

Four Decades of Progress in Monitoring and Modeling of Processes in the Soil-Plant-
Atmosphere System: Applications and Challenges

Prediction of soil texture distributions using VNIR-SWIR
reflectance spectroscopy

D. Curcio^a, G. Ciralo^b, F. D'Asaro^a, M. Minacapilli^{a*}

^a*Dipartimento di Scienze Agrarie e Forestali (SAF), Università di Palermo, Italy.*

^b*Dipartimento di Ingegneria Civile, Ambientale, Aerospaziale, dei Materiali (DICAM), Università di Palermo, Italy.*

Abstract

Reflectance spectroscopy provides an alternate method to non-destructively characterize key soil properties. Different approaches, including chemometrics techniques or specific absorption features, have been proposed to estimate soil properties from visible and near-infrared (VNIR, 400-1200 nm) and shortwave infrared (SWIR, 1200-2500 nm) reflectance domains. The main goal of this study was to test the performance of two distinct methods for soil texture estimation by VNIR-SWIR reflectance measurements: i) the Continuum Removal (CR) technique that was used to correlate specific spectral absorption features with clay, silt and sand content, and ii) the Partial Least-Squares Regression (PLSR) method, which is a classical statistical multivariate technique, that uses the full-spectrum data. At this aim, the surface reflectance of 100 soil samples collected from different sites in Sicily and covering a wide range of textures were measured in laboratory using an ASD FieldSpec Pro spectroradiometer (350-2500 nm). The results of our work indicated that the PLSR technique performed better than the CR approach. Particularly, the assessment of soil texture accuracy performed using root mean squared error (RMSE) and coefficient of determination (R^2) showed that the CR approach allowed to obtain a moderate prediction only for the clay texture fraction. Differently, using PLSR technique, the levels of accuracy resulted high for the clay fraction (RMSE=5.8%, $R^2=0.87$) and satisfactory for the sand (RMSE=7.7%, $R^2=0.80$) and silt fractions (RMSE=7.2%, $R^2=0.60$). Moreover the use of PLSR technique allowed to establish the “key wavelengths” of the investigated spectrum range that should be considered “essential” for the prediction of soil textures, suggesting the optimal settings for airborne or satellite sensors usable in the future for accurate mapping of soil textures.

© 2013 The Authors. Published by Elsevier B.V Open access under [CC BY-NC-ND license](https://creativecommons.org/licenses/by-nc-nd/4.0/).
Selection and/or peer-review under responsibility of the Scientific Committee of the conference

Keywords: Soil texture; Reflectance spectroscopy; Continuum Removal; Partial Least Squares Regression

* Mario Minacapilli. Tel.: +39-091-238-97075; fax:+39-091-484-035.
E-mail address: mario.minacapilli@unipa.it

1. Introduction

Soil texture is an important land environmental variable because it plays a key role in soil degradation and water transport processes, controlling soil quality and its productivity [1]. In Soil Science, methods for mapping soil texture at various scales are needed for several applications, but, typically, large numbers of samples must be collected and analyzed in order to adequately estimate the soil texture spatial variability. The complex and expensive conventional methods and field surveys are currently driving the scientific community to develop indirect estimation methods based on proximal and remote sensors (ground-based or airborne), including reflectance spectroscopy [2,3]. A number of approaches, including chemometrics techniques or specific absorption features, have been proposed to estimate soil properties from visible and near-infrared (VNIR, 400-1200 nm) and shortwave infrared (SWIR, 1200-2500 nm) reflectance domains. In the last two decades, these methods have focused on low-cost methods that can be used to characterize many soil variables (texture, CaCO_3 , water content, etc.) according to their reflectance in the 400 to 2500 nm wavelength range [4,5]. Quantitative soil spectral analysis by reflectance spectroscopy requires sophisticated techniques to discern the response of soil attributes using spectral characteristics. Various laboratory methods have been developed to relate the soil spectrum to soil attributes, but usually two methods are considered foremost, i.e. Continuum Removal (CR) and Partial Least-Squares Regression (PLSR).

