Mid Infrared spectroscopy and PLS to predict soil contaminant under different soil conditions

Ahmed A. Afifi and Refat A. Youssef

Soils and Water Use Dept., National Research Centre, El-Behouse st., Dokki, Giza, Egypt
a.afifinrc@gmail.com

Abstract: Soil contamination by naturally occurring and anthropogenic organic and inorganic chemicals is a serious human health and environmental problem in many industrialized and non-industrialzed nations. This process has dramatically increased in their extent and intensity over the last decades. Progressively, actions have been taken in order to evaluate and reduce the major threats that have already done devastation on soil conditions. The objective of the present study is to examine critically the suitability of Mid infrared reflectance spectroscopy (MIRS) as a tool for soil contamination assessment. A quicker method is developed based on a multivariate calibration procedure using partial least squares (PLS) regression to establish a relationship between reflectance spectra in the mid infrared (MIR) region and spectral of soil characteristics, that are inter-correlated with concentration levels of Cd and Pb. Several spectral pre-processing methods (normalization, multiplicative scatter correction (MSC), derivation, standard normal variant (SNV) transforms) were employed to improve the robustness and performance of the calibration models. The principal component analysis (PCA) was performed prior to MIR-PLS regression analysis that identified spectral outliers in the absorbance spectra of soil samples. Pearson correlation identified two elements (Cd, Pb). The obtained calibration models showed high regression (r) for Pb, and relatively high (r) for Cd ($r > 0.95$), while for Cd ($r>0.98$). Based on this result support the conclusion that mid-infrared spectroscopy could aid conventional method analyses of soils heavily contaminated with certain heavy metals after a robust model is developed.

[Ahmed A. Afifi and Refat A. Youssef. **Mid Infrared spectroscopy and PLS to predict soil contaminant under different soil conditions.** *Life Sci J* 2014;11(12):647-654]. (ISSN:1097-8135). http://www.lifesciencesite.com. 124

Keywords: MIR, PLS, Soil contaminant, soil condition, chemometric

1. Introduction:

NIRS utilizes wavelengths between 750 and 2500 nm, but this range is often extended to 400-2500 nm. Near infrared radiation is absorbed by different chemical bonds such as O-H, C-H, N-H, S-H and C=O. Absorption of NIR radiation results in bending, stretching, twisting and scissoring of the bonds (**WORKMAN, 1993 and WORKMAN, 2000)**. The amount of NIR radiation that is absorbed is determined by the nature and number of bonds present in the analyzed material. Hence, NIR spectra contain detailed information on the chemical composition of that material (**FOLEY, et al., 1998**).

The NIR spectra do not contain sharp and distinct peaks because they consist of overtones and combinations from primary absorption in the mid-infrared region. These overtones are anharmonic and impede interpretation of NIR spectra (**WORKMAN, 1993)**. Since a direct interpretation of NIR spectra of complex mixtures is extremely difficult, the application of NIRS for analysis of environmental materials requires a calibration procedure using sophisticated statistical techniques (**FOLEY, et al., 1998**).

Sample Preparation Requirements

The MIR spectra depend not only on chemical characteristics of the analyzed material but also on its moisture (due to strong absorption of water molecules at 1450 nm and 1930 nm) and particle size (**Foley, et al., 1998 and Casler and Shenk, 1985**). Therefore, in order to ensure reliable MIR measurements samples need to be dried carefully and ground to a consistent particle size. The latter can be achieved by using laboratory grinders with the same grinding performance.

Calibration and Validation Procedures

The most essential step in the calibration procedure is the selection of a proper sample set. The calibration sample set should cover the entire range of spectral variation in the whole population for which the calibration is being carried out. Spectra that differ significantly from the average spectrum should be rejected from the calibration set as outliers. To identify outlier samples to be rejected, the entire population is ranked in terms of Mahalanobis distance from the average spectrum. There are numerous methods of selecting samples to be rejected, the CENTER algorithm (**Shenk and Westerhaus, 1991**) being the most popular. Usually it is much easier to achieve large spectral data sets, whereas obtaining the reference data may be more time consuming. In order to minimize the size of calibration sample set, and thus the amount of necessary reference analyses, the samples in which spectra are very similar may be rejected. Removal of spectrally similar samples is based on the assumption that only one sample is

required to represent all samples in its neighbourhood (**Shenk and Westerhaus, 1991**). A select algorithm, based on the matrix of Mahalanobis distances between all pairs of spectra, can be used to identify neighbouring samples (**Shenk and Westerhaus, 1991**).

