





## Article

# Adapting Prediction Models to Bare Soil Fractional Cover for Extending Topsoil Clay Content Mapping Based on AVIRIS-NG Hyperspectral Data

Elizabeth Baby George <sup>1</sup>, Cécile Gomez <sup>2,3,\*</sup>  and Nagesh D. Kumar <sup>4</sup> 

<sup>1</sup> Department of Civil Engineering, Indian Institute of Science, Bengaluru 560012, India; elizabethg@iisc.ac.in

<sup>2</sup> LISAH, University of Montpellier, IRD, INRAE, Institut Agro, AgroParisTech, 34060 Montpellier, France

<sup>3</sup> Indo-French Cell for Water Sciences, IRD, Indian Institute of Science, Bengaluru 560012, India

<sup>4</sup> Department of Civil Engineering and Divecha Centre for Climate Change, Indian Institute of Science, Bengaluru 560012, India; nagesh@iisc.ac.in

\* Correspondence: cecile.gomez@ird.fr

**Abstract:** The deployment of remote sensing platforms has facilitated the mapping of soil properties to a great extent. However, the accuracy of these soil property estimates is compromised by the presence of non-soil cover, which introduces interference with the acquired reflectance spectra over pixels. Therefore, current soil property estimation by remote sensing is limited to bare soil pixels, which are identified based on spectral indices of vegetation. Our study proposes a composite mapping approach to extend the soil properties mapping beyond bare soil pixels, associated with an uncertainty map. The proposed approach first classified the pixels based on their bare soil fractional cover by spectral unmixing. Then, a specific regression model was built and applied to each bare soil fractional cover class to estimate clay content. Finally, the clay content maps created for each bare soil fractional cover class were mosaicked to create a composite map of clay content estimations. A bootstrap procedure was used to estimate the standard deviation of clay content predictions per bare soil fractional cover dataset, which represented the uncertainty of estimations. This study used a hyperspectral image acquired by the Airborne Visible/Infrared Imaging Spectrometer-Next Generation (AVIRIS-NG) sensor over cultivated fields in South India. The proposed approach provided modest performances in prediction ( $R^2_{val}$  ranging from 0.53 to 0.63) depending on the bare soil fractional cover class and showed a correct spatial pattern, regardless of the bare soil fraction classes. The model's performance was observed to increase with the adoption of higher bare soil fractional cover thresholds. The mapped area ranged from 10.4% for pixels with bare soil fractional cover  $>0.7$  to 52.7% for pixels with bare soil fractional cover  $>0.3$ . The approach thus extended the mapped surface by 42.4%, while maintaining acceptable prediction performances. Finally, the proposed approach could be adopted to extend the mapping capability of planned and current hyperspectral satellite missions.

**Keywords:** composite map; clay content; digital soil mapping; uncertainty; regression model



**Citation:** George, E.B.; Gomez, C.; Kumar, N.D. Adapting Prediction Models to Bare Soil Fractional Cover for Extending Topsoil Clay Content Mapping Based on AVIRIS-NG Hyperspectral Data. *Remote Sens.* **2024**, *16*, 1066. <https://doi.org/10.3390/rs16061066>

Academic Editors: Nikolaos L. Tsakiridis, Uta Heiden and Nikolaos Tziolas

Received: 26 January 2024

Revised: 7 March 2024

Accepted: 13 March 2024

Published: 18 March 2024



**Copyright:** © 2024 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<https://creativecommons.org/licenses/by/4.0/>).

## 1. Introduction

The achievement of sustainable land management requires a comprehensive understanding of earth's soil at high spatio-temporal resolutions. The information on several soil properties such as moisture content, nutrient availability, soil organic carbon (SOC) content and water holding capacity are crucial for policymaking and land resource management for ensuring global food security [1]. Soil texture, which refers to the relative proportions of sand, silt, and clay particles in a soil mass, is widely considered as a significant soil characteristic. Soil texture impacts on plant growth, crop yield, soil water infiltration and retention capacities, and absorption of nutrients by plants [2] and its mapping is crucial over large areas for cropland management. However, conventional methods of soil analysis

are expensive and time consuming and are consequently characterized by low spatial resolutions that fail to adapt the spatial variation exhibited by soil texture in a landscape.

To address this limitation, visible–near-infrared and short-wave infrared (VNIR–SWIR, 400 nm to 2500 nm) imaging spectroscopy has been extensively used for mapping several soil properties, including soil texture [3]. Several studies have estimated the clay content of the soil from its VNIR–SWIR reflectance spectra based on a diagnostic absorption signature near 2200 nm to 2300 nm induced by the vibrations associated with the OH bond and the OH–Al–OH bonds of kaolinite, montmorillonite, and illite minerals [4]. These studies typically employ continuum removal methods [5–7], multivariate regression models [8–10], or machine learning techniques [11–13] for estimating the clay content from VNIR–SWIR reflectance spectra.

A major limitation of soil spectroscopy from remote sensing platforms is that the mapping of soil properties, including clay content mapping, is limited to bare soil surfaces. The presence of non-soil cover affects the reflectance spectra acquired over the pixels and introduces a spectral mixing effect, which further affects the soil property estimated from the spectra [14–16]. This limitation is often encountered, particularly over cultivated areas that are typically covered by photosynthetic or non-photosynthetic vegetation during most of the year.

Some studies have attempted to address this issue by employing spectral unmixing to produce new spectra that are reasonably well representative of the pixel components. This approach, which takes advantage of data with high spectral resolution, was developed on hyperspectral VNIR–SWIR airborne imagery. Bartholomeus et al. [16] used a residual spectral unmixing approach on Airborne Hyperspectral Scanner-160 data to extract pure soil spectra over partially vegetated pixels and predict SOC content. Ouerghemmi et al. [17,18] used a blind source separation technique to extract pure soil spectra over partially vegetated pixels in a Hymap image and estimate their clay contents. Franceschini et al. [19] used a progressive spectral dataset selection based on unmixed soil fractions for soil property estimation using airborne AisaDUAL data.

With the availability of multi-temporal data from multispectral sensors such as Sentinel-2 and Landsat, several recent studies extracted bare soil pixels from multiple images over a common study site and combined them to generate a new composite image with a larger number of bare soil pixels [20–24]. These multi-temporal images could either be from the same sensor or from different sensors, over which an empirical selection of a spectral index threshold is applied to differentiate bare soil cover from non-soil cover [23]. Some spectral indices are commonly used to this extent, such as the Normalized Difference Vegetation Index (NDVI) for identifying green vegetation [21,22,25], Normalized Burn Ratio 2 (NBR2) for identifying dry vegetation [21,26,27], and the Bare Soil Index (BSI) for identifying bare soil [20,23]. Dematte et al. [21] proposed to use a multi temporal data mining approach to generate a Synthetic Soil Image (SYSI) by using the median of the spectral reflectance of bare soil pixels based on the Geospatial Soil Sensing System (GEOS3). This image provides a synthetic spectral reflectance of soils, and the approach has been adopted in several similar studies [13,28–31]. These approaches, creating bare soil composite images and synthetic imagery, enabled extended soil property mapping, including clay content, SOC, and cation exchange capacity [24,30,31]. Some studies have also explored the effect of factors such as the index used in bare soil selection [32], the length of the compositing period [33], and the pixelwise mosaicking approach used for generating the composite [34] on the prediction accuracy of estimated soil properties.

In this context, while the recent literature uses a temporal perspective to extend the soil properties mapping, our study proposes a new approach from a spatial perspective to (i) extend the soil properties mapping beyond bare soil pixels based on one hyperspectral VNIR–SWIR airborne image and (ii) provide an uncertainty map of the soil property estimations. This approach first classifies the pixels based on their bare soil fractional cover from the spectral unmixing process and then applies a specific regression model to each bare soil fractional cover class to estimate clay content along with the associated

uncertainty in prediction. The selected study area is in the southern part of the Indian subcontinent, where the crop fields are typically small-sized, scattered, and farmed with diverse crop varieties.