According to the CR technique the reflectance spectra are normalized allowing comparison of individual absorption features with respect to a common baseline [6]. This technique presents the advantage of targeting specific absorption features that should be persistent across scales and observation conditions. Moreover, the CR technique is computationally easier as compared to the chemometrics techniques, i.e. multivariate statistical techniques or PLSR, and it does not require complex pre-elaboration of spectra. Following the continuum removal application, absorption band depth values are calculated from VNIR and SWIR spectra to estimate mineral, rock, and soil properties. Specific absorption features of minerals, rocks, and soil properties, including clay and CaCO_3 , have been widely studied using the CR technique under laboratory conditions [7] and from remotely sensed data [8]. The PLSR technique is a standard multivariate statistical technique for spectral calibration and prediction of material properties [9,10], which uses the full investigated spectrum. The significant information contained in the spectral signature is concentrated in a few latent variables that are optimized to produce the best correlation with the desired properties of interest. Compared to other multivariate statistical techniques, PLSR is generally preferable because it is more understandable and the algorithm is computationally faster. A complete description of theoretical aspects of PLSR can be found in the specific literature [11]. A number of studies have successfully demonstrated the effectiveness of using reflectance spectroscopy to characterize soil of various pedo-climatic environments of the world. However, few studies have been carried out on Mediterranean soils [12,13].

In this general context, the main objective of this study was to assess the performance of VNIR and SWIR reflectance spectroscopy to predict soil texture for a set of 100 soil samples collected from several sites in Sicily. Particularly, the CR and PLSR techniques were applied and compared in order to define a laboratory operational protocol to predict soil texture.

2. Material and methods

Soil samples used in this research were collected from different sites in Sicily (Figure 1), as follows: 78 soil samples fall within three specific agricultural areas (23 from “Bompensiere”, 25 from “Dirillo”, 30 from Pietranera), while the remaining 22 soil samples (grouped under the heading “Other Soils”) were collected in various agricultural and/or forestry sites.

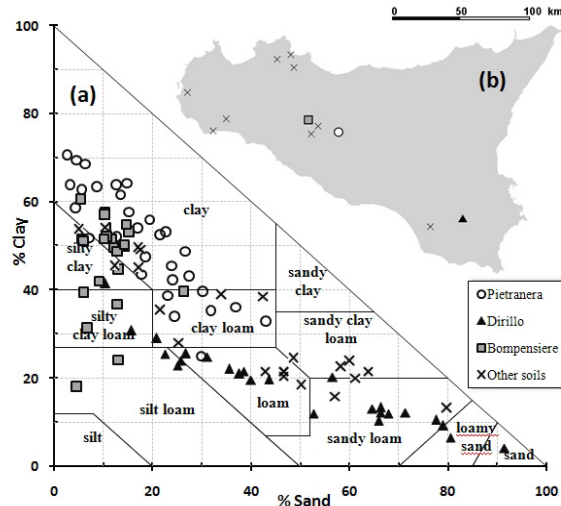


Fig. 1. (a) Texture classification according to USDA triangle and (b) geographic locations of the soil sampling sites.

Figure 1 reports the geographic locations of the sampling sites and the distribution of the 100 considered soil samples within the USDA texture triangle.

Texture was determined according to the Italian Official Methods for soil analysis [14]. In particular, clay and silt contents were determined by the hydrometer method, whereas total sand was determined by sieving.

In order to set up the spectral measurements, the sieved soil samples were packed inside circular black cells (5 cm diameter by 0.5 cm depth) and leveled with the edge of a spatula to obtain a smooth surface. A portable ASD FieldSpec-Pro spectroradiometer (manufactured by Analytical Spectral Device, Inc.) was employed for reflectance measurements. The instrument covers the VIS to SWIR wavelengths (350–2500 nm) with a sampling interval of 1.4 nm for the 350–1000 nm region, and 2 nm for the 1000–2500 nm region. Reflectance measurements were acquired in laboratory using a 25° foreoptic (Fig. 2). Soil samples were illuminated with two light sources (Lowel Light Pro, with JCV 14.5 V-50WC halogen lamp) positioned in the same plane, under a 45° illumination angle and from a distance of 20 cm. Spectral measurements were taken from nadir at 4 cm above the sample surface. Reflectance was calibrated against a white reference panel (Spectralon Diffuse Reflectance Panel). All spectral measurements were collected in a dark room to avoid contamination by stray light.

