APPROACHES TO ESTIMATE PHOSPHORUS AND POTASSIUM CONTENT OF WHEAT LEAVES
#9390

Yousra El-Mejjaouy[1,2 *], Benjamin Dumont[3], Abdallah Oukarroum[2], Benoît Mercatoris[1]
[1]University of Liège, Gembloux agro-Bio Tech, Biosystems Dynamics and Exchanges, TERRA Teaching and Research Centre, Gembloux, Belgium. [2] University Mohammed VI Polytechnic (UM6P) – AgoBioSciences, Plant Stress Physiology Laboratory, Benguerir, Morocco.
[3] University of Liège, Gembloux Agro-Bio Tech, Pant sciences / Crop science, TERRA Teaching and Research Centre, Gembloux, Belgium
e-mail: Yousra.elmejjaouy@uliege.be

ABSTRACT

The assessment of plant nutrient status to provide sufficient fertilization for rapid and continuous uptake by plants has been based on visual diagnosis in the field, which is quick but demands a lot of experience and has a low operability. Visible near-infrared spectroscopy (VNIS) has shown to be a quick, non-destructive, accurate, and cost-effective analytical method in precision agriculture. In this study, we assessed the potential of this technology to predict phosphorus and potassium content in the wheat leaves using different multivariate regression methods. The hyperspectral and the reference measurements were taken from wheat plant leaves grown in a long-term fertilization trial under contrasted concentrations of phosphorus and potassium. The leaf proximal and hyperspectral data were collected using an ASD FieldSpec4 spectroradiometer operating in the spectral range from 350 to 2500 nm. Before conducting the analysis, the leaves spectra were preprocessed with a Savitzky–Golay smooth filter and a Standard Normal Variate normalization method. A total of 60 samples, collected between flowering and maturity stages, combined with the preprocessed spectra were used to develop support vector regression (SVR), random forest (RF), and principal component regression (PCR) prediction models for estimating leaves phosphorus content (LPC) and leaves potassium content (LKC). The entire sample set was randomly split into a training set (70%) and a test set (30%), and the performances of the different prediction models were compared using normalized root mean square error (NRMSE) and coefficient of determination (R2) in both cross-validation and testing processes. The results showed that LPC prediction models outperformed the LKC models, with high accuracies (R^2) in cross-validation in the order of 0.84, 0.85, and 0.79 for SVR, PCR, and RF, respectively. For potassium, the coefficient of determination of cross-validation was 0.64, 0.59, and 0.54 for SVR, PCR, and RF, respectively. The highest validation results were returned by the RF algorithm for both LPC and LKC predictions, with moderate R^2 values equal to 0.56 and 0.53, respectively. In the RF model, phosphorus and potassium in wheat leaves can be predicted with errors of 19 and 13%, respectively.

Keywords: Phosphorus, Potassium, Visible Near Infrared Spectroscopy, Random Forest, Support Vector Regression, Principal Component regression

INTRODUCTION

The visible, near infrared, and short-wave infrared (Vis-NIR-SWIR) spectroscopy is an emerging technique that has been widely used for soil and plant analysis. The spectroscopy

technique uses the reflectance in the wavelength range from the visible to the SWIR for a non-destructive and effective plant characterization (Ge et al. 2019; Muehling et al. 2015). It is based on the interaction between plant leaves and the incident light in the Vis-NIR-SWIR spectral regions. In plant leaf cells, photosynthetic pigments such as chlorophylls, carotenoids, and anthocyanins absorb strongly in the VIS region. On the other hand, phosphorus (P) and potassium (K) are important and essential elements for plant growth and development and strongly involved in physiological processes (Jiaying et al. 2022). P plays an important role in energy metabolism and K is the most abundant cation in plants. Therefore, it is important to estimate and monitor the contents of P and K. Their deficiency affects pigments biosynthesis in wheat which allows application of Vis-NIR spectroscopy for non-destructive prediction of phosphorus and potassium in plants (Thornburg et al. 2020). The P and K were estimated using spectral reflectance and partial least square regression (PLSR) with an R^2 of 0.643 and 0.541, respectively. However, the support vector machine regression (SVMR) method had higher accuracy with R^2 values of 0.722 and 0.704 for P and K, respectively (Zhai et al. 2013). Using PLS, Mishra et al. (2021) accurately predicted the K content dried and prepared samples of pepper leaves, with an R^2 value of 0.82 and an RMSEP of 0.53%. The PLS Regression Analysis is a commonly used method to model Vis-NIR data but the non-linear regressions have been proved to handle better the non-linear relationships that exist between the response variable and predictor variables in soil and plants (Nawar and Mouazen, 2019; Zhai et al., 2013). Moreover, It has been demonstrated that the choice of calibration method can impact the measurement accuracy when using visible and near-infrared spectroscopy (Vasques et al. 2008). In this study, we aimed to estimate indirectly phosphorus and potassium content in wheat leaves using leaf-proximal hyperspectral data and different multivariate modeling techniques.

