفاده از دستگاه طیف‌سنج زمینی با دامنه طول موج ۴۰۰-۲۵۰۰ نانومتر انجام شد (۱۱۵ نمونه سطحی). پس از ثبت طیف‌ها، انواع روش‌های پیش‌پردازش، مورد ارزیابی قرار گرفت و سپس از رگرسیون حداقل مربعات جزئی (Partial Least Square Regression) و رگرسیون مولفه اصلی (PCR)، برای پیش‌بینی پارامترهای مورد نظر استفاده شد. جهت ارزیابی مدل، ۸۰ درصد داده‌ها برای کالیبراسیون مدل و ۲۰ درصد برای صحت‌سنجی مدل، به صورت تصادفی انتخاب شدند. پیش‌پردازش‌های مختلف از جمله مشتق اول و دوم، به همراه فیلتر ساواتزکی و گلای، متغیر نرمال استاندارد و تصحیح پخشیده چندگانه با هم مقایسه شدند. نتایج این بررسی نشان داد بهترین نتایج مدل‌سازی، مربوط به روش PLSR با پیش‌پردازش مشتق اول + فیلتر ساواتزکی و گلای برای برآورد گچ، کربنات و ماده آلی خاک است. با توجه به مقادیر درصد انحراف نسبی (RPD)، پیش‌بینی مدل، برای درصد گچ و ماده آلی در کلاس خوب و برای کربنات‌ها در کلاس ضعیف است.

کلید واژه‌ها: رگرسیون حداقل مربعات جزئی، رگرسیون مولفه اصلی، طیف‌سنجی، کربنات، گچ، ماده آلی

*نویسنده عهده‌دار مکاتبات: خرم آباد، دانشگاه لرستان- دانشکده کشاورزی، گروه علوم و مهندسی خاک.

۱- مقدمه

(Curcio et al., 2013; Gomez et al., 2008) با استفاده از طیفسنجی در محدوده مرئی و مادون قرمز نزدیک برآورد شده است. برای مدل سازی روابط بین پارامترهای طیفی و خصوصیات فیزیکوشیمیایی خاک‌ها، از روش‌های مختلفی استفاده می‌شود (Schwartz et al., 2011) که از آن جمله می‌توان رگرسیون مولفه اصلی (Martens and Næs., 1989) و رگرسیون حداقل مربعات جزئی (Wold et al., 1983) را نام برد. Summers et al (2011) با بررسی ۳۰۰ نمونه خاک در منطقه جنوب استرالیا به روش طیفسنجی مرئی مادون قرمز نزدیک، در برآورد کربن آلی، مقدار رس، اکسیدهای آهن و کربنات‌های خاک پرداختند. آنها نتیجه گرفتند که ضریب تبیین برای درصد رس، کربن آلی، اکسید آهن و کربنات‌های خاک بالاتر از ۰/۵ است و مقادیر درصد انحراف نسبی (RPD) قابل قبول گزارش شده است. (Khayamim et al (2015) نیز نتیجه گرفتند که مقادیر ضریب تبیین در گروه اعتبارسنجی برای ماده آلی، کربنات‌ها و گچ به ترتیب ۰/۶۱، ۰/۴۵ و ۰/۸ است و نشان داد که پیش بینی مدل، برای گچ کاملاً مناسب و برای ماده آلی قابل قبول است، درحالی‌که برای کربنات‌های خاک ضعیف است. Shiferaw and Hergarten (2014) داده‌های طیفی در محدوده VIS/NIR و روش‌های PCR و PLSR را برای تخمین کربن آلی خاک در اتیوپی به کار بردند. نتایج نشان داد مقادیر R²، RMSE، RPD مربوط به مدل PLSR، نسبت به مدل PCR برتری داشته است. به دلیل اینکه، این اندازه گیری‌ها، در زمان اندازه گیری‌های طیفی و به ویژه در صحرا، تحت تأثیر عوامل ناخواسته مانند رطوبت خاک، زبری سطح خاک، بقایای آلی درشت، آلودگی پروب به علت گرد و غبار، تغییر فاصله سنسور و خاک و تأثیرات پرتو افکنی جوی قرار می‌گیرند. بنابراین در بیشتر مطالعات، از روش‌های پیش‌پردازش مشتق اول و مشتق دوم (Gholizadeh et al., 2015)، تصحیح پخشیده چندگانه (MSC) (Rinnan et al., 2009) و

طیفسنجی مرئی- مادون قرمز نزدیک، یک روش توجیه‌پذیر از نظر زمان و هزینه است که برای بررسی ویژگی‌های خاک دارای قابلیت مناسب است (Cambou et al., 2015)، به‌طوریکه در ۲۰ سال گذشته، تمایل زیادی برای کاربرد طیف مرئی- مادون قرمز نزدیک، در برآورد ویژگی خاک وجود داشته است (Stenberg et al., 2010). طیفسنجی مرئی-مادون قرمز نزدیک، یک روش غیر مخرب، سریع، ارزان، با حداقل آماده‌سازی نمونه، بدون ضرر و تخریب برای محیط زیست، قابلیت برآورد چندین ویژگی خاک را دارد (Shepherd and Walsh., Stenberg et al., 2010) (Viscarrá Rossel and Behrens., 2010; 2002) و قابل استفاده در مزرعه و آزمایشگاه (Cambou et al., 2015) است. از آنجا که خاک نیز، همانند سایر اجسام موجود در طبیعت، بر اساس ترکیب و ساختار مولکولی خود، بازتاب مشخصی در نواحی مختلف امواج الکترومغناطیس و یا طول موج‌های گوناگون نشان می‌دهد، دارای سیمای طیفی مخصوص به خود است (McCoy., 2005). بنابراین می‌توان با کمک طیف‌های بازتابی اطلاعات فیزیکی، کانی‌شناسی و شیمیایی خاک را به دست آورد. برخی از ویژگی‌های خاک، مانند رنگ، بافت، کانی‌شناسی، ماده آلی، کربنات‌های آزاد، شوری، رطوبت و اکسید و هیدروکسیدهای آهن و منگنز در انعکاس بازتابی آن اثر می‌گذارند و سبب شکل‌گیری الگوی سیمای طیفی خاصی می‌شوند (Chen et al., 2014).

بر اساس بررسی‌های انجام شده، امکان استفاده از طیفسنجی به عنوان یک روش سریع، آسان، کم هزینه و پیشرفته برای پیش‌بینی ویژگی‌های خاک وجود دارد (Klement et al., 2013; Schwartz et al., 2011; Araújo et al., 2013). در این رابطه، گزارش‌هایی وجود دارد که نشان می‌دهد میزان کربنات‌های خاک (Ben-Dor and Banin, 1995; Lagacherie et al., 2008)، ماده آلی خاک (Liu et al., 2014)، گچ (Khayamim et al., 2015) و توزیع اندازه ذرات خاک

۲ میلی‌متری عبور داده شده و سپس مقدار ماده آلی خاک در نمونه‌های مورد مطالعه به روش والکی- بلک اندازه‌گیری شد. علاوه بر مقدار ماده آلی، درصد گچ و کربنات‌های خاک نیز به ترتیب با روش‌های آون و تیتراسیون برگشتی، مورد ارزیابی قرار گرفتند.

