Use of Sentinel-2 time-series images for classification and uncertainty analysis of inherent biophysical property: case of soil texture mapping

Cécile Gomez, Subramanian Dharumarajan, Jean-Baptiste Feret, Philippe Lagacherie, Laurent Ruiz, Muddu Sekhar

HAL Id: hal-02609202

https://hal.inrae.fr/hal-02609202

Submitted on 16 May 2020

Use of Sentinel-2 Time-Series Images for Classification and Uncertainty Analysis of Inherent Biophysical Property: Case of Soil Texture Mapping

Cécile Gomez [1,*] , **Subramanian Dharumarajan** [2] , **Jean-Baptiste Féret** [3] ,
Philippe Lagacherie [1] , **Laurent Ruiz** [4,5] **and Muddu Sekhar** [5,6]

1 LISAH, University Montpellier, IRD, INRA, Montpellier SupAgro, 34060 Montpellier, France;
philippe.lagacherie@inra.fr
2 ICAR-National Bureau of Soil Survey and Land Use Planning, Hebbal, Bangalore 440033, India;
sdharmag@gmail.com
3 TETIS, Irstea, AgroParisTech, CIRAD, CNRS, Université Montpellier, 34000 Montpellier, France;
jean-baptiste.feret@teledetection.fr
4 Géosciences Environnement Toulouse, Université Paul-Sabatier, IRD, CNRS, 31400 Toulouse, France;
laurent.ruiz@inra.fr
5 Indo-French Cell for Water Sciences, IRD, Indian Institute of Science, Bangalore 560012, India;
sekhar.muddu@gmail.com
6 Indian Institute of Science, Civil Engineering Department, Bangalore 560012, India
* Correspondence: cecile.gomez@ird.fr

Abstract: The Sentinel-2 mission of the European Space Agency (ESA) Copernicus program provides multispectral remote sensing data at decametric spatial resolution and high temporal resolution. The objective of this work is to evaluate the ability of Sentinel-2 time-series data to enable classification of an inherent biophysical property, in terms of accuracy and uncertainty estimation. The tested inherent biophysical property was the soil texture. Soil texture classification was performed on each individual Sentinel-2 image with a linear support vector machine. Two sources of uncertainty were studied: uncertainties due to the Sentinel-2 acquisition date and uncertainties due to the soil sample selection in the training dataset. The first uncertainty analysis was achieved by analyzing the diversity of classification results obtained from the time series of soil texture classifications, considering that the temporal resolution is akin to a repetition of spectral measurements. The second uncertainty analysis was achieved from each individual Sentinel-2 image, based on a bootstrapping procedure corresponding to 100 independent classifications obtained with different training data. The Simpson index was used to compute this diversity in the classification results. This work was carried out in an Indian cultivated region (84 km^2, part of Berambadi catchment, in the Karnataka state). It used a time-series of six Sentinel-2 images acquired from February to April 2017 and 130 soil surface samples, collected over the study area and characterized in terms of texture. The classification analysis showed the following: (i) each single-date image analysis resulted in moderate performances for soil texture classification, and (ii) high confusion was obtained between neighboring textural classes, and low confusion was obtained between remote textural classes. The uncertainty analysis showed that (i) the classification of remote textural classes (clay and sandy loam) was more certain than classifications of intermediate classes (sandy clay and sandy clay loam), (ii) a final soil textural map can be produced depending on the allowed uncertainty, and iii) a higher level of allowed uncertainty leads to increased bare soil coverage. These results illustrate the potential of Sentinel-2 for providing input for modeling environmental processes and crop management.

Keywords: time-series; Sentinel-2; soil texture; classification; uncertainty; Simpson index; bootstrap

1. Introduction

Soil provides key environmental functions such as food production, water storage and redistribution, pollutant filtering and carbon storage. The accurate characterization of soil properties over cultivated areas, including soil organic matter, soil texture or iron content, is essential for agricultural engineering work such as land consolidation, drainage management, soil erosion limitation and irrigation systems. This characterization provides valuable information to improve soil use and management actions over cultivated areas [1,2]. Environmental modeling increasingly contributes to decision making on soil management from local to global scales, but its operational use requires accurate and highly spatially referenced soil information as inputs. The most common way to obtain information on soil characterization is soil maps. However, the spatial resolutions of these maps are limited by the cost of collecting soil samples and measuring soil properties, and the spatial uncertainty of soil characterizations is usually unknown.

Laboratory Visible, Near-InfraRed and Short Wave InfraRed (VNIR-SWIR, 400–2500 nm) spectroscopy has become a good alternative to costly physical and chemical soil analysis for the estimation of a large range of soil properties, including soil organic matter, iron content and soil texture (e.g., [3,4]). This latter soil property refers to the composition of the soil in terms of the proportion of fine, medium, and large particles (clay, silt, and sand, respectively) in a specific soil mass and is one of the most studied properties by VNIR-SWIR data for three major reasons [4]. First, texture is an important physical property that plays a key role in soil functions, including fertility and water retention. Fine-textured soils (such as clayey soils) have more total pore space and individual pores with smaller diameters than coarse-textured soils (such as sandy soils). The adhesive force at the molecular scale between water and soil leads to higher levels of water retention in fine-textured soils than in coarse-textured soils. Second, soil texture influences both spectral intensity and absorption band depth in the VNIR-SWIR spectral domain [3,5]. A spectrum with a high content of clay fraction (fine particle sizes) will tend to have a higher albedo than a spectrum of sandy or loamy soil (coarse particle sizes). Third, soil texture does not vary in the short term, implying that the field collection of soil samples, soil texture analysis and VNIR-SWIR data acquisition may be carried out separately over time.

VNIR-SWIR hyperspectral airborne and satellite imagery has been shown to be a promising technology for quantitative soil surface property mapping (e.g., using a VNIR-SWIR hyperspectral airborne sensor [6–8]; or using a VNIR-SWIR hyperspectral satellite sensor [9,10]). VNIR-SWIR multispectral satellite imagery showed limited utility for quantitative soil surface property mapping applications compared to hyperspectral imagery. Indeed, the studies of quantitative soil texture mapping using hyperspectral data provided more accurate predictions [11,12] than those using multispectral data [13–15]. This could be due to (i) insufficient spectral information limiting prediction accuracy for soil properties [16] and (ii) the insufficient spatial resolution of VNIR-SWIR multispectral satellite sensors (e.g., 15 and 30 m for ASTER data depending on spectral bands, 30 m for LANDSAT data), which does not match with the patterns of interest required for the definition of soil structure [17]. Nevertheless, VNIR-SWIR multispectral satellite imagery has been successfully used for qualitative soil surface property mapping [18,19].

Whatever the VNIR-SWIR imaging sensor, the predictions performances of quantitative soil property are expressed using figure of merits such as coefficients of determination, root mean square errors or bias [20] and studying variograms of spatialized predicted properties [8]. In addition, whatever the VNIR-SWIR imaging sensor, the predictions performances of qualitative soil property are expressed using figure of merits such as the coefficient of determination and kappa and studying the confusion matrices. In addition, to our knowledge, only [21] has tried to estimate predictions uncertainties in addition to quantitative predictions performances.

With the successful launch of the Sentinel-2 satellites, the Copernicus program has provided global coverage of terrestrial surfaces with multispectral images, with a revisit time of five days, since 2017. Images are acquired with a spatial resolution of 10 m to 60 m on 13 spectral bands

from 400 to 2500 nm. Several studies have illustrated the relevance of individual Sentinel-2 images for environmental applications, including soil organic mapping [22,23], soil moisture mapping using the synergy of Sentinel-1 (radar) and Sentinel-2 data [24,25], and Mediterranean seagrass mapping [26]. The importance of the multitemporal dimension of Sentinel-2 data was also evidenced for environmental applications, including the detection of precursory motions before landslide failures [27], the evaluation of the extent of forest fires [28], the estimation of vegetation phenological stages [29], water use by cotton [30], and the improvement of crop type mapping [31]. In Soil Science, the multitemporal dimension may allow to (i) increase the probability of image acquisition in clear sky conditions during periods with consistent bare soil coverage over cultivated areas (e.g., between October and November in Mediterranean areas, during the plowing time) and (ii) provide several repetitions of spectral measurements of the surface. Combining time series data for extending bare soil coverage has been recently studied (e.g., [32,33]), and in our paper we propose another example of added value of using the time series data.

The objective of this work is to explore a Sentinel-2 time-series images for soil texture classification in terms of accuracy and uncertainty estimation. Supervised classifications are performed on each individual Sentinel-2 image, and two sources of uncertainty were studied. Uncertainties due to the training data were studied using a bootstrapping procedure applied to each individual image. In addition, uncertainties due to the Sentinel-2 acquisition date were studied from the frequency of classifications in the time-series classification, based on the hypothesis that soil texture does not vary in the short term. This work was carried out in an Indian cultivated region (Karnataka state).

