Evaluation of the spectral characteristics of five hyperspectral and multispectral sensors for soil organic carbon estimation in burned areas

Juanjo **Peón** [A,D], Susana **Fernández** [B], Carmen **Recondo** [A] and Javier F. **Calleja** [C]

[A] Area of Cartographic, Geodesic and Photogrammetric Engineering, Department of Mining Exploitation and Prospecting, University of Oviedo, Gonzalo Gutiérrez Quirós s/n, 33600 Mieres, Asturias, Spain.

[B] Department of Geology, University of Oviedo, Jesús Arias de Velasco s/n, 33005 Oviedo, Asturias, Spain.

[C] Department of Physics, Polytechnic School of Mieres, University of Oviedo, Gonzalo Gutiérrez Quirós s/n, 33600 Mieres, Asturias, Spain.

[D] Corresponding author. Email: juanjopeon@gmail.com

Abstract

Frequent wildfires in the NW of Spain affect soil organic matter. Soil properties can be estimated both spatially and temporally using remote sensing. A wide range of satellite and airborne hyperspectral/multispectral sensors are currently available. The spectral resolution varies significantly among sensors, making it difficult to identify the most suitable sensors and spectral regions for a specific application. This study aims at identifying the sensors and wavelengths with the greatest potential for topsoil organic C mapping. Total (TOC) and oxidizable organic carbon (OC) content were measured in 89 soil samples collected in a mountain region of NW Spain. Reflectance spectra of the samples (400-2450 nm) were resampled to the bands of five sensors: Compact Airborne Spectrographic Imager (CASI), Airborne Hyperspectral Scanner (AHS), Hyperion, Landsat 5 Thematic Mapper (TM), and Moderate Resolution Imaging Spectroradiometer (MODIS). Calibration models were obtained using partial least squares regression (PLSR). PLSR models proved to be effective for the hyperspectral sensors and also for the multispectral sensor MODIS (R^2 = 0.75-0.89), which suggests that hyperspectral capability is not required to accurately predict topsoil organic C. Models based on Landsat performed well, but with an error ~30-45% greater than that obtained for the hyperspectral sensors and MODIS.

Summary

The spectral characteristics of five hyperspectral and multispectralsensors were evaluated for topsoil organic carbon prediction in burned areas. The spectral resolution of both kind of sensors was suitable for the prediction. The most relevant spectral regions for topsoil carbon estimation were the visible and short-wave infrared.

Additional keywords: VIS-NIR-SWIR spectroscopy; PLSR; CASI; AHS; Hyperion; Landsat; MODIS.

Introduction

Fire is a major disturbance in many ecosystems around the world (Certini 2005; Santin *et al.* 2008; Santin and Doerr 2016). Wildfires in the Atlantic region of Spain usually show a relative high recurrence (Álvarez and Marquínez 2007). The increase of erosion following wildfires is widely documented (Wondzell and King 2003; Shakesby and Doerr 2006). Several processes are involved in the post-fire erosion, mainly erosion by overland flow and inter-rill erosion, and cause the loss of soil organic matter (SOM) (Canfield *et al.* 2005; Rumpel *et al.* 2006; Shakesby and Doerr 2006). Although the recovery of the vegetation cover is not complete in areas with high recurrence of wildfires (Fernandez *et al.* 2005; Marquinez *et al.* 2008), the rapid recovery of the vegetation due to the Atlantic climate characteristics makes it difficult to monitor soils affected by wildfires.

Development of reliable techniques is required to improve the knowledge on C stocks (Burras *et al.* 2001). On the one hand, the use of laboratory spectroscopy in the visible (VIS, 400-700 nm), near infrared (NIR, 700-1300 nm), and short wave infrared (SWIR, 1300-2500 nm), in combination with multivariate statistical methods such as partial least squares regression (PLSR), has been proven to be an accurate technique to estimate soil organic carbon (SOC) (Viscarra Rossel *et al.* 2006; Croft *et al.* 2012). On the other hand, remote sensing can provide estimates of the spatial and temporal distribution of C stocks at a regional scale. Remote sensing is well established as a cost-effective, rapid, and reproducible means of providing quantitative data on soil properties (Irons *et al.* 1989). Although remote sensing has been widely used to map soil types, there have been fewer studies on the direct quantification of topsoil SOC content because of several constraints, such as atmospheric absorption interfering with soil absorption features and low signal-to-noise ratios. Due to these constraints,

3

research has been mainly restricted to agricultural and bare soil surfaces (Selige *et al.* 2006; Gomez *et al.* 2012).

There has been an increasing number of studies which have used VIS-NIR-SWIR hyperspectral airborne imaging in topsoil mapping, but they still remain in the testing phases. SOC was estimated over agricultural areas with $R^2 > 0.74$ using PLSR and airborne hyperspectral data collected with HyMap (Selige *et al.* 2006; Patzold *et al.* 2008; Hbirkou *et al.* 2012). The potential of Airborne Hyperspectral Scanner (AHS) for SOC prediction was also proved in several studies (Stevens *et al.* 2008; Stevens *et al.* 2010). Fernández *et al.* (2016) also used AHS data to map SOC in a mountain area partially covered by vegetation and obtained $R^2 > 0.72$. Studies by Stevens *et al.* (2006) and Uno *et al.* (2005) who used the Compact Airborne Spectrographic Imager (CASI) over agricultural areas, obtained an R^2 of 0.85 and 0.75, respectively. DeTar *et al.* (2008) used data from AVNIR hyperspectral sensor to map several soil properties, but obtained less satisfactory results, for both SOM ($R^2 = 0.48$) and C ($R^2 = 0.27$).

There have been few studies on the use of satellite hyperspectral data for topsoil mapping because only one VIS-NIR-SWIR hyperspectral satellite sensor is currently available, Hyperion (Croft *et al.* 2012). Gomez *et al.* (2008) used Hyperion data to map SOC in agricultural fields and pastures, and obtained $R^2 = 0.51$. Using the same sensor, Lu *et al.* (2013) estimated SOC in bare soils with $R^2 = 0.63$. Relatively fewer studies have reported the use of satellite multispectral data for topsoil properties mapping. Among the sensors with medium spatial resolution, Landsat was the most used to map SOC and SOM: Huang *et al.* (2007) in glacial till soils, Jarmer *et al.* (2010) in non-agricultural arid and semi-arid areas, Nanni and Demattê (2006) in bare soils, and Peng *et al.* (2015) in an area mainly devoted to agriculture (R^2 of 0.46, 0.80-0.91, 0.51, and 0.46 respectively). Data from the Moderate Resolution Imaging Spectroradiometer

(MODIS) have been recently used to map SOC in Africa at a spatial resolution of 500 m (Vågen *et al.* 2016); the prediction models performed well, with $R^2 = 0.74$. Despite the constrains of remote sensing, relative strong predictions were obtained using airborne hyperspectral sensors, and moderately satisfactory with hyperspectral and multispectral satellite sensors (Croft *et al.* 2012).

