Iranian Journal of Soil and Water Research ISNN: 2423-7833
Homepage: http://ijswr.ut.ac.ir

Estmiating of Soil Particles Percentage Using Visible-Near Infra-Red (NIR) spectrometry in Semirom area, Isfahan

Fateme Rahmati[1]✉ | Saeid Hojati[2] | Kazem Rangzan[3] | Ahmad Landi[4]

1. Corresponding Author, Department of Soil Science, Faculty of Agriculture, Shahid Chamran University of Ahvaz, Ahvaz, Iran. E-mail: Fa.Rahmati25@gmail.com.
2. Department of Soil Science, Faculty of Agriculture, Shahid Chamran University of Ahvaz, Ahvaz, Iran. E-mail: s.hojati@scu.ac.ir.
3 Department of Remote Sensing and GIS, Faculty of Earth Science, Shahid Chamran University of Ahvaz, Ahvaz, Iran. E-mail: Kazemrangzan@scu.ac.ir.
4. Department of Soil Science, Faculty of Agriculture, Shahid Chamran University of Ahvaz, Ahvaz, Iran. E-mail: landi@scu.ac.ir.

Article Info

Article type: Research Article

Article history:

Received: May. 26, 2023

Revised: July. 13, 2023

Accepted: July. 19, 2023

Published online: Aug. 23, 2023

Keywords:
Artificial Neural Network,
Partial Least Squares Regression (PLSR),
Pre-processing methods,
Spectroscopy,
Support Vector Machine Regression.

ABSTRACT

The present research performed to estimate soil texture using visible near-infrared spectrometry in Semirom, Isfahan. A total number of 200 soil samples (0-10 cm) were collected from the Semirom area (51º 17' - 52º 3' E; 30º 42' - 31º 51' N), Isfahan. The samples were air dried and passed through a 2 mm sieve, and soil particles percentage was determined in the laboratory using hydrometry method. Reflectance spectra of all samples were measured using an ASD field spectrometer. Different pre-processing methods i.e., First Derivatives and Savitzky-Golay Filter, Multiplicative Scatter Correction and Standard Normal Variable were applied and performed on spectral data. The Partial Least Squares Regression, Support Vector Machine Regression and Artificial Neural Network models were used to estimate soil texture. The best result was obtained for Silt estimation, with excellent values of RPD >2, R^2 =0.98 and RMSE=1.08 using Artificial Neural Network model with MSC pre-processing technique. The results indicated the desirable capability of Artificial Neural Network model with MSC and SNV pre-processing techniques in estimating the Clay (RPD >2, R2=0.94 and RMSE=1.21) and Sand (RPD >2, R^2=0.84 and RMSE=6.24) contents of the soils, respectively. In general, based on the results of this study, VNIR spectroscopy was successful in estimating soil particles percentage and showed its potential for substituting laboratory analyses.

Cite this article: Rahmati, F., Hojati, S., Rangzan, K., & Landi, A. (2023).Estimating of Soil Particles Percentage Using Visible-Near Infra-Red (NIR) Spectrometry in Semirom Area, Isfahan, *Iranian Journal of Soil and Water Research*, 54 (6), 915-931. https://doi.org/10.22059/ijswr.2023.359898.669505

© The Author(s). Publisher: University of Tehran Press.

DOI: https://doi.org/10.22059/ijswr.2023.359898.669505

مجله تحقیقات آب و خاک ایران، دوره ۵۴، شماره ۶ شاپا: ۲۴۲۳-۷۸۳۳
Homepage: http://ijswr.ut.ac.ir

برآورد درصد ذرات خاک با استفاده از روش طیف‌سنجی مرئی-مادون‌قرمز نزدیک در منطقه سمیرم اصفهان

فاطمه رحمتی[۱]✉ | سعید حجتی[۲] | کاظم رنگزن[۳] | احمد لندی[۴]

۱. نویسنده مسئول، گروه علوم و مهندسی خاک، دانشکده کشاورزی، دانشگاه شهید چمران اهواز، اهواز، ایران. رایانامه: Fa.Rahmati25@gmail.com.
۲. گروه علوم و مهندسی خاک، دانشکده کشاورزی، دانشگاه شهید چمران اهواز، اهواز، ایران. رایانامه: s.hojati@scu.ac.ir.
۳. گروه سنجش از دور و GIS، دانشکده علوم زمین، دانشگاه شهید چمران اهواز، اهواز، ایران. رایانامه: Kazemrangzan@scu.ac.ir
۴. گروه علوم و مهندسی خاک، دانشکده کشاورزی، دانشگاه شهید چمران اهواز، اهواز، ایران. رایانامه: landi@scu.ac.ir.

اطلاعات مقاله

نوع مقاله: مقالهٔ پژوهشی

تاریخ دریافت: ۱۴۰۲/۳/۵
تاریخ بازنگری: ۱۴۰۲/۴/۲۲
تاریخ پذیرش: ۱۴۰۲/۴/۲۸
تاریخ انتشار: ۱۴۰۲/۶/۱

واژه‌های کلیدی:
روش‌های پیش‌پردازش،
رگرسیون حداقل مربعات جزئی،
شبکه عصبی مصنوعی،
رگرسیون ماشین بردار پشتیبان،
طیف‌سنجی.

چکیده

مطالعه حاضر با هدف برآورد درصد ذرات خاک با استفاده از روش طیف‌سنجی مرئی و مادون قرمز نزدیک در منطقه سمیرم استان اصفهان انجام بود. تعداد ۲۰۰ نمونه خاک سطحی (۱۰ سانتی‌متری) از منطقه سمیرم اصفهان (طول جغرافیایی ´۱۷ ° ۵۱ تا´۳° ۵۲ شرقی وعرض جغرافیایی ´۴۲ °۳۰ تا ´۵۱° ۳۱ شمالی) جمع‌آوری گردید. نمونه‌ها هواخشک شدند و از الک دو میلی‌متری عبور داده شدند و درصد ذرات خاک در آزمایشگاه با روش هیدرومتری تعیین شد. همچنین طیف‌سنجی نمونه‌های خاک با استفاده از دستگاه طیف‌سنج زمینی انجام گرفت. سپس روش‌های پیش‌پردازش مشتق اول با فیلتر ساویتزکی گلای، تصحیح پخشیده چندگانه و متغیر نرمال استاندارد بر روی طیف‌ها انجام شدند. برای برقراری ارتباط بین درصد ذرات خاک با ویژگی‌های طیفی آن از مدل‌های رگرسیون حداقل مربعات جزئی، ماشین بردار پشتیبان و شبکه عصبی استفاده گردید. بهترین نتیجه برای برآورد سیلت با استفاده از شبکه عصبی مصنوعی با روش پیش‌پردازش تصحیح پخشیده چندگانه با RPD (نسبت انحراف معیار به RMSE) بیشتر از ۲، R2=۰/۹۸ و کمترین مقدار RMSE=۱/۰۸ g/Kg به¬دست آمد. نتایج مطلوبی نیز برای مدل شبکه عصبی مصنوعی به ترتیب با روش‌های پیش‌پردازش تصحیح پخشیده چندگانه و متغیر نرمال استاندارد برای مقادیر رس (RPD بیشتر از ۲، R2=۰/۹۴ و کمترین مقدار RMSE=۱/۲۱ g/Kg) و شن (انحراف پیش‌بینی باقی‌مانده بیشتر از ۲، R2=۰/۸۴ و کمترین مقدار RMSE=۱/۰۸g/Kg) به¬دست آمد. به طور کلی، براساس نتایج این مطالعه، طیف‌سنجی مرئی مادون قرمز نزدیک در برآورد درصد ذرات خاک موفق بوده است و قابلیت جانشینی با روش‌های آزمایشگاهی را دارد.

استناد: رحمتی، فاطمه؛ حجتی، سسعید؛ رنگزن، کاظم؛ و لندی، احمد. (۱۴۰۲). برآورد درصد ذرات خاک با استفاده از روش طیف‌سنجی مرئی-مادون‌قرمز نزدیک در منطقه سمیرم اصفهان. *مجله تحقیقات آب و خاک ایران*، ۵۴ (۶)، ۹۱۵-۹۳۱. https://doi.org/10.22059/ijswr.2023.359898.669505

ناشر: مؤسسه انتشارات دانشگاه تهران.
© نویسندگان.
DOI: https://doi.org/10.22059/ijswr.2023.359898.669505