The samples selected for calibration should cover not only the entire spectral variation of the analyzed population but also the entire variability of the components or characteristics for which the calibration is carried out. Therefore, sometimes it is necessary to expand the calibration set by the samples not included in the calibration set chosen according to the center and select algorithms. calibration procedure relies on developing a regression equation between the absorbance spectra and the components or characteristics of interest. The most commonly used regression procedures include multiple linear regression (MLR), principal component regression (PCR), partial least square regression (PLS) and modified PLS regression (mPLS). The two latter methods are considered more powerful since – unlike the MLR – they use the entire spectral information. The other available approaches involve neural networks and wavelet theory (**Foley, et al., 1998**). Prior to the calibration the spectra should be corrected for a scatter by any available methods (e.g. Detrend and Standard Normal Variate or Multiplicative Scatter Correction). For development of calibration equations several mathematical treatments of spectra are usually used.

These include taking derivatives of 1st to 3rd order, defining the segment length over which the derivative is to be calculated and smoothing the spectra. Since there is no single "best" treatment for all variables and all sample types, usually the only way to optimize the mathematical treatment is to follow a trial and error procedure (**Couteaux, et al., 1998; Ludwig and Khanna, 2001 and Couteaux, et al., 2003**). In order to avoid overfitting, when using PCR, PLS or mPLS methods for model development, a procedure called crossvalidation should be used. The cross-validation approach enables us to determine an optimal number of terms or principal components to be included in the model. The calibration sample set is divided into several groups and a prediction is made for one group based on calibration equation developed from the remaining groups. This procedure is repeated until all groups are used for validation once. The residuals of the predictions are then pooled to calculate the standard error of cross-validation (SECV). The best model should have the smallest SECV. Although cross validation may help in selecting the number of PC or PLS components it should not be used blindly. Further statistical test should be applied to assess performance and relevance of the developed models and to detect irregularities in the data (eg. outliers) that may harm developed regressions. Detailed descriptions of calibration procedures are given in **Nas et al. (2002) and Wold et al. (2001)**.

Quality of the developed calibration equations is assessed in the validation stage. The validation sample set includes the samples for which the reference data were measured using classical methods. During validation the NIRS predicted values are regressed against the reference values. The quality criteria include correlation coefficient (r2) and regression coefficient (a) of linear regression measured against NIRS predicted values and standard error of prediction (SEP). SEP is calculated according to the equation:

$$SEP = \left[\sum_{i=1..n} (Yi - Xi).^2 (n-1).^{-1}\right].^{0.5}$$

where n is the number of samples, yi is the mean value of a constituent in sample *i* derived by the reference method, and xi is the NIRS predicted mean value for the sample *i*. The other commonly used quality parameters include ratio of standard deviation (SD) of the laboratory results to SEP (RPD) and the ratio of standard error of calibration (SEC) to SD. The use of RPD and SEC to SD ratio enables us to compare the accuracy of the models for constituents that are measured in different units.

Proper validation is a prerequisite for using the developed calibration models in routine analysis. It is essential to use for validating entirely independent samples, otherwise predictive accuracy of the developed models may be overestimated (**Brown, et al., 2006**). Furthermore, validation should simulate the intended model application. For instance, in regional models, randomly selected, independent, hold-out sites should be used for validation. When the number of available samples is too restrictive to carry out calibration and independent validation, the results of cross-validation may be used to assess the quality of calibration equations (**Martens and Dardenne, 1998**).

In cross-validation the quality criteria include SECV, values of r2 (reference vs MIRS predicted values), bias (mean of the MIRS predicted value less the mean of reference values), and the SD to SECV ratio (referred to as RPD or RSC) or standard error of calibration (SEC) to SD ratio (**Couteaux, et al., 2003**).

The objective of this study is to investigate the possibilities of predicting soil contaminant levels under different soil conditions (like clay and calcarous) from high-resolution reflectance spectra based on laboratory measurements of soil samples and multivariate calibration using partial least squares

(PLS) regression. Therefore, several data pre-processing methods are tested to verify what is most appropriate to produce acceptable results.

2. Methods:

2.1. Bulk soil and particle-size fraction

Tested soils were collected from abuMatameir area in the northern part of Egypt, representing a clay soil. In addition to, a calcareaous soil which has been collected from north-western coast of egypt. The study area is located in the mediteranean arid to semi arid zone. The annual mean temperature in the region is 25 to 30 °C and the annual average precipitation ranges from 10 to 60 mm. Soil sample consisted of amixture of nine subsamples from the surface to a depth of 50 cm. Soil sample was air-dried at room temperature and thoroughly mixed before chemical analyses.

The method used for the density fractionation of naturally occurring organic-mineral aggregates was loosely based on the method described by Huang et al. (2008). Briefly, 100 g of soil sample was placed into a 500-mL flask and 500 mL of saturated NaI solution (1.7 g/cm3) was then added to isolate the plant residues. The plant residues suspended on the liquid surface of the NaI solution were removed after the sample had been grounded into a paste with a glass stick, subjected to thorough agitation and then centrifuged.