The number and location of the bands through the spectrum as well as the bandwidth vary significantly among the different available sensors, and it is a challenge to decide the most suitable sensor and spectral region for mapping topsoil C. This research was aimed at identifying suitable sensors and spectral regions for estimating topsoil oxidizable (OC) and total organic carbon (TOC). Three hyperspectral (CASI, AHS, and Hyperion) and two multispectral sensors (Landsat 5 Thematic Mapper (TM), and MODIS) were evaluated in four spectral regions: VIS, NIR, SWIR, and VIS-NIR-SWIR. Reflectance spectra were simulated for each sensor by resampling the lab VIS-NIR-SWIR spectra of a set of soil samples, and then the resampled spectra were used to estimate SOC using PLSR.

Material and methods

Study area

In order to perform a valid method to monitor SOM in large burned areas of the Atlantic region of Spain, a training area of ~6000 ha located in the western part of the Cantabrian Range (NW Spain) was chosen (Fig. 1). The study area corresponds to a mountain region affected by frequent wildfires and is characterised by an Atlantic mild-humid climate, with a mean annual temperature of ~8°C and a mean annual precipitation of 1500 mm. The altitudes are in the range 200-1200 m. Dominant soil types are Lithosols, Regosols and Histosols (FAO 2014), with poor evolved profile,

stony, and with high SOM levels related to necromass from partially burned vegetation (Fernandez *et al.* 2005).

Soil collection, chemical analysis, and laboratory spectral measurements

Eighty-nine samples of the upper 5 cm of soil were collected in the heathery slopes of the study area in 2013 (Fig. 1). Samples were sieved to a size fraction < 2 mm and homogenised. Total C and oxidizable C (OC) content of the samples was determined using conventional methods; dry-combustion (Heiri *et al.* 2001) and Walkley-Black (Walkley and Black 1934), respectively. The total C corresponds to the Total Organic Carbon (TOC), as the soil parent material at the study area is non-calcareous (Fernandez *et al.* 2005). Samples were scanned in the lab using an ASD LabSpec® spectrometer with a spectral range of 350-2500 nm, a sampling interval of 1 nm, and a spectral resolution of 3 nm at 700 nm and 6 nm at 1400 nm and 2100 nm. A Hi-Brite contact probe with built-in light source and a spot size of 10 mm was used. The diffuse reflectance spectrum of each sample was obtained by averaging ten spectral measurements. Spectra were reduced to 400-2450 nm to remove noise at the edge of each spectrum.

Spectra resampling

Reflectance spectra were resampled to simulate the response in five sensors: CASI and AHS sensors owned by the Spanish National Institute for Aerospace Technology (INTA), Hyperion (Pearlman *et al.* 2003), Landsat 5 TM, and MODIS (Salomonson *et al.* 1989). The overall process is shown in the flowchart in Fig. 2. Landsat 5 TM was selected instead of more recent Landsat sensors because the availability of a long-term archive for this sensor, with more than 28 years of observations, make it very suitable for multitemporal analysis. For each sensor, bands which not overlap the lab spectra,

which are usually affected by noise, and those located at wavelengths with strong atmospheric absorption, were excluded. In the case of Hyperion, the 155 stable bands selected by Datt *et al.* (2003) were chosen. The number of selected bands, as well as the specifications of each sensor, are included in Table 1.

Spectral response libraries of each sensor were used to resample the reflectance spectra. Libraries for CASI and Hyperion were obtained using a Gaussian model with mean and standard deviation set to the band wavelength and Full Width at Half Maximum (FWHM), respectively. An example of the resampling process for a reflectance spectrum is presented in Fig. 3*a*, where it is shown the gradual degradation of the spectral resolution for each sensor. In the ASD spectrum in Fig. 3*a*, a strong absorption feature due to water can be seen at ∼1900 nm (Viscarra Rossel and Behrens 2010; Croft *et al.* 2012).

Modelling of SOC

The transformation log(1/reflectance) was applied to the reflectance spectra, as the ratio of performance to deviation (RPD) increased significantly with this conversion (Griffiths and Dahm 2007). TOC and OC were modelled using PLSR under leave-one-out cross-validation (Haaland and Thomas 1988; Wold *et al.* 2001; Viscarra Rossel *et al.* 2006). Akaike Information Criterion (AIC) was used to determine the optimal number of factors to be included in each model (Akaike 1969; Li *et al.* 2002). Further details of the calibration method can be found in Fernández *et al.* (2016).

For each soil property and sensor, four PLSR models were obtained for the following regions of the spectrum: VIS, NIR, SWIR, and VIS-NIR-SWIR (Fig. 2). In the case of Landsat 5 TM, only the model in the VIS-NIR-SWIR region was obtained due to the low number of bands of this sensor. Although the PLSR technique is mainly

suitable for sensors with a high number of bands, it was also used for Landsat in order to compare the results among sensors using the same modelling technique.

The important wavelengths of the lab-spectra for SOC prediction were identified by using both the PLSR coefficients (b) (Haaland and Thomas 1988) and the Variable Importance for Projection (VIP) (Wold *et al.* 2001), following the methodology proposed by Viscarra Rossel *et al.* (2008). Further details of VIP and its performace as a method for identifying relevant predictor variables is given by Chong and Jun (2005). According to Wold *et al.* (2001), the higher the b and VIP are, the more significant the variable is. The most important wavelengths were identified by setting thresholds for both b and VIP. The thresholds for b were based on their standard deviations (Viscarra Rossel *et al.* 2008), and the thresholds for VIP were set to 1, as suggested by Chong and Jun (2005). These thresholds were also used in other recent studies based on PLSR and VIS-NIR-SWIR spectra (Gomez *et al.* 2015).

Validation of SOC prediction models

Model quality was analysed using the root mean square error of cross-validation (RMSE$_{cv}$), the coefficient of determination (R$^2_{cv}$), and the RPD. RPD is the ratio between standard deviation of the calibration set and RMSE$_{cv}$. The model having the highest RPD is considered to be the best model. The model prediction ability was classified by Chang *et al.* (2001) in the following categories: good (RPD $>$ 2), intermediate (2 $>$ RPD $>$ 1.4), and no prediction capability (RPD $<$ 1.4).

Results and discussion

Summary statistics of SOC

Mean TOC and OC content of the 89 soil samples is 29.4% and 19.5%, respectively. Both soil properties show a wide range of variation (TOC: 4.3-63.2%, OC: 2.7-37.9%), which is representative of the study area and in accordance with results obtained in previous studies in the same area (Fernandez *et al.* 2005; Santin *et al.* 2008). The relative high values of C are related to the existence of paleo soils inherited from Pleistocene glaciations. These Histosols occupy paleo surfaces in the upper part of the hillslopes. The minimum values of C correspond to soils with extremely thin organic horizons, which origin is related to the degradation caused by frequent wildfires (Fernandez *et al.* 2005; Menéndez Duarte *et al.* 2008). The parent material of these soils is characterized by stony coarse regolith. TOC and OC are highly correlated and show a Pearson's correlation coefficient R > 0.9.

SOC prediction models based on laboratory spectra

Summary statistics of the PLSR models for SOC prediction are shown in Table 2. Accurate TOC and OC prediction models were obtained using the laboratory spectra (R^2_{cv} > 0.78 and RPD > 2.13). Between the two soil properties evaluated, OC was estimated with the greatest accuracy. According to Chang *et al.* (2001), we can state that lab spectroscopy can be a reliable technique to measure TOC and OC content. The results of these models are similar to those reported by Chang and Laird (2002) (R^2 = 0.89), Martin *et al.* (2002) (R^2 = 0.75, RPD = 1.97), and Mouazen *et al.* (2010) (R^2 = 0.8).

