

RAPID PREDICTION OF AVAILABLE K CONTENT IN SOIL USING NEAR-INFRARED SPECTROSCOPY

M. TODOROVA^{1,*} S. ATANASSOVA¹ M. GERGIKOVA¹ R. ILIEVA²

¹ Agricultural Faculty, Trakia University, Stara Zagora 6000, Bulgaria

² Institute of Soil Science, Sofia, Bulgaria

*e-mail: mtodorova@uni-sz.bg

Abstract: Soil NPK testing has been widely used for fertilizer recommendation of annual crops. Soil analyses of available K content by chemical methods are sufficiently accurate, but they are expensive, time consuming and labour intensive. Recently, fertilization recommendation to the farmers is based on reduced number of soil samples due to high price of analyses. For this reason a rapid and cost-effective soil analysis is needed for soil quality assessment. Near infrared spectroscopy (NIRS) could provide a possible alternative. The objective of this study was to investigate the possibilities of NIRS for prediction of available potassium content in different soil units. A total 191 samples from four soil type-Chernozems, Vertisols, Luvisols and Fluvisols were analyzed for available K by conventional chemical method. NIR spectra of all samples were obtained by using an InfraAlyzer 450 spectrophotometer within the range 1445-2348 nm and portable FQA-NIRGun scanning spectrophotometer in shortwave NIR range from 600 to 1100 nm. SIMCA- soft independent modeling of class analogy was performed to classify samples, according to soil type. MLR and PLS regression were used for calibration models development for available K determination. The best model was obtained for samples of Chernozems with correlation coefficient $R=0,94$, standard error of calibration $SEC=2,30$ mg/100g, and the ratio of the standard variation of the reference data to the SEC, indicating the performance of the calibration $RPD=3,1$. Accuracy of determination of K content for models for separated soil units, developed by spectral data in short-wave NIR region, was better to accuracy of models, based on filter-type instrument in classical NIR region. Correlation coefficients for the global calibrations containing the samples of all soil units' decrease and SEC or SEP increase compared to calibrations for separate soil units. The values of R were between 0,76 and 0,79 and RPD were between 1,5-1,7 for the both calibration and test set. Calibration models for each soil type increase accuracy of determination of available potassium content.

Key Words: Near infrared spectroscopy, Available potassium, Quantitative determination

1. INTRODUCTION

Soil nitrogen, phosphorus, and potassium (NPK) testing has been widely used for fertilizer recommendation of annual crops. Soil analyses of available potassium content by chemical methods are sufficiently accurate, but they are expensive, time consuming and labour intensive. Recently, fertilization recommendation to the farmers is based on reduced number of soil samples due to high price of analyses. For this reason a rapid and cost-effective soil analysis is needed for soil quality assessment. For more than 20 years, near infrared spectroscopy (NIRS) has been proved and utilized as a non-destructive, rapid low cost and environmentally-friendly quality analysis technique for many agricultural products. NIRS could provide a possible alternative for soil analysis. Additional advantage of NIRS was its field portability, making it most amenable to qualitative or directional decision making in on-the-go site-specific management practices.

Recently, successful application of near infrared spectroscopy for measuring soil properties, such as organic carbon, total nitrogen and clay content have been reported (Viscarra Rossel et al, 2006). Reported results for determination of total and available potassium in the soil by NIRS are comparatively small and controversial.

Successful determination of available K by NIR spectroscopy reported Daniel et al,(2003). The authors reported coefficient of determination $R^2 = 0.80$ between soil available potassium content and predicted by NIRS method. Li et al, (2007) used partial least squares PLS and artificial neural network

ANN techniques as calibration methods to relate NIR spectral data to the concentrations of available K. Reported coefficient of determination, based on calibration of PLS was 0.73 and the mean relative errors of PLS model was 7.40 %. Results of model based on ANN was $R^2 = 0.95$, and mean relative errors 7.87%, respectively.

Contrary results reported Yong et al, (2005) and Song and Yong, (2005) - an weak correlation ($R = 0,68-0,69$) between predicted soil available K and actual soil K, with high SEP - 25,05 mg.kg⁻¹.

The objective of this study was to investigate the possibilities of NIRS for prediction of available potassium content in different soil units.

