Prediction of Soil Available Phosphorus Using Visible–Near-Infrared Spectroscopy

Abstract

The physical, chemical, and mechanical properties of soil are crucial for implementing appropriate land management strategies. Among these, phosphorus is one of the most essential nutritional elements for plant growth. The study aimed to (i) evaluate the potential of visible–near-infrared (Vis–NIR) spectroscopy for predicting soil available phosphorus and (ii) investigate the effects of different spectral preprocessing techniques on prediction accuracy. A total of 150 samples were collected from tea farms in Lahijan County and from agricultural fields around the Tehran Refinery (Ismailabad) and the refinery south of Tehran (Ray City). Reflectance visible-near infrared (Vis-NIR) spectrometry for measurement and prediction of soil phosphorus content was studied. To this end, spectra of air-dried soil samples were acquired using a FieldSpec® 3 ASD spectroradiometer equipped with a contact probe, operating in interactance measurement mode over the spectral range of 350–2500 nm. The spectra were preprocessed using Savitzky–Golay smoothing, first (D1) and second (D2) derivatives, Multiplicative Scatter Correction (MSC), Moving Average (MA), and Standard Normal Variate (SNV) transformation. Five positions were measured on each soil sample, with four consecutive scans recorded at each position. The four scans at each position were first averaged, and the resulting five spectra were subsequently averaged to obtain one representative spectrum per soil sample. The results demonstrated that Vis–NIR spectroscopy is capable of predicting soil phosphorus content with satisfactory accuracy. The PLS model combined with SG + MSC preprocessing yielded the best predictive performance (R^2_c = 0.94, RMSEC = 13.04, R^2_{cv} = 0.91, RMSECV = 16.00, MAE=10.94, Bias=-2.96×10^{-5} and RPD = 3.28). Nevertheless, other preprocessing combinations also showed promising potential for phosphorus prediction, indicating the flexibility of the PLS model with different spectral transformations. These results represent laboratory performance for air-dried and sieved soil samples; independent validation under field-moist and in situ conditions is required before direct field application.

Keywords: *Partial least-squares regression, Reflectance Spectroscopy, Spectral pre-processing, Validation*

*Corresponding Author: Professor, Faculty of Agriculture, Tarbiat Modares University, Tehran, Iran. P. O. Box: 14115-336. E-mail: minaee@modares.ac.ir

ویرایش نشده

EXTENDED ABSTRACT

Introduction

Soil is a fundamental natural resource that supports food, fiber, and energy production and plays a key role in water regulation, contaminant filtration, nutrient cycling, carbon storage, and the maintenance of terrestrial ecosystems. The ability of soil to perform these functions depends on its physical, chemical, mechanical, and biological properties, all of which may vary considerably across space and time. Accurate assessment of these properties is therefore necessary for sustainable soil management, site-specific fertilizer application, environmental protection, and the implementation of precision agriculture practices.

Phosphorus is one of the essential macronutrients required for plant growth and development. It is involved in several physiological and biochemical processes, including energy transfer, photosynthesis, respiration, nucleic acid synthesis, root development, seed formation, and crop maturity. Insufficient phosphorus availability can restrict plant growth and reduce crop yield. In contrast, excessive application of phosphorus fertilizers can increase production costs, cause phosphorus accumulation in soil, and increase the risk of phosphorus loss to surface water and groundwater. Such losses may contribute to eutrophication and deterioration of water quality. Consequently, the accurate and timely determination of plant-available soil phosphorus is important for soil fertility assessment and the development of effective fertilizer management programs.

Conventional laboratory methods for measuring soil phosphorus generally provide reliable analytical results. Nevertheless, these methods are often time-consuming and expensive, require trained personnel, involve several sample preparation steps, and depend on the use of chemical reagents. Furthermore, the spatial variability of soil properties requires the collection and analysis of many samples, particularly when soil fertility is evaluated at the field or regional scale. The application of conventional analytical procedures to a large number of samples may therefore be technically and economically challenging. These limitations have encouraged researchers to develop rapid, non-destructive, cost-effective, and environmentally friendly alternatives for soil analysis.

Visible–near-infrared spectroscopy has emerged as a promising technique for rapid soil characterization. This method is based on the interaction between electromagnetic radiation and soil constituents in the visible and near-infrared regions (Nanni & Demattê, 2006; Shepherd & Walsh, 2002).

Soil spectral reflectance is affected by several properties, including soil-moisture and organic-matter contents, texture, particle-size distribution, carbonates, clay minerals, soil color, and iron and aluminum oxides. Calcium carbonate generally increases soil reflectance, whereas soil organic carbon and moisture often reduce reflectance intensity. Some soil properties have direct or relatively distinct absorption features in the Vis–NIR region. In contrast, available phosphorus does not normally exhibit a strong and specific absorption feature over this spectral range. Its prediction is therefore

generally indirect and depends on its relationships with spectrally active soil constituents, such as organic matter, clay minerals, metal oxides, moisture, and other physical and chemical properties.

Raw soil spectra may contain not only useful chemical and physical information, but also unwanted variability caused by instrumental noise, background effects, baseline shifts, light scattering, particle-size differences, and surface heterogeneity. Spectral preprocessing is consequently an important step in chemometric modeling. Preprocessing methods may reduce noise, correct scattering effects, remove baseline variation, enhance overlapping absorption features, and improve the relationship between spectral data and the target soil property. However, the effectiveness of a preprocessing method depends on the sample characteristics, measurement conditions, spectral range, target property, and modeling algorithm. Therefore, no single preprocessing technique can be regarded as universally optimal (Heise & Winzen, 2001; Jamshidi *et al.*, 2014).

Previous studies have demonstrated the potential of Vis–NIR spectroscopy for predicting several physical and chemical soil properties and nutrients (Kodaira & Shibusawa, 2013; Mirzaei *et al.*, 2022; Munnaf *et al.*, 2021; Peng *et al.*, 2019; Rasooli *et al.*, 2023). *Malley et al.* (1999) evaluated 877 soil samples and used principal component regression and partial least-squares regression to predict soil mineral nutrients, including phosphorus. Shao and He (2011) also investigated the predictability of soil nitrogen, phosphorus, and potassium contents using infrared spectroscopy. These studies indicated that soil phosphorus could be estimated spectroscopically, although prediction accuracy varied according to the spectral range, sample population, reference method, preprocessing procedure, and validation approach.

The objective of this study was to evaluate the capability of visible/near-infrared spectroscopy and investigate the effects of different spectral preprocessing methods on the accuracy of available soil phosphorus prediction. Partial least-squares regression was used to establish relationships between soil spectral data and reference phosphorus measurements. The performance of different preprocessing combinations was subsequently compared using calibration and prediction statistics.

Materials and Methods

A total of 150 soil samples were collected from contrasting agricultural environments in northern and central Iran. The dataset comprised 120 samples from four tea farms in Lahijan County, 15 samples from agricultural fields in Ismailabad, and 15 samples from agricultural fields in Ray City. The sampling interval within the selected farms was approximately 50 m. Soil samples were collected from a depth of 20–30 cm, transported to the laboratory, air-dried, gently crushed, and passed through a 2-mm sieve.

Available phosphorus was determined using the Olsen extraction method. Ten grams of soil were mixed with 20 mL of 0.5 mol L^{-1} sodium bicarbonate solution adjusted to pH 8.5 and shaken for 30 min. Following centrifugation and filtration, the phosphorus concentration was determined calorimetrically at 882 nm. Available phosphorus values were expressed in mg kg^{-1}.

Reflectance spectra were acquired using a FieldSpec 3 spectroradiometer equipped with a contact probe. Although the nominal operating range of the instrument was 350–2500 nm, only wavelengths between 400 and 2450 nm were retained because of the relatively high noise at the two ends of the spectral range. Five positions were measured on each soil sample, and four consecutive scans were recorded at each position. The four scans collected at each position were first averaged, and the resulting five spectra were subsequently averaged to obtain one representative spectrum for each sample. Thus, each soil sample constituted only one independent observation in the modeling procedure.

The evaluated preprocessing methods included moving average smoothing (MA), Savitzky–Golay smoothing (SG), multiplicative scatter correction (MSC), standard normal variate transformation (SNV), and first- and second-derivative transformations (D1 and D2). The investigated combinations were MA+MSC, MA+SNV, SG+MSC, SG+SNV, SG+D1, SG+D2, SG+D1+MSC, SG+D1+SNV, SG+D2+MSC, and SG+D2+SNV. A model based on spectra without preprocessing was also developed for comparison.

Principal component analysis was initially applied to the raw spectra to identify potentially anomalous samples. Eleven samples—seven from Lahijan, three from Ismailabad, and one from Ray—were identified as spectral outliers after their spectra and laboratory measurements were re-examined. Consequently, 139 samples were retained for PLS modeling. Model performance was evaluated using full cross-validation, in which each sample was omitted once, the model was fitted to the remaining samples, and the omitted sample was subsequently predicted.

The models were assessed using the calibration coefficient of determination) R_c^2 ,(root mean square error of calibration (RMSEC), cross-validation coefficient of determination) R_{CV}^2 ,(root mean square error of cross-validation (RMSECV), mean absolute error (MAE), bias, and the ratio of performance to deviation (RPD). Cross-validation residuals were calculated as the difference between the cross-validated predicted and reference values:

$$e_i = \hat{y}_{i,CV} - y_i$$

Positive and negative residuals therefore represented overestimation and underestimation, respectively.

Results and Discussion

Available phosphorus concentrations ranged from 1.018 to 234.68 mg kg^{-1}. According to the descriptive statistics presented in the original table, the mean phosphorus concentration was 88.63 mg kg^{-1}, and its standard deviation was 52.52 mg kg^{-1}. The wide concentration range and relatively high standard deviation indicated that the sample population contained adequate phosphorus variability for multivariate calibration.

The raw reflectance spectra showed noticeable differences in reflectance magnitude and spectral shape among the soil samples. These differences were likely related to variations in particle-size distribution, texture, organic carbon, residual moisture, carbonates, mineral composition, and iron and aluminum oxides. Because available phosphorus has no strong and exclusive absorption feature

in the Vis–NIR range, its successful prediction was most likely based on indirect covariance with these spectrally active soil constituents.

The PLS model developed using the SG+MSC preprocessing combination was identified in the original study as the selected model for predicting available soil phosphorus. The model used 10 latent variables and produced R_c^2 of 0.94, RMSEC of 13.04 mg kg⁻¹, R_{cv}^2 of 0.91, RMSECV of 16.00 mg kg⁻¹, MAE=10.94, Bias= -2.96×10^{-5} and RPD of 3.28. These statistics indicated that approximately 91% of the variation in the reference phosphorus values was explained during prediction or cross-validation. An RPD greater than 3 also suggested that the model had good potential for quantitative estimation within the studied sample population.

The MSC transformation was expected to reduce multiplicative and additive spectral variations resulting from light scattering and differences in particle size and sample surface characteristics. Savitzky–Golay smoothing reduced high-frequency noise while preserving the general shape of the spectra. Their combination therefore produced a comparatively accurate and stable model.

The application of derivative preprocessing did not consistently improve prediction performance. The SG+D1 model produced $R_{cv}^2 = 0.83$,RMSECV=21.89 mg kg⁻¹, and RPD=2.40. Although the SG+D1+SNV model achieved the highest calibration coefficient of determination ,$R_c^2 = 0.96$,its prediction statistics were substantially weaker, with $R_{cv}^2 = 0.82$,RMSECV=22.45 mg kg⁻¹, and RPD=2.34. This difference between calibration and prediction performance may indicate overfitting or the extraction of noise-related variation.

Second-derivative preprocessing had the most unfavorable effect on the models. The SG+D2 model produced $R_{cv}^2$of only 0.41 and RMSECV of 40.56 mg kg⁻¹. Similarly, SG+D2+MSC and SG+D2+SNV resulted in R_{cv}^2 values of 0.39 and RMSECV values above 41 mg kg⁻¹. Derivative transformations can improve the separation of overlapping bands and correct baseline effects, but they may also amplify spectral noise, particularly when smoothing parameters are not optimally selected. The poor performance of the second-derivative models suggests that these transformations were unsuitable for the present dataset or that their parameters required further optimization.

The results are comparable to those of Malley et al. (1999), who reported $R_p^2 = 0.92$ and a ratio of 3.44 for phosphorus prediction using PLS over the 400–2500 nm range. Although the present model produced a higher coefficient of determination than that reported by Shao and He (2011), the numerical results are not directly comparable because the studies differed in phosphorus range, reference extraction method, soil population, sample preparation, spectral configuration, and validation strategy. The comparison should therefore be regarded as contextual rather than as evidence of unequivocally superior model performance. In addition to preprocessing procedures and model-validation strategy, differences among studies may result from variations in soil type, phosphorus concentration range, reference analytical method, spectral range, and sample heterogeneity.