To analyze and exploit relationships between soil texture and VNIR/SWIR reflectance, two approaches were assessed: Continuum Removal (CR) and Partial Least-Squares Regression (PLSR).

With the CR technique, the reflectance spectra are normalized allowing comparison between absorption features, such as depth, depth position, area of the absorption, etc. [15]. These characteristics can be calculated from a continuum-removed spectrum to enhance absorption features in the recorded soil reflectance spectra. Clark and Roush [6] first suggested the CR analysis to isolate individual absorption features of interest. The so-called continuum of a generic spectrum is a convex “hull” of straight-line segments fitted over the top of the spectrum that connect local spectral maxima [6].

$$\hat{j}_i = b_0 + b_1 r_1 + b_2 r_2 + \dots + b_n r_n \quad (3)$$

in which b_0 to b_n are a set of coefficients (scores from PCs analysis) and r_1 to r_n are the reflectance values of each spectra after the 1st derivative pre-processing. By a linear regression of r_i versus j_i the set of regression coefficients of eq. (3) were obtained.

For both approaches (CR and PLSR), in order to quantify the model's performance, the following statistical parameters were calculated. First, the regression slope, the intercept and the determination coefficient (R^2) of the least square regression between observed (O_i) and predicted values (P_i), then, the overall accuracy of the model's output was assessed using the Mean Absolute Difference (MAD), the Root Mean Square Difference (RMSD) and the efficiency index of Nash and Sutcliffe (E) [24]:

$$MAD = \frac{1}{N} \left(\sum_{i=1}^N |P_i - O_i| \right) \quad (4)$$

$$RMSE = \sqrt{\left(\frac{1}{N} \sum_{i=1}^N (M_i - O_i)^2 \right)} \quad (5)$$

$$E = 1 - \frac{\sum_{i=1}^N (M_i - O_i)^2}{\sum_{i=1}^N (O_i - \bar{O})^2} \quad (6)$$

Statistical parameters were calculated both in calibration (N=70) and validation (N=30); to this aim, first of all, the data set for validation was randomly defined for both approaches (CR and PLSR). All data handling, regression calculations and PLSR analyses were performed using the Unscrambler 9.7 software [26].

3. Results and discussion

The spectra of all soil samples measured in laboratory and the average reflectance for each soil group are plotted in Figure 3. All spectra show the typical soil reflectance shape in each region of wavelength domain: reflectance is generally lower in the visible range (400-650 nm) and higher in the near infrared with specific absorbance bands around 1400, 1900 and 2200 nm.

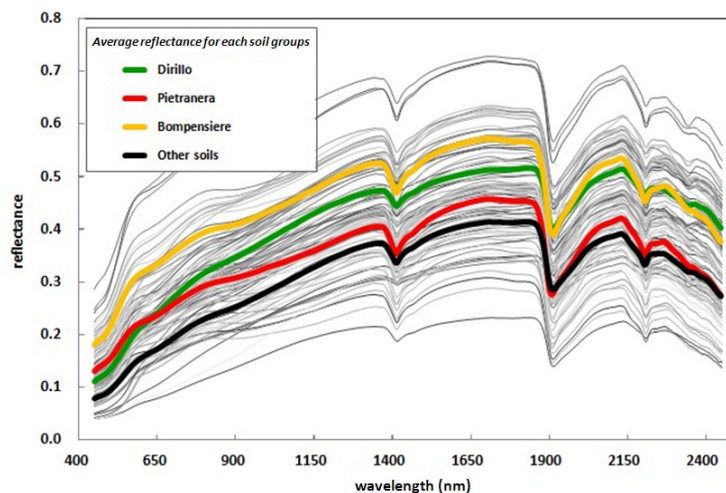


Fig. 3. Measured soil reflectance and average values for each soil group.