2. Materials

2.1. Study Area

The Berambadi catchment is a subcatchment of the South Gundal located in the Deccan Plateau of Southern India (Figure 1a) and covers 84 km². Our study area is located in the eastern part of the Berambadi catchment, which is covered by crop fields and corresponds to 60% of the catchment, while the western part is covered by forest (Figure 1). The Berambadi catchment belongs to the Kabini Critical Zone Observatory [34,35], which is part of the French Network of Critical Zone Observatories (OZCAR) [36]. The climate is tropical subhumid with an average rainfall of 800 mm/year and a PET of 1100 mm (aridity index P/PET of 0.7). The monsoon dynamics drive three main seasons: dry season (winter in January and February, summer from March to May), Kharif (southwest monsoon season, from June to September) and Rabi (north-east monsoon season, from October to December). Black soils (Vertisols and Vertic intergrades) are found mostly in the valley bottoms, while red soils (Ferralsols and Chromic Luvisols) cover the uplands and hillslopes [37].
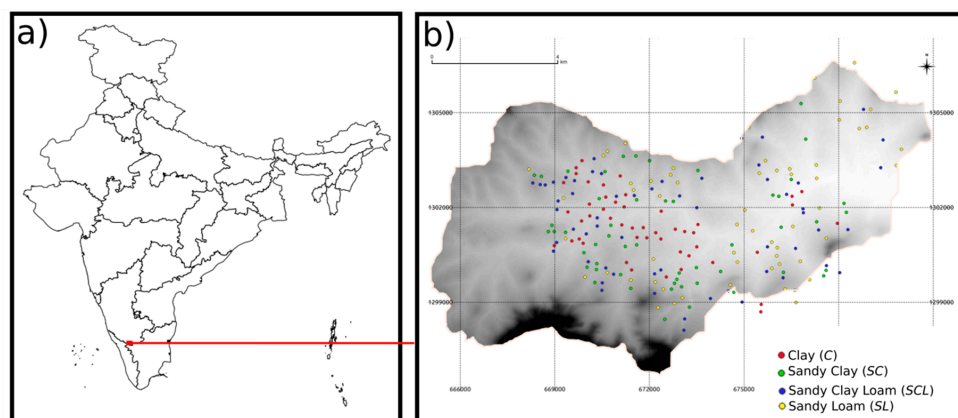

Figure 1. (**a**) Location of the Berambadi catchment in India (red point) and (**b**) the 216 collected soil surface samples plotted over a SRTM Digital Elevation Model, from black (low elevation) to white (high elevation) areas.

2.2. Sentinel-2 Images

Sentinel-2 satellites provide systematic global acquisitions of high spatial resolution multispectral imagery and are part of Copernicus program. Sentinel-2A was launched in 2015, followed by the Sentinel-2B launch in 2017. The combination of both satellites delivers a revisit period of up to five days under cloud-free conditions. The images include thirteen spectral bands in the VNIR-SWIR spectral domain, from 10 to 60 m of spatial resolution. Only ten bands remain after atmospheric correction, as the three bands acquired at 60 m spatial resolution (coastal at 443 nm, water vapor at 1375 nm and cirrus at 1376 nm) are used directly to perform atmospheric corrections and cloud detection. Here, atmospheric and topographic corrections were performed with Sen2Cor processor [38], and the level-2A images (bottom of atmosphere terrain corrected reflectance) were used to perform soil classification. The six spectral bands initially acquired with 20 m spatial resolution were resampled to 10 m.

Six Sentinel-2 images, including the Berambadi catchment during the dry season of 2017, were identified. Four images were acquired in February, under clear sky conditions (Figure 2). A fifth image was acquired on the 4th of April, under clear sky conditions and after a rainfall measured at 6 am over the Maddur Village (Figure 2). The sixth image was acquired on the 24th of April under clear sky conditions and after four days without rainfall (Figure 2).
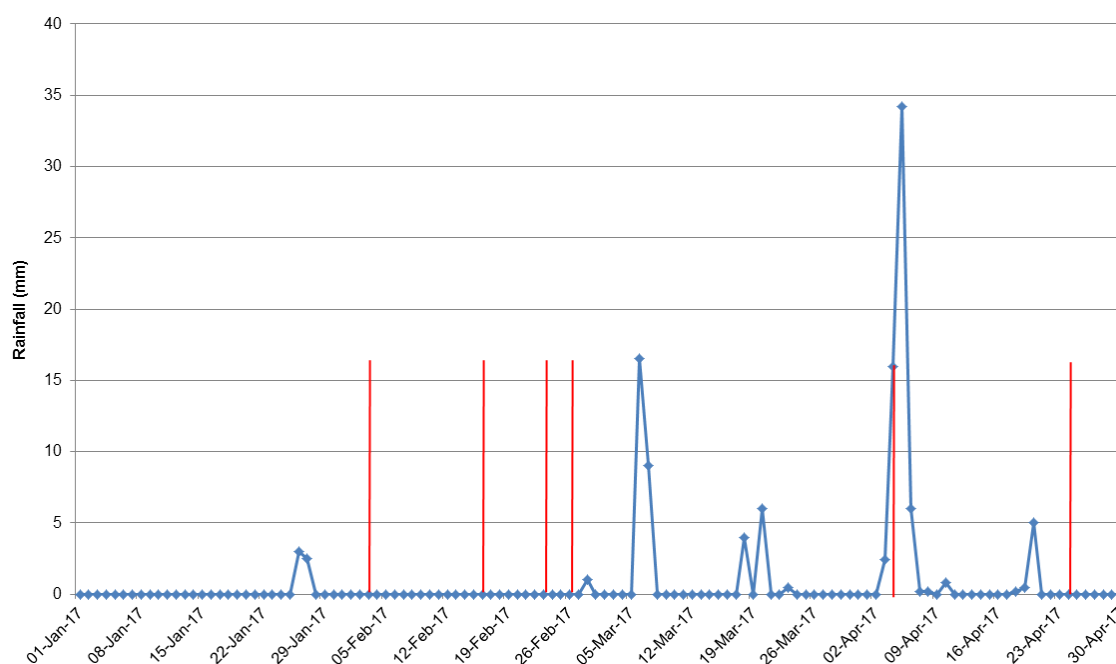
Figure 2. Rainfall (in mm) measured every day at 6 am, over Maddur village (Berambadi). Vertical red lines indicate the dates of Sentinel-2 images acquisition.

Once atmospherically and topographically corrected, the images were filtered to select only pixels corresponding to bare soil and to mask pixels corresponding to urban areas, bodies of water, crops and natural vegetation. Crops and natural vegetation were masked using the normalized difference vegetation index (NDVI) calculated using spectral bands at 842 nm and 665 nm [39,40]. A unique threshold for all images was fixed at 0.3, based on visual interpretation from an expert with extensive field knowledge. Urban areas and lakes were masked using a land use map available for the study area. Between 73% and 87% of the pixels were finally selected as bare soil in the Berambadi catchment, depending on the acquisition date. Overall, the six images shared 65% pixels corresponding to bare soil. For the sake of comparison among dates, this study focused on the pixels identified as bare soils and common to all images.

2.3. Soil Data

A total of 216 soil surface samples were collected over the Berambadi catchment (Figure 1b), including 176 soil surface samples collected in 2015 and 40 soil surface samples collected in 2017. The sampling was based on the soil physiographic relationship. Soil samples were collected in transects across the landforms. The major landforms are hills and hillslopes, gently sloping uplands and nearly level lowlands. Each sample was located with a GPS. The determination of soil textural classes was performed by experienced soil surveyors, using the feel method based on codified field tests [41] and many comparisons with soil analyses. Soil texture was classified into four classes: clay (C), sandy clay (SC), sandy clay loam (SCL) and sandy loam (SL).

Among the 216 soil surface samples associated with soil texture observation, 130 were located over bare soil pixels shared between the six images. The soil texture among these 130 samples was distributed as follows: 33% for SL, 27% for SCL, 23% for SC and 17% for C.

3. Methods

Soil texture classification was performed for each individual Sentinel-2 image over bare soil pixels shared between the six images. An uncertainty analysis was realized from each Sentinel-2 data, based on a bootstrapping procedure and from the time-series of soil texture classifications obtained from the time-series of Sentinel-2 images. Figure 3 illustrates the process flow. The classification model developments, training and validation partitions and prediction accuracy measurements were implemented in R using the caret package [42].

3.1. Classification of Individual Sentinel-2 Images

The spectral information from the pixels with corresponding soil texture measurements was extracted for each image (Figure 3a). For each classification, spectral information and corresponding textural class were then split into training (70%) and validation (30%) datasets using a stratified random sampling for training and validation data to follow a similar distribution of soil texture classes (Figure 3b).

Soil texture classification was performed using a linear support vector machine (Lin-SVM) classifier (Figure 3b), which is one of the most classical method for images classification [43–45]. The Lin-SVM was initially developed by [46] for binary classification and can be extended to multiclass problems. The SVM method consists of optimizing hyperplane separating classes by maximizing the margin between the classes' closest points, called support vectors. The best separating hyperplane can be found by solving a nonlinear convex problem that depends on a cost parameter C. This parameter C allows for reducing misclassifying in the training database while maximizing the margins between classes. Ten values ranging between 0 and 5 were tested for C, and the optimal value was defined for the best overall accuracy obtained when performing a 5-fold cross-validation on the training dataset. A detailed description of the Lin-SVM can be found in [47], and more recently in [43].