Conclusions and Recommendations

Vis–NIR spectroscopy combined with PLS regression showed good potential for predicting available

phosphorus in the studied air-dried soil samples. The SG+MSC model provided strong cross-validation performance with $R^2_{CV} = 0.91$,RMSECV = 16.00 mg kg-1 ,MAE = 10.94 mg kg-1 ,and RPD = 3.28. However, preprocessing did not universally improve prediction accuracy, and second-derivative transformations substantially reduced model performance. Residual analysis also revealed a relative tendency to underestimate phosphorus at the upper end of the concentration range despite an overall bias close to zero.

The developed model should presently be regarded as a complementary tool for rapid laboratory screening rather than a complete replacement for conventional reference analysis. Independent external validation using samples from additional geographic and pedological conditions is required before routine application. Future studies should also include more samples with high phosphorus concentrations, evaluate field-moist and in situ measurements, and compare PLS with nonlinear machine-learning approaches.

ویرایش استاری نشده

پیش‌بینی فسفر قابل جذب خاک با استفاده از طیف‌سنجی مرئی/ فروسرخ نزدیک

چکیده

ویژگی‌های فیزیکی، شیمیایی و مکانیکی خاک نقشی اساسی در اجرای راهبردهای مناسب مدیریت اراضی دارند. در میان این ویژگی‌ها، فسفر یکی از ضروری‌ترین عناصر غذایی برای رشد گیاه به شمار می‌رود. پژوهش حاضر با اهداف (۱) ارزیابی قابلیت طیف‌سنجی مرئی- فروسرخ نزدیک (Vis–NIR) در پیش‌بینی فسفر قابل‌دسترس خاک و (۲) بررسی تأثیر روش‌های مختلف پیش‌پردازش طیفی بر دقت پیش‌بینی انجام شد. بدین منظور، در مجموع ۱۵۰ نمونه از باغ‌های چای شهرستان لاهیجان و مزارع کشاورزی واقع در اطراف پالایشگاه تهران (اسماعیل‌آباد) و پالایشگاه جنوب تهران (شهر ری) جمع‌آوری شد. در این پژوهش، کاربرد طیف‌سنجی بازتابی در محدوده مرئی- فروسرخ نزدیک (Vis–NIR) برای اندازه‌گیری و پیش‌بینی مقدار فسفر خاک بررسی شد. برای این منظور، طیف نمونه‌های خاک خشک‌شده در هوا با استفاده از دستگاه طیف‌سنج FieldSpec® 3 ASD مجهز به پروب تماسی، در حالت اندازه‌گیری برهم‌کنشی و در محدوده طول موج ۳۵۰ تا ۲۵۰۰ نانومتر ثبت شد. طیف‌های به‌دست‌آمده با استفاده از روش‌های هموارسازی ساویتزکی—گولای، مشتق اول (D1)، مشتق دوم (D2)، تصحیح پراکنش چندگانه (MSC)، میانگین متحرک (MA) و تبدیل متغیر نرمال استاندارد (SNV) پیش‌پردازش شدند. از هر نمونه در پنج موقعیت و در هر موقعیت با چهار اسکن متوالی طیف‌برداری شد و میانگین این اندازه‌گیری‌ها به‌عنوان یک طیف نماینده برای هر نمونه در مدل‌سازی به کار رفت. نتایج نشان داد که طیف‌سنجی Vis–NIR قادر است مقدار فسفر خاک را با دقتی رضایت‌بخش پیش‌بینی کند. مدل حداقل مربعات جزئی (PLS) همراه با پیش‌پردازش ترکیبی SG + MSC بهترین عملکرد پیش‌بینی را ارائه داد (R^2_c=۰/۹۴، RMSEC=۱۳/۰۴، R^2_{cv} =۰/۹۱ ، RMSECV =۱۶/۰۰، MAE= ۱۰/۹۴، Bias = −۲/۹۶×۱۰$^{-۵}$ و RPD=۳/۲۸). بااین‌حال، سایر ترکیب‌های پیش‌پردازشی نیز قابلیت امیدوارکننده‌ای برای پیش‌بینی فسفر نشان دادند که بیانگر انعطاف‌پذیری مدل PLS در استفاده از تبدیل‌های طیفی مختلف است. این نتایج مربوط به نمونه‌های هواخشک و الک‌شده در شرایط کنترل‌شده آزمایشگاهی است و کاربرد مدل برای اندازه‌گیری مستقیم در مزرعه به اعتبارسنجی مستقل در خاک‌های مرطوب و شرایط درجا نیاز دارد.

کلیدواژه‌ها: *اعتبارسنجی، پیش‌پردازش طیفی، رگرسیون حداقل مربعات جزئی، طیف‌سنجی انعکاسی*

ویرایشناری نشده

مقدمه

خاک یکی از مهم‌ترین منابع طبیعی و بستر اصلی تولید غذا، الیاف و انرژی است و در تنظیم جریان آب، پالایش آلاینده‌ها، ذخیره عناصر غذایی و تثبیت کربن نقش اساسی دارد. توانایی خاک در انجام این وظایف به ویژگی‌های فیزیکی، شیمیایی و زیستی آن وابسته است؛ ویژگی‌هایی که در مقیاس‌های زمانی و مکانی تغییر می‌کنند. از این رو، تعیین دقیق ویژگی‌های خاک برای مدیریت پایدار اراضی، بهینه‌سازی مصرف نهاده‌ها و کاهش آثار نامطلوب زیست‌محیطی ضروری است.

فسفر یکی از عناصر غذایی پرمصرف و ضروری برای رشد گیاهان است و در فرایندهایی مانند انتقال انرژی، فتوسنتز، رشد ریشه، تشکیل بذر و بلوغ گیاه نقش دارد. کمبود فسفر می‌تواند رشد گیاه و عملکرد محصول را محدود کند؛ در مقابل، مصرف بیش از نیاز کودهای فسفری ممکن است موجب تجمع فسفر در خاک، افزایش هزینه تولید و آلودگی منابع آب شود. بنابراین، سنجش فسفر قابل‌جذب خاک از اجزای مهم برنامه‌های مدیریت حاصلخیزی و توصیه کودی است.

روش‌های مرجع اندازه‌گیری ویژگی‌های فیزیکی و شیمیایی خاک، با وجود دقت مناسب، معمولا زمان‌بر، پرهزینه و نیازمند آماده‌سازی نمونه و استفاده از مواد شیمیایی هستند. افزون بر این، تغییرپذیری مکانی خاک مستلزم برداشت و تجزیه تعداد زیادی نمونه است که اجرای روش‌های مرسوم را در مقیاس وسیع دشوار می‌سازد. از این رو، توسعه روش‌های سریع، غیرمخرب، کم‌هزینه و سازگار با محیط‌زیست برای ارزیابی ویژگی‌های خاک به یکی از محورهای مهم پژوهش در علوم خاک و کشاورزی دقیق تبدیل شده است (Walsh, & Shepherd 2002).

طیف‌سنجی مرئی - فروسرخ نزدیک یکی از روش‌های نوین برای ارزیابی ویژگی‌های خاک است. این فناوری ظرفیت مناسبی برای جایگزینی یا تکمیل روش‌های متداول آزمایشگاهی دارد. مطالعات مختلف نشان داده‌اند که داده‌های طیفی مرئی–فروسرخ نزدیک می‌توانند برای پیش‌بینی برخی ویژگی‌های فیزیکی، شیمیایی و زیستی خاک استفاده شوند (Shibusawa, 2013; Mirzaei *et al.*, & Kodaira Chodak, 2020; Ramirez-Lopez *et al.*, 2019; Rasooli *et* & Mouazen, 2018; Pudełko & 2022; Munnaf *et al.*, 2021; Nawar *al.*, 2023; Recena *et al.*, 2019).

با این حال، طیف‌های خاک علاوه بر اطلاعات مرتبط با ویژگی مورد مطالعه، تحت تأثیر عواملی مانند رطوبت، بافت، ماده آلی، رنگ، کانی‌ها، اکسیدهای آهن و آلومینیوم، پراکنش نور و نویز دستگاه قرار می‌گیرند. ازاین‌رو، استفاده از روش‌های مناسب پیش‌پردازش طیفی و مدل‌سازی چندمتغیره برای استخراج اطلاعات مفید و دستیابی به مدل‌های دقیق ضروری است.

بر این اساس، هدف پژوهش حاضر ارزیابی قابلیت طیف‌سنجی مرئی – فروسرخ نزدیک برای پیش‌بینی مقدار فسفر قابل‌جذب نمونه‌های خاک هواخشک و مقایسه تأثیر روش‌های مختلف پیش‌پردازش طیفی بر عملکرد مدل رگرسیون حداقل مربعات جزئی بود.

پیشینه پژوهش

طیف بازتابی خاک حاصل برهم‌کنش تابش الکترومغناطیسی با اجزای معدنی و آلی آن است. شدت و شکل منحنی بازتاب خاک می‌تواند تحت تأثیر بافت، رطوبت، ماده آلی، کربنات‌ها، اکسیدهای آهن، کانی‌های رسی و توزیع اندازه ذرات قرار گیرد. به‌طور معمول، افزایش کربنات کلسیم موجب افزایش بازتاب می‌شود؛ در حالی که ماده آلی و رطوبت خاک بازتاب طیفی را کاهش می‌دهند (Stenberg, 2010). بنابراین، حتی ویژگی‌هایی که جذب طیفی مستقیم و مشخصی ندارند، ممکن است از طریق همبستگی با سایر اجزای خاک به‌صورت غیرمستقیم پیش‌بینی شوند.

Nanni و Demattê (۲۰۰۶) نشان دادند که روش‌های بازتاب‌سنجی طیفی می‌توانند به‌عنوان مکمل روش‌های متداول تجزیه خاک استفاده شوند. هم‌چنین، Shepherd و Walsh (۲۰۰۲) با توسعه کتابخانه‌های طیفی خاک، ظرفیت طیف‌سنجی بازتابی را برای شناسایی و پیش‌بینی ویژگی‌های مختلف خاک نشان دادند. پیشرفت تجهیزات طیف‌سنجی و روش‌های کمومتریک نیز سبب شده است که استفاده از طیف‌سنجی مرئی - فروسرخ نزدیک در مقیاس آزمایشگاهی، مزرعه‌ای و برخط توسعه یابد.

در زمینه پیش‌بینی عناصر غذایی خاک، Malley *et al.* (۱۹۹۹) با بررسی ۸۷۷ نمونه خاک و استفاده از رگرسیون مؤلفه‌های اصلی و

رگرسیون حداقل مربعات جزئی، امکان پیش‌بینی برخی عناصر معدنی، از جمله فسفر، را ارزیابی کردند. Shao و He (۲۰۱۱) نیز از طیف‌سنجی فروسرخ برای پیش‌بینی نیتروژن، فسفر و پتاسیم خاک استفاده کردند و نشان دادند که مدل PLS می‌تواند برای تخمین این عناصر به کار رود، هرچند دقت مدل‌ها به ویژگی نمونه‌ها، محدوده طیفی و نحوه پردازش داده‌ها وابسته است. Sanogo *et al.* (۲۰۱۸) قابلیت طیف‌سنجی مرئی - فروسرخ نزدیک را برای برآورد کربن، پتاسیم و فسفر خاک بررسی کردند. Peng *et al.* (۲۰۱۹) نیز از طیف‌های بازتابی مرئی و فروسرخ نزدیک برای پیش‌بینی عناصر غذایی خاک استفاده کردند و بر اهمیت انتخاب طول موج‌ها و روش مدل‌سازی تأکید داشتند. Mohamed *et al.* (۲۰۲۰) با تلفیق داده‌های طیفی Vis–NIR و تصاویر ماهواره‌ای Landsat-8، نقشه‌برداری عناصر غذایی خاک را در شرایط خشک بررسی کردند.

پژوهش‌های جدیدتر نشان داده‌اند که دقت پیش‌بینی ویژگی‌های خاک علاوه بر مدل آماری، به کیفیت نمونه‌برداری، گستره تغییرات ویژگی هدف، رطوبت نمونه، روش انتخاب نمونه‌های واسنجی و نوع پیش‌پردازش وابسته است. برای نمونه، Mirzaei *et al* (۲۰۲۲) نشان دادند که کاهش اثر رطوبت می‌تواند دقت پیش‌بینی ویژگی‌های خاک را بهبود دهد. Munnaf *et al.* (۲۰۲۱) نیز با استفاده از رویکردهای ترکیبی داده‌کاوی، امکان پیش‌بینی برخط شاخص‌های اصلی کیفیت خاک را بررسی کردند.