۲-۲- آنالیز طیفی

به منظور اندازه‌گیری بازتاب طیفی نمونه‌های خاک، از دستگاه اسپکترو رادیومتر زمینی استفاده شد. نمونه‌های خاک، برای کاهش آثار مزاحم رطوبت و تداخل ساختاری خاک در رفتار طیفی نمونه‌ها، هوا خشک و الک شدند و سپس در یک پتری دیش به قطر تقریبی ده سانتی‌متر قرار گرفته و با کاردک، سطح آنها هموار شد و سپس جهت بررسی‌های طیفی، به اتاق تاریک مخصوص اندازه‌گیری طیفی منتقل شدند. هر نمونه، چهار بار (به ازای هر چرخش نمود درجه‌ای متوالی)، برای حذف آثار تغییر در هندسه تابش، مورد سنجش طیفی قرار گرفت. هندسه تابش، قادر است در تابع توزیع بازتاب دو جهته (BRDF)، اثر بگذارد؛ بنابراین لازم است آثار تغییر هندسه تابش، بر بازتاب‌های سنجش شده به کمترین میزان برسد. پیش از آغاز نمونه‌برداری طیفی از گزینه جریان تاریک (DC) برای افزایش نسبت سیگنال به نویز (SNR) و نیز از مرجع سفید اسپکترالون، به طور تقریبی به ازای هر نیم ساعت نمونه‌برداری طیفی برای کالیبره کردن انعکاس طیفی بهره برده شد. بدین طریق، نمونه‌های خاک اسکن شد و بازتاب‌های مطلق در دامنه طیفی ۲۵۰۰-۳۵۰ نانومتر و با قدرت تفکیک طیفی یک نانومتر، سبب تولید ۲۱۵۰ نقطه داده طیفی (SDP) به ازای هر نمونه خاک شد. منحنی‌های به دست آمده بلافاصله و به طور خودکار با استفاده از نرم افزار RS3 موجود بر روی رایانه قابل حمل متصل به دستگاه، میانگین‌گیری شده و به صورت یک منحنی طیفی، به نمایش در آمد.

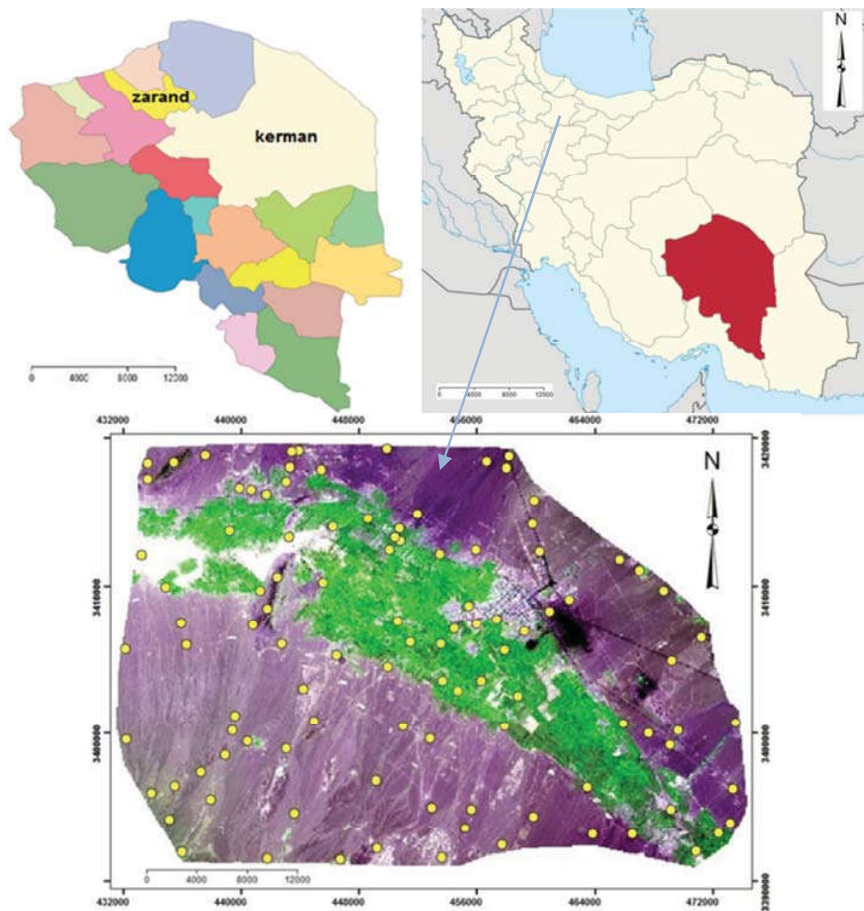
1- FieldSpec®3, ASD, FR, USA

متغیر نرمال استاندارد (SNV) (Barnes et al., 1989) به عنوان تکنیک‌های جدانپذیر در مدل‌سازی طیفی استفاده می‌شود. هدف این پژوهش، بررسی توانایی داده‌های طیفی در محدوده مرئی و مادون قرمز نزدیک در برآورد درصد گچ، کربنات و ماده آلی خاک با استفاده از روش‌های PLSR و PCR است.

۲- مواد و روش‌ها

۲-۱- تشریح محدوده مطالعاتی:

منطقه مورد مطالعه در جنوب شرق ایران، تقریباً در فاصله ۷۰ کیلومتری شهر کرمان، بین عرض جغرافیایی $30^{\circ} 40'$ تا 31° درجه شمالی و طول جغرافیایی $56^{\circ} 50'$ تا 56° درجه شرقی واقع شده است. این منطقه با مساحتی حدود ۱۰۰۰۰۰ هکتار، شهرستان زرند را نیز در بر می‌گیرد. شهرستان زرند، در شمال غرب استان کرمان قرار دارد و از شمال و شمال غرب با استان یزد (شهرستان بافق)، از شمال شرق و شرق با شهرستان راور، از جنوب شرق با شهرستان کرمان و از جنوب غرب و غرب با شهرستان رفسنجان هم‌مرز است. متوسط سالانه بارندگی، درجه حرارت و تبخیر و تعرق به ترتیب ۶۱ میلی‌متر، ۱۷ درجه سانتی‌گراد و ۱۷۵۰ میلی‌متر است. در منطقه مورد مطالعه، بر اساس تکنیک هایپرکیوب، محل ۱۱۵ پروفیل شناسایی و سپس نمونه‌برداری از افق‌ها انجام گرفت. روش هایپرکیوب، یک روش نمونه‌برداری برای نقشه‌برداری رقومی خاک، در مکان‌هایی است که نمونه قبلی خاک وجود ندارد (تنها اطلاعات کمکی وجود دارد). در واقع، یک طرح نمونه‌برداری مونت کارلوی محدود شده است. این روش، یک طرح تصادفی طبقه‌بندی شده‌ایست که باعث نمونه‌برداری موثر به کمک توزیع چند متغیر می‌شود؛ به این شکل که برای هر نقطه مطالعاتی، اطلاعات لازم، اعم از محل پروفیل در زمین نما، نوع واحد ژئومورفیک و نوع مواد مادری یادداشت برداری و از افق‌های هر پروفیل نمونه‌برداری شد. تمامی نمونه‌های خاک، پس از هوا خشک شدن در محیط آزمایشگاه به مدت ۷۲ ساعت، از الک



شکل ۱. موقعیت منطقه مورد مطالعاتی و نحوه پراکنش مکانی داده های خاک

مختص طیف و داده‌های فیزیکو شیمیایی در آن می‌توان نتایج مورد نظر را تخمین زد) انجام شد. سپس از روش‌های PLSR و PCR برای برآورد میزان گچ، کربنات و مواد آلی خاک استفاده شد.

