Application of Non-Linear Statistical Tools to a Novel Microwave Dipole Antenna Moisture Soil Sensor

Published Version:
Berardinelli, A., Luciani, G., Crescentini, M., Romani, A., Tartagni, M., Ragni, L. (2018). Application of Non-Linear Statistical Tools to a Novel Microwave Dipole Antenna Moisture Soil Sensor. SENSORS AND ACTUATORS. A, PHYSICAL, 282, 1-8 [10.1016/j.sna.2018.09.008].

Availability:
This version is available at: https://hdl.handle.net/11585/645055 since: 2019-02-04

Published:
DOI: http://doi.org/10.1016/j.sna.2018.09.008

Application of Non-Linear Statistical Tools to a Novel Microwave Dipole Antenna Moisture Soil Sensor

A. Berardinelli[a*], G. Luciani[b,c] M. Crescentini[b,d], A. Romani[b,d], M. Tartagni[b,d], L. Ragni[a,e]

[a]Department of Agricultural and Food Sciences, Alma Mater Studiorum, University of Bologna, P.zza Goidanich 60, 47521 Cesena (FC), Italy

[b]Department of Electrical, Electronic and Information Engineering "Guglielmo Marconi" - University of Bologna, Via Venezia 52, Cesena 47521, Italy

[c]Center for Industrial Research on ICT (CIRI ICT), University of Bologna, via Venezia 52, 47521 Cesena (FC), Italy.

[d] Advanced Research Centre on Electronic Systems (ARCES), University of Bologna, 40123 Bologna, Italy.

[e]Inter-Departmental Centre for Agri-Food Industrial Research, Alma Mater Studiorum, University of Bologna, P.zza Goidanich 60, 47521 Cesena (FC), Italy

*Corresponding author.

Abstract

In this paper we will show the boosting performance of nonlinear machine learning techniques applied to a novel soil moisture sensing approach. A probe consisting in a transmitting and a receiving dipole antenna was set up to indirect assess the moisture content (%) of three different types of soils (silty clay loam, river sand and lightweight expanded clay aggregate, LECA). Gain and phase signals acquired in the 1.0 GHz – 2.7 GHz frequency range were used to built predictive models based on linear PLS regression and on nonlinear Kernel-based orthogonal projections to latent structures (K-OPLS) algorithms. K-OPLS algorithm slightly increased the accuracy of the models built on the gain response on specific kind of soils with respect to classical linear PLS. However, the predictability increases significantly in the case where the models are built starting from a matrix containing all the considered soil samples (silty clay loam + river sand + LECA) achieving R^2 = 0.971 (RMSE = 1.4 %) when using K-OPLS non-linear model with respect to R^2 = 0.513 (RMSE = 6.1 %) obtained using linear PLS. Therefore, K-OPLS algorithm appears to give a significant improvement to modelling data where nonlinear behaviours occur.

Keywords. Dielectric spectroscopy; soil moisture content; linear and nonlinear multivariate data analysis; PLS; K-OPLS.

1. Introduction

The spectroscopic and time-domain analyses of the interaction between the electromagnetic wave and the agricultural soil is, at date, widely explored methods for the indirect assessment of its water content [1]. The acquired waveforms appeared to contain information related to different soil physico-chemical properties and the quantitative estimation accuracy is affected by two main factors as the used techniques and the statistical tools [2]. Examples of these techniques are Visible, Near and Infrared sensors [2,4], Theta probes, measuring apparent impedance at 100 MHz [5] and the Time Domain Reflectometry (TDR) [6], based on the analysis of the propagation time of the electromagnetic wave through a coaxial cable to a probe immersed in the medium (20 kHz - 1.5 GHz), a function of the soil dielectric permittivity.

Powerful multivariate data analysis tools able to relate two data matrix X (spectra acquired from several samples) and Y (the analytical properties), have played a big role in the development of the techniques [7].