2.2. Sorption of cadmium onto soil aggregates

Sorption obtained using a batch equilibrium technique at 25 °C in 50 ml acid-cleaned polystyrene centrifuge tubes. Aliquots of the dry and sieved soil aggregates and the bulk soil were accurately weighed and mixed with 0.01 mol/L $Ca(NO_3)_2$ as the background solution. Cadmium and Lead solutions with concentrations of 0, 20, 50, and 80 mg/L and 0, 100, 200, 300 mg/L respectively were added and the tubes were shaken on a variable speed reciprocal shaker at 100 strokes/min for 24 h. After shaking, the sample solutions were centrifuged at 5000 rpm for 3 min and then filtered. The supernatant was then used for analysis. Each curve consisted of four concentration points, and each point, including the blank, was run in triplicate. Concentrations of cadmium and lead in the supernatant were analyzed using atomic absorption. Because there was little adsorption by the flasks and no biodegradation, cadmium and lead sorbed by the sorbents and partitioned in the soil–water system.

2.3. Spectral measurements

Soil reflectance characteristics were investigated in two soils type (clay and calcareous).

Soil samples were dried at 80◦C for 24 h to standardise the moisture level. In order to minimise anisotropic scattering of light by soil aggregates of variable size, soils were grounded with a mortar and passed through a 2mm sieve.

The absorbance spectra of the soil samples were obtained by means of a MIRS, The spectra were measured between 400 and 4000 cm-1, using a resolution of about 4cm-1 bandwidth. The sample for irradiation is prepared by mixing the powdered sample with spectral grade KBr in the ratio of 1:20, KBr pellets having 13 mm diameter and 1 mm thickness are prepared under vacuum condition by applying a pressure of 10 torr on the stainless steel dies. The spectra were corrected for background absorption by division of a reference spectrum of KBr, and the final spectrum is an average based on the number of scans. The FT-IR measurements were performed using JASCO Model FT-IR spectrometer available at Soils and Water Use Department, National Research Centre, Egypt. To obtain a high signal/noise ratio 100 scans were accomplished for each sample.

3. Results and disscusion

1. Analysis of soil characteristics

Vegetation:	Wheat. barley. pea, potato, cabbage and lettuce.
Parent material	Nile alluvial deposits
Topography	Nearly level
Drainage	Moderately drained
Classification	*Vertisols, Typic Torrerts*
Hori. Depth (cm)	**Description**
Ap 0-20	Vcry dark greyish brown (10 YR 3/2. moist); strong, medium, sub-angular, blocky structure; very sticky, very plastic; strong effervescence with HCl; many fine roots 20 em in depth. many fine roots; diffuse boundary.
AC 20-40	Dark greyish brown (10 YR 4/2, moist); clay; strong. medium. sub-angular blocky structure; very sticky, very plastic; strong effervescence with HC!; many fine roots; diffuse boundary.
Ccn 40-60	Dark grayish brown (10 YR 4/2, moist); clay; strong, coarse, sub-angular blocky structure; very sticky, very plastic; strong effervesence with HCl; few CaCO. concretions; slickensides; few fine roots, clear boundary.
Clor 60-80	Very darkrayish brown (10 YF 3/2, moist); clay; strr-ig, coarse, sub-angular blocky structure; very sticky, very plastic; strc.ig effervescencewith HCI; few CaC03 concretions. distinct slickensides; some grey and rust mottles; diffuse boundary.
G2ro 80-120	As the above layer except the mottles increase with depth.

Table 2. Soil physical properties of the soil.

Depth cm	Mechanical composition			Texture	$CaCO_3$ %
	Sand 2-0.02 nm	Silt 0.02-0.002 nm	Clay <0.002nm		
0-20	15.6	29.7	54.7	Clay	3.3
20-40	15.3	30.8	53.9	Clay	3.3
40-60	15.1	30.3	54.6	Clay	2.7
60-80	12.2	32.4	55.4	Clay	3.5

Table 3. Soil chemical properties

Depth cm	EC dS/cm	pH	Cations meq/100g soil				Anions meq/100g soil			CEC Meq/100g soil	OM %
			Ca^{++}	Mg^{++}	Na^{+}	K^{+}	Cl^{-}	SO_4^{--}	HCO_3^{-}		
0-20	1.03	8.1	0.261	0.078	0.405	0.004	0.349	0.064	0.334	58.8	1.22
20-40	0.86	8.0	0.223	0.047	0.388	0.002	0.233	0.010	0.418	50.1	1.11
40-60	0.73	8.3	0.169	0.077	0.264	0.002	0.176	0.032	0.304	55.0	1.45
60-80	0.99	8.0	0.226	0.116	0.360	0.002	0.275	0.036	0.339	52.0	1.41

Table 4. Soil heavy metals analysis of the surface layer

Depth (cm)	Concentration (ppm)								
	Ni	Cu	Cd	Pb	Fe	Zn	Mn	Co	Cr
0-20	0.534	0.056	0.324	1.432	0.224	0.014	0.089	0.149	0.082

1. Relationships between Wavelengths and Chemical Structures

1.1. Spectra of the pure component

Linking particular MIR wavelengths to well-defined compounds is an extremely difficult task. Due to the broadband nature of MIR spectra, consisting of overlapping peaks, the individual chemical structures are not well resolved. Numerous constituents of analyzed materials absorb within the entire MIR region and the spectral information is repeated through successive overtones and combinations.