۲-۳- تحلیل آماری_طیفی

PLSR، روشی است که از طریق بهره‌گیری از تکنیک تحلیل مولفه‌های اصلی (PCA) و رگرسیون خطی چندمتغیره (MLR)^۳، بر مشکلات چند هم خطی‌گری

برای هر نمونه، ۵ تکرار ثبت شد. همچنین به منظور انجام پیش‌پردازش طیف‌ها، ابتدا دو بخش نویزی، در ابتدا و انتهای طیف‌ها در محدوده ۳۵۰-۴۰۰ و ۲۴۵۰-۲۵۰۰ نانومتر حذف شد. به علاوه، وقفه حاصل از تغییر دیتکتور در محدوده ۹۰۰ تا ۱۰۰۰ نانومتر نیز حذف شد. انواع روش‌های پیش‌پردازش، شامل مشتق اول، مشتق دوم به همراه فیلتر میانه و فیلتر ساویتزکی^۱ و گسلی^۲، تصحیح پخشیده چندگانه و متغییر نرمال استاندارد بر روی داده‌های طیفی به کمک نرم افزار Unscambler 10.3 (نرم افزاری که از اکسل، اطلاعات طیفی مورد نظر را دریافت می‌کند و با اعمال پیش‌پردازش‌ها و اجرای مدل‌های مختلف آماری،

1. Savitzky
2. Golay
3. Multiple Linear Regression

لازم به توضیح است که برای پیش بینی، داده‌ها به دو دسته آموزش (80 درصد از داده‌های نمونه برداری شده) و آزمون (20 درصد از داده‌های نمونه برداری شده) به طور تصادفی تقسیم شدند.

$$RMSE = \sqrt{\frac{1}{n} \sum_{i=1}^n (O_i - P_i)^2} \quad \text{رابطه (۱)}$$

$$R^2 = \left[\frac{\sum (X_e - \bar{X}_e)(X_o - \bar{X}_o)}{\sqrt{((X_e - \bar{X}_e)^2)(X_o - \bar{X}_o)^2}} \right]^2 \quad \text{رابطه (۲)}$$

$$RPD = SD/RMSE \quad \text{رابطه (۳)}$$

\bar{X} : متغیرهای کمی مشاهده شده، e : متغیرهای کمی تخمین زده، \bar{X}_e : میانگین متغیرهای مشاهده شده، \bar{X}_o : میانگین متغیرهای کمی برآورد شده. SD انحراف معیار مقادیر مشاهداتی و n تعداد داده‌ها است.

۳- نتایج و بحث

۳-۱- خلاصه آماری داده‌ها

جدول ۱، توصیف آماری ویژگی‌های خاک را در دو گروه کالیبراسیون و اعتبارسنجی نشان می‌دهد. میانگین مقادیر ماده آلی، کربنات‌ها و گچ در گروه کالیبراسیون به ترتیب ۰/۳، ۳۱/۲ و ۴/۷ درصد است. این مقادیر برای گروه اعتبارسنجی تقریباً مشابه است. انحراف معیار مقادیر ماده آلی، کربنات‌ها و گچ در گروه کالیبراسیون به ترتیب ۰/۳، ۱۷ و ۵/۴۵ درصد و در گروه اعتبارسنجی ۰/۳، ۱۷/۰۲ و ۵/۴۸ است. این مسئله به روشنی نشان می‌دهد که گروه اعتبارسنجی، نماینده مناسبی از مجموعه داده‌ها است.

و خود همبستگی، بین متغیرهای طیفی و نمونه‌ها، غلبه می‌کند و مجموعه‌ای از اجزا یا بردارهای پنهان را می‌یابد که سبب کاهش و یا تجزیه ماتریس متغیرهای طیفی تخمین گر (x) می‌شود و تا آنجا که امکان دارد سبب همبستگی بین متغیرهای پاسخ (y) با متغیرهای مستقل (x) می‌شود (Summers et al., 2011). در رگرسیون حداقل مربعات جزئی، هر دو ماتریس X و Y به طور همزمان، مدل می‌شوند، برای اینکه بتواند بردارهای پنهان در X را که می‌تواند Y را پیش بینی نماید، پیدا کند. در واقع هدف مدل PLSR، ساختن یک مدل خطی بر اساس رابطه $Y = Xb + E$ است که در آن Y متغیر برآورد شده توسط مدل در شکل ماتریس (متغیر $Y \times 1$) و X متغیر مستقل پیش‌گویی کننده است. همچنین b، ماتریسی شامل ضرایب رگرسیون و E ماتریسی از خطاهای باقیمانده است. رگرسیون مولفه اصلی، ترکیبی از رگرسیون خطی چندگانه (MLR) و تجزیه به مولفه‌های اصلی (PCA) است. تفاوت عمده PLSR با PCR این است که مؤلفه‌های اصلی، تنها با مقادیر متغیرهای X تعیین می‌شود، در صورتیکه در PLSR، مقادیر هر دو متغیرهای X و Y در ساخت مؤلفه‌ها تأثیر می‌گذارند (Camo., 2006).

۲-۴- اعتبارسنجی مدل‌ها

پارامترهایی که در این تحقیق برای اعتبارسنجی مدل استفاده شد: (۱) ریشه میانگین مربعات خطای پیش‌بینی است که این مقدار از رابطه RMSE به دست می‌آید. (۲) ضریب تبیین R^2 ، بین مقادیر اندازه‌گیری شده توسط روش‌های آزمایشگاهی و پیش‌بینی شده توسط مدل است (۳) انحراف پیش‌بینی باقیمانده (RPD) دقت و صحت مدل‌ها، مورد ارزیابی قرار گرفتند.