MATERIAL AND METHODS

The study was conducted on a long-term trial located in Gembloux-Belgium (50.564121, 4.698802), called the law of the minimum trial. It has been installed up and monitored the same way since 1896 with an objective to study for the long term the effect of nitrogen, phosphate, and potassium on field crop yields. In our study we focused on the plots presenting an interesting contrasts and variability of phosphorus and potassium contents. The area of interest is 5 sets of 10 microplots, each set represent a different fertilization treatment. The treatments are NPK (nitrogen, phosphorus, and potassium fertilization), PK (phosphorus and potassium fertilization), NK (nitrogen and potassium fertilization), NP (nitrogen and phosphorus fertilization), and 0 treatment (no application of the three macronutrients).

The data collection consisted of spectra acquisition using ASD FieldSpec4 spectroradiometer (Malvern Panalytical Ltd., Formerly Analytical Spectral Devices) with spectral range from 350 to 2500 nm and a spectral sampling of 1 nm. The reflectance measurements were done on the flag leaf. In parallel to spectra acquisition, biomass samples were taken from the different treatments for chemical analysis of P and K leaves content. The recorded spectra were preprocessed prior the multivariate analysis, the spectra were smoothed using Savitzky-Golay filter and normalized using the Standard Normal Variate. After the preprocessing, three predictive models were developed for P and K using support vector regression (SVR), random forest (RF), and principal component regression (PCR). The models' performances were evaluated using the statistical metrics such as coefficient of determination (R^2), normalized root mean square error (NRMSE), and mean absolute error (MAE).

RESULTS AND DISCUSSION

LPC and LKC prediction

The cross-validation yielded higher coefficient of determination, lower normalized root mean square error, and lower mean absolute error for the three predictive models, whereas the validation results were moderate (**Table 1**). The random forest model for predicting LPC had the highest predictive performance in the validation process, a coefficient of determination equal to 0.56, a NRMSE of 0.19, and a MAE equal to 0.71 mg/g, followed by the principal component regression, and the support vector regression model with respective performances (R^2_v = 0.534, $NRMSE_v$ = 0.202, MAE_v = 0.875) and (R^2_v = 0.503, $NRMSE_v$ = 0.203, MAE_v = 0.897). Compared to phosphorus, the developed models for estimating potassium content in wheat leaves had lower predictive performances in both cross-validation and validation processes. The highest performance was obtained by RF model, it predicted LKC with a coefficient of determination of 0.531, $NRMSE_v$ of 0.114, and MAE_v of 1.856 mg/g. The principal component regression had the lowest performance (R^2_v = 0.359, $NRMSE_v$ = 0.196, and MAE_v = 3.41 mg/g). The phosphorus predictive models fared better in cross-validation and validation compared to the potassium predictive models, as the phosphorus concentration in wheat leaves might be higher than the potassium concentration. The SVR had higher performances in cross-validation than in validation process. Xiong et al. (2020) attributed the low prediction performances of SVR model to the nonlinear principle of SVR that increases the complexity of the prediction model. On the other hand, random forest outperformed PLSR models in a study conducted by Nawar and Mouazen, (2019) to estimate soil total nitrogen using Vis-NIR spectroscopy suggesting that RF handles better the nonlinear relationships that exist between the response variable and predictor variables in soils. Based on Vis-NIR reflectance, SVRs outperformed PLSRs in a different comparison, with R^2 values of 0.722 and 0.704 for P and K, respectively, which has been explained by the ability of SVR to model the non-linearity between the reflectance data and plant biochemical variables (Zhai et al. 2013).

Table 1. Cross-validation and validation performances.