2. MATERIAL AND METHODS

2.1. Soil Samples

A total of 191 samples from the surface and subsurface horizon were collected from different part of Bulgaria during April - October, 2007. Soil types were Calcic Chernozems and Haplic Chernozems from North Bulgaria, Calcic Vertisols and Eutric Vertisols, Chromic Luvisols and Calcic Fluvisols (FAO, 1998) from South Bulgaria. These soil units are representative for soils in Bulgaria and are the most widespread soil units used in agriculture sector. The samples were analyzed for available potassium (K_2O) content by the Egner-Riem method.

2.2. Spectral Analysis

The spectral data of all air-dry samples were measured using two spectrometers: InfraAlyzer 450 - filter-type spectrophotometer within the range of 1445-2348 nm and portable scanning NIR instrument

FQA-NIRGun, (FANTEC, Japan) within a spectral range of 600-1100 nm. Two or three measurements were carried out for each soil sample using independent sampler cell fillings and then averaged to give one spectrum per sample. The absorbance was recorded as a log 1/R, where R is diffuse reflectance.

SIMCA- soft independent modeling of class analogy was performed to classify samples, according to soil types. SIMCA develops models for each class based on principal components analysis (PCA). PCA transforms the original data to a new coordinate system such that the greatest variance by any projection of the data comes to lie on the first coordinate (called the first principal component), the second greatest variance on the second coordinate, and so on. Once each class has its own model, new samples could be classified to one or another classes according to their spectra.

Multiple linear regression (MLR) was used for develop calibration models for available K content based on NIR spectra of samples obtained via InfraAlyzer 450 and partial least-square regression (PLS) for spectral data from FQA-NIRGun, respectively. Separate calibration equations were obtained for determination of available potassium (K_2O) for different sets of soil units. Additionally, global calibration equation was developed using two-third of all samples as calibration set and remaining samples as independent test set.

PLS regression and SIMCA modeling were carried out by Unscrambler 9.7 (CAMO Software AS, Norway).

3. RESULTS AND DISCUSSION

3.1. Quantitative Determination of K Content

Descriptive statistics and distribution of soil available potassium content for each soil type are shown in Figure 1. The range of K content varied from 6 to 58 mg/100g, depending on soil units. The range was the widest for Fluvisols and narrow for Luvisols samples. The average value of available potassium content was higher for Luvisols samples. Distribution of the samples according to available potassium content was not even. The samples with high available potassium content were a few for Vertisols, Chernozems and Fluvisols. There was relatively even distribution only for Luvisols.

The results of the quantitative determination of K_2O , mg/100 g content in different soil units by MLR and PLS are presented in Table 1. The obtained correlation coefficients between the measured and predicted values for available potassium were between 0,74 and 0,96. Models, based on InfraAlyzer 450 spectral data for Chernozems, Luvisols and Fluvisols showed R between 0,85 and 0,93 and RPD between 2,4 and 2,6 and could be classified as good models according to Rossel (2007). Models for same soil units, based on short-wave NIR region and PLS regression with R more than 0,94 and RPD were more than 3,0 could be classified as excellent. The statistical data of Vertisols' calibration equations in the both spectral regions showed- low correlation coefficient - 0,74 and 0,76, and RPD values - 2,2 and 2,4 could be classified as a good models.