## 2. Materials

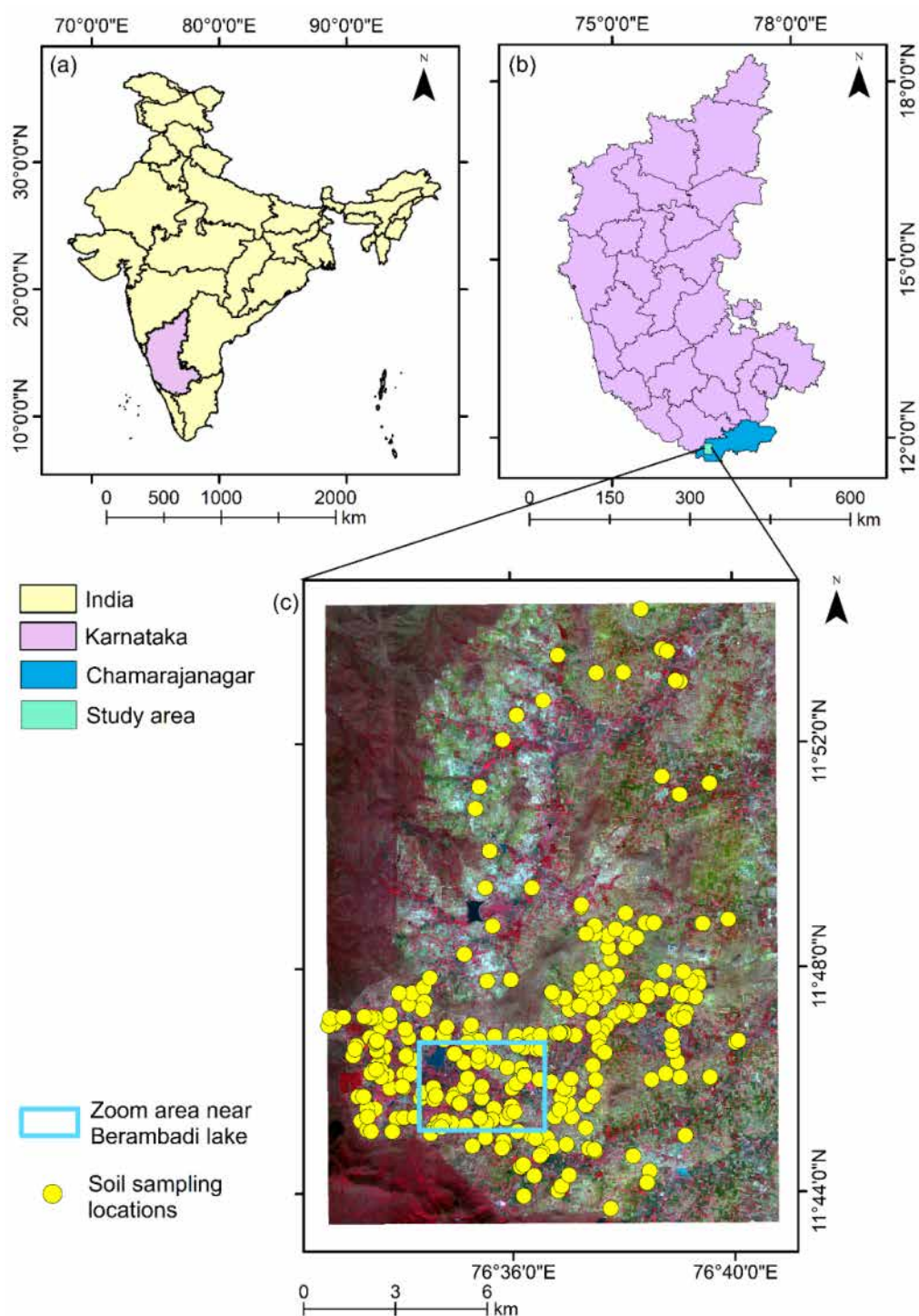
### 2.1. Study Area

The study area covers 300 km<sup>2</sup> in Gundlupet Taluk of the Indian state of Karnataka (Figure 1a). A large part of the Berambadi catchment which belongs to the Kabini Critical Zone Observatory (Assimilation of Multi-satellite data at Berambadi watershed for Hydrology And land Surface (AMBHAS), Service d'Observation—Bassins Versants Expérimentaux Tropicaux (SO-BVET), Tomer et al. [35]; Sekhar et al. [36]) and is a part of the Critical Zone Observatories: Research and Application network (OZCAR; Gaillardet et al. [37]), falls within the study area. The climate of the region is tropical sub-humid, and it receives rainfall of 800 mm annually. The study area is covered by both cultivated fields (where multi-cropping and intercropping are practiced) and forested lands. The size of the crop fields over the cultivated part varies from 0.01 to 9.3 hectares, with an average size of 1.2 hectares [38]. These are farmed with a large variety of seasonal crops, and farmers opt for irrigation using drip techniques, sprinklers, or furrow channels, depending on the availability of water and type of crop [38]. The cropping system in the region is regulated by three seasons: (1) kharif or the south-western monsoon season, from June to September, when the majority of the fields are cultivated; (2) rabi or the north-eastern monsoon, from October to January, when most of the fields are irrigated; and (3) the hot and dry summer season, from February to May, when only a few of the fields are cultivated where irrigation is viable [38].

The terrain of the study area comprises uplands with gentle slopes and lowlands that are nearly level, in addition to hills and hillslopes. The soils of the study area are mostly developed on granitic gneiss bedrock [39]. The region is characterized by red soil over uplands and hillslopes (Ferralsols and Chromic Luvisols), black soil over valley bottoms (Vertisols and Vertic intergrades), and rocky/weathered soil [39].

### 2.2. Soil Sample Collection and Clay Content Measurements

A total of 272 soil samples were collected over the study area during the rabi cropping season, with 180 samples in November 2019 and 92 in January 2020 (Figure 1c). As clay content is a perennial property in topsoil horizons, it was assumed that the clay content analyzed from soil samples collected in 2019 and 2020 and the clay content over the field in 2016 (during the AVIRIS-NG campaign) remained unchanged as previously assumed in many studies [5]. During the period of data collection, some of the crop fields in the study area were cultivated, harvested, or left bare. The strategy adopted for soil sampling was to characterize the heterogeneity of soils, crops, and landforms within the study area using stratified sampling based on both soils, landuse and geological maps, and human expertise. Since red soils located over uplands and hillslopes make up the majority of the study area [40], more samples (195 samples) were taken over them than over black soils, which are located over valley bottoms (77 samples). In order to represent the main landforms of the region, the sampling locations were dispersed throughout nearly level lowlands, moderately sloping uplands, and valley bottoms. The soil samples were collected from a depth of 0 to 15 cm. The samples thus collected were perceived to characterize the cultivated soil horizon as the process of tilling the fields prior to planting homogenizes the topsoil [41]. Each sample was a mixture of five sub-samples collected at the center and the four vertices, respectively, of a 10 m × 10 m square. These samples were assumed to be representative of an area of 5 m × 5 m, as the variation in clay content remains low within small areas [42,43]. A Garmin GPS instrument was used to record the location of each of the sampled plots.



**Figure 1.** Location of the study area in (a) the state of Karnataka in India (pink polygon) and (b) Chamarajanagar district (blue polygon) and (c) locations of the 272 topsoil samples (yellow dots) over the AVIRIS-NG False Color Composite (R-G-B as 782 nm, 662 nm, and 552 nm) with a zoom area near Berambadi lake (blue rectangle).

The clay content of the soil samples, i.e., the percentage of particles with diameter  $<0.002$  mm as per the USDA textural classification, was measured using Robinson's pipette method [44,45]. The samples were found to have clay contents that ranged from 5.8% to 62.2%, with a mean of 26.6%, a standard deviation of 13%, and a coefficient of variation of 48.8%. This soil dataset was previously used by George et al. [46].

### 2.3. AVIRIS-NG Hyperspectral Data

This work used a hyperspectral image acquired by the Airborne Visible/Infrared Imaging Spectrometer—Next Generation (AVIRIS-NG) imaging sensor over the study area on 10th of January 2016 from IST 12:00 to 13:45 h (Figure 1c). This acquisition was a part of the joint Indian Space Research Organization (ISRO)—National Aeronautics and Space Administration (NASA) hyperspectral science campaign [47] and corresponds to the image over the site ID 99, Muddur, Karnataka. The image was acquired under clear sky conditions. Since the study area had received rain 20 days prior to the date of image acquisition, the land surface can be considered dry at the time. The AVIRIS-NG data used in this study was obtained from the ISRO VEDAS portal (<https://vedas.sac.gov.in/>, accessed on 1 July 2021).

The AVIRIS-NG imaging spectrometer measures the reflected radiance in wavelengths in the 380–2510 nm range, at every 5 nm interval, thus resulting in a total of 425 spectral bands, with a ground sampling distance of 3.8 m for a flight altitude of 4.6 km and a high signal-to-noise ratio (SNR) > 2000 at 600 nm and > 1000 at 2200 nm [47]. The sensor had a field-of-view (FOV) of 34° and an instantaneous FOV (IFOV) of 1 milliradian [48,49]. The data had been ortho-rectified and atmospherically corrected using the ATmosphere REMoval algorithm (ATREM, Gao et al. [50]) to produce level 2 surface reflectance output, with values from 0 to 1 [47]. After eliminating water absorption bands and noisy bands, 260 spectral bands were retained for the modeling.

The images corresponding to nine flight lines were mosaicked to encompass the study area (Figure 1c). Water bodies in the study area were masked using an expert-calibrated threshold: pixels with a reflectance of less than 8% at 1663.67 nm (Band 258) were removed [43]. In addition, urban areas, forested areas towards the western end of the study area, and uncultivated hills were identified by visual inspection and field knowledge and were masked, and only the cultivated areas were kept for this work. This AVIRIS-NG data was previously described and used in George et al. [46].

## 3. Methods

The study proposed an approach to extend the soil properties mapping over non-soil pixels and provide an associated uncertainty map. The proposed approach first estimated the bare soil fractional cover of each pixel by spectral unmixing and then classified the pixels according to their bare soil fractional cover (Figure 2a). Specific regression models were then built for datasets with varying bare soil fractional cover thresholds to estimate the clay content associated with each pixel in the dataset (Figure 2b). A bootstrap procedure was used to estimate the standard deviation of clay content predictions per bare soil fractional cover dataset, which represented the uncertainty of estimations. Each regression model was then applied to the pixels belonging to the corresponding bare soil fractional cover class, providing a clay content map for the class (Figure 2c). Finally, the clay content maps created for each bare soil fractional cover class were mosaicked to create a composite map of clay content estimations, and the standard deviation of clay content predictions obtained through bootstrapping was used to create the associated uncertainty map (Figure 2d).

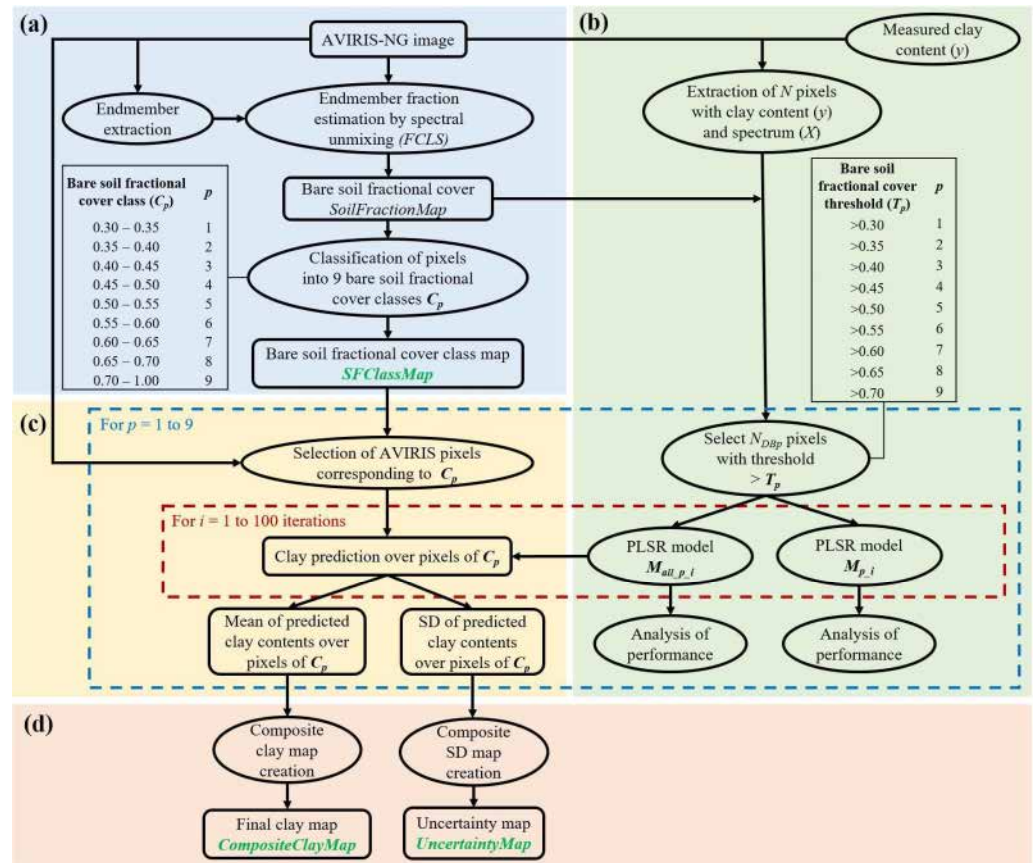
### 3.1. Spectral Unmixing for Bare Soil Fractional Cover Estimation

The spectral unmixing of this AVIRIS-NG image has been previously conducted by George et al. [46] (Figure 2a). Spectral unmixing is a process of deconvoluting the reflectance spectrum of a mixed pixel into its constituent pure component (also called ‘endmember’) spectra and the corresponding fractional covers of fractions (also called ‘abundances’) [51]. Based on the assumption that reflectance spectra of endmembers are linearly independent, the fully constrained least squares unmixing (FCLS) algorithm [52] was used to estimate the abundance of each endmember within each pixel [53]. The linear mixture model for a pixel with  $n$  spectral bands and  $p$  endmembers can be represented as:

$$y_i = M\alpha + \epsilon \quad (1)$$

where  $y_i$  is  $n \times 1$  spectral signature of the pixel  $i$ ,  $M = [M_1, \dots, M_p]$  is an  $n \times p$  matrix of endmember spectral reflectances,  $\alpha$  is the  $p \times 1$  abundance vector, and  $\epsilon$  is the measurement error [52]. The endmember abundances are then estimated by adopting the least squares technique to minimize the error  $\|y - M\alpha\|$  as follows:

$$\alpha = (M^T M)^{-1} M^T y \quad (2)$$



**Figure 2.** The workflow of the methodology with (a) generation of *SFClassMap*, (b) building the regression models  $M_p$  and  $M_{all\_p}$ , (c) application of the models on the AVIRIS-NG image, and (d) creation of the *CompositeClayMap* and *UncertaintyMap* (“SD” stands for standard deviation).