Nutrient	model	Cross- validation			Validation		
		R^2	NRMSE	MAE	R^2	NRMSE	MAE
LPC	SVR	0.84	0.156	0.739	0.503	0.203	0.897
	PCR	0.85	0.14	0.661	0.534	0.202	0.875
	RF	0.79	0.18	0.91	**0.56**	**0.192**	**0.713**
LKC	SVR	0.637	0.207	3.643	0.43	0.133	2.139
	PCR	0.591	0.243	3.715	0.359	0.196	3.41
	RF	0.542	0.241	4.29	**0.531**	**0.114**	**1.856**

Wavelength importance

Fig.1 shows the importance of each wavelength in explaining the variation in our phosphorus and potassium data. The most informative wavelengths have high importance values. The wavelength importance has overlapping peaks for SVR and PCR models for both phosphorus and potassium. For phosphorus, the SVR and PCR developed models have the most important wavelengths in the spectral ranges from 710 to 730 nm and in the SWIR region from 2022 to 2036 nm, whereas the most informative wavelengths are 716, 738, and 727 nm for random forest (Fig 1A). The most effective wavelengths for determining potassium levels using SVR and PCR were primarily in the blue part of the spectrum, specifically between 494-500 nm, 470-480 nm, and 2040-2060 nm. For the random forest model, the most informative wavelengths were at 1472, 1480, and 1485 nm. Similarly, Siedliska et al. (2021) found that

certain wavelengths in the red (715 and 723 nm) and short-wave infrared (2301 and 2332 nm) regions were important for identifying different levels of P treatment in strawberry plants. The absorption peaks at the red and far-red regions of the electromagnetic spectrum were primarily attributed to the chlorophyll a absorption. Pimstein et al. (2011) observed significant correlation in the spectral ranges 1400–1500 nm and 1900–2100 nm for phosphorus and suggested a spectral index based on wavelengths 1645 nm and 1715 nm for both potassium and phosphorus estimation. Malmir et al. (2020) reported important wavelengths for P prediction at 700–1000-nm region and higher β-coefficient values in the wavelength range 730–1000 nm for K prediction. Xiong et al. (2020) found several sensitive wavelengths in the region from 550 nm to 970 nm for potassium prediction in green leaves. The authors stated that the third overtone of the vibration of the C-H and O-H bonds is generally linked to the wavelength range of 730–900 nm.

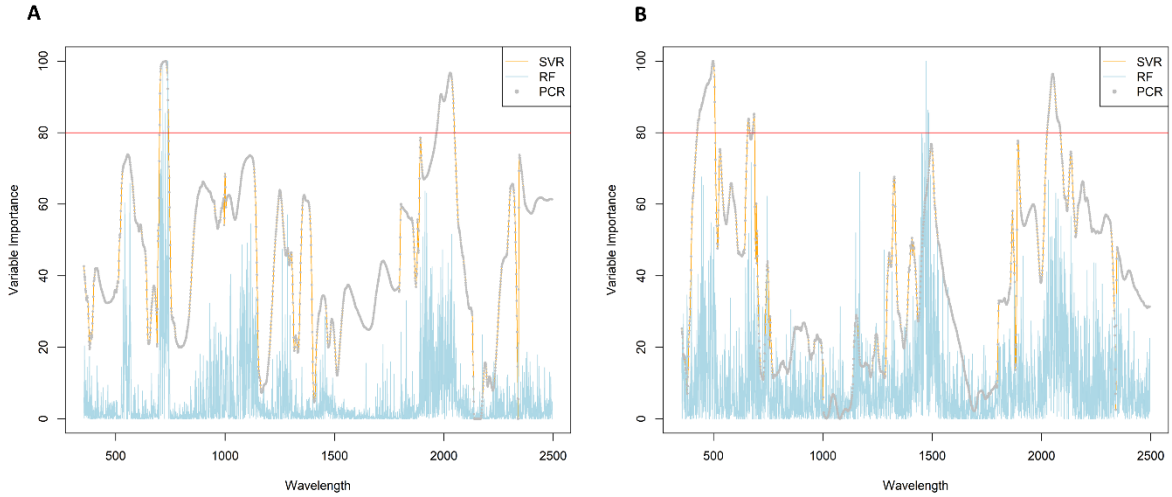

Fig. 1. Important wavelengths for the prediction of LPC (A) and LKC (B) content using support vector regression (SVR), random forest (RF), and principal component regression (PCR).