Originated around 1975, the widespread linear multivariate Partial Least Squares (PLS) regression is considered a standard procedure in chemometrics and it has been shown to be potential for extracting useful information starting from highly linearly correlated data coming from bioengineering indirect measurements. The tool uses a two-block quantitative PLS model based on a latent variable decomposition of X and Y variables keeping most of the variance of the explanatory variables. It is well known that PLS regression has proven to be extremely useful in situation when the number of observed variables is much higher than the numbers of acquired samples, typical situation with spectral data [8].

However, non linear behaviours are very frequent in biosystems, such as the light absorbance in milk, dependent on fat content [9], or the dielectric permittivity in microwave region, dependent on the soil moisture [10], just to cite a couple of examples. Samples variability and level of complexity of the matrices together with temperature fluctuations and interactions between sensor and product can negatively affect the robustness of PLS models and cause non linear behaviours as shown in different works conducted on quantitative assessment of fruits chemical properties through NIR spectral measurements [11-13]. Agricultural soil is a complex heterogeneous matrix characterised by organic (humus and different particulate residues) inorganic mineral fractions (proportions of sand, silt and clay particles), moisture and air [14]. Conversely, multivariate regression models based on non linear machine learning tools have shown significant improvements in the accuracy of the prediction of different physical and chemical properties of this complex matrix [4,7,15].

In order to improve robustness of PLS models in presence of non linearity, a considerable number of methods integrating non linear features within the linear PLS algorithm have been proposed. Quadratic PLS [16], smooth bivariate spline function [17], Neural Network PLS [18], Radial Basis Function (RBF) neural networks [19], and Kernel PLS (KPLS) [20] are some examples of the proposed machine learning implementation in PLS modelling. In KPLS the original X variables are transformed into a high-dimensional feature space by a non linear mapping. In this feature space a linear relationship can be appreciated and the PLS algorithm can then be performed; the feature space is defined after selecting a kernel function providing a similarity measure between pairs of spectra [21]. The accuracy of the KPLS algorithms were tested with images analysis generated by an airborne scanner with nine wavelength bands (from 500 to 10487 nm) [22], with UV-visible and FT-IR spectra for the prediction of different mixtures contents [23] and with NIR spectra for the prediction of apple sugar content [24] and for a rapid screening of water samples containing malathion [24].

Our approach is substantially different with respect to the above-mentioned electro-magnetic techniques. Differently from TDR, it is based on spectra analysis in the frequency domain instead of the time-domain. Then, in contrast with commonly used IR spectra techniques, we perform a spectral analysis of transfer functions involving microwaves. This ensures a better interaction with soil in terms of depth of penetration and also utilizes higher information content given by the phase. Finally, we use non-linear machine learning tools to boost the statistical inference of data.

A new probe in the dielectric sensors panorama characterised by a transmitting and a receiving dipole antenna was set up for the indirect assessment of the moisture content of different types of soils: silty clay loam soil, river sand and lightweight expanded clay aggregate (LECA). This innovative probe requires the previously drilling of the soil and then the inserting of the probe. With respect to traditional TRD probes it could less suffer for incomplete adherence of the soil to the sensor. In fact, the sensing is performed in a large portion of the volume surrounding the probe and any interference, such as air, can be removed by the powerful statistical analysis. Therefore, the information contained in both gain and phase signals acquired in the 1.0 GHz – 2.7 GHz frequency range, will be processed by using the Kernel-based orthogonal projections to latent structures (K-OPLS, an implementation of KPLS with a solution able to separate structured noise). Predictive models of the moisture content will be built starting from data sets characterised by the same soil type or starting from data sets containing all the analysed soil types.