Further, constraints were imposed for the estimated endmember abundances within a pixel to add up to unity and to be non-negative [52]:

$$\alpha_j \geq 0 \quad : \quad \forall p \quad (3)$$

$$\sum_1^p \alpha_j = 1 \quad (4)$$

Since the study area consisted of crop fields with actively growing vegetation or harvested fields that were left bare or had dry crop residue cover, George et al. [46] assumed that the AVIRIS-NG spectra could be modeled based on the three following endmembers: soil, photosynthetic vegetation (PV), and non-photosynthetic vegetation (NPV). The endmembers were derived directly from the AVIRIS-NG image as proposed by Plaza et al. [54], based on visual examination of the image, ground truth knowledge, and examining the reflectance spectra. Eight pure pixels belonging to each of the three classes (soil, PV, and NPV) were identified, and their respective average reflectances were considered the endmember spectra for unmixing. Finally, the map of the bare soil fractional cover, with values in pixels varying from 0 to 1, where 0 indicates no bare soil coverage and

1 indicates full bare soil coverage within a pixel, was produced over the study area [46]. The pixels in the bare soil fractional cover map were subsequently classified into nine classes  $C_p$  ranging from 0.3 to 1 (with  $p$  ranging from 1 to 9, Figure 2a), providing a map called *SFClassMap* (Supplementary Material Figure S1). The pixels with a bare soil fractional cover below 0.3 were not considered for further analysis as they corresponded to the lowest bare soil coverage, which implied the densest vegetation coverage (NPV or PV), and were subsequently masked in the *SFClassMap*.

### 3.2. PLSR Model

The Partial Least Squares Regression (PLSR) model was the quantitative statistical analysis technique used for estimating the clay content (response variable) from the AVIRIS-NG spectra (predictor variable). PLSR extracts latent predictor variables, which account for the maximum amount of variation in the response variables [55]. This technique is commonly used for soil property estimation through spectroscopy because of its ability to analyze data with noisy and collinear variables [56–58].

Before establishing the PLSR models, the AVIRIS-NG reflectance spectra were converted to “pseudo-absorbance” ( $\log [1/\text{reflectance}]$ ) to reduce nonlinearities [56], and noise reduction was achieved using the standard pretreatments: a second-degree Savitzky–Golay filter with a span of five was applied across the discontinued AVIRIS-NG spectra [59] and a mean centering and variance scaling were applied. Spectra with a Mahalanobis distance [60] greater than 3.5 were considered outliers and removed from the calibration set. The selection of the number of latent variables (LVs) was done such that the root mean squared error of prediction for the calibration dataset estimated through leave-one-out cross validation was minimized.

### 3.3. Dataset Preparation for PLSR Training and Validation

For estimating clay content from the AVIRIS-NG image, nine datasets, denoted as  $DB_p$  (with  $p$  ranging from 1 to 9), were created. Each dataset  $DB_p$  consisted of AVIRIS-NG pixels meeting two criteria:

- (1) The corresponding clay contents measured in the laboratory were available;
- (2) The bare soil fractional cover in the pixel was up to a specified threshold  $T_p$ . This threshold  $T_p$  ranged from  $>0.3$  to  $>0.7$  in increments of 0.05 (Figure 2b).

A first set of PLSR models, denoted as  $M_p$ , was trained from subsets of  $DB_p$  using an equal number of samples ( $N_{cal}$ ) in all calibration datasets. These PLSR models  $M_p$  were then validated from subsets of  $DB_p$  using an equal number of samples ( $N_{val}$ ) in all validation datasets (where  $N_{cal}$  and  $N_{val}$  are different).  $N_{cal}$  and  $N_{val}$  were limited by the most restrictive  $DB_p$  dataset, i.e.,  $DB_9$ , which corresponded to the bare soil fractional cover greater than 0.7 and consisted of 46 samples. By adopting an 80–20% split for calibration–validation datasets [61,62], each  $M_p$  model was calibrated on 36 samples ( $N_{cal}$ ) and validated on 10 samples ( $N_{val}$ ), belonging to the corresponding dataset  $DB_p$ . This equal number of samples in both the calibration ( $N_{cal} = 36$ ) and validation ( $N_{val} = 10$ ) datasets among the models allowed a robust comparison of the effect of bare soil fractional cover on the model performance.

A second set of PLSR models, denoted as  $M_{all-p}$ , were trained using the maximum number of available samples in the  $DB_p$  dataset, keeping aside  $N_{val}$  samples for validation. Similarly, we maintained consistent sample sizes  $N_{val}$  (where  $N_{val} = 10$ ) in validation datasets across all  $M_{all-p}$  models to facilitate performance comparisons.

### 3.4. Bootstrap Process

To provide a robust performance evaluation and standard deviation of predictions, a bootstrap procedure was adopted [63] (Figure 2b). Following previous soil spectroscopic studies that had adopted bootstrapping procedures [64,65], 100  $M_p$  models were built, each one denoted  $M_{p-i}$  (with  $i$  ranging from 1 to 100) (Figure 2b). For each of the  $M_{p-i}$  models, the

selection of 36 samples ( $N_{cal}$ ) for calibration and 10 samples ( $N_{val}$ ) for validation followed stratified random sampling based on the clay contents from the corresponding  $DB_p$  dataset.

Similarly, 100  $M_{all\_p}$  models were built, each one denoted by  $M_{all\_p\_i}$  (with  $i$  ranging from 1 to 100) (Figure 2b). For each of the  $M_{all\_p\_i}$  models, the selection of 10 samples ( $N_{val}$ ) for validation followed stratified random sampling based on the clay contents from the corresponding  $DB_p$  dataset. The remaining samples in the  $DB_p$  dataset were used for calibrating the  $M_{all\_p\_i}$  models.

It is to be noted that in both configurations, the models  $M_{all\_9}$  and  $M_9$  are the same, as they are calibrated and validated on the same number of samples.

### 3.5. Performance Evaluation

The performances of each of the PLSR models were evaluated using the coefficient of determination of validation ( $R_{val}^2$ ) and the root mean square error in prediction ( $RMSEP$ ) (Figure 2b), which were calculated as follows:

$$R_{val}^2 = 1 - \frac{\sum_{j=1}^n (\tilde{y}_j - y_j)^2}{\sum_{j=1}^n (y_j - \hat{y})^2} \quad (5)$$

$$RMSEP = \sqrt{\frac{\sum_{j=1}^n (\tilde{y}_j - y_j)^2}{n}} \quad (6)$$

where  $\tilde{y}_j$  is the  $j$ th estimated clay content,  $y_j$  is the  $j$ th measured clay content,  $\hat{y}$  is the average value of measured clay contents, and  $n$  is the number of samples. Further, the mean and standard deviation of the  $R_{val}^2$  and  $RMSEP$  of the 100  $M_{p\_i}$  models from the bootstrap procedure were calculated. Similarly, the mean and standard deviation of the  $R_{val}^2$  and  $RMSEP$  from the 100  $M_{all\_p\_i}$  models from the bootstrap procedure were also calculated (Figure 2b).

### 3.6. Clay Content Composite Mapping

The model  $M_{all\_p\_i}$  trained using pixels with bare soil fractional cover greater than a threshold  $T_p$ , was applied to all pixels belonging to class  $C_p$  in the  $SFClassMap$  for estimating the corresponding clay contents (Figure 2c). Thus, 100 regression models  $M_{all\_p\_i}$  were applied to each pixel belonging to class  $C_p$ , resulting in 100 clay content estimations, whose mean value represented the final clay content estimated at the pixel (Figure 2c). A final clay content composite map named *CompositeClayMap* was then created by mosaicking the clay content estimations obtained per bare soil fractional cover class  $C_p$  (Figure 2d).

Likewise, the standard deviation of the 100 clay content estimations obtained from bootstrapping ( $M_{all\_p\_i}$ ) was calculated, which represented the prediction uncertainty. The mosaicking of these standard deviation values generated the associated uncertainty map, named *UncertaintyMap* (Figure 2d).

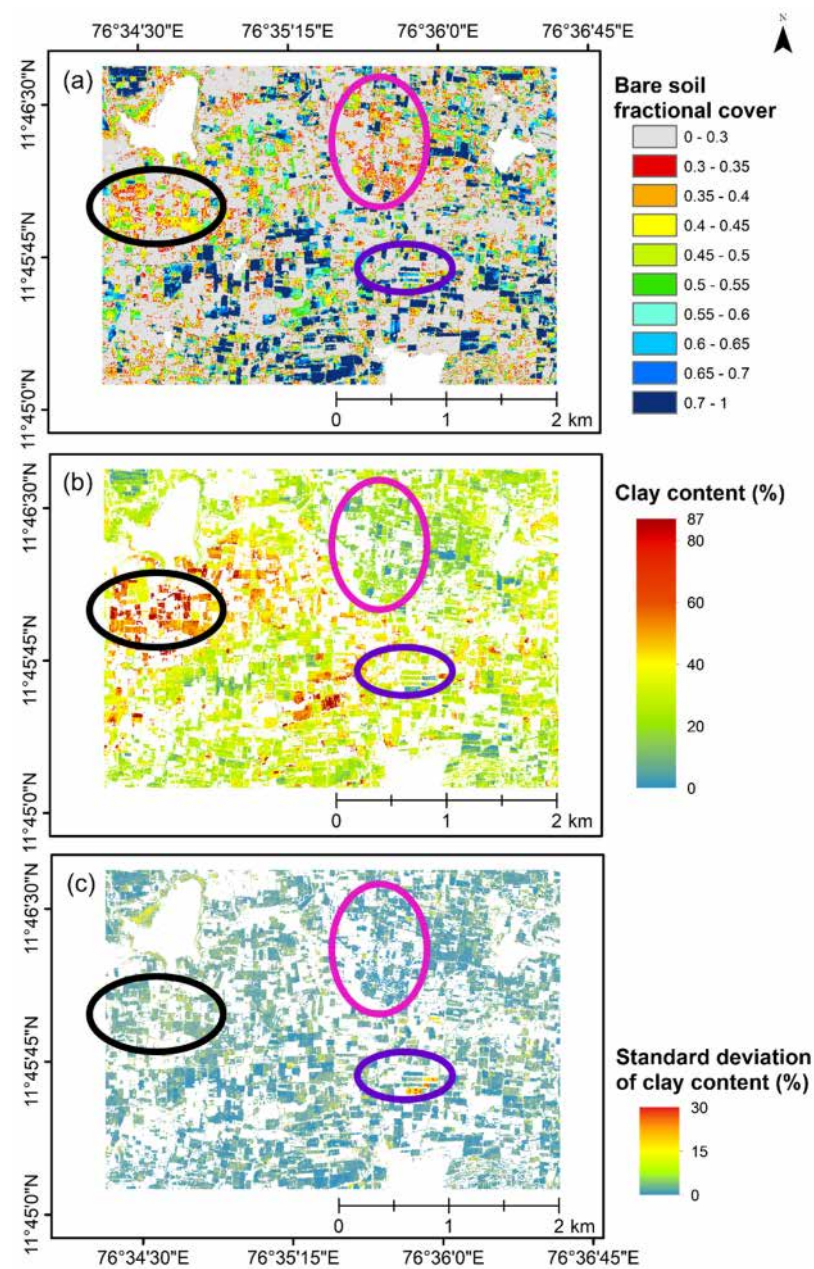
## 4. Results

### 4.1. Bare Soil Fractional Cover Class Map

The  $SFClassMap$  generated from the bare soil fractional cover values estimated through FCLS unmixing covered 52.7% of the study area (Figure 3a; Supplementary Material Figure S1). The percentage of the study area covered by the bare soil fractional cover classes varied from 2.8% for class  $C_8$  to 10.4% for class  $C_9$  (Table 1). The number of soil samples available for modeling per bare soil fractional cover class varied from 8 for  $C_7$  (bare soil fractional cover from 0.6 to 0.65) to 46 for  $C_9$  (bare soil fractional cover from 0.7 to 1) (Table 1). The number of soil samples  $N_{DB_p}$  available for modeling per bare soil fractional cover threshold ( $T_p$ ) varied from 46 for  $T_9$  (threshold > 0.7) to 167 for  $T_1$  (threshold > 0.3) (Table 2). So while 272 soil samples were initially available (Section 2.2), only 167 were finally used, as the remaining 107 samples were collected over surfaces associated with bare soil fractional



cover less than 0.3. Finally, the nine datasets  $DB_p$  (with  $p$  ranging from 1 to 9) had similar ranges and distributions of measured clay contents, regardless of the number of samples associated with them (Table 2).