مقدمه

یکی از مهمترین ویژگی‌های فیزیکی خاک، بافت خاک است. بافت خاک درصد نسبی مقادیر رس، شن و سیلت است و بر ظرفیت ذخیره کربن خاک، ظرفیت نگهداری عناصر تغذیه‌ای و حاصلخیزی خاک، فرسایش‌پذیری خاک، تخلخل خاک، ظرفیت ذخیره و نفوذپذیری آب، تهویه، جرم حجمی و پایداری خاکدانه مؤثر است (Wischmeier and Smith, 1978). برای تعیین درصد ذرات خاک در مقیاس‌های بزرگ با استفاده از روش‌های آزمایشگاهی (بخصوص در مواردی که حجم نمونه‌برداری زیاد باشد) وقت و هزینه زیادی مورد نیاز است. به همین دلیل توجه محققین به جایگزین کردن فن‌آوری‌های نوین مانند طیف‌سنجی به جای روش‌های آزمایشگاهی جلب شده است. طیف‌سنجی، تهیه منحنی بازتاب یک پدیده در محدوده‌ای از طول موج‌های مشخص است. از آن جا که خاک مانند سایر اجسام موجود در طبیعت براساس ترکیب و ساختار مولکولی خود، بازتاب مشخصی را در طول موج‌های مختلف نشان می‌دهد، بنابراین روش طیف‌سنجی مرئی و مادون قرمز نزدیک، می‌تواند روش مناسبی برای تعیین ویژگی‌های خاک باشد. روش طیف‌سنجی مرئی و مادون قرمز نزدیک یک روش سریع، کم هزینه، غیر مخرب و سازگار با محیط زیست، قابل تکرار و از روش‌های تجزیه و تحلیل با قابلیت تولید مجدد می‌باشد. به طور کلی در برآورد ویژگی‌های مختلف خاک به روش طیف سنجی، طیف وسیعی از تکنیک‌های چند متغیری توسط ویسکارا راسل و برنز (۲۰۱۰) پیشنهاد شده است. رگرسیون خطی چندگانه[1]، رگرسیون خطی گام به گام[2]، رگرسیون حداقل مربعات جزئی[3]، رگرسیون مؤلفه اصلی (PCA)[4]، ماشین‌های بردار پشتیبان[5]، جنگل تصادفی و شبکه عصبی مصنوعی[6] از جمله این روش‌ها هستند. مدل‌های رگرسیونی از جمله مدل رگرسیون حداقل مربعات جزئی، رایج‌ترین مدل برای تجزیه و تحلیل کمی ویژگی‌های خاک است و در بیشتر مطالعات طیف‌سنجی ویژگی‌های خاک از این مدل استفاده شده است. علت این امر دسترسی آسان به نرم‌افزار و درک نسبتاً آسان از رابطه بین ویژگی‌های خاک و بازتاب‌های طیفی مورد استفاده برای مدل‌های رگرسیونی است (Wold et al., 2001). با این حال در برخی از پژوهش‌ها، با استفاده از این روش‌ها، نتایج مطلوبی به دست نیامده است که دلیل آن خطی نبودن رابطه بین ویژگی‌های خاک با طیف‌های به دست آمده از خاک است (مهرابی گوهری و همکاران، ۱۳۹۹). برای حل این مشکل می‌توان از مدل‌هایی که قادر به برآورد ویژگی‌های خاک با استفاده از روابط غیر خطی بین متغیرها می‌باشند استفاده کرد. روش‌های مبتنی بر هوش مصنوعی می توانند الگوهای موجود بین پارامترهای ورودی و خروجی را بهتر استخراج نموده و تشخیص بهتری از روند تغییرات پارامترها داشته باشند چرا که توانایی قابل توجهی در استخراج پاسخ هدف از داده‌های پیچیده دارند. از بین این روش‌ها می‌توان به شبکه عصبی مصنوعی و ماشین بردار پشتیبان اشاره کرد. شبکه عصبی مصنوعی برخلاف روش‌های رگرسیونی طی روند آموزش، رابطه منطقی بین متغیرهای اولیه و ثانویه را کشف می‌کند و در این روش نیازی به استفاده از رابطه جبری بین متغیرها نیست (Sargent, 2001). با این حال، شبکه‌های عصبی مصنوعی مشکلات متعددی در هنگام تعمیم‌پذیری دارند که نتیجه آن ایجاد مدل‌هایی است که بیش از حد براساس داده‌های موجود برازش یافته‌اند. یکی از مدل‌های بسیار مشابه با شبکه عصبی پرسپترون چند لایه به ویژه در طبقه‌بندی دو کلاسه، مدل ماشین بردار پشتیبان است. در واقع، این دو مدل در کنار برخی تفاوت‌هایی که دارند، عملاً یک ساختار مشابه را با دو روش مختلف آموزش می‌دهند. در شبکه عصبی پرسپترون چند لایه، پارامترهای این مدل با کمینه‌سازی خطا تنظیم می‌شوند، اما در ماشین بردار پشتیبان، ریسک ناشی از عدم طبقه‌بندی صحیح به عنوان یک تابع هدف تعریف می‌شود و پارامترها نسبت به آن تنظیم و بهینه‌سازی می‌شوند. هنگام استفاده از این مدل برای برخی از مسائل، شاید معیار خطا به صفر هم برسد. اما در میان همه مدل‌هایی که دارای خطای صفر هستند، فقط این مدل کمترین ریسک عملیاتی را دارد. از این رو در برخی موارد خروجی ماشین بردار پشتیبان در کنار عملکرد بهتری که دارد، مقاومت بیشتری نیز نسبت به تغییر و نویز در داده‌ها از خود نشان خواهد داد. زیرا اساساً طوری طراحی و آموزش داده شده است که بتواند در مقابل چنین عدم اطمینانی مقاومت کند و عملکرد مطلوب خود را داشته باشد. (Vapnik and Vapnik, 1998). این پژوهش با هدف بررسی کارایی استفاده از روش طیف‌سنجی طیف‌سنجی مرئی و مادون قرمز نزدیک در برآورد درصد ذرات خاک سطحی در شهرستان سمیرم استان اصفهان صورت گرفته است. در این مطالعه، علاوه بر مدل رگرسیونی حداقل مربعات جزئی از دو مدل شبکه عصبی مصنوعی و ماشین بردار پشتیبان در برآورد درصد ذرات

1 Multiple liniear regression
2 Stepwise regression
3 Partial least squares regression
4 Principle Component Regression
5 Support Vector Machines
6 Artificial neural network

خاک استفاده شد و کارایی این سه مدل در برآورد درصد ذرات خاک با یکدیگر مقایسه گردیدند. همچنین نتایج به دست آمده از مدل‌ها با استفاده از روش‌های پیش‌پردازش و بدون استفاده از روش‌های پیش‌پردازش با یکدیگر مقایسه گردید و بهترین روش پیش‌پردازش تعیین گردید.

پیشینه پژوهش

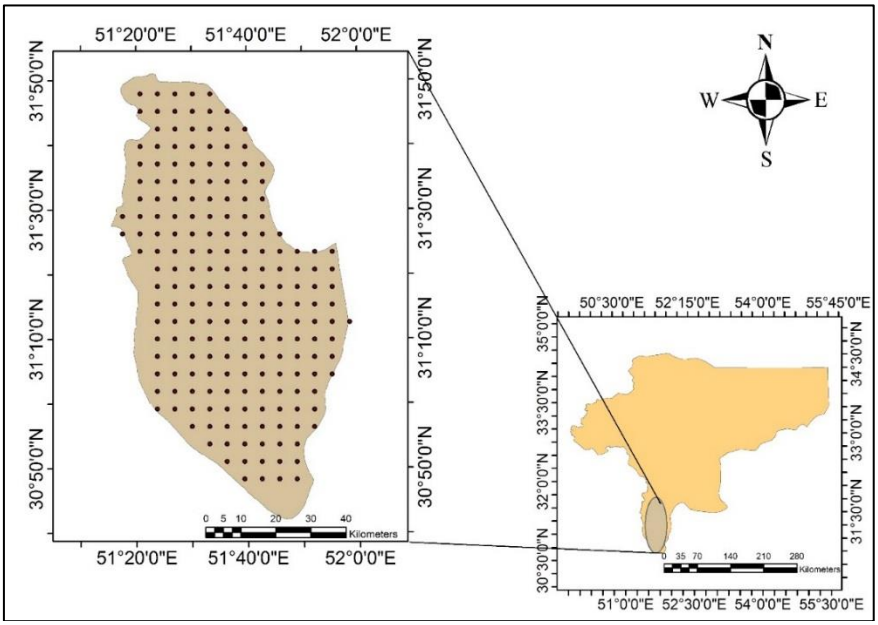
مروری بر پژوهش‌های اخیر در استفاده از طیف سنجی مرئی مادون قرمز نزدیک برای پیش‌بینی ویژگی‌های خاک نشان می‌دهد که می‌توان به این فناوری به عنوان یک گزینه کم هزینه و سریع جایگزین برای روش‌های استاندارد پیش‌بینی ویژگی‌های خاک با دقت قابل قبول اعتماد کرد (Ahmadi et al., 2021). Wang et al., (2018) برای برآورد هدایت الکتریکی خاک از روش طیف‌سنجی و دو مدل حداقل مربعات جزئی و جنگل تصادفی استفاده کردند. در این پژوهش، ضریب تبیین هدایت الکتریکی خاک با مدل جنگل تصادفی ۰/۹۳ به دست آمد. Lazaar et al., (2021) در پژوهشی در دشت تریفا واقع در شمال شرقی مراکش در ۱۰۰ نمونه خاک، درصد ذرات رس، سیلت و شن خاک را با استفاده از روش طیف‌سنجی مرئی مادون قرمز نزدیک تخمین زدند. این پژوهشگران درصد ذرات شن، سیلت و رس را به ترتیب با ضرایب تبیین ۰/۹۳، ۰/۸۷ و ۰/۵۳ و ریشه میانگین مربعات خطا ۳/۷۲، ۴/۵۵ و ۳/۷۲ برآورد کردند. رسولی و همکاران.، (۱۳۹۷) از روش طیف‌سنجی برای تخمین برخی ویژگی‌های ۱۵۰ نمونه مرکب خاک سطحی که از عمق ۲۰-۰ سانتی‌متری چهار کاربری مختلف در منطقه بردسیر در استان کرمان استفاده نمودند. در این پژوهش، مقادیر ضریب تبیین برای آهک، درصد رس، درصدشن، درصدسیلت، pH و کربن آلی به ترتیب ۰/۶۲، ۰/۳، ۰/۶۴، ۰/۶۶، ۰/۰۱ و ۰/۶۸ به دست آمد. بهترین نتایج، در برآورد میزان سیلت در روش شبکه عصبی مصنوعی با پیش‌پردازش تصحیح پخشیده چندگانه با ضریب تبیین (۰/۹۸) و ریشه دوم میانگین مربعات خطا (۱/۰۸) به دست آمد. با توجه به این که مقدار RPD (نسبت انحراف معیار به RMSE) در همه روش‌ها بیشتر از ۲ به دست آمد، پیش‌بینی همه مدل‌ها عالی بود. چترنور و همکاران (۱۳۹۸) در مطالعه‌ای، رفتار طیفی درصد رس، شن و سیلت در خاک‌های مستعد تولید ریز گرد استان خوزستان را ارزیابی کردند. در ابتدا فاکتورهای اصلی با مدل رگرسیون حداقل مربعات جزئی تعیین و سپس مدل جنگل تصادفی روی فاکتورهای تعیین شده اجرا گردید. در مرحله بعد عملکرد طیف اصلی و روش‌های پیش‌پردازش فیلتر ساویتزکی و گلای، فیلتر ساویتزکی و گلای به همراه مشتق اول، فیلتر ساویتزکی و گلای به همراه مشتق دوم، روش نرمال‌سازی استاندارد و روش حذف پیوستار در حذف نویز و افزایش دقت مدل مقایسه شدند. نتایج نشان داد که روش حذف پیوستار در دو ویژگی درصد رس (RPD_{CAL}=۱/۹۸) و درصد سیلت (RPD_{CAL}=۱/۶۵) و روش مشتق دوم برای درصد شن (RPD_{CAL}=۱/۹۷) بهترین عملکرد را داشته‌اند. مهرابی گوهری و همکاران (۱۳۹۹) مطالعه‌ای را با هدف پیش‌بینی بافت خاک با استفاده از اندازه‌گیری‌های طیفی و مدل‌های شبکه عصبی مصنوعی و رگرسیون حداقل مربعات جزئی انجام دادند. نتایج این بررسی نشان داد که شبکه عصبی مصنوعی نسبت به رگرسیون حداقل مربعات جزئی کارایی بهتری داشت. مدل رگرسیون مربعات جزئی دقت بسیار پایینی داشت (R^2~۰/۰-۱/۳). در حالی که، در روش شبکه عصبی مصنوعی مقدار R^2 به ترتیب برای رس، شن و سیلت ۰/۷، ۰/۷۶ و ۰/۷۳ به دست آمد و میانگین ریشه مربعات خطا به ترتیب ۱۴/۹، ۵/۵۴ و ۷/۰۱ گرم بر کیلوگرم براساس داده‌های آزمون (۲۰ درصد داده‌ها) به دست آمد که نشان دهنده دقت بالاتر و خطای کمتر مدل شبکه عصبی مصنوعی می‌باشد. همچنین نتیجه گرفتند که از آن‌جایی که رابطه بین درصد ذرات خاک و بازتاب طیفی خاک خطی نیست، به نظر می‌رسد روش شبکه عصبی مصنوعی برای بررسی و تجزیه و تحلیل رابطه بین اجزای بافت خاک و داده‌های طیفی مناسب باشد. رحمتی و همکاران.، (۱۴۰۱) در پژوهشی هدایت الکتریکی خاک، کربنات کلسیم و کربن آلی و واکنش خاک ۲۰۰ نمونه خاک سطحی منطقه سمیرم اصفهان را با استفاده از مدل‌های حداقل مربعات جزئی، رگرسیون مؤلفه اصلی، شبکه عصبی مصنوعی و رگرسیون ماشین بردار پشتیبان به روش طیف‌سنجی تخمین زدند. بهترین مدل‌ها در برآورد هدایت الکتریکی خاک، کربنات کلسیم و کربن آلی مدل حداقل مربعات جزئی و برای واکنش خاک مدل رگرسیون ماشین بردار پشتیبان و بهترین روش‌های پیش‌پردازش، روش‌های مشتق‌گیری به ترتیب با ضرایب تبیین ۰/۹۴، ۰/۸۸، ۰/۹ و ۰/۷۹ بودند و تمام برآوردها، کمترین RMSE را نسبت به روش‌های دیگر و RPD> ۲ داشتند.

