VIS-NIR Reflectance Spectroscopy as an Alternative Method for Rapid Estimation of Soil Available Potassium

Article *in* Journal of the Indian Society of Soil Science · December 2020

DOI: 10.5958/0974-0228.2021.00009.8

CITATIONS
0

READS
92

6 authors, including:

Bhabani Prasad Mondal
Indian Agricultural Research Institute
16 PUBLICATIONS **14** CITATIONS
SEE PROFILE

Bharpoor Singh Sekhon
Punjab Agricultural University
35 PUBLICATIONS **127** CITATIONS
SEE PROFILE

Arijit Barman
Central Soil Salinity Research Institute
20 PUBLICATIONS **77** CITATIONS
SEE PROFILE

Arghya Chattopadhyay
Banaras Hindu University
23 PUBLICATIONS **55** CITATIONS
SEE PROFILE

Journal of the Indian Society of Soil Science, Vol. 68, No. 3, pp 323-330 (2020)
DOI: 10.5958/0974-0228.2021.00009.8

VIS-NIR Reflectance Spectroscopy as an Alternative Method for Rapid Estimation of Soil Available Potassium

Bhabani Prasad Mondal[1]*, Bharpoor S. Sekhon, Priya Paul[2], Arijit Barman[3], Arghya Chattopadhyay[4] and Nilimesh Mridha[5]

Department of Soil Science, Punjab Agricultural University, Ludhiana, 141004, Punjab

Potassium (K) is an important macronutrient for crop plant and plays a crucial role in crop production. Therefore, accurate and rapid estimation of soil available K is necessary for judicious application of available K in an intensively cropped region. However, traditional soil chemical analysis for assessing soil available K is very much laborious, expensive and time consuming. The visible near-infrared (VIS-NIR) reflectance spectroscopy is considered as a promising alternative technique for rapid, non-destructive and ecofriendly estimation of available K and other soil properties. An experiment was carried out in an intensively cultivated region of Ludhiana district of Punjab to investigate the potential of VIS-NIR technique for accurate prediction of available K using multivariate model. A total of 170 georeferenced surface soil samples (0-15 cm) were collected from the study site for both chemical and spectral analysis of available K. A popular statistical technique namely, partial least square regression (PLSR) was employed to develop spectral model for K prediction. Important statistical diagnostics like coefficient of determination (R^2), root mean square error (RMSE) and residual prediction deviation (RPD) were used to evaluate the efficacy of prediction model. The results showed that the R^2 and RMSE and RPD values were 0.41, 0.09 and 1.44, respectively for independent validation dataset of PLSR model. The RPD value indicated acceptable prediction accuracy for soil available K with PLSR model. Comparatively lower performance of the studied prediction model could be ascribed to the less variation in the collected spectra of soil samples and the use of linear multivariate model. Therefore, the study suggested to explore advanced non-linear data mining techniques for achieving better prediction accuracy for soil available K.

Key words: Available potassium, reflectance spectroscopy, root mean square error, residual prediction deviation

Accurate estimation of soil available potassium (K) is crucial from plant nutritional point of view because it is one of the major plant nutrients and hence improving plant growth and yield. Less than required K-fertilizer use impairs crop yield, but more than required not only causes the 'luxury consumption' of K but also affecting the chemistry of micronutrients availability in soils. For a successful crop production, balanced amount of K-fertilizer application is also necessary along with other major nutrient containing fertilizers like nitrogen (N) and phosphorus (P). In practical application, the real time monitoring tools like SPAD meter, leaf colour chart (LCC) often enable the farmers to correct the deficiency of N during the crop growing seasons. The K-fertilizers are applied as basal dose to meet the crop requirement. However, the application of K-fertilizer is often neglected owing to K rich mineralogical make-up of the soil (Mondal and Sekhon 2019). Consequently, soil K level is continuously decreasing, which will bring a major threat to our farming system in near future. Thus, there is need to promote need-based application of K-fertilizers in the crop field. The recommendation rate for K-fertilizer needs to be further refined by giving priority to the soil spatial variability issues (Reza *et al.* 2019). Soil testing methods should be fast,

