

Article

Reducing the Influence of Soil Moisture on the Estimation of Clay from Hyperspectral Data: A Case Study Using Simulated PRISMA Data

Fabio Castaldi ¹, Angelo Palombo ², Simone Pascucci ², Stefano Pignatti ^{2,*}, Federico Santini ² and Raffaele Casa ¹

¹ Department of Agricultural and Forestry Sciences (DAFNE), Università degli Studi della Tuscia (DPV), Via San Camillo de Lellis, 01100 Viterbo, Italy; E-Mails: castaldi@unitus.it (F.C.); rcasa@unitus.it (R.C.)

² Institute of Methodologies for Environmental Analysis (IMAA), Consiglio Nazionale delle Ricerche (CNR), C.da S. Loja, 100, 85050 Tito, Italy; E-Mails: angelo.palombo@imaa.cnr.it (A.P.); simone.pascucci@imaa.cnr.it (S.P.); federico.santini@imaa.cnr.it (F.S.)

* Author to whom correspondence should be addressed; E-Mail: stefano.pignatti@cnr.it; Tel.: +39-06-4994-4023; Fax: +39-06-4993-4211.

Academic Editors: Eyal Ben-Dor, Lenio Soares Galvao and Prasad S. Thenkabail

Received: 14 September 2015 / Accepted: 12 November 2015 / Published: 19 November 2015

Abstract: Soil moisture hampers the estimation of soil variables such as clay content from remote and proximal sensing data, reducing the strength of the relevant spectral absorption features. In the present study, two different strategies have been evaluated for their ability to minimize the influence of soil moisture on clay estimation by using soil spectra acquired in a laboratory and by simulating satellite hyperspectral data. Simulated satellite data were obtained according to the spectral characteristics of the forthcoming hyperspectral imager on board of the Italian PRISMA satellite mission. The soil datasets were split into four groups according to the water content. For each soil moisture level a prediction model was applied, using either spectral indices or partial least squares regression (PLSR). Prediction models were either specifically developed for the soil moisture level or calibrated using synthetically dry soil spectra, generated from wet soil data. Synthetically dry spectra were obtained using a new technique based on the effects caused by soil moisture on the optical spectrum from 400 to 2400 nm. The estimation of soil clay content, when using different prediction models according to soil moisture, was slightly more accurate as compared to the use of synthetically dry soil spectra, both

employing clay indices and PLSR models. The results obtained in this study demonstrate that the *a priori* knowledge of the soil moisture class can reduce the error of clay estimation when using hyperspectral remote sensing data, such as those that will be provided by the PRISMA satellite mission in the near future.

Keywords: clay indices; remote sensing; proximal sensing; PLSR; satellite; spectral features

1. Introduction

Information on the variability of soil properties leads to an increased ability to monitor key physical and chemical processes, with agronomic and environmental management implications. In particular, soil texture is a key agro-ecosystem variable whose estimation is essential for a full knowledge of the potential of soils in terms of fertility and water holding capacity. For instance, the knowledge of soil texture variability is central for the implementation of site-specific farming management strategies that allow for a more efficient use of resources, such as water and fertilizers; therefore reducing costs and environmental impact.

Soil spectroscopy in the visible-near infrared (VNIR: 400–1300 nm) and short-wave infrared (SWIR: 1300–2500 nm) domains has proven to be a non-destructive, reproducible, and accurate analytical technique for the quantitative estimation of soil variables [1]. The spectral composition of energy reflected or emitted from the topsoil depends on the organic and mineral soil composition, the soil moisture content, and the optical-geometric scattering. The most important factors influencing the scattering effect are both the particle size and roughness. In the soil spectroscopy context, some soil variables have a direct spectral response, e.g., having typical absorption features related to chemical components [2]. The main chemical components, or chromophores, interacting with visible and infrared radiations are water, organic matter, clay minerals, iron oxides, carbonates, and salts. These are related to well-defined absorption features [2], typically associated with overtones of functional groups in the VNIR and especially in the SWIR spectral ranges. In particular, free water absorption features are detected at 1455 and 1915 nm that are caused by overtones of O-H and H-O-H stretch vibrations, while metal-OH bends and O-H stretch related to clay lattice both affect the absorption peaks centered at about 1415 and 2207 nm. Hyperspectral imagers allow the measurement of spectral reflectance for hundreds of narrow bands, thus being an attractive tool that can be used for estimating soil variables. The fine spectral resolution of hyperspectral imagers allows for distinguishing the spectral features related to the chromophores. Published research in which hyperspectral satellite data has been used to estimate soil variables is, however, still very rare since currently only the Hyperion satellite hyperspectral imager on board of NASA EO-1 is available for soil studies. Four satellites carrying hyperspectral imagers, operating in the VNIR-SWIR spectral range, are due to be launched in the near future: the German EnMap (Environmental Mapping and Analysis Program) [3], the Japanese HISUI (Hyperspectral Imager Suite) [4], the Italian PRISMA (PRecursore IperSpettrale della Missione Applicativa) [5], and the U.S. NASA HypSIIRI (Hyperspectral Infrared Imager) [6]. In particular, PRISMA, due to be launched in 2018, is a mission of the Italian Space Agency (ASI), which combines

a hyperspectral imager with a panchromatic, medium-resolution camera that will have the capability to collect 240 spectral bands in the wavelength range 400–2450 nm (VNIR-SWIR spectral ranges).

The use of hyperspectral remote sensing data for the estimation of soil variables is hampered by the need to acquire data on bare soil conditions and by the characteristics of existing satellite sensors, *i.e.*, the spectral and spatial resolution and the signal quality in the spectral regions affected by absorption peaks of the main soil variables [7]. Moreover, surface spectral data are also generally affected by the confounding effects of soil moisture and soil roughness. Soil moisture determines a reduction of the reflectance over the entire spectrum, because it has a strong influence on the amount and composition of reflected and emitted energy from the soil surface. This decrease in reflectance related to soil moisture is not linear and its magnitude varies depending on the spectral region and the soil type [7]. In particular, the strength of the most important absorption features related to the estimation of soil variables, such as clay, are greatly reduced with increasing levels of moisture [7–10]. For these reasons, soil moisture could be a limiting factor for the estimation of soil properties if its effects are not taken into account in the estimation models. Therefore, the *a priori* knowledge of soil moisture content, at the time of remote sensing imagery acquisition, could improve its estimation accuracy [9]. Many researchers have studied the influence of soil moisture on soil reflectance, but only a few have actually investigated the consequences of this factor on soil variable estimation, and particularly on clay content prediction [11–13]. In many cases, studies were only carried out in laboratory conditions [7,9,14–16]. The first derivative of soil spectra was used to understand the impact of variable soil moisture on soil organic carbon (SOC) prediction [10]. Specific prediction models were developed on soil samples having similar soil moisture contents, thereby obtaining a consistent improvement on SOC estimation accuracy [9]. This suggests that the *a priori* knowledge of soil moisture could improve the estimation accuracy of soil variables. Alternatively, the reduction of moisture effects was successfully obtained in spectra acquired in field conditions by [17] using a direct standardization (DS) of the spectra, using a technique called external parameter orthogonalization (EPO). This resulted in an improved prediction of clay and organic matter content from soils with unknown soil moisture content [12,13].

Satellite hyperspectral imagers have been employed to estimate soil moisture, in bare soil conditions, by using narrow band indices [18]. However, the low signal to noise ratio (SNR) in the SWIR band of the hyperspectral satellite sensors currently in operation (e.g., Hyperion) hampers the accuracy of the estimation. Therefore, the issue of soil moisture influence on clay estimation has been explored in this work, through the use of simulated satellite data.

Multivariate or hybrid techniques, such as PLSR (partial least squares regression), principal component regression, or regression-kriging, generally provide accurate estimations [19–21]. However, these techniques require extensive ground data, laboratory analyses, and spectral measurements and provide only local models, which cannot be applied to different conditions or areas. The different sensor characteristics (e.g., bandwidths, number of bands, wavelength position) make it necessary to perform a new calibration for each sensor [22] or even images acquired under different illumination and geometry of observation. Instead, the development of spectral indices using soil spectra collected in laboratory conditions would allow for obtaining more general models, albeit less accurate than multivariate models [22].

The objective of this work was to derive narrow band clay indices for the estimation of soil moisture and clay, developing a methodology to reduce the effect of soil moisture on clay estimation.

As a reference, results obtained with spectral clay indices were compared to those from PLSR models. To our knowledge, no previous research has investigated the possibility of estimating clay using narrow band spectral indices from realistically simulated satellite spectra.

Since the prediction of soil clay content from satellite images could be significantly improved by the estimation of soil moisture at acquisition time, we developed two different approaches to reduce the effects of soil moisture. The first approach is based on the calibration of clay indices selected according to different soil moisture levels, whereas in the second approach, equivalent dry soil spectra (hereafter referred to as synthetically dried) are reconstructed from the spectra of wet soil samples and then used in clay estimation models. Both approaches were developed and tested, using spectral datasets acquired under laboratory conditions. In order to verify the feasibility of the two approaches for hyperspectral satellite data, both methodologies were also applied to simulated PRISMA data.