2. Materials and methods

2.1 Probe and acquisition chain

The probe, designed to be inserted in the soil, assembles a transmitting (TX) and a receiving (RX) dipole antenna, spaced 50 mm, located in a 170 mm long PVC sealed pipe, with outer and inner diameter of 16 and 13 mm, respectively. Both TX and RX antennas consists of a ¼ of ring per pole. The dipole was mounted on a made of nylon ring and placed in the pipe rotated by 90° one with respect to the other to avoid direct coupling of the EM signal from transmitting to receiving antenna. A layout of the probe containing the dipoles is shown in Fig. 1a together with the particulars of the dipole antenna (b) and the probe inserted in the soil (c). The above described prototype was designed for moisture determination in the soil layer pertaining the secondary tillage. A longer probe, containing an array of antennas, suitably spaced, could be constructed for in depth stratified moisture assessment (Fig. 1d). The TX antenna was connected to a sweeper oscillator (HP8350B combined with the HP83592B plug in), by means of a power divider. The signal from the other output of the divider and that coming from the RX antenna were connected to a gain and phase comparator (Analog Devices AD8302) through a 20 dB attenuator. The outputs of the comparator give a measurement of both gain over a ±30 dB range, scaled to 30 mV/dB, and of phase between signals over a 0°–180° range, scaled to 10 mV/degree. The output of the comparator was connected to a sampling board (National instrument, DAQ USB-4431) with 24 bit of resolution and sampling frequency from 1 kS/s to 102 kS/s. The board was connected to the PC. LabVIEW software was used to display the spectrum and decimate the sampling frequency for reducing the number of data. A layout of the instrumental chain was depicted in Fig. 2. The sinusoidal oscillation (13 dBm) was linearly swept from a frequency of 1.0 GHz to 2.7 GHz in 60 s.

2.2 Soil samples

Waveform acquisition was conducted on three different soil samples: silty clay loam soil (collected from Romagna region agricultural soil, Italy), river sand (Bacchi S.P.A., Italy), and lightweight expanded clay aggregate (LECA) (Laterlite, Italy). According to USDA textural classification [26], the chosen materials are characterised by very distinct physical properties (textural classes). Silty clay loam soil is made of particles with the following size (s) distribution: s < 0.002 mm (34%); 0.002≤ s < 0.050 mm (49%); 0.050 ≤ s ≤ 2 mm (17%). Density of the silty clay soil (at 6% of moisture) is 1137 kg/m^3. River sand, classified as “sand”, consisting of grain with maximum size of 0.6 mm (density at 0.1% is 1371 kg/m^3). LECA consists of granules with dimensions from 4 to 10 mm (density at 0.2% of moisture is 380 kg/m^3) During measurements, silty clay loam samples were characterised by clods of size suitable for seeding operations.

For each soil type, five different hydration levels (moisture contents, %) were considered (standard error in brackets): 5.7 (0.02) %, 9.9(0.04)%, 15.9(0.21)%, 22.3(0.17)% and 27(0.50)% for silty clay loam soil; 0.1(0.003)%, 4.8 (0.04%, 9.6% (0.17), 14.5(0.41)% and 18.9 (0.38)% for river sand; 0.2 (0.01)%, 5.4(0.13)%, 13.1(0.24)%, 19.7(0.61)% and 28%(0.42) for LECA. The lowest level sample was assessed by a thermo-gravimetrical method [27]. Starting from this sample, the remaining ones were produced by adding a specific amount of deionised water. The highest level to produce was chosen taking into account the field capacity, evaluated on each type of soil according to Cihlar and Ulaby [28]. To allow the water diffusion, the samples were hermetically isolated for about 48 hours before the dielectric acquisitions.

2.3 Experimental set up

Each soil sample (type × hydration level) was placed (constant volume of 17600 cm^3) in a plastic cylindrical container (20 cm × 14.4 cm) (Fig. 1). Spectral acquisitions were conducted at constant temperature of 22°C (± 1°C); all the spectra were sampled at 100 Hz from 1.0 GHz to 2.7 GHz (260 kHz of resolution). For each moisture level, nine different acquisitions were conducted. These acquisitions were obtained by rotating the container with respect to its axis each time by an approximately constant angle (40°). Having the temperature a noticeable effect on the complex permittivity in the microwave range, a test was carried out on a silty clay loam with 16% of moisture at 2 ° C and 22 ° C. Five measurement repetitions for each temperature were carried out for this test.

