

Assessing the Visible-Near-Infrared Spectroscopy Method and PLSR and SVMR Methods in Modeling Organic Carbon and Total Neutralizing Value of Soil

ROKHSAR AKBARIFAZLI^{1,2}, TEIMOUR BABAEINEJAD^{2*}, NAVID GHANAVATI², AKBAR HASANI³, MOHAMAD SADEGH ASKARI³

1. Department of Soil Science, Khuzestan Science and Research Branch, Islamic Azad University, Ahvaz, Iran

2. Department of Soil Science, Ahvaz Branch, Islamic Azad University, Ahvaz, Iran

3. Department of soil science, faculty of Agricultural Zanjan, Zanjan, Iran

(Received: Oct. 7, 2020- Revised: Feb. 10, 2021- Accepted: Feb. 27, 2021)

ABSTRACT

For sustainable land management, it is necessary to understand the functions and characteristics of soils, and their spatial and temporal changes. Visible and near-infrared spectroscopy has a specific capability to identify and determine soil properties due to high accuracy and high-performance speed. The purpose of this study is to evaluate the accuracy of visible and near-infrared spectroscopy method in estimating soil organic matter and total neutralizing value. Therefore, 110 soil samples were collected from Khuzestan, Yazd and Tehran provinces, and spectral reflectance was performed using ASD FieldSpec3. The spectra obtained from the spectrometer were pre-processed using five methods including Savitzky-Golay filter (SG), the first derivative with the Savitzky-Golay filter (FD-SG), the second derivative with the Savitzky-Golay filter (SD-SG), the standard normal variate (SNV), and Multiplicative scatter correction (MSC). Also, the performance of PLSR and SVMR methods was compared in terms of soil organic carbon and total neutralizing value estimation. The results indicated that the PLSR model in estimating both organic carbon properties and total neutralizing value had higher accuracy compared to the SVR model. In estimation of soil organic carbon, PLSR method and MSC preprocessing method had the best performance ($R^2_{VAL}=0.59$, $RMSE_{VAL}=0.19$ and $PRD_{VAL}=1.47$) and the second derivative method had the least performance ($R^2_{VAL}=0.15$, $RMSE_{VAL}=0.27$ and $PRD_{VAL}=0.52$). Also for estimation of total neutralizing value, the first derivative preprocessing method had the best performance ($R^2_{VAL}=0.78$, $RMSE_{VAL}=5.70$ and $PRD_{VAL}=2.01$) and the second derivative method had the least performance ($R^2_{VAL}=0.1$, $RMSE_{VAL}=11.13$, and $PRD_{VAL}=0.31$). The key wavelengths were observed for soil organic matter in the range of 421- 612 nm and for total neutralizing value in the range of 2315- 2151 nm. This study showed that the Vis-NIR spectroscopy method, due to its physical basis and considering the influencing factors, as a large-scale model, makes it possible to evaluate and predict soil OC and TNV.

Keywords: Organic Carbon, Preprocessing, Regression, Spectroscopy, Total Neutralizing Value.

ارزیابی روش طیف‌سنجی امواج مرئی - مادون قرمز و روش‌های PLSR و SVMR در مدل‌سازی کربن آلی و کل مواد خنثی شوند خاک

رخسار اکبری فضلی^{۱*}، تیمور بابائی نژاد^{۲*}، نوید فنواتی^۱، اکبر حسینی^۲، محمد صادق عسکری^۲

۱. گروه خاکشناسی، پردیس علوم و تحقیقات خوزستان، دانشگاه آزاد اسلامی، اهواز، ایران

۲. گروه خاکشناسی، واحد اهواز، دانشگاه آزاد اسلامی، اهواز، ایران

۳. گروه خاکشناسی، دانشکده کشاورزی، دانشگاه زنجان، زنجان، ایران

(تاریخ دریافت: ۱۳۹۹/۷/۱۶ - تاریخ بازنگری: ۱۳۹۹/۱۱/۲۲ - تاریخ تصویب: ۱۳۹۹/۱۲/۹)

چکیده

برای مدیریت پایداری اراضی، اطلاع از فعالیت‌ها و خصوصیات خاک و تغییرات زمانی و مکانی آنها ضروری است. طیف‌سنجی امواج مرئی و مادون قرمز نزدیک به دلیل دقت و سرعت عمل بالا قابلیت ویژه‌ای در شناسایی و تعیین خصوصیات خاک دارد. هدف این مطالعه ارزیابی دقت روش طیف‌سنجی مرئی-مادون قرمز نزدیک در برآورد مقدار کربن آلی (OC) و کل مواد خنثی شونده خاک (TNV) خاک است. به این منظور تعداد ۱۱۰ نمونه خاک از استان‌های خوزستان، یزد و تهران تهیه و در آزمایشگاه طیف‌سنجی گردید. طیف به‌دست آمده از دستگاه طیف‌سنج با ۵ روش پیش‌پردازش فیلتر ساویتزکی گولای (SG)، مشتق اول همراه با ساویتزکی گولای (FD-SG)، مشتق دوم همراه با ساویتزکی گولای (SD-SG)، واریانس استاندارد نرمال (SNV)، تصحیح پخشیده چندگانه (MSC) اصلاح شد. همچنین عملکرد دو روش PLSR و SVMR در برآورد ویژگی‌های خاک مقایسه گردید. نتایج نشان دادند که مدل PLSR نسبت به مدل SVMR در برآورد OC و TNV دقت بالاتری دارد. دربرآورد OC، مدل PLSR و روش پیش‌پردازش MSC ($RPD_{VAL} = 1/47$ و $RMSE_{VAL} = 0/19$)، دربرآورد TNV، بهترین عملکرد و روش پیش‌پردازش SD-SG، ضعیف‌ترین عملکرد ($RPD_{VAL} = 0/52$ و $RMSE_{VAL} = 0/27$)، دربرآورد TNV، همچنین برای روش پیش‌پردازش (FD-SG) بهترین عملکرد ($RPD_{VAL} = 2/01$ و $RMSE_{VAL} = 0/70$)، دربرآورد TNV، ضعیف‌ترین عملکرد ($RPD_{VAL} = 0/31$ و $RMSE_{VAL} = 11/13$)، دربرآورد TNV، نشان داده است. طول موج کلیدی برای OC در محدوده ۴۲۱ و ۶۱۲ نانومتر و برای TNV در محدوده ۲۳۱۵ و ۲۱۵۱ نانومتر مشاهده گردید. این مطالعه نشان داد که روش طیف‌سنجی Vis-NIR به علت دارا بودن اساس فیزیکی و در نظر گرفتن فاکتورهای تاثیرگذار، به عنوان یک مدل بزرگ مقیاس، قابلیت مناسبی برای ارزیابی و پیش‌بینی OC و TNV خاک دارد.

واژه‌های کلیدی: طیف‌سنجی، رگرسیون، پیش‌پردازش، OC، کل مواد خنثی شونده.

مقدمه

توانایی تحلیل این تکنیک به تعداد دفعات جذب و گستردگی جذب امواج مرئی و مادون قرمز نزدیک به وسیله باندهای C-H، C-H، O-H و N-H وابسته است (Xuemei and Jianshe, 2013). تقریباً تمام ترکیبات اعم از آلی یا معدنی مانند OC و TNV که پیوند کووالانسی دارند طول موج‌های متفاوتی از امواج الکترومغناطیس در محدوده مادون قرمز را جذب می‌کنند، می‌توانند رفتار طیفی منحصری را ایجاد کنند. تخمین TNV با استفاده از طیف‌سنجی به دلیل اهمیتش در صنایع‌های مختلف و تاثیر بر خصوصیات خاک و OC با توجه به نقشی که در ارزیابی کیفیت خاک و تعدیل گرمایش جهانی دارد در دهه اخیر مورد

خاک یکی از مهم‌ترین منابع طبیعی زمین محسوب می‌شود که به دلیل تأثیرات اقلیمی بر مواد مادری و فعالیت‌های زیستی در گذر زمان، دارای خصوصیات مختلفی است (Tumsavas et al., 2019). برای مدیریت پایدار اراضی، اطلاع از فعالیت‌ها، خصوصیات و تغییرات زمانی و مکانی خاک ضروری به نظر می‌رسد (Florinsky et al., 2002; Tumsavas et al., 2019). از طریق تعیین خصوصیات فیزیکی، شیمیایی و بیولوژیکی خاک می‌توان پتانسیل و محدودیت‌ها موجود در کاربری اراضی کشاورزی و غیرکشاورزی را امکان‌پذیر ساخت (Tumsavas et al., 2019).

نمونه خاک جنگل خیرود را در پیش‌بینی ظرفیت تبادل کاتیونی بررسی کردند. آنها در مطالعه خود علاوه بر استفاده از روش PLSR از دو گروه از شاخص‌های باریک‌بند خاک در دو حالت طیف بازتابندگی و مشتق استفاده کردند. نتایج مطالعه آنها نشان داد که در روش PLSR استفاده از طیف خام نسبت به مشتق منجر به نتایج بهتری شد و در مقابل در روش شاخص‌های باریک‌بند استفاده از مشتق اول نسبت به طیف خام نتایج بهتری را به دنبال داشت. PLSR به عنوان روشی که توصیف‌کننده‌ی رابطه خطی بین یک مجموعه متغیر وابسته و یک مجموعه متغیر مستقل است، به صورت گسترده برای مدل‌سازی خصوصیات خاک براساس اطلاعات طیفی استفاده شده است (Wan *et al.*, 2019). این روش می‌تواند رابطه خطی بالقوه بین ویژگی‌های خالص خاک و بازتاب طیف در محدوده مرئی-مادون قرمز را شبیه‌سازی کند (Wold *et al.*, 2001a). SVMR یک روش رگرسیونی غیرخطی جدید است که به صورت کلی با توجه به سرعت بالا و عملکرد عالی در طبقه‌بندی و ارائه رگرسیون استفاده می‌شود (Wan *et al.*, 2019). این روش بر اساس تئوری مفاهیم آماری استوار است و جواب کلی را با کمینه کردن خطای ساختاری تعیین می‌کند (Wu and Vapnik, 1998).