روش‌شناسی پژوهش

منطقه مورد مطالعه

در این پژوهش منطقه مورد مطالعه، کل شهرستان سمیرم واقع در جنوب غربی استان اصفهان می‌باشد. مساحت شهرستان سمیرم ۵۲۲۴

کیلومتر مربع است و این شهرستان بین طول جغرافیایی´ ۱۷ ° ۵۱ تا´۳° ۵۲ شرقی وعرض جغرافیایی´ ۴۲° ۳۰ تا ´۵۱° ۳۱ شمالی قرار دارد. ارتفاع متوسط منطقه ۲۴۰۰ متر از سطح دریا است. براساس تقسیم‌بندی کوپن آب و هوای منطقه، سرد با تابستان‌های خشک و گرم و در ارتفاعات جنوبی سرد با تابستان‌های خشک و خنک می‌باشد. میانگین دمای سالیانه منطقه ۱۰/۶ درجه سانتی‌گراد است. میانگین بارندگی سالانه منطقه ۳۹۸/۱ میلی‌متر است. پوشش گیاهی منطقه عمدتاً استپی، علفزارها و درختان جنگلی است. سمیرم در منطقه رورانده زاگرس و زون سنندج سیرجان واقع شده است و از نظر چینه‌شناسی تشکیلات مربوط به دوران دوم زمین‌شناسی تا عهد حاضر در آن پراکندگی دارد قدیمی ترین رخنمون های سنگی در قالب شیل–شیست و آهک‌های دگرگون شده دوره ژوراسیک، نهشته‌های آهکی اربیتولین دار و مارن های قرمز، سبز و خاکستری به همراه کنگلومرا و شیل های رنگ کرتاسه–تشکیلات آسماری ائوسن و همچنین مارن، شیل، ماسه‌سنگ و کنگلومرا به همراه تشکیلات گچساران میوسن و تشکیلات کنگلومرای بختیاری پلیوسن و آبرفت‌های دوران چهارم شامل تراس‌های جدید و قدیمی کنار رودخانه‌ها و خاک‌های کشاورزی و مخروط‌افکنه‌ها از بارزترین شاخص‌های زمین‌شناسی بخش عمده محدوده مورد مطالعه می‌باشند. در محدوده شمال شرقی و شمال غرب و غرب خاک ها کم عمق، بدون تکامل پروفیلی و جوان و سنگریزه دار که بر روی مواد آهکی قرار گرفته است. منابع ارضی این محدوده را تپه‌های نسبتاً فرسایش یافته با قلل مدور و تشکیلات سنگ‌های آهکی و مارنی تشکیل می‌دهند. عمق زیاد و بافت سنگین تا خیلی سنگین خاک ها در دشت دامنه ای و رسوبات با شیب ملایم و سنگریزه های سطحی از مشخصات خاک و منابع اراضی این قسمت است. ضمن آن که خاک اطراف منطقه حنا روی دو گونه اراضی دشت دامنه‌ای و واریزه‌های سنگ‌دار تشکیل شده و از جهت عمق و بافت متفاوت است. خاک منطقه انتی‌سول، آلفی‌سول و اینسپتی‌سول می‌باشند (جلالیان، ۱۳۷۶). شکل ۱، موقعیت نقاط نمونه‌برداری در منطقه مورد مطالعه را نشان می‌دهد.



شکل ۱. موقعیت نقاط نمونه‌برداری در منطقه مورد مطالعه

نمونه‌برداری و آنالیزهای آزمایشگاهی

به منظور انجام این پژوهش تعداد ۲۰۰ نمونه خاک از لایه سطحی (۰–۱۰ سانتیمتر فوقانی) به روش نمونه‌برداری تصادفی منظم (Minasny and McBratney, 2006) برداشت شد. نمونه‌های جمع‌آوری شده ابتدا هوا خشک شده و سپس از الک دو میلی‌متری عبور داده شدند. اندازه‌گیری رطوبت خاک بوسیله خشک کردن نمونه‌ها در یک آون در دمای ۱۰۵ درجه سانتیگراد و به مدت ۲۴ ساعت صورت گرفت. سپس واکنش خاک در گل اشباع با دستگاه pHمتر، شوری خاک در عصاره اشباع و با استفاده از ECمتر (Richards, 1954)، بافت خاک با استفاده از روش هیدرومتری (Bouyoucos, 1951)، آهک به روش تیتراسیون برگشتی (Lanyon and Heald, 1982)، گچ به روش ترسیب با استون (Nelson, 1982) و میزان کربن آلی نمونه‌ها به روش اکسیداسیون تر و در مجاورت دی کرومات پتاسیم (Walkley and Black, 1934) اندازه‌گیری شدند.

طیف‌سنجی

در طیف‌سنجی، انرژی الکترومغناطیسی از یک منبع نور به یک پدیده برخورد می‌کند و بخشی از آن منعکس، بخشی جذب و بخشی دیگر عبور داده می‌شود (Clark, 1999). هر پدیده به دلیل ویژگی‌های خاص ترکیب مولکولی و ساختار خاص، رفتار طیفی منحصری را در این ناحیه از طیف الکترومغناطیس ایجاد می‌کند که اثرانگشت طیفی نامیده می‌شود (Viscarra Rossel et al., 2006). در این مطالعه، طیف‌ها در آزمایشگاه با استفاده از دستگاه طیف‌سنج زمینی FieldSpec 3 ASD (ساخت شرکت Inc, Analytical Spectral Devices) گروه علوم زمین دانشگاه شهید چمران اهواز در ناحیه مرئی–مادون قرمز نزدیک مادون قرمز طول موج کوتاه در محدوده ۳۵۰ تا ۲۵۰۰ نانومتر تهیه گردیدند (Janik et al., 2009). طیف‌ها با استفاده از نرم‌افزار ویو اسپک پرو (نسخه ۶,۲,۰) به صورت فایل متنی در آمدند. سپس برای تجزیه و تحلیل از نرم‌افزار آنسکرامبلر (نسخه ۱۰,۴) استفاده شد.

پیش‌پردازش طیف‌ها

در این پژوهش دو بخش دارای نویز در ابتدا و انتهای طیف‌ها در دامنه طول موجی بین ۳۵۰تا ۴۵۰ نانومتر و ۲۴۵۰ تا ۲۵۰۰ نانومتر حذف گردیدند. سپس مقادیر انعکاس به جذب تبدیل شد (معادله ۱). در این معادله A و R به ترتیب مقادیر جذب و انعکاس بر حسب درصد می‌باشد.

$$A=\mathrm{Log}\,(1/R)$$ رابطه ۱)

در مرحله بعد، به منظور حذف تأثیرات عوامل ناخواسته از روش‌های پیش پردازش استفاده شد. روش‌های پیش‌پردازش به منظور کاهش پخشیدگی نور، غیرخطی بودن و تأثیرات عوامل محیطی ناخواسته مانند زبری سطح خاک، بقایای آلی درشت، آلودگی پروب، تغییر فاصله سنسور و خاک و تأثیرات پرتوافکنی جوی از روش‌های پیش پردازش استفاده می‌شوند. نتایج پژوهش‌های مختلف در تخمین ویژگی‌های طیف‌سنجی نشان می‌دهد که روش‌های مختلف پیش‌پردازش تأثیر قابل توجهی بر روی مدل رگرسیونی دارند (Kuśnierek, 2011; Buddenbaum and Steffens, 2012; Gras et al., 2014; Nawar et al., 2016). نتایج برآورد ممکن است با استفاده از روش‌های مختلف پردازش آماری و مشتق گرفتن از طیف‌ها تغییر کند (Reeves et. al., 2002). روش‌های پیش‌پردازش مورد استفاده در این پژوهش شامل تصحیح پخشیده چندگانه[1]، متغیر نرمال استاندارد[2]، مشتق اول[3] به همراه فیلتر ساویتزکی گلای بودند. در این پژوهش برای اعمال فیلتر ساویتزکی گلای از چندجمله‌ای‌های درجه ۲ و تعداد نقاط همسایگی ۱۵ استفاده شد.

مدل‌سازی

مدل‌های مورد استفاده برای برآورد ذرات خاک مدل‌های رگرسیون حداقل مربعات جزئی، ماشین بردار پشتیبان و شبکه عصبی بودند. رگرسیون حداقل مربعات جزئی به علت دسترسی آسان به نرم‌افزار و درک آسان رابطه بین ویژگی‌های خاک و بازتاب‌های طیفی مدلی بسیار متداول است. این مدل سریع و به آسانی قابل تفسیر است. رگرسیون حداقل مربعات جزئی یکی از روش‌های با قابلیت در تجزیه و تحلیل داده‌ها می‌باشد که علت آن وابستگی کم به مقیاس اندازه‌گیری، حجم نمونه و توزیع باقی مانده و عدم وابستگی آن به فرض نرمال بودن داده‌ها است.

مدل رگرسیون حداقل مربعات جزئی ترکیبی از رگرسیون خطی چندگانه (MLR) و تجزیه به مؤلفه‌های اصلی[4] (PCA) است و از طریق یک رابطه خطی چندمتغیره، ارتباطی بین دو ماتریس x (داده‌های طیفی) و y (ویژگی خاک) ایجاد می‌کند (رابطه ۲). b_i بردار رگرسیونی و b_0 مرز جدا کننده می‌باشد (Wold et al., 2001).

$$y=b_i x+b_0$$ (رابطه ۲)

ماشین بردار پشتیبان یک روش یادگیری با نظارت و غیر پارامتریک است و می‌تواند بر مشکلات روش‌های دیگر که از گرایش بالقوه آن‌ها به بیش‌برازش اطلاعات باز می‌گردد غلبه کند. این مدل بر مبنای تئوری مفاهیم آماری[5] استوار است و برای برآوردهای رگرسیونی غیرخطی کاربرد دارد. ماشین بردار پشتیبان، خطای ساختاری نتایج را به حداقل می‌رساند (رابطه ۳ و ۴). در این مدل، تابعی مرتبط با متغیر

1. Multiplicative Scatter Correction
2. standard normal variate
3. First Derivative
4- Principal Component Analysis
5 Statistical learning theory

وابسته y که خود تابعی از چند متغیر x مستقل است، برآورد می‌شود. مشابه سایر مسائل رگرسیونی، فرض می شود رابطه میان متغیرهای مستقل و وابسته با تابع جبری مانند f(x) به علاوه مقداری خطای مجاز (ε) مشخص شود. W بردار ضرایب، Ø ثابت کرنل و b ثابت معادله است (Vapnik and Vapnik, 1998). محاسبات مربوط به مدل‌های رگرسیون حداقل مربعات جزئی ماشین بردار پشتیبان با استفاده از نرم‌افزارهای Unscrambler X 10.4 انجام شدند.

$$f(x)=W^T.Ø(x)+b$$ (رابطه ۳)

$$y=f(x)+noise$$ (رابطه ۴)