2. Materials and Methods

2.1. Soil Spectral Datasets

Datasets of spectral reflectances, acquired in the laboratory from an extensive set of soils, were used for the development of the estimation methods.

The first dataset (hereafter referred to as MAC), used for soil clay content estimation, includes spectra from 72 soil samples that were collected in the Maccarese farm (lat. 41°52'18"N, long. 12°14'05"E, alt. 8 m a.s.l.), near Rome (Central Italy), from the cultivated layer (0–30 cm soil depth) [23]. The soils in the area are Luvisols [24], which were formed from sediments of ancient coastal terraces from Pleistocene marine deposits and are the result of land reclamation works performed in the 1920s. For each sample we measured clay, sand, and silt contents (%) using the pipette method according to the USDA (United States Department of Agriculture) system [25]. The reproducibility of the pipette method for clay is assumed to have an error of 1% [26]. The clay content of the soil samples of MAC dataset varies between 22% and 56%. The mineralogical composition of the MAC soil samples includes quartz, feldspars, and phyllosilicates. Salts and hematite are also present. The phyllosilicates are represented by smectite, illite, and kaolinite.

The second dataset (hereafter referred to as SOLREFLIU [27]) is composed of 89 soil samples collected at a depth of 30 cm, coming from different sites in France and China. The clay content of the SOLREFLIU dataset varies between 0.2 and 66.8%. The mean calcium carbonate content is 5%. This dataset was only used to develop the procedures to reduce the effect of soil moisture on clay estimation (see Sections 2.2 and 2.3).

For both datasets, spectral radiance was acquired on each sample using an Analytical Spectral Device (ASD Inc., Boulder, CO, USA) Field Spec Pro spectroradiometer. The ASD spectroradiometer covers a spectral range of 350–2500 nm and was equipped with the ASD high-intensity contact probe with a 6.5 W quartz halogen lamp and a spot size of reflectance measurement of about 10 mm. The contact probe has a fixed spatial relationship between the detector fiber optic cable and the light source, allowing the measurement of a directional radiance value. The soil spectral radiances were acquired between 350 nm and 2500 nm with the contact probe using an integration time of 50 ms and a spectral resolution of 1 nm and then converted into absolute reflectance by using a NIST calibrated panel (Spectralon—reference

standard, S/N 5221—Sphere Optics Inc., Durham, NH, USA) to derive absolute reflectance spectra. The spectra were acquired by maintaining the same viewing geometry (nadir view) to minimize directional effects resulting from properties of the target with a directional set-up. Data at wavelengths less than 400 nm were discarded since they were too noisy.

The soil samples of both MAC and SOLREFLIU datasets had been air-dried and passed through a 2 mm sieve. They were then put individually into Petri dishes (little cylinders 8.7 cm in diameter, 1.6 cm in height) forming a layer considered as optically infinitely thick. Water was slowly poured down from the side of the box to reach full saturation. When free water disappeared from the soil surface (*i.e.*, 24 hours after wetting), the reflectance was measured. Measurements were performed four times during the drying process. In order to reduce the differences between the surface and the rest of the sample, Petri dishes were covered for two hours before the acquisition of spectra. The soil moisture content was then measured by weighing the samples after each ASD spectral acquisition and soil moisture was expressed as:

$$SM = \frac{W_w - W_d}{W_d} \quad (1)$$

Where W_w and W_d are, respectively, the weight of soil samples at wet and oven dry conditions. The main characteristics of the spectral datasets we used are reported in Table 1.

Table 1. Main characteristics of the MAC and SOLREFLIU spectral libraries.

	Spectral dataset	
	SOLREFLIU	MAC
N °samples	89	72
Type	Topsoil	Topsoil
Sampling area	France and China	Italy
Spectral measurement	Directional	Directional
Spectral range	350–2500	350–2500
Spectral resolution	1 nm	1 nm

2.2. Narrow Band Soil Moisture Indices—SMIR

As a preliminary step to the application of different clay estimation methods that take into account soil moisture, spectra from both MAC and SOLREFLIU datasets were used to develop reflectance indices related to the soil moisture content. Inspection of Pearson's correlation coefficient (r) between soil moisture and reflectance values for each wavelength within the range 400–2400 nm, allowed the identification of the spectral regions most affected by soil water content. These spectral regions were used to calibrate soil moisture indices, suitable for a satellite remote sensing context (hereafter referred to as SMIR), *i.e.*, avoiding those regions that would be influenced by atmospheric water vapor absorption, from normalized (Equation (2)) or simple ratios (Equation (3)), which could be subsequently applied to hyperspectral satellite data:

$$SMIR_A = \frac{R_{w1} - R_{w2}}{R_{w1} + R_{w2}} \quad (2)$$

$$SMIR_B = \frac{R_{w1}}{R_{w2}} \quad (3)$$

where R_{w1} and R_{w2} are the reflectance values at the wavelengths most correlated with soil moisture. In order to obtain more realistic relationships, we excluded from the data soil moisture values higher than the field capacity, since these are rarely found in agricultural soils. Linear or quadratic regression models were used to describe the relationship between each SMIR and the actual soil moisture values.

In addition to these indices, an index proposed in the literature, the Normalized Soil Moisture Index (NSMI) [18] was also tested on the same soil datasets used in this paper (MAC plus SOLREFLIU).

The models' accuracies were evaluated by examining the Root Mean Square Error (RMSE) and the Ratio of the Performance to the Interquartile Range (RPIQ) of the leave-one-out cross-validation. The equations used were:

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (x_o - x_p)^2}{n}} \quad (4)$$

$$RPIQ = \frac{IQ}{RMSE} \quad (5)$$

where, x_o and x_p are respectively the observed and predicted values, n the number of data pairs and IQ the inter-quartile range [28], *i.e.*, the difference between the values below which we can find 75% (Q3) and 25% (Q1) of the samples ($IQ = Q3 - Q1$). For all estimation models, the bias was also computed and its statistical significance was assessed using a Student t-test (significance: 95%).

2.3. Reducing the Effect of Soil Moisture from Laboratory Soil Spectra

Two different strategies (Figure 1) have been explored to reduce the influence of soil moisture on the estimation of clay. They are:

- (a) application of different prediction models according to the different soil moisture contents;
- (b) use of prediction models with synthetically dried soil spectra derived from wet soil spectra.

Both these approaches require the *a priori* knowledge of soil moisture content (see Section 2.2). Samples of MAC and SOLREFLIU datasets were divided into four subsets (*i.e.*, four soil moisture classes for each dataset) according to the gravimetric soil moisture content, and named as *dry* (D; soil moisture < 0.06), *little wet* (LW; 0.06 < soil moisture < 0.16), *wet* (W; 0.16 < soil moisture < 0.26), and *very wet* (VW; 0.26 < soil moisture < 0.36). Soil moisture values greater than 0.36 were excluded, because they corresponded to soil water content values higher than the field capacity of most agricultural soils [29]. We tested the two estimation approaches mentioned above on full spectra (limiting the spectral range from 400 to 2400 nm with 1 nm of spectral resolution) and on simulated PRISMA data, by using either narrow band clay indices or the PLSR technique (Figure 1).

2.3.1. Use of Different Prediction Models according to Soil Moisture Content

The entire MAC spectral dataset was first converted from reflectance to band depth values. In fact, the normalization of the spectra into their continuum and the subsequent transformation into band depth highlights the spectral features and makes a comparison possible among spectra acquired in different measurement conditions. Band depth values were obtained using continuum removal [30] to normalize the spectra over the entire spectral range (Band depth = 1 – continuum removed data). The correlation between band depth and clay for all wavelengths was then inspected, and the spectral

ranges having the highest Pearson's correlation coefficients were identified. Band depths valued at the most promising spectral wavelengths were then used to obtain normalized or simple ratio indices, according to Equation (2) or Equation (3). This process was carried out for each of the soil moisture classes in order to obtain different clay indices according to the soil moisture level. The results obtained from the indexes were then compared with those obtained from PLSR.