The b regression coefficients and VIP of the PLSR models based on the laboratory spectra are shown in Fig. 3*c* and 3*d*, respectively. The most important wavelengths for

TOC and OC prediction, identified by setting thresholds both in b and VIP, are presented in Table 3. In the VIS, the regression coefficients were similar for TOC and OC. Regions covering 400-413 nm (blue) and 556-570 nm (green) were significant in the TOC model. It is in accordance with Viscarra Rossel *et al.* (2006), who found important peaks for organic carbon prediction at 410 and 570 nm; and with Vasques *et al.* (2009), who also found positive regression coefficients between 380 and 430 nm. Krishnan *et al.* (1980) found 564 nm to be an optimal wavelength for SOM prediction. The region covering 471-526 nm was also important for TOC prediction, and it is consistent with results by Vasques *et al.* (2009), who also obtained significant negative regression coefficients in this region. Besides these regions, two ranges were important for OC prediction in the VIS: 651-668 nm, attributed to chlorophyll pigments at around 664 nm (Ben-Dor *et al.* 1997), and 692-703 nm (red). These results are in accordance with previous studies, in which the VIS region was also important for calibration of C and organic matter (Krishnan *et al.* 1980; Cozzolino and Morón 2006; Mouazen *et al.* 2007; Stevens *et al.* 2010).

In the NIR, no significant regions were identified, except the region 961-1000 nm, important for OC estimation. This region around 970 nm was also important in the PLSR models for organic matter prediction by Kooistra *et al.* (2003). Significant regions were found in the SWIR, for both TOC and OC (Table 3). The most important regions for TOC were: 2180-2239 nm, which was also identified by Morgan *et al.* (2009) and Vasques *et al.* (2009) as a relevant region and might be associated to biochar components (Bellon-Maurel and McBratney 2011); and 2261-2346 nm, also relevant in the models by Vasques *et al.* (2009). Two additional regions in the SWIR were significant for OC estimation: 1390-1431 nm, which was also identified by Kooistra *et al.* (2003) and may be attributed to absorption of O-H bonds (Viscarra Rossel *et al.*

2006); and 2041-2154 nm, also significant in the models by Vasques *et al.* (2009). In the latter region, at 2100 nm, McMorrow *et al.* (2004) noted absorption features related to cellulose. Many researches have also found the SWIR region to be important to estimate C, with the best range between 1650 and 2500 nm (Lee *et al.* 2009; Morgan *et al.* 2009; Vasques *et al.* 2009; Bellon-Maurel and McBratney 2011).

SOC prediction models based on resampled laboratory spectra

Lab-spectra were resampled for the different sensors and the effect of the spectral resampling on the estimation of TOC and OC was analysed. In the VIS, the results obtained with the three hyperspectral sensors (CASI, AHS, and Hyperion) and ASD were very similar, both in TOC (RPD = 2.06-2.12) and OC (RPD = 2.56-2.66) (Table 2). The best results in the VIS were obtained with AHS. The largest differences among sensors were observed in the NIR. The best performance was obtained with Hyperion both in TOC (RPD = 1.87) and OC (RPD = 2.18), and was similar to that obtained with ASD. The estimation with CASI and AHS had a lower accuracy, and was attributed to the distribution of their bands in the NIR, only located in the ~700-1050 nm region (Fig. 3*b*). In the NIR, CASI has more and narrower bands than Hyperion (CASI: 74 bands with FWHM of 6 nm, Hyperion: 57 bands with FWHM of 10-12 nm), however its distribution is less suitable for estimating SOC than it is in the case of Hyperion, which bands entirely cover the NIR region, between 700 and 1300 nm. Furthermore, CASI bands do not cover the significant region 1390-1431 nm (see Table 3 and Fig. 3*c*), which may explain the drop of accuracy. The results obtained in the SWIR with AHS and Hyperion were very similar to that obtained with ASD, for TOC (RPD = 2.11-2.24) and OC (RPD = 2.47).

The best estimation with the three hyperspectral sensors was obtained in the VIS-NIR-SWIR for both soil properties. The results in this region were similar for the hyperspectral sensors and ASD, with RPD = 2.11-2.28 for TOC and RPD = 2.75-2.97 for OC. These models provided accurate predictions according to the classification by Chang *et al.* (2001).

As for the multispectral sensors, accurate prediction models were obtained for MODIS in the VIS, with RPD = 2.00 for TOC and RPD = 2.68 for OC. The performance of these models was very similar to that obtained with ASD (Table 2). The models based on NIR bands of MODIS had a limited predictive power, both in TOC (RPD = 1.43) and OC (RPD = 1.77). A limited performance was also obtained for the models based on MODIS SWIR bands. The significant drop of accuracy observed when the entire lab spectra were resampled to the MODIS bands in the NIR and SWIR was attributed to the low number of bands in these regions: 7 in the NIR and only 3 in the SWIR. This drop of performance may also be attributed to the location of the bands and their bandwidth, which could be too wide to successfully detect some important absorption features for the estimation of SOC. Some significant regions for SOC prediction, such as 2041-2154 nm and 2287-2315 nm (see Table 3 and Fig. 3), are not entirely covered by MODIS bands. The performance of the models based on MODIS bands from the VIS-NIR-SWIR region are similar to that obtained with ASD, both in TOC and OC.

As for Landsat 5 TM, the accuracy of the models in the VIS-NIR-SWIR was lower than that obtained with ASD, but according to Chang *et al.* (2001) model for OC still provided accurate predictions, with RPD = 2.03. An intermediate prediction ability was observed in the model for estimating TOC (RPD = 1.62). The bands of Landsat 5 TM do not entirely cover important regions for SOC prediction, such as 961-1000 nm and

1390-1431 nm, and it could explain the drop of accuracy observed with this sensor. The wide band of Landsat at 2223 nm with bandwidth of 252 nm brings together the information of the three significant regions 2041-2154 nm, 2223-2231 nm, and 2287-2315 nm, and may not properly describe the variability in this region.

Considering only the spectral region with best results in Table 2, VIS-NIR-SWIR, no significant differences were observed between the accuracy of the models based on the entire lab spectra (ASD with 2051 bands) and those based on lab spectra resampled to the three hyperspectral sensors (38-155 bands) and MODIS (20 bands). Despite the wide difference in the number of bands between lab and these imaging spectrometers, both TOC and OC were successfully estimated using PLSR, with less than 10 factors. The lowest accuracy was obtained for the models based on Landsat, but these models still provided moderately satisfactory results, with an error ~30-45% greater than that obtained in the models based on lab spectra for TOC and OC.

Comparison with other SOC prediction models

SOC was accurately estimated under controlled lab conditions using PLSR. The validation results of the lab models (Table 2) were compared to the results obtained in previous studies in which actual airborne/satellite data were used to predict SOC at a regional scale. Studies were compared using the R^2 and RPD values obtained in the validation. The use of the RMSE was discarded, as it is not comparable among studies in which soil properties show a different range of variation and are expressed in different units.