شبکه‌های عصبی مصنوعی می‌توانند در مدل‌سازی سیستم‌های پیچیده به کار برده شوند و خود رابطه منطقی بین متغیرها را طی روند آموزش کشف می‌کنند (Almeida, 2002). پردازش در این شبکه‌ها با تعداد زیادی نرون انجام می‌شود. هر نرون در هر لایه با وزن‌هایی به عناصر لایه قبل و بعد متصل می‌شود (Karayiannis and Venetsanopouios, 1993). در این پژوهش مدل شبکه عصبی مصنوعی با استفاده از نرم‌افزار JMP با سه لایه مخفی برای ایجاد شبکه پرسپترون مورد استفاده قرار گرفت و تعداد نرون‌های لایه مخفی به صورت سعی و خطا به دست آمدند. شبکه عصبی مصنوعی دارای یک لایه ورودی، یک یا چند لایه مخفی و یک لایه خروجی است. در هر لایه تعدادی نرون وجود دارد که از طریق وزن‌ها به گره‌های بعدی متصل می‌شوند. تعداد نرون‌ها در لایه ورودی وابسته به تعداد پارامترهای ورودی و تعداد نرون در لایه خروجی وابسته به تعداد پارامتر خروجی است. به طور کلی تعداد نرون در لایه پنهان تابع ضابطه خاصی نبوده و تعیین تعداد مناسب آن فقط از طریق سعی و خطا امکان پذیر است. به عبارت دیگر تعداد بهینه نرون‌های مخفی زمانی حاصل می‌شود که شبکه عصبی بهترین پردازش را با کمترین خطا در هر دو مرحله آموزش و آزمون مدل انجام دهد (رابطه ۵). تابع f توسط کاربر انتخاب می‌شود و پارامترهای w (وزن) و b (ثابت معادله) قابل تنظیم است (Sargent, 2001).

$$a=f(wp+b)$$ (رابطه ۵)

ارزیابی مدل‌ها

به منظور واسنجی و اعتبارسنجی از روش اعتبارسنجی متقابل[1] استفاده شد. داده‌ها به دو گروه ۷۰٪ برای واسنجی و ۳۰٪ برای اعتبارسنجی مدل‌ها تقسیم شدند. در نهایت به منظور تعیین طول موج‌های بهینه، از روش آزمون عدم قطعیت مارتنز[2] استفاده شد. ارزیابی دقت نتایج با استفاده از آماره‌های ریشه میانگین مربعات خطا (رابطه ۶)، ضریب تبیین (رابطه ۷) و RPD [3] (رابطه ۸) انجام شد. در این روابط x_o مقدار اندازه‌گیری شده در آزمایشگاه، x_e مقدار برآورد شده به روش طیف‌سنجی، SD انحراف معیار و n تعداد داده‌ها است.

$$RMSE=[1/n\sum(x_o - x_e)]^{1/2}$$ (رابطه ۶)

$$R^2=[(\sum(x_e - \bar{x}_e) (x_o - \bar{x}_o))/(\sum(x_e - \bar{x}_e)^2(x_o - \bar{x}_o)^2)]$$ (رابطه ۷)

$$RPD=SD/RMSE$$ (رابطه ۸)

یافته‌های پژوهش

توصیف آماری درصد ویژگی‌های خاک اندازه‌گیری شده در آزمایشگاه در جدول ۱ نشان داده شده است. براساس آزمایش‌های انجام شده منطقه فاقد گچ بوده است.

جدول۱. توصیف آماری اجزای بافت خاک در منطقه مطالعاتی

ویژگی آماری	سیلت (٪)	شن (٪)	رس (٪)	کربن آلی	کربنات کلسیم معادل (٪)	واکنش خاک	هدایت الکتریکی خاک
میانگین	۳۹/۵۲	۳۱/۲۶	۲۷/۴۷	۰/۸۹	۳۱/۶	۷/۶	۰/۶
حداکثر	۵۶	۷۰	۱۵	۲/۱۷	۷۰	۸/۱۵	۱/۵۲
حداقل	۱۰	۵	۴۰	۰/۳	۵	۷	۰/۲۹
دامنه	۴۶	۶۵	۲۵	۱/۸۷	۶۵	۱/۱۵	۱/۲۳
انحراف معیار	۱۰/۰۸	۱۵/۷۱	۵/۱۴	۰/۱۹	۳/۸۷	۰/۱	۰/۱۷
ضریب تغییرات (٪)	۲۶	۵۰	۱۹	۲۱/۷	۱۲/۴	۱/۳۴	۲۸/۷۴

1. Cross validation
2. Martens' Uncertainty Test
3. Ratio of performance to deviation

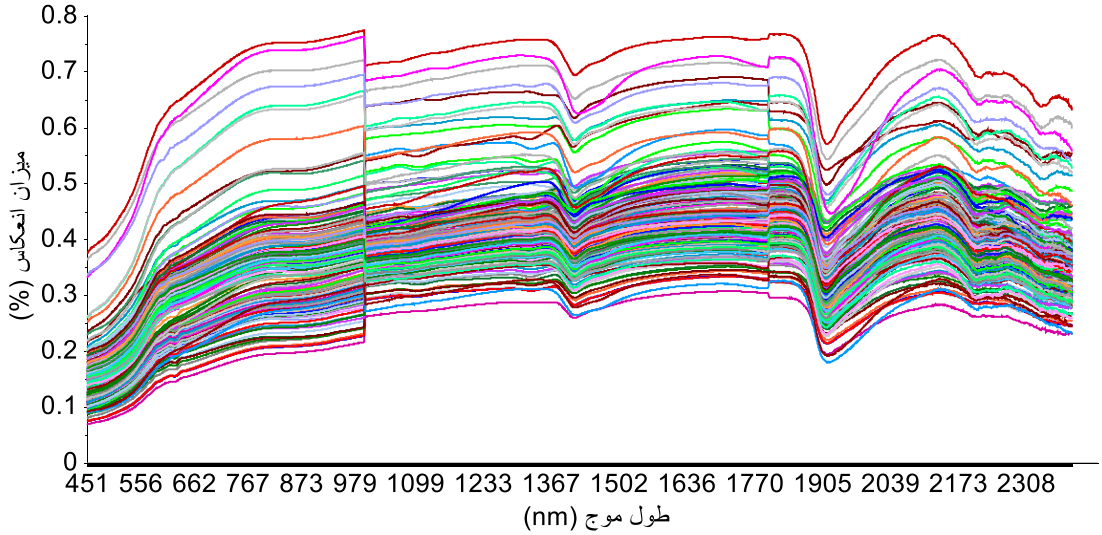
بر اساس طبقه‌بندی Wilding (1985) ضریب تغییرات کمتر از ۱۵ نشان دهنده تغییرپذیری کم، ضریب تغییرات ۱۵ تا ۳۵ درصد نشان دهنده تغییرپذیری متوسط و ضریب تغییرات بالاتر از ۳۵ درصد نشان دهنده تغییرات زیاد ویژگی‌های خاک در منطقه می‌باشد. بنابراین، تغییرات مقادیر رس و سیلت در منطقه متوسط و تغییرپذیری مقدار شن در منطقه زیاد می‌باشد.

جدول ۲. ضرایب همبستگی پیرسون بین ویژگی‌های خاک

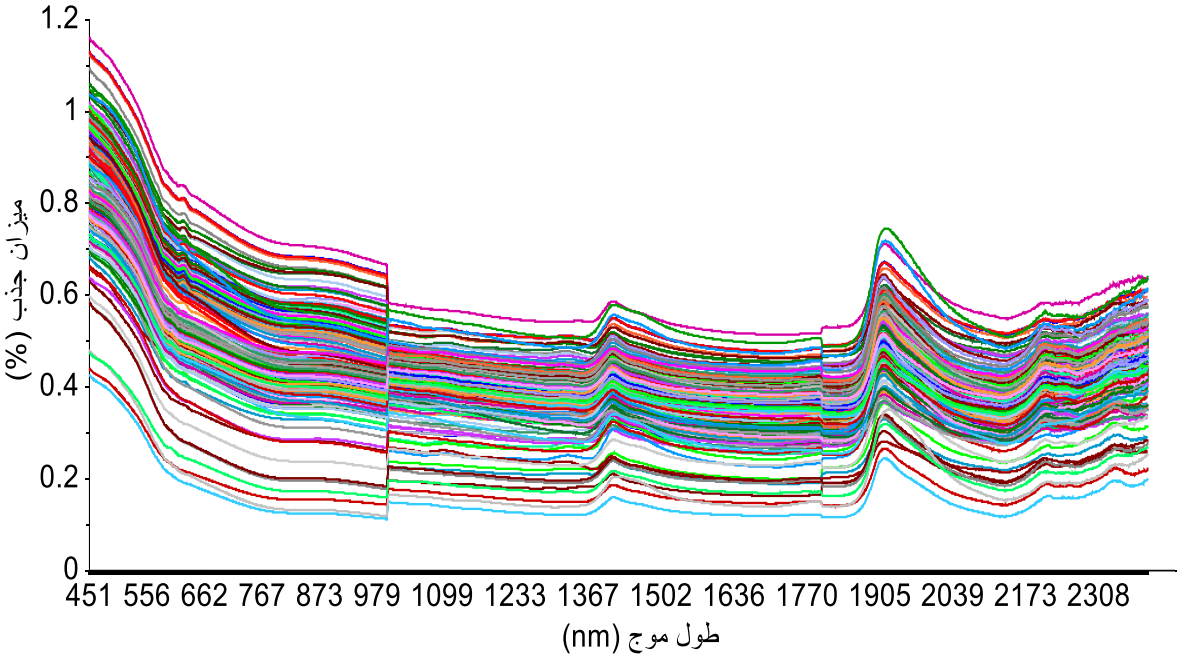
	درصد رطوبت حجمی	درصد کربنات کلسیم معادل	درصد رس	هدایت الکتریکی	درصد کربن آلی	واکنش خاک	درصد شن	درصد سیلت
درصد رطوبت حجمی	۱							
درصد کربنات کلسیم معادل	-۰/۷۷۹**	۱						
درصد رس	۰/۴۱۸**	-۰/۳۲۴**	۱					
هدایت الکتریکی	۰/۷۳۲**	-۰/۵۰۷**	-۰/۰۲۶	۱				
درصد کربن آلی	۰/۰۴۶	-۰/۱۰۱	-۰/۰۷۳	۰/۲۸۳**	۱			
واکنش خاک	-۰/۳۳**	-۰/۶۵۲**	-۰/۱۱۳	۰/۰۵۴	-۰/۲۲۳**	۱		
درصد شن	۰/۶۳۷**	-۰/۷۵۹**	-۰/۱۳۶	-۰/۵۶۱**	۰/۰۵	-۰/۶۱**	۱	
درصد سیلت	-۰/۷۷۸**	۰/۸۵۴**	-۰/۰۱۵	۰/۶۲۸**	-۰/۱۲	۰/۶۲**	-۰/۹۵۴**	۱

**. همبستگی در سطح ۰/۰۱ معنی دار است

ضرایب همبستگی پیرسون بین ویژگی‌های خاک در جدول ۲ نشان داده شده است. بر این اساس، بین درصد رس و درصد سیلت همبستگی منفی معنی داری در سطح ۰/۰۱ وجود دارد. همچنین بین درصد شن و درصد سیلت همبستگی منفی معنی‌داری در سطح ۰/۰۱ وجود دارد. بین درصد رس و درصد رطوبت حجمی همبستگی معنی دار مثبت در سطح ۰/۰۱ و بین درصد رس و درصد کربنات کلسیم معادل، هدایت الکتریکی، و واکنش خاک همبستگی منفی معنی‌داری در سطح ۰/۰۱ وجود دارد. بین درصد شن و درصد رطوبت حجمی در سطح ۰/۰۱ همبستگی مثبت و بین درصد شن و درصد کربنات کلسیم معادل همبستگی منفی در سطح ۰/۰۱ وجود دارد. بین درصد سیلت و درصد کربنات کلسیم معادل، هدایت الکتریکی و واکنش خاک، همبستگی مثبت معنی‌داری در سطح ۰/۰۱ و بین سیلت و درصد رطوبت حجمی همبستگی منفی معنی‌داری در سطح ۰/۰۱ وجود دارد. شکل ۲ و۳ به ترتیب نمودارهای جذب و انعکاس طیف‌های خام به دست آمده از نمونه‌های خاک را نشان می‌دهد. بازتاب خاک به طور کلی در محدوده مرئی پایین و و در محدوده فروسرخ بالا می باشد (Islam et al., 2003).