The prepared soils and pure chemicals have been characterizing by MIR spectra to build up the model reference which will be used to identify the polluted soils, as seen the figures (1 and 2).

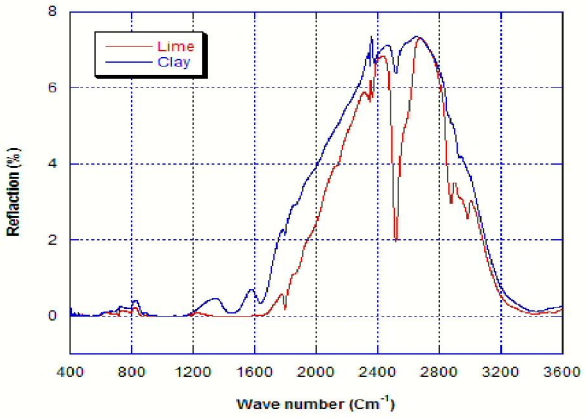


Figure (1) MIR of the pure soil samples

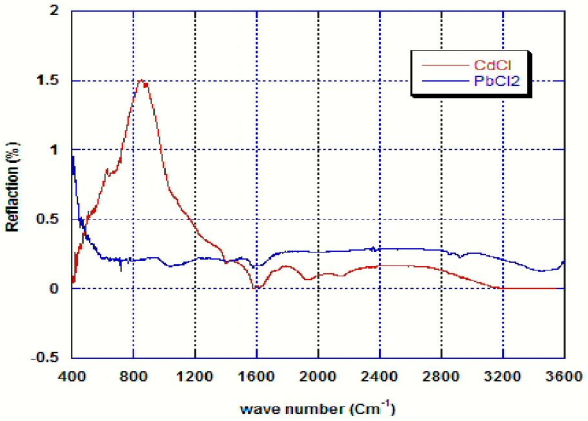


Figure (2) MIR of the pure chemicals

1.2. As seen in the prvious figures, it is cleary to see the deferences in spectra of all the pure compounds.

1.3. MIR spectral features

The spectra show several stretching and bending bands which were attributed to the mineral and organic components in soil. For example, the presence of hydroxyl (−OH) stretching and bending bands in the spectral region from 3750−3300 cm^{-1} and 950−820 cm^{-1} was possibly due to the presence of structural −OH groups in clay minerals and metal (Fe/Al) oxides (**Johnston and Aochi 1996**). Phyllosilicates can be distinguished into the −OH stretching (3750−3400 cm^{-1}), −OH bending (950−600 cm^{-1}), Si−O stretching (1200−700 cm^{-1}) and Si−O bending (600−400 cm^{-1}) vibrations (see Figure, 3)

(**Haberhauer et al. 1998**). The presence of a band at 3145 cm^{-1} in the spectra of samples tested in this study could possibly be attributed to −OH stretching in Fe oxides, as reported by **Haberhauer et al. (1998)** who also attributed the presence of −OH stretching band at 3140 cm^{-1} to Fe oxides. The spectral peaks observed in the region between 3695 and 3622 cm^{-1} could be attributed to the presence of −OH stretching band for kaolinite (**Janik et al. 2009**). The bands observed between 428 and 470 cm^{-1} in the spectra of soil samples were attributed to the Fe−O/Al−O bending vibrations (**Cornell and Schwertmann 2003; Ibrahim et al. 2008**).

The spectral bands in the region from 1510−1230 cm^{-1} and at 1730 cm^{-1} have been reported to be associated with the presence of aromatic and carboxylic (−COOH) functional groups (**Pirie et al. 2005**). The spectral band identified in the analysed soil samples at 1720 cm^{-1} in Figure, 3 could be associated with the presence of −COOH group in soil organic fraction. The bands in the spectral region from 2928−2852 cm^{-1} could be attributed to the aliphatic ($-CH_2$) stretching vibrations present in organic fraction in the studied soil samples. This is in agreement with those reported by **Janik et al., (2009)** and **Haberhauer et al. (1998)** for the presence of −CH2 functional group in spectra of soils (i.e. bands at 2930−2851 cm^{-1}).