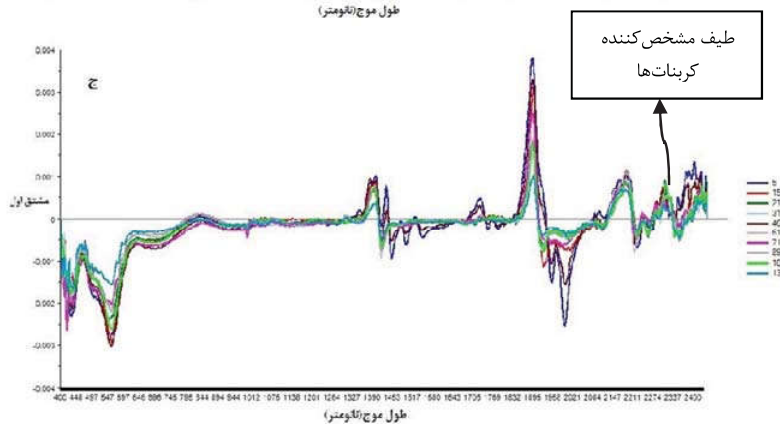
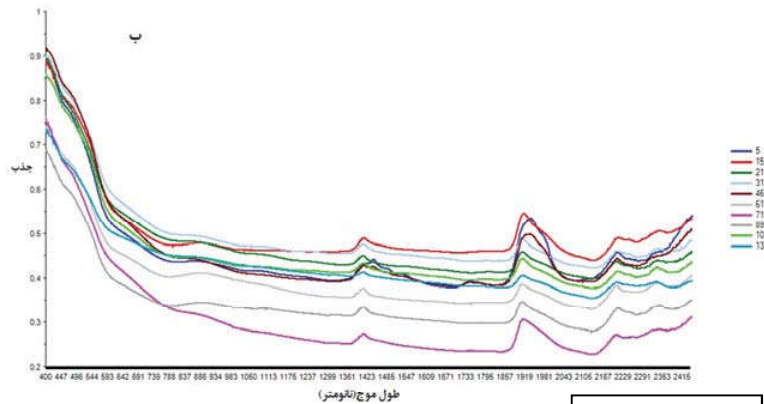
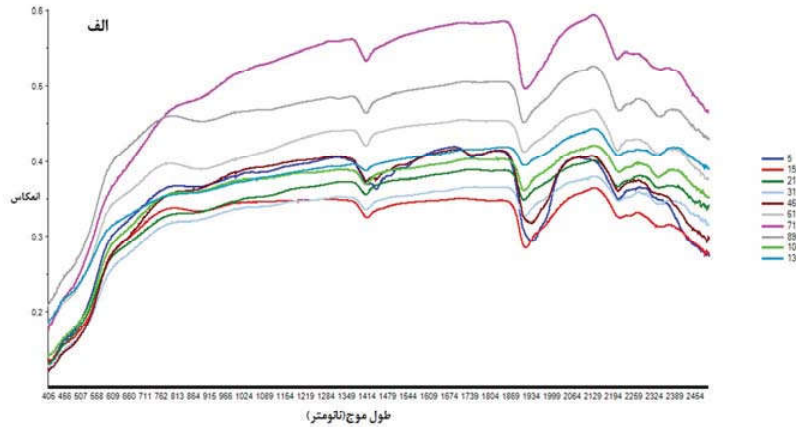
جدول ۱. خلاصه آماری خصوصیات شیمیایی خاک در منطقه زرد کرمان

اعتبارسنجی			کالیبراسیون			
ماده آلی	کربنات‌ها	گچ	ماده آلی	کربنات‌ها	گچ	
۰/۲۸	۲۷	۵/۳	۰/۳	۳۱/۲	۴/۷	میانگین
۱/۸	۶۵	۴۶	۱/۸	۷۶	۵۸	حداکثر
۰/۰۳	۴/۷	۲/۴	۰/۰۲	۵/۶	۳/۱	حداقل
۰/۳	۱۷/۰۲	۵/۴۸	۰/۳	۱۷/۰۰	۵/۴۵	انحراف معیار

۲-۳- ویژگی‌های طیف‌های خاک

طیف‌ها در همه خاک‌ها تقریباً مشابه است، به طوری‌که ویژگی‌های جذبی آب هیگروسکوپی و آب موجود در شبکه رس‌ها در طول موج‌های ۱۴۰۰، ۱۹۰۰ و ۲۲۰۷ نانومتر در همه طیف‌ها مشاهده می‌شود، البته

شکل ۲_الف، طیف انعکاسی محدوده مرئی - مادون قرمز نزدیک اندازه‌گیری شده در ده نمونه خاک، به صورت تصادفی را نشان می‌دهد. شکل کلی همه



شکل ۲. مقادیر انعکاس الف، جذب و ج) مشتق اول + فیلتر ساوانزکی-گلای، طیف‌های محدوده مرئی - مادون قرمز نزدیک در ده نمونه تصادفی خاک.

۰/۵۴ تا ۰/۶۷ در گروه کالیبراسیون و از ۰/۵۱ تا ۰/۶۵ در گروه اعتبارسنجی متغیر است. مقدار RMSE، نیز در دو گروه کالیبراسیون و اعتبارسنجی، به ترتیب از ۳/۴۶ تا ۴/۵۵ متغیر است. در پیش‌بینی درصد گچ توسط PLSR، بیشترین مقدار R2 و کمترین مقدار RMSE مربوط به مشتق اول + فیلتر ساویتزکی و گلای در هر دو گروه کالیبراسیون و اعتبارسنجی حاصل شد (جدول ۲). علاوه بر اینها، دقت مدل پیش‌بینی شده، توسط درصد انحراف نسبی نیز، مورد ارزیابی قرار می‌گیرد که از معیارهای سنجش مدل در مطالعات طیف‌سنجی است که به ۶ کلاس خیلی ضعیف ($RPD < 1$)، ضعیف (۱-۱/۴)، متوسط (۱/۴-۱/۸)، خوب (۱/۸-۲)، خیلی خوب (۲-۲/۵) و عالی ($RPD > 2.5$) طبقه بندی می‌شوند (Camo., 2006; McDowell et al., 2012). بیشترین مقادیر RPD برای داده‌های کالیبراسیون ۱/۸۴ و برای داده‌های اعتبارسنجی، برابر با ۱/۸۳ محاسبه شد. مقادیر درصد انحراف نسبی، بر اساس طبقه‌بندی (Chang et al., 2010; Mouazen et al., 2001)؛ برای درصد گچ، در کلاس خوب قرار می‌گیرد. در روش PCR، بهترین نتایج RPD، برای مشتق اول + فیلتر ساویتزکی و گلای حاصل شد که در کلاس ضعیف قرار می‌گیرد و روش‌های متغیر نرمال استاندارد و تصحیح پخشیده چندگانه، قادر به ایجاد ارتباط بین درصد گچ و داده‌های طیفی نبودند (جدول ۲).

اختلافاتی در میزان شدت انعکاس وجود دارد. ویژگی جذبی رس در ۲۲۰۰ نانومتر است. 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، چارچوب سیلیکات و کاتیون‌های اکتاهدرال، تتراهدرال و بین لایه ای است. اکسیدهای آهن در محدوده مرئی و کربنات‌ها در ۲۳۰۰ نانومتر قابل مشاهده هستند (Farmer and Russell., 1964). مقادیر انعکاس از طریق رابطه (۴) به جذب تبدیل شد. در رابطه (۴) R و A به ترتیب مقادیر جذب و انعکاس بر حسب درصد را نشان می‌دهند. شکل ۲ - نمودارهای جذب طیف مرئی - مادون قرمز ده نمونه خاک را نشان می‌دهد.

$$A = \log(1/R) \quad (4)$$

۳-۳- پیش‌بینی ویژگی‌های خاک

جدول‌های ۲ تا ۴، مقادیر ضرایب تبیین (R^2)، ریشه حداقل مربعات خطا (RMSE) و درصد انحراف نسبی (RPD) حاصل از مدل‌سازی توسط رگرسیون حداقل مربعات جزئی (PLSR) و رگرسیون مولفه اصلی (PCR) را برای سه پارامتر ماده آلی، کربنات‌ها و گچ خاک‌های زرد استان کرمان را نشان می‌دهد. در روش PLSR، نتایج پیش‌بینی درصد گچ نشان داد که مقادیر R2 از

جدول ۲. نتایج آماری روش‌های PLSR و PCR و روش‌های مختلف پیش پردازش طیفی در تعیین درصد گچ خاک