شکل ۲. نمودار میزان انعکاس طیف‌های خام به دست آمده از همه نمونه‌های خاک



شکل ۳. نمودار میزان جذب طیف‌های خام به دست آمده از همه نمونه‌های خاک

در این مدل، با استفاده از فاکتورهای بهینه مدل، معادلات رگرسیونی بین ویژگی‌های خاک با فاکتورهای بهینه در مدل رگرسیون خطی چندگانه به دست آمدند. فاکتورهای بهینه در مدل رگرسیون حداقل مربعات جزئی برای درصد ذرات خاک، در جدول ۳ نشان داده شده‌اند.

جدول ۳. معادلات رگرسیونی بین ذرات خاک با فاکتورهای بهینه در مدل رگرسیون حداقل مربعات جزئی

ویژگی‌های خاک	روش پیش‌پردازش	فاکتور بهینه	معادله رگرسیونی
رس (٪)	مشتق اول+فیلتر ساویتزکی گلای	۸	$Y=27.82+11.41X_1+655.66X_2+305.83X_3+699.735X_4+1281.48X_5+1575.13X_6+660.52X_7+725.04X_8+3.49$
	تصحیح پخشیده چندگانه	۸	$Y=27.82+1.32X_1+10.96X_2+7.67X_3+10.75X_4+10.42X_5+26.93X_6+27.87X_7+41.20X_8+6.01$
	متغیر نرمال استاندارد	۷	$Y=27.82+0.18X_1+1.49X_2+0.99X_3+0.98X_4+0.93X_5+4.36X_6+4.73X_7+6.44$
شن (٪)	مشتق اول+فیلتر ساویتزکی گلای	۳	$Y=27.33+626.18X_1+876.55X_2+2156.91X_3+57.53$
	تصحیح پخشیده چندگانه	۶	$Y=27.33+9.95X_1+14.63X_2+10.43X_3+28.48X_4+50.38X_5+29.92X_6+48.45$
	متغیر نرمال استاندارد	۶	$Y=27.33+1.31X_1+1.8X_2+1.45X_3+3.46X_4+6.5X_5+3.44X_6+48.27$
سیلت (٪)	مشتق اول+فیلتر ساویتزکی گلای	۴	$Y=42.18+440.62X_1-488.86X_2+1800.53X_3+678.29X_4+12.47$
	تصحیح پخشیده چندگانه	۲	$Y=42.18+7.4X_1+4.68X_2+16.75$
	متغیر نرمال استاندارد	۲	$Y=42.18+0.97X_1+0.67X_2+16.33$

نتایج مدل‌سازی رس، شن و سیلت خاک با استفاده از آماره‌های مقدار ریشه میانگین مربعات خطا (RMSE) (معادله ۱)، ضریب تبیین (معادله ۲) و RPD بررسی گردید. جداول ۴، ۵ و ۶ نتایج مدل‌سازی رس، شن و سیلت خاک درمنطقه مورد مطالعه را نشان می‌دهند. شکل ۴ رابطه بین درصد رس اندازه‌گیری شده و پیش‌بینی شده با طیف خاک در مدل PLSR و با روش پیش‌پردازش تصحیح پخشیده چندگانه را نشان می‌دهد.

جدول ۴. نتایج مدل‌سازی درصد رس درمنطقه مورد مطالعه

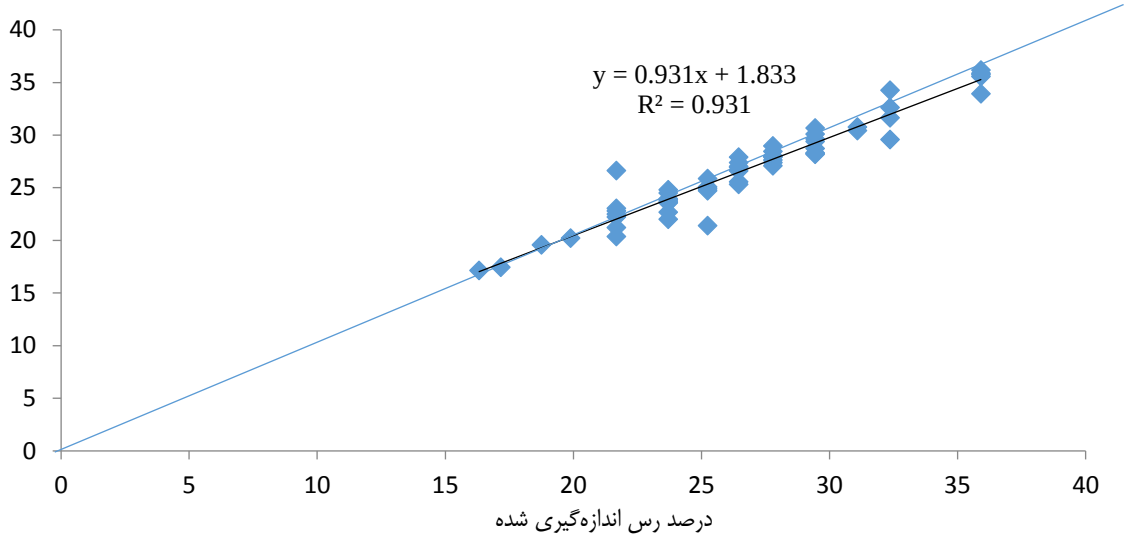
ویژگی‌های خاک	مدل	روش پیش‌پردازش	کالیبراسیون			اعتبارسنجی		
			R^2	RMSE	RPD	R^2	RMSE (g/Kg)	RPD
رس	رگرسیون حداقل مربعات جزئی	بدون پیش‌پردازش	۰/۷۵	۲/۵۳	۱/۹۳	۰/۶۹	۲/۸۴	۱/۷۲
		مشتق اول+فیلتر ساویتزکی گلای	۰/۸۷	۱/۸۱	۲/۷۱	۰/۹۱	۱/۲۹	۳/۸
		تصحیح پخشیده چندگانه	۰/۷۸	۲/۳۷	۲/۱	۰/۹۳	۱/۲۳	۳/۴
		متغیر نرمال استاندارد	۰/۷۶	۲/۴۶	۱/۹۹	۰/۹۲	۱/۳۳	۳/۷
	ماشین بردار پشتیبان	بدون پیش‌پردازش	۰/۱۸	۴/۷۱	۱/۰۴	۰/۱۱	۴/۹	۱
		مشتق اول+فیلتر ساویتزکی گلای	۰/۸	۲/۳۶	۲/۱	۰/۸۸	۱/۹۷	۲/۴۹
		تصحیح پخشیده چندگانه	۰/۷	۲/۸۲	۱/۷۴	۰/۶۶	۳/۱۱	۱/۵۸
		متغیر نرمال استاندارد	۰/۷۱	۲/۷۶	۱/۷۸	۰/۶۵	۳/۱۴	۱/۵۶
	شبکه عصبی مصنوعی	بدون پیش‌پردازش	۰/۱	۴/۱	۱/۱۹	۰/۱	۴/۱	۱/۱۹
		مشتق اول+فیلتر ساویتزکی گلای	۰/۷۷	۲/۱۲	۲/۳۱	۰/۶۷	۳/۲۵	۱/۵۱
		تصحیح پخشیده چندگانه	۰/۹۴	۱/۲۱	۳/۷۷	۰/۶۸	۲/۶۵	۱/۸۵
		متغیر نرمال استاندارد	۰/۹۳	۱/۳	۳/۷۷	۰/۷۵	۲/۵۱	۱/۹۵

جدول ۵. نتایج مدل‌سازی درصد شن درمنطقه مورد مطالعه

ویژگی‌های خاک	مدل	روش پیش‌پردازش	کالیبراسیون			اعتبارسنجی		
			R^2	RMSE	RPD	R^2	RMSE (g/Kg)	RPD
شن	رگرسیون حداقل مربعات جزئی	بدون پیش‌پردازش	۰/۹۱	۴/۵۴	۳/۲۹	۰/۹	۴/۷۹	۳/۱۲
		مشتق اول+فیلتر ساویتزکی گلای	۰/۷۶	۷/۴۸	۲	۰/۸۷	۲/۷۵	۵/۴۵
		تصحیح پخشیده چندگانه	۰/۸۱	۶/۷۸	۲/۲۱	۰/۹۵	۱/۶۳	۹/۱۹
		متغیر نرمال استاندارد	۰/۸۱	۶/۷۷	۲/۲۱	۰/۹۴	۱/۸	۸/۳۲
	ماشین بردار پشتیبان	بدون پیش‌پردازش	۰/۸۳	۶/۴۱	۲/۳۳	۰/۸۱	۶/۷	۲/۲۳
		مشتق اول+فیلتر ساویتزکی گلای	۰/۸۳	۶/۳۷	۲/۳۵	۰/۹۲	۲/۳۵	۶/۳۷
		تصحیح پخشیده چندگانه	۰/۷۹	۷/۱۱	۲/۱۱	۰/۸۲	۳/۴۳	۴/۳۷
		متغیر نرمال استاندارد	۰/۸	۶/۹۱	۲/۱۷	۰/۸۱	۳/۴۸	۴/۳
	شبکه عصبی مصنوعی	بدون پیش‌پردازش	۰/۸	۶/۱	۲/۴۵	۰/۸	۶/۱	۲/۴۵
		مشتق اول+فیلتر ساویتزکی گلای	۰/۸۴	۶/۲۷	۲/۳۹	۰/۸۱	۵/۸۸	۲/۵۵
		تصحیح پخشیده چندگانه	۰/۷۱	۸/۰۴	۱/۸۶	۰/۷۳	۷/۷	۱/۹۵
		متغیر نرمال استاندارد	۰/۸۴	۶/۲۴	۲/۴	۰/۸۳	۵/۶۱	۲/۶۷