2.4 Data analysis

For each type of soil and for both gain and phase spectra, data were arranged in a 45 (samples) × 6536 (independent variables, gain or phase) X matrix and in a 45 (samples) × 1 (dependent variable, moisture content %) vector Y column. Six (3 soil types × gain or phase) predictive models of the moisture content (%) were obtained and discussed by means of PLS regression analysis of gain or phase signals (MatlabR2016b, Statistics and Machine Learning Toolbox). *N*-fold cross validation (random segments with about 14% of the samples in each segment) was used to validate multivariate models and accuracy was expressed in terms of coefficient of determination R^2 and Root Mean Square Error (RMSE). In *N*-fold cross validation, models are subsequently built from the remaining samples (n-1 segments). According to Esbensen [29], *N*-fold cross validation can be considered a powerful solution when there is a relative number of objects in the training data set.

The optimal number of PLS components was selected by analysing the validation residual variance [29].

In order to improve the accuracy of the PLS prediction in presence of possible non linear behaviours, the Kernel-based Orthogonal Projections to Latent Structures (K-OPLS) (K-OPLS package for MATLAB, http://kopls.sourceforge.net/) was explored. The algorithm constructs a regression model for predicting the dependent variable Y (moisture content, %) by using the information contained in a Kernel matrix.

Orthogonal projection to latent structures (O-PLS) can improve the model interpretation and reduce its complexity by removing systematic variability in X that is not correlated with Y or, in mathematical terms, that is orthogonal to Y [30]. O-PLS can be considered as a combination of PLS algorithm with a pre-processing tool and relative models are often characterised, for a given accuracy, by a lower number of PLS components [31].

The transformation to higher dimensional spaces is performed by using a kernel function k (x,y). Particularly, the K-OPLS models were fitted using the Gaussian kernel function:

$$k(x,y) = \exp(-\|x-y\|^2 / 2\sigma^2)$$

where σ parameter was selected according to a traditional approach by performing a grid search over a defined parameter interval; at each setting, *n*-fold cross validation (random segments with about 14% of the samples in each) was used to identify the parameter value yielding the error minimisation [32]. *N*-fold cross validation (random segments with about 14% of the samples in each segment) and errors plot were also used to identify the optimal number of predictive PLS components and Y-orthogonal components.

The potentiality of PLS and K-OPLS algorithms were also explored with the aim of building predictive models of soil moisture content independently from the soil type. To this purpose, two 135 (silty clay loam soil + river sand + LECA samples) × 6536 (independent variables, gain or phase) X matrices and two 135 (silty clay loam soil + river sand + LECA samples) × 1 (dependent variable, moisture content %) vector Y columns were processed.

3. Results

As an example, the acquired waveforms (average of 9 acquisitions for each hydration level) are reported in Figure 3 in the 1.10 – 1.36 GHz frequency range for both gain and phase signals. For all the analysed soil types, different moisture contents appeared to produce differences in the gain and phase spectrum. As can be observed for the gain spectra, river sand gain are characterised by significant variations and by a behaviour that can be evidently non linear. The waveform acquired at temperature of 2°C and 22°C for a silty clay soil with 16 % of moisture were shown in Figure 4.

From this figure, shifts of the waveforms are observable from 1.18 GHz to 1.28 GHz and from 1.30 GHz, for gain and phase respectively. More evident differences were found in other spectral regions (but not used for statistical analysis). All these difference can be attributed to the dependence of the loss factor and dielectric constant on the temperature, but they are plausibly influenced by the complex water-soil chemical-physical interactions. On the contrary, several spectral regions appeared influenced by moisture content but not (or minimally) by the temperature, so the moisture prediction algorithms could take advantage of this feature.

Main results of the PLS and K-OPLS regressions are summarised in Table 1 for gain and phase models characterised by the same soil type and for those built starting from all the three analysed soils. The models accuracy is described in terms of R^2 and RMSE obtained by performing *n*-fold cross validations in the 1.10 – 1.36 GHz frequency range, resulted to produce, in general, the best performances.