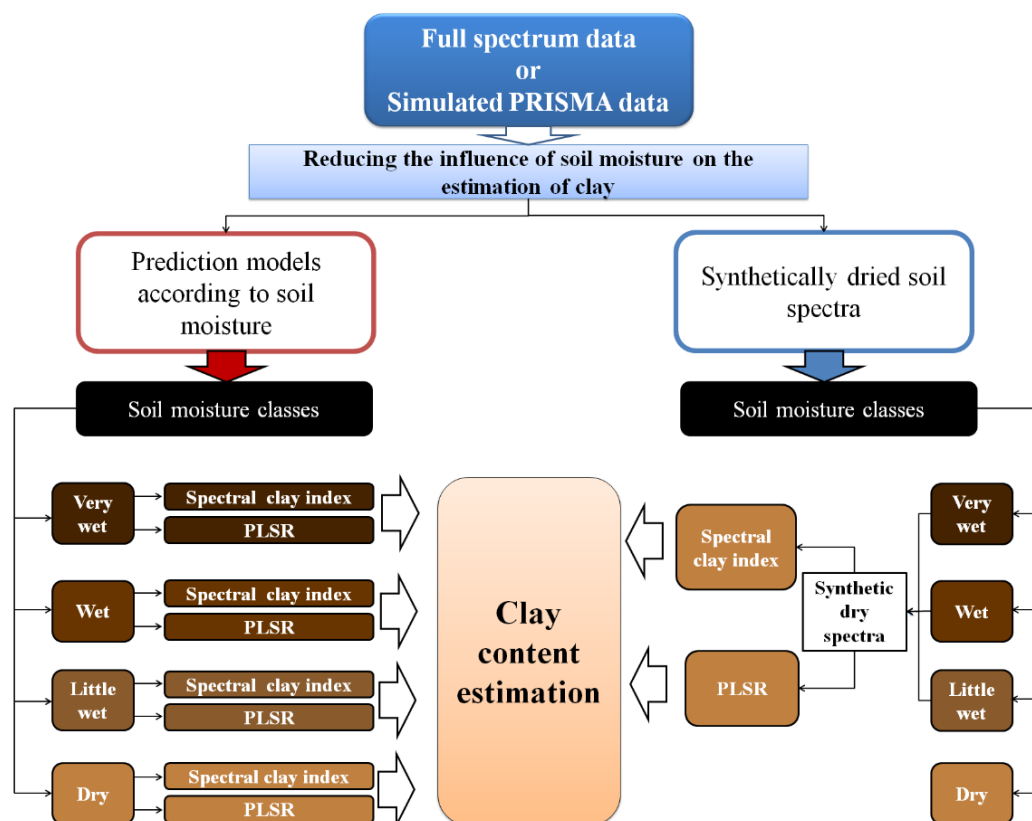


Figure 1. Flow chart concerning the estimation of clay estimation using different prediction models according to soil moisture content or using reconstructed (synthetically) dry soil spectra.

PLSR is a multivariate statistical technique [31] that allows the prediction of a dependent variable from a large number of predictors, by selecting a restricted number of latent variables having the strongest relationship to the dependent variable (clay). To obtain reliable models, we selected the number of latent variables, which explained at least 90% of the predictor variance.

The performance of the model of each soil moisture class was assessed using a leave-one-out cross-validation, and the estimation accuracy was evaluated by using RMSE and RPIQ. Only cross-validation results for the best indices have been reported.

2.3.2. Use of Prediction Models Calibrated with Synthetically Dried Soil Spectra

In addition to the previous method, the use of synthetically dried soil spectra obtained from wet soil spectra (Figure 1) was attempted, in order to calibrate clay prediction models less sensitive to the influence of soil moisture.

For this purpose, we computed the difference for each sample of MAC and SOLREFLIU datasets between the band depth values of a dry spectrum and the band depth values of the corresponding wet spectra (LW, W, and VW), thus obtaining three band depth differences. These band depth differences, for each moisture class i , were then averaged over all the samples of the dataset to obtain a mean difference \bar{x}_i and its standard deviation σ_i . Then, randomly extracted values from a normal distribution having \bar{x}_i as the mean and $\frac{\sigma_i}{3}$ as the standard deviation were added to the band depth spectra of the wet (LW, W, and VW) samples (BD_{wet}). We used one-third of the standard deviation to avoid extreme values from the distribution tails. This procedure allowed us to obtain synthetically dried band depth (BD) soil spectra (BD_{syn_dry}),

$$BD_{syn_dry} = BD_{wet} + N(\bar{x}_i, \frac{\sigma_i}{3}) \quad (6)$$

The second term of the equation indicates the drawing from a normal distribution with mean \bar{x}_i and standard deviation $\frac{\sigma_i}{3}$. The SOLREFLIU dataset includes spectra of different soil types having a larger spectral variability than the MAC dataset. For this reason, in order to test the effective generalization potential of this technique, we obtained the synthetically dried spectra for MAC adding the correction factors (second term of Equation (6)) obtained from either the SOLREFLIU or MAC datasets. The Savitzky-Golay smoothing filter was then applied (filter order: 3; filter length: 21) by using the *signal* package (OCTAVE-Forge project <http://octave.sf.net>) in the R software [32].

The indices and the PLSR models were calibrated on dry soil samples and were then applied to naturally dry and synthetically dried laboratory band depth spectra. The mean RMSE and RPIQ values (obtained by leave-one-out cross-validation) of 100 iterations were used to compare the estimation accuracy of clay indices and PLSR models.

2.4. PRISMA Simulated Data

In order to obtain PRISMA simulated data, we took into account the wavelengths (band center), the full width at half maximum (FWHM), and the estimated Noise Equivalent Delta-Radiance (NEDR) of PRISMA's satellite imager (as provided by the manufacturer, see [5]). The PRISMA payload design is based on a pushbroom type observation concept providing hyperspectral imagery (~250 bands) at a spatial resolution of 30 m on a swath of 30 km. The PRISMA instrument is based on a prism spectrometer concept and consists of the Hyp/Pan, optical head. The laboratory spectral reflectance acquired at all soil moisture levels and the synthetically dried soil spectra of MAC dataset (expressed as reflectance) were resampled according to PRISMA wavelengths and FWHM. The at-sensor radiance was calculated by applying the inverse equation of the atmospheric correction as expressed in Equation (7).

$$Rad(\lambda) = \frac{\rho(\lambda)\tau_{\uparrow}(\lambda)[E_0(\lambda)\cos(\theta_z)\tau_{\downarrow}(\lambda) + E_{\downarrow}(\lambda)] + \pi L_{\uparrow}(\lambda)}{\pi} \quad (7)$$

where, $Rad(\lambda)$ is at-sensor radiance, $\rho(\lambda)$ is the reflectance of soil samples, $\tau_{\uparrow}(\lambda)$ is the atmospheric transmittance from the ground to the sensor, $E_0(\lambda)$ is the solar spectral irradiance at the top of the atmosphere (TOA), θ_z is the solar zenith angle, $\tau_{\downarrow}(\lambda)$ is the atmospheric transmittance from TOA to the ground, $E_{\downarrow}(\lambda)$ is the at-surface spectral diffuse irradiance, and $L_{\uparrow}(\lambda)$ is the upwelling radiance in the

sensor's field of view (*i.e.*, the path radiance). We then added to the $Rad(\lambda)$ a Gaussian noise having zero mean and a standard deviation corresponding to the expected PRISMA NEDR (Equation (8)).

$$Rad(\lambda)_{noisy} = Rad(\lambda) + N(0, NEDR) \quad (8)$$

Finally, in order to re-obtain the PRISMA ground reflectance $\rho(\lambda)_{PRISMA}$, we applied the direct equation used for the atmospheric correction (Equation (9)).

$$\rho(\lambda)_{PRISMA} = \frac{\pi[Rad(\lambda)_{noisy} - \pi L_{\uparrow}(\lambda)]}{\tau_{\uparrow}(\lambda)[E_0(\lambda)\cos(\theta_z)\tau_{\downarrow}(\lambda) + E_{\downarrow}(\lambda)]} \quad (9)$$

The $\tau_{\uparrow}(\lambda)$, $E_0(\lambda)$, $\tau_{\downarrow}(\lambda)$, $E_{\downarrow}(\lambda)$ and $L_{\uparrow}(\lambda)$ in both Equations (5) and (7) were simulated using the MODTRAN radiative transfer code [33] by varying the input model parameters for summer, autumn, and winter conditions. This is expected to influence the SNR of the data according to the solar zenith angle. These simulated PRISMA data were used to test the predictive models of clay, both applying spectral indices and PLSR models.

3. Results and Discussion

3.1. Soil Moisture Estimation

3.1.1. Laboratory Dataset

The correlation between gravimetric soil moisture content of the MAC and SOLREFLIU samples and reflectance values highlighted the wavelengths most affected by water content (Figure 2). We detected four wavelengths in the NIR-SWIR region having the highest r value and that are potentially suitable for satellite data applications, *i.e.*, less affected by atmospheric effects as compared to the typical water absorption bands centered at 1455 and 1915 nm. We tested these wavelengths for the development of normalized or simple ratio soil moisture indices.