Accurate SOC prediction models were obtained in this study using the wavelengths associated with CASI (R^2 = 0.78-0.87 and RPD = 2.11-2.75). Stevens *et al.* (2006) and Uno *et al.* (2005) used CASI data to estimate SOC/SOM in agricultural fields and

13

obtained similar R^2 values (R^2 = 0.75-0.85) and lower RPD values (RPD = 1.57-1.86) than those obtained in this study.

Using the wavebands associated with AHS, we predicted SOC with R^2 = 0.81-0.89 and RPD = 2.28-2.97. Stevens *et al.* (2008) estimated SOC in agricultural fields using PLSR, under lab conditions (R^2 = 0.76 and RPD = 2.03) and also at a regional scale using AHS data (R^2 = 0.54 and RPD = 1.47). The error obtained using AHS data was approximately two to three times greater than the value obtained in the lab. Stevens *et al.* (2010) used AHS data to estimate SOC over agricultural fields and obtained R^2 = 0.72 and RPD = 1.89. Stevens *et al.* (2012) obtained better results in the SOC prediction with AHS using a large dataset and performing local models for the different regions and soil textures in the study area (R^2 = 0.79 and RPD = 2.33). SOC maps of our study area were obtained using AHS in a recent work by Fernández *et al.* (2016). The accuracy of these maps was satisfactory, with R^2 = 0.72-0.73, RPD = 1.89-1.92, and an error similar to that obtained in the lab.

SOC prediction models based on lab spectra resampled to Hyperion showed R^2 = 0.79-0.87 and RPD = 2.17-2.80 (good prediction ability). These results are better than those obtained in previous studies based on actual Hyperion data: Gomez *et al.* (2008) in agricultural fields and pastures (R^2 = 0.51 and RPD = 1.43), and Lu *et al.* (2013) in agricultural fields (R^2 = 0.63 and RPD = 1.65). The models obtained in both studies provided an intermediate prediction ability. The performance of the models by Zhang *et al.* (2013), who also used Hyperion data to estimate SOC/SOM in agricultural fields, was intermediate, with R^2 = 0.48-0.74 and RPD = 1.48-1.91. Jaber *et al.* (2011) used Hyperion data to predict SOC in agricultural fields, pastures and forests, but their models had a low prediction ability (RPD = 0.68). Models using Hyperion spectra were less accurate than models based on lab spectra and it was attributed to the presence of

different covers within the same pixel (spectral mixing), and to the noise present in the image (Gomez *et al.* 2008; Jaber *et al.* 2011; Lu *et al.* 2013).

Models obtained in this study using the wavelengths associated with Landsat showed a good-intermediate prediction ability, with R^2 = 0.62-0.75 and RPD = 1.62-2.03. Comparable or even better R^2 values were found in the previous studies in which Landsat data were used to estimate SOC/SOM. Nanni and Demattê (2006) estimated SOM in bare soils using lab spectra resampled to Landsat 5 TM (R^2 = 0.53) and also using actual Landsat data (R^2 = 0.51). Huang *et al.* (2007) estimated soil total C content in glacial till soils and obtained R^2 = 0.46 using Landsat reflectance data, and R^2 = 0.62 when topographical variables were included in the model. Peng *et al.* (2015) used Landsat to estimate SOC in an area mainly devoted to agriculture and obtained R^2 = 0.46 and RPD = 1.3. SOC was accurately estimated in non-agricultural arid and semi-arid areas using Landsat 5 TM data, with R^2 = 0.80-0.91 (Jarmer *et al.* 2010).

SOC prediction models obtained in this study for MODIS showed R^2 = 0.77-0.88 and RPD =2.01-2.95. Vågen *et al.* (2016) used actual MODIS data to map SOC in Africa at a spatial resolution of 500 m and obtained a similar R^2 value (R^2 = 0.74).

Models for SOC prediction obtained in previous studies based on actual airborne/satellite data were typically less accurate than models obtained in this study using simulated spectra. The drop in the RPD values was on average 29.4% for CASI, 27.6% for AHS, and 42.5% for Hyperion. These differences between lab and field predictions were attributed to a wide range of factors including the presence of noise in airborne/satellite data, the soil surface conditions such as the roughness and moisture content, the presence of vegetation and different covers within the same pixel, and the lack of a proper atmospheric correction. Among the hyperspectral sensors, the largest differences were observed in Hyperion and they were related to the low signal-to-ratio

of this sensor (Pearlman *et al.* 2003) and to its spatial resolution of 30 m. Based on the results showed in this section, the hyperspectral sensor with greatest potential to be transferred from lab to field scale is AHS.

As for the multispectral sensors, SOC models based on actual reflectance spectra showed R^2 values on average 14.6% lower than models based on lab spectra resampled for Landsat 5 TM, and 9.2% lower for MODIS. Although better results are expected using hyperspectral data, moderately satisfactory predictions were reported in the literature for both satellite multispectral sensors, which can be very valuable to map topsoil C in homogeneous and extensive bare soil areas, as well as in large areas affected by wildfires.

Conclusions

Topsoil organic carbon can be accurately estimated using PLSR and VIS-NIR-SWIR spectra of soils. A high number of narrow contiguous bands is not required to obtain accurate predictions of SOC, as the PLSR models proved to be effective not only for hyperspectral sensors, but also for multispectral sensors such as MODIS. The results suggest that the drop of accuracy in the estimation of SOC using satellite or airborne sensors seems to be more related to low signal-to-noise ratios or an inadequate atmospheric correction than to a lack of hyperspectral bands. Although previous studies based on multispectral sensors provided promising estimations of topsoil organic carbon in bare soil areas, further research on their application for soil properties mapping is required.

The location of the bands through the spectrum seems to be more important to effectively estimate SOC than the number of bands or their bandwidth. The most

significant regions for SOC prediction were the VIS and the SWIR, mainly between 2000 and 2450 nm. AHS showed the greatest potential to accurately estimate SOC.

Acknowledgments

References

Akaike H (1969) Fitting autoregressive models for prediction. *Annals of the Institute of Statistical Mathematics* **21**, 243-247. doi:10.1007/bf02532251

Álvarez MA, Marquínez J (Eds) (2007) 'Impacto de los incendios forestales en Asturias. Análisis de los últimos 30 años.' (Principado de Asturias, INDUROT, KRK Ediciones: Oviedo, España)

Bellon-Maurel V, McBratney A (2011) Near-infrared (NIR) and mid-infrared (MIR) spectroscopic techniques for assessing the amount of carbon stock in soils - Critical review and research perspectives. *Soil Biology & Biochemistry* **43**, 1398-1410. doi:10.1016/j.soilbio.2011.02.019

Ben-Dor E, Inbar Y, Chen Y (1997) The reflectance spectra of organic matter in the visible near-infrared and short wave infrared region (400-2500 nm) during a controlled decomposition process. *Remote Sensing of Environment* **61**, 1-15. doi:10.1016/s0034-4257(96)00120-4

Burras L, Cheng HH, Kimble JM, Kissel DE, Lal R, Luxmoore RJ, Mausbach MJ, Rice CW, Uehara G, Wilding LD (2001) Carbon Sequestration: Position of the Soil Science Society of America. (SSSA: Madison, Wisconsin)

Canfield HE, Wilson CJ, Lane LJ, Crowell KJ, Thomas WA (2005) Modeling scour and deposition in ephemeral channels after wildfire. *Catena* **61**, 273-291. doi:10.1016/j.catena.2005.03.013

Certini G (2005) Effects of fire on properties of forest soils: a review. *Oecologia* **143**, 1-10. doi:10.1007/s00442-004-1788-8

Chang CW, Laird DA (2002) Near-infrared reflectance spectroscopic analysis of soil C and N. *Soil Science* **167**, 110-116. doi:10.1097/00010694-200202000-00003

Chang CW, Laird DA, Mausbach MJ, Hurburgh CR (2001) Near-infrared reflectance spectroscopy-principal components regression analyses of soil properties. *Soil Science Society of America Journal* **65**, 480-490.