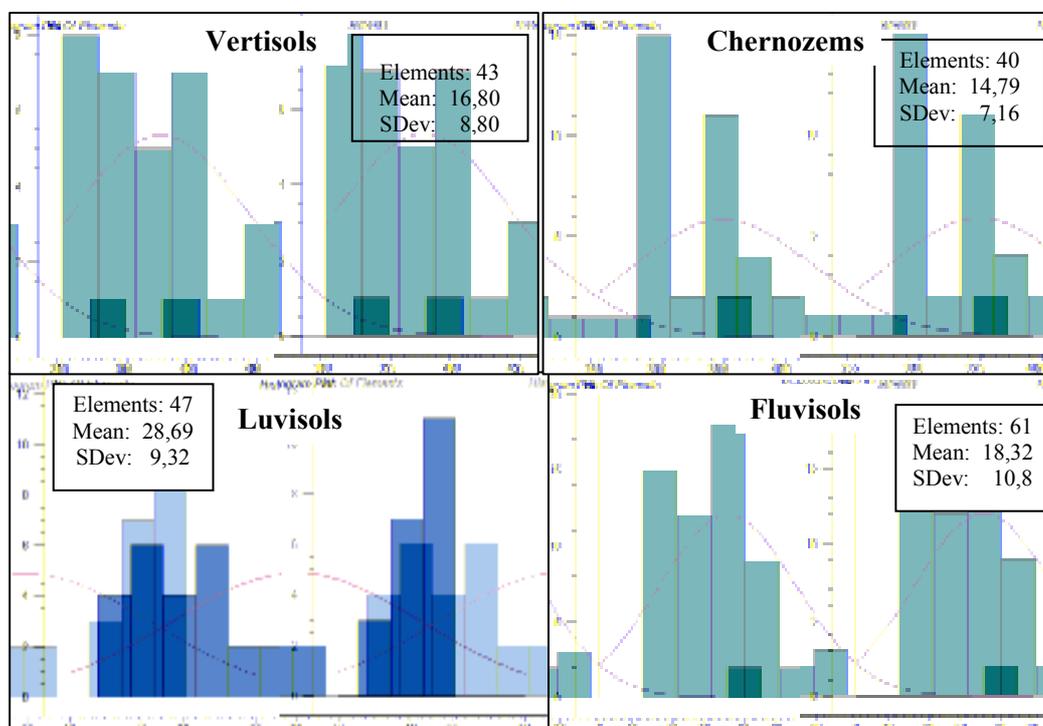


Figure 1. Histogram, mean and standard deviation (SD) of K_2O , mg/100 g content, in different soil units and data set

Table 1. Statistical parameters of the calibration equations for NIRS prediction of K₂O, mg/100 g in different soil units

Soil units	InfraAlyzer 450 MLR			FQA-NIRGun PLS		
	SEC	R	RPD	SEC	R	RPD
Vertisols	4,07	0,74	2,2	3,7	0,76	2,4
Chernozems	2,76	0,85	2,6	2,3	0,94	3,1
Luvisols	3,66	0,93	2,5	3,1	0,96	3,0
Fluvisols	4,41	0,89	2,4	3,4	0,94	3,2

SEC – standard error of calibration, SEP – standard error of prediction,
R – coefficient of multiple correlation, RPD – ratio between SD and SEC

Accuracy of determination of available potassium in all soil units for models, developed by spectral data obtained using FQA-NIRGun instrument in short-wave NIR region, was better compared to an accuracy of models, based on InfraAlyzer 450 spectral data. The reason for obtained better accuracy of determination using short-wave NIR region was probably much more spectral information, available and used for developing of calibration models by PLS, in comparison with spectral information at 6 or 7 wavelengths from total 19, included in models based on InfraAlyzer 450 data. Figure 2 and 3 graphically illustrate the relationships between determined and NIR spectroscopy predicted values of available potassium in Chernozems and Luvisols samples. Correlation coefficients for the global calibrations for all soil units' decrease and SEC or SEP increase compared to calibrations for separate soil units. (Table 2) The values of R were between 0,76 and 0,79 and RPD were between 1,5-1,7 for the both calibration and test set. Separate calibration models for each soil type increase accuracy of determination of available potassium content.

The most important spectral information for determination of available potassium content in soil samples in short-wave NIR region were found to be at 748 nm, in the region 791-800 nm, at 977, 987, from 1030 to 1040 nm, and 1070 nm. The most often used spectral data in calibration equation, based on InfraAlyzer 450 data were at 1680, 1940, 2230, 2270, 2310, 2336 and 2348 nm.

Soil potassium exists in solution, exchangeable and nonexchangeable forms that are in dynamic

equilibrium with each other. Exchangeable K⁺ held on the negatively charged sites clay minerals in the soil. The most prevalent clay minerals are the layered aluminosilicates. Their crystals are composed of two basic structural units, namely: a tetrahedron of oxygen atoms surrounding a central cation, usually Si⁴⁺, and an octahedron of oxygen atoms or hydroxyl groups surrounding a larger cation usually Al³⁺ or Mg²⁺. The tetrahedral are joined at their basal corners and the octahedral are joined along their edges by means of shared oxygen atoms.