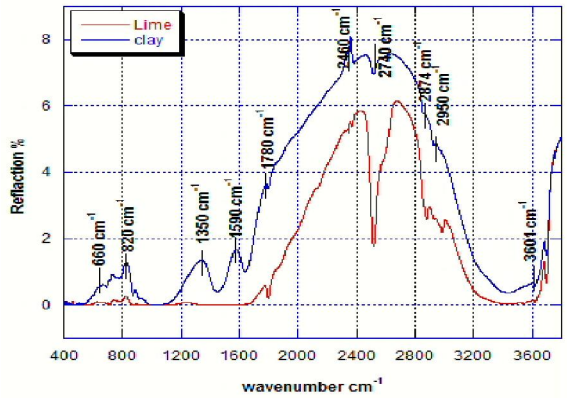


Figure (3) MIR spectral of the clay and calcaereous soils

In this study, the author used FTIR spectroscopy to investigate sorption complexes of pollutants on the surface of soils. Cd adsorption was identified due to the presence of spectral bands at 450 cm^{-1} and between 550 and 1200cm^{-1} due to the adsorption of Cd on the surface of soil. While the adsorption of Pb was identified in the spectral band between 400-850 cm^{-1} in the both types of the soils. However, the spectral bands for the Mn, Cu and Mo were 400-750, 450-1100, and 460-850 cm^{-1} respectively.

1.4. Building up the references of Cd and Pb on clayey soil

Three concentration in Cd and Pb were used on clay soil as (20, 50, 80 ppm) and (100, 200, 300 ppm) respectively. As shown in figure (4) the Cd absorbed on the clay will be recognized at the band ranges from 400 to 1600 cm^{-1}. As well as, it can be recognized that the effect of the concentration of Cd, where it increase the broading or sharpness in the spectra peaks.

Regarding the Pb effect of the peak on the clay soil, **the** the Pb absorbed on the clay will be recognized at the band ranges from 400 to 1500 cm^{-1}. As well as, the Pb concentration will affect the spectra at the range form 1550 to 3250 cm^{-1}. As seen in figure (5).

1.5. Building up the references of Cd and Pb on Calcareous soils

For the effect of the lime content on the absorption and Cd and Pb, it is shown, in figure (6) for Cd and figure (7) for Pb.

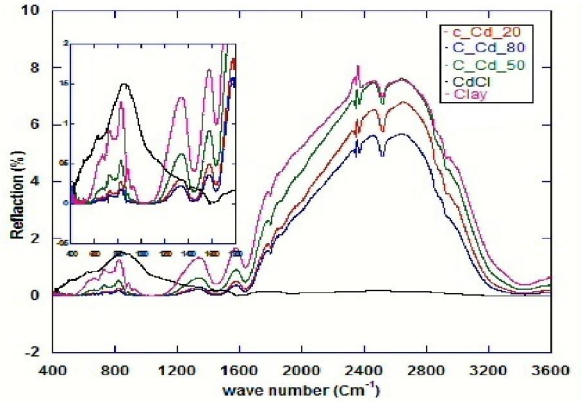


Figure (4) MIR spectra of clay soils under different cd concentration compared with the pure chemical spectra

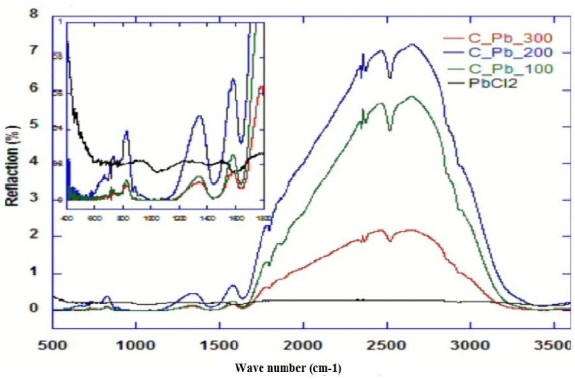


Figure (5) MIR spectra of clay soils under different Pb concentration compared with the pure chemical spectra

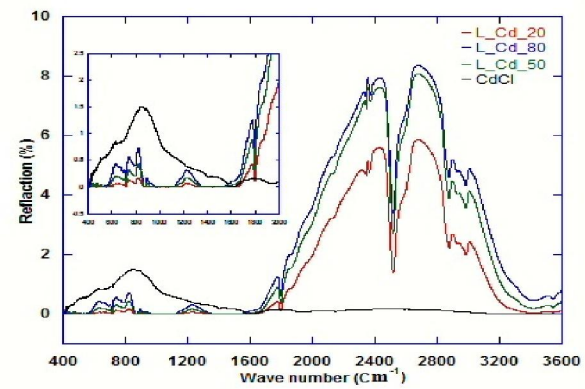


Figure (6) MIR spectra of calcareous soils under different cd concentration compared with the pure chemical spectra

The spectal range for Cd and Pb at the calcaerous soils between 400 and 1800 cm^{-1}. While, for the effect of the concentration of the heavy metals used in the study can be seen also in the range 1800 to 1600 cm^{-1}.