پیش‌پردازش طیفی برای کاهش اثر پراکنش، تصحیح خط مبنا، هموارسازی طیف، حذف نویز و تفکیک بهتر باندهای جذبی استفاده می‌شود. با وجود این، یک روش پیش‌پردازش واحد را نمی‌توان برای همه نمونه‌ها و ویژگی‌ها به‌عنوان روش برتر معرفی کرد؛ زیرا عملکرد آن به نوع ماده، دستگاه، شرایط اندازه‌گیری، گستره طول موج و ویژگی هدف بستگی دارد (Heise & Winzen, 2001; Jamshidi *et al.*, 2014). مطالعات آقائی سعدی و همکاران (۲۰۱۸، ۲۰۲۱) نیز نشان داده‌اند که انتخاب روش پیش‌پردازش مناسب می‌تواند بر دقت مدل‌های مبتنی بر طیف‌سنجی Vis–NIR اثر قابل‌توجهی داشته باشد.

با توجه به نتایج متفاوت مطالعات پیشین و اثر ویژگی‌های منطقه‌ای خاک بر عملکرد مدل‌های طیفی، ارزیابی ترکیب‌های مختلف پیش‌پردازش برای مجموعه‌ای از نمونه‌های دارای تنوع فیزیکی و شیمیایی ضروری است. نوآوری پژوهش حاضر را می‌توان در استفاده هم‌زمان از نمونه‌های خاک مناطق مرطوب و خشک، بررسی ترکیب‌های متعدد پیش‌پردازش و مقایسه عملکرد آن‌ها برای پیش‌بینی فسفر قابل‌جذب خاک دانست.

مواد و روش‌ها

موقعیت جغرافیایی مناطق مورد مطالعه

منطقه مورد مطالعه شامل بخشی از مزارع چای شهرستان لاهیجان و هم‌چنین مزارع اطراف پالایشگاه تهران (اسماعیل‌آباد) و تصفیه‌خانه جنوب شهر تهران (شهر ری) است (شکل۱). مزارع چای لاهیجان در عرض جغرافیایی ′۳۷°۱۲,۷۸۵ تا ′۳۷°۱۳,۰۱۹ و طول جغرافیایی ′۴۹°۵۷,۰۴۵ تا ′۴۹°۵۷,۷۲۱ قرار دارند و ارتفاع متوسط آن‌ها ۴- متر از سطح دریا است. مزارع اطراف پالایشگاه تهران و تصفیه‌خانه جنوب شهر تهران در عرض جغرافیایی ′۳۵°۳۲,۷۵۸ تا ′۳۵°۳۲,۷۹۱ و طول جغرافیایی ′۵۱°۲۴,۵۲۴ تا ′۵۱°۲۴,۵۵۶ با ارتفاع متوسط ۱۰۳۸ متر از سطح دریا قرار گرفته‌اند. این مناطق به دلیل تنوع داده‌های آزمایشگاهی و پراکندگی مناسب انتخاب شده‌اند؛ به گونه‌ای که خاک مزارع چای لاهیجان دارای فسفر و رطوبت بالا است، در حالی که خاک مزارع جنوب تهران شرایط متفاوتی دارد. بر اساس اطلاعات ایستگاه‌های هواشناسی، متوسط بارندگی سالانه در محدوده لاهیجان ۹۲۴/۱ میلی‌متر و در محدوده اسماعیل‌آباد و شهر ری ۹۵/۶ میلی‌متر است. همچنین، متوسط دمای سالانه در لاهیجان ۲۲/۸۴ درجه سلسیوس و در مناطق جنوب تهران ۲۵/۴۹ درجه سلسیوس گزارش شده است.

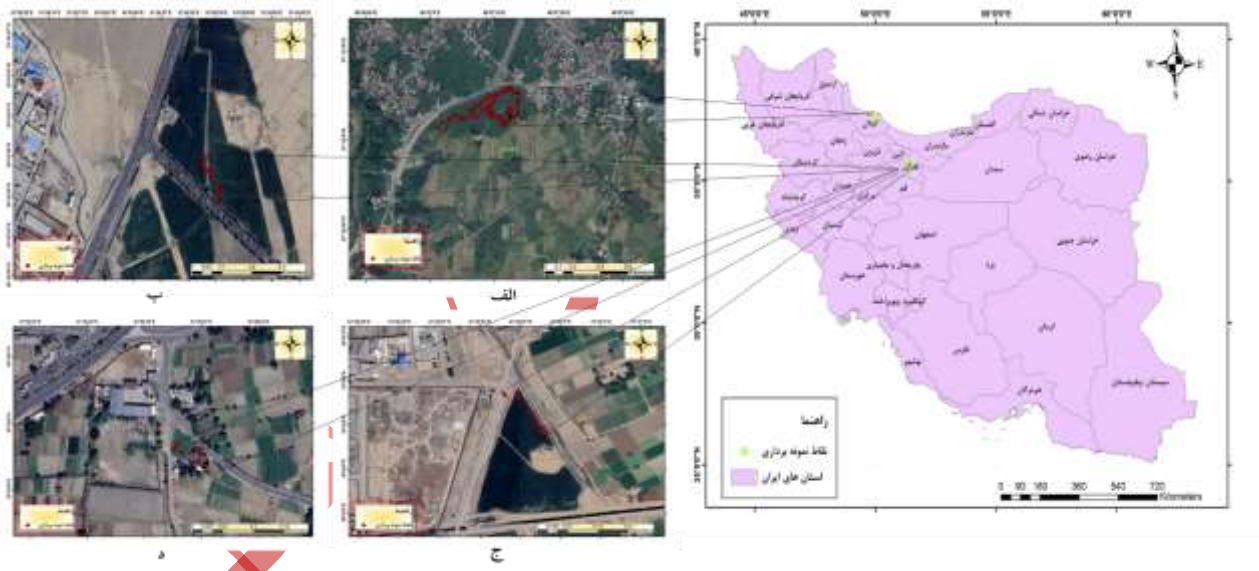
نمونه‌برداری خاک

در این پژوهش، در مجموع ۱۵۰ نمونه خاک جمع‌آوری شد که شامل ۱۲۰ نمونه از مزارع چای لاهیجان، ۱۵ نمونه از مزارع اسماعیل‌آباد و ۱۵ نمونه از مزارع شهر ری بود. در شهرستان لاهیجان نمونه‌ها از ۴ مزرعه مختلف جمع‌آوری شدند که فاصله نمونه‌ها ۵۰ متر در نظر گرفته شد. در مناطق جنوب تهران هم نمونه‌ها از سه مزرعه مختلف با فاصله ۵۰ متر جمع‌آوری شدند. نمونه‌های خاک از لایه ۲۰ تا ۳۰

سانتی‌متری، به‌عنوان لایه سطحی تحت تأثیر مدیریت زراعی و کاربرد کود فسفری، برداشت شدند. موقعیت جغرافیایی نمونه‌های مورد مطالعه به کمک سامانه مکان‌یاب جهانی (GPS)، شناسایی و ثبت گردید. نمونه‌های خاک پس از انتقال به آزمایشگاه، هوا خشک و به نرمی کوبیده و با عبور از الک ۲ میلی‌متری برای تعیین ویژگی‌های فیزیکی و شیمیایی آماده شد. علاوه بر فسفر قابل‌جذب، کربن آلی، pH و برخی دیگر از ویژگی‌های فیزیکی و شیمیایی خاک نیز اندازه‌گیری شدند. بااین‌حال، از آنجا که هدف پژوهش حاضر بررسی قابلیت طیف‌سنجی مرئی– فروسرخ نزدیک و اثر پیش‌پردازش‌های طیفی بر پیش‌بینی فسفر قابل‌جذب بود، نتایج تفصیلی سایر ویژگی‌ها در این پژوهش ارائه نشد. لازم به ذکر است که این ویژگی‌ها به‌عنوان متغیرهای مستقل وارد مدل PLS نشدند و مدل‌سازی تنها براساس طیف‌های Vis–NIR و مقادیر مرجع فسفر قابل‌جذب انجام شد. توزیع تجمعی اندازه ذرات خاک به روش هیدرومتر بایکاس، جرم ویژه حقیقی خاک بوسیله پیکنومتر و جرم ویژه ظاهری به روش کلوخه پارافین اندازه‌گیری شد. تخلخل خاک نیز از طریق محاسبه نسبت جرم ویژه ظاهری به جرم ویژه حقیقی و بهره‌گیری از رابطه ۱ محاسبه گردید.

$$f\ (\%) = (1 - \frac{\rho_b}{\rho_s}) \times 100$$ رابطه ۱)

که در آن، f تخلخل خاک (درصد)، ρ_b جرم ویژه ظاهری (g/cm^3) و ρ_s جرم ویژه حقیقی (g/cm^3) است. جدول ۱ توصیف آماری برخی ویژگی‌های فیزیکی و شیمیایی نمونه خاک‌های مورد مطالعه را نشان می‌دهد.



شکل ۱. موقعیت جغرافیایی مناطق مورد مطالعه و نقاط نمونه‌برداری، (الف) شهرستان لاهیجان (ب)، (ج) و (د) مناطق جنوب شهر تهران

جدول ۱. توصیف آماری برخی ویژگی‌های فیزیکی و شیمیایی نمونه‌های خاک

ویژگی‌ها	کمینه	بیشینه	میانگین	انحراف معیار
رس (درصد)	۱۶	۳۸	۲۱/۷۶	۶/۱۵
سیلت(درصد)	۱۸	۳۶	۲۸/۱۸	۳/۲۶
شن(درصد)	۳۰	۶۰	۵۰/۰۵	۸/۰۶
جرم ویژه ظاهری (g/cm^3)	۱/۰۷	۱/۹۳	۱/۴۳	۰/۱۳
جرم ویژه حقیقی (g/cm^3)	۲/۱۴	۲/۸۹	۲/۶۰۵	۰/۱۴۴
تخلخل (درصد)	۲۷	۵۶	۴۵	۳/۸
فسفر (ppm)	۱/۰۱۸	۲۳۴/۶۸	۸۸/۶۳	۵۲/۵۲

طیف‌سنجی نمونه‌های خاک

اندازه‌گیری‌های طیف‌سنجی Vis–NIR از نمونه‌های خاک در مد برهم‌کنش به صورت تماسی و با استفاده از دستگاه طیف‌سنج FieldSpec3 (ASD, Inc. USA) و کاوشگر تماسی[1] مدل A122300 در محدوده طیفی ۲۵۰۰–۳۵۰ نانومتر انجام شد. این طیف‌سنج دارای تکفام‌ساز نوع منشوری و شامل سه آشکارساز در دامنه‌های طول‌موج ۱۰۰۰–۳۵۰ دارای ۵۱۲ عنصر از نوع آرایه سیلیکونی، ۱۸۳۰–۱۰۰۰ از نوع InGaAs و ۲۵۰۰–۱۸۳۰ نانومتر از نوع InGaAs با درجه حساسیت متفاوتی نسبت به تغییرات دما و رطوبت است. از منبع نور هالوژن–تنگستن مدل High Intensity Contact Probe A122300 (ASD, Inc. USA) دارای توان خروجی ۶/۵ وات و ولتاژ DC ۱۸–۱۲ ولت و یک کاوشگر فیبر نوری تک‌شاخه مدل Ceram Optic 106622 استفاده گردید. حسگر فیبر نوری که به وسیله یک کابل با طول ۱/۲ متر بازتاب نوری را به طیف سنج منتقل می‌کند، قابلیت انتقال اطلاعات را به کامپیوتر برای نمایش و ذخیره سازی طیف‌های به دست‌آمده در نرم‌افزار View Spec Pro 6.0.15 ، از طریق پورت USB دارد. برای هر نمونه، پنج موقعیت متفاوت روی سطح نمونه خاک انتخاب شد و در هر موقعیت چهار اسکن متوالی ثبت گردید؛ بنابراین، برای هر نمونه ۲۰ اسکن خام به دست آمد. ابتدا چهار اسکن هر موقعیت میانگین‌گیری شد و سپس میانگین طیف‌های پنج موقعیت به‌عنوان یک طیف نماینده برای آن نمونه محاسبه شد. در نتیجه، هر نمونه خاک فقط یک مشاهده مستقل در مدل سازی داشت. پس از ایجاد چیدمان و پیش از طیف سنجی، ابتدا طیف تیره[2](D) با بستن درپوش فیبر نوری و خاموش‌کردن منبع نور و طیف مرجع[3](Ref) با استفاده از یک صفحه سفید مبنا از جنس تفلون، تعریف و ذخیره شدند.