مدل	پیش پردازش	R2	RMSE	RPD	R2	RMSE	RPD
		کالیبراسیون	کالیبراسیون	کالیبراسیون	اعتبارسنجی	اعتبارسنجی	اعتبارسنجی
PCR	جذب بدون پیش پردازش	-	-	-	-	-	-
	فیلتر ساویتزکی و گلای	۰/۲۷	۶/۲۸	۰/۸۷	۰/۲۲	۶/۳۱	۰/۸۶
	متغیر نرمال استاندارد	-	-	-	-	-	-
	تصحیح پخشیده چندگانه	-	-	-	-	-	-
PLSR	مشتق اول + فیلتر ساویتزکی و گلای	۰/۳۱	۵/۲۳	۱/۰۴	۰/۲۹	۵/۱۶	۱/۰۶
	مشتق دوم + فیلتر ساویتزکی و گلای	۰/۳۰	۵/۲۵	۰/۰۴	۰/۲۷	۵/۱۹	۱/۰۵
	جذب بدون پیش پردازش	۰/۶۳	۲/۹۸	۱/۸۳	۰/۶۱	۳/۰۳	۱/۸۱
	فیلتر ساویتزکی و گلای	۰/۶۶	۲/۹۸	۱/۸۳	۰/۶۳	۳/۰۳	۱/۸۱
	متغیر نرمال استاندارد	۰/۵۹	۳/۱۸	۱/۷۱	۰/۵۶	۳/۲۶	۱/۶۸
	تصحیح پخشیده چندگانه	۰/۵۴	۳/۱	۱/۷۶	۰/۵۱	۳/۲۲	۱/۷
	مشتق اول + فیلتر ساویتزکی و گلای	۰/۶۷	۲/۹۶	۱/۸۴	۰/۶۵	۲/۹۹	۱/۸۳
	مشتق دوم + فیلتر ساویتزکی و گلای	۰/۶۱	۲/۹۹	۱/۸۲	۰/۵۷	۳/۰۴	۱/۸

در برآورد درصد ماده آلی خاک با روش PLSR، مقدار R2 در دو گروه کالیبراسیون و اعتبارسنجی به ترتیب از ۰/۵۵ تا ۰/۷۱ و ۰/۵۲ تا ۰/۶۹ متغیر است (جدول ۴). بیشترین مقدار R2 و کمترین مقدار RMSE، در روش مشتق اول + فیلتر ساویتزکی و گلای در دو گروه کالیبراسیون و اعتبارسنجی به دست آمد. مقادیر RPD، برای گروه کالیبراسیون برابر با ۱/۷۷ و برای گروه اعتبارسنجی برابر با ۱/۶۷ به دست آمد (جدول ۴) که بر اساس طبقه‌بندی (Chang et al 2010)؛ (2001)؛ al (2001)؛ هر دو در کلاس خوب قرار می‌گیرند. در روش PCR، بهترین نتایج RPD برای مشتق اول + فیلتر ساویتزکی و گلای حاصل شد که در کلاس ضعیف قرار می‌گیرد و روش‌های متغیر نرمال استاندارد و تصحیح پخشیده چندگانه، قادر به ایجاد ارتباط بین درصد گچ و داده‌های طیفی نبودند (جدول ۴).

در پیش‌بینی درصد کربنات خاک با روش PLSR، مقدار R2 در دو گروه کالیبراسیون و اعتبارسنجی، به ترتیب از ۰/۵۱ تا ۰/۵۷ و از ۰/۴۸ تا ۰/۵۶ متغیر است (جدول ۳). کمترین مقدار RMSE و بیشترین مقدار R2 در روش مشتق اول + فیلتر ساویتزکی و گلای در دو گروه کالیبراسیون و اعتبارسنجی به دست آمد. مقادیر RPD، برای گروه کالیبراسیون، برابر با ۱/۵۵ و برای گروه اعتبارسنجی برابر با ۱/۳۶ به دست آمد (جدول ۳) که بر اساس طبقه‌بندی (Mouazen et al 2010)؛ (2001)؛ Chang et al (2001)؛ به ترتیب در کلاس خوب و متوسط قرار می‌گیرند. در روش PCR، پیش‌پردازش تصحیح پخشیده چندگانه، بیشترین مقدار R2 و کمترین مقدار RMSE را برای داده‌های کالیبراسیون و اعتبارسنجی نشان داد که کارایی مدل در هر دو حالت ضعیف است (جدول ۳).

جدول ۳. نتایج آماری روش‌های PLSR و PCR و روش‌های مختلف پیش‌پردازش طیفی در تعیین درصد کربنات خاک

مدل	پیش پردازش	R2	RMSE	RPD	R2	RMSE	RPD
		کالیبراسیون	کالیبراسیون	کالیبراسیون	اعتبارسنجی	اعتبارسنجی	اعتبارسنجی
	جذب بدون پیش پردازش	۰/۲۶	۱۷/۲۵	۰/۹۷	۰/۲۴	۱۸/۲	۰/۹۳
	فیلتر ساویتزکی و گلای	۰/۳	۱۶/۳۶	۱/۰۴	۰/۲۳	۱۵/۲۷	۱/۱۰
PCR	متغیر نرمال استاندارد	۰/۲۳	۱۵/۴۶	۱/۰۹	۰/۲۲	۱۵/۷	۱/۰۸
	تصحیح پخشیده چندگانه	۰/۳۵	۱۲/۵۰	۱/۳۶	۰/۳۲	۱۲/۵۹	۱/۳۵
	مشتق اول + فیلتر ساویتزکی و گلای	۰/۲۷	۱۵/۲۳	۱/۱۲	۰/۲۴	۱۵/۱۷	۱/۱۲
	مشتق دوم + فیلتر ساویتزکی و گلای	۰/۲۵	۱۵/۳۰	۱/۱۱	۰/۲۴	۱۵/۱۴	۱/۱۲
	جذب بدون پیش پردازش	۰/۵۳	۱۱/۴۵	۱/۴۸	۰/۵۰	۱۲/۵۹	۱/۳۵
	فیلتر ساویتزکی و گلای	۰/۵۵	۱۱/۳۰	۱/۵	۰/۵۳	۱۲/۵۶	۱/۳۵
PLSR	متغیر نرمال استاندارد	۰/۵۱	۱۲/۵۰	۱/۳۶	۰/۴۹	۱۳/۵۵	۱/۲۵
	تصحیح پخشیده چندگانه	۰/۵۲	۱۲/۳	۱/۳۸	۰/۴۸	۱۳/۵۰	۱/۲۶
	مشتق اول + فیلتر ساویتزکی و گلای	۰/۵۸	۱۰/۹۱	۱/۵۵	۰/۵۶	۱۲/۵۰	۱/۳۶
	مشتق دوم + فیلتر ساویتزکی و گلای	۰/۵۷	۱۲/۳۱	۱/۳۷	۰/۵۵	۱۲/۸	۱/۳۳

جدول ۴. نتایج آماری روش‌های PLSR و PCR و روش‌های مختلف پیش‌پردازش طیفی در تعیین درصد ماده آلی خاک