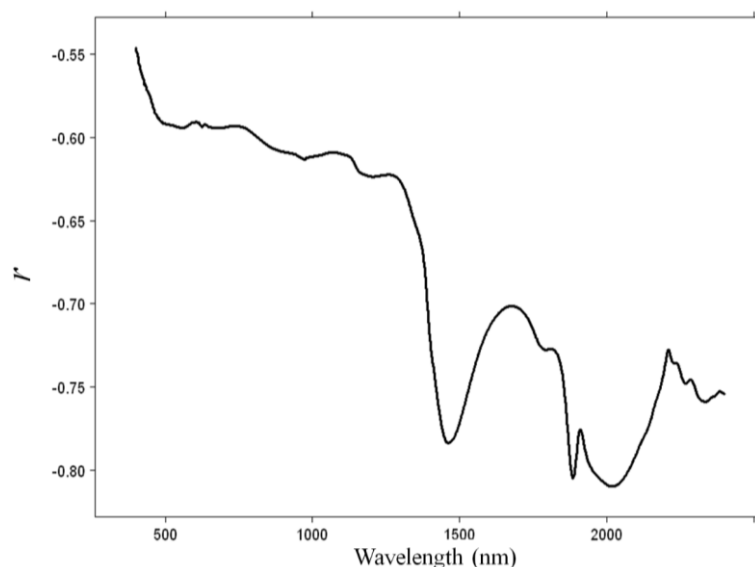


Figure 2. Pearson's correlation coefficient (r) between soil moisture content and reflectance values of MAC and SOLREFLIU spectral datasets as a function of the wavelength.

Table 2 shows the results for the two best normalized or simple ratio indices (Equations (2) and (3)), named as SMIR_A and SMIR_B, providing the highest r value, alongside results using the NSMI index [18]. The SMIR_A index shows a strong positive correlation ($r = 0.89$), while SMIR_B shows a strong negative correlation ($r = -0.88$) with soil moisture. The quadratic regression model between SMIR_A and soil moisture provided an RMSE = 0.05, and similar statistics were found for SMIR_B (RMSE = 0.05). The plots of measured and predicted (by regression models) soil moisture values are shown in Figure 3. NSMI index, used as reference, shows a strong correlation with soil moisture, with similar performance as SMIR_B and SMIR_A. In fact, the quadratic regression models between NSMI and soil moisture provided the same RMSE value of SMIR_A and SMIR_B. The estimation of soil moisture by these indices was not affected by bias, since it was not significantly different from zero for all indices. The wavelengths involved in SMIR_A and NSMI are very similar. These two normalized indices use wavelengths close to the absorption bands of water and, in particular, those corresponding to the water absorption peaks' shoulders (1770–1800 nm and 2100–2119 nm). The SMIR_B index utilizes the ratio between the right shoulder of the water peak centered at 1455 nm and the left shoulder of the water peak centered at 1915 nm.

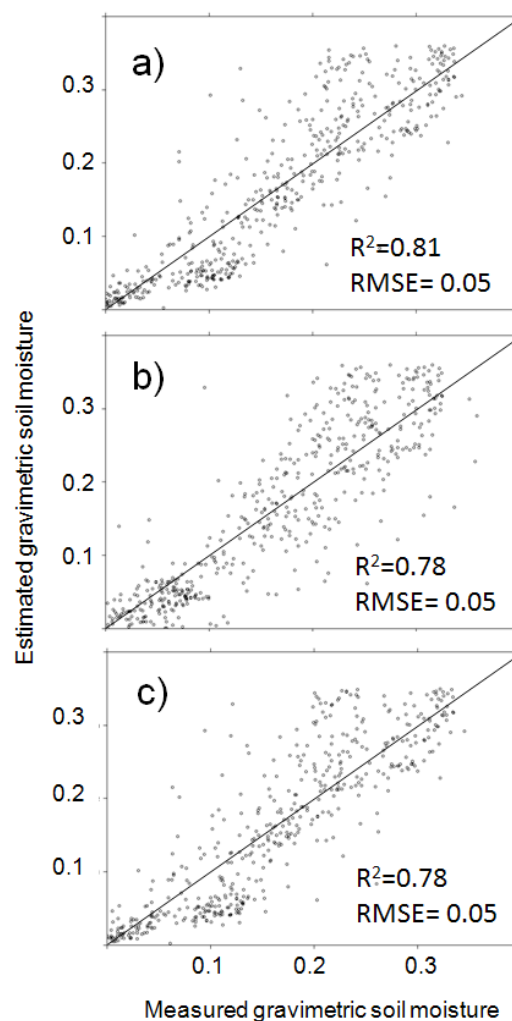


Figure 3. Plots of measured vs. estimated gravimetric soil moisture content by using (a) SMIR_A, (b) SMIR_B and (c) NSMI narrow band indices. In the graphs the 1:1 lines are reported.

Table 2. Cross-validation results of novel soil moisture indices (SMIR_A and SMIR_B) and the normalized soil moisture index (NSMI), using data from MAC+SOLREFLIU datasets [13]. The equations of the soil moisture indices, Pearson's correlation coefficient (r), the coefficients of the quadratic regressions (Soil moisture = $a + b \cdot \text{index} + c \cdot \text{index}^2$), the root mean square error (RMSE) and the ratio of the performance to the interquartile range (RPIQ) are reported.

Soil Moisture Index	Formula	r	a	b	c	RMSE (%)	RPIQ
SMIR_A	$(R_{1770} - R_{2100}) / (R_{1770} + R_{2100})$	0.89	0.03	1.63	-1.89	0.05	4.25
SMIR_B	R_{1506} / R_{1770}	-0.88	0.48	0.24	-0.75	0.05	4.25
NSMI	$(R_{1800} - R_{2119}) / (R_{1800} + R_{2119})$	0.88	0.50	-1.67	1.46	0.05	4.25

3.1.2. Simulated PRISMA Data

The application of the indices reported in Table 2 to simulated PRISMA data provided very similar results (Table 3) to those obtained from full spectra data (Table 2) for all simulated atmospheric conditions. SMIR_A and NSMI indices seemed not to suffer from the lower spectral resolution and the addition of atmospheric effects, showing identical statistics as those of Table 2 for these scenarios. The results obtained with the SMIR_B were only slightly worse than those obtained using the full spectra. No bias was detected for the soil moisture prediction models using simulated PRISMA data.

Table 3. Results from the application of SMIR and NSMI indices for the estimation of soil moisture from simulated PRISMA data, with the addition of simulated atmospheric effects for different seasonal scenarios.

Atmosphere	Statistics	SMIR_A	SMIR_B	NSMI
Summer	r	0.89	-0.81	0.88
	R^2	0.8	0.67	0.78
	RMSE (%)	0.05	0.06	0.05
	RPIQ	4.25	3.54	4.25
Autumn	r	0.89	-0.86	0.88
	R^2	0.8	0.74	0.78
	RMSE (%)	0.05	0.05	0.05
	RPIQ	4.25	4.25	4.25
Winter	r	0.89	-0.86	0.88
	R^2	0.8	0.75	0.78
	RMSE (%)	0.05	0.06	0.05
	RPIQ	4.25	3.54	4.25

3.2. Clay Content Estimation Considering Soil Moisture

3.2.1. Clay Estimation from Full Spectra

Pearson's correlation coefficients between clay content and band depth were quite different among the different soil moisture classes of the MAC dataset. Although, for all soil moisture classes, high correlation coefficients were found between 2170 and 2270 nm and at 2360 nm, other correlation

peaks were detected at 530 nm for the LW class, at 1340 nm for the W class and at 1680 nm for the VW class (Figure 4). We used the four wavelengths showing the highest correlation between clay and band depth for each soil moisture class to derive spectral clay indices according to Equation (2) or Equation (3). The clay indices of VW, W, and D classes showed r values higher or equal to 0.7, while the r value of LW class was lower than 0.7 (Table 4). In the extreme soil moisture classes (VW and D), the RMSE values were lower than 6% (Table 5) and RPIQ was higher than 2.8, while for the other two classes (W and LW) RMSE was higher than 7% and RPIQ lower than 2.3. The mean RMSE value of the four soil moisture classes, *i.e.*, considering the different clay indices according to soil moisture content, was 6.7%. In the hypothesis that a spectral index for soil clay content estimation is obtained from dry soil samples, its application to moist soil data will entail a decrease of accuracy. When applying the D index to all the soil moisture classes, the mean RMSE was higher than 9% (Table 5). This shows that the *a priori* knowledge of the soil moisture class, for the selection the optimal clay index, decreases the RMSE by 27% (6.7% instead of 9.3%) increasing the accuracy of soil clay content estimation.