For classical linear PLS models, highest R^2 values emerged for phase acquisitions: 0.980 (RMSE = 1.1 %), 0.983 (RMSE = 0.9 %), and 0.981 (RMSE = 1.4 %) respectively for silty clay loam, river sand, and LECA soils. Phase spectra appeared to be better linearly correlated to the dielectric characteristics for all the three soil samples. Gain acquisitions on these soils are respectively characterised by R^2 values of 0.970 (RMSE = 1.4 %), 0.960 (RMSE = 1.4 %), and 0.975 (RMSE = 1.6 %).

Passing from PLS to K-OPLS results, a significant improvement in the regression can be observed in a single case, for river sand gain model (R^2 = 0.988, RMSE = 0.7 %). In this case, the K-OPLS algorithm seemed to better interpret nonlinear variability observed among the five moisture levels. As expected, all K-OPLS models are characterised by only one predictive PLS component. One predictive PLS component appeared to be sufficient to discriminate river sand samples according to the moisture content as shown in Figure 5a plotting the K-OPLS predictive score vector Tp against the first Y-orthogonal score vector To. The hydration discriminatory direction is clearly described by the predictive PLS component while systematic and linearly independent variations are modelled by the first Y-orthogonal ones.

Predicted versus observed values of the moisture content (%) for river sand gain spectra obtained from K-OPLS is given in Figure 5b (*N*-cross validation).

Respect to classical linear PLS regressions, K-OPLS phase models appeared to be characterised by a significant lower accuracy for all soil types: 0.867 (RMSE = 2.3 %), 0.851 (RMSE = 2.0 %), and 0.946 (RMSE = 2.2 %) respectively for silty clay loam, river sand, and LECA soils.

For predictive models built starting from matrices including all soil types samples (silty clay loam + river sand + LECA), the nonlinear multivariate approach K-OPLS clearly better interprets the information correlated to the moisture content, respect to PLS regressions. Furthermore, the K-OPLS algorithm seems to overcame problems related to systematic variations due to differences among soil types. In *N*-fold cross validation, R^2 values of 0.971 (RMSE = 1.4 %) and 0.909 (RMSE = 1.6 %) were observed respectively for gain and phase acquisitions. Predicted versus observed moisture (%) are shown in Figure 6 for both gain and phase K-OPLS models. By performing PLS regression analysis, significantly lower R^2 values were calculated: 0.513 (RMSE = 6.1 %) and 0.553 (RMSE = 5.8 %) respectively for gain and phase acquisitions (Fig. 7). By plotting scores vectors (first component Tp-Up score plot including all soil types samples) obtained from both PLS and K-OPLS gain models, a linear correlation can be better appreciated for K-OPLS plot respect to PLS ones (Fig. 8).

The indirect prediction of the moisture content independently from the soil physical and chemical characteristics it is not easily obtainable by traditional linear models. An example is provided by Yin et al. [33] where a combination of 4 different soils can produce an R^2 value of 0.642 (RMSE up to 9.26 % according to the soil type and for a range of 0 – 52 %) starting from a NIR reflectance sensor. R^2 value of about 0.973 was also shown by Zanetti et al. [34] by using the apparent dielectric constant (K_a) obtained from TDR waveforms and a combination of different physical properties as input variables (bulk density, sand, silt, clay, and organic matter content) of ANN models.

4. Conclusions

A novel approach for soil moisture sensing based on non-linear machine learning tools applied to microwave spectra has been presented. A cylindrical dipole antenna probe operating in the low frequency microwave region, together with gain and phase spectral data processed by linear and nonlinear PLS statistical tools, shown to be a promising technique for soil moisture determination. Validation R^2 values for basic PLS were from 0.960 (RMSE = 1.4 %) to 0.983 (RMSE = 0.9%), depending on kind of soil and used spectra (gain or phase). Advanced K-OPLS algorithm allows to greatly improve the prediction accuracy independently on the kind of soil (R^2 = 0.971, RMSE = 1.4%, gain data). The current probe could be developed for moisture determination at several depths by equipping the sensor with an array of dipole antennas. Temperature can have a crucial influence on the measured waveforms, so the calibration dataset will have to take into account this parameter.