اندازه‌گیری‌های مرجع

فسفر یکی از مهم‌ترین عناصر غذایی مورد نیاز گیاهان است. برای اندازه‌گیری این عنصر روش اولسن[4] به کار برده شد (Kuo, 1996). در این روش از محلول بی‌کربنات سدیم ۰/۵ مولار با pH=۸/۵ استفاده گردید. در این روش ۱۰ گرم نمونه خاک با ۲۰ میلی‌لیتر محلول بی‌کربنات سدیم برای ۳۰ دقیقه درون شیکر تکان داده شد. سپس نمونه به منظور جداسازی فاز محلول از ذرات معدنی با دستگاه گریز از مرکز چرخانده شد و برای به دست آوردن عصاره شفاف فیلتر گردید. ۵ میلی‌لیتر عصاره نمونه با ۵ میلی‌لیتر معرف رنگی رقیق شده و جذب محلول با بهره‌گیری از یک اسپکتروفتومتر در ۸۸۲ نانومتر قرائت گردید. تبدیل مقدار فسفر عصاره ($mg\ L^{-1}$) به مقدار فسفر قابل جذب خاک ($mg\ kg^{-1}$) به کمک رابطه ۲ انجام شد.

$$P\ (mg\ kg^{-1}) = [(A \times B \times R \times 1000)/E] \quad \text{رابطه ۲)}$$

که در آن، A قرائت عصاره نمونه ($mg\ L^{-1}$)، B حجم عصاره (L)، R نسبت آون–خشک به هوا–خشک یا نسبت آون–خشک به رطوبت مزرعه، ۱۰۰۰ ضریب تبدیل به kg و E وزن نمونه (g) است.

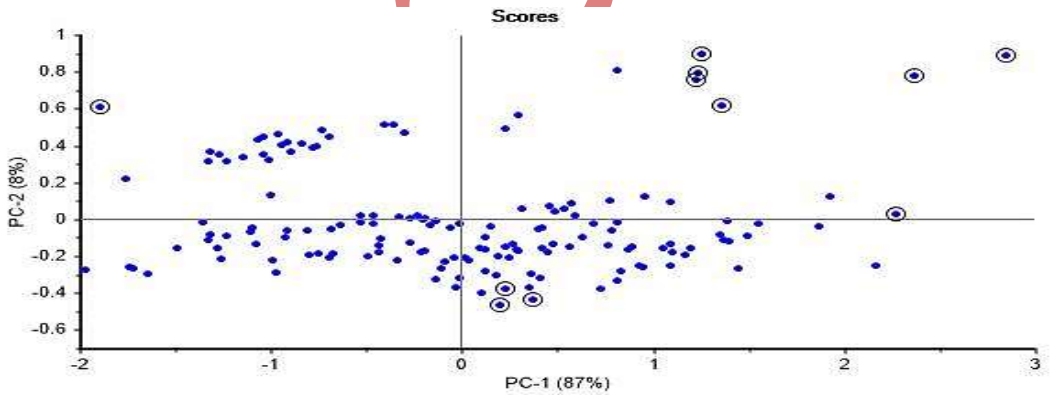
پیش‌پردازش داده‌ها و مدلسازی

برای حذف اطلاعات ناخواسته، پس‌زمینه و انواع نویز و دستیابی به مدل‌های واسنجی دقیق و قابل اعتماد، پیش‌پردازش داده‌های طیفی پیش از تدوین مدل‌های رگرسیون لازم است. برای این منظور، روش‌های پیش‌پردازش بسیاری بر پایه شیوه‌های ریاضی و آماری گسترش یافته و یا در حال تدوین هستند که هر یک کاربرد ویژه‌ای دارند و انتخاب یک روش پیش‌پردازش به‌عنوان بهترین روش که منجر به دقیق‌ترین نتایج مدل‌های پیش‌بینی پارامترهای مختلف در نمونه‌های مختلف شود، امکان‌پذیر نیست و به شرایط و امکانات آزمایش‌ها، نوع و ویژگی‌های نمونه و پارامتر مورد اندازه‌گیری بستگی دارد (Aghaei Sadi *et al.*, 2021; Heise & Winzen, 2001;

1. Contact Probe
2. Dark
3. Reference
4. Olsen

ویرایش اندازه‌گیری نشده

(Jamshidi *et al.*, 2014). بر این اساس، به منظور آماده‌سازی و پیش‌پردازش طیف‌ها، از برخی روش‌های هموارسازی[1] و کاهش نویز (میانگین‌گیری متحرک[2](MA) و ساویتزکی-گولای[3](SG))، نرمال‌سازی[4](تصحیح پراکنش افزاینده[5](MSC) و توزیع نرمال استاندارد[6] (SNV))، افزایش قدرت تفکیک طیف (مشتق‌های اول و دوم (D1 و D2)) استفاده شد. هم‌چنین اثر ترکیب روش‌های مختلف پیش‌پردازش بر دقت مدل‌های چندمتغیره تدوین‌شده برای سنجش مقدار فسفر نمونه‌های خاک مورد مطالعه قرار گرفت. هموارسازی ساویتزکی-گولای با پنجره ۳ نقطه‌ای و چندجمله‌ای مرتبه ۲، مشتق اول و دوم با استفاده از پنجره ۳ نقطه‌ای و چندجمله‌ای مرتبه ۲ و میانگین‌گیری متحرک با پنجره ۳ نقطه‌ای به کار رفت. MSC و SNV با تنظیمات پیش‌فرض نرم‌افزار Unscrambler X 10.4 اجرا شدند. برای اجرای این روش‌ها نرم‌افزارهای Unscrambler X 10.4 و View Spec Pro 6.0.15 استفاده شد. برای ایجاد ارتباط بین مقدار فسفر اندازه‌گیری شده و داده‌های طیفی از روش‌های مدل سازی رگرسیون حداقل مربعات جزئی (PLS) بهره‌گیری شد. در این راستا، نمونه‌های پرت[7] با بهره‌گیری از روش تجزیه مولفه‌های اصلی[8](PCA)، مشخص و حذف شدند (شکل ۲). تحلیل مؤلفه‌های اصلی روی طیف‌های خام در محدوده ۴۰۰۰ تا ۲۴۵۰ نانومتر انجام شد. نمونه‌های مشکوک براساس Hotelling's T^2 در سطح اطمینان ۹۵ درصد شناسایی شدند. از ۱۵۰ نمونه اولیه، ۱۱ نمونه، شامل ۷ نمونه از لاهیجان، ۳ نمونه از اسماعیل‌آباد و ۱ نمونه از شهر ری، مشکوک تشخیص داده شدند. طیف و نتیجه آزمایشگاهی این نمونه‌ها مجدد بررسی شد و به عنوان داده‌های پرت در نظر گرفته شدند. سپس، ۱۳۹ نمونه باقی‌مانده با روش Full Cross Validation مدل‌سازی و اعتبارسنجی شدند. در این روش، در هر تکرار یک نمونه از مجموعه داده کنار گذاشته شد، مدل با استفاده از $n-1$ نمونه باقی‌مانده تشکیل شد و مقدار فسفر نمونه کنارگذاشته شده پیش‌بینی گردید. این فرایند تا زمانی ادامه یافت که هر نمونه یک‌بار در مرحله اعتبارسنجی قرار گیرد.



شکل ۲. تشخیص داده‌های پرت از نظر طیفی با بهره‌گیری از روش تجزیه مولفه‌های اصلی (PCA)

در نهایت مدل‌های تدوین شده بر اساس خطای مدل‌ها به صورت ریشه میانگین مربعات خطای واسنجی[9](RMSEC)، ریشه میانگین مربعات خطای اعتبارسنجی متقابل[10](RMSECV)، ضریب تبیین واسنجی[11](R^2_c)، ضریب تبیین اعتبارسنجی متقابل[12](R^2_{cv})، MAE،

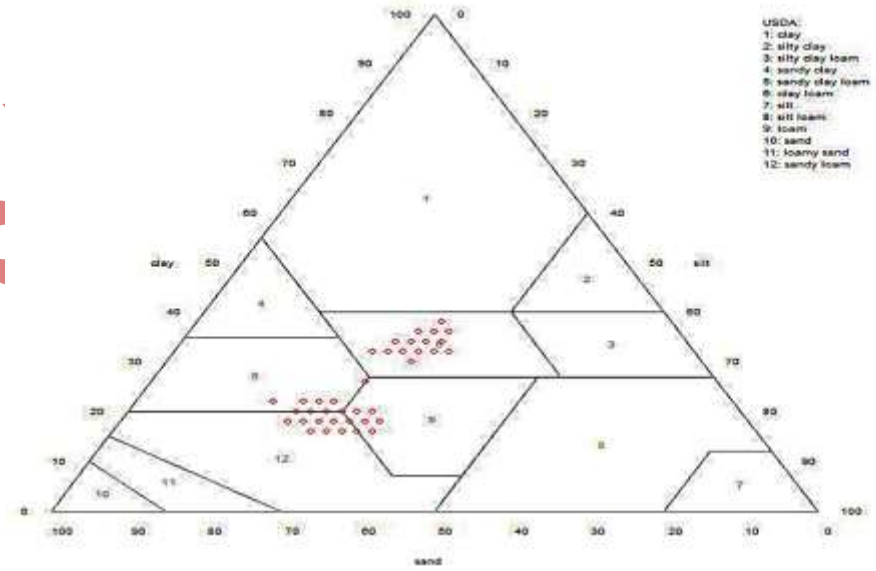
1. Smoothing
2. Moving Average
3. Savitzky-Golay
4. Normalizing
5. Multiplicative Scatter Correction
6. Standard Normal Variate
7. Outliers
8. Principal Component Analysis
9. Root Mean Square Error of Calibration
10. Root Mean Square Error of Cross-Validation
11. Correlation Coefficient of Calibration
12. Correlation Coefficient of Cross Validation

Bias و شاخص RPD اعتبارسنجی شدند. در این پژوهش، تعداد LVs بر پایه‌ی داشتن RMSECV و RMSEC کم‌تر، R_{cv}^2، R_c^2 و RPD بیش‌تر انجام شد. کم‌تر بودن اختلاف بین RMSEC و RMSECV نیز در مرحله‌ی بعد مد نظر قرار گرفت (Aghaei Sadi *et al.*, 2018). به دلیل وجود نویز در ابتدا و انتهای بازه طیف‌سنجی از اطلاعات محدوده ۴۰۰–۲۴۵۰ nm استفاده شد.