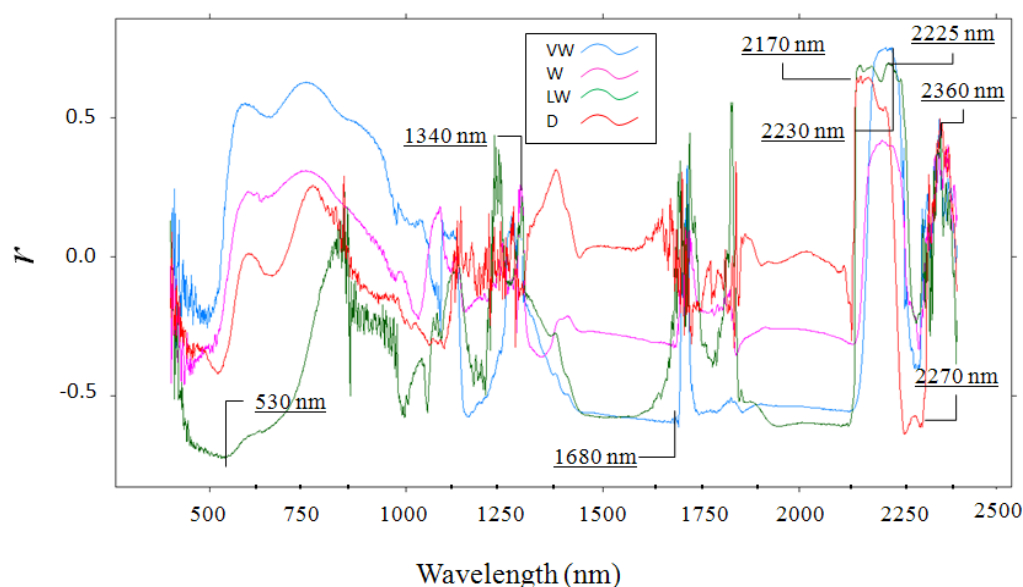


Figure 4. Pearson's correlation coefficient between clay content and band depth spectral values for the four soil moisture classes of the MAC dataset (VW: very wet; W: wet; LW: low wet; D: dry).

Table 4. Band Depth (BD) clay indices for classes having different soil moisture (SM) level (D = dry, LW = little wet; W = wet, VW = very wet) and Pearson's correlation coefficient (r) between indices and clay content.

SM class	Clay index	r
VW	$(BD_{2230} - BD_{1680}) / (BD_{2230} + BD_{1680})$	0.70
W	$(BD_{1340} - BD_{2360}) / (BD_{1340} + BD_{2360})$	-0.73
LW	BD_{530} / BD_{2225}	0.65
D	$(BD_{2170} - BD_{2270}) / (BD_{2170} + BD_{2270})$	0.75

The PLSR models developed for clay estimation showed a better accuracy than spectral indices for each soil moisture level. The best results were reached for the class D (Table 5). Again, when applying the PLSR model calibrated on dry samples to all soil moisture classes, we obtained very poor accuracy (RMSE = 15.7%; RPIQ = 1.07) and, only in this case, the clay estimation was biased (*i.e.*, the difference between measured and predicted values was significantly higher than zero).

Table 5. Results from full spectral data using two different approaches (see text) to reduce the influence of soil moisture (SM) on clay estimation with two different clay estimation techniques (spectral indices and PLSR).

Dataset	SM Reduction Method	SM Class	Clay Estimation Method			
			Spectral Indices		PLSR	
			RMSE (%)	RPIQ	RMSE (%)	RPIQ
MAC	Different spectral indices according to SM level ^(a)	VW	5.7	2.96	4.7	3.59
		W	7.4	2.28	6.4	2.64
		LW	7.8	2.16	5.0	3.37
		D	6.0	2.81	3.9	4.33
		Mean	6.7	2.52	5.0	3.37
		D model to all (mean value)	9.3	1.81	15.7	1.07
MAC synthetically dried using MAC ^(b)	Synthetically dried spectra ^(c)	VW	7.9	2.13	7.6	2.22
		W	6.5	2.58	7.9	2.14
		LW	6.2	2.73	6.2	2.72
		D	6.0	2.79	3.9	4.33
		Mean	6.7	2.52	6.4	2.64
MAC synthetically dried using SOLREFLIU ^(d)	Synthetically dried spectra	VW	9.0	1.87	8.0	2.11
		W	8.2	2.07	7.9	2.14
		LW	6.6	2.58	6.5	2.60
		D	6.0	2.79	4.7	3.59
		Mean	7.4	2.27	6.8	2.49

^(a) the indices used for each class are reported in Table 4; ^(b) equivalent dry band depth spectra obtained from wet samples using Equation (6) with the second term of Equation (6) calculated from the MAC dataset; ^(c) the spectral clay index used for one is reported in Table 4 for class D; ^(d) equivalent dry band depth spectra obtained from wet samples using Equation (6) with the second term of Equation (6) calculated from the SOLREFLIU dataset.

The typical absorption features related to the clay lattice of kaolinite and montmorillonite are located near 1400, 1900, and 2200 nm [34], whereas illite has also two absorption features near 2300 and 2400 nm (Figure 5). Since the clay spectral features near 1400 and 1900 nm correspond to the atmospheric water absorption features, they cannot be used in remote sensing. The investigation on the wavelengths most correlated with clay content at different moisture levels, showed that the most suitable spectral features are strongly affected by soil moisture. It is likely that water interacts with spectral features of other chromophores correlated to the clay amount in soils. The feature at 530 nm of the LW index is due to the ferrous and ferric iron oxides (hematite). The feature at 1340 nm in the W index corresponds to the left shoulder of the water absorption peak at 1400 nm, while the feature used in the VW index at 1680 nm matches the highest reflectance values on the left shoulders of the water

absorption peak at 1900 nm. The increasing amount of water in the soil samples masks the typical clay absorption features, especially in the medium soil moisture classes (W and LW). This might explain why the best estimation accuracies were reached using dry and very wet samples (Table 5).

The index for dry samples exploits two absorption features. The first is typical of kaolinite (2170 nm), the second (near 2200 nm) is characteristic of illite, montmorillonite, and kaolinite (Figure 5). A little addition of water (LW samples) brings about the masking of the first kaolinite feature, highlighting the iron oxides feature (near 530 nm). Concerning the index for W samples, the relation between clay and water content (1340 nm) and the absorption peaks of illite (2360 nm) was exploited. The indices for VW samples use the typical features of all clay minerals (near 2200 nm). Also the PLSR models showed the best estimation statistic for D and VW subsets [18] confirming an improvement of clay estimation accuracy using soil spectra having high soil moisture content as compared to spectra with lower moisture. Thus, both high and low soil moisture data seem to provide the best conditions for the estimation of clay from hyperspectral remote sensing.

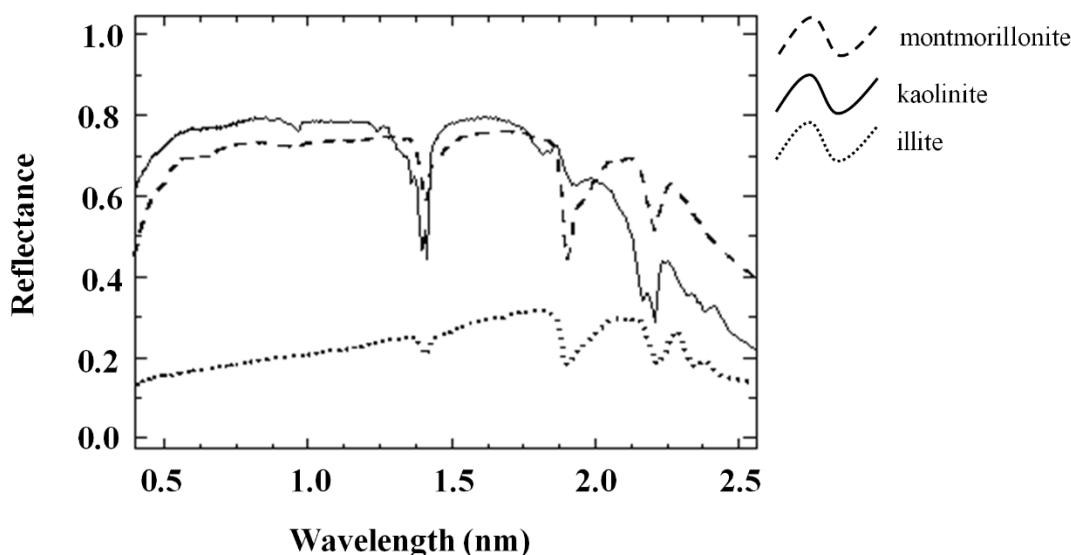


Figure 5. Typical spectra of montmorillonite, illite, and kaolinite obtained from the spectral library [24] (<http://speclab.cr.usgs.gov/spectral.lib04/spectral-lib04.html>).

We then tested the procedure of bringing back the wet spectra to a simulated dry state for clay estimation on the MAC dataset. The band depth values of the synthetically dried spectra did not overlap exactly with the band depth values of the actual dry samples over the entire spectrum (Figure 6b). There was, however, a good overlap in the SWIR spectral region, especially at wavelengths employed in the clay dry index (2170 and 2270 nm). In Figure 6, the two wavelengths that were used for the clay index are shown. When applying the clay index obtained for the dry class (Table 4, class D) on synthetically dried spectra, clay content was estimated with increasing RMSE and decreasing RPIQ values, as a function of the initial moisture content of soil samples (Table 5). Results obtained from LW and W samples provided, however, a RMSE and RPIQ very similar to D samples (Table 5). This methodology does not seem to be very accurate for soil spectra characterized by high moisture content (VW), since in this case the RMSE was quite high and RPIQ was rather low (Table 5). The mean RMSE and RPIQ values obtained using this procedure from all soil moisture classes were respectively

6.7% and 2.52, thus providing similar results as those obtained using the other approach, *i.e.*, different indices according to the soil moisture class. Better results were obtained using the synthetically dry spectra for developing PLSR models (Table 5). Similarly to the application of indices, the highest RPIQ values were reached for LW samples (RMSE=6.2%; RPIQ=2.72). The mean RPIQ values obtained in this work, both from indices and PLSR models (Table 5), are very similar to those reached by [12] using the EPO algorithm and PLSR.