The accuracy and kappa coefficients were used to measure the performance of the classifications (Figure 3b). Accuracy is commonly measured as the percentage of pixels correctly classified in the validation set. The Kappa coefficient compares the observed accuracy with an expected accuracy (random chance). The kappa statistic is used to evaluate both a single classifier and classifiers among themselves, which is an important evaluation, especially on an imbalanced data set. A kappa coefficient of 1 indicates perfect classification, and a kappa coefficient of 0 corresponds to a random classification [48]. Based on [49], kappa values greater than 0.80 represent strong agreement, Kappa values between 0.4 and 0.8 represent moderate agreement, and Kappa values below 0.4 represent poor agreement. Finally, the confusion matrix was also calculated from each classification model (Figure 3b).
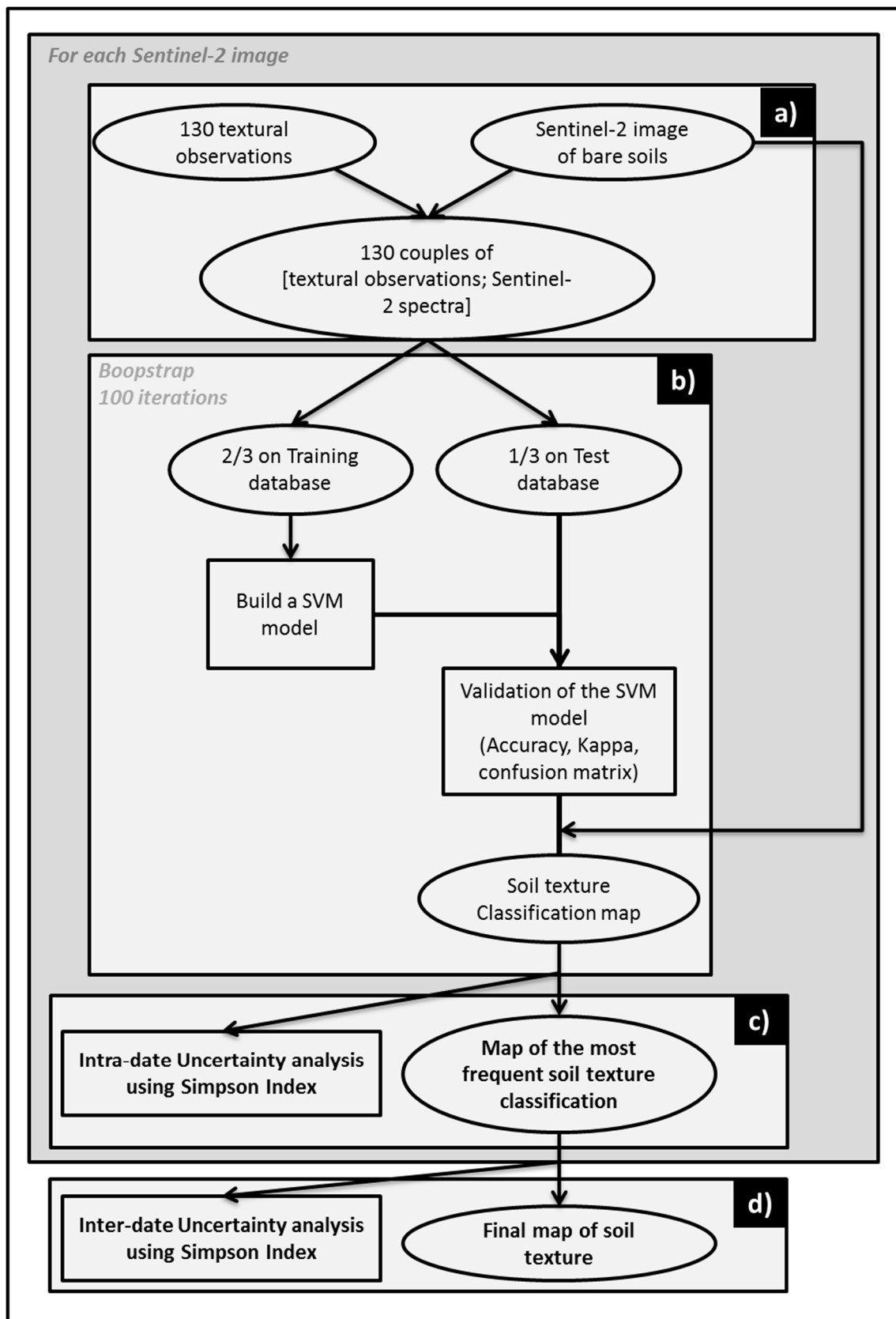
Figure 3. Workflow of the soil texture classification associated with the uncertainty analysis. (**a**) Extraction of spectral information associated to field soil texture, (**b**) classification process, (**c**) intra-date uncertainty analysis and intra-date soil texture map and (**d**) inter-date uncertainty analysis and final soil texture map.

3.2. Expressions of Uncertainty

As previously investigated by [21,50], a bootstrap procedure was used to study prediction uncertainties due to the training data. The bootstrapping involves repeated stratified random sampling for training and validation data to follow a similar distribution of soil texture classes, with replacement of the available data [51]. By repeating the process of sampling and applying the model, probability distributions of the prediction realizations are generated for each pixel. Here, a bootstrap procedure was applied to the original dataset to define 100 sets of training and validation subsets for each individual image. From each bootstrap sample, a classification model was fitted and applied to generate a soil classification (Figure 3b). After the bootstrap procedure, a single-date map of the most frequent class obtained for each pixel of each Sentinel-2 image was produced (Figure 3c), providing a time-series classification.

The frequency of classifications in the time-series classification was used for studying prediction uncertainties due to the Sentinel-2 acquisition date (Figure 3d). A final map of the most frequent class obtained for each pixel over the time-series classification was produced (Figure 3d).

The diversity of classifications was measured for each pixel by computing Simpson's diversity index, corresponding to the diversity of classes assigned to each pixel. Simpson's diversity index (SI) is one of the well-known indices used to measure a degree of diversity [52] and is calculated as follows:

$$SI = 1 - \sum_{t=1}^{t=k} p_t^2 \tag{1}$$

where p_t is the proportion of classification of class t, and k is the number of classes. SI reaches a minimum value (SI = 0) when all classification results are identical, so the proportion of this given class is 1 and the proportion of the other classes is 0. SI reaches its maximum when all classes are equally distributed, which corresponds to SI = $(1 - 1/k)$. In our case, SI was calculated for each bare soil pixel of each Sentinel-2 image and after the 100 iterations to study uncertainties due to the training data (Figure 3c). SI was calculated for each bare soil pixel of the time-series of Sentinel-2 images to study uncertainties due to the Sentinel-2 acquisition date (Figure 3d).

4. Results

4.1. Classification Model Performance

The classification models derived from the Sentinel-2 image acquired on the 24th of April 2017 outperformed other models by providing moderate accuracy (mean accuracy = 0.5) and fair agreement (mean kappa = 0.31) (Figure 4). These classification models provided a low confusion rate of classification for samples corresponding to remote textural classes in terms of texture (C and SL), and the highest confusion rate was obtained between neighboring textural classes (e.g., SCL and SL) (Table 1). The prediction ability for the SCL class systematically showed a poor score of correct predictions (Table 1).

The classification corresponding to the five other images resulted in lower performances (for both accuracy and kappa) than the classification obtained with the image acquired on the 24th of April 2017, with the lowest performances obtained when using the image acquired on the 4th of April 2017 (Figure 4). The confusion matrices showed similar patterns among dates, with low confusion rates of classification between remote textural classes (C and SL) and frequent confusion between neighboring textural classes (e.g., C and SL) (Table 1).

Finally, the standard deviation of classification performances at each date over the 100 models expressed in accuracy and kappa coefficients was similar among the six images: the standard deviation reached 0.05 for accuracy and 0.08 for Kappa (Figure 4).
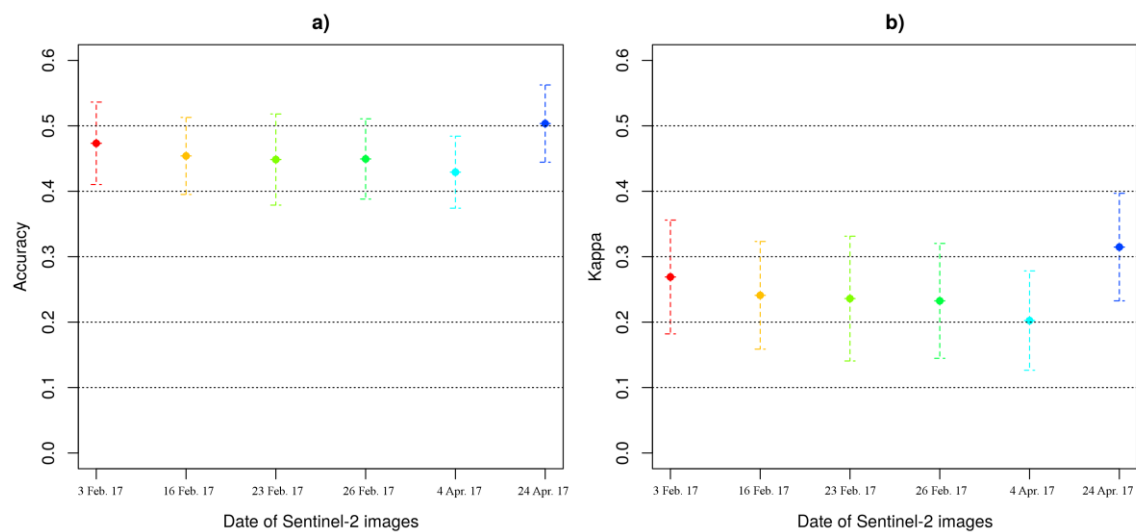
Figure 4. (**a**) Accuracy and (**b**) kappa coefficient obtained for 100 classification models built using each Sentinel-2 image.

Table 1. Mean of Confusion Matrixes (expressed in %) obtained by the 100 classification models from each image. Each row of the confusion matrix represents the instances in a predicted class while each column represents the instances in an observed class.

3 Feb. 2017		Reference				16 Feb. 2017		Reference			
		C	SC	SCL	SL			C	SC	SCL	SL
Prediction	C	67.3	9.3	8.3	0	Prediction	C	68.2	8.9	12.1	0.4
	SC	0.5	1.4	4.1	1.5		SC	1.5	1.3	2.7	2.1
	SCL	20.3	46.1	42.9	23.1		SCL	19.3	40.1	38.2	24.7
	SL	11.8	43.1	44.7	75.4		SL	11	49.7	47	72.8
23 Feb. 2017		**Reference**				**26 Feb. 2017**		**Reference**			
		C	SC	SCL	SL			C	SC	SCL	SL
Prediction	C	66.3	7.2	13.3	1	Prediction	C	64.2	7.8	11.1	1
	SC	1.5	1.6	3.9	2.4		SC	1.7	2.3	3.7	1.3
	SCL	21.3	45.6	36.2	22.8		SCL	27.7	42	32	19.8
	SL	10.8	45.7	46.6	73.8		SL	6.5	47.9	53.2	77.8
4 Apr. 2017		**Reference**				**24 Apr. 2017**		**Reference**			
		C	SC	SCL	SL			C	SC	SCL	SL
Prediction	C	64.3	14	12.3	1.6	Prediction	C	70.8	12.9	6.9	4.2
	SC	0.8	1.1	4.8	2.7		SC	1.2	1.4	3.4	0.9
	SCL	12.3	27.9	23.1	15.9		SCL	15.7	41.8	51.3	18.7
	SL	22.5	57	59.8	79.8		SL	12.3	43.9	38.4	76.2