In our study, phosphorus models fared better than potassium and random forest analysis had the best predictive performances in validation for both phosphorus and potassium. In addition, a few variables gave very high variable importance values in RF models compared to SVR and PCR models, this suggests that our proposed RF models mainly focus on the important variables for modeling while neglecting the influence of noisy variables.

REFERENCES

Ge, Y., Atefi, A., Zhang, H., Miao, C., Ramamurthy, R.K., 2019. High - throughput analysis of leaf physiological and chemical traits with VIS – NIR – SWIR spectroscopy : a case study with a maize diversity panel. Plant Methods 1–12. https://doi.org/10.1186/s13007-019-0450-8

Jiaying, M., Tingting, C., Jie, L., Weimeng, F., Baohua, F., Guangyan, L., Hubo, L., Juncai, L., Zhihai, W., Longxing, T., Guanfu, F., 2022. Functions of Nitrogen, Phosphorus and Potassium in Energy Status and Their Influences on Rice Growth and Development. Rice Sci. 29, 166–178. https://doi.org/10.1016/J.RSCI.2022.01.005

Malmir, M., Tahmasbian, I., Xu, Z., Farrar, M.B., Bai, S.H., 2020. Prediction of macronutrients in plant leaves using chemometric analysis and wavelength selection. J. Soils Sediments 20, 249–259. https://doi.org/10.1007/s11368-019-02418-z

Mishra, P., Herrmann, I., Angileri, M., 2021. Improved prediction of potassium and nitrogen

in dried bell pepper leaves with visible and near-infrared spectroscopy utilising wavelength selection techniques. Talanta 225, 121971. https://doi.org/10.1016/J.TALANTA.2020.121971

Muehling, K.H., Goldbach, H., Olfs, H.-W., Husted, S., Van Maarschalkerweerd, M., 2015. Recent developments in fast spectroscopy for plant mineral analysis. https://doi.org/10.3389/fpls.2015.00169

Nawar, S., Mouazen, A.M., 2019. On-line vis-NIR spectroscopy prediction of soil organic carbon using machine learning. Soil Tillage Res. 190, 120–127. https://doi.org/10.1016/j.still.2019.03.006

Pimstein, A., Karnieli, A., Bansal, S.K., Bonfil, D.J., 2011. Exploring remotely sensed technologies for monitoring wheat potassium and phosphorus using field spectroscopy. F. Crop. Res. 121, 125–135. https://doi.org/10.1016/J.FCR.2010.12.001

Siedliska, A., Baranowski, P., Pastuszka-Woźniak, J., Zubik, M., Krzyszczak, J., 2021. Identification of plant leaf phosphorus content at different growth stages based on hyperspectral reflectance. BMC Plant Biol. 21. https://doi.org/10.1186/s12870-020-02807-4

Thornburg, T.E., Liu, J., Li, Q., Xue, H., Wang, G., Li, L., Fontana, J.E., Davis, K.E., Liu, W., Zhang, B., Zhang, Z., Liu, M., Pan, X., 2020. Potassium Deficiency Significantly Affected Plant Growth and Development as Well as microRNA-Mediated Mechanism in Wheat (Triticum aestivum L.). Front. Plant Sci. 11, 1219. https://doi.org/10.3389/FPLS.2020.01219/BIBTEX

Vasques, G.M., Grunwald, S., Sickman, J.O., 2008. Comparison of multivariate methods for inferential modeling of soil carbon using visible/near-infrared spectra. Geoderma 146, 14–25. https://doi.org/10.1016/J.GEODERMA.2008.04.007

Xiong, Y., Ohashi, S., Nakano, K., Jiang, W., Takizawa, K., Iijima, K., Maniwara, P., 2020. Quantification of potassium concentration with Vis-SWNIR spectroscopy in fresh lettuce. J. Innov. Opt. Health Sci. 13. https://doi.org/10.1142/S1793545820500297/ASSET/IMAGES/LARGE/S1793545820500297FIGF5.JPEG

Zhai, Y., Cui, L., Zhou, X., Gao, Y., Fei, T., Gao, W., 2013. Estimation of nitrogen, phosphorus, and potassium contents in the leaves of different plants using laboratory-based visible and nearinfrared reflectance spectroscopy: comparison of partial least-square regression and support vector machine regression methods. Int. J. Remote Sens. 34, 2502–2518. https://doi.org/10.1080/01431161.2012.746484