References

[1] N. Romano, Soil moisture at local scale: measurements and simulations, J. Hydrol. 516 (2014) 6-20.

[2] S.L. S.U., D.N. Singh, M. Shojaei Baghini, A critical review of soil moisture measurement, Measurement 54 (2014) 92-105.

[3] R.A. Viscarra Rossel, D.J.J. Walvoort, A.B. McBratney, L.J. Janik, J.O. Skjemstad, Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties, Geoderma 131 (2006) 59-75.

[4] A. Morellos, X-E. Pantazi, D. Moshou, T. Alexandridis, R. Whetton, G. Tziotzios, J. Wiebensohn, R. Bill, A.B. Mouazen, Machine learning based prediction of soil total nitrogen, organic carbon and moisture content by using VIS-NIR spectroscopy, Biosyst. Eng. 152 (2016) 104-116.

[5] D.A. Robinson, C.M.K. Gardner, J.D. Cooper, Measurement of relative permittivity in sandy soils using TDR, capacitance and theta probes: comparison, including the effects of bulk soil electrical conductivity, J. Hydrol. 223 (1999) 198-211.

[6] D.A. Robinson, S.B. Jones, J.M. Wraith, D. Or, S.P. Friedman, A review of advances in dielectric and electrical conductivity measurement in soils using time domain reflectometry, Vadose Zone J. 2 (2003) 444-475.

[7] A.M. Mouazen, B. Kuang, J. De Baerdemaeker, H. Ramon, Comparison among principal component, partial least squares and back propagation neural network analyses for accuracy of measurement of selected soil properties with visible near infrared spectroscopy, Geoderma 158 (2010) 23-31.

[8] S. Wold, M. Sjöström, L. Eriksson, PLS-regression: a basic tool of chemometrics, Chemometr. Intell. Lab. Syst. 58 (2001) 109-130.

[9] L. Ragni, E. Iaccheri, C. Cevoli, A. Berardinelli, Spectral-sensitive pulsed photometry to predict the fat content of commercialized milk, J. Food Eng. 171 (2016), 95-101.

[10] A. Szypłowska, J. Szerement, A. Lewandowski, M. Kafarski, A. Wilczek, W. Skierucha, Impact of soil salinity on the relation between soil moisture and dielectric permittivity, in: Proceedings of the 12th International Conference on Electromagnetic Wave Interaction with Water and Moist Substances, Lublin, Poland 4-7 June 2018, pp. 14-16.

[11] A. Peirs, J. Tirry, B. Verlinden, P. Darius, B.M. Nicolai, Effect of biological variability on the robustness of NIR models for soluble solids content of apples, Postharvest Biol. Technol. 28 (2003) 269-280.

[12] A. Peirs, N. Scheerlinck, B.M. Nicolai, Temperature compensation for near infrared reflectance measurement of apple fruit soluble solids contents, Postharvest Biol. Technol. 30 (2003) 233-248.

[13] F. Chauchard, R. Cogdill, S. Roussel, J.M. Roger, V. Bellon-Maurel, Application of LS-SVM to non-linear phenomena in NIR spectroscopy: development of a robust and portable sensor for acidity prediction in grapes, Chemometr. Intell. Lab. Syst. 71 (2004) 141-150.

[14] J.O. Skjemstad, P. Clarke, A. Golchin, J.M. Oades, Characterization of soil organic matter by solid-state 13C NMR spectroscopy, in: G. Gadish, K.E. Giller (Eds.), Driven by nature: plant litter quality and decomposition, CAB International, Wellington (UK), 1997, pp. 253-271.