*Corresponding author (Email: bpmondal27@gmail.com)
Present address
[1]Division of Agricultural Physics, [2]Division of Plant Physiology, ICAR-Indian Agricultural Research Institute, New Delhi, 110012
[3]Division of Natural Resource Management, ICAR-CSSRI, Karnal, 132001, Haryana
[4]Department of Soil Science and Agricultural Chemistry, IAS, BHU, Varanasi, 221005, Uttar Pradesh
[5]ICAR-National Institute of Natural Fibre Engineering and Technology, Kolkata, 700040, West Bengal

economical and address the spatial variability issue by allowing more intensive sampling. However, the conventional chemical analysis for estimating soil available K is laborious, expensive, destructive in nature. Therefore, an alternative technique is required for successful quantification of soil available K for its optimum use in the crop field. Diffuse reflectance spectroscopy technique especially visible-near infrared (VIS-NIR) spectroscopy is considered as a potential alternative technique to the expensive traditional laboratory-based chemical analysis for rapid and accurate estimation of soil available K and other soil properties (Viscarra Rosssel *et al*. 2006). This technique provides rapid, reliable, non-destructive, reproducible and eco-friendly estimation of soil available K both in lab-based and in-situ conditions (Grass *et al.* 2014). This method is used to obtain the spectral data in the spectral region of 350-2500 nm form the studied soil sample through interacting the electromagnetic wave with the sample constituents (Pasquini 2003). Depending on the composition of soil samples, light energy is absorbed or reflected by the individual molecular bond such as C-H, N-H, O-H, *etc*. through bending and stretching mechanisms based on the energy difference between two successive quantum energy levels (Knadel *et al.* 2017). Nonetheless, the soil VIS-NIR spectra are largely non-specific and produce broad absorption features in these spectral regions due to the weak overtones and the vibrations of those different functional groups (Viscarra Rossel and Behrens 2010). Therefore, multivariate statistical techniques are required to extract those absorption features mathematically in order to correlate them with the measured soil property (here soil available K) for spectral model development (Stenberg *et al.* 2010). Among the various multivariate techniques, the partial least square regression (PLSR) is the most commonly used linear regression technique, efficient in handing multicollinearity in the dataset (Viscarra Rosssel *et al*. 2006). The PLSR often produces acceptable accuracy in major soil property prediction by adding more latent variables (LVs) in the dataset (Thissen *et al.* 2004). Several statistical indices like coefficient of determination (R^2), root mean square error (RMSE), residual prediction deviation (RPD) *etc*. can be utilized to evaluate the predictive performance of the newly developed spectral model (Sarathjith *et al.* 2016). Several researchers have used PLSR technique for prediction of various soil properties including available, exchangeable and total K from VIS-NIR spectra throughout the world (Grass *et al.* 2014). However, in India this technique is quite new and literature show that very few investigations have been carried out in India for quantitative estimation of soil properties using VIS-NIR reflectance spectroscopy (Mondal and Sekhon 2019).

Keeping the above-mentioned things in view, a systematic study was carried out to evaluate the potential of the VIS-NIR reflectance spectroscopy for the prediction of soil available K through PLSR technique in an intensively cultivated region of Ludhiana district of Punjab, India.

Materials and Methods

Study Area and Soil Sampling

The study area was located near Ladian village of Ludhiana district of Punjab (are 30°57'330" N and 75°47'675" E) at 250 m above mean sea level. A total of 170 georeferenced surface soil samples (0-15 cm) were collected from a farm of the study area supporting berseem-based, rice-wheat, poplar-wheat based agroforestry systems for last 10 years. The study site can be characterized by semi-arid climatic conditions with mean annual temperature of 45-50 °C in summer and 1-2 °C in winter. The average annual rainfall is 700-800 mm. The study area belongs to the Indo-Gangetic alluvium plain of Quaternary age and it has sandy-loam to clayey soil texture with a neutral to alkaline soil pH (7.6-8.7). The area supports a wide variety of agricultural crops. Rice-wheat is the predominant cropping system of this study area.

Collected soil samples were air-dried and ground with wooden pestle and mortar and passed through 2 mm sieve before analysis. These processed soil samples were divided into two major parts; one part was used for laboratory-based soil chemical analysis (soil available K estimation) and another part for spectral analysis.

Chemical Analysis

Estimation of soil available K was carried out by the neutral normal ammonium acetate method (Schollenberger and Simon 1945). Soil was treated with neutral 1 *N* ammonium acetate solution, adjusted to pH 7.0 and shaken for 5 min with a mechanical shaker. After shaking, the extract was filtered and the concentration of K present in filtrate was determined with the help of a flame photometer (Elico CL 361).