4.2. Single-Date Classification Maps

The most frequent textural classes after the bootstrap process provided a single-date classification of soil texture. The Pearson coefficient of correlation between these six single-date classifications was superior at 0.74 (Table 2). Soil texture from valley bottoms was mainly classified as C, while soil texture from hillslopes was mainly classified as SL and SCL (Figures 1b and 5). The distribution of the most and least common textural classes estimated by the 100 classification models over the study area was also similar among Sentinel-2 images; the SL class was dominant among the estimated classes, while SC was almost never attributed to pixels (Figure 6).
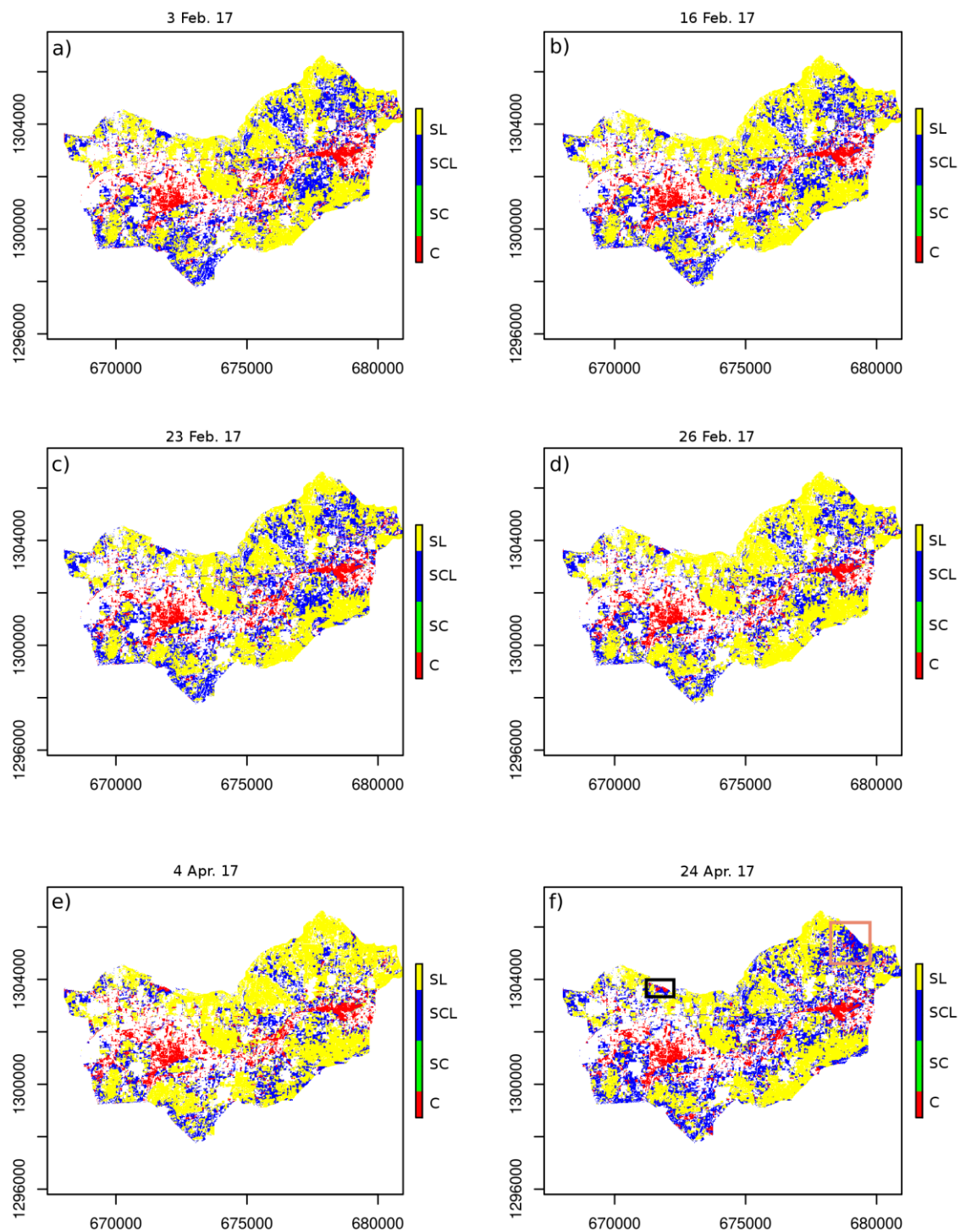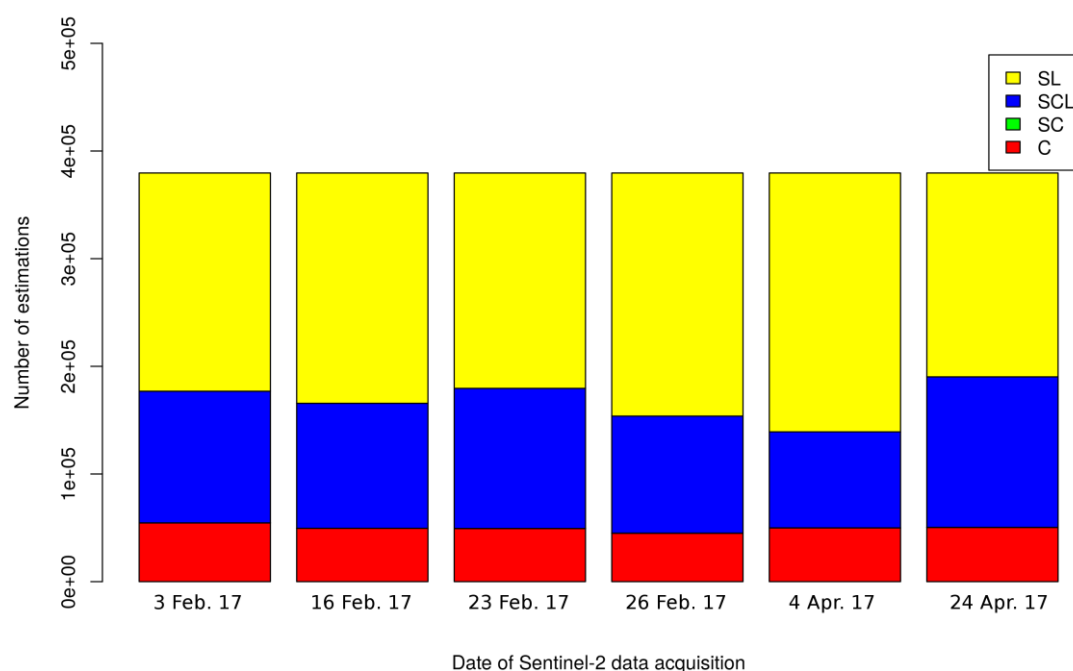
Figure 5. Maps of the most frequent textural class among the 100 estimations obtained from the image acquired on (**a**) 3 Feb. 2017, (**b**) 16 Feb. 2017, (**c**) 23 Feb. 2017, (**d**) 26 Feb. 2017, (**e**) 4 Apr. 2017 and (**f**) 24 Apr. 2017. The white areas correspond to the masked mixed pixels.

The image acquired on the 24th of April 2017, associated with the most accurate classification performance (Figure 4a), provided the most specific classification with Pearson coefficients of correlation inferior to 0.8 with other dates (Table 2, Figure 5f). On the northeast part of the study area, poorly covered by field samples, the image acquired on 24th of April 2017 provided a classification of a mixture of SCL and C classes (orange rectangle, Figure 5f), whereas this same area was mostly classified as a mixture of SCL and SL classes from the four images acquired in February (Figure 5a–d).

In addition, in the northwest part of the study area, for which two soil samples were collected and analyzed as the SC class, the image acquired on the 24th of April 2017 provided a classification of mixed C and SCL classes (black rectangle, Figure 5f), whereas this same area was mostly classified as SCL class based on the four images acquired in February (Figure 5a–d).

Table 2. Pearson coefficient of correlation (R) between the classification (most frequent class after the bootstrap process) obtained at each date.

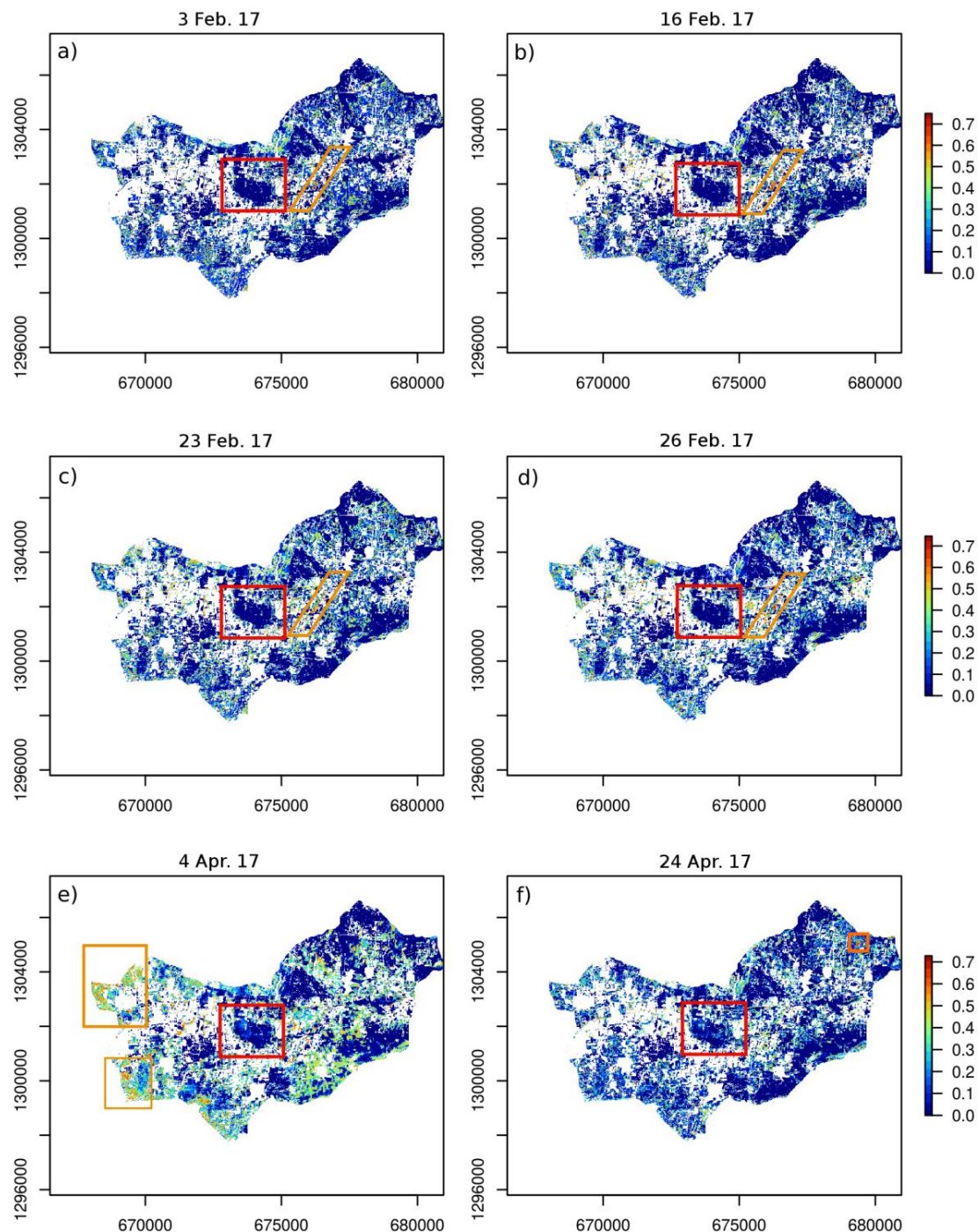|	3 Feb. 2017	16 Feb. 2017	23 Feb. 2017	26 Feb. 2017	4 Apr. 2017	24 Apr. 2017
3 Feb. 2017	1	0.89	0.89	0.84	0.75	0.74
16 Feb. 2017		1	0.94	0.92	0.79	0.76
23 Feb. 2017			1	0.91	0.80	0.78
26 Feb. 2017				1	0.81	0.77
4 Apr. 2017					1	0.78