Chong IG, Jun CH (2005) Performance of some variable selection methods when multicollinearity is present. *Chemometrics and Intelligent Laboratory Systems* **78**, 103-112. doi:10.1016/j.chemolab.2004.12.011

Cozzolino D, Morón A (2006) Potential of near-infrared reflectance spectroscopy and chemometrics to predict soil organic carbon fractions. *Soil and Tillage Research* **85**, 78-85. doi:10.1016/j.still.2004.12.006

Croft H, Kuhn NJ, Anderson K (2012) On the use of remote sensing techniques for monitoring spatio-temporal soil organic carbon dynamics in agricultural systems. *Catena* **94**, 64-74. doi:10.1016/j.catena.2012.01.001

Datt B, McVicar TR, Van Niel TG, Jupp DLB, Pearlman JS (2003) Preprocessing EO-1 Hyperion hyperspectral data to support the application of agricultural indexes. *IEEE Transactions on Geoscience and Remote Sensing* **41**, 1246-1259. doi:10.1109/TGRS.2003.813206

DeTar WR, Chesson JH, Penner JV, Ojala JC (2008) Detection of soil properties with airborne hyperspectral measurements of bare fields. *Transactions of the ASABE* **51**, 463-470.

FAO (2014) World reference base for soil resources 2014. World Soil Resources Report 106. (Food and Agriculture Organization: Rome)

Fernandez S, Marquinez J, Duarte RM (2005) A susceptibility model for post wildfire soil erosion in a temperate oceanic mountain area of Spain. *Catena* **61**, 256-272. doi:10.1016/j.catena.2005.03.006

Fernández S, Peón J, Recondo C, Calleja JF, Guerrero C (2016) Spatial modelling of organic carbon in burned mountain soils using hyperspectral images, field datasets, and NIR spectroscopy (Cantabrian Range; NW Spain). *Land Degradation & Development* **27**, 1479-1488. doi:10.1002/ldr.2452

Gomez C, Lagacherie P, Coulouma G (2012) Regional predictions of eight common soil properties and their spatial structures from hyperspectral Vis-NIR data. *Geoderma* **189**, 176-185. doi:10.1016/j.geoderma.2012.05.023

Gomez C, Oltra-Carrió R, Bacha S, Lagacherie P, Briottet X (2015) Evaluating the sensitivity of clay content prediction to atmospheric effects and degradation of image spatial resolution using Hyperspectral VNIR/SWIR imagery. *Remote Sensing of Environment* **164**, 1-15. doi:10.1016/j.rse.2015.02.019

Gomez C, Viscarra Rossel RA, McBratney AB (2008) Soil organic carbon prediction by hyperspectral remote sensing and field vis-NIR spectroscopy: An Australian case study. *Geoderma* **146**, 403-411. doi:10.1016/j.geoderma.2008.06.011

Griffiths PR, Dahm DJ (2007) Continuum and Discontinuum Theories of Diffuse Reflection. In 'Handbook of Near-Infrared Analysis, Third Edition.' (Eds DA Burns, EW Ciurczak.) pp. 21-64. (CRC Press: Boca Raton, Florida)

Haaland DM, Thomas EV (1988) Partial least-squares methods for spectral analyses. 1. Relation to other quantitative calibration methods and the extraction of qualitative information. *Analytical Chemistry* **60**, 1193-1202. doi:10.1021/ac00162a020

Hbirkou C, Paetzold S, Mahlein A-K, Welp G (2012) Airborne hyperspectral imaging of spatial soil organic carbon heterogeneity at the field-scale. *Geoderma* **175**, 21-28. doi:10.1016/j.geoderma.2012.01.017

Heiri O, Lotter AF, Lemcke G (2001) Loss on ignition as a method for estimating organic and carbonate content in sediments: Reproducibility and comparability of results. *Journal of Paleolimnology* **25**, 101-110. doi:10.1023/A:1008119611481

Huang X, Senthilkumar S, Kravchenko A, Thelen K, Qi J (2007) Total carbon mapping in glacial till soils using near-infrared spectroscopy, Landsat imagery and topographical information. *Geoderma* **141**, 34-42. doi:10.1016/j.geoderma.2007.04.023

Irons JR, Weismiller RA, Petersen GW (1989) Soil Reflectance. In 'Theory and Applications of Optical Remote Sensing.' (Ed. G Asrar.) pp. 66–106. (John Wiley and Sons: New York, USA)

Jaber SM, Lant CL, Al-Qinna MI (2011) Estimating spatial variations in soil organic carbon using satellite hyperspectral data and map algebra. *International Journal of Remote Sensing* **32**, 5077-5103. doi:10.1080/01431161.2010.494637

Jarmer T, Hill J, Lavée H, Sarah P (2010) Mapping topsoil organic carbon in non-agricultural semi-arid and arid ecosystems of Israel. *Photogrammetric Engineering and Remote Sensing* **76**, 85-94.

Kooistra L, Wanders J, Epema GF, Leuven RSEW, Wehrens R, Buydens LMC (2003) The potential of field spectroscopy for the assessment of sediment properties in river floodplains. *Analytica Chimica Acta* **484**, 189-200. doi:10.1016/S0003-2670(03)00331-3

Krishnan P, Alexander JD, Butler BJ, Hummel JW (1980) Reflectance Technique for Predicting Soil Organic Matter. *Soil Science Society of America Journal* **44**, 1282-1285. doi:10.2136/sssaj1980.03615995004400060030x

Lee KS, Lee DH, Sudduth KA, Chung SO, Kitchen NR, Drummond ST (2009) Wavelength identification and diffuse reflectance estimation for surface and profile soil properties. *Transactions of the ASABE* **52**, 683-695.

Li B, Morris J, Martin EB (2002) Model selection for partial least squares regression. *Chemometrics and Intelligent Laboratory Systems* **64**, 79-89. doi:10.1016/S0169-7439(02)00051-5

Lu P, Wang L, Niu Z, Li L, Zhang W (2013) Prediction of soil properties using laboratory VIS-NIR spectroscopy and Hyperion imagery. *Journal of Geochemical Exploration* **132**, 26-33. doi:10.1016/j.gexplo.2013.04.003

Marquinez J, Wozniak E, Fernandez S, Martinez R (2008) Analysis of the evolution of soil erosion classes using multitemporal Landsat imagery. *International Journal of Wildland Fire* **17**, 549-558. doi:10.1071/wf06138

Martin PD, Malley DF, Manning G, Fuller L (2002) Determination of soil organic carbon and nitrogen at the field level using near-infrared spectroscopy. *Canadian Journal of Soil Science* **82**, 413-422.