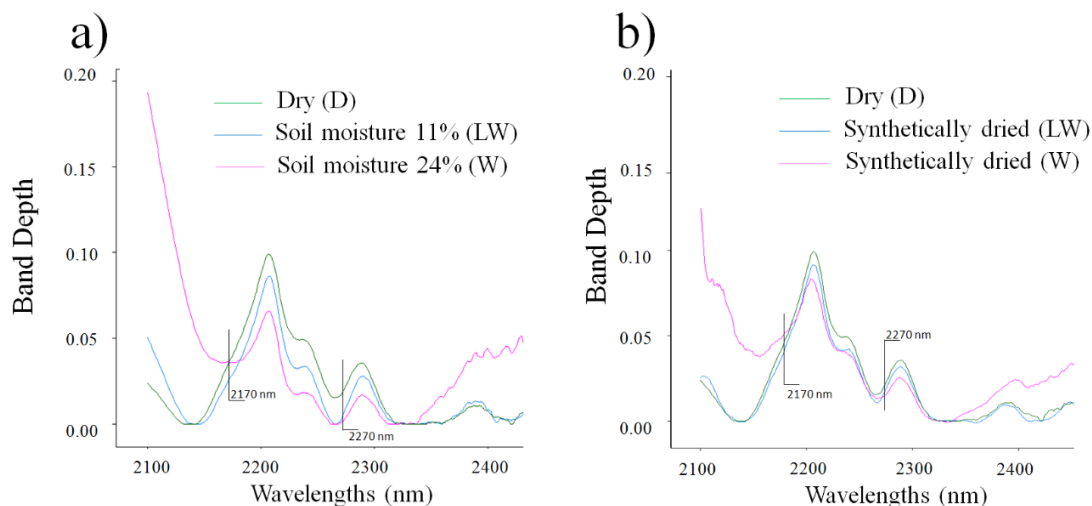


Figure 6. (a) Band depth spectra of a soil sample from the MAC dataset at different moisture levels; (b) band depth of the same sample as (a) in which wet soil data were "synthetically dried" using the procedure described in Section 2.3.2. The vertical bars represent the wavelengths used for the clay dry index (SM class D).

Although the SOLREFLIU dataset has fewer spectra than other regional or continental spectral datasets, such as LUCAS [35], it still includes spectra of different soil types having a much larger spectral variability than the MAC dataset. For this reason, in order to test the effective capability of the synthetically dry spectra technique, we obtained synthetically dried spectra by adding the correction parameters obtained from an extensive soil spectral dataset (SOLREFLIU) to a more local dataset (MAC). The purpose was to compare this procedure to the previous one in which correction parameters were obtained from the MAC dataset. In this case the RMSE and RPIQ mean values were slightly worse than those achieved using MAC data to obtain correction parameters, according to Equation (6), both by using indices or PLSR (Table 5).

Dry soil samples are generally used to test the capability of soil spectroscopy to estimate soil variables under laboratory conditions [2], but when working in field conditions using remote or proximal sensing, it is not possible to have dry soil nor uniform soil moisture content within the study area. Many researchers investigated the capability of the PLSR technique to estimate clay from the spectra of dry soil samples, showing statistics very similar to those obtained in this work (RMSE=3.9%; Table 5). For example, [36] and [19] obtained a RMSE of leave-one-out cross-validation of 3.1%, whilst [37] reported a RMSE of 3.8%. Although less accurate than PLSR, the results we reached using clay indices are quite good. In fact, the RMSE (6%) and RPIQ (2.8) are very similar to the statistics obtained in other works from PLSR or other techniques [38,39]. Although the

range of clay content values in the MAC dataset is quite large, the results have been obtained using only one soil group (LUVISOL). For this reason the estimation models shown in this work could mainly have a local validity. The capability of narrow band clay indices according to soil moisture needs to be further verified on different soil groups and geographical locations.

3.2.2. Clay Estimation from Simulated PRISMA Data

The clay indices applied to simulated PRISMA spectra provided less accurate results in terms of RMSE and RPIQ (Table 6) as compared to those obtained from full laboratory spectra (Table 4). The reduction of spectral resolution and the addition of noise and atmospheric effects lead to a slight decrease in accuracy of clay estimation using indices. Simulated PRISMA data provided, for autumn and summer by using the MODTRAN atmospheric radiative transfer model parameterization, a mean RMSE value of 7.3% and RPIQ of 2.3, whereas using the full spectra data, RMSE and RPIQ values were 6.8% and 2.5%, respectively. Additionally, these simulations support the advantage of the *a priori* knowledge of soil moisture. In fact, the results obtained by applying an index calibrated on dry soil data to all moisture classes showed very high RMSE values (Table 6) and in this case a significant bias was detected, which did not happen for estimation models according to moisture level. The results obtained by PLSR were statistically similar across seasons. The best performance for PLSR models, as well as for indices, was obtained by using the summer and autumn parameterizations. The estimation accuracy drops sharply when applying the PLSR model, calibrated on dry samples, to the whole dataset, as shown by RMSE values between 10.4% (winter) and 15.8% (summer). Additionally, for PLSR models, significant bias was detected only when the model for dry soils was applied to the whole dataset. PLSR models showed slightly better statistics as compared to those obtained when using indices, both for non-corrected spectra and synthetically dried spectra, however this difference could be irrelevant considering the error of 1% of the laboratory method to estimate clay [26]. The estimation of clay by using different prediction models (PLSR or indices) according to soil moisture, showed a slightly better accuracy than the prediction using synthetically dried spectra (Figure 7). The use of indices with Synthetically Dried Soil Spectra (SDSS) brings about a decrease of estimation accuracy, when moving from synthetically dried spectra obtained from soil samples having little soil moisture values to more wet samples, whereas using the PLSR technique, the differences between soil moisture classes are less noticeable. In this case, the difference between PLSR and indices in terms of RMSE was often higher than 1%, and thus higher than the error of reproducibility of the clay analysis method. The best results when using synthetically dry spectra were obtained for the LW class. The residuals of the prediction models are not correlated with the soil moisture content for both index and PLSR models, as shown in Figure 8. The lack of correlation between residuals and soil moisture was observed for both strategies used in this work (prediction model according moisture level and synthetic drying process). This confirms the robustness of these techniques to reduce the influence of soil moisture on clay estimation despite the non-linear effect of soil moisture on reflectance spectra [40].

The results obtained by using simulated PRISMA data cannot be compared to previous literature because of the lack of similar satellite hyperspectral remote sensors. Only Hyperion (on board of the EO-1 platform of NASA, USA) has a comparable spectral and spatial resolution, however, owing to the low SNR ratio of this sensor between 1900 and 2500 nm [41], the estimation of clay content is

strongly hampered. Using Hyperion data, [42] obtained similar RPIQ values to those shown in this paper (*i.e.*, 2.40) when employing spatial techniques (*i.e.*, linear mixed effect models), while when using PLSR, the RPIQ was drastically reduced to 1.6. Accordingly, the soil clay content estimation from simulated PRISMA data could be improved by using predictive models that exploit the spatial correlation between ground samples

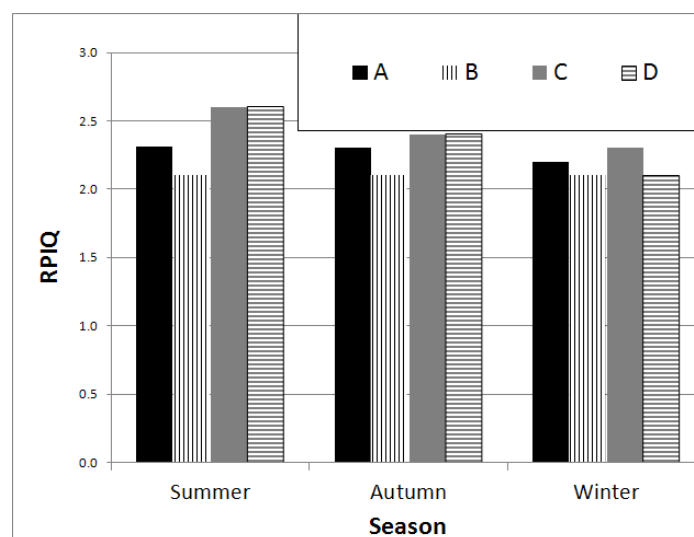


Figure 7. Bar plot of mean ratio of performance to inter-quartile range (RPIQ) values for clay estimation using simulated PRISMA data by means of different models according to soil moisture content both with clay indices (A) and PLSR (C), and by the use of synthetically dried spectra both with clay indices (B) and PLSR (D).