Khayamim *et al.*, (2018) با استفاده از ۱۰۰ نمونه خاک در منطقه اصفهان توانایی روش طیفسنجی مرئی-مادون قرمز نزدیک را در تعیین مقدار کانی‌های رسی بررسی کردند. در مطالعه آنها از دو روش رگرسیون حداقل مربعات جزئی و طیف‌های حذف پیوستار برای مدل‌سازی استفاده شد. نتایج نشان داد که روش رگرسیون حداقل مربعات جزئی نسبت به روش حذف پیوستار دارای دقت بیشتری در پیش‌بینی کانی‌های خاک است. Rasooli *et al.*, (2018) در مطالعه خود قابلیت استفاده از روش طیفسنجی مرئی-مادون قرمز نزدیک را در برآورد مقادیر OC، درصد آهک، درصد شن، درصد سیلت، درصد رس و pH، ۱۵۰ نمونه خاک منطقه بردسیر کرمان را بررسی کردند. نتایج مطالعه آنها نشان داد که براساس شاخص PRD پیش‌بینی مدل برای درصد شن و سیلت کاملاً مناسب و برای OC و آهک مناسب و برای درصد رس و pH ضعیف بوده است. (Chatrenour *et al.*, 2019a) برای مدل‌سازی شوری خاک اراضی مستعد تولید ریزگرد در استان خوزستان از ترکیب روش طیفسنجی مرئی-مادون قرمز و سه روش رگرسیون حداقل مربعات جزئی، شبکه عصبی مصنوعی و مدل جنگل تصادفی استفاده کردند. نتایج مطالعه آنها نشان داد که ترکیب دو روش رگرسیون حداقل مربعات جزئی و شبکه عصبی مصنوعی در برآورد شوری خاک دارای دقت مناسبی

توجه واقع گردیده است (Xie and Li, 2018). مواد آلی خاک، یک شاخص مهم در ارزیابی حاصلخیزی و کیفیت خاک و همچنین شاخص مهمی در تعیین سطح کیفیت خاک مزارع کشاورزی است (Hong *et al.*, 2019a; Rial *et al.*, 2017). مواد آلی می‌تواند باعث بهبود خصوصیات فیزیکی و شیمیایی خاک شوند. همچنین مواد آلی نقشی حیاتی در ارزیابی‌های زیست‌محیطی و چرخه جهانی کربن ایفا می‌کند (Powlson *et al.*, 2011). در نتیجه پایش سریع و بادقت میزان مواد آلی خاک، گامی مهم برای تخمین میزان ذخیره آن و در نتیجه تعیین حاصلخیزی خاک نمود (Hong *et al.*, 2019a). در روش مرسوم، مقدار ماده آلی خاک از طریق نمونه‌برداری از مزارع و آنالیزهای شیمیایی آزمایشگاهی برآورد می‌شود. روش‌های آزمایشگاهی دقت بالایی دارند اما بسیار زمان‌بر و پرهزینه هستند. (Tumsavaş *et al.*, 2019). پیشرفت‌های تکنیک طیفسنجی امواج مرئی و مادون قرمز نزدیک قابلیت ویژه‌ای در شناسایی و تعیین خصوصیات خاک از خود نشان داده است (Ben-Dor and Banin, 1994; Hong *et al.*, 2019a; Stoner and Baumgardner, 1981; Viscarra Rossel *et al.*, 2016). این تکنیک به دلیل کارایی زمانی، سهولت در استفاده و آماده‌سازی حداقل نمونه، به سرعت در حال ظهور است که می‌تواند هزینه و زمان آنالیز را کاهش دهد (Hong *et al.*, 2019a; Tumsavaş *et al.*, 2019). توانایی طیفسنجی Vis-NIR برای توصیف خصوصیات خاک با اجزای فعال طیفی خاک مانند اکسید آهن، کانی‌های رسی و ماده آلی مرتبط است (Rossel and Behrens, 2010; Viscarra Rossel *et al.*, 2016; Xie and Li, 2018). مطالعات گسترده‌ای در داخل و خارج کشور از روش طیفسنجی امواج مرئی و مادون قرمز نزدیک برای شناسایی مشخصات فیزیکی و شیمیایی خاک استفاده شده است. از روش طیفسنجی امواج مرئی و مادون قرمز در بررسی، برآورد و مدل‌سازی pH (Ng *et al.*, 2019) و (Wan *et al.*, 2019)، درصد رس و شن (Rossel *et al.*, 2009) و (Tumsavaş *et al.*, 2019)، درصد مواد آلی (de Santana *et al.*, 2019) و (Hong *et al.*, 2019b)، درصد OC خاک‌های مناطق شهری (Rial *et al.*, 2017)، مقدار NPK خاک (Coutinho *et al.*, 2018) و (Hu, X.Y., 2013) حاصلخیزی خاک مزارع برنج (Johnson *et al.*, 2019) CEC، (Ng *et al.*, 2019) نیتروژن کل (Sorenson *et al.*, 2018)، شوری خاک (Fariftrh *et al.*, 2007) و آلودگی‌های فلزات سنگین (Jiang *et al.*, 2018) استفاده کردند. (Mousavi *et al.*, 2018) قابلیت بازتاب طیفی ۴۵

OC و TNV خاک شده است. برای مدل سازی و ارتباط بین بازتاب نمونه های خاک و مقدار ماده آلی و کل مواد خنثی شونده از دو روش PLSR و SVMR استفاده شده است. همچنین برای بررسی اثر انواع روش های پیش پردازش بر دقت مدل سازی از ۵ روش فیلتر ساویتزکی گولای (SG)، مشتق اول همراه با ساویتزکی گولای (FD-SG)، مشتق دوم همراه با ساویتزکی گولای (SD-SG)، واریانس استاندارد نرمال (SNV)، تصحیح پخشیده چندگانه (MSC) استفاده شد.

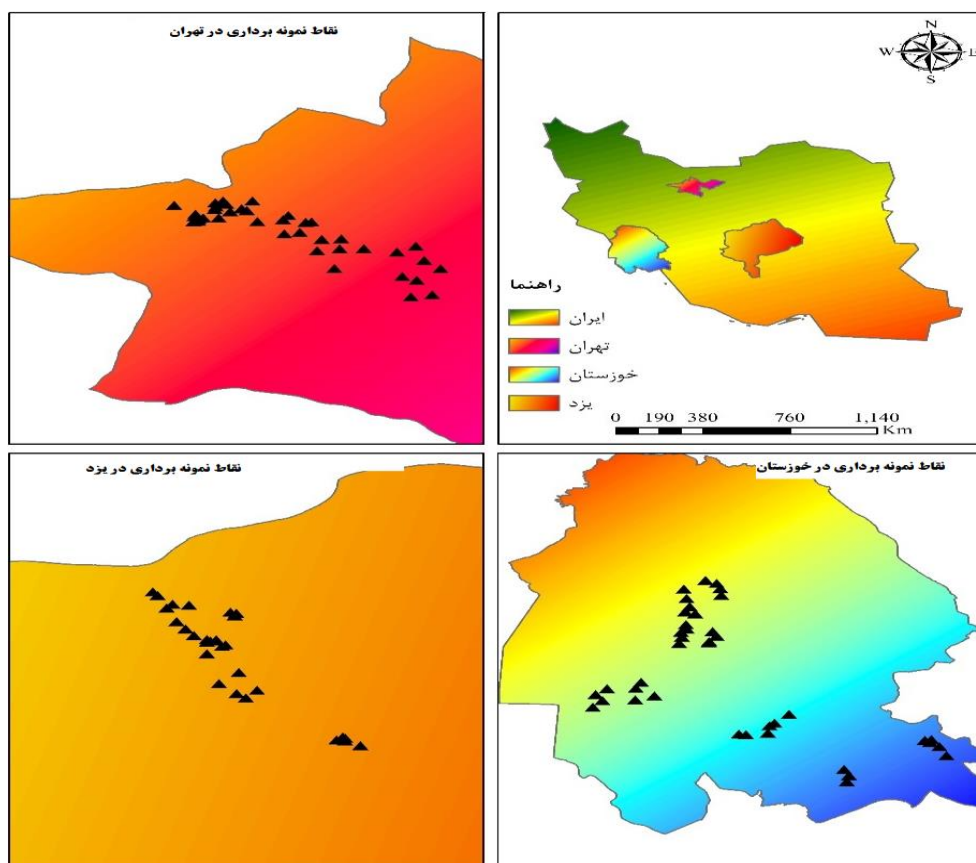
مواد و روش ها

منطقه مورد مطالعه

در این مطالعه به منظور تنوع در خاک ها و برآورد بهتر ویژگی های خاک، نعداد ۱۱۰ نمونه خاک در بخش هایی از سه استان تهران (۴۰ نمونه)، خوزستان (۴۲ نمونه) و یزد (۲۸ نمونه) برای طیف سنجی انتخاب گردید. ابتدا نمونه های خاک هوا خشک، سپس کوبیده و الک شدند. اندازه گیری TNV از روش تیتراسیون (Bashour and Sayegh, 2007) و OC از روش والکلی- بلک (Walkly and Black, 1934) تعیین شدند.