Figure 6. Distribution of the most frequent textural class estimated over bare soils pixels at each acquisition date.

4.3. Uncertainties due to the Training Data Set

The highest uncertainty, i.e., the highest values of SI calculated from the bootstrap process, were obtained using the Sentinel-2 image acquired on the 4th of April 2017 (Figure 7e). The high values of the SI obtained from this date were clustered in the western part of the study area over hillslopes close to the forest (orange rectangles in Figure 7e). Moreover, even if urban pixels were masked in the pretreatment of the Sentinel-2 data (Section 2.2), the main road was not identified in the land use map of the catchment; thus, it was not masked and was associated with high SI (Figure 7e).
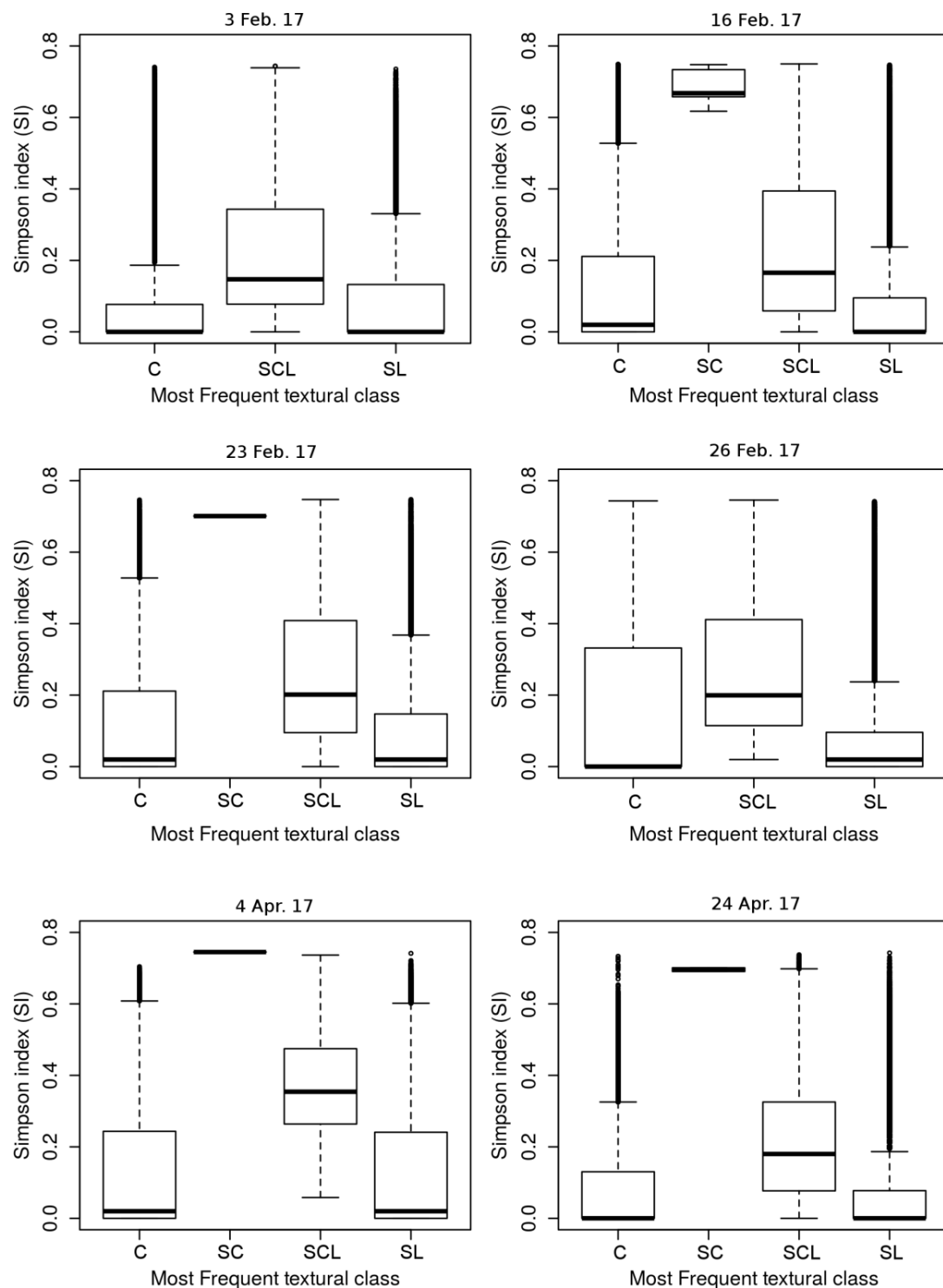
The lowest values of SI calculated from the bootstrap process were obtained using the Sentinel-2 image acquired on the 24th of April 2017 (Figure 7f). As obtained from the 4th of April 2017, the main road pixels are associated with high SI (Figure 7f). A wasteland located in the eastern part of the catchment is also associated with high values of SI (orange rectangle in Figure 7e). From the 4 Sentinel-2 images acquired on February, the highest SI values were obtained over a bottom valley (orange parallelogram in Figure 7a–d). Finally, some areas were associated with low SI values regardless of the Sentinel-2 image used (e.g., located by the red rectangle on Figure 7).

Figure 7. Maps of the Simpson's diversity index obtained from the image acquired on (**a**) 3 Feb. 2017, (**b**) 16 Feb. 2017, (**c**) 23 Feb. 2017, (**d**) 26 Feb. 2017, (**e**) 4 Apr. 2017 and (**f**) 24 Apr. 2017. The white areas correspond to the masked mixed pixels.

For each pixel and at each date, the SI values were expressed in regard to the most frequent textural class after the bootstrap process (Figure 8). Lower SI values were obtained consistently among images for pixels classified as C and SL, the two remote textural classes (Figure 8). This consistency in the classification process for repetitions on a given image and among images suggests lower uncertainty for these classes. Conversely, the SC class appeared as the most uncertain class for all images (Figure 8), as SI calculated on pixels attributed to this SC class was mostly superior to 0.6, and for two Sentinel-2

dates, the SC class was never affected as the most frequent class over the 100 iterations. Pixels attributed to the SCL class showed intermediate uncertainty based on SI (Figure 8).

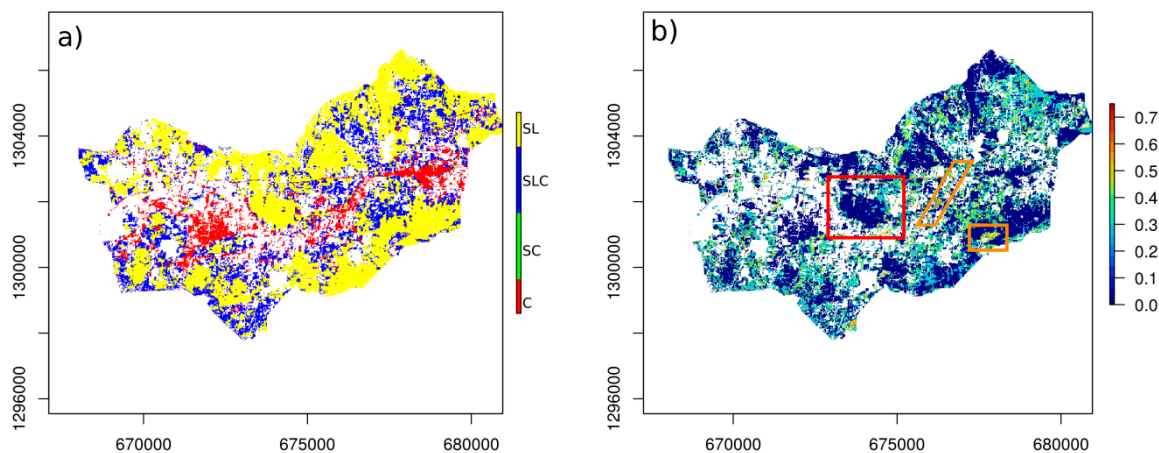**Figure 8.** Boxplot of Simpson's diversity index per class (the most frequent textural class in the set of 100 estimations) and per image.