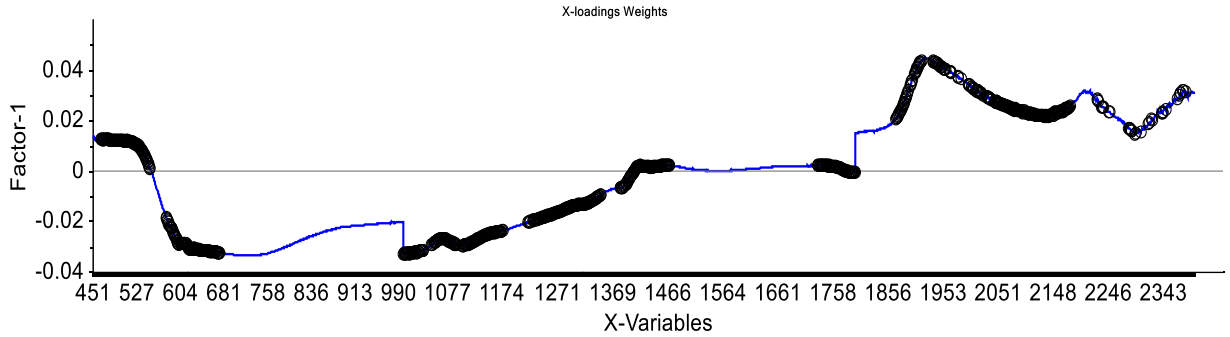
جدول ۶- نتایج مدل‌سازی درصد سیلت درمنطقه مورد مطالعه

ویژگی‌های خاک	مدل	روش پیش‌پردازش	کالیبراسیون			اعتبارسنجی		
			R^2	RMSE	RPD	R^2	RMSE (g/Kg)	RPD
سیلت	رگرسیون حداقل مربعات جزئی	بدون پیش‌پردازش	۰/۹۷	۱/۸۳	۵/۳۳	۰/۹۶	۱/۹۱	۵/۱
		مشتق اول+فیلتر ساویتزکی گلای	۰/۸۸	۳/۴۷	۲/۸۱	۰/۹۵	۱/۰۸	۹/۰۴
		تصحیح پخشیده چندگانه	۰/۸۴	۴/۰۵	۲/۴۱	۰/۸۷	۱/۶۸	۵/۸۱
		متغیر نرمال استاندارد	۰/۸۴	۳/۹۹	۲/۴۵	۰/۸۷	۱/۶۷	۵/۸۴
	ماشین بردار پشتیبان	بدون پیش‌پردازش	۰/۹۴	۲/۶۵	۳/۶۸	۰/۹۲	۲/۸۵	۳/۴۲
		مشتق اول+فیلتر ساویتزکی گلای	۰/۹۲	۲/۸۵	۳/۴۲	۰/۸۷	۱/۸	۵/۴۲
		تصحیح پخشیده چندگانه	۰/۹	۳/۲۴	۳/۰۱	۰/۸۷	۱/۷	۵/۷۴
		متغیر نرمال استاندارد	۰/۹	۳/۱۹	۳/۰۶	۰/۸۷	۱/۶۹	۵/۷۸
	شبکه عصبی مصنوعی	بدون پیش‌پردازش	۰/۹	۱/۱	۸/۸۷	۰/۹	۱/۱	۸/۸۷
		مشتق اول+فیلتر ساویتزکی گلای	۰/۸۵	۳/۷۲	۲/۶۲	۰/۷۴	۴/۷۶	۲/۰۵
		تصحیح پخشیده چندگانه	۰/۹۸	۱/۰۸	۹/۰۴	۰/۹۲	۳/۱۳	۳/۱۲
		متغیر نرمال استاندارد	۰/۸۳	۴/۲۲	۲/۳۱	۰/۷۴	۴/۰۹	۲/۳۹

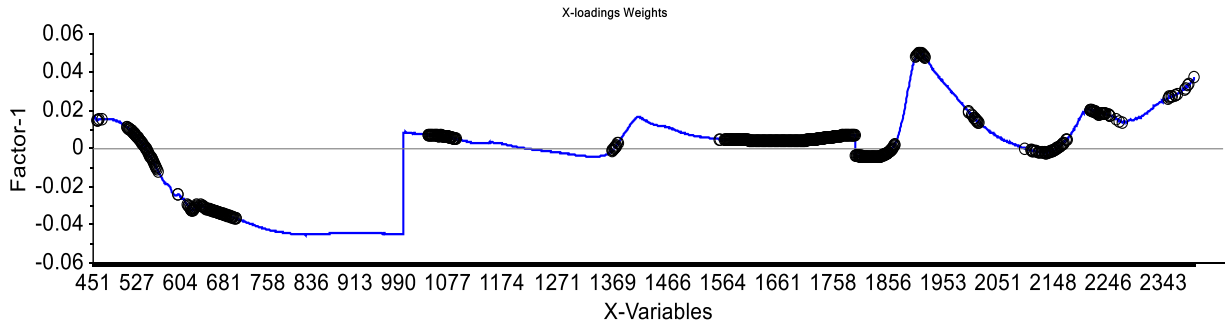


شکل ۴. رابطه بین درصد رس اندازه‌گیری شده و پیش‌بینی شده با طیف خاک در مدل PLSR و با روش پیش‌پردازش تصحیح پخشیده چندگانه

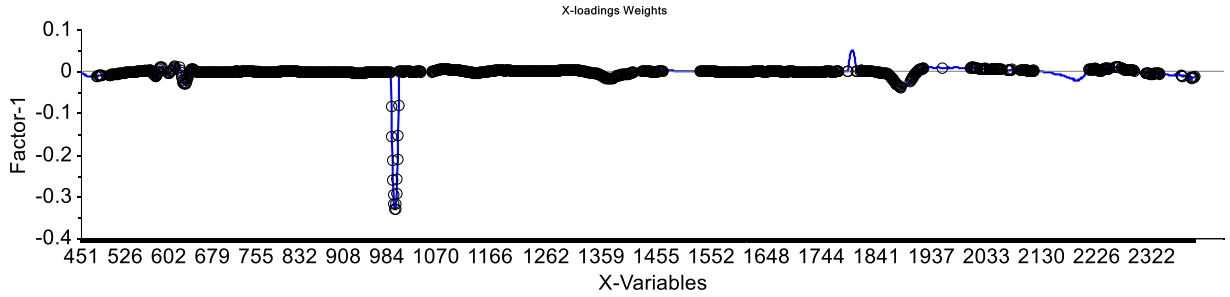
شکل‌های ۵، ۶ و ۷ به ترتیب همبستگی رس و شن با طیف خاک در هر طول موج در روش رگرسیون حداقل مربعات جزئی و روش پیش‌پردازش تصحیح پخشیده چندگانه و همبستگی سیلت با با طیف خاک در هر طول موج در روش رگرسیون حداقل مربعات جزئی و روش پیش‌پردازش تصحیح پخشیده چندگانه را نشان می‌دهند. نقاط مشخص شده، نواحی مهم طیف در برآورد درصد ذرات خاک را نشان می‌دهد. در این پژوهش در مدل حداقل مربعات جزئی، نتایج آزمون عدم قطعیت مارتین طول موج‌های مهم در برآورد رس، شن و سیلت خاک با استفاده از مدل حداقل مربعات جزئی و انواع روش‌های پیش‌پردازش را نشان داد. این باندها مربوط به OH آب در طول موج ۱۴۰۰ و ۱۹۰۰ نانومتر و Al-OH شبکه رس در طول موج‌های ۲۲۰۰ و ۲۲۵۰ نانومتر هستند (Islam et. al., 2003).



شکل ۵. همبستگی رس با طیف خاک در هر طول موج در روش رگرسیون حداقل مربعات جزئی و روش پیش‌پردازش تصحیح پخشیده چندگانه



شکل ۶. همبستگی شن با با طیف خاک در هر طول موج در روش رگرسیون حداقل مربعات جزئی و روش پیش‌پردازش تصحیح پخشیده چندگانه



شکل ۷. همبستگی سیلت با با طیف خاک در هر طول موج در روش رگرسیون حداقل مربعات جزئی و روش مشتق اول و فیلتر ساویتزکی گلای

بحث

هرچه مقدار ریشه میانگین مربعات خطا (RMSE) کمتر و ضریب تبیین بیشتری باشد، مدل برآورد ویژگی خاک را بهتر نشان می‌دهد. ضریب تبیین بین ۰/۶۶ و ۰/۸ دقت پیش‌بینی کم مدل را نشان می‌دهد. مقدار ضریب تبیین بین ۰/۸۱ تا ۰/۹ پیش‌بینی خوب مدل را نشان می‌دهد و ضریب تبیین بیش از ۰/۹ بیانگر پیش‌بینی عالی مدل است (Mali et al., 2004). RPD کمتر از ۱/۵ نشان دهنده پیش‌بینی ضعیف مدل است. اگر RPD بین ۱/۵ تا ۲ باشد، پیش‌بینی مدل قابل قبول و RPD بیشتر از ۲ پیش‌بینی عالی مدل را نشان می‌دهد (Summers et al., 2011).

رس

با توجه به جدول ۴ به طور کلی نتایج نشان داده است که در برآورد درصد رس، در هر دو گروه واسنجی و اعتبارسنجی، با استفاده از روش‌های پیش‌پردازش طیف‌ها ضرایب تبیین بالاتر و مقدار ریشه دوم میانگین مربعات خطا کمتری و در نتیجه نتایج بهتری نسبت به زمانی که از روش‌های پیش‌پردازش استفاده نشد، به دست آمده است. نتایج پژوهش‌های مختلف نشان دهنده تأثیر روش‌های پیش‌پردازش بر نتایج مدل رگرسیونی می‌باشد (Kuśnierek, 2011; Gras et al., 2014; Nawar et al., 2016). Kuśnierek (2011) نتیجه گرفت دقت مدل‌سازی با استفاده از روش‌های مختلف پیش‌پردازش در مقایسه با شرایط عدم استفاده از روش‌های پیش‌پردازش به میزان ۳۰ درصد افزایش یافت. همچنین Nawar et al., (2016) در تحقیقات خود بهترین نتایج را با استفاده از پیش‌پردازش طیف‌ها در مقایسه با شرایط بدون پیش‌پردازش به دست آوردند.

در گروه واسنجی، بهترین نتایج برآورد درصد رس در روش شبکه عصبی مصنوعی با پیش‌پردازش تصحیح پخشیده چندگانه با بیشترین مقدار ضریب تبیین (۰/۹۴) و کمترین مقدار ریشه دوم میانگین مربعات خطا (۱/۲۱) به دست آمد. با توجه به این که مقدار RPD در مدل حداقل مربعات جزئی و پیش‌پردازش‌های مشتق اول و فیلتر ساویتزکی گلای و تصحیح پخشیده چندگانه، مدل ماشین بردار پشتیبان و پیش‌پردازش مشتق اول و فیلتر ساویتزکی گلای و در مدل شبکه عصبی مصنوعی با هر سه روش پیش‌پردازش پیش‌پردازش‌های مشتق اول و فیلتر ساویتزکی گلای، تصحیح پخشیده چندگانه و متغیر نرمال استاندارد بیشتر از ۲ به دست آمد، پیش‌بینی این مدل‌ها عالی بود. در حالی که RPD در مدل حداقل مربعات جزئی بدون پیش‌پردازش و پیش‌پردازش متغیر نرمال استاندارد، مدل ماشین بردار پشتیبان و پیش‌پردازش تصحیح پخشیده چندگانه و متغیر نرمال استاندارد بین ۱/۵ تا ۲ به دست آمد. بنابراین دقت پیش‌بینی این مدل‌ها قابل قبول بوده است. در مدل ماشین بردار پشتیبان و شبکه عصبی مصنوعی بدون پیش‌پردازش با RPD زیر ۱/۵، پیش‌بینی مدل ضعیف بوده است.