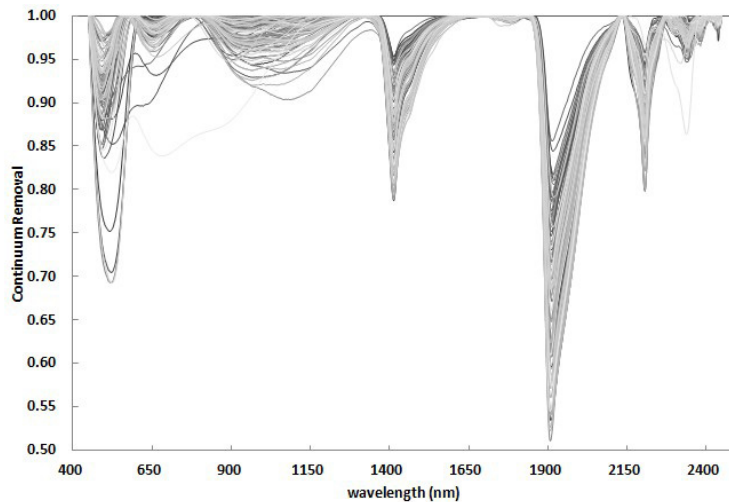


Fig. 4. Continuum-removed spectra obtained from measured reflectance of soil samples.

In terms of average reflectance the only differences between the four groups of soils are the higher reflectance values of “Bompensiere” and “Dirillo” soils. This circumstance can be ascribed to lighter colour of “Bompensiere” and “Dirillo” soils compared to the others groups.

The absorption features in the infrared region can be better analyzed using CR reflectance spectra (Fig. 4). Particularly, as noticed in literature [19,20,9,21] the dominant absorption features near 1400 nm, (CR_{1400}), and 1900 nm, (CR_{1900}), should generally be related to the presence of both hydration and crystallization water. Moreover, since CR_{1400} should also be related to the -OH vibration of water in silicates, it could be used to indirectly explain the percentage of the sand fraction that is generally constituted by these minerals.

Differently, as proved mainly in specific literature [4,5,18,17], the absorption peak at 2200 nm should be ascribed to the -OH-Al bending modes, as observed in various clay-soils. Therefore CR_{2200} can be directly used to explain clay fraction from spectral data. On the basis of these considerations in our study the CR values at 1400, (CR_{1400}), and 2200 nm, (CR_{2200}), extracted from continuum-removed spectral data-set were linearly correlated to measured clay and sand fractions (%) to verify the reliability of eq.(1) and (2) in the prediction of soil texture.

A statistical description of the performance of the CR approach to predict soil texture is summarized in Table 1 for both calibration and validation of the datasets. As can be observed, the CR approach allows to obtain a moderate-level performance only for clay prediction, characterized by RMSE values of 13.2 and 8.7 for the calibration and validation data set, respectively.

Table 1. Statistics of soil texture fraction obtained using the CR approach

	<i>Calibration</i> N = 70				<i>Validation</i> N = 30			
	R ²	MAD	RMSE	E	R ²	MAD	RMSE	E
Clay	0.47	10.5	13.2	0.47	0.68	7.2	8.7	0.67
Silt*	0.01	12.0	14.8	-0.36	0.012	11.4	13.4	-0.91
Sand	0.36	15.5	18.9	0.36	0.17	15.8	18.1	-0.15