In regression equations for determination of available potassium content in soil predominated spectral information, connected with absorption of O-H and metals-OH groups. Absorption at 748 nm, 977 nm, 987 nm and from 1030 to 1040 nm is corresponding to overtones of O-H vibrations. The NIR absorption at 1940 nm due to the combinations of the H-O-H bend with the O-H stretches vibrations is typical for a water-rich mineral, like montmorillonite-AL₂O₃2SiO₂nH₂O, MgOAl₂O₃4SiO₂H₂O. (Clark et al, 1990; Popova,1990). According to Gaydon et al, (2009), absorption in the region from 2170 to 2230 nm is associated with Al-OH group, and at 2240 nm possibly indicated the presence of Si-OH. The NIR absorption at 2270 nm and 2310 nm could be connected with the combination Al-OH bend plus O-H stretch vibrations. Metals-OH bend plus O-H stretch combination near 2200 nm and 2300 nm are diagnostic absorption features in clay mineral identification (Clark et al, 1990).

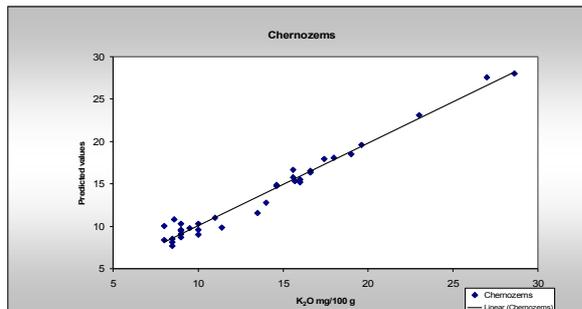


Figure 2. Relation between actual and NIRS predicted values of available potassium in Chernozems samples

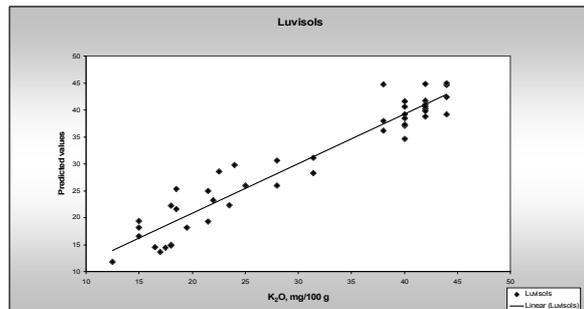


Figure 3. Relation between actual and NIRS predicted values of available potassium in Luvisols samples

Table 2. Statistical data of the calibration equations and validation statistics for NIRS prediction of K₂O mg/100 g in all examined samples

	InfraAlyzer 450			FQA-NIRGun		
	MLR		RPD	PLS		RPD
All soil samples	SEC/SEP	R/r		SEC/SEP	R/r	
Calibration set	5.9	0.79	1.7	6.3	0.78	1.6
Test set	6.5	0.76	1.5	6.2	0.78	1.6

SEC – standard error of calibration, SEP – standard error of prediction, R – coefficient of multiple correlation, r – correlation coefficient between actual and NIRS predicted values, RPD – ratio between SD and SEC or SEP

The spectral information in calibration equations for determination of available K confirmed that determination of potassium in soil was indirect and based on relationship between potassium and clay minerals in soils.

3.2. Classification of Soil Samples According To Soil Units

Four classes were formed according to soil units as follow: Chernozems class, Fluvisols class, Luvisols class and Vertisols class. Spectral data obtained by analysis of tested soil samples by FQA-NIRGun were used for classification. Spectra of ten samples from each class were used as validation data set. The remaining samples were used as calibration data set for development of SIMCA models. Obtained SIMCA models with included 10 principal components correct classified all samples from calibration data set. Interclass distance between soil classes varied between 3.08 and 25.15. Greater the distance between 2 classes the greater is the difference in composition of samples belonging to those classes. As a rule of thumb, a distance of over 3 indicates that the samples are well separated. The highest interclass distance was found between class Fluvisols and class Chernozems – 25.15 and between class Fluvisols and class Luvisols – 13.24, respectively. Obtained results showed that specific differences exist between spectral data of soils from different soil units in short-wave NIR region.

Obtained models were tested using spectra of samples from validation set and results were presented in table 3. At the 5% significance level the most of the samples from validation set were recognized by their type correct class. Only two Vertisols samples were not recognized as belong to any soil class.