مدل	پیش‌پردازش	R2 کالیبراسیون	RMSE کالیبراسیون	RPD کالیبراسیون	R2 اعتبارسنجی	RMSE اعتبارسنجی	RPD اعتبارسنجی
	جذب بدون پیش‌پردازش	۰/۳۰	۰/۲۷	۱/۱۱	۰/۲۶	۰/۲۸	۱/۰۷
	فیلتر ساویتزکی و گلای	۰/۳۳	۰/۲۵	۱/۲	۰/۳۰	۰/۲۶	۱/۱۵
PCR	متغیر نرمال استاندارد	-	-	-	-	-	-
	تصحیح پخشیده چندگانه	-	-	-	-	-	-
	مشتق اول+فیلتر ساویتزکی و گلای	۰/۳۵	۰/۲۳	۱/۳	۰/۳۳	۰/۲۵	۱/۲
	مشتق دوم+فیلتر ساویتزکی و گلای	۰/۳۰	۰/۲۴	۱/۲۵	۰/۲۹	۰/۲۵	۱/۲
	جذب بدون پیش‌پردازش	۰/۶۴	۰/۱۹	۱/۵۸	۰/۶۲	۰/۱۹	۱/۵۸
	فیلتر ساویتزکی و گلای	۰/۶۶	۰/۱۹	۱/۵۸	۰/۶۴	۰/۱۸	۱/۶۶
PLSR	متغیر نرمال استاندارد	۰/۵۷	۰/۲۴	۱/۲۵	۰/۵۲	۰/۲۴	۱/۲۵
	تصحیح پخشیده چندگانه	۰/۵۵	۰/۲۲	۱/۳۶	۰/۵۲	۰/۲۳	۱/۳۳
	مشتق اول+فیلتر ساویتزکی و گلای	۰/۷۱	۰/۱۷	۱/۷۷	۰/۶۹	۰/۱۸	۱/۶۷
	مشتق دوم+فیلتر ساویتزکی و گلای	۰/۵۷	۰/۲۳	۱/۳	۰/۵۵	۰/۲۱	۱/۴۲

۳-۴- ارزیابی صحت داده‌های طیفی در برآورد خصوصیات شیمیایی خاک

برای صحت‌سنجی مدل‌های ایجاد شده توسط روش‌های PLSR و PCR، با روش‌های مختلف پیش‌پردازش مقادیر ضریب تبیین (R2PRE) و ریشه مربعات خطا (RMSEP) برای ۳۰ درصد داده‌ها که از قبل جدا شده بودند، تعیین شد (جدول ۵). برای تعیین درصد گچ، کربنات و ماده آلی خاک از روش PLSR، همراه با پیش‌پردازش مشتق اول+فیلتر ساویتزکی و گلای استفاده شد. نتایج صحت‌سنجی نشان داد مقادیر R2PRE، برای درصد گچ، کربنات و ماده آلی به ترتیب برابر ۰/۶۵، ۰/۵۲ و ۰/۷۱ و مقدار RMSEP به ترتیب ۳/۴۵، ۱۱/۵۸ و ۰/۳ به دست آمد.

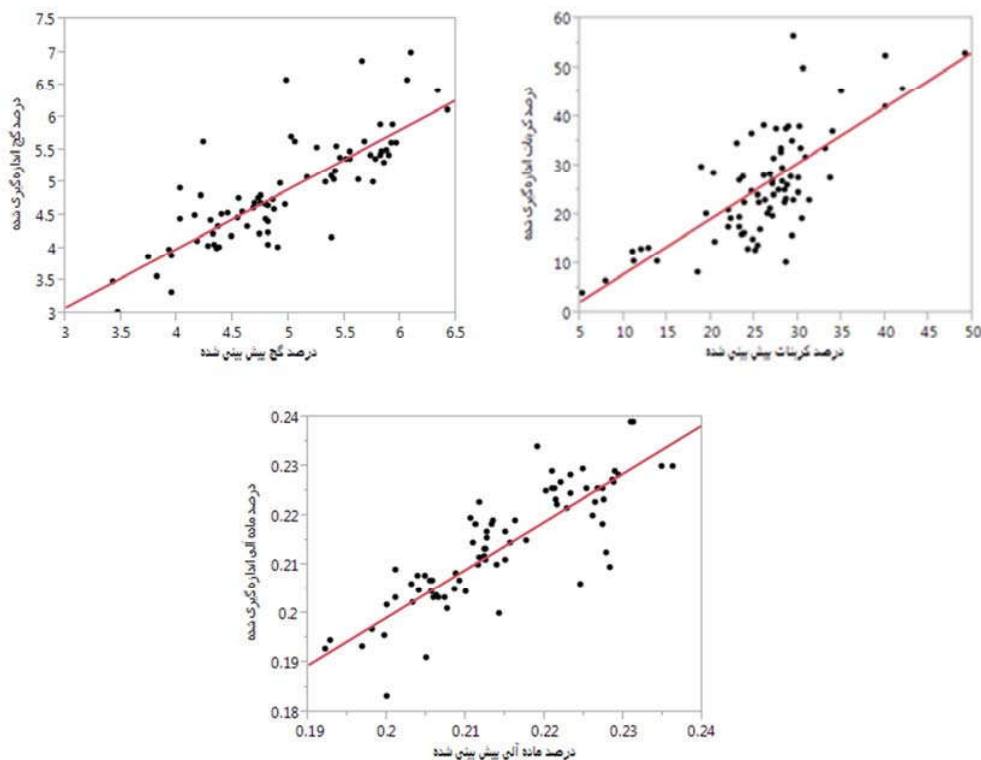
همانطور که در شکل ۳ دیده می‌شود، پیش‌بینی مدل رگرسیونی، برای درصد گچ موجود در خاک به خوبی توسط مدل رگرسیونی پیش‌بینی شده است؛ اما در مورد کربنات‌ها مدل مناسب نبود و ضعیف عمل کرد. همچنین پیش‌بینی مدل رگرسیونی برای ماده آلی، تا حد زیادی قابل قبول و مقادیر پیش‌بینی شده توسط مدل، مشابه مقادیر اندازه‌گیری شده است. مطالعات طیف‌سنجی مرئی - مادون قرمز نزدیک زیادی به ارزیابی ویژگی‌های خاک مربوط است، از جمله Lee et al (2009)، مقدار RPD را برای برآورد کربن آلی توسط مدل PLSR، ۲/۸۷ بدست آورده‌اند. Khayamim et al (2015) نتیجه گرفتند که مقادیر ضریب تبیین در گروه اعتبارسنجی برای ماده آلی، کربنات‌ها و گچ به

جدول ۵. نتایج صحت سنجی مدل سازی برای گچ، کربنات و ماده آلی خاک (%)

متغیر	مدل	پیش پردازش	R2PRE	RMSEP
گچ (%)	PLSR	مشتق اول+فیلتر ساویتزکی و گلای	۰/۶۵	۳/۴۵
کربنات (%)	PLSR	مشتق اول+فیلتر ساویتزکی و گلای	۰/۵۲	۱۱/۵۸
ماده آلی (%)	PLSR	مشتق اول+فیلتر ساویتزکی و گلای	۰/۷۱	۰/۳

است. (Stenberg et al (2010) معتقدند که تنوع نتایج برای تخمین ماده آلی دلایل متفاوتی دارد از جمله: اگرچه مواد آلی ویژگی های جذبی خاصی در محدوده مادون قرمز نزدیک دارند، ولی در اغلب موارد این ویژگی های جذبی ضعیف اند. بنابراین در خاک هایی که تنوع کانی شناسی بالایی وجود دارد و یا مقادیر شن در آنها زیاد است، ویژگی های جذبی ماده آلی به علت پخش نور با حضور کوارتز پوشیده می شود. بنابراین، تخمین ماده آلی در این نوع خاک ها ضعیف خواهد بود.