* Silt fraction was obtained as residual term from clay and sand predictions

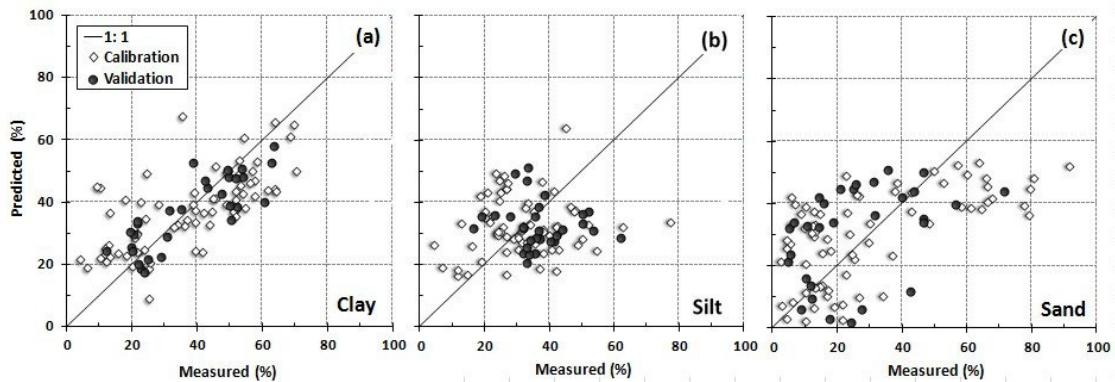


Fig. 5. Measured vs. Predicted soil texture using the Continuum-Removal approach.

This level of performance is corroborated by the scatterplots shown in Figure 5 where measured and predicted soil texture fractions were compared. The low level of sand-fraction prediction and consequently of silt fraction (derived as difference) could be ascribed to an overlap of some different effect that perturbs the absorption features at 1400 nm, as the presence of -OH-Al and/or -OH-Fe [4, 5].

A statistical description of the performance of the PLSR technique to predict soil texture is summarized in Table 2. As can be clearly observed, the PLSR technique is significantly more accurate than the CR approach for all soil texture classes. A very good result was obtained for clay, as highlighted by statistics parameters obtained from the validation data set ($R^2 = 0.87$, MAD = 4.5%, RMSE = 5.8%, E=0.855). A good prediction result was obtained also for sand ($R^2 = 0.80$, MAD = 6.2%, RMSE = 7.7%, E=0.793), whereas for silt the prediction value was only moderate ($R^2 = 0.60$, MAD = 5.3%, RMSE = 7.2%, E=0.444).

The good results obtained using the PLSR technique is clearly proved by the scatterplots shown in Figure 6. In this case, the distribution of points in the scatterplots suggests to correlate the lesser accuracy of silt prediction obtained from validation data set to the reduced range variability of predicted values compared to the measured ones.

The better results obtained by using the PLSR method are clearly due to the fact that PLSR takes advantage of the use of the entire spectral signature and not only specific absorption features in given wavelength as does the CR approach. Therefore, PLSR exploits other correlations between soil texture and surrogate soil properties that are also related to spectra. To examine the importance of these different causes, we used the so-called “loading weights” implemented in the PLSR algorithm [9] to optimize the choice of best PCs for predicting the dependent variable (soil texture). Loading weights show how PLSR latent variables were constructed from the predictors that, in our case, are represented by the reflectance at various wavebands. In our case the loading-weight distributions for each modelled variable were presented in Figure 7.

Table 2. Statistics of soil texture fraction obtained using the PLSR approach

	<i>Calibration</i> N = 70				<i>Validation</i> N = 30			
	R^2	MAD	RMSE	E	R^2	MAD	RMSE	E
Clay	0.87	5.0	6.6	0.87	0.87	4.5	5.8	0.85
Silt	0.82	4.5	5.4	0.82	0.60	5.3	7.2	0.44
Sand	0.89	6.4	8.0	0.89	0.80	6.2	7.7	0.793

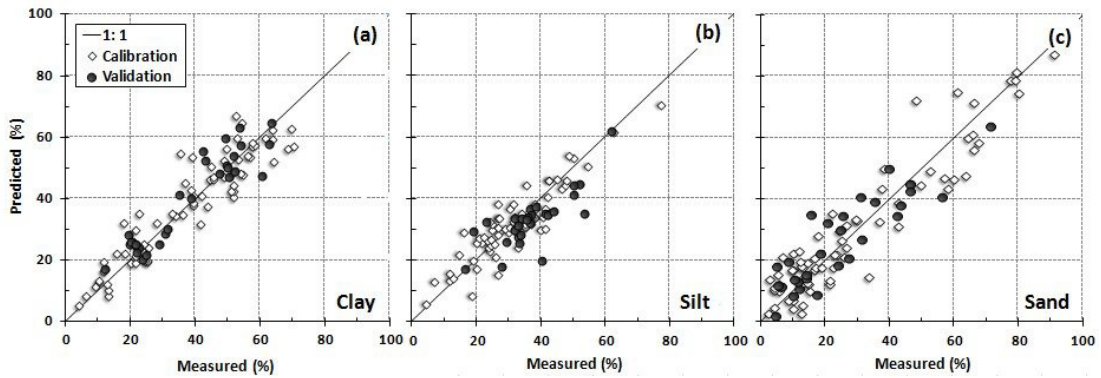


Fig. 6. Measured vs. Predicted soil texture using the PLSR technique.