McMorrow JM, Cutler MEJ, Evans MG, Al-Roichdi A (2004) Hyperspectral indices for characterizing upland peat composition. *International Journal of Remote Sensing* **25**, 313-325. doi:10.1080/01431160310001117065

Menéndez Duarte R, Wozniak E, Recondo C, Cabo C, Marquinez J, Fernandez S (2008) Estimation of surface roughness and stone cover in burnt soils using SAR images. *Catena* **74**, 264-272. doi:10.1016/j.catena.2008.05.003

Morgan CLS, Waiser TH, Brown DJ, Hallmark CT (2009) Simulated in situ characterization of soil organic and inorganic carbon with visible near-infrared diffuse reflectance spectroscopy. *Geoderma* **151**, 249-256. doi:10.1016/j.geoderma.2009.04.010

Mouazen AM, Kuang B, De Baerdemaeker J, Ramon H (2010) Comparison among principal component, partial least squares and back propagation neural network analyses for accuracy of measurement of selected soil properties with visible and near infrared spectroscopy. *Geoderma* **158**, 23-31. doi:10.1016/j.geoderma.2010.03.001

Mouazen AM, Maleki MR, De Baerdemaeker J, Ramon H (2007) On-line measurement of some selected soil properties using a VIS-NIR sensor. *Soil and Tillage Research* **93**, 13-27. doi:10.1016/j.still.2006.03.009

Nanni MR, Demattê JAM (2006) Spectral reflectance methodology in comparison to traditional soil analysis. *Soil Science Society of America Journal* **70**, 393-407. doi:10.2136/sssaj2003.0285

Patzold S, Mertens FM, Bornemann L, Koleczek B, Franke J, Feilhauer H, Welp G (2008) Soil heterogeneity at the field scale: a challenge for precision crop protection. *Precision Agriculture* **9**, 367-390. doi:10.1007/s11119-008-9077-x

Pearlman JS, Barry PS, Segal CC, Shepanski J, Beiso D, Carman SL (2003) Hyperion, a space-based imaging spectrometer. *IEEE Transactions on Geoscience and Remote Sensing* **41**, 1160-1173. doi:10.1109/TGRS.2003.815018

Peng Y, Xiong X, Adhikari K, Knadel M, Grunwald S, Greve MH (2015) Modeling Soil Organic Carbon at Regional Scale by Combining Multi-Spectral Images with Laboratory Spectra. *PLoS One* **10**, e0142295. doi:10.1371/journal.pone.0142295

Rumpel C, Chaplot V, Planchon O, Bernadou J, Valentin C, Mariotti A (2006) Preferential erosion of black carbon on steep slopes with slash and burn agriculture. *Catena* **65**, 30-40. doi:10.1016/j.catena.2005.09.005

Salomonson VV, Barnes WL, Maymon PW, Montgomery HE, Ostrow H (1989) MODIS: advanced facility instrument for studies of the Earth as a system. Geoscience and Remote Sensing. *IEEE Transactions on Geoscience and Remote Sensing* **27**, 145-153. doi:10.1109/36.20292

Santin C, Doerr SH (2016) Fire effects on soils: the human dimension. *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* **371**, doi:10.1098/rstb.2015.0171

Santin C, Knicker H, Fernandez S, Menendez-Duarte R, Alvarez MA (2008) Wildfires influence on soil organic matter in an Atlantic mountainous region (NW of Spain). *Catena* **74**, 286-295. doi:10.1016/j.catena.2008.01.001

Selige T, Boehner J, Schmidhalter U (2006) High resolution topsoil mapping using hyperspectral image and field data in multivariate regression modeling procedures. *Geoderma* **136**, 235-244. doi:10.1016/j.geoderma.2006.03.050

Shakesby RA, Doerr SH (2006) Wildfire as a hydrological and geomorphological agent. *Earth-Science Reviews* **74**, 269-307. doi:10.1016/j.earscirev.2005.10.006

Stevens A, Miralles I, van Wesemael B (2012) Soil organic carbon predictions by airborne imaging spectroscopy: Comparing cross-validation and validation. *Soil Science Society of America Journal* **76**, 2174-2183. doi:10.2136/sssaj2012.0054

Stevens A, Udelhoven T, Denis A, Tychon B, Lioy R, Hoffmann L, van Wesemael B (2010) Measuring soil organic carbon in croplands at regional scale using airborne imaging spectroscopy. *Geoderma* **158**, 32-45. doi:10.1016/j.geoderma.2009.11.032

Stevens A, van Wesemael B, Bartholomeus H, Rosillon D, Tychon B, Ben-Dor E (2008) Laboratory, field and airborne spectroscopy for monitoring organic carbon content in agricultural soils. *Geoderma* **144**, 395-404. doi:10.1016/j.geoderma.2007.12.009

Stevens A, Van Wesemael B, Vandenschrick G, Toure S, Tychon B (2006) Detection of carbon stock change in agricultural soils using spectroscopic techniques. *Soil Science Society of America Journal* **70**, 844-850. doi:10.2136/sssaj2005.0025

Uno Y, Prasher SO, Patel RM, Strachan IB, Pattey E, Karimi Y (2005) Development of field-scale soil organic matter content estimation models in Eastern Canada using airborne hyperspectral imagery. *Canadian Biosystems Engineering / Le Genie des biosystems au Canada* **47**, 9-14.

Vågen T-G, Winowiecki LA, Tondoh JE, Desta LT, Gumbricht T (2016) Mapping of soil properties and land degradation risk in Africa using MODIS reflectance. *Geoderma* **263**, 216-225. doi:10.1016/j.geoderma.2015.06.023

Vasques GM, Grunwald S, Sickman JO (2009) Modeling of Soil Organic Carbon Fractions Using Visible-Near-Infrared Spectroscopy. *Soil Science Society of America Journal* **73**, 176-184. doi:10.2136/sssaj2008.0015

Viscarra Rossel RA, Behrens T (2010) Using data mining to model and interpret soil diffuse reflectance spectra. *Geoderma* **158**, 46-54. doi:10.1016/j.geoderma.2009.12.025

Viscarra Rossel RA, Jeon YS, Odeh IOA, McBratney AB (2008) Using a legacy soil sample to develop a mid-IR spectral library. *Soil Research* **46**, 1-16.

Viscarra Rossel RA, Walvoort DJJ, McBratney AB, Janik LJ, Skjemstad JO (2006) Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for

simultaneous assessment of various soil properties. *Geoderma* **131**, 59-75. doi:10.1016/j.geoderma.2005.03.007

Walkley A, Black IA (1934) An examination of the Degtjareff method for determining soil organic matter, and a proposed modification of the chromic acid titration method. *Soil Science* **37**, 29-38.