Spectral Analysis

Field spec spectroradiometer (PAN analytical Spectral Devices Inc., Boulder, Colorado, USA, ASD)

was used to obtain the diffuse reflectance spectra of the air-dried and finely ground soil samples in the spectral range of 350-2500 nm at 1 nm sampling interval, yielding 2151 spectral data points. Each soil sample was placed in a petridis, and a contact probe (PANalytical Inc. Colorado, USA) having in-built high intensity light source (6.5 W halogen lamp), connected with spectroradiometer through fiber-optic cable was utilized to measure the spectra from the soil sample. The contact probe facilitates the direct measurement of spectra from the surface of studied soil samples, and thus eliminating the noise and other light interferences from the outside. Before taking the spectral measurement, the sensor of the instrument was calibrated by capturing white reflectance spectra using standard spectralon, made up with barium sulphate ($BaSO_4$). Average of thirty scanning spectra were recorded for each sample to reduce the noise level in the spectral data. The upper surface of each soil sample was levelled-off with a spatula for proper spectral measurement of the sample.

Spectral Preprocessing

All kinds of spectral preprocessing like spectral smoothening, first order derivative of reflectance spectra for deriving more spectral information were carried out in the Unscrambler 10.4 software (CAMO Technologies, Inc. Norway). First the huge number of spectral data points (*i.e.* 2151) were reduced by resampling at 10 nm interval due to high spectral collinearity among the contiguous data points. Spectral data were preprocessed using Savitzky-Golay (Savitzky and Golay 1964) filtering and smoothening algorithm with first order derivative in order to reduce the base line variations and to derive more spectral information for better modeling purpose (Xu *et al.* 2018).

Descriptive Statistical Analysis

Descriptive statistical analysis of soil available K was done using a statistical software SAS and also in excel spreadsheet. Several statistical indices like mean, minimum, maximum, median, per cent coefficient of variation (% CV), skewness and kurtosis were computed for characterizing the distribution of soil available K.

Spectral Model Development

Wold *et al.* (1983) developed partial least square regression (PLSR) technique for quantitative analysis of reflectance spectroscopic data used in various applications (Viscarra Rossel and Behrens 2010). It is the most commonly used linear multivariate technique for spectral modeling (Viscarra Rossel *et al.* 2006). It has the potential to obviate the multicollinearity of the dataset. It is also capable of handling small to medium level non linearity, present in the dataset and thus it can improve the predictive performance of the model through computing and adding more LVs (successive orthogonal factors) in each regression step which maximizes the covariance between predictor and response variables (Thissen *et al.* 2004). In this study, PLSR technique was employed to develop and calibrate the spectral model through correlating the derivative reflectance values with lab-measured soil available K, in Unscrambler 10.4 software (CAMO Technologies, Inc. Norway). In order to develop the spectral model, the whole dataset (n=170) was randomly divided into two subsets *viz*. calibration and validation datasets. Two-third (n=113) of the datasets were used to calibrate the model and rest one-third data points (n=57) were used to validate the model independently. A leave-one-out cross-validation technique was also used in the calibration dataset to observe the prediction accuracy of the model within the calibration dataset itself and to avoid underfitting or overfitting of the dataset (Viscarra Rossel *et al.* 2006). The independent validation was also performed to test the accuracy of the model using one-third of the dataset mentioned above. The important wavelengths for better K prediction were identified based on the regression coefficients values, obtained from the calibration models of PLSR. Those wavelengths were selected, where peak maxima (either positive or negative) were observed in the plot of regression coefficients vs wavelength (Viscarra Rossel and Behrens 2010). These key wavelengths help in generalizing the interpretability of the spectral model.

Model Evaluation

Three statistical diagnostics namely, coefficient of determination (R^2) (Equation 1), root mean square error (RMSE) (Equation 2) and residual prediction deviation (RPD) (Equation 3) were used to evaluate the prediction accuracy of the newly developed spectral model using PLSR technique. The mathematical expressions of these evaluation indices are described in terms of equation *viz*.

$$R^2 = 1 - \left(\frac{\sum_i^n (Y_{pred} - Y_{meas})^2}{\sum_i^n (Y_i - Y_{mean})^2} \right) \quad \ldots(1)$$

Table 1. Descriptive statistics of soil available potassium (kg ha^{-1}) in the study area