4.4. Uncertainties due to the Image Acquisition Date

The lower values of SI calculated from the set of six classifications were obtained over pixels for which the most frequent textural classes among this set of six classifications are C and SL (Figure 9a,b).

The pixels for which the most frequent textural class among this set of six classifications is SLC are associated with an intermediate SI (Figure 9a,b). No pixels are classified as SC from the set of six classifications, suggesting high uncertainty for this class (Figure 9a). As obtained from single-date classifications, the main road pixels are associated with high SI (Figure 9b); high SI values are also obtained over a bottom valley (orange parallelogram in Figures 7a–d and 9b), and some areas are associated with low SI values (e.g., located by the red rectangle on Figures 7 and 9b). Some other regions are associated with higher SI than those obtained from single-date classifications (e.g., located by the orange rectangle on Figure 9b).

Figure 9. Map of (**a**) the most frequent textural class and (**b**) the Simpson's diversity index, both calculated from the time series of six classifications. The white areas correspond to the masked mixed pixels.

4.5. Toward the Use of an Uncertainty Threshold

The estimated uncertainties using the SI index can be used as thresholds for filtering the pixels that are considered too uncertain for being documented by a predicted soil texture. This loss in pixels having predictions may be compensated by better prediction performances for the pixels that remain predicted. We verified this both for SI calculated from the set of six single data classifications and for SI calculated from the set of 100 bootstrapped classifications of the best image (24th of April).

As expected, the gradients of variations of uncertainty and amount of classified bare soil are opposite. From the set of the six single-date classifications, 57% of bare soil pixels are classified if an SI of 0 is allowed (black triangles on Figure 10), and the classification accuracy associated with these bare soil pixels is estimated to be 0.63 (blue triangles on Figure 10). A nearly full coverage of classified bare soil pixels (93%) could be obtained for a maximum SI of 0.44, which corresponds to a class identified four times among classification processes performed over the six Sentinel-2 images (black triangles on Figure 10). The classification accuracy associated with this maximum SI of 0.44 is estimated to be 0.6 (blue triangles on Figure 10).

From the set of the 100 classifications obtained by bootstrap from the image acquired on the 24th of April, 39% of classified bare soil pixels are classified if an SI of 0 is allowed (black circles on Figure 10), and the associated classification accuracy is estimated to be 0.73 (blue circles on Figure 10). A nearly full coverage of classified bare soil pixels (93%) could be obtained for a maximum SI of 0.46 (black circles on Figure 10). The classification accuracy associated with this maximum SI of 0.46 is estimated to be 0.56 (blue circles on Figure 10).
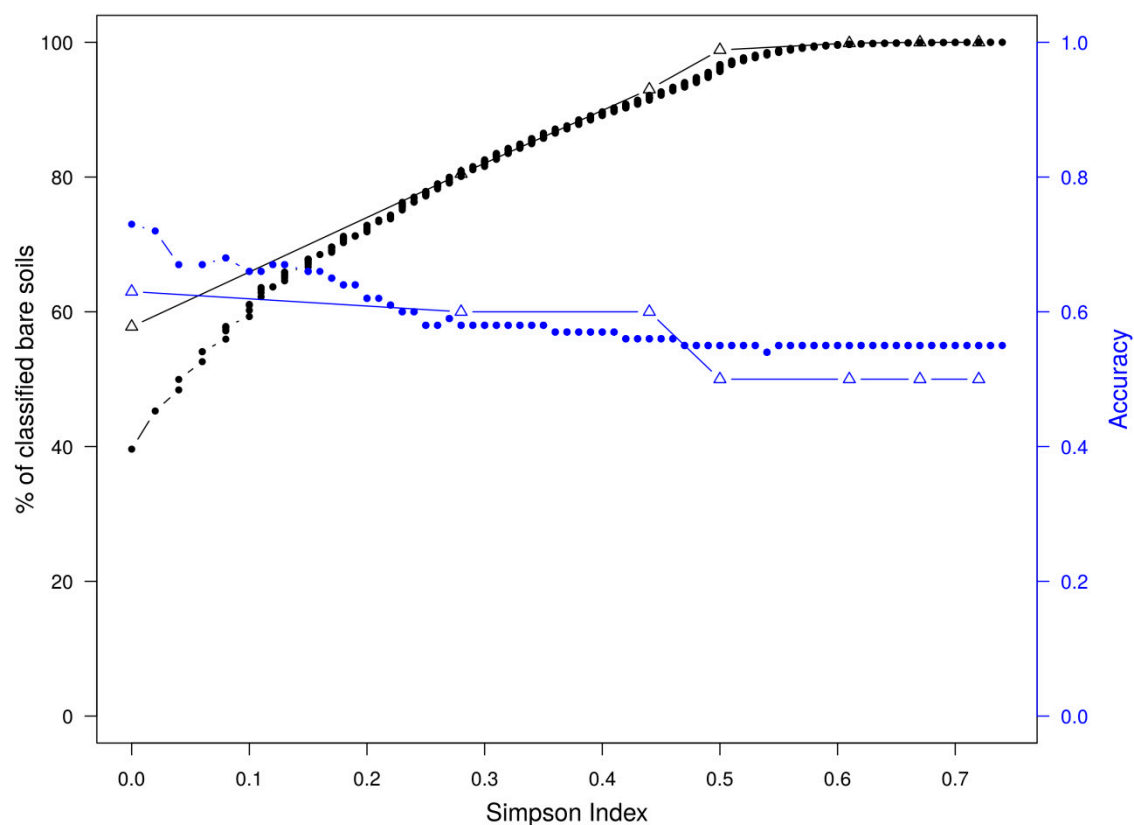
Figure 10. Proportion of classified bare soil pixels and classification performance (accuracy), versus a maximal value of SI allowed. Triangles correspond to results obtained from the set of six classifications. Circles correspond to results obtained from the bootstrap process applied to the image acquired on 24th of April.

Globally, if the allowed SI is inferior to 0.3, the set of the 100 classifications obtained by bootstrap from the image acquired on 24th of April provided better accuracy but lower classified coverage than the set of the six single-date classifications (Figure 10). From an SI of 0.3, the classified coverage remains the same, regardless of the set of classifications used. With an SI of 0.3 to 0.46, the accuracy obtained from the set of the 100 classifications obtained by bootstrap from the image acquired on the 24th of April is lower than the accuracy obtained from the set of the six single-date classifications. From an SI of 0.46, this tendency is opposite: the accuracy obtained from the set of the 100 classifications obtained by bootstrap from the image acquired on the 24th of April is higher than the accuracy obtained from the set of the six single-date classifications (Figure 10).

5. Discussion

Classification Performance by Individual Sentinel-2 Data

The classification models provided low to moderate performances for soil texture classification depending on the date of acquisition of the Sentinel-2 image used (accuracy ranging from 0.43 to 0.5). These performances were lower than those obtained by [18], who reported an accuracy of approximately 0.58 to 0.64 for soil texture classification using Landsat-5 TM images and lower than those obtained by [19], who reported an overall accuracy of 0.64 for soil texture classification and using Landsat-5 TM images.

The best performance of soil texture classification was obtained using the image acquired on the 24th of April 2017, and lower performance was obtained when using images acquired in February (Figure 4), which may be explained by higher soil moisture in the winter season. The lowest

performance, obtained when using the image acquired on the 4th of April 2017 (Figure 4) compared to the 24th of April 2017, may also be explained by soil moisture resulting from rainfall that occurred days before the acquisition (Figure 2), impacting the soil reflectance and reducing the spectral dissimilarity among classes. These results are in agreement with the results obtained by [53], who described a progressive decrease of performance for soil organic matter estimation from the drier to the wetter subsets. Soil surface roughness and crop residue over fields may also contribute to the poor discrimination among textural classes in addition to soil surface moisture, as suggested by [18], who attempted to classify soil texture from Landsat images.

Soil texture classification maps produced in our study showed good discrimination ability for (i) hillslopes and uplands, which correspond to coarse soil texture (sandy loam) due to erosion processes, and (ii) valleys, which correspond to finer soil texture (clay) due to deposition processes. Low confusion was obtained between the C and SL classes corresponding to remote textural classes (Table 1), and high confusion was obtained between neighboring textural classes, i.e., between C and SC and between SL and SCL, independent of the acquisition time of the Sentinel-2 image (Table 1). Dematte et al. [19] also succeeded in discriminating remote textural classes (clayey from sandy soils) using Landsat-5 TM images. As neighboring textural classes (e.g., C and SC) require almost similar types of management and behave similarly in crop production, whereas remote textural classes require different types of management, soil resource management policies may strongly benefit from the capacity of Sentinel-2 data to accurately identify these soil texture properties over regional scales.

How Training Dataset Selection and Acquisition Date Are Sources of Uncertainties

The bootstrap procedure provided an uncertainty expression in class prediction for each Sentinel-2 image and each bare soil pixel due to soil sample selection in the training dataset. Despite the scarcity of studies about uncertainty estimations of soil property prediction, the bootstrap procedure seems to be the most commonly used way to estimate it, as tested by [21,50,54]. These previous works estimated quantitative soil properties (soil organic carbon for [50,54]; clay content for [21]), so their uncertainties are represented by prediction variances. As we studied the soil texture class, which is a qualitative soil property, our uncertainties are represented by Simpson's diversity index.

Classification models seem to be highly sensitive to the soil sample selection in the training dataset, regardless of the Sentinel-2 image used (Figure 7). As an example, only 53% of the bare soil pixels were classified in the same textural class after the 100 bootstrap applied to the image acquired on the 24th of April 2017 (black circles on Figure 10). Regardless of the Sentinel-2 image used, low uncertainties were mostly obtained for pixels associated with the remote textural classes C and SL, while high uncertainties were mostly obtained for pixels associated with the class SC (Figure 8). Thus, if one classification is produced from one soil sample selection, the risk of misclassification of class SC is high, while the risk of misclassification of classes C and SL is lower. The high uncertainty associated with the textural class SC cannot be explained by an underrepresentation of class SC, as it corresponds to 30% of the training samples. The image acquired on the 4th of April 2017, which provided the lowest classification performance (Figure 4a), was associated with the highest uncertainties (Figure 7e). Therefore, this uncertainties map focuses attention on the use of a soil textural classification obtained from an image probably impacted by soil moisture resulting from rainfall that occurred days before the acquisition (Figure 2).