ترتیب ۰/۴۵، ۰/۶۱ و ۰/۸ است، همچنین مقادیر ریشه حداقل مربعات خطا نیز به ترتیب ۱۲/۵۱، ۰/۳ و ۳/۴۳ درصد است. نتایج (Khayamim et al(2015 نشان داد که پیش بینی مدل برای گچ، کاملاً مناسب و برای ماده آلی قابل قبول است؛ درحالی که برای کربنات های خاک ضعیف است. (Aïchi et al(2009 نتیجه گرفتند که در تخمین مقدار کربن آلی خاک، تطابق بالایی بین مقادیر اندازه گیری شده و پیش بینی شده، وجود دارد و ضریب تبیین، ریشه حداقل مربعات خطای تخمین و انحراف نسبی پیش بینی به ترتیب ۰/۳۶، ۰/۹۱ و ۳/۴ بوده



شکل ۳. مقادیر اندازه گیری شده در مقابل مقادیر پیش بینی شده گچ، کربنات و ماده آلی با استفاده از مدل PLSR

در محدوده مرئی و مادون قرمز نزدیک (۴۰۰-۲۵۰۰) می‌توان درصد گچ، کربنات و ماده آلی خاک را تخمین زد. با توجه به مقادیر RPD، پیش‌بینی مدل برای درصد گچ و ماده آلی مناسب بوده و این در حالیست که پیش‌بینی مدل برای درصد کربنات‌های خاک ضعیف است. بنابراین با توجه به نتایج فوق می‌توان دریافت که روش طیف‌سنجی مرئی-مادون قرمز نزدیک، یک روش سریع، غیرمخرب و با حداقل آماده‌سازی نمونه است که با انجام یک آزمایش می‌تواند چندین ویژگی خاک را پیش‌بینی کند. همچنین دقت پیش‌بینی پارامترهای خاک برای خاک‌های کشورمان نیز تا حد زیادی قابل قبول است. همچنین نتایج نشان داد که پیش‌پردازش مشتق اول + فیلتر ساواتزکی و گلای، بهترین ضریب تبیین و کمترین خطا را برای برآورد گچ، کربنات و ماده آلی خاک داشت. بنابراین، یکی از مهمترین عوامل تأثیرگذار بر دقت مدل‌سازی با استفاده از داده‌های طیفی، انجام پیش‌پردازش بر روی داده‌های طیفی خاک و انتخاب مناسب‌ترین روش پیش‌پردازش است.

۴- منابع

- Aïchi, H.Y., Fouad, C., Walter, R.A., Viscarra Rossel, Z.L., Chabaane and Sanaa M., 2009, **Regional predictions of soil organic carbon content from spectral reflectance measurements**, Biosys. Eng. 104: 442-446.
- Araújo, S.R., Demattê, J.A.M., Franceschini, M.H.D., Rizzo, R., Stenberg, B. & Wetterlind, J., 2013, **Improving the predictive performance of a national vis-NIR spectroscopic library by comparing clustering data transformation, and data-mining calibration techniques**, p. 2431-2440. Anais XVI Simpósio Brasileiro de Sensoriamento Remoto, 13 - 18 April. 2013. Foz do Iguaçu, PR, Brazil.
- Ben-Dor, E. & Banin, A., 1995, **Near-infrared analysis (NIRA) as a method to simultaneously evaluate spectral featureless constituents in soils**, Soil Science, 159:259-270.

به‌طور کلی، براساس نتایج به دست آمده، روش PLSR، در مقایسه با روش PCR، دارای کارایی بالاتری است و بهترین نتایج مدل‌سازی با پیش‌پردازش مشتق اول + فیلتر ساواتزکی و گلای برای برآورد گچ، کربنات و ماده آلی خاک به دست آمد. دیگر پژوهشگران نیز نتایج مشابهی، مبنی بر نتایج بهتر روش PLSR در پیش‌بینی ویژگی‌های خاک با استفاده از داده‌های طیفی در محدوده مرئی و مادون قرمز نزدیک (۴۰۰=۲۵۰۰) به دست آوردند (Shiferaw and Hergarten., 2014; Lee et al., 2009; Viscarra Rossel et al., 2007). Debae et al (2010) از روش طیف‌سنجی مرئی-مادون قرمز نزدیک، برای تخمین مقدار کربن آلی خاک، اسیدیته خاک و مقادیر منیزیم، پتاسیم، فسفر و درصد رس خاک سطحی استفاده کردند. نتایج آنها نشان داد بهترین ضریب تبیین برای درصد کربن آلی خاک، ۰/۶۵ و منیزیم خاک ۰/۶۹ است. در حالیکه ضرایب تبیین، برای سایر پارامترها ضعیف بود. (Summers et al (2011) با بررسی ۳۰۰ خاک در منطقه جنوب استرالیا به ارزیابی روش طیف‌سنجی مرئی مادون قرمز نزدیک، در برآورد کربن آلی، مقدار رس، اکسیدهای آهن و کربنات های خاک پرداختند. نتایج آنها نشان داد که ضریب تبیین برای درصد رس، مقدار کربن آلی، اکسیدهای آهن و کربنات های خاک به ترتیب ۰/۶۹، ۰/۶۱، ۰/۵۷ و ۰/۶۶ است. همچنین مقایسه روش‌های مختلف پیش‌پردازش برای گچ، کربنات و ماده آلی نشان داد که دقت مدل‌سازی با به‌کارگیری انواع روش‌های پیش‌پردازش، بر روی داده‌های طیفی می‌تواند افزایش یابد. (Sharma et و Gholizadeh et al (2015) al(2014) نیز نتایج مشابهی به دست آوردند، به طور کلی طیف‌سنجی می‌تواند برای تعیین خصوصیات خاک به کار رود.