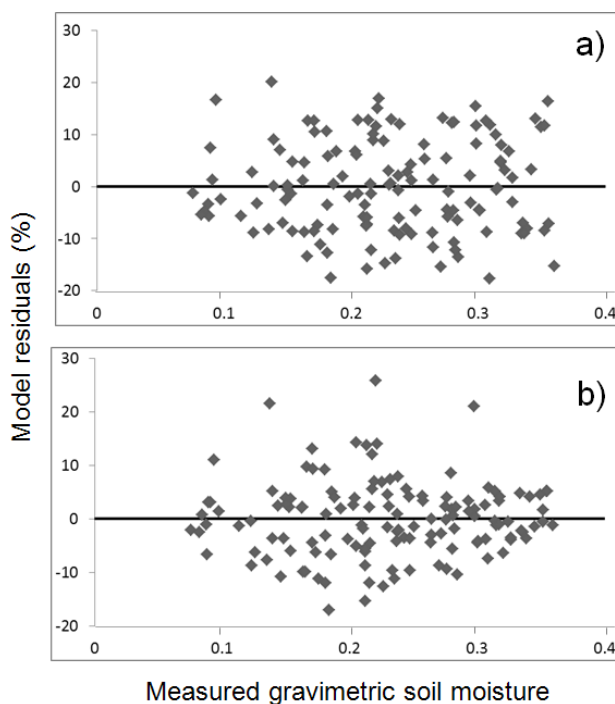


Figure 8. Plots measured gravimetric soil moisture *vs.* residuals of the estimation models using simulated PRISMA data with summer parameterization, both for (a) index and (b) PLSR method.

Table 6. Results from simulated PRISMA data using two different approaches for reducing the effect of soil moisture (SM) and two different clay estimation techniques (indices and PLSR). Atmospheric effect was added using seasonal parameterizations of the radiative transfer model MODTRAN.

SM reduction method	SM class	Clay estimation method											
		Spectral indices						PLSR					
		Summer		Autumn		Winter		Summer		Autumn		Winter	
		RMSE (%)	RPIQ	RMSE (%)	RPIQ	RMSE (%)	RPIQ	RMSE (%)	RPIQ	RMSE (%)	RPIQ	RMSE (%)	RPIQ
Different spectral indices according to SM level ^(a)	VW	6.6	2.6	6.3	2.7	6.3	2.7	7.1	2.4	6.7	2.5	7.4	2.3
	W	8.2	2.1	8.5	2	10.6	1.6	6.3	2.7	6.8	2.5	9.2	1.8
	LW	8.5	2	8.2	2.1	7.8	2.2	9.5	1.8	9.1	1.9	9.5	1.8
	D	5.9	2.9	6	2.8	5.9	2.9	4.7	3.6	5.8	2.9	7.3	2.3
	Mean	7.3	2.3	7.3	2.3	7.7	2.2	6.9	2.6	7.1	2.4	8.4	2.3
	D to all (mean)	13.2	1.3	13.3	1.3	11.9	1.4	15.8	1.1	15.6	1.1	10.4	1.6
Synthetically dried spectra ^(b)	VW	9.8	1.7	9.8	1.7	9.8	1.7	7.6	2.2	7.8	2.2	8.6	2
	W	9.2	1.8	9.2	1.8	9.2	1.8	7.4	2.3	7.2	2.3	8.5	2
	LW	7.7	2.2	7.7	2.2	7.7	2.2	7.1	2.4	7.2	2.3	7.7	2.2
	D	5.9	2.9	6	2.8	5.9	2.9	4.7	3.6	5.8	2.9	7.3	2.3
	Mean	8.2	2.1	8.2	2.1	8.2	2.1	6.7	2.6	7	2.4	8	2.1

^(a) the indices used for each class are reported in Table 4; ^(b) the spectral clay index used the one is reported in Table 4 for class D.

4. Conclusions

The results obtained in this study confirm that the *a priori* knowledge of soil moisture can significantly reduce the estimation error of clay using hyperspectral data. The use of narrow band clay indices according to soil moisture is a novel approach potentially of interest for proximal and remote sensing applications. In order to apply clay indices, precise quantitative information on soil moisture might not be required, since it will be sufficient to assign data to a rather broad soil moisture class: for instance we used classes with a 10% soil moisture range. Therefore, the use of clay indices appears to be very attractive when applied to hyperspectral data provided by satellite (remote sensing) or on-the-go sensors (proximal sensing).

The two approaches tested to reduce the influence of soil moisture, *i.e.*, the use of different prediction models according to soil moisture content or the use of models calibrated by using synthetically dried soil spectra data set, were both effective in improving clay content retrieval, both using narrow band clay indices and PLSR models. The estimation of clay when using different prediction models according to soil moisture was, however, slightly more accurate as compared to the use of synthetically dry spectra, both employing clay indices (mean RMSE = 6.7%; mean RPIQ = 2.52) and PLSR models (mean RMSE = 5%; mean RPIQ = 3.37). Even though the spectral clay indices showed lower estimation accuracy than PLSR, indices generally ensure a higher adaptability to different soil conditions, being based on the specific spectral absorption features of the clays.

The use of simulated PRISMA data brought about a decrease in soil clay content estimation accuracy as compared with the full spectral datasets data, both for indices (mean RMSE = 7.3%; mean RPIQ = 2.3) and PLSR (mean RMSE = 6.9%; mean RPIQ = 2.6) models. The *a priori* knowledge of soil moisture, for the selection the optimal clay index on the basis of the soil moisture class, decreased by 27% the RMSE values and increased the clay content estimation accuracy by using lab soil datasets.

The estimation accuracy of soil moisture content, as well as that of clay, is mainly affected by the SWIR spectral region and this is particularly evident around the water absorption features. Thus, in order to obtain acceptable estimation accuracies of soil variables from hyperspectral remote sensing data, high SNR values in the SWIR region are required.

Although the process that we used to obtain simulated PRISMA data was rigorous, taking into account the sensor signal to noise ratio and different atmospheric effects, the different techniques proposed here should be tested on real hyperspectral satellite data with similar performances in real field conditions, in which, for instance, soil roughness or the presence of crop residues would add further errors.

Acknowledgements

The research has been co-funded by the Italian Space Agency (ASI) with grant I/019/11/10 (PRISMA Mission) and by the Italian Ministry of Agricultural Policy with grant APREINF. The authors would like to thank Frédéric Baret (INRA, Avignon, France), who made the INRA SOILREFLIU spectral dataset available.

The authors would also thank the colleagues C. Belviso and L. Medici for the assistance in the mineralogical analysis of the samples.

Author Contributions

Raffaele Casa conceived the project and was project PI. Fabio Castaldi and Angelo Palombo, designed, performed and supported the experiments. Federico Santini and Stefano Pignatti analyzed the data. Fabio Castaldi and Simone Pascucci performed the imagery processing, statistical analysis and modeling. All authors contributed in paper writing.

Conflicts of Interest

The authors declare no conflict of interest.

References

1. Viscarra Rossel, R.A.; Walvoort, D.J.J.; McBratney, A.B.; Janik, L.J.; Skjemstad, J.O. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. *Geoderma* **2006**, *131*, 59–75.
2. Ben-Dor, E.; Chabrillat, S.; Dematte, J.A.M.; Taylor, G.R.; Hill, J.; Whiting, M.L.; Sommer, S. Using imaging spectroscopy to study soil properties. *Remote Sens. Environ.* **2009**, *113*, S38–S55.
3. Richter, K.; Hank, T.; Atzberger, C.; Locherer, M.; Mauser, W. Regularization strategies for agricultural monitoring: The EnMAP vegetation analyzer (AVA). In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, Munich, Germany, 22–27 July 2012; pp. 6613–6616.
4. Tanii, J.; Iwasaki, A.; Kawashima, T.; Inada, H. Results of evaluation model of Hyperspectral Imager Suite (HISUI). In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, Munich, Germany, 22–27 July 2012; pp. 131–134.
5. Pignatti, S.; Acito, N.; Amato, U.; Casa, R.; de Bonis, R.; Diani, M.; Laneve, G.; Matteoli, S.; Palombo, A.; Pascucci, S.; *et al.* Development of algorithms and products for supporting the Italian hyperspectral PRISMA mission: The SAP4PRISMA project. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, Munich, Germany, 22–27 July 2012; pp. 127–130.
6. Houborg, R.; Anderson, M.; Gao, F.; Schull, M.; Cammalleri, C. Monitoring water and carbon fluxes at fine spatial scales using HypsIRI-like measurements. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, Munich, Germany, 22–27 July 2012; pp. 7302–7305.
7. Lobell, D.B.; Asner, G.P. Moisture effects on soil reflectance. *Soil Sci. Soc. Am. J.* **2002**, *66*, 722–727.
8. Morgan, C.L.S.; Waiser, T.H.; Brown, D.J.; Hallmark, C.T. Simulated *in situ* characterization of soil organic and inorganic carbon with visible near-infrared diffuse reflectance spectroscopy. *Geoderma* **2009**, *151*, 249–256.
9. Nocita, M.; Stevens, A.; Noon, C.; van Wesemael, B. Prediction of soil organic carbon for different levels of soil moisture using Vis-NIR spectroscopy. *Geoderma* **2013**, *199*, 37–42.
10. Wu, C.-Y.; Jacobson, A.R.; Laba, M.; Baveye, P.C. Alleviating Moisture Content Effects on the Visible Near-Infrared Diffuse-Reflectance Sensing of Soils. *Soil Sci.* **2009**, *174*, 456–465.
11. Waiser, T.H.; Morgan, C.L.S.; Brown, D.J.; Hallmark, C.T. *In Situ* Characterization of Soil Clay Content with Visible Near-Infrared Diffuse Reflectance Spectroscopy. *Soil Sci. Soc. Am. J.* **2007**, *71*, 389–396.