Based on the hypothesis that soil texture does not significantly change within months in natural conditions, the Sentinel-2 time-series provided an uncertainty expression in class prediction for each bare soil pixel due to the acquisition date. Classification models seem to be highly sensitive to the acquisition date, as only 57% of the bare soil pixels were classified in the same textural class among the time-series classification (black triangles on Figure 10). The lowest uncertainties were obtained for pixels associated with the remote textural classes C and SL. Inversely, the class SC was never predicted by the six images over the same bare soil pixels. This confirms that pixels for which soil texture is identified as SC based on individual images should be considered very carefully.

Uncertainties due to the acquisition date and soil sample selection in the training dataset are theoretically independent. Despite this independence, some specific areas were always associated with low uncertainties (red rectangle, Figures 7 and 9b). Conversely, other areas were associated with different levels of uncertainty depending on the images used (orange rectangles, Figures 7 and 9b). Therefore, this study demonstrated that uncertainties may be important, variable and spatially heterogeneous. However, hierarchical uncertainties cannot be proposed based on these results. The number of soil samples (130), the number of Sentinel-2 images (6) and the time period (4 months) might serve as limitations on clear hierarchical uncertainties analysis.

Additional Uncertainties and Inaccuracies

The classification accuracy over a pixel *i* may depend on the composition of this pixel *i* and the composition of pixels of the training database. As greater fraction of materials other than the soil (such as the dry and green vegetation) are covering these pixels (*i* and the ones used to train the classification model), the higher would be the importance of perturbation factors of VNIR-SWIR spectral measurements [55,56]. Therefore, the step of vegetated pixels mask is crucial. In VNIR-SWIR hyperspectral studies, both NDVI and Cellulose Absorption Index (CAI) spectral indexes [57] are used to identify green and dry vegetation pixels, respectively, and mask them [58]. In VNIR-SWIR multispectral studies, including using the Sentinel-2 sensor, the CAI may not be calculated as it needs spectral information measured at 2000, 2100 and 2200 nm. Spectral information measured at 2100 and 2200 nm are included in one band of Sentinel-2 (centered at 2190 nm with a band width of 180 nm), and the 2000 nm wavelength is not covered by Sentinel-2. So a part of inaccuracy and uncertainty may be due to crop residue coverage which cannot be removed from Sentinel-2 data.

The classification accuracy may also depend on the accuracy of class determination on training data. In this study, the determination of soil textural classes was performed on the field by experienced soil surveyors, using the feel method based on codified field tests [41]. Expertise of soil surveyors is more uncertain than laboratory routine process based on physico-chemical analysis. Confusion between soil textural classes may happen, in particular between neighboring textural classes. Among 60 soil textural classes determined from both physico-chemical analysis and the feel method over soil samples collected on 17 soil profiles on the Berambadi area, 23.3 % are misclassified by the feel method (Table 3). In addition, the higher confusion happens between both neighboring textural classes SC and SCL (Table 3). This high confusion between SC and SCL determined by the feel method may explain the high difficulty of the SVM model to correctly classify this class SC whatever the acquisition date and the training dataset. So the use of physico-chemical analysis, instead of the feel method, to determine the soil textural classes is likely to improve the capacity of Sentinel-2 data for mapping soil textural classes. Importantly, in terms of uncertainties, the confusion between SC and SCL by the feel method may also explain the high uncertainties associated with the predicted class SC after each bootstrap. So our uncertainty maps provide valuable information for avoiding misclassification due to the reference data.

Table 3. Confusion matrix obtained from 60 soil textural classes determined from both physico-chemical analysis and the feel method over soil samples collected on 17 soil profiles on the Berambadi area.

		Determined by Physico-Chemical Analysis				
		C	CL	SC	SCL	SL
Determined by feel Method	C	14	0	4	0	0
	CL	0	0	0	0	0
	SC	0	0	2	6	0
	SCL	0	1	0	20	3
	SL	0	0	0	0	10

How Uncertainty and Classification Estimations May be Used Together

Uncertainty estimation accompanying biophysical property estimation by remote sensing is a well-researched topic (e.g., [21,59,60]). The Global Soil Map consortium, which aims to make a new digital soil map of the world at fine resolution, proposes to use emerging technologies such as optical remote sensing for soil mapping and predicting soil properties [61]. Uncertainty estimations must be provided in association with soil property estimations [62]. Therefore, the proposed approach providing an estimation of two uncertainties coupled to a soil property classification is in line with actual research in soil science.

One viable option is to use the whole classification (over all bare soil pixels) coupled to the whole uncertainty maps. Another strategy could be finding a compromise between the proportion of classified pixels and the model performance by choosing an acceptable uncertainty threshold. A high proportion of classified pixels may be obtained in accepting high uncertainties in the classification results, and inversely, a low proportion of classified pixels may be obtained in accepting low uncertainties in the classification results (Figure 10). The uncertainty threshold may depend on the use of the soil textural map as an input in a crop model, as a covariable in a digital soil mapping approach or for management decisions.

6. Conclusions

Considering the low temporal variation of soil properties, classification or estimation of soil surface properties based on VNIR-SWIR remote sensing data are usually performed by using only one acquisition during a dry period. This paper shows the benefits of using Sentinel-2 data for classification of soil surface texture as our soil texture classification maps showed good discrimination ability for hillslopes and uplands which correspond to coarse soil texture due to erosion processes, and for valleys which correspond to finer soil texture due to deposition processes. In addition, this work highlights the impact of both the training dataset and the Sentinel-2 acquisition date on soil textural soil classifications. Indeed, both the bootstrap process and Sentinel-2 time series allow the production of uncertainty maps associated with soil property classifications, showing that remote textural classes are mapped with a high confidence level, whereas intermediate soil texture classes are mapped with lower confidence.

This kind of uncertainty map will potentially help to better use soil property maps obtained by VNIR-SWIR remote sensing data by providing confidence levels for soil property maps used as input data for environmental models (e.g., crop models) or in digital soil mapping approaches. Finally, this approach using Sentinel-2 time-series to produce soil surface texture classification with associated uncertainty maps will be tested on other soil properties characterized by limited short-term variations, such as carbonates and iron.

References

1. Baumgardner, M.F. Soil databases. In *Handbook of Soil Sciences: Resource Management and Environmental Impacts*; Huang, P.M., Li, Y., Sumner, M.E., Eds.; CRC Press: Boca Raton, FL, USA, 2011; pp. 21–35.

2. Borrelli, P.; Ballabio, C.; Panagos, P.; Montanarella, L. Wind erosion susceptibility of European soils. *Geoderma* **2014**, *232–234*, 471–478. [CrossRef]

3. Ben-Dor, E.; Banin, A. Near infrared analysis (NIRA) as a simultaneously method to evaluate spectral featureless constituents in soils. *Soil Sci.* **1995**, *159*, 259–269. [CrossRef]

4. 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]

5. Baumgardner, M.F.; Silva, L.F.; Biehl, L.L.; Stoner, E.R. Reflectance properties of soils. *Adv. Agron.* **1985**, *38*, 1–44.

6. 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]

7. Stevens, A.; Udelhoven, T.; Denis, A.; Tychon, B.; Lioy, R.; Hoffmann, L.; van Wesemael, B. Measuring soil organic carbon in croplands at regional scale using airborne imaging spectroscopy. *Geoderma* **2010**, *158*, 1–2. [CrossRef]

8. Gomez, C.; Coulouma, G.; Lagacherie, P. Regional predictions of eight common soil properties and their spatial structures from hyperspectral Vis–NIR data. *Geoderma* **2012**, *189–190*, 176–185. [CrossRef]

9. Gomez, C.; Viscarra Rossel, R.A.; McBratney, A.B. Soil organic carbon prediction by hyperspectral remote sensing and field VNIR/SWIR spectroscopy: An Australian case study. *Geoderma* **2008**, *146*, 403–411. [CrossRef]

10. Lu, P.; Wang, L.; Niu, Z.; Li, L.; Zhang, W. Prediction of soil properties using laboratory VIS–NIR spectroscopy and Hyperion imagery. *J. Geochem. Explor.* **2013**, *132*, 26–33. [CrossRef]

11. Selige, T.; Bohner, J.; Schmidhalter, U. High resolution topsoil mapping using hyperspectral image and field data in multivariate regression modeling procedures. *Geoderma* **2006**, *136*, 235–244. [CrossRef]

12. Lagacherie, P.; Baret, F.; Feret, J.-B.; Madeira Netto, J.; Robbez-Masson, J.-M. Estimation of soil clay and calcium carbonate using laboratory, field and airborne hyperspectral measurements. *Remote Sens. Environ.* **2008**, *112*, 825–835. [CrossRef]

13. Shabou, M.; Mougenot, B.; Chabaane, Z.L.; Walter, C.; Boulet, G.; Aissa, N.B.; Zribi, M. Soil Clay Content Mapping Using a Time Series of Landsat TM Data in Semi-Arid Lands. *Remote Sens.* **2015**, *7*, 6059–6078. [CrossRef]

14. Liao, K.; Xu, S.; Wu, J.; Zhu, Q. Spatial estimation of surface soil texture using remote sensing data. *Soil Sci. Plant Nutr.* **2013**, *59*, 488–500. [CrossRef]

15. Vaudour, E.; Gomez, C.; Fouad, Y.; Lagacherie, P. Sentinel-2 image capacities to predict common topsoil properties of temperate and Mediterranean agroecosystems. *Remote Sens. Environ.* **2019**, *223*, 21–33. [CrossRef]

16. Gomez, C.; Adeline, K.; Bacha, S.; Driessen, B.; Gorretta, N.; Lagacherie, P.; Roger, J.M.; Briottet, X. Sensitivity of clay content prediction to spectral configuration of VNIR/SWIR imaging data, from multispectral to hyperspectral scenarios. *Remote Sens. Environ.* **2018**, *204*, 18–30. [CrossRef]