است. Chatrenour *et al.*, (2019b) در برآورد مقدار کل مواد خنثی شونده، گچ، OC و نیتروژن کل خاک مناطق مستعد تولید ریزگرد استان خوزستان از روش طیف سنجی مرئی-مادون قرمز نزدیک و روش های PLSR و ترکیب PCR و PLSR استفاده کردند. همچنین در مطالعه آنها از دو روش پیش پردازش مشتق اول و دوم استفاده شد. نتایج مطالعه آنها نشان داد که مدل PLSR در حالت پیش پردازش مشتق دوم با کاهش نویز، بیشترین دقت را برای کل مواد خنثی شونده، OC و نیتروژن با ضریب تبیین به ترتیب: ۰/۹۵ و ۰/۹۲ و ۰/۸۱ داشت. برای گچ بیشترین دقت در حالت مشتق اول با ضریب تبیین ۰/۸۷ بود.

همان طور که بیان شد، در سال های اخیر از روش طیف سنجی مرئی-مادون قرمز نزدیک برای برآورد خصوصیات مختلف خاک استفاده شده است. اما در کشور ایران اکثر مطالعات انجام شده، محدود به یک منطقه بوده است و در کمتر مطالعه ای به بررسی کارایی روش طیف سنجی در مدل سازی نمونه های خاک مناطق مختلف با سطوح متفاوت پارامترهای فیزیکی و شیمیایی اقدام شده است. لذا در این مطالعه با استفاده از نمونه های خاک برداشت شده از سه استان تهران، خوزستان و یزد اقدام به ارزیابی دقت روش طیف سنجی مرئی-مادون قرمز نزدیک در برآورد مقدار



شکل ۱- موقعیت استان های منتخب در کشور و موقعیت نقاط نمونه برداری

تا ۴۵۰ و ۲۵۰۰-۲۴۵۰ نانومتر قرار دارند حذف گردید. به علاوه دو وقفه حاصل از تغییر آشکارساز^۱ در محدوده ۱۰۰۰ و ۱۸۳۰ نانومتر نیز حذف شدند (Viscarra Rossel *et al.*, 2009). در ادامه از پنج روش پیش پردازش فیلتر ساویتزکی گولای^۲ (SD-SG) (به صورت تابع چندجمله‌ای درجه دوم و نیز تعداد ۲۳ نقطه هموارساز)، مشتق اول همراه با فیلتر ساویتزکی گولای (FD-SG)، مشتق دوم همراه با فیلتر ساویتزکی گولای (SD-SG)، واریانس استاندارد نرمال (SNV) و تصحیح پخشیدگی چندگانه (MSC) استفاده شد.

روش‌های PLSR و SVMR

در این مطالعه از دو رگرسیون حداقل مربعات جزئی^۳ (PLSR) و ماشین بردار پشتیبان^۴ (SVMR) برای مدل سازی رابطه بین مقدار ماده آلی و کل مواد خنثی شونده با امواج طیفی مرئی و مادون قرمز نمونه‌های خاک استفاده شد. PLSR روشی است که رابطه خطی بین یک مجموعه متغیر وابسته و یک مجموعه متغیر مستقل را توصیف می‌کند. ایده اصلی روش PLSR استخراج متغیرهای پیش‌بینی کننده متعامد و یا نهفته برای شناسایی هر چه بیشتر تغییرات متغیر(های) وابسته است (Farifteh *et al.*, 2007). از این رگرسیون خطی چندگانه به صورت گسترده برای مدل سازی خصوصیات خاک براساس اطلاعات طیفی استفاده شده است (Wan *et al.*, 2019).

$$Y = XW^* \times C' + F = XB + F \quad (\text{رابطه ۱})$$

Y: متغیر وابسته، X: متغیر (های) مستقل، W*: ماتریس ضرایب، C': ضرایب، F: ماتریس باقی مانده‌ها و B: ضرایب رگرسیون PLS هستند (Wold *et al.*, 2001a). جزئیات بیشتر در مورد روش PLSR را می‌توان در مطالعات (Wold *et al.*, 2001b) و (Geladi and Kowalski, 1986) یافت.

SVMR یک روش رگرسیونی غیرخطی جدید است که به صورت کلی با توجه به سرعت بالا و عملکرد عالی در طبقه‌بندی و ارائه رگرسیون استفاده می‌شود (Wan *et al.*, 2019). این روش بر اساس تئوری مفاهیم آماری^۵ (SLT) استوار است و جواب کلی را با کمینه کردن خطای ساختاری تعیین می‌کند (Wu and Vapnik, 1998). مجموعه توابع به منظور پیش‌بینی تعیین می‌شود و F(x) به گونه‌ای تعیین می‌شود که X بیشترین حاشیه را از مقادیر آموزشی Y داشته باشد. در این روش منحنی با ضخامت ϵ ، با کمترین خطای برآورد به داده‌ها برازش داده می‌شود (Smola and Scholkopf, 2004).

اندازه‌گیری بازتاب طیفی Vis-NIR نمونه‌های خاک

در شرایط آزمایشگاهی و با استفاده از طیفسنج آزمایشگاهی ASD FieldSpec3 (ساخت شرکت Analytical Spectral Devices, Inc) دانشگاه تربیت مدرس در محدوده طول موج بین ۳۵۰ تا ۲۵۰۰ نانومتر بازتاب طیفی نمونه‌ها برآورد شد. نمونه‌های خاک در فاصله ۲۰ سانتی متری لامپ دستگاه قرار گرفتند و برای جلوگیری از بازتاب بخشی، نور با زاویه قابل تنظیم ۲۰ درجه نسبت به عمود به نمونه‌ها تابانده و بازتاب از طریق گیرنده دستگاه طیف سنج ثبت گردید. برای حذف اثر توزیع اندازه ذرات بر طیف و تعدیل طیف نمونه‌ها، پتری‌دیش حاوی خاک، ۹۰ درجه چرخانده و اسکن نمونه‌ها در چهار موقعیت انجام شد. در نهایت میانگین چهار طیف برای هر نمونه تعیین شد. داده‌های بازتاب طیفی با استفاده از نرم‌افزار Viewspect به داده‌های با فرمت ACSII تبدیل و جهت ورود به نرم‌افزار The UnscramblerX (Han *et al.*, 2019b; Chatrenour *et al.*, 2019) آماده شدند (2019).

روش‌های پیش پردازش

طیفسنجی امواج مرئی و مادون قرمز نزدیک یک مکمل سریع و غیرمخرب برای روش‌های مرسوم بر نظارت مشخصات خاک است (Hong *et al.*, 2019b; Viscarra Rossel *et al.*, 2016). با این حال در عمل ممکن است عوامل بسیاری طیف خاک را تحت تأثیر قرار دهند و طیف Vis-NIR به دست آمده ممکن است شامل اطلاعات محیطی و نویز باشد که عملکرد مدل متأثر از آنها باشد (Hong *et al.*, 2019b; Stenberg *et al.*, 2010). علاوه بر این، همبستگی بین امضاهای طیفی و خصوصیات خاک در طول موج‌های خاص ممکن است با توجه به پیچیدگی‌های بین نور و مشخصات خاک و وجود همپوشانی بین دسته موج‌ها در طیف Vis-NIR که از اجزای شیمیایی خاک حاصل می‌شوند، پنهان شوند. بهترین روش برای غلبه بر این اثرات جانبی، استفاده از استراتژی‌های پیش پردازش برای حذف و یا کاهش نویزهای ناخواسته سیگنال، برجسته کردن مشخصات طیفی مطلوب و بهینه‌سازی استخراج اطلاعات طیفی مفید است. برخی از مطالعات گذشته نشان داده‌اند که پیش پردازش‌های طیفی یکی از اجزا مهم تحلیل مدل سازی چندگانه و باعث بهبود عملکرد پیش‌گویی مدل‌های می‌شوند (Hong *et al.*, 2019b). در این مطالعه برای انجام پیش پردازش، ابتدا دو بخش داری آشفستگی طیفی در ابتدا و انتهای طیف که به ترتیب در محدوده بین ۳۵۰

$$F(x) = W \times X + B \quad (\text{رابطه ۲})$$

در این روابط W وزن بردار X و B مقدار بایاس است.

معیارهای ارزیابی

در این مطالعه برای ارزیابی دقت مدل‌ها و روش‌های پیش‌پردازش مختلف در تخمین مقادیر OC و TNV از معیار نسبت عملکرد به انحراف^۱ (RPD) که در اکثر مطالعات مشابه به کار رفته، استفاده شده است. اما جهت ارزیابی دقیق‌تر نتایج، دو معیار آماری مرسوم دیگر از جمله ضریب تبیین^۲ (R^2)، ریشه میانگین مربعات خطا^۳ (RMSE) محاسبه و ارائه شده است.

$$R^2 = \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \quad (\text{رابطه ۳})$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{n}} \quad (\text{رابطه ۴})$$

$$RPD = \frac{SD}{RMSE \times \sqrt{n(n-1)}} \quad (\text{رابطه ۵})$$

در روابط ۳ تا ۵، \hat{Y} : مقدار پیش‌بینی شده، \bar{Y} : میانگین مقادیر مشاهده‌ای، y : مقادیر مشاهده‌ای و n : تعداد مشاهدات است (Hu, 2013). برای شاخص PRD شش کلاس مختلف در نظر گرفته می‌شود که در جدول (۱) ارائه شده است (Hong et al., 2019a).