نتایج

ویژگی‌های فیزیکی و شیمیایی خاک

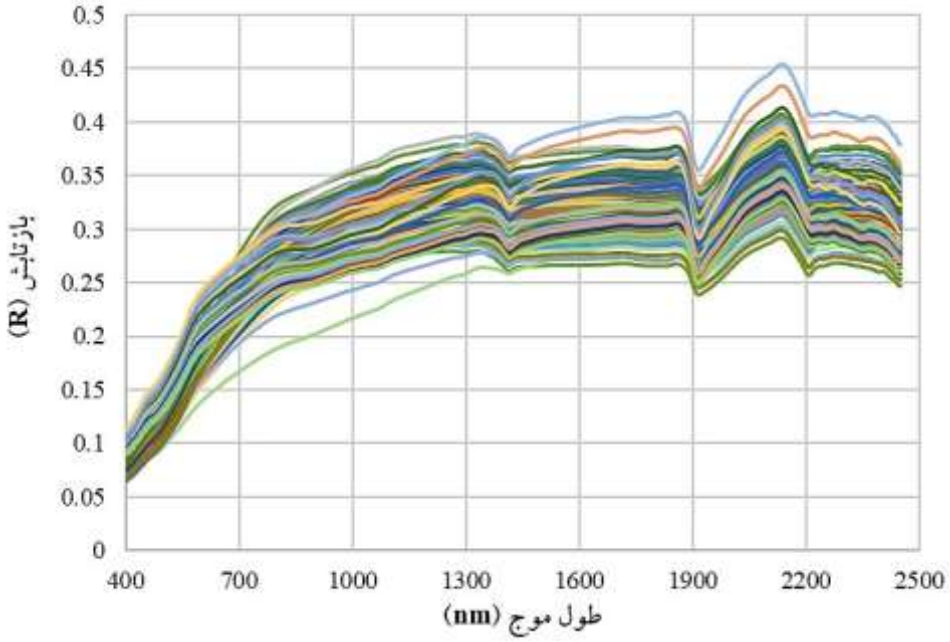
در شکل ۳ توزیع کلاس‌های بافتی نمونه خاک‌های مورد مطالعه روی مثلث بافت خاک ارائه شده است. با توجه به شکل ۳، نمونه خاک‌های مورد مطالعه از نظر کلاس بافتی چهار کلاس بافتی لوم، لوم رسی، لوم شنی و لوم رسی شنی را در بر می‌گیرد. با توجه به جدول ۱ مقادیر رس، سیلت و شن نمونه خاک‌های مورد مطالعه در محدوده ۳۸–۱۶، ۳۶–۱۸ و ۶۰–۳۰ در صد ا ست. همان‌گونه که در جدول ۱ مشاهده می‌شود جرم ویژه ظاهری و حقیقی نمونه‌های خاک به ترتیب در محدوده ۱/۹۳–۱/۰۷ و ۲/۸۹–۲/۱۴ گرم بر سانتی‌متر مکعب متغیر بوده و میانگین آن‌ها نیز به ترتیب ۱/۴۳ و ۲/۶۰۵ گرم بر سانتی‌متر مکعب ا ست. بررسی نتایج نشان می‌دهد که مقدار جرم ویژه ظاهری و حقیقی نمونه‌های خاک مزارع چای لاهیجان نسبت به نمونه‌های جنوب تهران کم‌تر ا ست که می‌تواند به دلیل مقدار زیاد فسفر و کربن آلی در این مزارع باشد. هم‌چنین تنوع در نوع اراضی و شیب منطقه و تغییر عمق می‌تواند مقدار جرم ویژه ظاهری و حقیقی خاک را تغییر دهد. تفاوت جرم ویژه ظاهری میان مناطق ممکن است با تغییرات بافت، ماده آلی، ساختمان خاک، کاربری زمین و مدیریت زراعی مرتبط باشد. با توجه به اینکه تمام عوامل مؤثر در این پژوهش اندازه‌گیری نشدند، نمی‌توان علت این تفاوت را به‌طور قطعی تعیین کرد. جرم ویژه حقیقی خاک معمولا در محدوده ۲/۷۵–۲/۶ گرم بر سانتی‌متر مکعب تغییر می‌کند و به طور میانگین ۲/۶۵ گرم بر سانتی‌متر مکعب در نظر گرفته می شود و تحت شرایط خاص تغییر می‌یابد. میانگین مقدار تخلخل ۴۵ در صد به د ست آمد. نتایج نشان می‌دهند که مقدار تخلخل در بین خاک‌ها کاملا متفاوت ا ست؛ در خاک‌های در شت بافت بین ۲۷ تا ۴۰ در صد و در خاک‌های ریز بافت بین ۴۰ تا ۵۶ در صد متغیر ا ست. مقدار متو سط فسفر خاک‌ها در مناطق مورد مطالعه برابر ۸۸/۶۳ ppm اندازه‌گیری شد که این مقدار با توجه به نقش آن در بهبود ویژگی‌های فیزیکی، شیمیایی و بیولوژیک خاک قابل قبول است. بااین‌حال، ارزیابی کفایت زراعی این مقدار به نوع محصول، ویژگی‌های خاک و حدود بحرانی توصیه شده برای روش اولسن وابسته است؛ ازاین‌رو، در پژوهش حاضر این مقدار صرفا به‌عنوان مشخصه آماری جمعیت نمونه گزارش شده است.



شکل ۳. توزیع کلاس‌های بافتی نمونه خاک‌های مورد مطالعه

رفتار طیفی نمونه‌های خاک

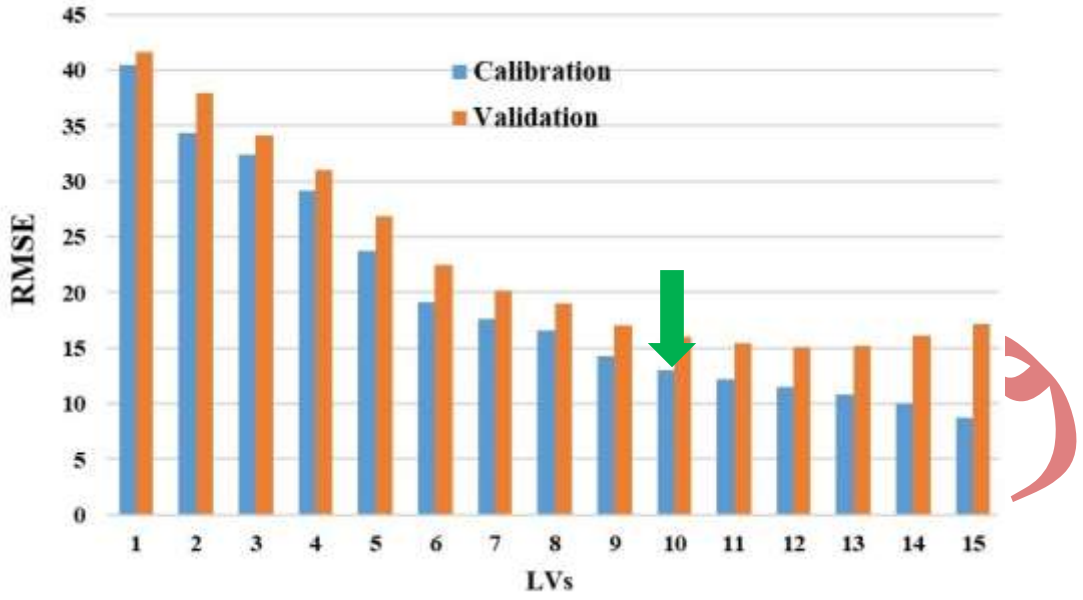
منحنی‌های خام بازتاب نمونه‌های خاک در شکل ۴ ارائه شده‌اند. تفاوت میان منحنی‌های بازتاب می‌تواند تحت تأثیر تغییرات بافت، ماده آلی، رطوبت باقی‌مانده، کربنات‌ها، کانی‌های رسی، اکسیدهای آهن و سایر اجزای خاک قرار گرفته باشد. رنگ، رطوبت، کربن آلی، توزیع اندازه ذرات و اکسیدهای آهن و آلومینیوم از مهم‌ترین عوامل مؤثر بر بازتاب طیفی خاک هستند (Stenberg, 2010). اثر این عوامل بر بازتاب خاک در مطالعات طیف سنجی پیشین گزارش شده است. بااین‌حال، در پژوهش حاضر سهم مستقل هر یک از این ویژگی‌ها در تغییرات طیفی و پیش‌بینی فسفر تعیین نشد؛ بنابراین، این عوامل صرفا به‌عنوان سازوکارهای احتمالی مؤثر بر پاسخ طیفی مطرح می‌شوند.



شکل ۴. میانگین منحنی‌های خام بازتاب طیفی نمونه‌های خاک

عملکرد مدل‌های پیش‌بینی فسفر

نتایج مدل PLS بر پایه طیف‌های Vis–NIR و ترکیب‌های مختلف پیش‌پردازش در جدول ۲ ارائه شده است. نمودارهای RMSECV و RMSEC در هر LVs در شکل ۵ نشان داده شده است. LVs بهینه یا مطلوب در روی شکل مشخص شده است که بیانگر این است که LVs موردنظر کم‌ترین مقدار RMSECV و RMSEC و کم‌ترین اختلاف بین RMSECV و RMSEC را دارد.



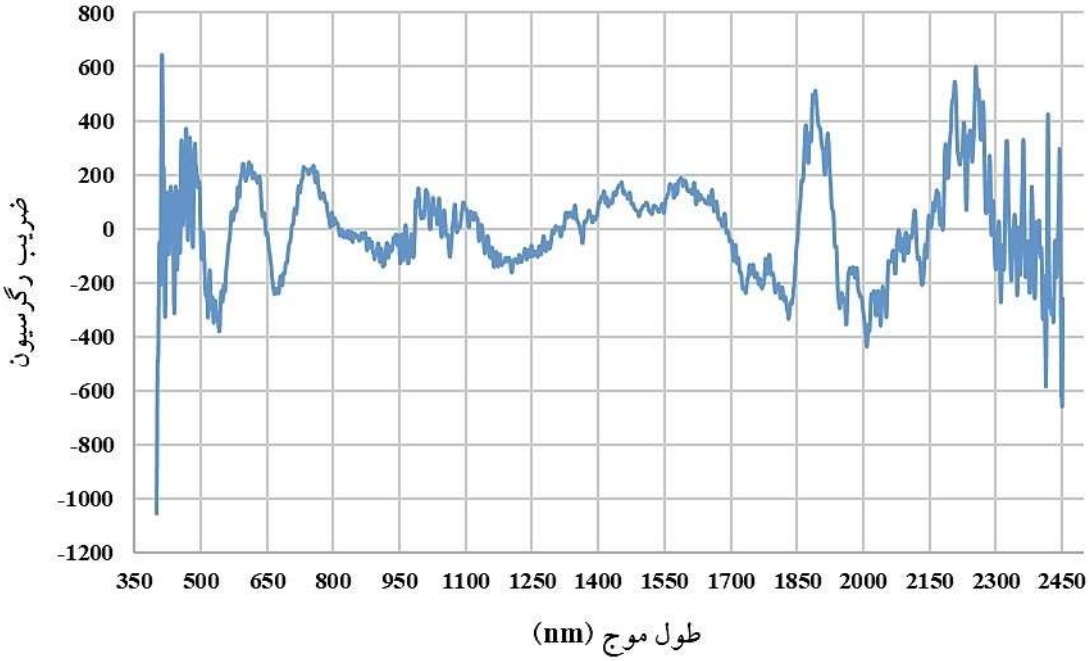
شکل ۵. نمودارهای RMSECV و RMSEC در هر LVs برای مدل PLS تدوین شده بر پایه طیف‌های NIR و پیش‌پردازش SG+MSC برای سنجش میزان فسفر نمونه‌های خاک هواخشک

جدول ۲. نتایج مدل‌های پیش‌بینی فسفر نمونه‌های خاک

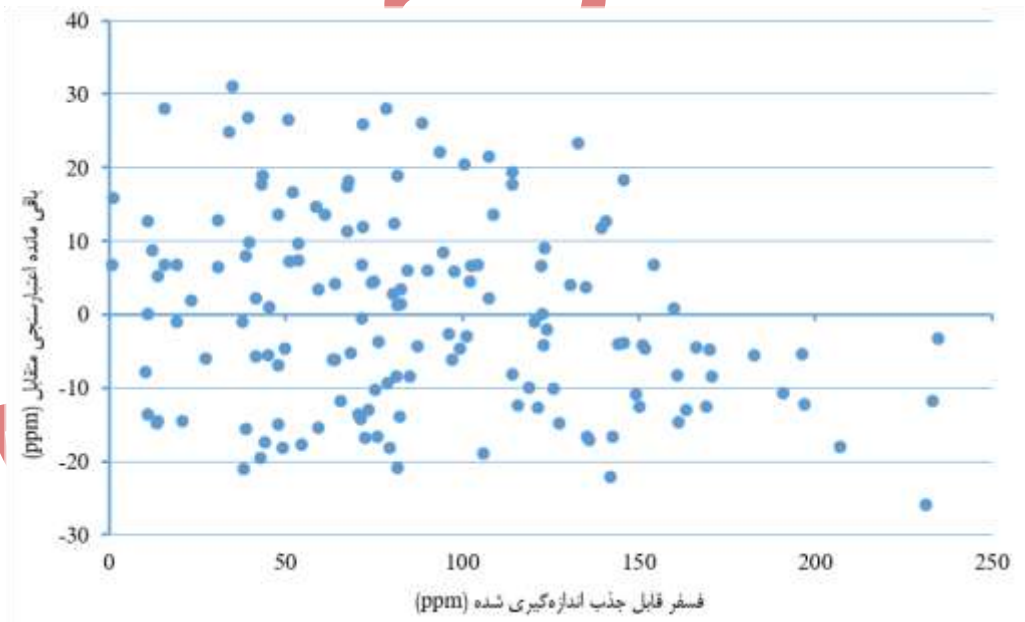
پیش‌پردازش	مدل	LVs	RMSEC	R_c^2	RMSECV	R_{cv}^2	MAE	Bias	RPD
بدون پیش‌پردازش	PLS	۱۱	۱۲/۶۰	۰/۹۴	۱۵/۸۵	۰/۹۱	۹/۵	$-۵/۳\times۱۰^{-۵}$	۳/۷۳
MA+MSC	PLS	۱۰	۱۳/۹۲	۰/۹۳	۱۷/۰۷	۰/۹۰	۱۱/۰۱	$۱/۳۱\times۱۰^{-۶}$	۳/۰۸
MA+SNV	PLS	۱۱	۱۴/۴۹	۰/۹۲	۱۷/۶۶	۰/۸۹	۱۰/۷۵	$۱/۴۸\times۱۰^{-۵}$	۲/۹۷
SG+MSC	**PLS**	**۱۰**	**۱۳/۰۴**	**۰/۹۴**	**۱۶/۰۰**	**۰/۹۱**	**۱۰/۹۴**	$\mathbf{-۲/۹۶\times۱۰^{-۵}}$	**۳/۲۸**
SG+SNV	PLS	۱۱	۱۴/۴۸	۰/۹۲	۱۷/۶۹	۰/۸۹	۱۰/۷	$۸/۴۸\times۱۰^{-۷}$	۲/۹۷
SG+D1	PLS	۶	۱۲/۵۴	۰/۹۴	۲۱/۸۹	۰/۸۳	۱۰/۰۹	$۸/۱\times۱۰^{-۷}$	۲/۴۰
SG+D2	PLS	۵	۱۹/۴۸	۰/۸۶	۴۰/۵۶	۰/۴۱	۱۴/۹۶	$-۲/۳۸\times۱۰^{-۶}$	۱/۲۹
SG+D1+MSC	PLS	۶	۱۳/۶۴	۰/۹۳	۲۳/۱۳	۰/۸۱	۱۱/۶۶	$۲/۱۸\times۱۰^{-۶}$	۲/۲۷
SG+D1+SNV	PLS	۸	۱۰/۲۵	۰/۹۶	۲۲/۴۵	۰/۸۲	۱۱/۶۹	$-۱/۸\times۱۰^{-۶}$	۲/۳۴
SG+D2+MSC	PLS	۵	۲۰/۱۴	۰/۸۵	۴۱/۲۰	۰/۳۹	۱۷/۲۶	$۶/۶۳\times۱۰^{-۶}$	۱/۲۷
SG+D2+SNV	PLS	۵	۲۰/۱۸	۰/۸۵	۴۱/۲۵	۰/۳۹	۱۵/۳۵	$۲/۴۵\times۱۰^{-۶}$	۱/۲۷