در گروه اعتبارسنجی، بهترین نتایج برآورد درصد رس در روش رگرسیون حداقل مربعات جزئی با روش پیش‌پردازش تصحیح پخشیده چندگانه با بیشترین مقدار ضریب تبیین (۰/۹۳) و کمترین مقدار ریشه دوم میانگین مربعات خطا (۱/۲۳) به دست آمد. همچنین با توجه به این که مقدار RPD در مدل حداقل مربعات جزئی و سه روش پیش‌پردازش و ماشین بردار پشتیبان و پیش‌پردازش مشتق اول و فیلتر ساویتزکی گلای بیشتر از ۲ به دست آمد، پیش‌بینی این مدل‌ها عالی بود. پیش‌بینی مدل حداقل مربعات جزئی بدون پیش‌پردازش، مدل ماشین بردار پشتیبان و پیش‌پردازش تصحیح پخشیده چندگانه و متغیر نرمال استاندارد و شبکه عصبی مصنوعی با پیش‌پردازش تصحیح پخشیده چندگانه با RPD بین ۱/۵ تا ۲ قابل قبول به دست آمد. در مدل ماشین بردار پشتیبان و شبکه عصبی مصنوعی بدون پیش‌پردازش با RPD زیر ۱/۵، پیش‌بینی مدل ضعیف بوده است.

شن

در گروه واسنجی، بهترین نتایج برآورد درصد شن در روش حداقل مربعات جزئی بدون پیش‌پردازش با مقدار ضریب تبیین (۰/۹۱) و مقدار ریشه دوم میانگین مربعات خطا (۴/۵۴) به دست آمد. با توجه به این که مقدار درصد انحراف نسبی در روش شبکه عصبی مصنوعی با پیش‌پردازش تصحیح پخشیده چندگانه ۱/۵ تا ۲ به دست آمد بنابراین دقت پیش‌بینی این مدل قابل قبول بوده است. در حالی که RPD در سایر روش‌ها بیش از ۲ بود که نشان دهنده برای پیش‌بینی کاملاً مناسب این مدل‌ها هستند.

در گروه اعتبارسنجی، بهترین نتایج برآورد درصد شن در روش رگرسیون حداقل مربعات جزئی با پیش‌پردازش تصحیح پخشیده چندگانه با مقدار ضریب تبیین (۰/۹۵) و مقدار ریشه دوم میانگین مربعات خطا (۱/۶۳) به دست آمد. با توجه به این که مقدار درصد انحراف نسبی در روش شبکه عصبی مصنوعی با پیش‌پردازش تصحیح پخشیده چندگانه ۱/۵ تا ۲ به دست آمد. بنابراین دقت پیش‌بینی این مدل قابل قبول بوده است. در حالی که RPD در سایر روش‌ها بیش از ۲ بود که نشان دهنده برای پیش‌بینی عالی این مدل‌ها هستند.

سیلت

نتایج برآورد درصد سیلت در گروه واسنجی نشان داد بهترین نتایج در روش شبکه عصبی مصنوعی با پیش‌پردازش تصحیح پخشیده چندگانه با بیشترین مقدار ضریب تبیین (۰/۹۸) و کمترین مقدار ریشه دوم میانگین مربعات خطا (۱/۰۸) به دست آمده است. با توجه به این که مقدار RPD در همه روش‌ها بیشتر از ۲ به دست آمد، پیش‌بینی همه مدل‌ها عالی بود.

نتایج برآورد درصد سیلت در گروه اعتبارسنجی نشان داد اگر چه در روش رگرسیون حداقل مربعات جزئی بدون پیش‌پردازش بیشترین مقدار ضریب تبیین (۰/۹۶) با مقدار ریشه دوم میانگین مربعات خطا (۱/۹۱) به دست آمد، در روش رگرسیون حداقل مربعات جزئی و پیش‌پردازش مشتق اول و فیلتر ساویتزکی گلای با مقدار ضریب تبیین (۰/۹۵)، کمترین مقدار ریشه دوم میانگین مربعات خطا (۱/۰۸) به دست آمد. Reeves et al., (2002) در پژوهشی بیان کردند که نتایج تخمین ممکن است بر مبنای روش‌های مختلف پردازش آماری داده‌ها و مشتقات طیف‌ها تغییر کند. فرایند مشتق اول برای حذف اثرات اندازه ذرات و پراکنش نوری مناسب است (Tsai. and Philpot, 1998).همچنین با توجه به این که مقدار RPD در همه مدل‌ها و روش‌های پیش‌پردازش بیشتر از ۲ به دست آمد، پیش‌بینی همه مدل‌ها عالی بود.

مقایسه عملکرد مدل در گروه واسنجی و اعتبارسنجی

به طور کلی در این مطالعه، در برآورد درصد رس و سیلت در گروه واسنجی، مدل شبکه عصبی مصنوعی با روش پیش‌پردازش متغیر نرمال استاندارد و در برآورد درصد شن مدل رگرسیون حداقل مربعات جزئی بدون پیش‌پردازش دقت بالاتری را نسبت به بقیه مدل‌ها نشان داد. Mouazen et al., (2010) به این نتیجه رسیدند که بهترین دقت پیش‌بینی کربن آلی خاک با شبکه عصبی مصنوعی به دست آمده است. Zhao et. al., (2009) نتایج قابل قبولی از شبکه عصبی برای پیش‌بینی بافت خاک به دست آوردند. Zhu et al., (2007) نتیجه گرفتند که مدل شبکه عصبی مصنوعی نسبت به رگرسیون خطی چند متغیره و مدل‌های دیگر با رابطه توانی کارایی بهتری در پیش‌بینی غلظت رسوب داشته است.

بهترین مدل برای برآورد رس، سیلت و شن در گروه اعتبارسنجی، مدل رگرسیون حداقل مربعات جزئی بوده است. همچنین بهترین روش پیش‌پردازش در گروه اعتبارسنجی برای برآورد درصد رس و شن، روش تصحیح پخشیده چندگانه و برای سیلت روش مشتق اول و فیلتر ساویتزکی گلای بود. Curcio et al., (2013) از دو روش حذف پیوسته و رگرسیون حداقل مربعات جزئی برای برآورد درصد ذرات خاک استفاده کردند. نتایج نشان داد که روش رگرسیون حداقل مربعات جزئی به ترتیب با ضرایب تبیین ۰/۸۷، ۰/۶ و ۰/۸ و مجذور میانگین مربعات خطا ۵/۸، ۷/۲ و ۷/۷ دقت قابل قبولی در برآورد رس، سیلت و شن داشته است. Dotto et al., (2014) برای برآورد درصد ذرات خاک از روش طیف‌سنجی و رگرسیون خطی چندگانه استفاده کردند و برای شن و سیلت ضرایب تبیین ۰/۷۲ و ۰/۷۷ را به دست آوردند. Silva et al., (2016) برای تعیین درصد ذرات خاک از روش طیف‌سنجی با روش رگرسیون حداقل مربعات جزئی و روش پیش‌پردازش مشتق دوم استفاده کردند و برای برآورد درصد ذرات شن، سیلت و رس به ترتیب ضرایب تبیین ۰/۳، ۰/۵۹ و ۰/۶۹ و RMSE ۵/۴۷، ۵/۱۸ و ۵/۳۹ را به دست آوردند. Lacerda et al., (2016) برای پیش‌بینی بافت خاک از روش رگرسیون حداقل مربعات جزئی استفاده کردند و مقدار ضریب تبیین را برای شن و رس به ترتیب برابر با ۰/۹۶ و ۰/۹۳ به دست آوردند.

نتیجه‌گیری و پیشنهادها

در این مطالعه، با توجه به نتایج بر اساس آماره‌های مقدار ریشه میانگین مربعات خطا، ضریب تبیین و شاخص انحراف نسبی در برآورد درصد ذرات خاک به روش طیف‌سنجی، به طور کلی بهترین مدل برای برآورد رس، سیلت و شن براساس گروه اعتبارسنجی، مدل رگرسیون حداقل مربعات جزئی بوده است. همچنین بهترین روش پیش‌پردازش در گروه اعتبارسنجی، برای برآورد درصد رس و شن روش تصحیح پخشیده چندگانه و برای سیلت روش مشتق اول و فیلتر ساویتزکی گلای بود. نتایج نشان داده است که در برآورد درصد رس، در هر دو گروه واسنجی و اعتبارسنجی، با استفاده از روش‌های پیش‌پردازش طیف‌ها ضرایب تبیین بالاتر و مقدار ریشه دوم میانگین مربعات خطا کمتری و در نتیجه نتایج بهتری نسبت به زمانی که از روش‌های پیش‌پردازش استفاده نشد، به دست آمده است. به طور کلی نتایج این پژوهش کارایی روش طیف‌سنجی مرئی مادون قرمز نزدیک را در برآورد درصد ذرات خاک نشان می‌دهد. با توجه به این که نوع مدل‌ها و روش پیش‌پردازش بر نتایج حاصل از روش طیف‌سنجی تأثیر داشته است، پیشنهاد می‌شود که مدل‌های دیگر و روش‌های دیگر پیش‌پردازش طیف‌ها برای برآورد اجزای بافت خاک در روش طیف‌سنجی استفاده شوند و نتایج حاصل با هم مقایسه گردد.

سپاسگزاری

هزینه‌های اجرای این پژوهش توسط معاونت پژوهشی دانشگاه شهید چمران اهواز (پژوهانه شماره SCU.AS1401.365) پرداخت شده است.

"هیچ‌گونه تعارض منافع بین نویسندگان وجود ندارد"

منابع

جلالیان، احمد (۱۳۷۶). *مطالعات تعیین منابع و قابلیت اراضی منطقه سمیرم*. وزارت جهاد سازندگی، جهاد سازندگی استان اصفهان. ۱۶۰-۱۶۵.

چترنور، منصور.، لندی، احمد.، فرخیان فیروزی، احمد.، نوروزی، علی اکبر. و بهرامی، حسینعلی. مدل‌سازی رفتار طیفی بافت خاک در کانون‌های ریزگرد استان خوزستان با استفاده از تصاویر ابر طیفی و مدل جنگل تصادفی. *زمین شناسی کاربردی پیشرفته*. ۹(۴): ۴۷۹-۴۶۶.

رحمتی، فاطمه؛ حجتی، سعید؛ رنگزن، کاظم و لندی، احمد (۱۴۰۱). بررسی کارایی روش طیف‌سنجی مرئی-مادون قرمز نزدیک در تخمین برخی ویژگی‌های خاک منطقه‌ی سمیرم اصفهان. *نشریه آب و خاک*. ۳۶ (۲): ۲۸۳-۳۰۰.

رسولی، نجمه؛ فرپور، محمد هادی؛ خیامیم، فاطمه؛ و رنجبر، حجت‌الله (۱۳۹۷). پیش‌بینی برخی ویژگی‌های خاک به روش طیف‌سنجی مرئی-مادون قرمز نزدیک در منطقه بردسیر کرمان. *نشریه پژوهش‌های خاک (علوم خاک وآب)*، الف، ۳(۲): ۲۳۱-۲۴۳.

مهرابی گوهری، الهام.، متین فر، حمیدرضا.، تقی زاده مهرجردی، روح‌الله. و جعفری، اعظم. ۱۳۹۹. قابلیت طیف‌سنجی مرئی مادون قرمزنزدیک (VIS-NIR) در پیش‌بینی درصد ذرات خاک با استفاده از شبکه عصبی مصنوعی و رگرسیون حداقل مربعات جزئی. *نشریه آب و خاک (علوم و صنایع کشاورزی)*، شماره ۳۴(۳): ۶۳۵-۶۲۳.