جدول ۱- طبقه‌بندی کارایی و دقت مدل‌های به کار رفته براساس شاخص

*PRD					
مقدار PRD	PRD	>PRD	>PRD	>PRD	>PRD
≥ ۲/۵	≥ ۲	≥ ۱/۸	≥ ۱/۴	≥ ۱	PRD
PRD	>PRD	>PRD	>PRD	>PRD	PRD
	۲/۵	۲	۱/۸	۱/۴	
وضعیت	خیلی ضعیف	ضعیف	نسبتاً خوب	خوب	خیلی خوب
	عالی	خوب	خیلی خوب	خوب	خیلی خوب

* RPD معیار نسبت عملکرد به انحراف برای ارزیابی دقت مدل‌ها و روش‌های پیش‌پردازش

نتایج و بحث

نتایج جدول (۲) مشخصات آماری دو پارامتر OC و TNV در نمونه‌های خاک برداشت شده در نقاط نمونه برداری شده را نشان داده است. بطوریکه میانگین OC و TNV خاک‌ها به ترتیب برابر ۰/۲۵ و ۱۸/۵۴ تعیین گردید.

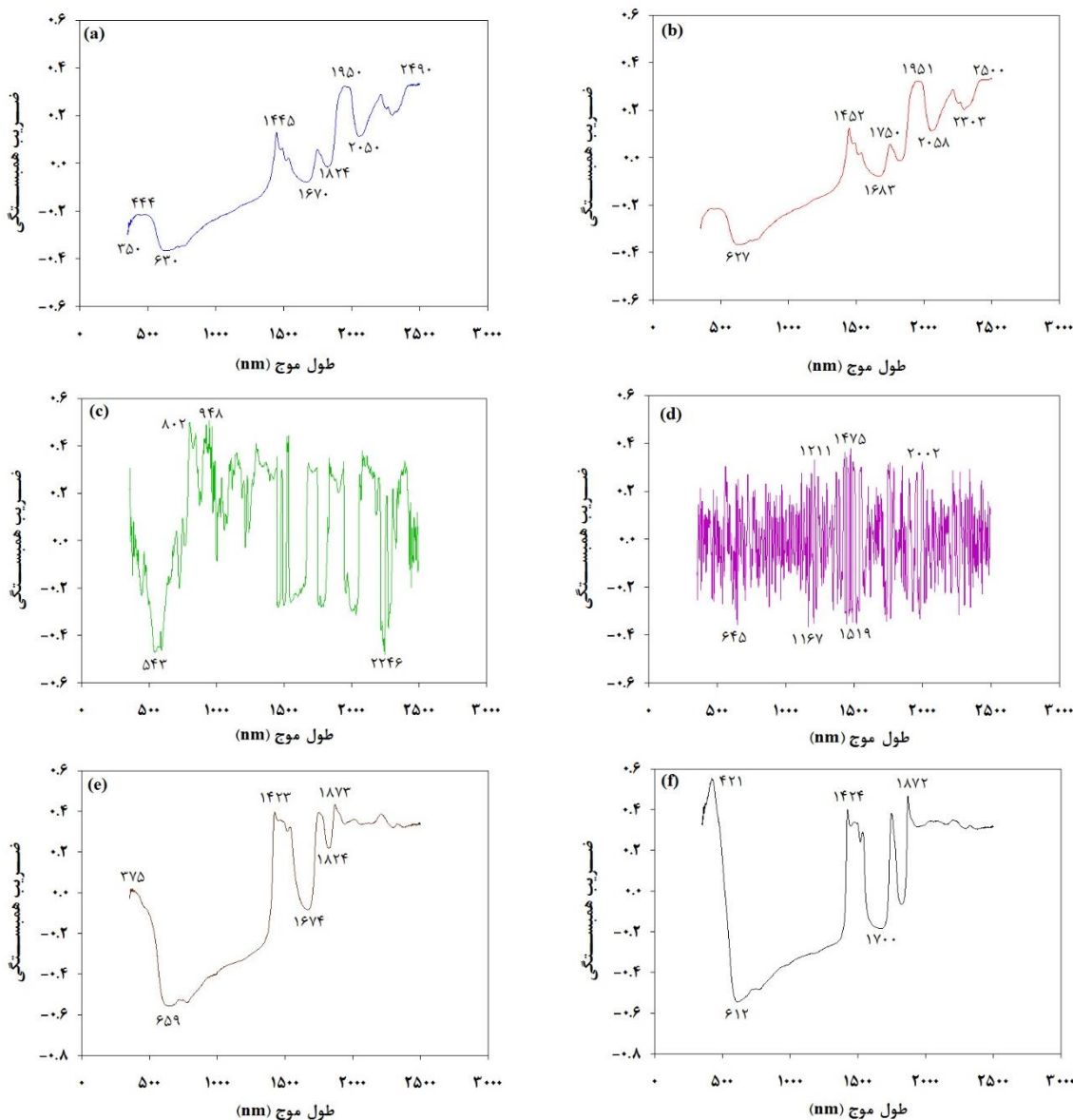
جدول ۲- مشخصات آماری OC و TNV اندازه‌گیری شده خاک

ویژگی خاک	بیشینه	میانه	کمینه	میانگین	ضریب تغییرات	انحراف معیار
OC	۱/۲۶	۰/۱۶	۰/۰۱	۰/۲۵	۰/۳۶	۰/۲۷
TNV	۴۴/۲۳	۱۶/۸۷	۱/۰۲	۱۸/۵۴	۰/۲۷	۹/۵۵

TNV: کل مواد خنثی شونده خاک، OC: کربن آلی

از یک روند مشخص برخوردار نیستند و بین محدوده طول موج ۳۰۰ تا ۲۵۰۰ نانومتر دارای نوسان هستند. نمودار همبستگی بین بازتاب و OC در دو روش پیش‌پردازش SMV ((۲(e)) و MSC ((۲(f)) در محدوده طول موج ۶۵۰ تا ۲۵۰۰ نانومتر تقریباً دارای رفتاری مشابه هستند. همچنین در روش MSC بیشترین همبستگی در طول موج ۴۲۱ و ۶۱۲ نانومتر مشاهده شده است. به دلیل کم بودن میزان بازتاب خاک در محدوده مرئی همبستگی منفی و با افزایش طول موج همبستگی در محدوده مادون قرمز مثبت می‌گردد، که این به دلیل تاثیر عواملی مانند مقدار رس و سیلت و نوع مدیریت زراعی بر نگهداری و افزایش میزان مواد آلی خاک است که منتج به ایجاد عوارض جذبی ناشی از OH موجود در شبکه‌های رس می‌گردد (Landi et al., 2019).