The discriminating power is a measure of variable importance, in this case spectral data, which contribute to the development of classification models. A value close to 0 indicated low discrimination ability in a variable, while a value much larger than 1 implied high discrimination power. The most significant spectral data for discrimination between investigated

classes by SIMCA models were found to be at wavelengths 600, 626, 706, 752, 787, 961, 1024 and 1080 nm. Probably the differences in the visible region around at 600, 626 and 706 nm were caused from colour differences in soil units. Absorption at 753, 961 and 1080 nm is corresponding to overtones of O-H vibrations. Another important wavelength in discriminating power plot at 787 could be associated with absorption of N-H stretch vibration, third overtone and at 1024 nm with N-H primary aromatic amine bands, respectively. These finding showed that discrimination between classes was based on differences in soil composition, mainly clay and organic matter content.

Results from classification procedure showed possibilities of near-infrared spectroscopy in combination with multivariate chemometrics technique such as SIMCA for fast classification of soil samples according to their types. That will allow choosing the most suitable calibration equation and will increase an accuracy of determination of available potassium content in soil.

4. CONCLUSION

The results in the present study demonstrate potential of near infrared spectroscopy for developing a fast and cost-effective analysis of available potassium soil content. Accuracy of available potassium content determination depends on soil types, spectral region and available spectral information. Dividing the samples into groups according to their soil types improved accuracy of NIR prediction of available K content. Accuracy of determination of available K in all soil units for models, developed by spectral data in short-wave NIR region, was better compared to accuracy of models, based on filter-type instrument in classical NIR region. Near infrared spectroscopy in combination with multivariate chemometrics technique offers an alternative approach to traditional methods with large potentials for a rapid and reliable soil analysis.

Table 3. Results for SIMCA models, validation set.

Soil units	Chemozems class	Fluvisol class	Luvisol class	Vertisol class	No match
Chernozems	10	0	0	0	0
Fluvisols	0	10	0	0	0
Luvisols	0	0	10	0	0
Vertisols	0	0	0	8	2

5. ACKNOWLEDGEMENT

The authors gratefully acknowledge the financial support of National Science Fund - project USS-401/2006.

6. REFERENCES

- Clark, R.N., King, T.V.V., Klejwa, M., Swayze, G., Vergo, N., 1990. High spectral resolution reflectance spectroscopy of minerals. *Journal of Geophysical Research*, 95: 12653-12680.
- Daniel, K., Tripathi, N., Honda, K., 2003. Artificial neural network analysis of laboratory and in situ spectra for the estimation of macronutrients in soils of Lop Buri (Thailand). *Australian Journal of Soil Research*, 41: 47-59.
- Gaydon, J.W., Glass, H.J., Pascoe, R., 2009. Method for near infrared sensor-based sorting of a copper ore. *Journal of Near Infrared Spectroscopy*, 17: 177-194.
- Li, W., Zhang, S., Zhang, Q., Dong, C., Zhang, S., 2007. Rapid prediction of available N, P and K content in soil using near-infrared reflectance spectroscopy. *Transactions of the Chinese Society of Agricultural Engineering*, 23: 55-59.
- Popova, R., 1990. Comparative chemical-mineralogical characteristics of the heavy clay soils from the Soviet Union and Bulgaria. Ph.D Thesis, Moskow State University, Moskva.
- Song, H., He, A.Y., 2005. A new approach to detect soil nutrient content based on NIR Spectroscopy technique. *Conf. Proc. IEEE-EMBS. Engineering in Medicine and Biology Society*, pp: 3149-3152.
- Viscarra Rossel, R.A., Walvoort, D.J.J., McBratney, A.B., Janik, L.J., Skjemstad, J.O., 2006. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma*, 131(1/2): 59-75.
- Viscarra Rossel, R.A., 2007. Robust modelling of soil diffuse reflectance spectra by "bagging-partial least squares regression". *J. Near Infrared Spectrosc.*, 15: 39-47.
- World Reference Base for Soil Resources, 1998. FAO.
- Yong, He., Song, H., Garsia Pereira, A., Hernandez Gomez, A., 2005. A new approach to predict N, P, K and OM content in a loamy mixed soil by using near infrared reflectance spectroscopy. *ICIC, part I, LNCS 3644*, 859-867.