مدل بدون پیش‌پردازش و مدل SG+MSC عملکرد بسیار مشابهی داشتند. انتخاب مدل SG+MSC به این دلیل است که در LVs کم‌تر نسبت به مدل بدون پیش‌پردازش، نتایج نزدیک به مدل بدون پیش پردازش حاصل شده است. این مدل با ۱۰ متغیر نهفته، مقادیر R^2_c =۰/۹۴، RMSEC=۱۳/۰۴، R^2_{cv} =۰/۹۱ ، RMSECV =۱۶/۰۰، MAE= ۱۰/۹۴، Bias = $-۲/۹۶\times۱۰^{-۵}$ و RPD= ۳/۲۸ را به دست آورد. علامت منفی bias نشان داد که مدل به‌طور متوسط دارای گرایش به کم‌برآورد است ولی چون مقدار Bias نزدیک به صفر است یعنی مدل در سطح کلی فاقد گرایش نظام‌مند به بیش‌برآورد یا کم‌برآورد فسفر بود. استفاده از مشتق دوم موجب کاهش محسوس

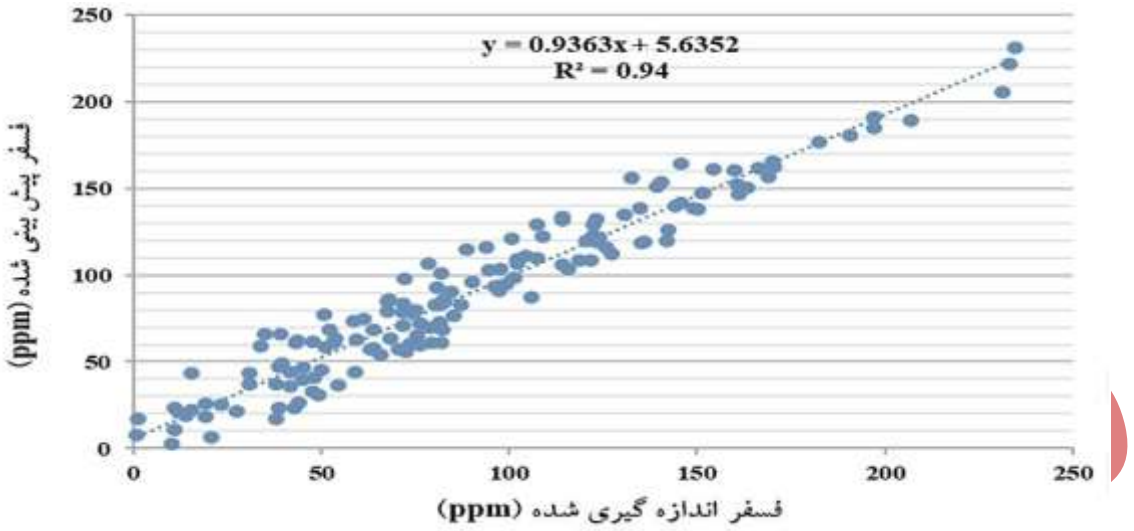
عملکرد مدل شـــد؛ به‌طوری‌که مقدار R^2_{cv} در مدل‌های SG+D2، SG+D2+MSC و SG+D2+SNV به‌ترتیب ۰/۴۱، ۰/۳۹ و ۰/۳۹ بود. بیش‌ترین ضریب تبیین وا سنجی متعلق به ترکیب SG+D1+SNV بود، اما افزایش RMSECV و کاهش R^2_{cv} ن شان داد که برازش بهتر در مرحله واسنجی الزاما به پیش‌بینی بهتر منجر نشده است.



شکل ۶. ضرایب رگرسیونی مدل منتخب برای پیش‌بینی فسفر خاک



شکل ۷. باقی‌مانده‌های حاصل از اعتبارسنجی متقابل کامل مدل PLS-SG+MSC در برابر مقادیر اندازه‌گیری‌شده فسفر قابل‌جذب خاک



شکل ۸. مقادیر پیش‌بینی‌شده فسفر در مقابل مقادیر اندازه‌گیری‌شده

بحث

نتایج این پژوهش نشان داد که طیف‌سنجی مرئی–فروسرخ نزدیک در محدوده ۴۰۰ تا ۲۴۵۰ نانومتر، همراه با رگرسیون حداقل مربعات جزئی، ظرفیت مناسبی برای برآورد فسفر قابل‌جذب خاک‌های مورد مطالعه دارد. بااین‌حال، تفسیر این قابلیت باید با توجه به ماهیت پاسخ طیفی فسفر انجام شود. برخلاف برخی اجزای خاک مانند مادهٔ آلی، آب، کانی‌های رسی و کربنات‌ها که دارای باندهای جذبی نسبتاً مشخص در ناحیهٔ مرئی و فروسرخ نزدیک هستند، فسفر قابل‌جذب معمولاً فاقد باند جذبی مستقیم و قوی در این محدوده طیفی است. بررسی ضرایب رگرسیونی مدل منتخب نشان داد که نواحی با ضرایب بالاتر بیش‌ترین سهم را در پیش‌بینی داشتند. نواحی مرئی ممکن است با رنگ خاک، مادهٔ آلی و اکسیدهای آهن مرتبط باشند، درحالی‌که تغییرات محدوده فروسرخ نزدیک می‌تواند تحت تأثیر گروه‌های O–H، کانی‌های رسی، آب پیوندی و ترکیبات آلی قرار گیرد. بااین‌حال، این انتساب‌ها باید با احتیاط تفسیر شوند؛ زیرا ضرایب PLS در حضور طول موج‌های شدیداً همبسته لزوماً نشان‌دهندهٔ رابطهٔ علّی نیستند و فسفر قابل‌جذب نیز باند جذب مستقیم و اختصاصی شناخته‌شده‌ای در این محدوده ندارد. بنابراین، طول موج‌های مهم احتمالاً بازتاب هم‌تغییری فسفر با اجزای فعال طیفی خاک هستند (Stenberg *et al.*, 2010; Recena *et al.*, 2019; Peng *et al.*, 2019). ازاین‌رو، دقت بالای مدل به‌دست‌آمده در پژوهش حاضر احتمالاً نه تنها ناشی از پاسخ مستقیم فسفر، بلکه حاصل ثبت هم‌زمان تغییرات مجموعه‌ای از خصوصیات خاک مرتبط با نگهداشت، تثبیت و فراهمی فسفر بوده است.

دامنهٔ وسیع فسفر قابل‌جذب نمونه‌ها، از ۱/۰۱۸ تا ۲۳۴/۶۸ میلی‌گرم بر کیلوگرم، یکی از عوامل مؤثر در عملکرد مطلوب مدل محسوب می‌شود. وجود تغییرپذیری کافی در متغیر هدف، امکان استخراج مؤلفه‌های پنهان مرتبط با فسفر را برای مدل PLS فراهم می‌کند. نمونه‌ها از دو محیط پدولوژیکی و اقلیمی متفاوت، یعنی مزارع مرطوب و اسیدی چای در لاهیجان و اراضی خشک‌تر جنوب تهران، جمع‌آوری شدند. این تفاوت احتمالاً باعث ایجاد تنوع در بافت، مادهٔ آلی، کانی‌شناسی، اکسیدهای فلزی، کربنات‌ها و سابقهٔ مدیریت کود شده و در نتیجه، فضای طیفی و شیمیایی نسبتاً گسترده‌ای برای واسنجی ایجاد کرده است. مطالعات پیشین نیز نشان داده‌اند که گسترش دامنهٔ تغییرات ویژگی هدف و افزایش تنوع نمونه‌های واسنجی می‌تواند عملکرد مدل‌های Vis–NIR را بهبود دهد، مشروط بر آنکه ساختار داده‌ها به‌درستی اعتبارسنجی شود (Shepherd & Walsh, 2002; Nawar & Mouazen, 2018; Ramirez-Lopez *et al.*, 2019).

باید توجه داشت که ترکیب نمونه‌های چند منطقه‌ای می‌تواند علاوه بر افزایش استحکام مدل، خطر ایجاد همبستگی‌های مکانی یا منطقه‌ای کاذب را نیز افزایش دهد. برای مثال، اگر نمونه‌های لاهیجان به‌طور هم‌زمان دارای فسفر، رطوبت باقیمانده، مادهٔ آلی و تیرگی

بیش‌تر باشند، مدل ممکن است بخشی از تغییرات فسفر را از طریق تفاوت کلی میان دو منطقه تشخیص دهد، نه از طریق روابطی که در هر منطقه نیز به‌طور مستقل پایدار باشند. بنابراین، اعتبارسنجی تصادفی یا اعتبارسنجی متقابل کامل ممکن است عملکرد مدل را خوش‌بینانه‌تر از کاربرد آن در مناطق جدید نشان دهد (Ramirez-Lopez *et al.*, 2019; Munnaf *et al.*, 2021).

از میان روش‌های بررسی‌شده، ترکیب هموارسازی ساویتزکی–گولای و تصحیح پراکنش چندگانه، یعنی SG+MSC، عملکرد مناسبی نشان داد و مقادیر R_c^2 = ۰/۹۴، R_{cv}^2 = ۰/۹۱، RMSEC برابر با ۱۳/۰۴ و RMSECV برابر با ۱۶/۰۰ میلی‌گرم بر کیلوگرم را ایجاد کرد. مقدار MAE مدل SG+MSC برابر ۱۰/۹۴ میلی‌گرم بر کیلوگرم و bias برابر $-۲/۹۶\times۱۰^{-۵}$ بود. علامت منفی bias نشان داد که مدل به‌طور متوسط دارای گرایش به کم‌برآورداست ولی چون مقدار Bias نزدیک به صفر است یعنی مدل در سطح کلی فاقد گرایش نظام‌مند به بیش‌برآورد یا کم‌برآورد فسفر بود. مقدار RPD برابر با ۳/۲۸ نیز نشان می‌دهد که مدل از نظر توان تفکیک و برآورد کمّی در محدوده مناسبی قرار دارد. عملکرد مطلوب SG را می‌توان به کاهش نوسان‌های فرکانس بالا و حفظ شکل عمومی باندهای طیفی نسبت داد. در مقابل، MSC تغییرات چندگانه و ضربی ناشی از پراکنش نور را اصلاح می‌کند؛ تغییراتی که ممکن است بر اثر تفاوت در اندازه ذرات، فشردگی نمونه، ناهمواری سطح و نحوه تماس کاوشگر با خاک ایجاد شوند. ترکیب این دو روش احتمالا باعث شده است بخش بیش‌تری از تغییرات طیفی مرتبط با ترکیب شیمیایی خاک حفظ و بخشی از تغییرات فیزیکی نامرتبط حذف شود. این نتیجه با مبانی کمومتریک طیف سنجی سازگار است؛ زیرا پیش‌پردازش مناسب باید نسبت سیگنال به نویز را افزایش دهد، بدون آنکه اطلاعات شیمیایی مفید را حذف یا ساختار نویز را تقویت کند (Heise & Winzen, 2001). بااین‌حال، هیچ روش پیش‌پردازشی را نمی‌توان به‌طور عمومی برای تمام خاک‌ها و همه ویژگی‌ها بهترین روش دانست. انتخاب روش مناسب به نوع خاک، دامنه طیفی، وضوح دستگاه، شرایط آماده‌سازی نمونه، ویژگی هدف و الگوریتم مدل‌سازی بستگی دارد. نتایج پژوهش حاضر نیز این موضوع را تأیید می‌کند؛ زیرا پیش‌پردازش‌های مختلف عملکردهای بسیار متفاوتی ایجاد کردند.