Statistical parameters	Values
Minimum	134
Maximum	302
Mean	179
Median	176
Coefficient of Variation (%CV)	17.7
Skewness	1.86
Kurtosis	5.54

$$RMSE = \sqrt{\frac{\sum_i^n (Y_{pred} - Y_{meas})^2}{n-1}} \quad \ldots(2)$$

$$RPD = SD/RMSEP \quad \ldots(3)$$

Where, Y_{pred} indicates the predicted values, Y_{mean} refers to mean of measured values, Y_{meas} refers to measured values and n is the number of measured or predicted values (number of samples) and SD stands for standard deviation. Generally, higher R^2 value indicates a good prediction accuracy of the model. On the other hand, lower RMSE value suggests a better prediction model. Viscarra Rossel *et al.* (2006) proposed a criterion to evaluate the prediction accuracy of the spectral model based on RPD values. According to this criterion, the RPD values of less than 1, between 1.0-1.4, 1.4-1.8, 1.8-2.0, 2.0-2.5 and greater than 2.5 indicates very poor, poor, fair, good, very good and excellent model performances, respectively.

Results and Discussion

Descriptive Statistics for Soil Available K

The summary of the descriptive statistical analysis for soil available K (kg ha^{-1}) is represented in table 1. The statistical summary showed that the available K varied from 134 to 302 kg ha^{-1} with a mean value of 179 kg ha^{-1} (CV of 17.7%). The average distribution of available K was also higher and it could be ascribed to higher clay content of the soil. Wilding *et al.* (1985) proposed a criterion to determine the variability of the dataset based on CV values. According to it, CV values more than 35%, in between 15-35% and less than 15% correspond to the high, medium and low levels of variability of the soil property. Thus, the classical statistical analysis showed a moderate variability of available K in terms %CV. The statistical summary also indicated a positive skewed distribution of available K. Medium range variability of soil available K sometimes indicate moderate performance of the prediction model.

Spectral Reflectance Characteristics of Soil Samples

The spectral reflectance behavior of soil samples was represented through spectral reflectance curve (Fig. 1a).

The overall shape of the spectral curve is similar due to presence of some similar types of spectrally active constituents called chromophores in the soil (Xu *et al.* 2018). However, variation in the absorption features of the spectral curve at different wavelengths could be attributed to the heterogeneity of the soil systems and its properties. The spectral response is very much sensitive to the changes of soil moisture content and the organic matter content. Water absorption bands at 1400 and 1900 nm and hydroxyl (-OH) absorption bands at 2200 nm are two major characteristic absorption features in this spectral reflectance characteristics curve (Fig. 1a). Such absorption bands at those specific wavelengths decrease the spectral reflectance values. The presence of interlayer water molecules in phyllosilicate minerals like kaolinite, smectite along with hydroxyl group of aluminium (Al) or iron (Fe) and their stretching and vibrations bring significant absorption in the spectral curve at higher wavelength regions at approximately 2204 and 2280 nm (Dematte *et al.* 2017). The first derivative reflectance spectra (Fig. 1b) were also calculated for deriving nuanced spectral

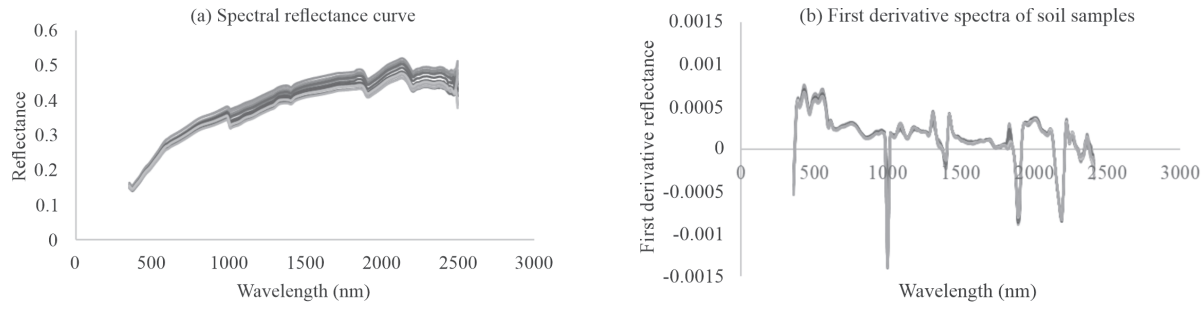

Fig. 1. Spectral reflectance characteristics of soil samples (a) reflectance curve, (b) first derivative reflectance curve