**Figure 3.** (a) Bare soil fractional cover classes map (*SFClassMap*), (b) composite clay content map (*CompositeClayMap*; mean clay content from 100 iterations), and (c) the associated map of uncertainties (*UncertaintyMap*; standard deviation from 100 iterations) of a sub-part of the study area (blue rectangle in Figure 1c). The pink and black ellipses highlight areas characterized by red and black soils, respectively. The purple ellipse highlights an area with high uncertainty.

**Table 1.** Area covered by each of the bare soil fractional cover classes obtained by the FCLS unmixing.

Bare Soil Fractional Cover	Class ( $C_p$ )	Number of Available Soil Samples	Area (%)
0.30–0.35	$C_1$	17	7.3
0.35–0.40	$C_2$	18	7.4
0.40–0.45	$C_3$	20	6.9
0.45–0.50	$C_4$	14	5.9
0.50–0.55	$C_5$	13	4.9
0.55–0.60	$C_6$	14	3.9
0.60–0.65	$C_7$	8	3.2
0.65–0.70	$C_8$	17	2.8
0.70–1.00	$C_9$	46	10.4
0.30–1.00		167	52.7

**Table 2.** Details of the available soil samples associated with each bare soil fractional cover threshold  $T_p$ .

Dataset ( $DB_p$ )	Bare Soil Fractional Cover Threshold ( $T_p$ )	Number of Available Samples ( $N_{DB_p}$ )	Minimum Clay Content (%)	Maximum Clay Content (%)	Mean Clay Content (%)	Standard Deviation of Clay Content (%)	
$DB_1$	$T_1$	>0.30	167	5.81	59.61	26.21	12.72
$DB_2$	$T_2$	>0.35	150	5.83	59.61	25.70	12.53
$DB_3$	$T_3$	>0.40	132	5.83	59.61	26.10	12.61
$DB_4$	$T_4$	>0.45	112	5.83	59.61	26.70	12.50
$DB_5$	$T_5$	>0.50	98	5.83	59.61	27.12	12.43
$DB_6$	$T_6$	>0.55	85	6.96	58.91	26.60	11.47
$DB_7$	$T_7$	>0.60	71	6.96	49.04	26.01	11.16
$DB_8$	$T_8$	>0.65	63	6.96	49.04	25.48	10.99
$DB_9$	$T_9$	>0.70	46	6.96	47.30	24.91	9.88

#### 4.2. Effect of Bare Soil Fractional Cover on Clay Content Estimation

The first set of regression models  $M_p$  built from the bootstrap procedure with the same number of calibration data ( $N_{cal} = 36$ ) yielded modest to good performances (Table 3). The best performance was obtained for the model  $M_9$ , which considered the higher bare soil fractional cover (>0.7) (an  $R^2_{val}$  of 0.63 and an  $RMSEP$  of 6.13% as the mean of 100 iterations), while the worst performance was obtained for the model  $M_1$ , which considered the lowest bare soil fractional cover (>0.3) (an  $R^2_{val}$  of 0.46 and an  $RMSEP$  of 9.27% as the mean of 100 iterations) (Table 3). As expected, an increase in bare soil fractional cover in the data used for training the PLSR models involved an improvement in clay content estimation accuracy (Table 3). This increase in performance may be attributed to a better selection of bare soil pixels by the stricter threshold, which implies a better selection of soil spectra unaffected by vegetation.

**Table 3.** Performances of  $M_p$  models for clay content estimation.

Bare Soil Fractional Cover	Dataset ( $DB_p$ )	PLSR Model ( $M_p$ )	$N_{cal}$	$N_{val}$	Validation (Mean $\pm$ Standard Deviation over 100 Iterations)	
					$R^2_{val}$	$RMSEP$ (%)
>0.30	$DB_1$	$M_1$	36	10	0.46 $\pm$ 0.22	9.27 $\pm$ 2.00
>0.35	$DB_2$	$M_2$	36	10	0.47 $\pm$ 0.20	9.16 $\pm$ 2.23
>0.40	$DB_3$	$M_3$	36	10	0.49 $\pm$ 0.20	9.01 $\pm$ 1.90
>0.45	$DB_4$	$M_4$	36	10	0.49 $\pm$ 0.19	9.04 $\pm$ 1.95
>0.50	$DB_5$	$M_5$	36	10	0.50 $\pm$ 0.19	8.63 $\pm$ 1.69
>0.55	$DB_6$	$M_6$	36	10	0.50 $\pm$ 0.19	8.04 $\pm$ 1.90
>0.60	$DB_7$	$M_7$	36	10	0.52 $\pm$ 0.18	7.53 $\pm$ 1.51
>0.65	$DB_8$	$M_8$	36	10	0.59 $\pm$ 0.17	7.00 $\pm$ 1.41
>0.70	$DB_9$	$M_9^*$	36	10	0.63 $\pm$ 0.13	6.13 $\pm$ 1.09

\*  $M_9 = M_{all\_9}$ .

The standard deviation of  $R_{val}^2$  from the 100 iterations ranged from 0.13 for a bare soil fractional cover higher than 0.7 to 0.22 for a bare soil fractional cover higher than 0.3 (Table 3). The standard deviation of  $RMSEP$  from the 100 iterations ranged from 1.09% for a bare soil fractional cover exceeding 0.7 to 2.23% for a bare soil fractional cover higher than 0.35 (Table 3). Finally, an increase in the bare soil fractional cover threshold involved a decrease in the standard deviation of clay content predictions (Table 3).

The second set of regression models  $M_{all_p}$  was trained from the maximum number of available samples ( $N_{cal\_all_p}$ ) per bare soil fractional cover threshold, keeping aside  $N_{val}$  samples for validation (where  $N_{val} = 10$ ). Therefore, the number of calibration samples ( $N_{cal\_all_p}$ ) ranged from 36 to 157 (Table 4). These models also yielded modest to good performances, with  $R_{val}^2$  (as the mean of 100 iterations) from 0.53 (obtained for bare soil fractional cover  $>0.3$ ,  $>0.45$ , and  $>0.5$ ), to 0.63 (for bare soil fractional cover  $> 0.7$ ) and  $RMSEP$  (as the mean of 100 iterations) from 6.13% (for bare soil fractional cover  $> 0.7$ ) to 8.30% (for bare soil fractional cover  $> 0.3$ ) (Table 4). As observed with the  $M_p$  models, the performance of  $M_{all_p}$  in clay content estimation improved with an increase in bare soil fractional cover.

**Table 4.** Performances of  $M_{all_p}$  models for clay content estimation.

Bare Soil Fractional Cover	Dataset ( $DB_p$ )	PLSR Model ( $M_{all_p}$ )	$N_{cal\_all_p}$	$N_{val}$	Validation (Mean $\pm$ Standard Deviation over 100 Iterations)	
					$R_{val}^2$	$RMSEP$ (%)
$>0.30$	$DB_1$	$M_{all\_1}$	157	10	$0.53 \pm 0.20$	$8.30 \pm 1.96$
$>0.35$	$DB_2$	$M_{all\_2}$	140	10	$0.54 \pm 0.21$	$8.20 \pm 2.00$
$>0.40$	$DB_3$	$M_{all\_3}$	122	10	$0.54 \pm 0.18$	$8.19 \pm 1.78$
$>0.45$	$DB_4$	$M_{all\_4}$	102	10	$0.53 \pm 0.19$	$8.19 \pm 1.72$
$>0.50$	$DB_5$	$M_{all\_5}$	88	10	$0.53 \pm 0.21$	$7.90 \pm 1.86$
$>0.55$	$DB_6$	$M_{all\_6}$	75	10	$0.55 \pm 0.17$	$7.26 \pm 1.72$
$>0.60$	$DB_7$	$M_{all\_7}$	61	10	$0.59 \pm 0.14$	$6.97 \pm 1.29$
$>0.65$	$DB_8$	$M_{all\_8}$	53	10	$0.61 \pm 0.16$	$6.50 \pm 1.37$
$>0.70$	$DB_9$	$M_{all\_9}^*$	36	10	$0.63 \pm 0.13$	$6.13 \pm 1.09$

\*  $M_{all\_9} = M_9$ .

For the  $M_{all_p}$  models, the standard deviation of  $R_{val}^2$  from the 100 iterations ranged from 0.13 for bare soil fractional cover  $> 0.7$  to 0.21 for bare soil fractional cover  $>0.35$  and  $>0.5$  (Table 4). The standard deviation  $RMSEP$  for the  $M_{all_p}$  models from the 100 iterations ranged from 1.09% for bare soil fractional cover  $> 0.7$  to 2.00% for bare soil fractional cover  $> 0.35$  (Table 4). Finally, an increase in bare soil fractional cover involved a decrease in the standard deviation of clay content estimation, as observed for the  $M_p$  models (Tables 3 and 4).

The  $M_{all_p}$  models thus exhibited a similar trend in PLSR model performance for clay content estimation as the  $M_p$  models, i.e., an increase in  $R_{val}^2$  and a decrease in  $RMSEP$  as the bare soil fractional cover increased. It was also observed that the  $M_{all_p}$  models exhibited better performances than the  $M_p$  models for the same  $p$  (Tables 3 and 4). Thus, using a greater number of samples for training a PLSR model improved its performance in clay content estimation.