12. Ge, Y.; Morgan, C.L.S.; Ackerson, J.P. VisNIR spectra of dried ground soils predict properties of soils scanned moist and intact. *Geoderma* **2014**, *221–222*, 61–69.
13. Ackerson, J.P.; Demattê J.A.M.; Morgan, C.L.S. Predicting clay content on field-moist intact tropical soils using a dried, ground VisNIR library with external parameter orthogonalization. *Geoderma* **2015**, *259–260*, 196–204.
14. Rienzi, E.A.; Mijatovic, B.; Mueller, T.G.; Matocha, C.J.; Sikora, F.J.; Castrignanò, A. Prediction of soil organic carbon under varying moisture levels using reflectance spectroscopy. *Soil Sci. Soc. Am. J.* **2014**, *78*, 958–967.
15. Minasny, B.; McBratney, A.B.; Bellon-Maurel, V.; Roger, J.-M.; Gobrecht, A.; Ferrand, L.; Joalland, S. Removing the effect of soil moisture from NIR diffuse reflectance spectra for the prediction of soil organic carbon. *Geoderma* **2011**, *167–168*, 118–124.
16. Kuang, B.; Mouazen, A.M. Effect of spiking strategy and ratio on calibration of on-line visible and near infrared soil sensor for measurement in European farms. *Soil Tillage Res.* **2013**, *128*, 125–136.
17. Ji, W.; Viscarra Rossel, R.A.; Shi, Z. Accounting for the effects of water and the environment on proximally sensed vis-NIR soil spectra and their calibrations. *Eur. J. Soil Sci.* **2015**, *66*, 555–565.
18. Haubrock, S.-N.; Chabrilat, S.; Lemmnitz, C.; Kaufmann, H. Surface soil moisture quantification models from reflectance data under field conditions. *Int. J. Remote Sens.* **2008**, *29*, 3–29.
19. 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.
20. Casa, R.; Castaldi, F.; Pascucci, S.; Basso, B.; Pignatti, S. Geophysical and hyperspectral data fusion techniques for in-field estimation of soil properties. *Vadose Zone J.* **2013**, doi:10.2136/vzj2012.0201.
21. Pascucci, S.; Casa, R.; Belviso, C.; Palombo, A.; Pignatti, S.; Castaldi, F. Estimation of soil organic carbon from airborne hyperspectral thermal infrared data: A case study. *Eur. J. Soil Sci.* **2014**, *65*, 865–875.
22. Li, D.; Chen, X.; Peng, Z.; Chen, S.; Chen, W.; Han, L.; Li, Y. Prediction of soil organic matter content in a litchi orchard of South China using spectral indices. *Soil Tillage Res.* **2012**, *123*, 78–86.
23. Casa, R.; Castaldi, F.; Pascucci, S.; Palombo, A.; Pignatti, S. A comparison of sensor resolution and calibration strategies for soil texture estimation from hyperspectral remote sensing. *Geoderma* **2013**, *197–198*, 17–26.
24. FAO-ISRIC-ISSS. *World Reference Base for Soil Resources*; World Soil Resources Report 84. Food and Agriculture Organisation: Rome, Italy, 1998.
25. Soil Survey Staff. *Keys to Soil Taxonomy*, 12th ed.; USDA-Natural Resources Conservation Service: Washington, DC, USA, 2014.
26. Beuselinck, L.; Govers, G.; Poesen, J.; Degraer, G.; Froyen, L. Grain-size analysis by laser diffractometry: Comparison with the sieve-pipette method. *Catena* **1998**, *32*, 193–208.
27. Liu, W.; Baret, F.; Gu, X.; Zhang, B.; Tong, Q.; Zheng, L. Evaluation of methods for soil surface moisture estimation from reflectance data. *Int. J. Remote Sens.* **2003**, *24*, 2069–2083.
28. Bellon-Maurel, V.; Fernandez-Ahumada, E.; Palagos, B.; Roger, J.-M.; McBratney, A. Critical review of chemometric indicators commonly used for assessing the quality of the prediction of soil attributes by NIR spectroscopy. *TrAC Trends Anal. Chem.* **2010**, *29*, 1073–1081.

29. Saxton, K.E.; Rawls, W.J. Soil water characteristic estimates by texture and organic matter for hydrologic solutions. *Soil Sci. Soc. Am. J.* **2006**, *70*, 1569–1578.
30. Clark, R.N.; Roush, T.L. Reflectance spectroscopy: Quantitative analysis techniques for remote sensing applications. *J. Geophys. Res.* **1984**, *89*, 6329–6340.
31. Wold, S.; Sjöström, M.; Eriksson, L. PLS-regression: A basic tool of chemometrics. *Chemom. Intell. Lab. Syst.* **2001**, *58*, 109–130.
32. R Development Core Team. *R: A Language and Environment for Statistical Computing*; R Foundation for Statistical Computing: Vienna, Austria, 2011.
33. Berk, A.; Bernstein, L.S.; Anderson, G.P.; Acharya, P.K.; Robertson, D.C.; Chetwynd, J.H.; Adler-Golden, S.M. MODTRAN Cloud and Multiple Scattering Upgrades with Application to AVIRIS. *Remote Sens. Environ.* **1998**, *65*, 367–375.
34. Clark, R.N. Chapter 1: Spectroscopy of rocks and minerals, and principles of spectroscopy. In *Manual of Remote Sensing, Remote Sensing for the Earth Sciences*; Rencz, A.N., Ed.; John Wiley and Sons: New York, NY, USA, 1999; Volume 3, pp. 3–58.
35. Tóth, G.; Jones, A.; Montanarella, L. The LUCAS topsoil database and derived information on the regional variability of cropland topsoil properties in the European Union. *Environ. Monit. Assess.* **2013**, *185*, 7409–7425.
36. Summers, D.; Lewis, M.; Ostendorf, B.; Chittleborough, D. Visible near-infrared reflectance spectroscopy as a predictive indicator of soil properties. *Ecol. Indic.* **2011**, *11*, 123–131.
37. Bilgili, A.V.; Akbas, F.; van Es, H.M. Combined use of hyperspectral VNIR reflectance spectroscopy and kriging to predict soil variables spatially. *Precis. Agric.* **2010**, *12*, 395–420.
38. Lagacherie, P.; Baret, F.; Feret, J.-B.; Madeira Netto, J.; Robbez-Masson, J.M. Estimation of soil clay and calcium carbonate using laboratory, field and airborne hyperspectral measurements. *Remote Sens. Environ.* **2008**, *112*, 825–835.
39. Viscarra Rossel, R.A.; Cattle, S.R.; Ortega, A.; Fouad, Y. *In situ* measurements of soil colour, mineral composition and clay content by vis–NIR spectroscopy. *Geoderma* **2009**, *150*, 253–266.
40. Ackerson, C.; Morgan, L.S.; Ge, Y.; McBratney, A.; Minasny, B. A tale of two projections: Removing the effects of water content and intactness from *in situ* VisNIR spectra using external parameter orthogonalization and direct standardization. In Proceedings of the Pedometrics 2015, Cordoba, Spain, 14–17 September 2015; p. 70.
41. Folkman, M.A.; Pearlman, J.; Liao, L.B.; Jarecke, P.J. EO-1/Hyperion hyperspectral imager design, development, characterization, and calibration. *Proc. SPIE* **2001**, doi:10.1117/12.417022.
42. Castaldi, F.; Casa, R.; Castrignanò, A.; Pascucci, S.; Palombo, A.; Pignatti, S. Estimation of soil properties at the field scale from satellite data: A comparison between spatial and non-spatial techniques. *Eur. J. Soil Sci.* **2014**, *65*, 842–851.