The loading-factor distributions highlight the most significant spectral bands to properly predict soil texture, because wavelengths with small loading weights are considered less important than those with large ones [9]. The distributions of loading values plotted in Figure 7 confirmed that the wavelengths at 1400, 1900 and 2200 nm used in CR approach are sensitive for soil texture assessment. Moreover, the high significance of other spectral regions, i.e. VIS, is evident and this justifies the higher performance of the PLSR approach. Following this concept we attempt to determine the “key wavelengths” of the investigated spectrum that should be considered “essential” for the prediction of soil textures. To this aim we selected a reduced number of spectral bands characterized by loading weights greater than a threshold value that we set equal to 0.02. In this way we selected only the twenty “key wavelengths” labelled in Figure 7. We successively reapplied PLSR using only the selected “key wavelengths”. The results of this final simulation were summarized in Table 3 and Figure 8. As can be observed, despite the reduced number of spectral bands, the performance of the PLSR technique in prediction of soil texture was essentially unchanged respect to those obtained using the full spectrum.

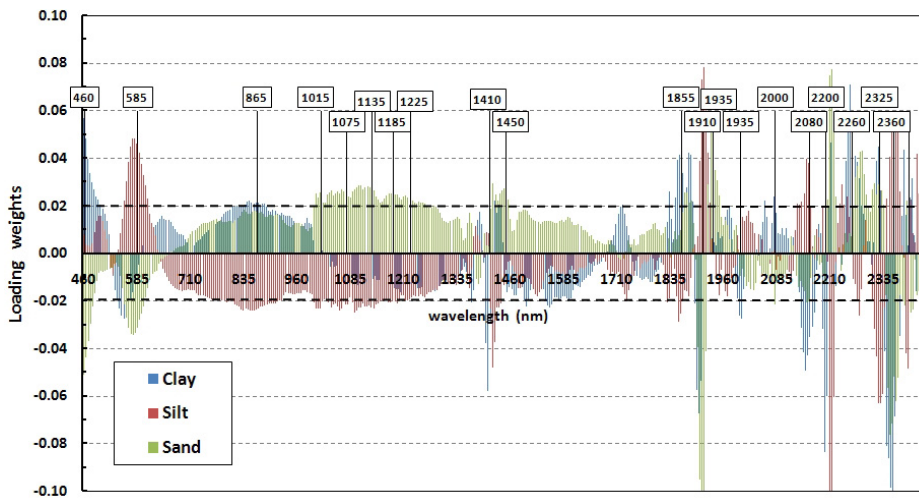


Fig. 7. Loading weights over wave bands for each predicted variables. In the plot the “keys wavelengths” with higher loading weights are also labelled.

Table 3. Statistics of soil texture fractions obtained using PLSR with the only twenty “keys wavelengths”

	<i>Calibration</i> N = 70				<i>Validation</i> N = 30			
	R ²	MAD	RMSE	E	R ²	MAD	RMSE	E
Clay	0.88	5.0	6.3	0.881	0.86	4.7	6.0	0.842
Silt	0.69	5.9	7.1	0.691	0.45	6.2	8.2	0.283
Sand	0.84	7.6	9.4	0.844	0.70	7.8	9.5	0.684