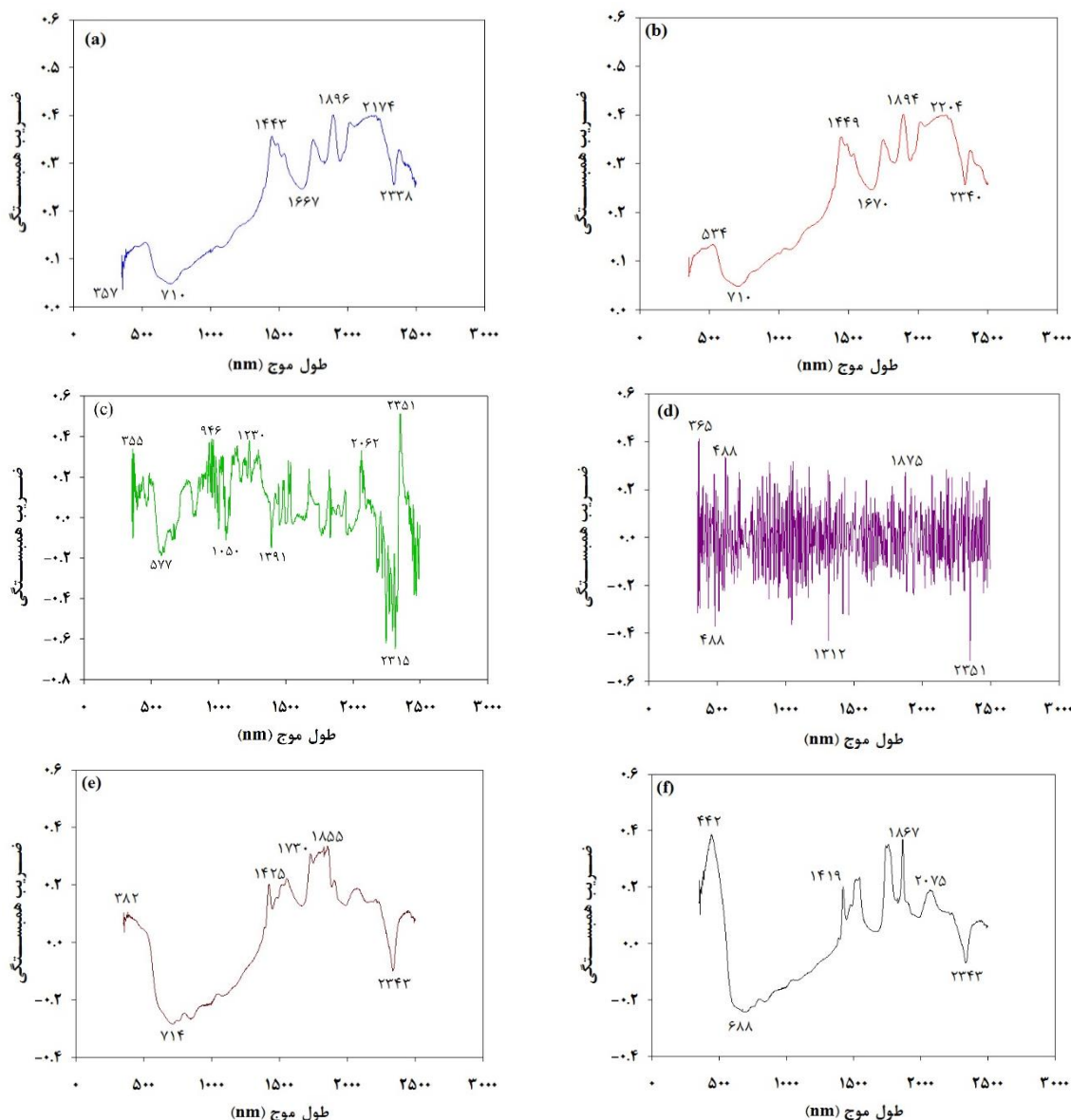
نمودارهای ضریب همبستگی بین بازتاب خاک با OC برای طیف اصلی ((۲(a)) و پنج روش پیش‌پردازش SG ((۲(b))، (FD- شکل (۲) نشان داده شده است. همان‌طور که در شکل (۲(a)) مشخص است در محدوده طول موج ۳۰۰ تا ۱۱۰۰ نانومتر و ۱۱۰۰ تا ۲۵۰۰ نانومتر بین بازتاب خاک در طیف اصلی و OC به ترتیب همبستگی منفی و مثبت وجود دارد. در طیف اصلی بیشترین همبستگی مثبت بین بازتاب خاک و OC در دو ناحیه ۱۹۵۰ نانومتر و ۲۴۹۰ نانومتر شناسایی شده است. در روش پیش‌پردازش SG ((۲(b)) تقریباً نتایجی مشابه با طیف اصلی به دست آمده است. در دو روش پیش‌پردازش (FD-SG) ((۲(c)) و (SD-SG) ((۲(d))، نمودارهای همبستگی بین بازتاب خاک و OC



شکل ۲- نمودارهای ضریب همبستگی OC نمونه‌های خاک با طیف اصلی (a)، فیلتر ساویتزکی گولای (b)، مشتق اول (c)، مشتق دوم (d)، واریانس استاندارد نرمال (e) و تصحیح پخشیده چندگانه (f)

شکل (۳(d)) SG، نمودارهای همبستگی بین بازتاب خاک و TNV از یک روند مشخص برخوردار نیستند و بین محدوده طول موج ۳۰۰ تا ۲۵۰۰ نانومتر دارای نوسان هستند. در روش پیش‌پردازش مشتق اول، بیشترین همبستگی مثبت بین بازتاب خاک و TNV مربوط به محدوده‌های ۲۳۵۱، ۲۰۶۲، ۱۲۳۰، ۹۴۶ و ۳۵۵ نانومتر است. نمودار همبستگی بین بازتاب و TNV در دو روش پیش‌پردازش SMV (۳(e)) و MSC (۳(f)) تقریباً دارای رفتاری مشابه هستند. علاوه بر این در روش‌های پیش‌پردازش SMV و MSC بیشترین همبستگی به ترتیب مربوط به نواحی ۱۸۵۵ و ۴۴۲ نانومتر است.

شکل (۳) نمودارهای ضریب همبستگی بین بازتاب خاک با TNV نمونه‌های خاک برای طیف اصلی و پنج روش پیش‌پردازش SG (۳(b))، FD-SG (۳(c))، SD-SG (۳(d))، SMV (۳(e)) و MSC (۳(f)) را نشان می‌دهد. براساس شکل (۳(a)) مشخص شد که همبستگی بین بازتاب طیفی و TNV در محدوده طول موج ۳۰۰ تا ۲۵۰۰ نانومتر مثبت است و بیشترین همبستگی در محدوده‌های ۱۴۴۳، ۱۸۹۶ و ۲۱۷۴ نانومتر وجود دارد. نمودار همبستگی بین بازتاب خاک و TNV در روش پیش‌پردازش SG (۳(b)) تقریباً مشابه با طیف اصلی به دست آمده است و بیشترین همبستگی در محدوده‌های طیفی ۱۴۴۹، ۱۸۹۴ و ۲۲۰۴ نانومتر مشاهده شد. در دو روش پیش‌پردازش (FD-SG) (۳(c)) و (SD-



شکل ۳- نمودارهای ضریب همبستگی TNV نمونه‌های خاک با طیف اصلی (a)، فیلتر ساویتزکی گولای (b)، مشتق اول (c)، مشتق دوم (d)، واریانس استاندارد نرمال (e) و تصحیح پخشیده چندگانه (f)

کردند. (Nocita *et al.*, 2014) در مطالعه خود نشان دادند که در طیف اصلی طول موج‌های محدوده ۶۸۰-۵۵۰ نانومتر به محتوای OC خاک حساس هستند. همچنین در مطالعه آنها مشخص شد که طول موج ۶۰۰ نانومتر با OC خاک همبستگی دارد. نتایج تحقیقات Hong *et al.*, (2019a) نشان دادند که بهترین طول موج برای برآورد OC خاک در محدوده ناحیه نورمرئی (۴۰۰ تا ۷۸۰ نانومتر) است که بازتاب طیفی آن به‌طور عمده تابعی از کروموفورها^۱ و تیرگی اسید هیومیک است. همچنین بیان کردند که برای OC خاک طول موج‌های مهمی در نواحی ۸۲۰ و ۸۳۰ نانومتر که به گروه C-H، و در ۹۳۰، ۹۴۰ و ۹۵۰ نانومتر که به

بررسی مطالعات گذشته نشان می‌دهد که میزان بازتاب خاک با افزایش OC کاهش و با افزایش کربنات کلسیم افزایش می‌یابد. در خاک‌هایی که میزان OC بیشتر از ۲ باشد، ماده آلی نقش پوششی را خواهد داشت و در برآورد سایر ویژگی‌های خاک باعث ایجاد خطا می‌شود (Stoner and Baumgardner, 1981). Nawar *et al.*, (2016) در مطالعه خود همبستگی مثبت و منفی بین ویژگی‌های خاک با طیف را در طول موج‌های مختلف نشان دادند. آنها بالاترین همبستگی بین OC و طیف اصلی را در طول موج ۶۰۰ نانومتر و برای طیف‌های پیش‌پردازش شده با روش‌های حذف پیوسته، مشتق اول و مشتق دوم در ۱۹۰۰ نانومتر مشاهده

R^2_{CAL} کمترین دقت را نشان داده است. مقایسه دو روش PLSR و SVMR در برآورد OC نشان داد که مدل PLSR در تمامی طیفها به جز (SD-SG)، دقت بالاتری را نسبت به مدل SVMR داشته است، در مقایسه عملکرد کلی شش نوع طیف در دو مدل PLSR و SVMR، در مرحله واسنجی روش پیش‌پردازش (MSC) در رگرسیون PLSR بهترین دقت برآورد را برای OC خاک داشته است.

در مرحله صحت‌سنجی، رگرسیون PLSR با روش پیش‌پردازش MSC ($RPD_{VAL} = 1/47$) و $RMSE_{VAL} = 0/19$ بهترین عملکرد و در مقابل با روش پیش‌پردازش (SD-SG) ضعیف‌ترین عملکرد ($RPD_{VAL} = 0/52$) و $RMSE_{VAL} = 0/27$ را نشان داده است. همچنین دو روش پیش‌پردازش SG و (FD-SG) دقتی تقریباً مشابه با طیف اصلی داشته درحالی‌که روش پیش‌پردازش (SD-SG) دقت کمتری نسبت به طیف اصلی نشان داده است. در مرحله صحت‌سنجی رگرسیون SVMR روش پیش‌پردازش (FD-SG) دارای بهترین دقت ($RPD_{VAL} = 1/44$) و $RMSE_{VAL} = 0/21$ و در مقابل روش پیش‌پردازش SNV دارای کمترین دقت ($RPD_{VAL} = 0/91$) و $RMSE_{VAL} = 0/24$ است. همچنین دو روش پیش‌پردازش (SD-SG) و SNV نسبت به طیف اصلی دقت کمتری داشته‌اند.

بر اساس نتایج مدل‌سازی ماده آلی خاک با استفاده از دو روش رگرسیونی SVMR و PLSR و بازتاب طیفی اصلی و بازتاب طیفی به‌دست آمده از پنج روش پیش‌پردازش، مشاهده شد که صرفاً استفاده از پیش‌پردازش سبب افزایش دقت برآورد OC خاک نشده است. به‌طور مثال در هر دو روش مدل‌سازی، روش پیش‌پردازش (SD-SG) دقت کمتری را نسبت به (FD-SG) نشان داده است.