#### 4.3. Clay Content Composite Map and Associated Uncertainties Map

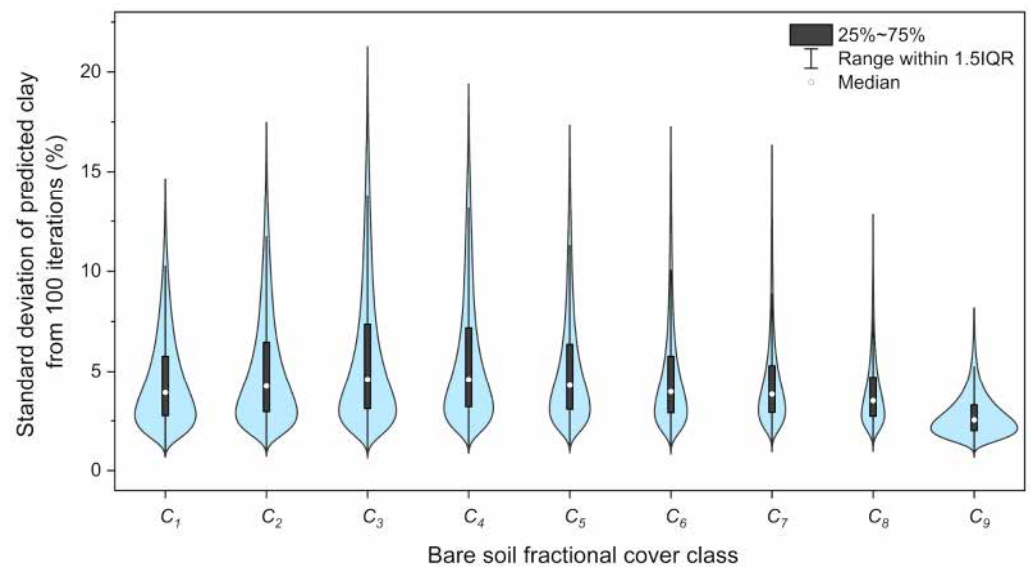
To take advantage of the maximum number of soil samples in regression models and benefit from the best model performances, the  $M_{all_p-i}$  regression models were used for clay content mapping. The 100 regression models  $M_{all_p-i}$  were applied to generate 100 clay content estimations for each pixel in the AVIRIS-NG image, based on their bare soil fractional cover class  $C_p$  (Figure 2c). The mean of these 100 values was reported as the clay content at the pixel, and the mosaicking of these estimations produced the clay content composite map, *CompositeClayMap* (Supplementary Material Figure S2). This approach allowed mapping over 52.7% of the study area, whereas a single model  $M_{all\_9}$  would have allowed mapping over only 10.4% of the study area (Table 1).

The spatial pattern of the *CompositeClayMap* was coherent with the soilscape trend of the region. High clay content values were estimated over the black soils corresponding to Vertisol and Vertic intergrades, found in the valley bottom around the riverbeds (with a mean clay content of around 40%), while lower clay contents were estimated over the red soils corresponding to Ferralsol and Chromic Luvisols, found in the uplands and hillslopes (with a mean clay content < 40%). The region towards the east of the Berambadi lake that is characterized by red soils was rightly identified with 10 to 30% of the estimated clay content (Figure 3b, pink ellipse). Similarly, the black soils present to the south of the lake were identified with higher clay contents of around 40 to 60% in the *CompositeClayMap* (Figure 3b, black ellipse). These observations were consistent with the pedological map and knowledge of the study area. The clay content composite map also compared well with the soil texture classification by Gomez et al. [66] and clay content mapping by George et al. [46] over the study area using Sentinel-2 and AVIRIS-NG images, respectively.

This coherent spatial pattern for clay content was observed from *CompositeClayMap*, regardless of the bare soil fractional cover classes (Figure 3a,b). Even the lower bare soil fractional cover classes ( $C_1$  to  $C_6$ ) with relatively higher *RMSEP* and lower  $R^2_{val}$  were able to discriminate between the red and black soil in the study area. This implied that even over low bare soil fractional cover, the PLSR models were able to provide the correct spatial pattern of predicted clay values, which can be useful for supplementing the pedological knowledge at the regional scale.

The standard deviation of the predicted clay contents over 100 iterations, which represents the uncertainty of prediction in a pixel over the entire area, had a mean of 4.62% (Supplementary Material Figure S3). In these, 0.01% of the pixels spread across the study area without exhibiting any spatial pattern were associated with high uncertainties (>30%) and considered outliers and excluded from the analysis. Over the study area, pixels associated with high uncertainties (standard deviation from 15 to 30%) were located over the forested region to the north-west (Supplementary Material Figure S1), where bare soil fractional cover was estimated to vary from 0.3 to 0.45, which could possibly be an overestimation of the bare soil fraction in this area. Over the cultivated area, both red and black soils (Figure 3b, pink and black ellipses, respectively) were associated with lower uncertainties (standard deviation < 10%), except over some specific agricultural fields that had high uncertainties (e.g., Figure 3c, purple ellipse), which indicated that clay estimations over these specific fields must be considered with caution.

The mean values of the uncertainty in prediction associated with each bare soil fractional cover class varied from 2.81% for  $C_9$  to 5.78% for  $C_3$ , while the median varied from 2.54% for  $C_9$  to 4.6% for  $C_3$  and  $C_4$  (Figure 4). Violin plots for each bare soil fractional cover class  $C_p$  were generated to compare the distribution of the uncertainties associated with the class (Figure 4). The higher bare soil fractional cover class  $C_9$  was associated with a lower range of uncertainty as compared to the lower bare soil fractional cover classes (Figure 4). Finally, though there was a statistically significant difference between the uncertainties of bare soil fractional cover classes as determined by one-way ANOVA ( $F(8, 1.1 \times 10^7) = 188,787.14, p < 0.0001$ ), there was no correlation between the bare soil fraction classes (Figure 3a) and the standard deviations (Figure 3c) of predictions over the entire mapped study area ( $R^2$  of 0.08,  $p < 0.0001$ ).



**Figure 4.** Standard deviation of clay content estimations (from 100 iterations) in the bootstrap procedure. The thin line, bar, and white dot show the 95% confidence interval, interquartile range, and median, respectively.

## 5. Discussion

### 5.1. Bare Soil Fractional Cover Estimation from Spectral Unmixing

The choice of a suitable collection of endmembers may have a significant impact on the result of spectral unmixing [54]. The endmembers used for unmixing can be selected either from a spectral library of known materials (acquired from a laboratory or field) or directly from the image [67]. In this study, the three endmembers (soil, PV, and NPV) were derived from the AVIRIS-NG image since the spectra of these image endmembers would have the same spatial and spectral resolutions and similar acquisition conditions as the unmixing input image [54].

Estimating the number of endmembers to be used for unmixing is an application dependent and subjective task [68]. Incorporating any of the available endmember variability reduction techniques, such as iterative mixture analysis, spectral feature selection, spectral transformations, etc., could account for the temporal and spatial variability between and among endmembers [69]. While the study area encompasses cultivated lands featuring a variety of crops at different growth stages, leading to intra-class spectral variability for vegetation endmembers (PV and NPV), we have simplified their representation by assigning each a single spectral signature. Similarly, despite the Berambadi catchment being characterized by both red and black soil, a singular soil endmember was chosen. It would be worth exploring the inclusion of two soil endmembers (black soil and red soil) and multiple PV and NPV endmembers in the unmixing process to better capture the intra-class spectral variability across this complex and heterogeneous study area.

### 5.2. Model Performance over the Highest Bare Soil Fractional Cover

The PLSR models  $M_9$ , calibrated using samples corresponding to the highest bare soil fractional cover class ( $>0.7$ ), provided accurate predictions for topsoil clay content (Tables 3 and 4). These model performances fell within the range reported in the literature dedicated to bare soil pixels ( $R^2$  from 0.53 to 0.75). It slightly surpassed the performance for clay content estimation ( $R^2_{val}$  of 0.55) achieved by Casa et al. [8] using PLSR modeling based on airborne hyperspectral MIVIS data. It also surpassed the performance for clay content estimation ( $R^2_{val}$  from 0.39 to 0.57) achieved by Castaldi et al. [9] using PLSR modeling and resampled spectra from various hyperspectral satellite sensors, including EO-1 Hyperion, PRISMA, HypIRI, and EnMAP, with the addition of noise and atmospheric effects. However, this performance was slightly lower compared to some studies employing

the PLSR model and concentrating on bare soil pixels, such as Selige et al. [70] reporting an  $R_{cv}^2$  of 0.71 from HyMap, Gomez et al. [43] reporting an  $R_{val}^2$  of 0.75 from AisaDual data, and Nouri et al. [10] reporting an  $R_{test}^2$  of 0.72 from HyMap data. Finally, it was in accordance with the performance obtained by Gomez et al. [5], reporting an  $R_{cv}^2$  of 0.64 using HyMap hyperspectral data for clay content estimation.

The large range of performances of topsoil clay content prediction using hyperspectral airborne sensors reported in the literature could be attributed to several factors, such as the difference in the number of soil samples used to train the regression models [11], the soil surface roughness and moisture [23], the pedological context and soil diversity in each study area, the residual PV or NPV cover on the selected bare soil pixels [3], and the variability in spatial resolution [43].

### 5.3. Model Performances Depending on Bare Soil Fractional Cover

The performance in clay content estimations decreased slightly in terms of  $R_{val}^2$  and  $RMSEP$  as the bare soil fractional cover decreased for both  $M_p$  and  $M_{all_p}$  models (Tables 3 and 4). This result is consistent with the literature, as a better selection of bare soil pixels by the stricter threshold would ensure lesser spectral influence from non-soil components [23].

The  $M_{all_p}$  models exhibited satisfactory performances even at lower bare soil fractional cover, with  $R_{val}^2$  consistently exceeding 0.5. This could be attributed to the utilization of the maximum number of soil samples by these models in their calibration. These results were consistent with those of Dvorakova et al. [71], where models predicting SOC showed only slight variations across a Cellulose Absorption Index (CAI) threshold ranging from 0.00 to 3.00. In that study, the  $R^2$  values increased from a CAI threshold of 0.00 to reach the maximum at 0.75 ( $R^2 = 0.59$ ), then decreased until a threshold of 1.25, and finally remained relatively stable for CAI thresholds between 1.50 and 3.00 ( $R^2$  around 0.47) [71]. However, this stability in performance contrasts with the observations by Bartholomeus et al. [16], who conducted simulations on the impact of PV on SOC estimation over maize fields and reported significant inaccuracies beyond 5% of PV cover. Similarly, Ouerghemmi et al. [17] reported similar results for clay content estimation over vineyards, noting a significant decrease in accuracy beyond 15% of PV cover.

Finally, following the work of Dutta et al. [11] analyzing the impact of the number of samples in the calibration dataset on model performance or the work of Vaudour et al. [23] analyzing the impact of soil roughness and moisture on soil property predictions, future research could evaluate the impact of these factors in the proposed approach. It is also to be noted that while each model  $M_{all_p}$  was built using pixels of classes  $C_p$  to  $C_9$  to obtain the best possible performance for clay content modeling, it was applied only to pixels of the unique class  $C_p$ . Increasing the spatial sampling density to have more soil samples per bare soil fractional cover class would allow extensive studies on the proposed composite mapping approach by building and testing regression models on each of the bare soil fractional cover class  $C_p$ .

### 5.4. Clay Content Composite Mapping Approach

The accuracy of the spatialized clay predictions in this study with  $R_{val}^2$ , ranging from 0.53 to 0.63 depending on the bare soil fractional cover classes (Table 4), fell within the range of accuracy obtained from bare soil composite maps in the literature. For example, Gasmi et al. [72] obtained  $R_{val}^2$  of 0.64 for clay prediction using a bare soil composite map based on reflectance median along multispectral Landsat-TM time-series images, while Sorenson et al. [31] obtained  $R_{val}^2$  of 0.44 for clay prediction using a bare soil composite map based on Landsat-5 time-series images combined with legacy data.