Wold S, Sjöström M, Eriksson L (2001) PLS-regression: a basic tool of chemometrics. *Chemometrics and Intelligent Laboratory Systems* **58**, 109-130. doi:10.1016/S0169-7439(01)00155-1

Wondzell SM, King JG (2003) Postfire erosional processes in the Pacific Northwest and Rocky Mountain regions. *Forest Ecology and Management* **178**, 75-87. doi:10.1016/s0378-1127(03)00054-9

Zhang T, Li L, Zheng B (2013) Estimation of agricultural soil properties with imaging and laboratory spectroscopy. *Journal of Applied Remote Sensing* **7**, 73587. doi:10.1117/1.JRS.7.073587

Table 1. Sensor specifications and number of selected bands in each spectral region

Sensors are ranked in increasing order of pixel size. HYP: Hyperion; L5TM: Landsat 5 TM; MOD: MODIS. [A]Bandwidth values correspond to the Full Width at Half Maximum (FWHM). [B]Two bands of MODIS are acquired at a spatial resolution of 250 m, 5 at 500 m, and the remaining bands at 1000 m

Specifications	Lab spectra	Resampled lab spectra				
	ASD	CASI	AHS	HYP	L5TM	MOD
Sensor type	Lab	Airborne	Airborne	Satellite	Satellite	Satellite
Spectral range (nm)	400-2450	407-1053	443-2326	427-2355	485-2223	413-2130
Number of bands:						
VIS (400-700 nm)	301	62	9	27	3	10
NIR (700-1300 nm)	600	74	10	57	1	7
SWIR (1300-2450 nm)	1150		19	71	2	3
VIS-NIR-SWIR	2051	136	38	155	6	20
Bandwidth (nm)[A]	3-6	6	17-85	10-12	66-252	10-50
Spatial resolution (m)		1.5	5	30	30	1000[B]
Swath width (km)		1.5	4.1	7.5	185	2330
Temporal resolution (days)				200	16	1

Table 2. Summary statistics of the PLSR models for soil organic C prediction based on laboratory spectra
HYP: Hyperion; L5TM: Landsat 5 TM; MOD: MODIS. n: 89. NF: number of latent variables or PLSR factors

Soil property	Spectral region	Parameter	Lab spectra	Resampled lab spectra				
			ASD	CASI	AHS	HYP	L5TM	MOD
TOC (%)	VIS	NF	6	6	6	6		6
		AIC	186.6	188.1	185.8	187.1		190.7
		R^2_{cv}	0.77	0.76	0.78	0.77		0.75
		RMSE$_{cv}$ (%)	7.11	7.24	7.05	7.15		7.44
		RPD	2.10	2.06	2.12	2.09		2.00
	NIR	NF	6	8	7	6		6
		AIC	196.6	211.0	214.8	196.8		220.7
		R^2_{cv}	0.72	0.65	0.59	0.72		0.51
		RMSE$_{cv}$ (%)	7.96	8.95	9.54	7.97		10.44
		RPD	1.87	1.67	1.56	1.87		1.43
	SWIR	NF	4		6	5		3
		AIC	181.5		180.9	184.1		211.3
		R^2_{cv}	0.78		0.80	0.77		0.55
		RMSE$_{cv}$ (%)	7.02		6.67	7.07		10.04
		RPD	2.12		2.24	2.11		1.49
	VIS-NIR-SWIR	NF	7	8	7	7	6	7
		AIC	187.3	189.9	181.2	185.7	209.8	192.4
		R^2_{cv}	0.78	0.78	0.81	0.79	0.62	0.75
		RMSE$_{cv}$ (%)	7.01	7.06	6.55	6.88	9.23	7.42
		RPD	2.13	2.11	2.28	2.17	1.62	2.01
OC (%)	VIS	NF	5	5	6	5		7
		AIC	133.3	134.3	132.6	133.0		134.0
		R^2_{cv}	0.85	0.85	0.86	0.85		0.86
		RMSE$_{cv}$ (%)	4.00	4.04	3.88	3.98		3.85
		RPD	2.58	2.56	2.66	2.59		2.68
	NIR	NF	11	8	8	10		6
		AIC	153.6	164.2	164.4	158.6		168.9
		R^2_{cv}	0.82	0.75	0.74	0.79		0.68
		RMSE$_{cv}$ (%)	4.39	5.29	5.30	4.75		5.83
		RPD	2.35	1.95	1.95	2.18		1.77
	SWIR	NF	4		4	4		3
		AIC	133.5		135.3	135.2		169.4
		R^2_{cv}	0.84		0.84	0.84		0.63
		RMSE$_{cv}$ (%)	4.10		4.18	4.18		6.27
		RPD	2.52		2.47	2.47		1.65
	VIS-NIR-SWIR	NF	9	8	7	7	6	9
		AIC	130.2	133.7	124.8	130.3	157.0	129.5
		R^2_{cv}	0.88	0.87	0.89	0.87	0.75	0.88
		RMSE$_{cv}$ (%)	3.53	3.75	3.47	3.69	5.10	3.50
		RPD	2.93	2.75	2.97	2.80	2.03	2.95

Table 3. Most significant spectral regions for TOC and OC prediction obtained from PLSR models based on laboratory spectra (ASD)

For each spectral region, the sign of the b PLSR regression coefficients is shown in brackets

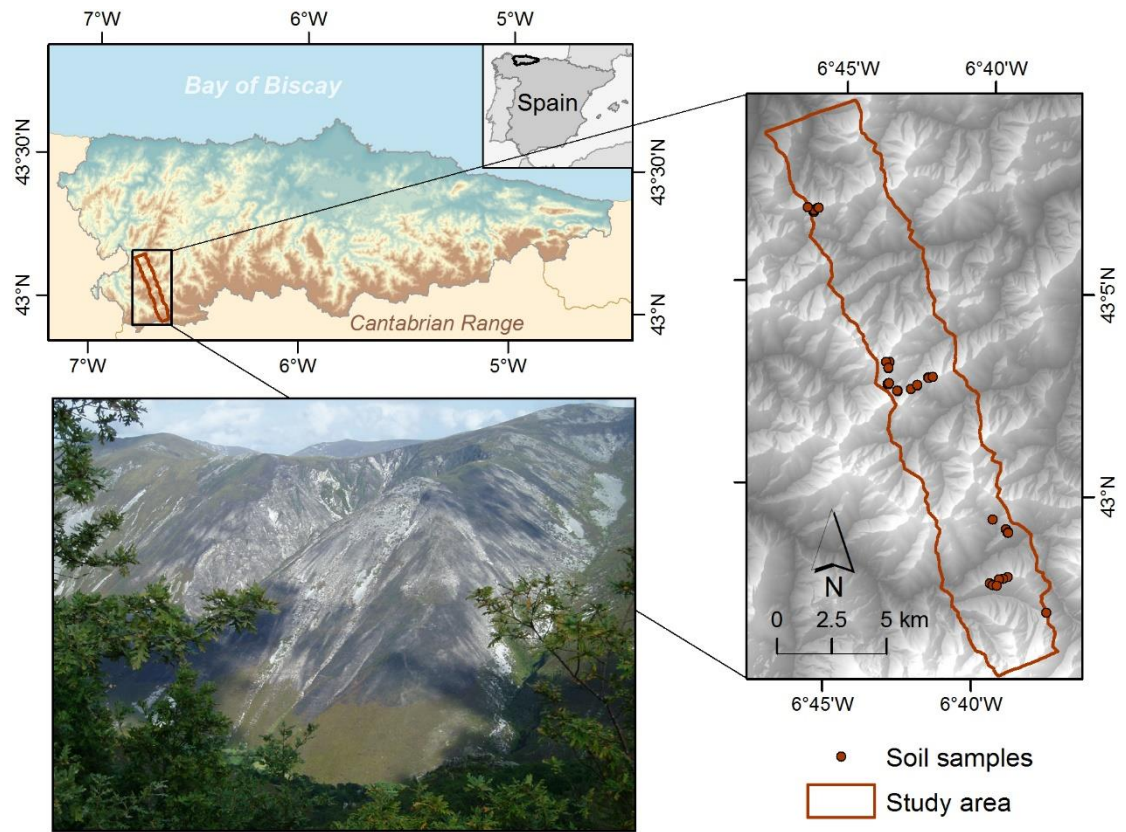| Spectral region	Wavelength (nm)	
	TOC	OC
VIS	400-413 (+)	400-409 (-)
	420-425 (+)	420-452 (+)
	471-526 (-)	475-534 (-)
	556-570 (+)	559-563 (+)
		574-616 (+)
		651-668 (-)
		692-703 (-)
NIR		961-1000 (+)
SWIR	1434-1476 (+)	1390-1431 (-)
	1805-1814 (-)	1906-1908 (-)
	1819-1892 (-)	2041-2154 (+)
	1919-1930 (+)	2223-2231 (-)
	2001-2023 (-)	2287-2315 (+)
	2180-2239 (-)	2353-2365 (-)
	2261-2346 (+)	2381-2385 (-)
		2425-2444 (-)