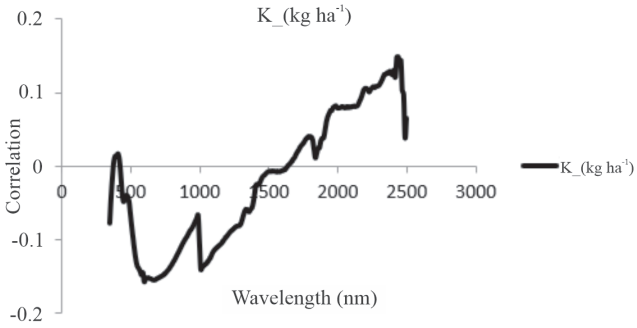

Fig. 2. Correlation between soil available K and reflectance at various wavelengths

information for better modeling purpose. Actually the first derivative reflectance indicates the changes in spectral reflectance values at per unit change in wavelength. Therefore, the spectral derivatives in conjunction with other normalization treatments were applied to reduce light scattering and baseline variations in the collected spectra and to enhance the spectral features (Xu *et al.* 2018).

Correlation Between Reflectance Spectra and Available K

The statistical correlation analysis was performed to determine the correlation between reflectance spectra and soil available K. Statistical correlation between available K and reflectance spectra was graphically represented (Fig. 2). Available K exhibited both positive and negative correlation with the spectra in the spectral region of 350-2500 nm. The correlation graph showed positive correlation around 400 nm and above 1500 nm. However, between 400 nm to 1500 nm the correlation between reflectance spectra and soil available K was negative (Fig. 2). In this study, the correlation coefficient values between reflectance spectra and the soil available K were ranged from -0.15 to 0.17.

Performance of PLSR Model for Lab-based Available K Prediction

The PLSR technique was applied to develop the spectral model for the prediction of soil available K through correlating first derivative reflectance spectra with measured soil K. Actually, PLSR model was used to calibrate and validate the spectroscopic data against measured soil available K. In calibration dataset, the first derivative reflectance spectra were calibrated with soil available K utilizing PLSR model. Some statistical indices *viz*. number of explanatory factors *i.e.* latent variables, R^2, RMSE values were chosen to specify the nature of distribution of the calibration dataset. They also reflected that how the model was well trained with the calibration dataset. A leave one out cross validation was carried out using the same dataset to check the accuracy within the calibration dataset. The scatter plot of model-predicted *vs* lab-measured available K in calibration and cross validation datasets was illustrated simultaneously in the same figure (Fig. 3). The logarithmic transformations were applied to available K data for normalizing the data before calibration model development.

The number of LVs or factors used to develop calibration model were nine. The R^2 and RMSE values for calibration dataset were 0.77 and 0.04,

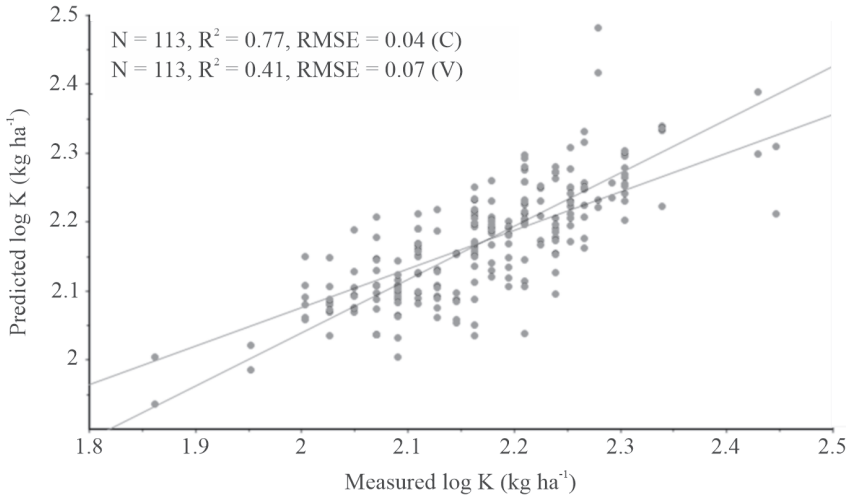

Fig. 3. Scatter plot of measured *vs* VIS-NIR spectra predicted soil available K content using PLSR model; (C) calibration and (V) cross validation