در جدول (۴) نتایج مدل‌سازی TNV نمونه‌های خاک با استفاده از دو روش PLSR و SVMR و ۶ بازتاب طیفی ارائه شده است. با توجه به نتایج ارائه شده در جدول (۴)، رگرسیون PLSR براساس روش پیش‌پردازش (FD-SG) بهترین عملکرد ($RPD_{CAL} = 3/83$) و $RMSE_{CAL} = 2/37$ ، $RMSE_{CAL} = 7/27$ و $RPD_{CAL} = 0/8$ (SD-SG) و براساس روش پیش‌پردازش (SD-SG) کمترین دقت برآورد را دارد. در واقع در بین تمامی روش‌های پیش‌پردازش فقط روش (FD-SG) دقت بیشتری را نسبت به طیف اصلی نشان داد و سایر روش‌های پیش‌پردازش دقت کمتری داشته‌اند. با توجه به نتایج مدل‌سازی TNV با استفاده از رگرسیون SVMR مشخص شد که روش پیش‌پردازش (FD-SG) در گروه واسنجی بهترین عملکرد ($RPD_{CAL} = 1/41$) و $RMSE_{CAL} = 6/31$ و $RMSE_{CAL} = 0/71$ و روش

آب آزاد و گوئیت، و ۱۰۰۰ نانومتر که به N-H، و محدوده ۱۴۰۰ نانومتر به O-H، در محدوده‌های ۱۹۲۰ و ۲۳۸۰ به ترتیب به اسمکتیت و اسید هیومیک وابسته هستند.

Harrison, (2012) در مطالعه خود همبستگی گچ خالص با طیف خاک در طول موج‌های ۱۲۰۰، ۱۴۰۰-۱۶۰۰، ۱۷۵۰، ۱۹۴۵، ۲۱۰۰-۲۲۰۰ و ۲۴۰۰ را گزارش کرده است. ویژگی‌های جذبی طیف خاک در طول موج ۱۶۰۰، ۲۳۵۰، ۲۵۰۰، ۲۰۰۰ و ۱۹۰۰ می‌تواند ناشی از انواع کربنات‌ها باشد. طول موج ۵۴۰، ۵۷۰، ۱۴۳۰، ۱۴۶۰، ۱۷۴۰، ۱۸۷۰، ۱۹۴۰، ۲۰۱۰، ۲۳۵۰ و ۲۴۱۰ با شوری خاک ناشی از انواع کاتیون‌ها و آنیون‌ها همبستگی دارد همچنین در اغلب مطالعه‌ها مشخص شده است طول موج‌های ۵۴۰، ۱۷۴۰، ۲۰۱۰ و ۲۳۵۰ نانومتر به درجه شوری خاک حساسیت بیشتری دارند (Xu et al., 2016).

مدل‌سازی ویژگی‌های خاک و مقایسه تأثیر روش‌های مختلف پیش‌پردازش بر عملکرد مدل‌سازی

برای مدل‌سازی دو پارامتر OC و TNV نمونه‌های خاک و همچنین ارزیابی کارایی روش‌های مختلف پیش‌پردازش بر مدل‌سازی، با استفاده از دو روش رگرسیونی SVMR و PLSR، ۱۲ مدل توسعه یافت و کارایی این مدل‌ها با معیار ارزیابی نسبت عملکرد به انحراف (RPD) ارزیابی شد و جهت مقایسه بهتر نیز مقادیر دو معیار ارزیابی ضریب تبیین (R^2) و ریشه میانگین مربعات خطا (RMSE) نیز ارائه شده است.

نتایج برآورد OC خاک با استفاده از دو رگرسیون SVMR و PLSR و شش نوع طیف بازتابی در جدول (۳) ارائه شده است. همان‌طور که مشخص است در دوره واسنجی روش PLSR دو روش پیش‌پردازش (FD-SG) و (MSC) در مقایسه با طیف اصلی باعث افزایش شاخص PRD و دقت مدل PLSR در برآورد OC خاک شده‌اند. اما در مقابل روش پیش‌پردازش SG تغییری ایجاد نکرد و سایر روش‌های پیش‌پردازش باعث کاهش دقت مدل شده‌اند و دقت کمتری را نسبت به طیف اصلی داشته‌اند. نتایج مقایسه انواع روش‌های پیش‌پردازش در گروه واسنجی مدل PLSR نشان داد که، این مدل براساس روش پیش‌پردازش (MSC) بهترین عملکرد ($RPD_{CAL} = 4/42$) و $RMSE_{CAL} = 0/06$ ، $RMSE_{CAL} = 0/94$ و در مقابل براساس روش پیش‌پردازش (SD-SG) ($RPD_{CAL} = 0/51$) و $RMSE_{CAL} = 0/24$ ، $RMSE_{CAL} = 0/2$ کمترین دقت برآورد را داشته است.

با توجه به نتایج ارائه شده در جدول (۳) روش SVMR براساس روش پیش‌پردازش (FD-SG)، بهترین عملکرد ($RPD_{CAL} = 1/53$) و $RMSE_{CAL} = 0/18$ ، $RMSE_{CAL} = 0/71$ و براساس روش پیش‌پردازش SG ($RPD_{CAL} = 0/3$) و $RMSE_{CAL} = 0/30$ ، $RMSE_{CAL} = 0/25$

پیش‌پردازش (MSC) کمترین دقت ($R^2_{CAL} = ۰/۲$ ، $RMSE_{CAL} = ۸/۶۸$ و $RPD_{CAL} = ۱/۰۳$) را دارد. مقایسه دو رگرسیون PLSR و SVMR در برآورد TNV نشان داد که PLSR در تمامی طیف‌ها به جز (SD-SG)، دقت بالاتری را نسبت به SVMR داشته است و مقایسه عملکرد کلی شش نوع طیف در دو رگرسیون PLSR و SVMR مشخص کرد که در مرحله واسنجی روش (FD-SG) بهترین دقت برآورد را برای TNV خاک داشته است. براساس نتایج مرحله صحت‌سنجی رگرسیون PLSR، روش پیش‌پردازش (FD-SG) بهترین دقت ($R^2_{VAL} = ۰/۷۸$ ، $RMSE_{VAL} = ۰/۷۰$ و $RPD_{VAL} = ۲/۰۱$) و روش (SD-SG) کمترین دقت ($R^2_{VAL} = ۰/۱۱۳$ ، $RMSE_{VAL} = ۱۱/۱۳$ و $RPD_{VAL} = ۰/۳۱$) را دارد. همچنین با توجه به معیار PRD مشخص شد که تمامی انواع طیف‌ها دقت برآورد ضعیفی ($RPD_{VAL} < ۱/۱۳$) را داشته‌اند. علاوه براین در برآورد TNV، روش PLSR نسبت به SVMR عملکرد بهتری داشته است.

جدول ۳- نتایج معیارهای ارزیابی مدل‌سازی OC در دو دوره توسعه مدل و صحت‌سنجی

مرحله						روش پیش‌پردازش	روش مدل‌سازی
صحت‌سنجی			توسعه مدل (واسنجی)				
RPD	RMSE	R^2	RPD	RMSE	R^2		
۰/۹۹	۰/۲۰	۰/۵۶	۱/۲۴	۰/۱۷	۰/۶۰	ROW	PLSR
۰/۹۹	۰/۲۰	۰/۵۵	۱/۲۴	۰/۱۷	۰/۵۹	SG	
۰/۹۹	۰/۲۲	۰/۴۴	۲/۲۵	۰/۱۱	۰/۸۲	(FD-SG)	
۰/۵۲	۰/۲۷	۰/۱۵	۰/۵۱	۰/۲۴	۰/۲۰	(SD-SG)	
۰/۸۶	۰/۲۰	۰/۵۴	۱/۱۱	۰/۱۸	۰/۵۴	SNV	
۱/۴۷	۰/۱۹	۰/۵۹	۴/۴۲	۰/۰۶	۰/۹۴	MSC	
۱/۲۱	۰/۲۵	۰/۱۸	۱/۱۴	۰/۲۴	۰/۳۱	ROW	SVMR
۱/۲۱	۰/۲۸	۰/۱۰	۰/۳	۰/۳۰	۰/۲۵	SG	
۱/۴۴	۰/۲۱	۰/۵۳	۱/۵۳	۰/۱۸	۰/۷۱	(FD-SG)	
۱/۰۴	۰/۲۹	۰/۰۵	۱/۱۰	۰/۲۵	۰/۳۵	(SD-SG)	
۰/۹۱	۰/۳۳	۰/۲۴	۰/۶۰	۰/۴۶	۰/۲۲	SNV	
۱/۲۶	۰/۲۴	۰/۳۷	۱/۱۹	۰/۲۳	۰/۴۷	MSC	

دو رگرسیون حداقل مربعات جزئی (PLSR) و ماشین بردار پشتیبان (SVMR) و روش پیش‌پردازش فیلتر ساویتزکی گولای (SG)، مشتق اول همراه با فیلتر ساویتزکی گولای (FD-SG)، مشتق دوم همراه با فیلتر ساویتزکی گولای (SD-SG)، واریانس استاندارد نرمال (SNV) و تصحیح پخشیدگی چندگانه (MSC) استفاده شد. ریشه میانگین مربعات خطا (RMSE)، معیار نسبت عملکرد به انحراف (RPD)، ضریب تبیین (R^2)