Despite our modest prediction accuracies, the proposed clay content composite map showed a correct soil pattern, regardless of the bare soil fractional cover classes, where clay-rich areas were identified over black soils found in the valley bottoms around the riverbeds and low clay content areas were identified over red soils found in the uplands

and hillslopes. While exercising caution in the interpretation of clay predictions at the local scale, such as the field level, our model demonstrates its utility in providing detailed and nuanced soil pattern recognition at the regional scale. This is particularly evident in its ability to discriminate between black and red soils, offering valuable insights into the heterogeneity of the soil landscape.

The standard deviation of clay predictions along the bootstrap process was considered an indicator of the prediction uncertainty (Figure 3c). The bootstrap, as proposed by Breiman [73], measures the uncertainty of predictions by generating several models based on different subsets that are simulated from a single dataset. Our approach to estimating and mapping the uncertainties of prediction follows the work of Brodský et al. [74], Gomez et al. [43], and more recently, Castaldi et al. [34]. Brodský et al. [74] proposed a stochastic simulation (100 realizations) to assess the uncertainty as the standard deviation of predictions, Gomez et al. [43] proposed a bootstrap procedure (999 iterations) to obtain the pixel-wise variance of predictions. Castaldi et al. [34] proposed to estimate the pixel-wise uncertainty as the ratio between the standard deviation and the mean of predictions. While both standard deviation and variance reflect the variability in a distribution, the standard deviation is expressed in the same unit as the original data (% of clay content, in this case), which can be useful for evaluating both the prediction and uncertainty maps. As these prediction uncertainties were uncorrelated to the bare soil fractional cover (Figure 3a,c), they cannot be attributed to the bare soil fractional cover. The prediction uncertainties could also be related to soil roughness and soil moisture, which may impact the reflectance spectra and, in turn, the predictions, as shown by Ben-Dor et al. [75], Wu et al. [76], and Denis et al. [77]. In the context of this study, the soil surface roughness and soil moisture might be highly variable as they depend on the type and intensity of the irrigation and tillage operations, respectively, and the study area is characterized by small sized crop fields and frequent crop rotation throughout the year [38]. So, as demonstrated by Vaudour et al. [23] and Urbina-Salazar et al. [33], future studies would have to consider the influence of several spectral perturbing factors and environmental parameters that would affect the soil property to be estimated, including soil surface characteristics such as soil moisture and roughness.

##### *5.5. Comparison with Other Approaches to Extend Predictions*

To overcome the issue of spectral mixing in imaging soil spectroscopy, several approaches have been proposed to extend the mapping capability over bare soil pixels. A first approach consists of (i) extracting the spectrum of “pure soils” from a mixed spectrum of soil and vegetation and then (ii) using it to predict the targeted soil properties [16–18]. A second approach consists of (i) creating a bare soil composite map from multi-temporal data, acquired from a single sensor such as Sentinel-2 [23,33] or Landsat [20,78] or multiple sensors [24,30] and then (ii) using it to predict the targeted soil properties [23,24]. These studies synthesize bare soil composites using the median or mean spectral reflectance of bare soil pixels identified using spectral indices like NDVI, BSI, and NBR2 [20–22,33,72]. While most of these studies used multispectral satellite data to create a bare soil composite map from multi-temporal data, Diek et al. [79] produced a bare soil composite map from multi-temporal hyperspectral data acquired from the Airborne Prism EXperiment (APEX) imaging sensor.

In contrast, our study used a single AVIRIS-NG image and thus benefited from the relatively finer spatial and spectral resolution inherent in airborne hyperspectral data. This finer spatial resolution of AVIRIS-NG data (4 m) is useful for mapping soil properties over small-sized fields with highly diverse vegetation types, whereas the spatial resolution of Sentinel-2 or Landsat data (10–30 m) might present limitations in capturing soil pixels in such fragmented landscapes. Additionally, the finer spectral resolution of the hyperspectral AVIRIS-NG data (5 nm) is useful in identifying dry vegetation (which is often a major component covering crop fields, especially after harvest), while the limited number of spectral bands and coarse spectral resolution of Sentinel-2 or Landsat may be insufficient [80].

Finally, airborne hyperspectral data have limited availability due to the expenses associated with the flight and the requirement of clear sky conditions for airborne platforms. The availability of data from new generation spaceborne hyperspectral sensors like the Italian PRecursores IperSpettrale della Missione Applicativa (PRISMA) [81], American Hyperspectral InfraRed Imager (HyspIRI) [82], Japanese Hyperspectral Imager Suite (HISUI) [83], and German Environmental Mapping and Analysis Program (EnMAP) [84] could provide multi-temporal hyperspectral observations. Combining our proposed approach with the ones based on bare soil composite images using multi-temporal data [23,71,79] could significantly increase soil property estimation coverage and help provide associated prediction uncertainties.

## 6. Conclusions

This study proposed a novel clay content composite mapping approach for extending topsoil clay content from an airborne AVIRIS-NG hyperspectral image over a heterogeneous agricultural study area. The approach classified the image pixels according to their bare soil fractional cover determined by spectral unmixing and assigned specific regression models to each bare soil fractional cover class to estimate the clay content along with the uncertainty in prediction. It was observed that the model performance increases with the adoption of higher bare soil fractional cover thresholds in terms of an increase in  $R^2$  and a decrease in  $RMSEP$  values. Despite the modest prediction accuracy, the clay content composite map showed a correct spatial pattern regardless of the bare soil fractional cover classes. The potential mapped area in terms of clay content in this study ranged from 10.4% to 52.7% with a bare soil fraction threshold of  $>0.7$  and  $>0.3$ , respectively, and the compositing approach allowed an extension of the mapped surface by 42.42%. The proposed approach could be adopted to extend the mapping capability of planned and current hyperspectral satellite missions that have relatively coarser spatial resolutions, which amplifies the spectral mixing problem. In such cases, a combination of the proposed approach and multi-temporal images could significantly increase both the extent of the mapped area and associated prediction performances. Further, additional research will be needed to mitigate the influence of factors that affect the spectral reflectance of soils, such as moisture and surface roughness, so as to improve the prediction performance.

**Supplementary Materials:** The following supporting information can be downloaded at: <https://www.mdpi.com/article/10.3390/rs16061066/s1>, Figure S1: The map of bare soil fractional cover classes using spectral unmixing (*SFClassMap*) for the entire study area. Figure S2: The clay content composite map *CompositeClayMap* (mean clay content from 100 iterations) for the entire study area. Figure S3: The map of uncertainties associated with the clay content estimation *UncertaintyMap* (standard deviation from 100 iterations) for the entire study area.

**Author Contributions:** Conceptualization, E.B.G., C.G. and N.D.K.; methodology, C.G. and N.D.K.; software, E.B.G.; formal analysis, E.B.G.; resources, C.G. and N.D.K.; data curation, E.B.G.; writing—original draft preparation, E.B.G.; writing—review and editing, C.G. and N.D.K.; supervision, N.D.K.; funding acquisition, C.G. and N.D.K. All authors have read and agreed to the published version of the manuscript.

**Funding:** This research received no external funding.

**Data Availability Statement:** Most data presented in this study are available upon reasonable request from the corresponding author. The AVIRIS-NG image used in this study can be downloaded from the ISRO VEDAS portal at <https://vedas.sac.gov.in/> (accessed on 1 July 2021).

**Acknowledgments:** The authors are indebted to ICAR—NBSS & LUP Bangalore, India, for soil sample collection and laboratory analysis. This research was supported by the project ATCHA ANR-16-CE03-0006 and the Programme National de Télédétection Spatiale (PNTS, <http://www.insu.cnrs.fr/pnts>, accessed on 1 April 2020), grant no. PNTS-2019-5. The first author wishes to acknowledge the Grantham Fellowship provided by the Divecha Centre for Climate Change, IISc, Bangalore, to conduct this research. The Kabini Critical Zone Observatory (AMBHAS, BVET, Sekhar et al. [36]; Tomer et al. [35]; [www.ambhas.com](http://www.ambhas.com); <https://mtropics.obsmp.fr/>, accessed on 1 September 2021), which is part of the



OZCAR network (Gaillardet et al. 2018; <http://www.ozcar-ri.org/ozcar/>, accessed on 1 September 2021), is also acknowledged.

**Conflicts of Interest:** The authors declare no conflicts of interest.

## References

- Mulder, V.L.; De Bruin, S.; Schaepman, M.E.; Mayr, T.R. The use of remote sensing in soil and terrain mapping—A review. *Geoderma* **2011**, *162*, 1–19. [[CrossRef](#)]
- Brown, R.B. *Soil Texture [Fact Sheet]*; University of Florida Institute of Food and Agricultural Sciences: Gainesville, FL, USA, 2003.
- Chabrilat, S.; Ben-Dor, E.; Cierniewski, J.; Gomez, C.; Schmid, T.; van Wesemael, B. Imaging spectroscopy for soil mapping and monitoring. *Surv. Geophys.* **2019**, *40*, 361–399. [[CrossRef](#)]
- Chabrilat, S.; Goetz, A.F.; Krosley, L.; Olsen, H.W. Use of hyperspectral images in the identification and mapping of expansive clay soils and the role of spatial resolution. *Remote Sens. Environ.* **2002**, *82*, 431–445. [[CrossRef](#)]
- 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. [[CrossRef](#)]
- Lagacherie, P.; Baret, F.; Feret, J.B.; Netto, J.M.; 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. [[CrossRef](#)]
- 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. [[CrossRef](#)]
- 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*, 17–26. [[CrossRef](#)]
- Castaldi, F.; Palombo, A.; Santini, F.; Pascucci, S.; Pignatti, S.; Casa, R. Evaluation of the potential of the current and forthcoming multispectral and hyperspectral imagers to estimate soil texture and organic carbon. *Remote Sens. Environ.* **2016**, *179*, 54–65. [[CrossRef](#)]
- Nouri, M.; Gomez, C.; Gorretta, N.; Roger, J.M. Clay content mapping from airborne hyperspectral Vis-NIR data by transferring a laboratory regression model. *Geoderma* **2017**, *298*, 54–66. [[CrossRef](#)]
- Dutta, D.; Goodwell, A.E.; Kumar, P.; Garvey, J.E.; Darmody, R.G.; Berretta, D.P.; Greenberg, J.A. On the feasibility of characterizing soil properties from aviris data. *IEEE Trans. Geosci. Remote Sens.* **2015**, *53*, 5133–5147. [[CrossRef](#)]
- Liu, L.; Ji, M.; Buchroithner, M. Transfer learning for soil spectroscopy based on convolutional neural networks and its application in soil clay content mapping using hyperspectral imagery. *Sensors* **2018**, *18*, 3169. [[CrossRef](#)] [[PubMed](#)]
- Bellinaso, H.; Silvero, N.E.; Ruiz, L.F.C.; Amorim, M.T.A.; Rosin, N.A.; de Sousa, M.W.; de Sousa, G.P.B.; Sepulveda, L.M.A.; de Queiroz, L.G.; Nanni, M.R.; et al. Clay content prediction using spectra data collected from the ground to space platforms in a smallholder tropical area. *Geoderma* **2021**, *399*, 115116. [[CrossRef](#)]
- Demattê, J.A.M.; Huete, A.R.; Ferreira, L.G.; Nanni, M.R.; Alves, M.C.; Fiorio, P.R. Methodology for bare soil detection and discrimination by Landsat TM Image. *Open Remote Sens. J.* **2009**, *2*, 24–35. [[CrossRef](#)]
- Plaza, A.; Benediktsson, J.A.; Boardman, J.W.; Brazile, J.; Bruzzone, L.; Camps-Valls, G.; Chanussot, J.; Fauvel, M.; Gamba, P.; Gualtieri, A.; et al. Recent advances in techniques for hyperspectral image processing. *Remote Sens. Environ.* **2009**, *113*, 110–122. [[CrossRef](#)]
- Bartholomeus, H.; Kooistra, L.; Stevens, A.; van Leeuwen, M.; van Wesemael, B.; Ben-Dor, E.; Tychon, B. Soil organic carbon mapping of partially vegetated agricultural fields with imaging spectroscopy. *Int. J. Appl. Earth Obs. Geoinf.* **2011**, *13*, 81–88. [[CrossRef](#)]
- Ouerghemmi, W.; Gomez, C.; Naceur, S.; Lagacherie, P. Applying blind source separation on hyperspectral data for clay content estimation over partially vegetated surfaces. *Geoderma* **2011**, *163*, 227–237. [[CrossRef](#)]
- Ouerghemmi, W.; Gomez, C.; Naceur, S.; Lagacherie, P. Semi-blind source separation for the estimation of the clay content over semi-vegetated areas using VNIR/SWIR hyperspectral airborne data. *Remote Sens. Environ.* **2016**, *181*, 251–263. [[CrossRef](#)]
- Franceschini, M.H.D.; Demattê, J.A.M.; da Silva, T.F.; Vicente, L.E.; Bartholomeus, H.; de Souza, F.C.R. Prediction of soil properties using imaging spectroscopy: Considering fractional vegetation cover to improve accuracy. *Int. J. Appl. Earth Obs. Geoinf.* **2015**, *38*, 358–370. [[CrossRef](#)]
- Diek, S.; Fornallaz, F.; Schaepman, M.E.; De Jong, R. Barest pixel composite for agricultural areas using Landsat time series. *Remote Sens.* **2017**, *9*, 1245. [[CrossRef](#)]
- Demattê, J.A.M.; Fongaro, C.T.; Rizzo, R.; Safanelli, J.L. Geospatial Soil Sensing System (GEOS3): A powerful data mining procedure to retrieve soil spectral reflectance from satellite images. *Remote Sens. Environ.* **2018**, *212*, 161–175. [[CrossRef](#)]
- Rogge, D.; Bauer, A.; Zeidler, J.; Mueller, A.; Esch, T.; Heiden, U. Building an exposed soil composite processor (SCMaP) for mapping spatial and temporal characteristics of soils with Landsat imagery (1984–2014). *Remote Sens. Environ.* **2018**, *205*, 1–17. [[CrossRef](#)]
- Vaudour, E.; Gomez, C.; Lagacherie, P.; Loiseau, T.; Baghdadi, N.; Urbina-Salazar, D.; Loubet, B.; Arrouays, D. Temporal mosaicking approaches of Sentinel-2 images for extending topsoil organic carbon content mapping in croplands. *Int. J. Appl. Earth Obs. Geoinf.* **2021**, *96*, 102277. [[CrossRef](#)]