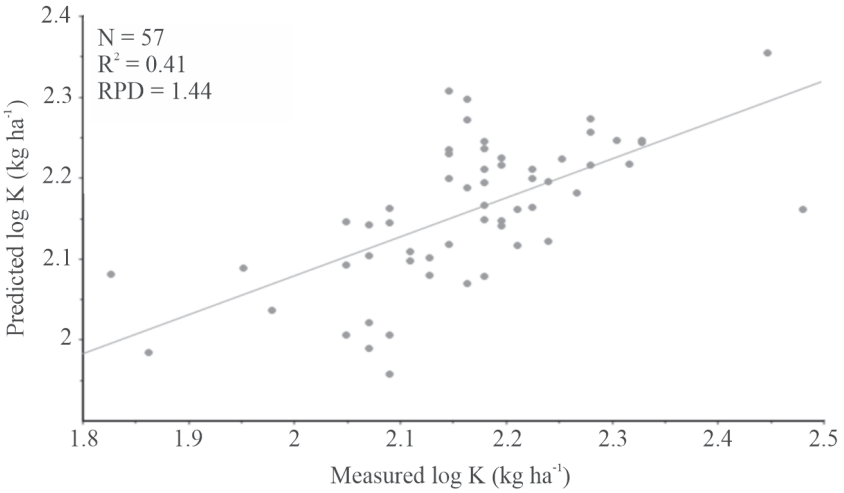

Fig. 4. Scatter plot of measured vs predicted values of soil available K content (logarithmic transformation applied) in independent validation dataset

respectively. Higher R^2 and lower RMSE values indicate better model performance during the training of the spectral model. During cross validation, the R^2 and RMSE values corresponded to 0.41 and 0.07, indicated moderate performance of the model. In case of independent validation (using 57 samples), it maintained same accuracy (R^2 = 0.41, RMSE = 0.09 and RPD = 1.44). The scatter plot for independent validation of soil available K was portrayed in fig. 4.

Such type of similar prediction accuracies was also reported by Xu *et al.* (2018), who obtained R^2 = 0.58, RMSE = 3.84 and RPD = 1.52 with 9 LVs using PLSR model. The prediction accuracies of our validation dataset could be categorized as 'fair' prediction for available K with an acceptable accuracy as per the criteria proposed by Viscarra Rossel *et al.* (2006). Moderate performance of spectral model for soil available K prediction could be attributed to various factors such as the form of existence of the nutrient in the soil system and the nature of multivariate model used. The successful prediction of any soil property depends on the aforementioned factors (Wenjun *et al.* 2014). Total form of any nutrients like total N, total P or total K can be better predicted than their available forms because the amount of available forms of those nutrients are very less than their total forms (say available N accounts only 5% of total N and for available K *i.e.* soil solution K + exchangeable K consist of 2% of total K), which have very slight influences on the absorption characteristics of the VIS-NIR spectra (Wenjun *et al.* 2014). Therefore, little effect of available K on the absorption behavior of soil diffuse reflectance spectra is one of the primary reasons to get lower prediction accuracy by the model. Besides, several researchers also reported very poor prediction performance of both available and total K (Wenjun *et al.* 2014). They mentioned that not only soil K but also the prediction accuracy of soil P would be lower as these two soil parameters did not have direct spectral relationships with the VIS-NIR region of the spectra. Selection of an appropriate multivariate technique is often a prerequisite to obtain a better performance for soil property prediction (Viscarra Rossel and Behrens 2010). The performance of non-linear model such as support vector machine regression (SVMR), random forest (RF), multivariate adaptive regression splines (MARS), artificial neural network (ANN) often perform better than linear models like PLSR and principal component regressions (PCR) due to existence of non-linear spectral relationship between soil spectral data and soil attributes (Araújo *et al.* 2014). Qi *et al.* (2018) also reported better prediction of soil N, P and K using the above mentioned advanced non-linear data mining techniques. Therefore, our study suggests to explore these alternative non-linear techniques to obtain better prediction accuracies for soil available K. Besides, the prediction capability of a model is often determined by the physical and chemical constituents *i.e.* chromophores of soil systems (Shepherd and Walsh 2002). Heterogeneity of the soil system and the nature of the dataset used affect the prediction performance of the spectral model (Nawar and Mouazen 2017). Relatively less promising performance of the studied model could also be