مدل‌های مبتنی بر مشتق دوم، شامل SG+D2، SG+D2+MSC و SG+D2+SNV، ضعیف‌ترین نتایج را نشان دادند؛ به‌طوری‌که R_{cv}^2 به حدود ۰/۳۹ تا ۰/۴۱ کاهش یافت و RMSECV به بیش از ۴۰ میلی‌گرم بر کیلوگرم رسید. مشتق دوم می‌تواند باندهای هم‌پوشان را تفکیک و خط پایه را حذف کند، اما هم‌زمان نسبت به نویز بسیار حساس است. ازآنجاکه فسفر در محدوده Vis–NIR پاسخ جذبی مستقیم و شدیدی ندارد، استفاده از مشتق دوم احتمالا سیگنال‌های ضعیف و غیرمستقیم مرتبط با فسفر را کاهش داده و نوسان‌های دستگاه یا تغییرات جزئی سطح نمونه را تقویت کرده است. تعداد نسبتا محدود نمونه‌ها نیز ممکن است حساسیت مدل‌های مشتقی را نسبت به بیش‌برازش افزایش داده باشد. کاهش تعداد متغیرهای پنهان در مدل‌های مشتق دوم، همراه با افت شدید دقت پیش‌بینی، نشان می‌دهد که کاهش پیچیدگی مدل لزوما به معنی افزایش قابلیت تعمیم نیست و ممکن است بخشی از اطلاعات مفید طیفی حذف شده باشد.

نتایج مربوط به مشتق اول نیز نشان داد که افزایش تفکیک طیفی همیشه به بهبود پیش‌بینی منجر نمی شود. مدل SG+D1 دارای مقدار R_{cv}^2 = ۰/۸۳ و RMSECV برابر با ۲۱/۸۹ میلی‌گرم بر کیلوگرم بود که در مقایسه با SG+MSC ضعیف‌تر است. افزودن MSC یا SNV به مشتق اول نیز بهبود قابل‌توجهی ایجاد نکرد. این یافته نشان می‌دهد که برای داده‌های این پژوهش، اصلاح اثر پراکنش و هموارسازی طیف‌ها از برجسته سازی تغییرات شیب طیفی مؤثرتر بوده است. مشابه بودن عملکرد MSC و SNV در برخی ترکیب‌ها نیز قابل انتظار است؛ زیرا هر دو روش برای کاهش اثر پراکنش و اختلاف خط پایه استفاده می شوند، اگرچه سازوکار ریاضی آن‌ها یکسان نیست.

عملکرد مناسب PLS را می‌توان به توانایی این روش در مدیریت تعداد زیاد طول موج‌های هم‌بسته و استخراج تعداد محدودی متغیر پنهان مرتبط با پاسخ فسفر نسبت داد. طیف‌های Vis–NIR دارای هم‌خطی شدید هستند و استفاده مستقیم از روش‌های رگرسیون معمولی می‌تواند به ضرایب ناپایدار منجر شود. PLS با استفاده هم‌زمان از ساختار ماتریس طیفی و متغیر پاسخ، مؤلفه‌هایی را استخراج می‌کند که بیش‌ترین ارتباط را با ویژگی هدف دارند. در مدل منتخب، استفاده از ۱۰ متغیر پنهان نشان می‌دهد که رابطه فسفر با طیف نسبتا پیچیده و چندعاملی بوده است. بااین‌حال، تعداد متغیرهای پنهان باید با اعتبارسنجی مناسب تعیین شود؛ زیرا افزایش بیش‌ازحد آن‌ها

می‌تواند نویز و تغییرات اختصاصی مجموعه واسنجی را وارد مدل کند.

مقدار $R^2_{cv} = ۰/۹۱$ نشان می‌دهد که مدل منتخب توانسته است بخش زیادی از تغییرات فسفر را توضیح دهد. هم‌چنین نزدیک بودن R^2_c و R^2_p، همراه با اختلاف نسبتا محدود RMSEC و RMSECV، نشانه‌ای از نبود بیش‌برازش شدید در چارچوب اعتبارسنجی انجام‌شده است. مقدار RPD بیش از ۳ نیز معمولا بیانگر قابلیت مناسب مدل برای برآورد کمّی و تفکیک نمونه‌های دارای سطوح مختلف ویژگی هدف است. بااین‌حال، حدود تفسیر شاخص‌هایی مانند RPD نباید به صورت مطلق در نظر گرفته شود؛ زیرا مقدار آن‌ها به توزیع داده‌ها، انحراف معیار متغیر مرجع و طراحی مجموعه اعتبارسنجی وابسته است. بنابراین، ارزیابی مدل باید بر اساس مجموعه‌ای از شاخص‌ها و نه تنها یک معیار انجام شود.

افزون بر شاخص‌های ضریب تبیین و RMSE، برای مدل SG+MSC مقدار MAE برابر ۱۰/۹۴ میلی‌گرم بر کیلوگرم و bias برابر $-۲/۹۶\times۱۰^{-۵}$ میلی‌گرم بر کیلوگرم به دست آمد. مقدار بسیار نزدیک به صفر bias نشان داد که میانگین کلی خطاهای جهت‌دار مدل تقریباً صفر بود. بااین‌حال، از آنجا که خطاهای مثبت و منفی می‌توانند یکدیگر را خنثی کنند، الگوی باقیمانده‌ها نیز در برابر مقادیر مرجع فسفر بررسی شد. همان‌گونه که در شکل ۷ مشاهده می شود، باقیمانده‌ها در بخش پایین و میانی دامنه فسفر در دو طرف خط صفر پراکنده بودند. با افزایش غلظت فسفر، جهت باقیمانده‌ها عمدتا منفی شد؛ به‌طوری‌که در نمونه‌های دارای فسفر بیش از حدود ۱۵۰ میلی‌گرم بر کیلوگرم، بیش‌تر مقادیر پیش‌بینی شده کم‌تر از مقادیر مرجع بودند. بنابراین، مدل در انتهای بالای دامنه فسفر گرایش نسبی به کم‌برآورد نشان داد. در مقابل، در غلظت‌های پایین‌تر، علاوه بر کم‌برآورد، تعدادی بیش‌برآورد قابل‌توجه نیز مشاهده شد. گرایش به کم‌برآورد مقادیر بالا ممکن است به تعداد محدودتر نمونه‌ها در انتهای بالای دامنه، توزیع نامتوازن مقادیر مرجع و تمایل مدل خطی PLS به نزدیک‌کردن پیش‌بینی مقادیر حدی به میانگین مجموعه واسنجی مربوط باشد. بنابراین، با وجود مناسب بودن شاخص‌های کلی مدل، پیش‌بینی مقادیر بسیار بالای فسفر باید با احتیاط تفسیر شود. افزایش تعداد نمونه‌ها در انتهای بالای دامنه و استفاده از اعتبارسنجی مستقل یا مدل‌های غیرخطی می‌تواند در پژوهش‌های آینده به ارزیابی و کاهش این محدودیت کمک کند.

یافته‌های پژوهش حاضر با نتایج Malley *et al.* (۱۹۹۹) همسو است؛ آن‌ها نیز نشان دادند که طیف‌سنجی فروسرخ نزدیک می‌تواند برای برآورد برخی عناصر معدنی خاک، از جمله فسفر، استفاده شود. نتایج حاضر از نظر R^2_p به نتایج گزارش شده توسط Malley *et al.* نزدیک است، اما باید توجه داشت که تفاوت در روش استخراج فسفر، دامنه غلظت، نوع خاک و روش اعتبارسنجی، مقایسه مستقیم مطالعات را محدود می‌کند. عملکرد مدل حاضر از نتایج Shao & He (۲۰۱۱) بهتر بود؛ در مطالعه آن‌ها پیش‌بینی فسفر با دقت پایین‌تری انجام شد. این اختلاف می‌تواند ناشی از استفاده از محدوده طیفی گسترده‌تر، تنوع بیشتر نمونه‌ها، آماده‌سازی یکنواخت خاک‌های هواخشک و انتخاب روش مناسب پیش‌پردازش در پژوهش حاضر باشد. اگرچه ضریب تعیین مدل حاضر از مقدار گزارش‌شده در برخی پژوهش‌ها بیش‌تر بود، مقایسه مستقیم این اعداد به دلیل تفاوت در دامنه فسفر، روش استخراج مرجع، نوع و تنوع خاک، آماده‌سازی نمونه، محدوده طیفی و طراحی اعتبارسنجی مناسب نیست. ازاین‌رو، این مقایسه صرفا برای ارائه زمینه پژوهشی انجام شد و به‌معنای برتری قطعی مدل حاضر نیست.

Peng *et al.* (۲۰۱۹) نیز قابلیت Vis–NIR را برای پیش‌بینی عناصر غذایی خاک گزارش کردند، اما تأکید داشتند که دقت پیش‌بینی ویژگی‌های فاقد باند جذبی مستقیم، به روابط آن‌ها با سایر اجزای طیف‌فعال خاک وابسته است. Recena *et al.* (۲۰۱۹) نشان دادند که طیف‌سنجی Vis–NIR ممکن است عملکرد یا وضعیت کلی حاصلخیزی خاک را بهتر از برخی شاخص‌های عملیاتی فراهمی عناصر پیش‌بینی کند. این موضوع برای تفسیر نتایج فسفر اهمیت دارد؛ زیرا فسفر اولسن یک شاخص استخراج‌شونده و وابسته به روش آزمایشگاهی است و لزوما یک جزء طیف‌فعال مستقل محسوب نمی شود. بنابراین، مدل طیفی ممکن است بیش از آنکه صرفا مقدار شیمیایی فسفر استخراج‌شده را شناسایی کند، مجموعه شرایط خاک مؤثر بر فراهمی فسفر را بازنمایی کرده باشد.

مطالعه Mohamed *et al.* (۲۰۲۰) نشان داد که تلفیق طیف‌سنجی Vis–NIR با داده‌های سنجش‌ازدور می‌تواند برای نقشه‌برداری عناصر غذایی خاک در شرایط خشک مفید باشد. هم‌چنین Munnaf *et al.* (۲۰۲۱) گزارش کردند که ترکیب روش‌های داده‌کاوی با طیف‌سنجی Vis–NIR می‌تواند پیش‌بینی برخط شاخص‌های کیفیت خاک را بهبود دهد. این نتایج، امکان توسعه مدل حاضر از شرایط

آزمایشگاهی به سامانه‌های مزرعه‌ای و نقشه‌برداری مکانی فسفر را مطرح می‌کنند؛ هرچند انتقال مدل‌های آزمایشگاهی به شرایط صحرایی با چالش‌هایی مانند تغییر رطوبت، زبری سطح، بقایای گیاهی، نور محیط و تغییر فاصله حسگر تا خاک همراه است.