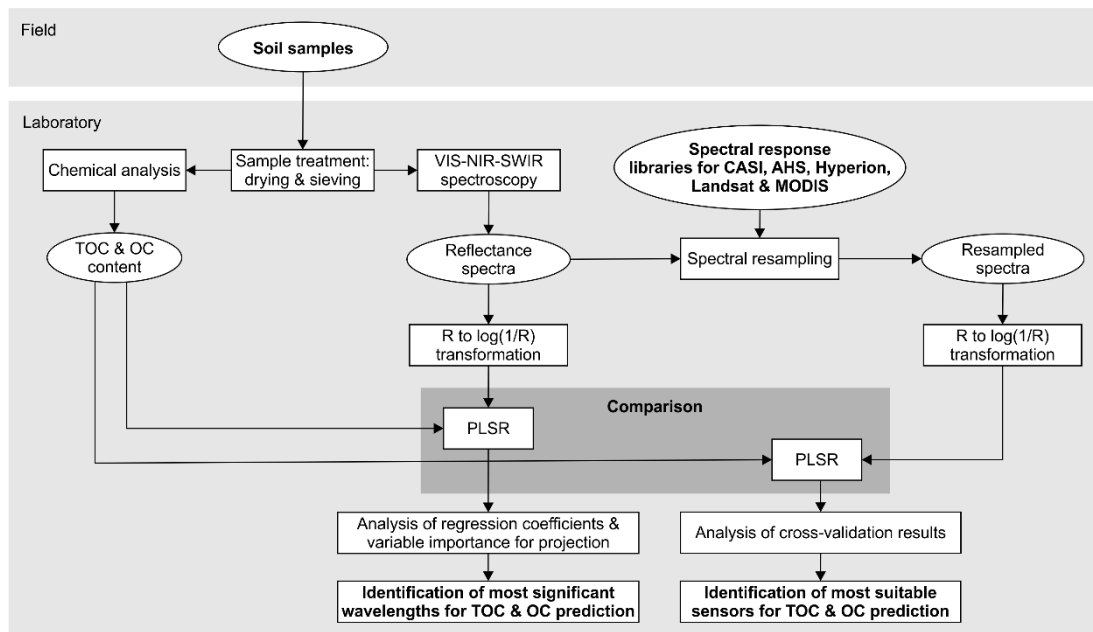**Fig. 1.** Location and general view of the study area (NW Spain).
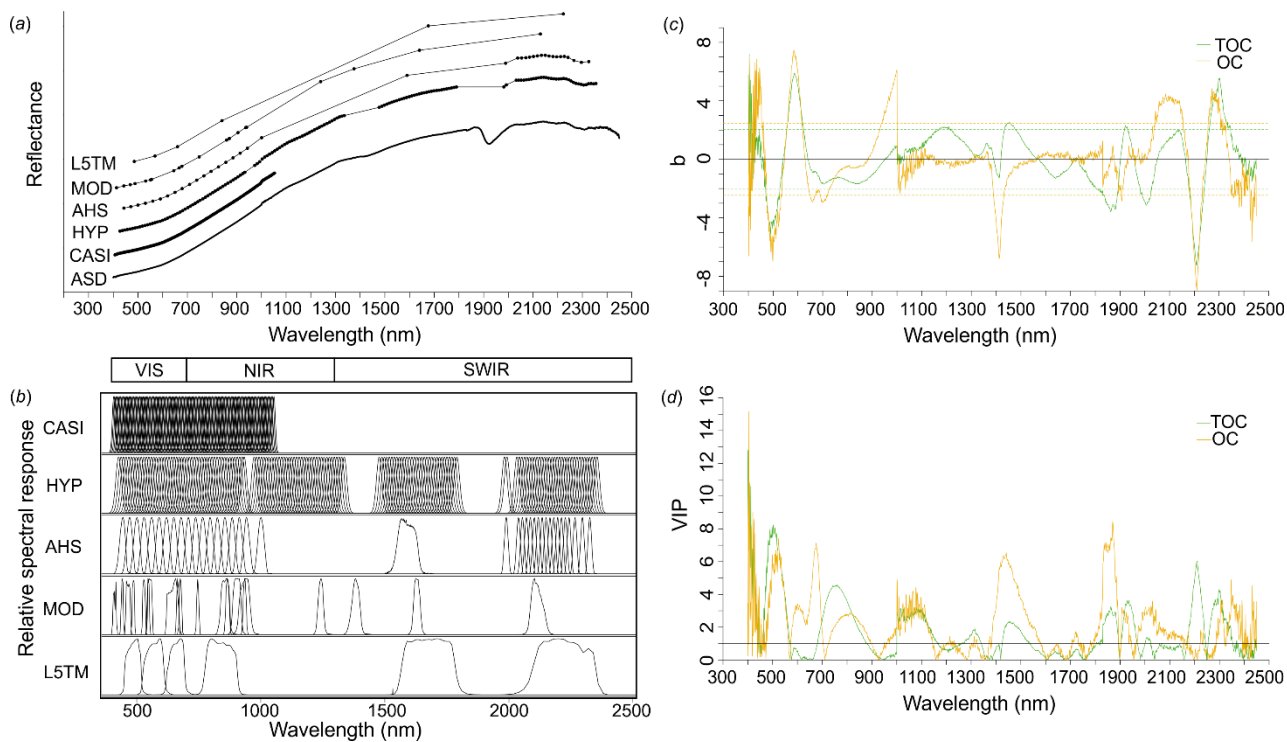
Fig. 2. Processing flow chart.

Fig. 3. Spectral resampling of the ASD reflectance spectrum of a soil sample for the different sensors (*a*). Note that, for sake of clarity, the resampled spectra plots have been given an offset. HYP: Hyperion; MOD: MODIS; L5TM: Landsat 5 TM. Band spectral responses of the sensors (*b*). Regression coefficients of PLSR models for SOC estimation based on the lab spectra measured with ASD (*c*), and plot of the Variable Importance for Projection (*d*). The dotted horizontal lines in (*c*) represent the thresholds based on the standard deviation of both TOC and OC. The black solid horizontal line in (*d*) represents the threshold for VIP, set at 1.