۳-۵- نتیجه گیری کلی

نتایج این مطالعه نشان داد با استفاده از داده‌های طیفی

- Barnes, R.J., Dhanoa, M.S. & Lister, S.J., 1989, **Standard Normal Variate Transformation and De-trending of Near-Infrared Diffuse Reflectance Spectra**, Applied Spectroscopy, 43:772-777.
- Camo, AS Press ., 2006, **The Unscrambler Tutorials**. CAMO Software Research & Development Team, CAMO Software, NedreVollgate 8, N-0158, Oslo, Norway, retrieved on March 20/2012 from <http://www.camo.com/>
- 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, 151–159.
- 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 Science Society of America Journal, 65:480–490.
- Chen, L., Sheng-lu, Z., Shao-hua, W., Qing, Z. & Qi, D., 2014, **Spectral Response of Different Eroded Soils in Subtropical China: A Case Study in Changting County, China**, Journal of Materials Science, 11: 697-707.
- Curcio, D., Ciraolo, G., D'Asaro, F. & Minacapillia, M., 2013, **Prediction of soil texture distributions using VNIRSWIR reflectance spectroscopy**, Procedia Environmental Sciences, 19:494 – 503.
- Debaene, G., Niedźwiecki, J. & Pecio, A., 2010, **Visible and near-infrared spectrophotometer for soil analysis: preliminary results**, Pol. J. Agronom. 3: 3–9.
- Farmer, V.C. & Russell, J.D., 1964, **The infrared spectra of layer silicates**, Spectrochim. Acta. 20: 1149-1173.
- Gomez, C., Rossel, V. & McBratney, B., 2008, **Soil organic carbon prediction by hyperspectral remote sensing and field vis-NIR spectroscopy: An Australian case study**, Geoderma, 146:403–411.
- Gholizadeh, A., Boruvkai, L., Saberioon, M.M., Kozaki, J., Vasati, R. & Nemeki, K., 2015, **Comparing Different Data Preprocessing**, Soil & Water Res., 10:218–227.
- Khayamim, F., Khademi, H., Stenberg, B. & 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.
- Klement, A., Jaksik, O., Kodesova, R., Drabek, O. & Boruvka, L., 2013, **Application of VNIR diffuse reflectance spectroscopy to estimate soil organic carbon content, and content of different forms of iron and manganese**, Geophysical Research Abstracts, 15:10846-1.
- Lagacherie, P., Baret, F., Feret, J.B., Netto, J.M. & Robbez-Masson, J.M., 2008, **Estimation of soil clay and calcium carbonate using laboratory, field and airborne hyperspectral measurements**, Remote Sensing of Environment, 112:825–835.
- Lee, K.S., Lee, D.H., Sudduth, K.A., Chung, S.O., Kitchen, N.R. & Drummond, S.T., 2009, **Wavelength identification and diffuse reflectance estimation for surface and profile soil properties**, American Society of Agricultural and Biological Engineers, 52:683-695.
- Liu, Y., Jiang, Q., Fei, T., Wang, J., Shi, T., Guo, K., Li, X. & Chen, Y., 2014, **Transferability of a Visible and Near-Infrared Model for Soil Organic Matter Estimation in Riparian Landscapes**, Remote Sensing, 6:4305-4322.
- Martens, H. & Næs, T., 1989, **Multivariate Calibration**, John Wiley & Sons: Chichester, United Kingdom.
- McCoy, R.M., 2005, **Field Methods in Remote Sensing**, A Division of Guilford Publications, Inc. Spring, New York, U.S, pp. 67-87.
- McDowell, M.L., Bruland, G.L., Deenik, J.L., Grunwald, S. & Knox, N.M., 2012, **Soil total carbon analysis in Hawaiian soils with visible, near-infrared and mid-infrared diffuse reflectance spectroscopy**, Geoderma, 189:312-320.

- Mouazen, A.M., Kuang, B., DeBaerdemaeker, J. & Ramon, H., 2010, **Comparison between principal components, partial least squares and artificial neural network analyses for accuracy of measurement of selected soil properties with visible and near infrared 455 spectroscopy**, *Geoderma*, 158:23-31.
- Rinnan, A., Berg, F.V.D. & Engelsen, S.B., 2009, **Review of the most common pre-processing techniques for near-infrared spectra**, *Trends in Analytical Chemistry*, 28: 1201-1222.
- Schwartz, G., Eshel, G. & Ben-Dor, E., 2011, **Reflectance spectroscopy as a tool for monitoring contaminated soils**, p. 67-90. In S. Pascucci (ed.) *Soil Contamination*, In Technology.
- Sharma, S., Goodarzi, M., Ramon, H. & Saeys, W., 2014, **Performance evaluation of preprocessing techniques utilizing expert information in multivariate calibration**, *Talanta*, 121:105-12.
- Shepherd, K.D. & Walsh, M.G., 2002, **Development of reflectance spectral libraries for characterization of soil properties**, *Soil Sci. Soc. Am. J.* 66, 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, 163–215.
- Shiferaw, A., and Hergarten, Ch., 2014, **Visible near infra-red (VisNIR) spectroscopy for predicting soil organic carbon in Ethiopia**, *Journal of Ecology and the Natural Environment*, 6:126-139.
- Summers, D., Lewis, M., Ostendorf, B. & Chittleborough, D., 2011, **Visible near-infrared reflectance spectroscopy as a predictive indicator of soil properties**, *Ecological Indicators*, 11:123-131.
- Stenberg, B., Jonsson, A. & Börjesson, T., 2002, **Near infrared technology for soil analysis with implications for precision agriculture**, PP. 279–284. In: A. Davies and R. Cho, (Eds.), *Near Infrared Spectroscopy: Proceeding of 10th International Conference*. NIR Publications, Chichester, UK.
- Viscarra Rossel, R.A. & Behrens, T., 2010, **Using data mining to model and interpret soil diffuse reflectance spectra**, *Geoderma* 158, 46–54.
- Viscarra Rossel, R.A., Taylor, H.J. & McBratney, A.B., 2007, **Multivariate calibration of hyperspectral γ -ray energy spectra for proximal soil sensing**, *European Journal of Soil Science*, 58:343-353.
- Wold, S., Martens, H. & Wold, H., 1983, **The multivariate calibration problem in chemistry solved by the PLS method**, *Matrix Pencils*.



Comparing Different Statistical Models and Pre-processing Techniques for Estimation several chemical properties of the soil Using VNIR/SWIR Spectrum

Mehrabi Gohari, E^۱., Matinfar, H^{۲*}., Jafari, A^۳., Taghizadeh-Mehrjardi, R^۴., Khayamim, F.

۱. PHD Student Department of Soil Science, College of Agriculture, Lorestan University, Lorestan, Iran and Faculty, Department of Agriculture, Payame Noor University.

۲. Associate professor, Department of Soil Science, College of Agriculture, Lorestan University, Lorestan, Iran

۳. Assistant Professor, Department of Soil Science, Agriculture Faculty, Shahid Bahonar University of Kerman, Kerman, Iran

۴. Assistant Professor, Faculty of Agriculture and Natural Resources, Ardakan University, Yazd, Iran.

۵. PHD Graduate, Dept of soil science, Faculty of Agriculture, Isfahan university of Technology, Isfahan, Iran

Abstract

Visible and Near-Infrared (VNIR) and Short-Wave Infrared (SWIR) reflectance spectroscopy (۴۰۰-۲۴۵۰ nm), which are at least as costly and time-consuming, are widely used in the estimation of physical and chemical properties of the soil. The purpose of this study was to investigate the ability of this method to estimate the amount of organic matter, carbonates and gypsum content of soil surface. In the present study, ۱۱۵ profiles were identified based on the Hypercube technique, and the horizons were sampled and the amount of organic matter, carbonates and gypsum content were measured by standard methods. Reflectance spectra of all samples were measured using an ASD field-portable spectrometer in the laboratory. Soil samples were divided into two random groups (۸۰% and ۲۰%) for calibration and validation of models. PLSR and PCR models and different pre-processing methods i.e. First (FD) and Second Derivatives (SD), Multiplicative Scatter Correction (MSC) and Standard Normal Variate (SNV) were applied and compared to estimate texture elements. The highest RPD of calibration and validation were obtained for PLSR with First derivative of reflectance+ Savitzky_Golay filter pre-processing technique which was classified as a good for the amount of organic matter and gypsum and was classified as a poorly for the amount of carbonates.

Keywords: Partial Least Squares Regression (PLSR), Principal Component Regression (PCR),