رطوبت یکی از مهم‌ترین عوامل محدودکننده انتقال مدل‌های طیفی خاک از آزمایشگاه به مزرعه است. آب دارای باندهای جذبی قوی در نزدیکی ۱۴۰۰ و ۱۹۰۰ نانومتر است و می‌تواند باندهای ضعیف مربوط به سایر اجزای خاک را بپوشاند. در این پژوهش، استفاده از نمونه‌های هواخشک و الک شده بخشی از تغییرات ناشی از رطوبت و اندازه ذرات را کنترل کرده و احتمالا به افزایش دقت مدل کمک کرده است. بااین‌حال، این شرایط استاندارد ممکن است عملکرد مدل در خاک مرطوب مزرعه را بازنمایی نکند. Mirzaei *et al.* (۲۰۲۲) نشان دادند که روش حذف مؤلفه‌های متعامد پارامتر خارجی می‌تواند اثر رطوبت را کاهش داده و پایداری پیش‌بینی ویژگی‌های خاک را بهبود دهد. بنابراین، برای توسعه کاربرد برخط مدل حاضر، پیشنهاد می‌شود نمونه‌هایی با سطوح مختلف رطوبت وارد واسنجی شوند یا از روش‌هایی مانند تصحیح انتقال، EPO، به‌روزرسانی مدل و استانداردسازی دستگاه استفاده شود.

Rasooli *et al.* (۲۰۲۳) نیز نشان دادند که Vis–NIR می‌تواند به‌عنوان روشی سازگار با محیط‌زیست برای پایش تغییرات پدومحیطی و ارزیابی خاک در محیط‌های متنوع ایران استفاده شود. همسویی نتایج آن مطالعه با پژوهش حاضر، اهمیت توسعه کتابخانه‌های طیفی منطقه‌ای و ملی را برجسته می‌کند. کتابخانه‌های طیفی با پوشش دامنه وسیعی از خاک‌ها، اقلیم‌ها و شیوه‌های مدیریت می‌توانند هزینه اندازه‌گیری‌های مرجع را کاهش دهند؛ با این‌حال، کاربرد آن‌ها مستلزم انتخاب هوشمند نمونه‌های واسنجی، کنترل تغییرات میان دستگاه‌ها و به‌روزرسانی محلی مدل است.

یکی دیگر از نقاط قوت پژوهش، استفاده از پنج ناحیه مختلف هر نمونه و چند تکرار طیفی است. این روش می‌تواند ناهمگنی سطح نمونه را کاهش دهد و طیف نماینده‌تری از هر خاک تولید کند. در مجموع، نتایج نشان می‌دهد طیف سنجی Vis–NIR می‌تواند به‌عنوان یک روش سریع، غیرمخرب و کم‌مصرف برای غربالگری و برآورد فسفر قابل‌جذب خاک استفاده شود. این فناوری می‌تواند تعداد تجزیه‌های مرجع مورد نیاز را کاهش دهد و در برنامه‌های کشاورزی دقیق، مدیریت کود فسفری و تهیه نقشه‌های حاصلخیزی خاک کاربرد داشته باشد. با این‌حال، جایگزینی کامل روش اولسن تنها زمانی قابل توصیه است که عملکرد مدل با مجموعه‌ای کاملا مستقل، خاک‌های متعلق به مناطق جدید، شرایط رطوبتی متفاوت و دستگاه‌های دیگر تأیید شود.

نتیجه‌گیری و پیشنهادها

نتایج نشان داد که طیف‌سنجی مرئی–فروسرخ نزدیک، همراه با مدل‌سازی رگرسیون حداقل مربعات جزئی، قابلیت مناسبی برای پیش‌بینی فسفر قابل‌جذب نمونه‌های خاک هواخشک دارد. مدل مبتنی بر پیش‌پردازش SG+MSC توانست فسفر خاک را با R^2_{cv} =۰/۹۱، RMSECV=۱۶/۰۰، MAE= ۱۰/۹۴، Bias = −۲/۹۶×۱۰$^{-۵}$ و RPD= ۳/۲۸ پیش‌بینی کند. در مقابل، مشتق دوم به دلیل تقویت نویز و تضعیف روابط غیرمستقیم فسفر با طیف، دقت مدل را به‌طور محسوسی کاهش داد. این یافته‌ها بیانگر ظرفیت طیف سنجی Vis–NIR برای کاهش زمان، هزینه و مصرف مواد شیمیایی در تجزیه تعداد زیادی نمونه خاک است. لازم است تأکید شود که مدل حاضر با نمونه‌های هواخشک، خردشده و الک‌شده و تحت شرایط کنترل‌شده آزمایشگاهی توسعه یافته است. خشک‌کردن و الک‌کردن، اثر رطوبت و ناهمگنی سطح را کاهش می‌دهد و ممکن است عملکرد مدل را نسبت به شرایط مزرعه بهبود دهد. بنابراین، نتایج حاضر نباید به‌عنوان تأیید کاربرد مستقیم حسگر در مزرعه تفسیر شوند.

با توجه به نتایج، پیشنهادهای زیر ارائه می‌شود:

۱. تعداد نمونه‌ها و تنوع خاک‌ها از نظر اقلیم، بافت، ماده آلی و مدیریت زراعی افزایش یابد.
۲. اثر رطوبت خاک بر عملکرد مدل در شرایط مزرعه‌ای بررسی شود.
۳. روش PLS با الگوریتم‌های غیرخطی مانند ماشین بردار پشتیبان، جنگل تصادفی، شبکه عصبی و روش‌های یادگیری عمیق مقایسه شود.
۴. روش‌های انتخاب طول موج برای کاهش تعداد متغیرها و توسعه سامانه‌های کم‌هزینه و قابل‌حمل بررسی شوند.

۵. مدل‌های جداگانه منطقه‌ای با مدل کلی مقایسه شوند تا اثر تفاوت خاک‌های لاهیجان و جنوب تهران مشخص شود.

منابع

آقایی سعدی، مهرداد؛ مینایی، سعید و جمشیدی، بهاره (۱۴۰۰). مطالعه اثر پیش پردازش‌های مختلف طیفی بر مدل‌های پیش‌بینی غیرمخرب عیار چغندرقند و طراحی سامانه تشخیص مبتنی بر طیف سنجی مرئی/ فروسرخ نزدیک. *پژوهش‌های مکانیک ماشینهای کشاورزی*. ۱۰ (۱)، ۸۱-۶۹. https://jrmam.sku.ac.ir/article_10015.html.

آقایی سعدی، مهرداد؛ مینایی، سعید؛ جمشیدی، بهاره و عبداللهیان نوقابی، محمد (۱۳۹۷). سنجش غیرمخرب عیار چغندرقند با بهره‌گیری از ترکیب طیف‌سنجی فروسرخ نزدیک (NIR) با روش‌های شیمی‌سنجی. *مهندسی بیوسیستم ایران*. ۴۹ (۱)، ۱۸-۹. htpps://doi: 10.22059/ijbse.2017.208413.664785.

Aghaei Sadi, M., Minaei, S., & Jamshidi, B. (2021). Effect of spectral Pre-processing Methods on developed models for nondestructive prediction of sugar content in Sugar Beet and design of a detection system based on Vis/NIR spectroscopy. *Journal of Researches in Mechanics of Agricultural Machinery*, *10(1)*, 69–81. https://jrmam.sku.ac.ir/article_10015.html. (In Persian).

Aghaei Sadi, M., Minaei, S., Jamshidi, B., & Abdollahian Noghabi, M. (2018). Non-Destructive Evaluation of Sugar content Using a Combination of Near-Infrared Spectroscopy (NIRS) and Chemometrics Methods. *Iranian Journal of Biosystems Engineering*, *49(1)*, 9-18. htpps://doi: 10.22059/ijbse.2017.208413.664785. (In Persian).

Heise, H. M, & Winzen, R. (2001). Chemometrics in near-infrared spectroscopy. In H. W. Siesler, Y. Ozaki, S. Kawata , &H. M. Heise (Eds.). *Near-infrared spectroscopy: Principles, instruments, applications* (pp. 125–162). Wiley-VCH . https://doi.org/10.1002/9783527612666.ch07

Jamshidi, B., Minaei, S., Mohajerani, E., & Ghassemian, H. (2014). Prediction of soluble solids in oranges using visible/near-infrared spectroscopy: effect of peel. *International Journal of Food Properties*, *17(7)*, 1460-1468. https://doi.org/10.1080/10942912.2012.717332.

Kodaira, M., & Shibusawa, S. (2013). Using a mobile real-time soil visible-near infrared sensor for high resolution soil property mapping. *Geoderma*, *199*, 64–79. https://doi.org/10.1016/j.geoderma.2012.09.007.

Kuo, S., 1996. Phosphorus. In: D.L. Sparks, A.L. Page, P.A. Helmke, & R.H. Loeppert (Editors), Methods of Soil Analysis Part 3—Chemical Methods. SSSA Book Series. *Soil Science Society of America, American Society of Agronomy*, Madison, WI, pp. 869-919.

Malley, D. F., Yesmin, L., Wray, D., & Edwards, S. (1999). Application of near-infrared spectroscopy in analysis of soil mineral nutrients. *Communications in Soil Science and Plant Analysis*, *30(7-8)*, 999-1012.

Mirzaei, S., Ali Darvishi Boloorani, Hossein Ali Bahrami, Seyed Kazem Alavipanah, Alijafar Mousivand, & Abdul M. Mouazen. (2022). Minimising the effect of moisture on soil property prediction accuracy using external parameter orthogonalization. *Soil and Tillage Research*, 215l 105225. https://doi.org/10.1016/j.still.2021.105225.

Mohamed, E. S., Baroudy, A. A. E., El-Beshbeshy, T., Emam, M., Belal, A. A., Elfadaly, A., ... & Lasaponara, R. (2020). Vis-nir spectroscopy and satellite landsat-8 oli data to map soil nutrients in arid conditions: A case study of the northwest coast of egypt. *Remote Sensing*, *12(22)*, 3716.

Munnaf, M. A., Guerrero, A., Nawar, S., Haesaert, G., Van Meirvenne, M., & Mouazen, A. M. (2021). A combined data mining approach for on-line prediction of key soil quality indicators by Vis-NIR spectroscopy. *Soil and Tillage Research*, 205(June 2020), 104808. https://doi.org/10.1016/j.still.2020.104808.

Nanni, M. R., & Demattê, J. A. M. (2006). Spectral reflectance methodology in comparison to traditional soil analysis. *Soil Science Society of America Journal*, *70(2)*, 393–407. https://doi.org/10.2136/sssaj2003.0285.

Nawar, S., & Mouazen, A. M. (2018). Optimal sample selection for measurement of soil organic carbon using on-line vis-NIR spectroscopy. *Computers and Electronics in Agriculture*, 151(May), 469–477. https://doi.org/10.1016/j.compag.2018.06.042.

Peng, Y., Zhao, L., Hu, Y., Wang, G., Wang, L., & Liu, Z. (2019). Prediction of soil nutrient contents using visible and near-infrared reflectance spectroscopy. *ISPRS international journal of geo-information*, *8(10)*, 437.

Pudełko, A., & Chodak, M. (2020). Estimation of total nitrogen and organic carbon contents in mine soils with NIR reflectance spectroscopy and various chemometric methods. *Geoderma*, 368(February). https://doi.org/10.1016/j.geoderma.2020.114306.

Ramirez-Lopez, L., Wadoux, A. C., Franceschini, M. H., Terra, F. S., Marques, K. P. P., Sayão, V. M., & Demattê, J. A. M. (2019). Robust soil mapping at the farm scale with vis–NIR spectroscopy. *European Journal of Soil Science*, *70*(2), 378-393.

Rasooli, N., Farpoor, M.H., Mahmoodabadi, M., & Esfandiarpour-Boroujeni, I. (2023). Vis-NIR spectroscopy as an eco-friendly method for monitoring pedoenvironmental variations and pedological assessments in Lut Watershed, Central Iran. *Soil and Tillage Research*, 233, 105808. https://doi.org/10.1016/j.still.2023.105808.
Recena, R., Fernández-Cabanás, V. M., & Delgado, A. (2019). Soil fertility assessment by Vis-NIR spectroscopy: Predicting soil functioning rather than availability indices. *Geoderma*, *337*, 368-374.
Sanogo, S., Konare, H., Ba, A., & Diakite, T. (2018). Capability of visible-near infrared spectroscopy in estimating soils carbon, potassium and phosphorus. *Optics and Photonics Journal*, *8*(05), 123.
Shao, Y., & He, Y. (2011). Nitrogen, phosphorus, and potassium prediction in soils, using infrared spectroscopy. *Soil Research*, *49*(2), 166-172.
Shepherd, K. D., & Walsh, M. G. (2002). Development of reflectance spectral libraries for characterization of soil properties. *Soil Science Society of America Journal*, 66(3), 988–998. https://doi.org/10.2136/sssaj2002.9880.
Stenberg, B. (2010). Effects of soil sample pretreatments and standardised rewetting as interacted with sand classes on Vis-NIR predictions of clay and soil organic carbon. *Geoderma*, *158(1–2)*, 15–22. https://doi.org/10.1016/j.geoderma.2010.04.008.

ویراستاری نشده