17. Gomez, C.; Oltra Carrio, R.; Lagacherie, P.; Bacha, S.; Briottet, X. Sensitivity of soil property prediction obtained from Hyperspectral Vis-NIR imagery to atmospheric effects and degradation in image spatial resolutions. *Remote Sens. Environ.* **2015**, *164*, 1–15. [CrossRef]

18. Zhai, Y.; Thomasson, J.A.; Boggess, J.E.; Sui, R. Soil texture classification with artificial neural networks operating on remote sensing data. *Comput. Electron. Agric.* **2006**, *54*, 53–68. [CrossRef]

19. Dematte, J.A.M.; Alves, M.R.; da Terra, F.S.; Bosquilia, R.W.D.; Fongaro, C.T.; da Barros, P.P.S. Is It Possible to Classify Topsoil Texture Using a Sensor Located 800 km Away from the Surface? *Rev. Bras. De Ciência Do Solo* **2016**, *40*. [CrossRef]

20. Bellon-Maurel, V.; Fernandez-Ahumada, E.; Palagos, B.; Roger, J.M.; McBratney, A. Prediction of soil attributes by NIR spectroscopy. A critical review of chemometric indicators commonly used for assessing the quality of the prediction. *TrAC Trends Anal. Chem.* **2010**, *29*, 1073–1081. [CrossRef]

21. Gomez, C.; Drost, A.P.A.; Roger, J.M. Analysis of the uncertainties affecting predictions of clay contents from VNIR/SWIR hyperspectral data. *Remote Sens. Environ.* **2015**, *156*, 58–70. [CrossRef]

22. Gholizadeh, A.; Žižala, D.; Saberioon, M.; Borůvka, L. Soil organic carbon and texture retrieving and mapping using proximal, airborne and Sentinel-2 spectral imaging. *Remote Sens. Environ.* **2018**, *218*, 89–103. [CrossRef]

23. Castaldi, F.; Hueni, A.; Chabrillat, 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]

24. Gao, Q.; Zribi, M.; Escorihuela, M.J.; Baghdadi, N. Synergetic Use of Sentinel-1 and Sentinel-2 Data for Soil Moisture Mapping at 100 m Resolution. *Sensors* **2017**, *17*, 1966. [CrossRef] [PubMed]

25. El Hajj, M.; Baghdadi, N.; Zribi, M.; Bazzi, H. Synergic use of Sentinel-1 and Sentinel-2 images for operational soil moisture mapping at high spatial resolution over agricultural areas. *Remote Sens.* **2017**, *9*, 28. [CrossRef]

26. Traganos, D.; Reinartz, P. Mapping Mediterranean seagrasses with Sentinel-2 imagery. *Mar. Pollut. Bull.* **2018**, *134*, 197–209. [CrossRef] [PubMed]

27. Lacroix, P.; Bièvre, G.; Pathier, E.; Kniess, U.; Jongmans, D. Use of Sentinel-2 images for the detection of precursory motions before landslide failures. *Remote Sens. Environ.* **2018**, *215*, 507–516. [CrossRef]

28. Navarro, G.; Caballero, I.; Silva, G.; Parra, P.-C.; Vázquez, Á.; Caldeira, R. Evaluation of forest fire on Madeira Island using sentinel-2A MSI imagery. *Int. J. Appl. Earth Obs. Geoinf.* **2017**, *58*, 97–106. [CrossRef]

29. Vrieling, A.; Meroni, M.; Darvishzadeh, R.; Skidmore, A.K.; Wang, T.; Zurita-Milla, R.; Oosterbeek, K.; O'Connor, B.; Paganini, M. Vegetation phenology from Sentinel-2 and field cameras for a Dutch barrier island. *Remote Sens. Environ.* **2018**, *215*, 517–529. [CrossRef]

30. Rozenstein, O.; Haymann, N.; Kaplan, G.; Tanny, J. Estimating cotton water consumption using a time series of Sentinel-2 imagery. *Agric. Water Manag.* **2018**, *207*, 44–52. [CrossRef]

31. Vuolo, F.; Neuwirth, M.; Immitzer, M.; Atzberger, C.; Ng, W.T. How much does multi-temporal Sentinel-2 data improve crop type classification? *Int. J. Appl. Earth Obs. Geoinf.* **2018**, *72*, 122–130. [CrossRef]

32. 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]

33. 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]

34. 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]

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. Gaillardet, J.; Braud, I.; Hankard, F.; Anquetin, S.; Bour, O.; Dorfliger, 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*, 180067. [CrossRef]

37. Barbiero, L.; Kumar, M.S.M.; Violette, A.; Oliva, P.; Braun, J.J.; Kumar, C.; Furian, S.; Babic, M.; Riotte, J.; Valles, V. Ferrolysis induced soil transformation by natural drainage in vertisols of sub-humid South India. *Geoderma* **2010**, *156*, 173–188. [CrossRef]

38. Louis, J.; Debaecker, V.; Pflug, B.; Main-Knorn, M.; Bieniarz, J.; Mueller-Wilm, U.; Cadau, E.; Gascon, F. Sentinel-2 Sen2Cor: L2A Processor for Users. In Proceedings of the Living Planet Symposium, Prague, Czech Republic, 9–13 May 2016; pp. 1–8.

39. Rouse, J.W.; Haas, R.H.; Schell, J.A.; Deering, W.D. Monitoring vegetation systems in the Great Plains with ERTS. *Third ERTS Symp.* **1973**, *NASA SP-351*, 309–317.

40. Frampton, W.J.; Dash, J.; Watmough, G.; Milton, E.J. Evaluating the capabilities of Sentinel-2 for quantitative estimation of biophysical variables in vegetation. *ISPRS J. Photogram. Remote Sens.* **2013**, *82*, 83–92. [CrossRef]

41. USDA. *Soil Survey Manual*; Handbook No. 18; USDA: Washington, DC, USA, 2017; ISBN PS57-D36-39BA.

42. Kuhn, M.; Wing, J.; Weston, S.; Williams, A.; Keefer, C.; Engelhardt, A.; Cooper, T.; Mayer, Z.; Kenkel, B.; R Core Team; et al. *Caret: Classification and Regression Training*; R Core Team: Vienna, Austria, 2016.

43. Kovačević, M.; Bajat, B.; Gajić, B. Soil type classification and estimation of soil properties using support vector machines. *Geoderma* **2010**, *154*, 340–347. [CrossRef]

44. Pal, M.; Mather, P.M. Support vector machines for classification in remote sensing. *Int. J. Remote Sens.* **2005**, *26*, 1007–1011. [CrossRef]

45. Mountrakis, G.; Im, J.; Ogole, C. Support vector machines in remote sensing: A review. *ISPRS J. Photogram. Remote Sens.* **2011**, *66*, 247–259. [CrossRef]

46. Cortes, C.; Vapnik, V. Support-vector networks. *Mach. Learn.* **1995**, *20*, 273–297. [CrossRef]

47. Burges, C.J.C. *A Tutorial on Support Vector Machines for Pattern Recognition*; Kluwer Academic Publishers: Boston, MA, USA, 1998; 43p.

48. Congalton, R. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sens. Environ.* **1991**, *37*, 35–46. [CrossRef]

49. Congalton, R.G.; Green, K. *Assessing the Accuracy of Remotely Sensed Data: Principles and Practices*; CRC/Taylor & Francis: Boca Raton, FL, USA, 1988.

50. 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]

51. Efron, B.; Tibshirani, R. *An Introduction to the Bootstrap*; Chapman and Hall: London, UK, 1993.

52. Simpson, E.H. Measurement of diversity. *Nature* **1949**, *163*, 688. [CrossRef]

53. Nocita, M.; Stevens, A.; Noon, C.; van Wesemael, B. Prediction of soil organic carbon for different levels of soil moisture using Vis-NIR spectroscopy. *Geoderma* **2013**, *199*, 37–42. [CrossRef]

54. 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]

55. Bartholomeus, H.; Kooistra, L.; Stevens, A.; Van Leeuwen, M.; VanWesemael, 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]

56. 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]

57. Nagler, P.L.; Daughtry, C.S.T.; Goward, S.N. Plant litter and soil reflectance. *Remote Sens. Environ.* **2000**, *71*, 207–215. [CrossRef]

58. Madeira Netto, J.S.R.; Robbez-Masson, J.M.; Martins, E. Visible–NIR hyperspectral imagery for discriminating soil types in the La Peyne watershed (France). In *Digital Soil Mapping: An Introductory Perspective*; Lagacherie, P., Mc Bratney, A.B., Voltz, M., Eds.; Elsevier: Amsterdam, The Netherlands, 2007.

59. Rocchini, D.; Foody, G.M.; Nagendra, H.; Ricotta, C.; Anand, M.; He, K.S.; Amici, V.; Kleinschmit, B.; Förster, M.; Schmidtlein, S.; et al. Uncertainty in ecosystem mapping by remote sensing. *Comput. Geosci.* **2013**, *50*, 128–135. [CrossRef]

60. Cockx, K.; Van de Voorde, T.; Canters, F. Quantifying uncertainty in remote sensing-based urban land-use mapping. *Int. J. Appl. Earth Observ. Geoinf.* **2014**, *31*, 154–166. [CrossRef]

61. Lagacherie, P.; Gomez, C. Vis-NIR-SWIR Remote Sensing Products as New Soil Data for Digital Soil Mapping. In *Pedometrics. Progress in Soil Science*; Mcbratney, A., Minasny, B., Stockm, U., Eds.; Springer: Cham, Switzerland, 2018.

62. Arrouays, D.; Grundy, M.G.; Hartemink, A.E.; Hempel, J.W.; Heuvelink, G.B.M.; Hong, Y.S.; Lagacherie, P.; Lelyk, G.; McBratney, A.B.; McKenzie, N.J.; et al. GlobalSoilMap: Toward a Fine-Resolution Global Grid of Soil Properties. *Adv. Agron.* **2014**, *125*, 93–134.