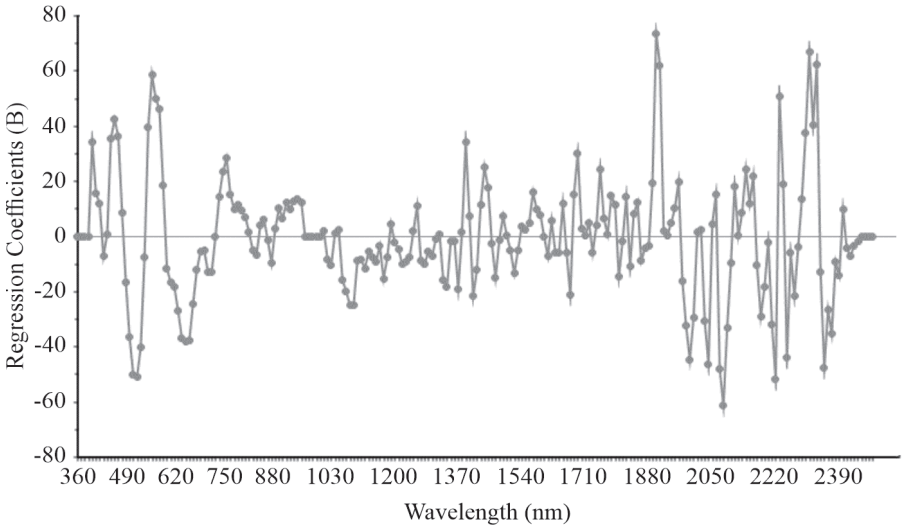

Fig. 5. Partial least square regression coefficients (beta coefficients) for soil available K at different wavelengths

ascribed to the less variation in soil reflectance spectra, measured by spectroradiometer.

Important Wavebands for Soil Available K Prediction

The important wavebands, where better spectral responses appear for a particular parameter should be identified for obtaining better prediction accuracy of the spectral model. This method is also known as feature selection technique (Viscarra Rossel and Behrens 2010). The feature selection technique increases the interpretability of the multivariate models for a particular soil property prediction. Such technique helps in improving the model performance through removing the uninformative variables from the dataset and retaining the informative variables in the dataset (Xu *et al.* 2018). The feature selection or important waveband selection can be done following various methods such as sensitivity analysis, PLS-variable importance in the projection (VIP) method, beta regression of PLSR (PLS-BETA method) *etc*. (Chong and Jun 2005). Among these techniques, PLS-BETA method was applied here to identify the key spectral wavelengths for soil available K prediction. According to this method, the key wavelengths were spotted out based on the regression coefficient values. The wavelength, corresponded to higher regression coefficient value (either positive or negative) was designated as key spectral wavelength for available K prediction.

The significant spectral wavebands for soil available K prediction were laid around 370, 490, 500, 555, 1370, 1880, 2060, 2220 and 2285 nm (Fig. 5).

These findings were also supported by Xu *et al.* (2018), who reported similar types of spectral bands for soil K prediction. As this parameter does not possess any direct relationship with VIS-NIR spectra, the statistical techniques are the only ways to recognize the crucial bands for soil available K prediction using spectroscopic technique.

Conclusions

To conclude, the present study highlighted the feasibility of using diffuse reflectance spectroscopy particularly VIS-NIR spectroscopy for rapid soil available K prediction. It showed that PLSR technique could provide fair prediction of available K with an acceptable accuracy (R^2 =0.41, RMSE= 0.09, RPD= 1.44). The key spectral wavelengths for soil available K estimation such as 370, 490, 500, 555, 1370, 1880, 2060, 2220 and 2285 nm were identified. Moderate performance of the spectral model could be attributed to less variation in the collected soil spectra and the use of linear multivariate model. Therefore, the study suggested to employ non-linear multivariate models like SVMR, RF, ANN *etc*. to achieve better prediction accuracy for soil available K.

References

Araújo, S.R., Wetterlind, J., Demattê, J.A.M. and Stenberg, B. (2014) Improving the prediction performance of a large tropical VIS-NIR spectroscopic soil library from Brazil by clustering into smaller subsets or use of data mining calibration techniques. *European Journal of Soil Science* **65**, 718-729.

Chong, I.G. and Jun, C.H. (2005) Performance of some variable selection methods when multicollinearity is present. *Chemometrics and Intelligent Laboratory Systems* **78**, 103-112.

Demattê, J.A.M, Horák-Terra, I., Beirigo, R.M, Terra, F da S., Marques, K.P.P, Fongaro, C.T, Silva, A.C, and Vidal-Torrado, P. (2017) Genesis and properties of wetland soils by VIS-NIR-SWIR as a technique for environmental monitoring. *Journal of Environmental Management* **197**, 50-62.