[15] B. Stenberg, R.A. Viscarra Rossel, A.M. Mouazen, J. Wetterlind, Visible, and near infrared spectroscopy in soil science, in: D. L. Sparks (Ed.), Advances in Agronomy, Academic Press, Burlington, 2010, pp. 163-215.

[16] S. Wold, N. Kettaneh-Wold, B. Skagerberg, Non-linear PLS modelling, Chemometr. Intell. Lab. Syst. 7 (1989) 53-65.

[17] S. Wold, Non-linear Partial Least Squares Modelling II. Spline Inner Function, Chemometr. Intell. Lab. Syst. 14 (1992) 71-84.

[18] G. Baffi, E.B. Martin, A.J. Morris, Non-linear projection to latent structures revisited (the neural network PLS algorithm), Comput. Chem. Engin. 23 (1999) 1293-1307.

[19] D.J.H. Wilson, G.W. Irwin, G. Lightbody, Nonlinear PLS modelling Using Radial Basis Functions, American Control Conference, Albuquerque, New Mexico, June 4-6,1997.

[20] R. Rosipal, L.J. Trejo, Kernel partial least squares regression in Reproducing Kernel Hilbert Space, J. Mach. Learn. Res. 2 (2001) 97-123.

[21] R. Rosipal, Kernel partial least squares for nonlinear regression and discrimination, Neural Netw. World 13 (2003) 291-300.

[22] F. Lindgreen, P. Geladi, S. Wold, Kernel-based PLS regression; Cross-validation and applications to spectral data, J. Chemom. 8 (1994) 337-389.

[23] S. Rännar, P. Geladi, F. Lindgren, S. Wold, A PLS kernel algorithm for data set with many variables and few objects. Part II: Cross validation, missing data and examples, J. Chemom. 9 (1995) 459-470.

[24] B.M. Nicolaï, K.I. Theron, J. Lammertyn, Kernel PLS regression on wavelet transformed NIR spectra for prediction of sugar content of apple, Chemometr. Intell. Lab. Syst. 85 (2007) 243-252.

[25] C. Gu, B. Xiang, Y. Su, J. Xu, Near-Infrared spectroscopy coupled with Kernel Partial Least Squares-Discriminant Analysis for rapid screening water containing malathion, Am. J. Analyt. Chem. 4 (2013) 111-116.

[26] Soil Survey Staff, Soil survey manual, Soil Conservation Service U.S. Department of Agriculture Handbook, 18, 1993.

[27] ASTM D2216, Standard Test Methods for Laboratory Determination of Moisture (Moisture) Content of Soil, ASTM International, West Conshohocken, PA, 2008.

[28] J. Cihlar, F.T. Ulaby, Dielectric properties of soils as a function of moisture content, Kansas Univ. Center for Research, Inc. 1974, NASA-CR-141868.

[29] K. H. Esbensen, Multivariate Data Analysis – in practice. An introduction to multivariate data analysis and experimental design, 5th edition, CAMO Software AS, 1994.

[30] J. Trygg, S. Wold, Orthogonal projections to latent structures (O-PLS), J. Chemometr. 16 (2002) 119-128.

[31] D.J. Biagioni, D.P. Astling, P. Graf, M.F. Davis, Orthogonal projection to latent structures solution properties or chemometrics and systems biology data, J. Chemom. 25 (2011) 514-525.

[32] M. Bylesjö, M. Rantalainen, J.K. Nicholson, E. Holmes, J. Trygg, K-OPLS package: Kernel-based orthogonal projections to latent structures for prediction and interpretation in feature space. BMC Bioinformatics, 9 (2008) 106.

[33] Z. Yin, T. Lei, Q. Yan, Z. Chen, Y. Dong, A near-infrared reflectance sensor for soil surface moisture measurement, Comput. Electron. Agric. 99 (2013) 101-107.

[34] S.S. Zanetti, R.A. Cecílio, V.H. Silva, E.G. Alves, General calibration of TDR to assess the moisture of tropical soils using artificial neural networks, J. Hydrol. 530 (2015) 657-666.