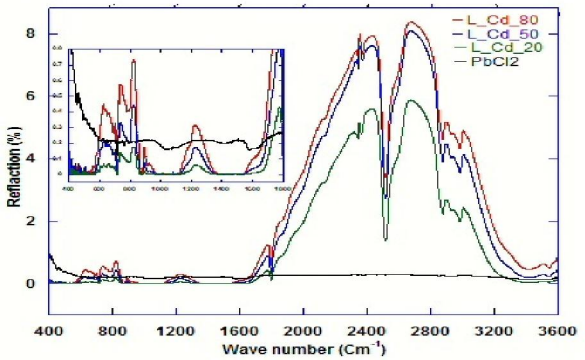


Figure (7) MIR spectra of calcareous soils under different Pb concentration compared with the pure chemical spectra

Chemometric Methods:-

When handling a large data set in NIRS/MIRS, we must not only look for the typical peak or the significant peak. But, it is better to perform multifactor analysis. We should realize that two spectra can be different just because of a shift in baseline. So, we add the possible methods for correcting change in baseline (e.g. multiplicative scatter correction, Standard Normal Variate, derivative of first and second order).

PCA was carried out to detect the presence of any spectral outliers inthe spectral data, prior to develop a prediction model using PLS regression. In PCA, the eigenspectra and their respective eigenvalues are calculated. The number of components in the data are then reduced to a smaller group of principal components, termed as PCA loadings, Figures (8 and 9).

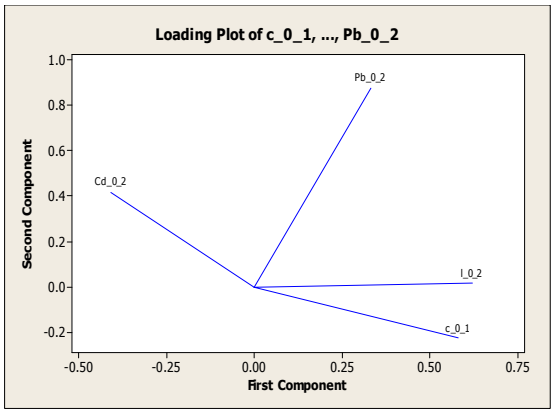


Figure (8) Loading plot of PC

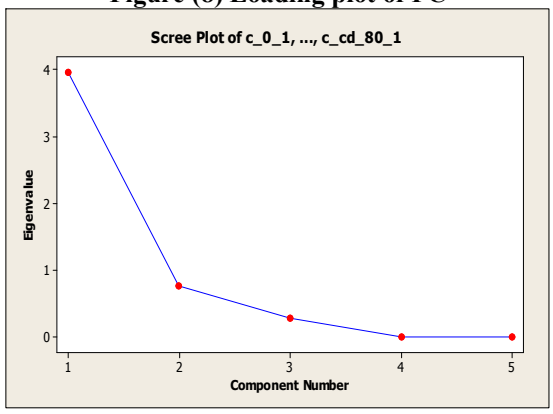


Figure (9) Eigen value of PC

PCA performed on the (first derivative pre-treated) reflactance spectra of all soil samples $n = 49$). The green filled circles are the 3 spectral outliers in the soil samples, Figure (10).

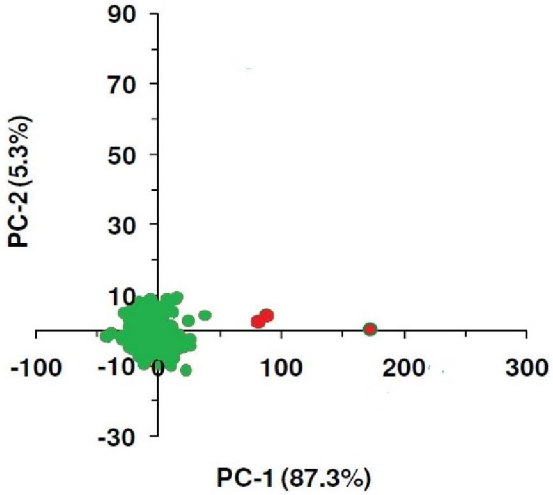


Figure (10) PCA

Previous studies indicate that PCA is a useful tool for the identification of spectral outliers in the absorbance spectra of the samples and can be employed to increase the quality of the prediction

model (**Pirie et al. 2005**). A PCA was performed on the (first derivative pretreated) absorbance spectra of all soil samples which enabled to detect 3 spectral outliers in the data (see Figure 10, red colour filled circles. The absorbance spectra of these (spectral outlier) samples were included in the preliminary calibration model (data not shown), however, excluded from the PLS analysis performed to develop a final calibration model.