جدول ۴- نتایج معیارهای ارزیابی مدل‌سازی TNV در دو دوره توسعه و صحت‌سنجی

مرحله						روش پیش‌پردازش	روش مدل‌سازی
صحت‌سنجی			توسعه (واسنجی)				
RPD	RMSE	R^2	RPD	RMSE	R^2		
۱/۹۸	۵/۳۳	۰/۷۲	۲/۴۸	۳/۵۱	۰/۸۶	ROW	PLSR
۱/۹۶	۵/۳۸	۰/۷۱	۲/۴۰	۳/۶۰	۰/۸۴	SG	
۲/۰۱	۵/۷۰	۰/۷۸	۳/۸۳	۲/۳۷	۰/۹۳	(FD-SG)	
۰/۳۱	۱۱/۱۳	۰/۱۰	۰/۸۰	۷/۲۷	۰/۳۸	(SD-SG)	
۱/۶۰	۵/۸۶	۰/۶۶	۱/۸۳	۴/۵۰	۰/۷۶	SNV	
۱/۸۲	۵/۳۴	۰/۷۲	۱/۹۹	۴/۲۱	۰/۷۹	MSC	
۱/۱۳	۹/۷۰	۰/۰۹	۱/۰۷	۸/۳۱	۰/۲۵	ROW	SVMR
۱/۲۱	۹/۰۸	۰/۱۰	۱/۰۷	۸/۳۰	۰/۲۵	SG	
۱/۳۶	۸/۰۴	۰/۴۴	۱/۴۱	۶/۳۱	۰/۷۱	(FD-SG)	
۱/۱۴	۹/۶۴	۰/۰۸	۱/۱۹	۷/۴۵	۰/۵۰	(SD-SG)	
۱/۱۹	۹/۲۳	۰/۰۷	۱/۰۴	۸/۵۳	۰/۲۶	SNV	
۱/۱۷	۹/۳۸	۰/۰۴	۱/۰۳	۸/۶۸	۰/۲۰	MSC	

دو رگرسیون حداقل مربعات جزئی (PLSR) و ماشین بردار پشتیبان (SVMR) و پنج روش پیش‌پردازش فیلتر ساویتزکی گولای (SG)، مشتق اول همراه با فیلتر ساویتزکی گولای (FD-SG)، مشتق دوم همراه با فیلتر ساویتزکی گولای (SD-SG)، واریانس استاندارد نرمال (SNV) و تصحیح پخشیدگی چندگانه (MSC) استفاده شد. ریشه میانگین مربعات خطا (RMSE)، معیار نسبت عملکرد به انحراف (RPD)، ضریب تبیین (R^2)

گروه واسنجی نسبت به صحت‌سنجی دارای ضریب تبیین بیشتر

مقادیر ضرایب تبیین (R^2) در هر دو مدل نشان داد که

مطابقت دارد. تفاوت در تغییرات محدوده غلظت عناصر و تفاوت خاک‌های منطقه مطالعاتی که سبب افزایش نویز در طیف خاک‌ها می‌شود نیز در نتایج برآورد مدل مؤثر است. همچنین نتایج مطالعه حاضر با نتایج مطالعه Nawar *et al.*, (2016) که با انواع پیش‌پردازش مشتق اول و مشتق دوم دقت قابل قبولی ($R^2 > 0.50$) را برای OC به دست آوردند تطابق دارد.

نتیجه‌گیری

در این مطالعه پتانسیل روش طیفسنجی امواج مرئی-مادون قرمز و استفاده از دو روش رگرسیونی SVMR و PLSR در مدل‌سازی OC و TNV مورد بررسی قرار گرفت. همچنین برای بررسی اثر انواع مختلف روش‌های پیش‌پردازش بازتاب طیفی بر عملکرد مدل‌ها از پنج روش پیش‌پردازش SG، (FD-SG)، (SD-SG)، SNV و MSC استفاده شد. در طیف اصلی بیشترین همبستگی مثبت بین بازتاب خاک و ماده آلی در دو ناحیه ۱۹۵۰ نانومتر و ۲۴۹۰ نانومتر شناسایی شد. همچنین در طیف اصلی همبستگی کل مواد خنثی‌شونده خاک و بازتاب طیفی در محدوده طول موج ۳۰۰ تا ۲۵۰۰ نانومتر مثبت و بیشترین همبستگی در محدوده-های ۱۴۴۳، ۱۸۹۶ و ۲۱۷۴ نانومتر وجود دارد. همچنین طول موج‌های کلیدی ماده آلی و کل مواد خنثی‌شونده به ترتیب در محدوده (۴۲۱ و ۶۱۲) نانومتر و (۲۳۱۵ و ۲۱۵۱) نانومتر مشاهده گردید. در مدل‌سازی OC خاک با استفاده از دو روش PLSR و SVMR بهترین عملکرد از دو روش پیش‌پردازش MSC و (FD-SG) به دست آمد و همچنین در مدل‌سازی TNV نیز روش پیش‌پردازش (FD-SG) در هر دو مدل رگرسیونی بهترین عملکرد را نشان داد. همچنین نتایج به دست آمده نشان داد که نوع پیش‌پردازش تاثیر قابل توجهی در عملکرد دو روش رگرسیونی PLSR و SVMR در مدل‌سازی پارامترهای مورد نظر دارد. علاوه بر این مشخص شد که در مدل‌سازی OC و TNV روش PLSR نسبت به SVMR دقت بیشتری داشت.

"هیچ‌گونه تعارض منافع توسط نویسندگان وجود ندارد"

REFERENCES

- Bashour, I. I. and Sayegh, A. H. (2007). Methods of analysis for soils of arid and semi-arid regions (p. 119). Rome, Italy: *Food and Agriculture Organization of the United Nations*.
- Ben-Dor, E. and Banin, A. (1994). Visible and near-infrared (0.4–1.1 μm) analysis of arid and semiarid soils. *Remote Sensing of Environment*, 48(3), 261-274.
- Ben-Dor, E. and Banin, A. (1995). Near-infrared analysis as a rapid method to simultaneously evaluate several soil properties. *Soil Science Society of America Journal*, 59(2), 364-372.

و ریشه میانگین مربعات خطای کمتری است. براساس مطالعه Wilding, (1985) داده‌هایی با ضریب تغییرات (CV) بین ۰/۱۵-۰/۳۵ با گستردگی متوسط در نظر گرفته می‌شوند و براساس مطالعات Nawar *et al.*, (2016) و Kuang and Mouazen, دامنه تغییرات غلظت ویژگی‌های خاک عامل مهمی در دقت برآورد مدل‌های رگرسیونی است. در این مطالعه انحراف معیار دو پارامتر OC و TNV به ترتیب ۰/۲۷ و ۰/۹۵ برآورد شد. تغییرات موجود در مجموعه داده‌ها، از جمله انواع خاک‌ها نیز در عملکرد مدل اثرگذار است (Wang *et al.*, 2010). با افزایش انحراف از معیار داده‌ها، مقدار R^2 کاهش و RMSE روند افزایشی را نشان می‌دهد. نتایج مطالعه حاضر با مطالعه Ben-Dor (1995) and Banin, تطابق داشت که در مطالعه آنها ضرایب تبیین مدل‌سازی سه پارامتر کربنات کلسیم، ماده آلی و مقدار رس به ترتیب ۰/۶۹، ۰/۵۵ و ۰/۸۶ و خطای استاندارد به ترتیب ۱/۱/۶، ۱/۳۴ و ۱۰/۳ برآورد شد (Ben-Dor and Banin, 1995).

همچنین نتایج به دست آمده از مطالعات Nawar (2016) *et al.*, و Hassani *et al.*, (2014) که با استناد به معیار ارزیابی PRD، برای پارامترهایی همچون کربنات کلسیم ($PRD=2/86$) و گچ ($PRD=2/65$) عملکرد عالی و برای ماده آلی عملکرد قابل قبولی ($PRD=1/64$) را مشاهده کردند، مشابه به مطالعه حاضر است. Hong *et al.*, (2019b) و Hong *et al.*, (2019a) در مطالعات خود ضریب تبیین مدل‌سازی ماده آلی خاک با استفاده از روش PLSR و طیف اصلی را ۰/۶۸ به دست آوردند که با نتیجه این مطالعه مشابه است. نتایج برآورد مدل PLSR در TNV و OC نشان داد که روش پیش‌پردازش (FD-SG) دقت بیشتری نسبت به طیف اصلی داشته و در مقابل روش پیش‌پردازش (SD-SG) دقت کمتری دارد. در واقع با مشتق گرفتن از طیف اصلی دقت برآورد تا حدودی افزایش یافته است و در مشتق دوم به شدت کاهش پیدا کرده است؛ و دلیل آن را می‌توان به کاهش شیب پایه‌ای و خطی‌سازی طیف و تولید نویز مرتبط دانست که با نتایج مطالعه Nawar *et al.*, (2016) و Khayamim *et al.*, (2015) و

- Chaternour, M., Landi A., Farrokhian firouzi A., Noroozi A., and Bahrami, H. (2020). Application of hyperspectral images in quantification of soil gypsum in center areas of Khuzestan province prone to dust generation. *Applied Soil Research*, 8(3):1-13.(In Farsi)
- Chatrenour, M., Landi, A., Farrokhian Firouzi, A., Noroozi, A. and Bahrami, H. (2019a). Modeling Soil Salinity in Khuzestan Lands Susceptible for Dust Production Using Visible-Near Infrared Spectroscopic Method. *Iranian Journal of Soil Water Research*, 50(8), 1951-