Gras, J.P., Barthès, B.G., Mahaut, B. and Trupin, S. (2014) Best practices for obtaining and processing field visible and near infrared (VNIR) spectra of top soils. *Geoderma* **214-215**, 126-134.

Knadel, M., Gislum, R., Hermansen, C., Peng, Y., Moldrup, P., de Jonge, L.W. and Greve, M.H. (2017) Comparing predictive ability of laser-induced breakdown spectroscopy to visible near-infrared spectroscopy for soil property determination. *Biosystem Engineering* **156**, 157-172.

Mondal, B.P. and Sekhon, B.S. (2019) Using diffuse reflectance spectroscopy for assessment of soil phosphorus status of an intensively cropped region. *Agricultural Research Journal* **56**, 657-661.

Nawar, S. and Mouazen, A.M. (2017) Predictive performance of mobile vis-near infrared spectroscopy for key soil properties at different geographical scales by using spiking and data mining techniques. *Catena* **151**, 118-129.

Pasquini, C. (2003) Near infrared spectroscopy: Fundamentals, practical aspects and analytical applications. *Journal of the Brazilian Chemical Society* **14**, 198-219.

Qi, H., Paz-Kagan, T., Karnieli, A., Jin, X. and Li, S. (2018) Evaluating calibrating methods for predicting soil available nutrients using hyperspectral VNIR data. *Soil and Tillage Research* **175**, 267-275.

Reza, S.K., Ray, P., Ramachandran, S., Bandyopadhyay, S., Mukhopadhyay, S., Sah, K.D., Nayak, D.C., Singh, S.K. and Ray, S.K. (2019) Spatial distribution of soil nitrogen, phosphorus and potassium contents and stocks in humid subtropical North-eastern India. *Journal of the Indian Society of Soil Science* **67**, 12-20.

Sarathjith, M., Das, B., Wani, S., Sahrawat, K.L. and Gupta, A. (2016) Comparison of data mining approaches for estimating soil nutrient contents using diffuse reflectance spectroscopy. *Current Science* **110**, 1031-1037.

Savitzky, A. and Golay, M.J.E. (1964) Smoothing and differentiation of data by simplified least squares procedure. *Analytical Chemistry* **36**, 1627-1639.

Schollenberger, C.J. and Simon, R.H. (1945) Determination of exchange capacity and exchangeable bases in soil by ammonium acetate method. *Soil Science* **59**, 13-24.

Shepherd, K.D. and Walsh, M.G. (2002) Development of reflectance spectral libraries for characterization of soil properties. *Soil Science Society of America Journal* **66**, 988-998.

Stenberg, B., Viscarra Rossel, R.A., Mouazen, A.M. and Wetterlind, J. (2010). Visible and near infrared spectroscopy in soil science. *Advances in Agronomy* **107**, 163-215.

Thissen, U., Pepers, M., Ustun, B., Melssen, W.J. and Buydens, L.M.C. (2004) Comparing support vector machines to PLS for spectral regression applications. *Chemometrics and Intelligent Laboratory Systems* **73**, 169-179.

Viscarra Rossel, R.A., Walvoort, D.J.J., McBratney, A.B., Janik, L.J. and 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.

Viscarra Rossel, R.A. and Behrens, T. (2010) Using data mining to model and interpret soil diffuse reflectance spectra. *Geoderma* **158**, 46-54.

Wenjun, J., Zhou, S., Jingyi, H. and Shuo, L. (2014) In situ measurement of some soil properties in paddy soil using visible and near-infrared spectroscopy. *PloS one* **9**, 105708.

Wilding, L.P. (1985) *Spatial variability: it's documentation, accommodation and implication to soil surveys*. In *Soil Spatial Variability*. Nielsen, D.R., Bouma, J. (Eds.), Pudoc, Wageningen, The Netherlands, pp. 166-194.

Wold, S., Sjöström, M. and Eriksson, L. (2001) PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems* **58**, 109-130.

Xu, S., Zhao, Y., Wang, M. and Shi, X. (2018) Comparison of multivariate methods for estimating selected soil properties from intact soil cores of paddy fields by VIS-NIR spectroscopy. *Geoderma* **310**, 29-43.

Received 21 August 2020; Accepted 30 September 2020