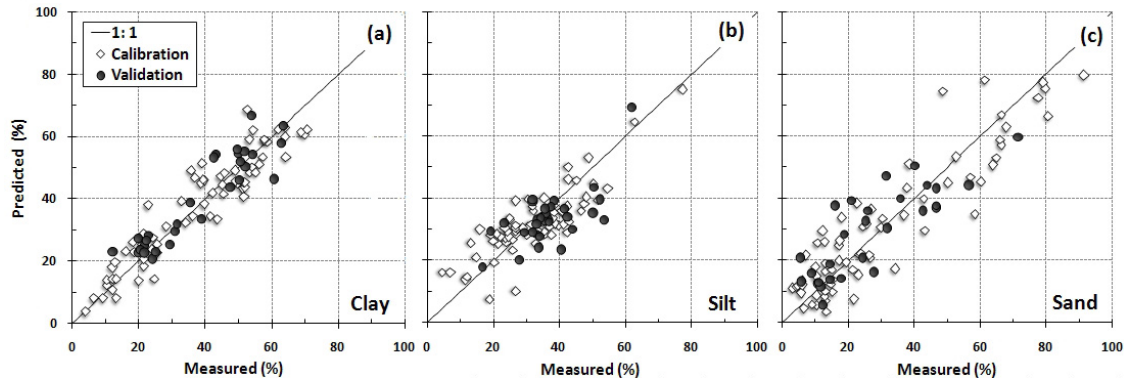


Fig. 8. Measured vs. Predicted soil texture using the PLSR technique with the only twenty “keys wavelengths”.

Considering that at the present a new generation of remote sensors provides hyperspectral data, sensed with an improved spectral resolution [2], our results can be considered a starting point to check the optimal settings usable in the future for accurate mapping of soil textures from remote platforms. To this aim further investigations will be focused on the use of spectroscopy (remotely and/or proximally sensed) for in-situ predictions of soil texture.

4. Conclusion

The main objective of this research was to test the performance of VNIR-SWIR reflectance spectroscopy to estimate soil texture. Two distinct methods were checked: i) the Continuum Removal (CR) approach and ii) the Partial Least-Squares Regression (PLSR) method. To this aim a laboratory experiment was set up using a set of spectral measurements performed on a set of 100 soil samples collected in Sicily.

The results suggest that a satisfactory level of prediction of soil texture can be obtained using only the PLSR technique, whereas a moderate level of prediction was obtained only for clay content using the CR approach. Furthermore, PLSR allowed to assess the “key wavelengths” of the entire spectrum that should be considered “essential” for the prediction of soil textures, suggesting the optimal settings for airborne or satellite sensors usable in the future for accurate mappings of soil textures. Further investigation is planned to assess the use of spectroscopy (remotely and/or proximally sensed) for in-situ predictions of soil texture.

Future applications will also be focused on evaluating the spatial distributions of soil properties using remotely sensed VNIR-SWIR data to test the methodology over larger areas and with mixed surfaces.

Acknowledgements

This work has been funded by Università degli Studi di Palermo and Progetto “*Metodologie innovative per la caratterizzazione idraulica e la valutazione della qualità fisica dei suoli siciliani, CISS*”, supervised by Prof. V. Bagarello. The Authors are grateful to Prof. M. Iovino and Prof. G. Provenzano for having provided the texture data-base and the soil samples used in this investigation.