24. Gasmí, A.; Gomez, C.; Chehbouni, A.; Dhiba, D.; Elfil, H. Satellite Multi-Sensor Data Fusion for Soil Clay Mapping Based on the Spectral Index and Spectral Bands Approaches. *Remote Sens.* **2022**, *14*, 1103. [[CrossRef](#)]
25. Loiseau, T.; Chen, S.; Mulder, V.L.; Dobarco, M.R.; Richer-de-Forges, A.C.; Lehmann, S.; Bourennane, H.; Saby, N.P.; Martin, M.P.; Vaudour, E. Satellite data integration for soil clay content modelling at a national scale. *Int. J. Appl. Earth Obs. Geoinf.* **2019**, *82*, 101905. [[CrossRef](#)]
26. Gallo, B.C.; Demattê, J.A.M.; Rizzo, R.; Safanelli, J.L.; Mendes, W.D.S.; Lepsch, I.F.; Sato, M.V.; Romero, D.J.; Lacerda, M.P. Multi-temporal satellite images on topsoil attribute quantification and the relationship with soil classes and geology. *Remote Sens.* **2018**, *10*, 1571. [[CrossRef](#)]
27. Castaldi, F.; Hueni, A.; Chabrilat, S.; Ward, K.; Buttafuoco, G.; Bomans, B.; Vreys, K.; Brell, M.; van Wesemael, B. Evaluating the capability of the Sentinel 2 data for soil organic carbon prediction in croplands. *ISPRS J. Photogramm. Remote Sens.* **2019**, *147*, 267–282. [[CrossRef](#)]
28. Demattê, J.A.M.; Safanelli, J.L.; Poppiel, R.R.; Rizzo, R.; Silvero, N.E.Q.; Mendes, W.D.S.; Bonfatti, B.R.; Dotto, A.C.; Salazar, D.F.U.; Mello, F.A.D.O.; et al. Bare earth's surface spectra as a proxy for soil resource monitoring. *Sci. Rep.* **2020**, *10*, 4461. [[CrossRef](#)]
29. Safanelli, J.L.; Chabrilat, S.; Ben-Dor, E.; Demattê, J.A.M. Multispectral models from bare soil composites for mapping topsoil properties over Europe. *Remote Sens.* **2020**, *12*, 1369. [[CrossRef](#)]
30. Silvero, N.E.Q.; Demattê, J.A.M.; Amorim, M.T.A.; dos Santos, N.V.; Rizzo, R.; Safanelli, J.L.; Poppiel, R.R.; de Sousa, M.W.; Bonfatti, B.R. Soil variability and quantification based on Sentinel-2 and Landsat-8 bare soil images: A comparison. *Remote Sens. Environ.* **2021**, *252*, 112117. [[CrossRef](#)]
31. Sorenson, P.T.; Shirliff, S.J.; Bedard-Haughn, A.K. Predictive soil mapping using historic bare soil composite imagery and legacy soil survey data. *Geoderma* **2021**, *401*, 115316. [[CrossRef](#)]
32. Zepp, S.; Heiden, U.; Bachmann, M.; Möller, M.; Wiesmeier, M.; van Wesemael, B. Optimized bare soil compositing for soil organic carbon prediction of topsoil croplands in Bavaria using Landsat. *ISPRS J. Photogramm. Remote Sens.* **2023**, *202*, 287–302. [[CrossRef](#)]
33. Urbina-Salazar, D.; Vaudour, E.; Richer-de-Forges, A.C.; Chen, S.; Martelet, G.; Baghdadi, N.; Arrouays, D. Sentinel-2 and Sentinel-1 Bare Soil Temporal Mosaics of 6-year Periods for Soil Organic Carbon Content Mapping in Central France. *Remote Sens.* **2023**, *15*, 2410. [[CrossRef](#)]
34. Castaldi, F.; Koparan, M.H.; Wetterlind, J.; Źydelski, R.; Vinci, I.; Savaş, A.Ö.; Vaudour, E. Assessing the capability of Sentinel-2 time-series to estimate soil organic carbon and clay content at local scale in croplands. *ISPRS J. Photogramm. Remote Sens.* **2023**, *199*, 40–60. [[CrossRef](#)]
35. Tomer, S.K.; Al Bitar, A.; Sekhar, M.; Zribi, M.; Bandyopadhyay, S.; Sreelash, K.; Sharma, A.K.; Corgne, S.; Kerr, Y. Retrieval and multi-scale validation of soil moisture from multi-temporal SAR data in a semi-arid tropical region. *Remote Sens.* **2015**, *7*, 8128–8153. [[CrossRef](#)]
36. Sekhar, M.; Riotte, J.; Ruiz, L.; Jouquet, P.; Braun, J.J. Influences of climate and agriculture on water and biogeochemical cycles: Kabini critical zone observatory. *Proc. Indian Natl. Sci. Acad.* **2016**, *82*, 833–846. [[CrossRef](#)]
37. Gaillardet, J.; Braud, I.; Hankard, F.; Anquetin, S.; Bour, O.; Dorflinger, N.; de Dreuzy, J.R.; Galle, S.; Galy, C.; Gogo, S.; et al. OZCAR: The French network of critical zone observatories. *Vadose Zone J.* **2018**, *17*, 1–24. [[CrossRef](#)]
38. Robert, M.; Thomas, A.; Sekhar, M.; Badiger, S.; Ruiz, L.; Willaume, M.; Leenhardt, D.; Bergez, J.E. Farm typology in the Berambadi Watershed (India): Farming systems are determined by farm size and access to groundwater. *Water* **2017**, *9*, 51. [[CrossRef](#)]
39. Barbiero, L.; Parate, H.R.; Desclotres, M.; Bost, A.; Furian, S.; Kumar, M.M.; Kumar, C.; Braun, J.J. Using a structural approach to identify relationships between soil and erosion in a semi-humid forested area, South India. *Catena* **2007**, *70*, 313–329. [[CrossRef](#)]
40. Gomez, C.; Dharumarajan, S.; Lagacherie, P.; Riotte, J.; Ferrant, S.; Sekhar, M.; Ruiz, L. Mapping of tank silt application using Sentinel-2 images over the Berambadi catchment (India). *Geoderma Reg.* **2021**, *25*, e00389. [[CrossRef](#)]
41. Lagacherie, P.; Gomez, C. Vis-NIR-SWIR remote sensing products as new soil data for digital soil mapping. In *Pedometrics*; McBratney, A.B., Minasny, B., Stockmann, U., Eds.; Springer: Cham, Switzerland, 2018; pp. 415–437. [[CrossRef](#)]
42. Ben-Dor, E.; Patkin, K.; Banin, A.; Karnieli, A. Mapping of several soil properties using DAIS-7915 hyperspectral scanner data—A case study over clayey soils in Israel. *Int. J. Remote Sens.* **2002**, *23*, 1043–1062. [[CrossRef](#)]
43. Gomez, C.; Oltra-Carrió, R.; Bacha, S.; Lagacherie, P.; Briottet, X. Evaluating the sensitivity of clay content prediction to atmospheric effects and degradation of image spatial resolution using Hyperspectral VNIR/SWIR imagery. *Remote Sens. Environ.* **2015**, *164*, 1–15. [[CrossRef](#)]
44. Robinson, G.W. A new method for the mechanical analysis of soils and other dispersions. *J. Agric. Sci.* **1992**, *12*, 306–321. [[CrossRef](#)]
45. Gee, G.W.; Or, D. 2.4 Particle-size analysis. In *Methods of Soil Analysis, Part 4: Physical Methods*; Dane, J.H., Top, C.G., Eds.; John Wiley and Sons: Madison, WI, USA, 2002; pp. 255–293. [[CrossRef](#)]
46. George, E.B.; Gomez, C.; Nagesh Kumar, D.; Dharumarajan, S.; Lalitha, M. Impact of bare soil pixels identification on clay content mapping using airborne hyperspectral AVIRIS-NG data: Spectral Indices versus Spectral Unmixing. *Geocarto Int.* **2022**, *37*, 15912–15934. [[CrossRef](#)]
47. Bhattacharya, B.K.; Green, R.O.; Rao, S.; Saxena, M.; Sharma, S.; Kumar, K.A.; Srinivasulu, P.; Sharma, S.; Dhar, D.; Bandyopadhyay, S.; et al. An overview of AVIRIS-NG airborne hyperspectral science campaign over India. *Curr. Sci.* **2019**, *116*, 1082–1088. [[CrossRef](#)]
48. Hamlin, L.; Green, R.O.; Mouroulis, P.; Eastwood, M.; Wilson, D.; Dudik, M.; Paine, C. Imaging spectrometer science measurements for terrestrial ecology: AVIRIS and new developments. In *Proceedings of the 2011 Aerospace Conference*, Washington, DC, USA, 5–12 March 2011; IEEE Computer Society: Washington, DC, USA, 2011; pp. 1–7. [[CrossRef](#)]