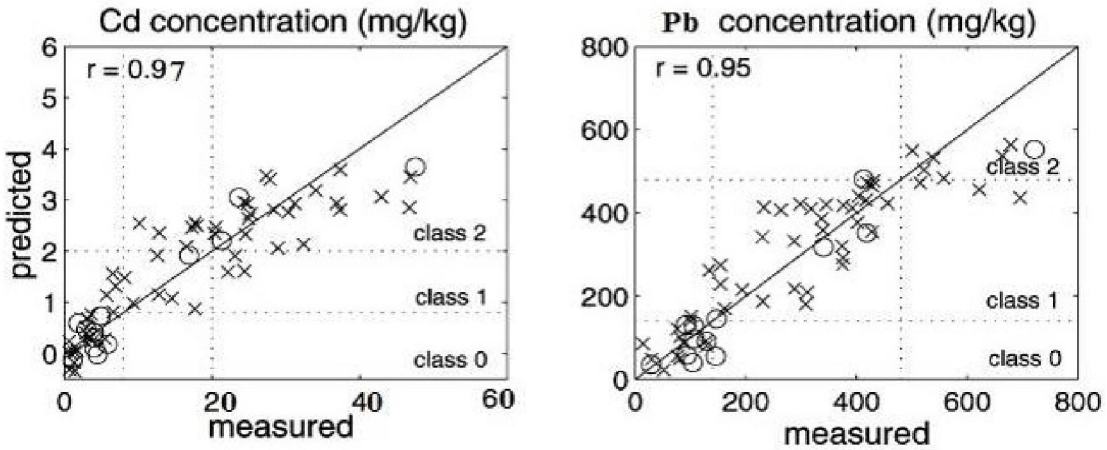
PLS cross-validation prediction

The calibration model was developed using leave-one-out cross validation (LOOCV) method in PLS regression analysis. The LOOCV procedure estimates the prediction error by removing samples one by one from the calibration samples data and predicting them as unknown samples using the remaining samples in the data. The PLS model training process continues until the minimum prediction residual sum of squares (PRESS) is attained, which is used to choose the optimal number of PLS factors in order to diminish the possibility of overfitting the model (**Haaland and Thomas 1988**).

After excluding the spectral outliers, the remaining samples were randomly split into two sets. One (calibration set) set of samples was used for constructing MIR-PLS calibration model and the second set was used for the validation of the calibration model, referred to as validation set, Figure (11).

It is also important to estimate the true errors in the prediction model and validate it using a set of unknown soil samples, i.e. validation set (n = 49), as mentioned earlier. The prediction equation obtained from the calibration model was applied to the validation set in order to predict the concentration of Pb and Cd in soil in these samples. To evaluate the efficiency of the prediction-model, predicted Pb and Cd concentration values were plotted against the measured soil pollutants concentrations. The following statistics was used to assess the prediction ability of calibration and validation models including, the coefficient of determination in calibration (R^2c) and validation (R^2v), the standard error of cross validation (SECV) in calibration, the standard error of prediction (SEP) in validation, and the residual prediction deviation (RPD) in calibration (RPDc) and validation (RPDv) models (**Pirie et al. 2005; Islam et al. 2003**).

As the obtained results table (5) it si clearly to see that the clay and Cd have RSQ more than 97% prediction while, Pb and Lime has RSQ more than 95%.



With MIRS (Mid infrared) and with scatter correction and derivative of 1st order (pre-treatment SNV): the adjustments are not enhanced. Table (5)

Figure (11) predicted value and measured of soil pollutant

Table 5: PLS cross-validation prediction

Constituent	Type	N	Mean	Est. Max	Est. Min	Est. Max	SEC	RSQ	SECV	1-VR
C	1	49	55.8394	140.6643	0.0000	140.6643	4.2949	0.9769	5.0926	0.9674
L	1	49	54.7794	141.7252	0.0000	141.7252	7.1990	0.9383	9.5422	0.8909
Pb	1	49	53.4286	137.8047	0.0000	137.8047	6.3651	0.9488	8.4569	0.9092
Cd	1	49	54.9635	139.4352	0.0000	139.4352	4.0267	0.9795	4.7814	0.9710

Conclusions:

Reflectance spectroscopy is a simple and nondestructive analytical method that can be used to predict not only spectral active constituents but also trace elements, which are spectrally featureless.

This study showed that high resolution spectra of soil samples can be used for predicting the soil contamination levels in this material. Soil spectroscopy in the MIR region with a PLS model is shown to be a very promising method for the determination of both soil properties and metal concentrations in soils.