1962. (In Farsi)
- Chatrenour, M., Landi, A., Farrokhian Firouzi, A., Noroozi, A. and Bahrami, H. (2019b). Investigating soil Properties in susceptible areas of dust production in Khuzestan province by visible and wave-near infrared spectroscopy. *Journal of Watershed Engineering and Management*, 12(3), 670-685. (In Farsi)
- Coutinho, M. F. B., Alari, F.O., Ferreira, M. C., Amaral, R. (2019). Green methodology for soil organic matter analysis using a national near infrared spectral library in tandem with learning machine. *Geoderma*, 338, 401-409.
- De Santana, F. B., de Giuseppe, L.O., de Souza, M. A., Poppi, R. J. (2019). Green methodology for soil organic matter analysis using a national near infrared spectral library in tandem with learning machine. *Science of the Total Environment*, 658, 895-900.
- Farifteh, J., Van der Meer, F., Atzberger, C. and Carranza, E.J.M. (2007). Quantitative analysis of salt-affected soil reflectance spectra: A comparison of two adaptive methods (PLSR and ANN). *Remote Sensing of Environment*, 110, 59-78.
- Florinsky, I.V., Eilers, R.G., Manning, G.R. and Fuller, L.G. (2002). Prediction of soil properties by digital terrain. *Environmental Modelling and Software*, 17, 295-311.
- Geladi, P. and Kowalski, B.R. (1986). Partial least-squares regression: a tutorial. *Analytica Chimica Acta*, 185, 1-17.
- Han, Y., Tang, Z., Sun, J., Xing, X., Zhang, M. and Cheng, J. (2019). Heavy metals in soil contaminated through e-waste processing activities in a recycling area: Implications for risk management. *Process Safety and Environmental Protection*, 125, 189-196.
- Harrison, T.N. (2012). Experimental VNIR reflectance spectroscopy of gypsum dehydration: Investigating the gypsum to bassanite transition. *American Mineralogist*, 97(4), 598-609.
- Hassani, A., Bahrami, H., Noroozi, A. and Oustan, S. (2014). Visible-near infrared reflectance spectroscopy for assessment of soil properties in gypseous and calcareous soils. *Journal of Watershed Engineering and Management*, 6(2), 125-138.
- Hong, Y., Chen, S., Liu, Y., Zhang, Y., Yu, L., Chen, Y., Cheng, H. and Liu, Y. (2019a). Combination of fractional order derivative and memory-based learning algorithm to improve the estimation accuracy of soil organic matter by visible and near-infrared spectroscopy. *Catena*, 174, 104-116.
- Hong, Y., Liu, Y., Chen, Y., Liu, Y., Yu, L., Liu, Y. and Cheng, H. (2019b). Application of fractional-order derivative in the quantitative estimation of soil organic matter content through visible and near-infrared spectroscopy. *Geoderma*, 337, 758-769.
- Hu, X.Y. (2013). Application of visible/near-infrared spectra in modeling of soil total phosphorus. *Pedosphere*, 23(4), 417-421.
- Khayamim, F., Khademi, H. and Ayoubi, S. (2018). Capability and limitations of clay minerals estimation in surface soils of the Isfahan province by Vis-NIR Spectroscopy. *Iranian Journal of Soil Research*, 32, 129-140.
- Khayamim, F., Khademi, H., Stenberg, B. and Wetterlind, J. (2015). Capability of vis-NIR spectroscopy to predict selected chemical soil properties in Isfahan province. *Isfahan University of Technology*, 19(72), 81-92. (In Farsi)
- Kuang, B. and Mouazen, A.M. (2011). Calibration of visible and near infrared spectroscopy for soil analysis at the field scale on three European farms. *European Journal of Soil Science*, 62(4), 629-636.
- Mousavi, F., Abdi, E., Ghalandarzadeh, A., Bahrami, H., Majnounian, B. and Mirzaei, S. (2018). Estimate of soil cation exchange capacity using reflectance spectrometry. *Journal of Forest Research and Development*, 4, 347-361.
- Nawar, S., Buddenbaum, H., Hill, J., Kozak, J. and Mouazen, A.M., (2016). Estimating the soil clay content and organic matter by means of different calibration methods of vis-NIR diffuse reflectance spectroscopy. *Soil Tillage Reserach*, 155, 510-522.
- Ng, H. Y., Loka, V., Pooi, C. K. (2019). Treatment and hybrid modeling of domestic reverse osmosis concentrate using biological activated carbon. *Desalination*, 468, 114047.
- Nocita, M., Stevens, A., Toth, G., Panagos, P., van Wesemael, B. and Montanarella, L. (2014). Prediction of soil organic carbon content by diffuse reflectance spectroscopy using a local partial least square regression approach. *Soil Biology and Biochemistry*, 68, 337-347.
- Powlson, D.S., Brookes, P.C., Whitmore, A.P., Goulding, K.W.T., Hopkins, D.W. (2011). Soil Organic Matters. *European Journal of Soil Science*, 62, 1-4.
- Rasooli, N., Farpoor, M.H., Khayamim, F. and Ranjbar, H. (2018). Prediction of selected soil properties using visible and near infrared spectroscopy in bardsir area, Kerman province. *Iranian Journal of Soil Research*, 32, 231-243.
- Rial, M., Cortizas, A. M. and Rodríguez-Lado, L. (2017). Understanding the spatial distribution of factors controlling topsoil organic carbon content in European soils. *Science of the Total Environment*, 609, 1411-1422.
- Rossel, R. V. and Behrens, T. (2010). Using data mining to model and interpret soil diffuse reflectance spectra. *Geoderma*, 158(1-2), 46-54.
- Rossel, R.V., Behrens, T., Ben-Dor, E., Brown, D.J., Demattê, J.A.M., Shepherd, K.D., ... and Aichi, H. (2016). A global spectral library to characterize the world's soil. *Earth-Science Reviews*, 155, 198-230.
- Rossel, R.V., Cattle, S.R., Ortega, A. and Fouad, Y. (2009). In situ measurements of soil colour,

- mineral composition and clay content by vis-NIR spectroscopy. *Geoderma*, 150(3-4), 253-266.
- Smola, A. J. and Scholkopf, B. (2004). A tutorial on support vector regression, *Statistics and Computing*, 14, 199-222.
- Stenberg, B., Viscarra Rossel, R.A., Mouazen, A.M. and Wetterlind, J. (2010). Visible and Near Infrared Spectroscopy in Soil Science, In Donald L. Sparks, editor: *Advances in Agronomy*, Vol. 107, Burlington: Academic Press, 2010, pp. 163-215.
- Stoner, E.R. and Baumgardner, M.F. (1981). Characteristic variations in reflectance of surface soils. *Soil Science Society of America Journal*, 45(6), 1161-1165.
- Tumsavas, Z., Tekin, Y., Ulusoy, Y. and Mouazen, A.M. (2019). Prediction and mapping of soil clay and sand contents using visible and near-infrared spectroscopy. *Biosystems Engineering*, 177, 90-100.
- Vapnik, V.N. (1998). Statistical learning theory. Adaptive and learning systems for signal processing. *Communications and Control*, 2, 1-740.
- Viscarra Rossel, R.A., Behrens, T., Ben-Dor, E., Brown, D.J., ... and Ji, W. (2016). A global spectral library to characterize the world's soil. *Earth-Science Reviews*, 155, 198-230.
- Viscarra Rossel, R.A., Cattle, S.R., Ortega, A., Fouad, Y., (2009). In situ measurements of soil colour, mineral composition and clay content by vis-NIR spectroscopy. *Geoderma*, 150, 253-266.
- Walkly, A. and Black, I.A. (1934). An examination of the Dugtijaraff method for determining soil organic matter and proposed modification of the chronic and titration method. *Soil Science*, 37, 23-38.
- Wan, M., Qu, M., Hu, W., Li, W., Zhang, C., Cheng, H. and Huang, B. (2019). Estimation of soil pH using PXRf spectrometry and Vis-NIR spectroscopy for rapid environmental risk assessment of soil heavy metals. *Process Safety and Environmental Protection*, 132, 73-81.
- Wang, J., He, T., Lv, C., Chen, Y. and Jian, W. (2010). Mapping soil organic matter based on land degradation spectral response units using Hyperion images. *International Journal of Applied Earth Observation and Geoinformation*, 12, 171-180.
- Wilding, L.P. (1985). Spatial variability: its documentation, accommodation and implication to soil surveys. In *Soil spatial variability*, Las Vegas NV, 30 November-1 December 1984 (pp. 166-194).
- Wold, S. (2001a). M. Sjostrom, L. Eriksson. *Chemom. Intell. Lab. Syst.* 58, 109-130. Wu, Y., Vapnik, V.N., 1999. *Statistical Learning Theory, Technometrics*.
- Wold, S., Sjostrom, M., Eriksson, L., (2001b). PLS-regression: A basic tool of chemometrics. *Chemom. Intell. Lab. Syst.* 58, 109-130.
- Xie, X. L. and Li, A.B. (2018). Identification of soil profile classes using depth-weighted visible-near-infrared spectral reflectance. *Geoderma*, 325, 90-101.
- Xu, C., Zeng, W., Huang, J., Wu, J. and Van Leeuwen, W. J. (2016). Prediction of soil moisture content and soil salt concentration from hyperspectral laboratory and field data. *Remote Sensing*, 8(1), 42.
- Xuemei, L., and L. Jianshe. (2013). Measurement of soil properties using visible and short wave-near infrared spectroscopy and multivariate calibration. *Measurement*. 46(10):3808 -3814.