References

- [1] Hillel D. *Applications of Soil Physics*. Academic Press Inc. 1980.
- [2] Ben-Dor E, Chabrilat S, Demattè JAM, Taylor GR, Hill J, Whiting ML, Sommer S. *Using imaging spectroscopy to study soil properties*. *Remote Sensing of Environment* 2009; **113**:38-55.
- [3] Brown DJ, Shepherd K D, Walsh MG, Mays MD, Reinsch TG.. Global soil characterization with VNIR diffuse reflectance spectroscopy. *Geoderma* 2006; **132**: 273–290.
- [4] Viscarra Rossel RA, Walvoort DJJ, McBratney AB, Janik LJ, Skjemstad JO.. Visible, near-infrared, mid-infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma* 2006; **131**: 59–75.
- [5] Viscarra Rossel RA, Cattle SR, Ortega A, Fouad Y. In situ measurements of soil colour, mineral composition and clay content by vis–NIR spectroscopy. *Geoderma* 2009; **150**:253–266.
- [6] Clark RN, Roush TL.. Reflectance spectroscopy: quantitative analysis techniques for remote sensing applications. *Journal of Geophysical Research*, 1984; **89**:6329–6340.
- [7] Gaffey SJ.. Spectral reflectance of carbonate minerals in the visible and near infrared (0.35–2.55 μm): Calcite, aragonite and dolomite. *American Mineralogist* 1986; **71**:151–162.
- [8] Chabrilat S, Goetz AFH, Krosley L, Olsen HW.. Use of hyperspectral images in the identification and mapping of expansive clay soils and the role of spatial resolution. *Remote Sensing of Environment* 2002; **82**:431–445.
- [9] Chang C-W, Laird DA. Near-infrared reflectance spectroscopic analysis of soil C and N. *Soil Science* 2002; **167**(2):110–116.
- [10] McCarty GW, Reeves III, JB, Reeves VB, Follett RF, Kimble JM.. Mid-infrared and near-infrared diffuse reflectance spectroscopy for soil carbon measurements. *Soil Science Society of America Journal* 2002; **66**:640–646.
- [11] Martens H, Næs T. *Multivariate Calibration*. John Wiley and Sons, Chichester, 1989 p. 419.
- [12] Leone AP, Calabrò G, Coppola E, Maffei C, Menenti M, Tosca M, Vella M, Buondonno A. Prediction of soil properties with VIS-NIR-SWIR reflectance spectroscopy and artificial neural networks. A case study. *Advances in GeoEcology* 2008; **39**:689-702.
- [13] Leone AP, Viscarra-Rossel RA, Buondonno A., *Prediction of Soil Properties with PLSR and vis-NIR Spectroscopy: Application to Mediterranean Soils from Southern Italy*. *Current Analytical Chemistry*, 2012; **8**:283-299.
- [14] MIPAF, Ministero delle Politiche Agricole e Forestali, 2000. *Metodi di Analisi Chimica del Suolo*, Franco Angeli ed: Milano.
- [15] Chabrilat S, Goetz AFH, Krosley L, Olsen HW.. Use of hyperspectral images in the identification and mapping of expansive clay soils and the role of spatial resolution. *Remote Sensing of Environment* 2002; **82**:431-445.
- [16] Huang Z, Turner BJ, Dury S J, Wallis IR, Foley WJ. Estimating foliage nitrogen concentration from HYMAP data using continuum removal analysis. *Remote Sens. Environ.* 2004; **93**(1-2):18-29.
- [17] Gomez C, Lagacherie P, Coulouma G. Continuum removal versus PLSR method for clay and calcium carbonate content estimation from laboratory and airborne hyperspectral measurements. *Geoderma*, 2008; **148**:141–148.
- [18] Bonfils P. *Carte pédologique de France au 1/100000ème. feuille de Lodève*. INRASESCPF, 1993.
- [19] Hunt GR, Salisbury JW. Visible and near-infrared spectra of minerals and rocks. I. Silicate minerals. *Modern Geology* 1970; **1**(4): 283–300.
- [20] Bishop JL, Pieters CM, Edwards JO. Infrared spectroscopic analyses on the nature of water in montmorillonite. *Clay and Clay Minerals* 1994; **42**:702–716.
- [21] Sorensen LK, Dalgaard S. Determination of clay and other soil properties by near infrared spectroscopy. *Soil Science Society of America Journal* 2005; **9**:159–167.
- [22] McCarty GW, Reeves III, JB, Reeves VB, Follett RF, Kimble JM.. Mid-infrared and near-infrared diffuse reflectance spectroscopy for soil carbon measurements. *Soil Science Society of America Journal*, 2002; **66**:640–646.
- [23] Ehsani MR, Upadhyaya SK, Slaughter D. *A NIR technique for rapid determination of soil mineral nitrogen*. *Precis. Agri.*, 1999; **1**:217-234.
- [24] Esbensen KH. *Multivariate data analysis in practice*. Corvallis: CAMO, 598 pp. 2000.
- [25] Nash JE, Sutcliffe JV. River flow forecasting through conceptual models part I – A discussion of principles. *Journal of Hydrology* 1970; **10**(3):282-290.
- [26] CAMO Inc. *The Unscrambler user manual*. CAMO Inc, Corvallis, OR. 1998.