Mid-infrared spectroscopy exhibited potential utility in detection of metal-contaminated soils. Results indicate that MIR may be used for quantitative measurements of metals in diverse soils if the calibration is developed from soils within the region.

Application of the calibration models for soil quality characterisation could be achieved. Types of soil and pollutant can be recognized from the spectra with acurately also the concentrations.

Corresponding author:
Ahmed A. Afifi
Soils and Water Use Dept.. National Research Centre, Doki, Giza, Egypt. a.afifinrc@gmail.com

References

1. Brown D.J., Bricklemyer R.S., Miller P.R. Validation requirements for diffuse reflectance soil characterization models with case study of VNIR soil C prediction in Montana. Geoderma 129, 251, 2006.
2. Casler M.D., Shenk J.S. Effect of sample grinding on forage quality estimates of smooth bromegrass clones. Crop Sci. 25, 167, 1985.
3. Cornell RM, Schwertmann U (2003) The iron oxides: structure, properties, reactions, occurrences, and uses. Completely rev. and extended edn. Wiley-VCH, Weinheim.
4. Couteaux M.M., Berg B., Rovira P. Near infrared reflectance spectroscopy for determination of organic matter fractions including microbial biomass in coniferous forest soils. Soil Biol. Biochem. 35, 1587, 2003.
5. Couteaux M.M., Mctternan K.B., Berg B., Szuberla D., Dardenne P., Bottner P. Chemical composition and carbon mineralization potential of Scots pine needles at different stages of decomposition. Soil Biol. Biochem. 30, 583, 1998.
6. Foley W.J., Mcilwee A., Lawler I., Aragones L., Woolnough A.P., Berding N. Ecological applications of near infrared reflectance spectroscopy – a tool for rapid, cost-effective prediction of the composition of plant and animal tissues and aspects of animal performance. Oecologia 116, 293, 1998.
7. Haaland DM, Thomas EV (1988) Partial least-squares methods for spectral analyses. 2. Application to simulated and glass spectral data. Anal Chem 60 (11):1202-1208.
8. Haberhauer G, Rafferty B, Strebl F, Gerzback MH (1998) Comparison of the forest soil litter derived from three different sites at various decompositional stages using FTIR spectroscopy. Geoderma 83:331-342.
9. Huang, Q., Li, F.S., Xiao, R.,Wang, Q.H., Tan,W.J., 2008. Characterization of organo-mineral aggregates of chernozem in Northeast China and their adsorption behavior to phenanthrene.
10. Ibrahim M, Hameed AJ, Jalbout A (2008) Molecular spectroscopic study of River Nile sediment in the greater Cairo region. Appl Spectrosc 62 (3):306-311.
11. Islam, K., Singh, B., & McBratney, A. 2003. Simultaneous estimation of several soil propertie by ultra-violet, visible, and near-infrared reflectance spectroscopy. *Australian Journal of Soil Research* 41, 1101–1114.
12. Janik LJ, Merry RH, Forrester ST, Lanyon DM, Rawson A (2009) A rapid prediction of soil water retention using Mid Infrared spectroscopy. Soil Sci Soc Am J 71:507-514.
13. Johnston CT, Aochi CT (1996) Fourier transform infrared and raman spectroscopy. In: Sparks DL (ed) Methods of Soil Analysis, Part 3: Chemical Methods. Soil Science Society of America, Madison, WI, USA, pp 269-321.
14. Ludwig B., Khanna P.K. Use of near infrared spectroscopy to determine inorganic and organic carbon fractions in soil and litter. In: Assessment Methods for soil Carbon. Advances in Soil Science. pp 361 – 370, 2001.
15. Martens H.A., Dardenne P. Validation and verification of regression in small data sets. Chemom. Intell. Lab. Syst. 44, 99, 1998.
16. Nas T., Isaksson T., Fearn T., Davies T. A user friendly guide to multivariate calibration and classification. NIR Publications, Chichester, UK. 2002.
17. Pirie A, Singh B, Islam K (2005) Ultra-violet, visible, near-infrared, and mid-infrared diffuse reflectance spectroscopic techniques to predict several soil properties. Aust J Soil Res 43 (6):713-721.
18. Shenk J.S., Westerhaus M.O. Population definition, sample selection, and calibration procedure for near infrared reflectance spectroscopy. Crop Sci. 31, 469, 1991.
19. Wold S., Sjostrom M., Eriksson L. PLS-regression: a basic tool of chemometrics. Chemom. Intell. Lab. Syst. 58, 109, 2001.
20. Workman J. Handbook of Organic Compounds: NIR, IR, Raman, and UV-Vis Spectra Featuring Polymers and Surfactants. Vol 1, Academic Press, pp 77 – 197, 2000.
21. Workman Jr J.J. A brief review of the near infrared measurement technique. NIR news 4, 8, 1993.

11/11/2014