49. Thorpe, A.K.; Frankenberg, C.; Aubrey, A.D.; Roberts, D.A.; Nottrott, A.A.; Rahn, T.A.; Sauer, J.A.; Dubey, M.K.; Costigan, K.R.; Arata, C.; et al. Mapping methane concentrations from a controlled release experiment using the next generation airborne visible/infrared imaging spectrometer (AVIRIS-NG). *Remote Sens. Environ.* **2016**, *179*, 104–115. [[CrossRef](#)]
50. Gao, B.C.; Heidebrecht, K.B.; Goetz, A.F. Derivation of scaled surface reflectances from AVIRIS data. *Remote Sens. Environ.* **1993**, *44*, 165–178. [[CrossRef](#)]
51. Keshava, N.; Mustard, J.F. Spectral unmixing. *IEEE Signal Process. Mag.* **2002**, *19*, 44–57. [[CrossRef](#)]
52. Heinz, D.C.; Chang, C.I. Fully constrained least squares linear spectral mixture analysis method for material quantification in hyperspectral imagery. *IEEE Trans. Geosci. Remote Sens.* **2001**, *39*, 529–545. [[CrossRef](#)]
53. Chi, J.; Crawford, M.M. Spectral unmixing-based crop residue estimation using hyperspectral remote sensing data: A case study at Purdue university. *IEEE J. Sel. Top. Appl. Earth Obs. Remote Sens.* **2014**, *7*, 2531–2539. [[CrossRef](#)]
54. Plaza, A.; Martínez, P.; Pérez, R.; Plaza, J. A quantitative and comparative analysis of endmember extraction algorithms from hyperspectral data. *IEEE Trans. Geosci. Remote Sens.* **2004**, *42*, 650–663. [[CrossRef](#)]
55. Wold, S.; Sjöström, M.; Eriksson, L. PLS-regression: A basic tool of chemometrics. *Chemom. Intell. Lab. Syst.* **2001**, *58*, 109–130. [[CrossRef](#)]
56. 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. [[CrossRef](#)]
57. Peón, J.; Recondo, C.; Fernández, S.F.; Calleja, J.; De Miguel, E.; Carretero, L. Prediction of topsoil organic carbon using airborne and satellite hyperspectral imagery. *Remote Sens.* **2017**, *9*, 1211. [[CrossRef](#)]
58. Ward, K.J.; Chabrilat, S.; Brell, M.; Castaldi, F.; Spengler, D.; Foerster, S. Mapping soil organic carbon for airborne and simulated EnMAP imagery using the LUCAS soil database and a Local PLSR. *Remote Sens.* **2020**, *12*, 3451. [[CrossRef](#)]
59. Savitzky, A.; Golay, M.J. Smoothing and differentiation of data by simplified least squares procedures. *Anal. Chem.* **1964**, *36*, 1627–1639. [[CrossRef](#)]
60. Mark, H.L.; Tunnell, D. Qualitative near-infrared reflectance analysis using Mahalanobis distances. *Anal. Chem.* **1985**, *57*, 1449–1456. [[CrossRef](#)]
61. Mendes, W.D.S.; Boechat, C.L.; Gualberto, A.V.S.; Barbosa, R.S.; Silva, Y.J.A.B.D.; Saraiva, P.C.; Sena, A.F.S.D.; Duarte, L.D.S.L. Soil spectral library of Piauí State using machine learning for laboratory analysis in Northeastern Brazil. *Rev. Bras. Ciência Solo* **2021**, *45*, e0200115. [[CrossRef](#)]
62. Cambou, A.; Barthès, B.G.; Moulin, P.; Chauvin, L.; Masse, D.; Chevallier, T.; Chapuis-Lardy, L. Prediction of soil carbon and nitrogen contents using visible and near infrared diffuse reflectance spectroscopy in varying salt-affected soils in Sine Saloum (Senegal). *Catena* **2022**, *212*, 106075. [[CrossRef](#)]
63. Efron, B.; Tibshirani, R.J. *An Introduction to the Bootstrap*; Chapman and Hall: New York, NY, USA, 1994. [[CrossRef](#)]
64. Viscarra-Rossel, R.A.; Hicks, W.S. Soil organic carbon and its fractions estimated by visible–near infrared transfer functions. *Eur. J. Soil Sci.* **2015**, *66*, 438–450. [[CrossRef](#)]
65. Vaudour, E.; Gilliot, J.M.; Bel, L.; Lefevre, J.; Chehdi, K. Regional prediction of soil organic carbon content over temperate croplands using visible near-infrared airborne hyperspectral imagery and synchronous field spectra. *Int. J. Appl. Earth Obs. Geoinf.* **2016**, *49*, 24–38. [[CrossRef](#)]
66. Gomez, C.; Dharumarajan, S.; Féret, J.B.; Lagacherie, P.; Ruiz, L.; Sekhar, M. Use of sentinel-2 time-series images for classification and uncertainty analysis of inherent biophysical property: Case of soil texture mapping. *Remote Sens.* **2019**, *11*, 565. [[CrossRef](#)]
67. Roberts, D.A.; Batista, G.T.; Pereira, J.L.G.; Waller, E.; Nelson, B.W. Change identification using multitemporal spectral mixture analysis: Applications in eastern Amazonia. In *Remote Sensing Change Detection: Environmental Monitoring Methods and Applications*; Lunetta, R.S., Elvidge, C.D., Eds.; University of Michigan Press: Ann Arbor, MI, USA, 1998; pp. 137–161.
68. Drumetz, L.; Tochon, G.; Chanussot, J.; Jutten, C. Estimating the number of endmembers to use in spectral unmixing of hyperspectral data with collaborative sparsity. In Proceedings of the 13th International Conference on Latent Variable Analysis and Signal Separation, Grenoble, France, 21–23 February 2017; Tichavsky, P., Babaie-Zadeh, M., Michel, O., Thirion-Moreau, N., Eds.; Springer: Cham, Switzerland, 2017; pp. 381–391. [[CrossRef](#)]
69. Somers, B.; Asner, G.P.; Tits, L.; Coppin, P. Endmember variability in spectral mixture analysis: A review. *Remote Sens. Environ.* **2011**, *115*, 1603–1616. [[CrossRef](#)]
70. Selige, T.; Böhner, J.; Schmidhalter, U. High resolution topsoil mapping using hyperspectral image and field data in multivariate regression modelling procedures. *Geoderma* **2006**, *136*, 235–244. [[CrossRef](#)]
71. Dvorakova, K.; Shi, P.; Limbourg, Q.; van Wesemael, B. Soil organic carbon mapping from remote sensing: The effect of crop residues. *Remote Sens.* **2020**, *12*, 1913. [[CrossRef](#)]
72. Gasmí, A.; Gomez, C.; Lagacherie, P.; Zouari, H.; Laamrani, A.; Chehbouni, A. Mean spectral reflectance from bare soil pixels along a Landsat-TM time series to increase both the prediction accuracy of soil clay content and mapping coverage. *Geoderma* **2021**, *388*, 114864. [[CrossRef](#)]
73. Breiman, L. Bagging predictors. *Mach. Learn.* **1996**, *24*, 123–140. [[CrossRef](#)]
74. Brodský, L.; Vašát, R.; Klement, A.; Zádorová, T.; Jakšík, O. Uncertainty propagation in VNIR reflectance spectroscopy soil organic carbon mapping. *Geoderma* **2013**, *199*, 54–63. [[CrossRef](#)]
75. Ben-Dor, E.; Chabrilat, S.; Demattê, 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. [[CrossRef](#)]

76. Wu, C.Y.; Jacobson, A.R.; Laba, M.; Baveye, P.C. Accounting for surface roughness effects in the near-infrared reflectance sensing of soils. *Geoderma* **2009**, *152*, 171–180. [[CrossRef](#)]
77. Denis, A.; Stevens, A.; van Wesemael, B.; Udelhoven, T.; Tychon, B. Soil organic carbon assessment by field and airborne spectrometry in bare croplands: Accounting for soil surface roughness. *Geoderma* **2014**, *226–227*, 94–102. [[CrossRef](#)]
78. Demattê, J.A.M.; Alves, M.R.; Terra, F.D.S.; Bosquilia, R.W.D.; Fongaro, C.T.; Barros, P.P.D.S. Is it possible to classify topsoil texture using a sensor located 800 km away from the surface? *Rev. Bras. Ciência Solo* **2016**, *40*. [[CrossRef](#)]
79. Diek, S.; Schaepman, M.E.; De Jong, R. Creating multi-temporal composites of airborne imaging spectroscopy data in support of digital soil mapping. *Remote Sens.* **2016**, *8*, 906. [[CrossRef](#)]
80. Vaudour, E.; Gomez, C.; Loiseau, T.; Baghdadi, N.; Loubet, B.; Arrouays, D.; Ali, L.; Lagacherie, P. The Impact of Acquisition Date on the Prediction Performance of Topsoil Organic Carbon from Sentinel-2 for Croplands. *Remote Sens.* **2019**, *11*, 2143. [[CrossRef](#)]
81. Loizzo, R.; Daraio, M.; Guarini, R.; Longo, F.; Lorusso, R.; Dini, L.; Lopinto, E. Prisma mission status and perspective. In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2019, Yokohama, Japan, July 28–2 August 2019; IEEE: Yokohama, Japan, 2019; pp. 4503–4506. [[CrossRef](#)]
82. Lee, C.M.; Cable, M.L.; Hook, S.J.; Green, R.O.; Ustin, S.L.; Mandl, D.J.; Middleton, E.M. An introduction to the NASA Hyperspectral InfraRed Imager (HypIRI) mission and preparatory activities. *Remote Sens. Environ.* **2015**, *167*, 6–19. [[CrossRef](#)]
83. Matsunaga, T.; Iwasaki, A.; Tsuchida, S.; Iwao, K.; Tanii, J.; Kashimura, O.; Nakamura, R.; Yamamoto, H.; Kato, S.; Obata, K.; et al. Current status of hyperspectral imager suite (HISUI) onboard International Space Station (ISS). In Proceedings of the IEEE International Geoscience and Remote Sensing Symposium (IGARSS 2017), Worth, TX, USA, 23–28 July 2017; pp. 443–446. [[CrossRef](#)]
84. Kaufmann, H.; Segl, K.; Guanter, L.; Hofer, S.; Foerster, K.P.; Stuffer, T.; Mueller, A.; Richter, R.; Bach, H.; Hostert, P.; et al. Environmental mapping and analysis program (EnMAP)—Recent advances and status. Proceedings of IEEE International Geoscience and Remote Sensing Symposium, IGARSS 2008, Boston, MA, USA, 7–11 July 2008; IEEE: Boston, MA, USA, 2008; pp. 109–112. [[CrossRef](#)]

**Disclaimer/Publisher’s Note:** The statements, opinions and data contained in all publications are solely those of the individual author(s) and contributor(s) and not of MDPI and/or the editor(s). MDPI and/or the editor(s) disclaim responsibility for any injury to people or property resulting from any ideas, methods, instructions or products referred to in the content.