REFERENCES

Ahmadi, A.; Emami, M.; Daccache, A. & He, L. (2021). Soil Properties Prediction for Precision Agriculture Using Visible and Near-Infrared Spectroscopy: A Systematic Review and Meta-Analysis. *Agronomy*, 11(433): 1-14.

Almeida, J., & Predictive, S. (2002). Non-linear modeling of complex data by artificial neural networks. *Current Opinion in Biotechnology*, 13: 72-6.

Bouyoucos, G. J. (1951). A recalibration of hydrometer method for making mechanical analysis of soil. *Agronomy*, 43: 434-438.

Camo, A. 1998. The Unscrambler User Manual. CAMO ASA Norway.

Chaternour, M., Landi, A., Farrokhian Firouzi, A., Noroozi, A.A., Bahrami, H.A. 2020. Spectral behavior modeling of soil texture over dust center of Khuzestan Province using hyperspectral images and Random Forest (RF) model. *Advanced applied Geology*, 9 (4): 466-479.

Chin, W. W., Marcolin, B., & Newsted, P. (1996). A partial least squares latent variable modeling approach for measuring interaction effects: results from a Monte Carlo simulation study and voice mail emotion/adoption study, In proceeding of the 17th international conference on information systems, 16-18 Dec. 1996, Cleveland, Ohio, 21-41.

Clark, R.N. 1999. Spectroscopy of rocks and minerals, principles of spectroscopy. John Wiley and Sons.

Curcio, D. G., Ciraolo, F. & Minacapilli, M. (2013). Prediction of soil texture distributions using VNIR-SWIR reflectance spectroscopy. *Procedia Environmental Sciences*, 19: 494-503.

Dotto, A. C., Dalmolin, R. S. D., Pedron, F. d. A., Caten, A. T., & Ruiz, L. F. C., (2014). Digital mapping of soil properties: particle size and soil organic matter by diffuse reflectance spectroscopy. *Revista Brasileira de Ciência do Solo*, 38: 6. 1663-1671.

Gras J.P., Barthès, B.G. Mahaut B. & Trupin. S. (2014). Best practices for obtaining and processing field visible and near infrared (VNIR) spectra of topsoil. *Geoderma*. 215: 126–134.

Islam, K., Singh, B., & McBratney, A. (2003). Simultaneous estimation of several soilproperties by ultraviolet, visible and near infrared reflectance spectroscopy. *Australian Journal of Soil Research*, 41: 1193–1202.

Jalalian, A. (1997). *The studies of land resources and capability determination in Semirom area*. The Ministry of Jahad Sazandegi, Isfahan Province. (in Persian)

Janik, L. J., Forrester, S. T. & Rawson, A. (2009). The prediction of soil chemical and physical properties from mid-infrared spectroscopy and combined partial least-squares regression and neural networks (PLS-NN) analysis. *Chemometrics and Intelligent Laboratory Systems*, 97: 2. 179-188.

Karayiannis N. B., & Venetsanopouios A. N. (1993). *Artificial Neural Network: learning algorithms, performance evaluation, & application*. Kluwer academic publisher. boston.

Kuśnierek, K. (2011). Pre-processing of soil visible and near infrared spectra taken in laboratory and field conditions to improve the within-field soil organic carbon multivariate calibration, The Second Global Workshop on Proximal Soil Sensing, Montreal, Canada, 100-103.

Lacerda, M. P. C., Demattê, J. A. M., Sato, M. V., Fongaro, C. T., Gallo, B. C., & Souza, A. B. (2016). Tropical Texture Determination by Proximal Sensing Using a Regional Spectral Library and Its Relationship with Soil Classification. *Remote Sensing*, 8: 701. 1-20.

Lazaar, A., Pradhan, B., Naiji, Z., Gourfi, A., El Hammouti, K., Andich, K., & Monir, A. (2021). The manifestation of VIS-NIRS spectroscopy data to predict and map soil texture in the Triffa plain (Morocco). *Kuwait Journal of Science*, 48: 1. 127-137.

Mehrabi Gohari1, E. Matinfar, H.R. Taghizadeh-Mehrjardi, R.A. & Jafari, A. 2022. Visible-Near Infrared (VIS-NIR) Spectrophotometry in Predicting Soil Particle Percentage Using Artificial Neural Network and Partial Least Squares Regression. *Journal of Water and Soil*, 34(3): 623-635.

Lanyon, L.E. & Heald W. R. (1982). Magnesium, calcium, strontium and barium. P. 247-260. In: A. L., Page et al. (ed.), Methods of Soil Analysis. Part2, Agron. Monogr. ASA and SSSA, Madison, Wisconsin.

Minasny, B., & McBratney, A. B. (2006). A conditioned Latin hypercube method for sampling in the presence of ancillary information. *Computers and Geosciences*, 32: 9. 1378-1388.

Mouazen, A. M., Kuang, B., D. E. Baerdemaeker & Ramon, H. (2010). Comparison between principal component, partial least squares and back propagation neural network analyses for accuracy of measurement of selected soil properties with visible and near infrared spectroscopy. *Geoderma*, 158: 23-3.

Nelson, R.E. (1982). Carbonate and gypsum. In: A. L. Page et al. (eds.), Methods of Soil Analysis (2nd ed). Part 2, Agronomy Monogaraph. No: 9. ASA and SSSA, Madison, Wisconsin. 181-196.

Nawar, S., Buddenbaum, H., Hill, J., Kozak, J., & Mouazen, A. (2016). Estimating the soil clay content and organic matter by means of different calibration methods of vis-NIR diffuse reflectance spectroscopy. *Soil and Tillage Research*, 155: 510-522.

Richards, L.A. (1954). Diagnosis and Improvement of Saline-Alkali Soils.US Departent of Agriculture, Washington DC.

Rahmati, F., Hojati, S., Rangzan K., & Landi, A. (2022). Investigating the Efficiency of Visible-Near Infra-Red (NIR) Spectrometry to Estimate Selected Soil Properties in Semirom Area, Isfahan. *Journal of Water and Soil*, 36(2): 283-300. (In Persian)

Rasooli, N., Farpoor M., 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(2): 231-243. (in Persian)

Reeves, J., McCarty G. & Mimmo, T. (2002). The potential of diffuse reflectance spectroscopy for the determination of carbon inventories in soils. *Environmental Pollution*, 116: 277–284.

Rinnan, A., Van den Berg, F., & Engelsen, S. B. (2009). Review of the most common preprocessing techniques for near-infrared spectra. *TrAC Trends in Analytical Chemistry*, 28: 1201–1222.

Tsai, F. and W. Philpot. (1998). Derivative analysis of hyperspectral data. *Remote Sensing Environment*, 66: 41–51.

Sargent D.J. (2001). Comparison of artificial neural networks with other statistical approaches: results from medical data sets. *Cancer*, 91:1636-42.

Silva, E. B., Ten Caten, A., Dalmolin, R. S. D., Dotto, A. C., Silva, W. C., & Giasson, E. (2016). Estimating Soil Texture from a Limited Region of the Visible/Near-Infrared Spectrum Digital Soil Morphometrics, *Springer*, 73-87.

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.

Vapnik, V. and Vapnik, V. (1998). *Statistical learning theory*. Wiley. New York. 156-160.

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: 59–75.

Walkley, A., & Black, I.A. (1934). An examination of the Degtjareff method for determining organic carbon in soils: effect of variations in digestion conditions and of inorganic soil constituents. *Soil Science*, 63: 251-263.

Wilding, L. (1985). Soil Spatial variability: Its documentation, accommodation, and implication to soil surveys. In Soil Spatial Variability. D.R. Nielson and J. Bouma (eds), Pudo, Wagenigen, the Netherlands, 166-194.

Wischmeier, W. H. & Smith, D.D. (1978). *Predicting Rainfall Erosion Losses, a Guide to Conservation Planning*. Agriculture Handbook No. 537. U.S. Department of Agriculture, Washington, DC.

Wold S., Sjostrom M. & Eriksson L. 2001. PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratoary Systems*, 58: 109–130.

Zhao, Z., Chow, T. L., Rees H. W., Yang, Q., Xing, Z., & Meng, F. (2009). Predict soil texture distributions using an artificial neural network model. *Computers and Electronics Agrriculture*, 65: 36-48.

Zhu, Y. M., Lu X. X., & Zhou Y. (2007). Suspended sediment flux modeling with artificial neural network: An example of the Longchuanjiang River in the Upper Yangtze Catchment, China. *Geomorphology*, 84: 111–125.

Estmiating of Soil Particles Percentage Using Visible-Near Infra-Red (NIR) spectrometry in Semirom area, Isfahan

EXTENDED ABSTRACT

Introduction

Soil texture, describing the relative proportion of sand, silt and clay in the mineral phase of soils is a major determinant of its water storage capacity and permeability, aeration, bulk density, aggregate stability and carbon storage capacity. Soil texture measurement on large scales using experimental methods can be extremely time-consuming and expensive, especially when dealing with a high spatial sampling density. Soil Visible and Near-Infra Red (V-NIR) reflectance spectroscopy has proven to be a fast, cost-effective, nondestructive, environmental-friendly, repeatable, and reproducible analytical technique. V-NIR reflectance spectroscopy has been used more than 30 years to predict an extensive variety of soil properties like organic and inorganic carbon, nitrogen, organic carbon, moisture, texture and salinity. The objective of this study was to estimate soil texture using visible near-infrared and short-wave Infrared (SWIR) reflectance spectroscopy (350-2500 nm) in Semirom, Isfahan. In this study, the best predictions of all the soil particles percentage, model and pre-processing technique were also determined. The Partial Least Squares Regression (PLSR), Support Vector Machine Regression and Artificial Neural Network models were also compared to estimate soil texture.

Materials and Methods

A total number of 200 soil samples (0-10 cm) were collected from the Semirom area (51º 17' - 52º 3' E; 30º 42' - 31º 51' N), Isfahan. The samples were air dried and passed through a 2 mm sieve, and soil particles percentage was determined in the laboratory. Reflectance spectra of all samples were measured using an ASD field spectrometer. Different pre-processing methods i.e., First (FD) Derivatives and Savitzky-Golay Filter, Multiplicative Scatter Correction (MSC) and Standard Normal Variable (SNV) were applied and were performed on spectral data. The Partial Least Squares Regression (PLSR), Support Vector Machine Regression and Artificial Neural Network models were used to estimate soil texture. The selection of the best model was done according to the value of the Residual Prediction Deviation (RPD), the coefficient of determination (R^2), and the Root MeanSquare Error (RMSE).

Results and Discussion

Coefficient of Variation (CV) values indicated that the variability of clay and silt were medium. However, the variability of Sand was high. The soil property of best Result was Silt, with excellent values of RPD >2, R^2 =0.98 and RMSE=1.08 using Artificial Neural Network model with MSC pre-processing technique. The results indicated the desirable capability of Artificial Neural Network model with MSC and SNV pre-processing techniques in estimating the Clay (RPD >2, R^2=0.94 and RMSE=1.21) and Sand (RPD >2, R^2=0.84 and RMSE=6.24) contents of the soils, respectively.

Conclusions

In general, based on the results of this study, VNIR spectroscopy was successful in estimating soil particles percentage and showed its potential for substituting laboratory analyses. Further, spectroscopy could be considered as a simple, fast, and low-cost method in predicting soil properties.

Keywords: Artificial Neural Network, Partial Least Squares Regression (PLSR), Pre-processing methods, Spectroscopy, Support Vector Machine Regression.