Mapping of tank silt application using Sentinel-2 images over the Berambadi catchment (India)

Cécile Gomez, S. Dharumarajan, Philippe Lagacherie, Jean Riotte, Sylvain Ferrant, M. Sekhar, Laurent Ruiz

https://doi.org/10.1016/j.geodrs.2021.e00389

Accepted Proof.

Mapping of tank silt application using Sentinel-2 images over the Berambadi catchment (India).

Gomez[1,2,*] C., Dharumarajan[3] S., Lagacherie[1] P., Riotte[2,4] J., Ferrant[5] S., Sekhar[2,6] M., Ruiz[2,4,7] L.

[1] LISAH, Univ. Montpellier, IRD, INRAE, Institut Agro, Montpellier, France

[2] Indo-French Cell for Water Sciences, IRD, Indian Institute of Science, Bangalore, India

[3] ICAR-National Bureau of Soil Survey and Land Use Planning, Hebbal, Bangalore, India

[4] GET, Université Paul-Sabatier, IRD, CNRS, Toulouse, France

[5] CESBIO, Université Paul-Sabatier, CNRS, CNES, IRD, INRAE, Toulouse, France

[6] Indian Institute of Science, Civil Engineering Department, Bangalore, India

[7] SAS, INRAE, Institut Agro, Rennes, France

[*] Corresponding author. cecile.gomez@ird.fr

Highlights

- We studied an Indian age-old practice: a black-colored tank silt application over red-colored soils
- Soil color maps were produced by supervised classification method using Sentinel-2 images
- Soil color changes between two S2 images were used as proxy to identify tank silt applications
- The proposed approach allows mapping soil colors with correct accuracy
- This anthropogenic soil modification was identified over 202 fields

Abstract

Mapping soil properties is becoming more and more challenging due to the increase in anthropogenic modification of the landscape, calling for new methods to identify these changes. A striking example of anthropogenic modifications of soil properties is the widespread practice in South India of applying large quantities of silt from dry river dams (or “tanks”) to agricultural fields. Whereas several studies have demonstrated the interest of tank silt for soil fertility, no assessment of the actual extent of this age-old traditional practice exists. Over South-Indian pedological context, this practice is characterized by an application of black-colored tank silt to red-colored soils such as Ferralsols. The objective of this work was to evaluate the usefulness of Sentinel-2 images for mapping tank silt applications, hypothesizing that observed changes in soil surface color can be a proxy for tank silt application. We used data collected in a cultivated watershed in South India including 217 soil surface samples characterized in terms of Munsell color. We used two Sentinel-2 images acquired on February and April 2017. The surface soil color over each Sentinel-2 image was classified into two soil types (“Black” and “Red” soils). A change of soil color from “Red” in February 2017 to “Black” in April 2017 was attributed to tank silt application. Soil color changes were analyzed accounting for possible surface soil moisture changes. The proposed methodology was based on a well-balanced Calibration data created from the initial imbalanced Calibration dataset thanks to the Synthetic Minority Over-sampling Technique (SMOTE) methodology, coupled to the Cost-Sensitive Classification And Regression Trees (Cost-Sensitive CART) algorithm. To estimate the uncertainties of *i)* the two-class classification at each date and *ii)* the change of soil color from “Red” to “Black”, a bootstrap procedure was used providing fifty two-class classifications for each Sentinel-2 image. The results showed that 1) the CART method allowed to classify the “Red” and “Black” soil with correct overall accuracy from both Sentinel-2 images, 2) the tank silt application was identified over 202 fields and 3) the soil color changes were not related to a surface soil moisture change between both dates. With the actual availability of the Sentinel-2 and the past availability of the LANDSAT satellite imageries, this study may open a way toward a simple and accurate method for delivering tank silt application mapping and so to study and possibly quantify retroactively this farmer practice.

Keywords: tank silt application; Vertisols; Ferralsols; soil color classification; CART; SMOTE; Sentinel-2; India.

1. Introduction

The soil has an integral part to play in the global environmental sustainability challenges of food security, water security, climate stability, biodiversity, and ecosystem service delivery (McBratney et al, 2014) and accurate mapping of soil properties is crucial for adequate management both at global and very local levels. For many years, digital soil mapping represents an alternative to conventional soil survey, to produce soil predictions maps and associated uncertainties, exploiting large sets of spatial data (e.g., Digital Elevation Model, optic satellite data) combined to recent or ancillary data (McBratney et al., 2003).

To face an increasing demand of food, water scarcity and soil health degradation, farmers adapt their crop types, crop rotations, water supplies and practices (Hardaker, 2004) and these land use practices linked to soil management may determine frequently changes in soil class (Dazzi and Lo Papa, 2015), as the Solonchaks developed from Cambisols in arid environments because of irrigation with saline water, or Regosols derived from the truncation of Cambisols due to soil erosion.

India is highlighted as one of the most risk-prone countries for water scarcity, declining soil fertility through land degradation and climate change impacts (Roberts, 2001). A traditional "tank system", composed of cascades of reservoirs along valleys, has been used for centuries not only to harvest water for irrigation but also to trap sediments to limit erosion losses at the catchment scale by restoring tank-trapped sediments to agricultural fields (Gunnell and Krishnamurthy, 2003). While this latter practice used to be mostly limited to fields in the vicinity of tanks, the recent development of motorization (excavators and tractors) and of governments-implemented vast programs of tanks desilting (DHAN Foundation) have given this age-old practice a new momentum. However, to our knowledge, no estimate of the temporal and spatial evolutions of this is available.

While the primary objective of tank desilting is to increase the water storage capacity of these tanks, several studies have demonstrated the interest of restoring tank silt to agricultural fields in India, as it improves soil fertility (DHAN Foundation; Karanam et al., 2008; Osman et al., 2009; Obi Reddy et al., 2017), specifically increasing soil water holding capacity (Deshmukh et al., 2019), organic carbon and available nutrient status (Patil et al., 2017). Moreover, these benefits of tank sediments for the soil quality has been also demonstrated in others parts of the word such as in Peru (Walter et al., 2012) and Poland (Baran et al., 2019).

The potential of this technique is large in India: The number of irrigation tanks is estimated around 0.3 million all over India (Reddy et al., 2018), with 35% located in south India (Tamil Nadu, Karnataka and Andhra Pradesh) (Narayanamoorthy A., 2007). From the 1990s, tank irrigation declined sharply at the expense of groundwater irrigation, which accounted for almost 60% of the irrigated area after only 10 years (Shah et al., 2003), thanks to the development of submersible pump technology. However, tank irrigation is still one of the major strategies for adapting to rainfall variability (CWC, 2010; Reddy et al., 2018), and has a great potential provided sustainable solution are found to resolve the problem of silting. In three of the southern Indian states (Andhra Pradesh, Telangana, Karnataka and Tamil Nadu), about 140 000 tanks are silted up and their rehabilitation could both increase irrigation potential and improve soil health (DHAN Foundation).

Over pedological contexts characterized by Vertisols, Ferralsols and Chromic Luvisols like over the Deccan Plateau, tank silt is characterized by a black color and is applied to red-colored soils such as Ferralsols and Chromic Luvisols. Therefore, tank silt applications are easily recognizable over landscapes as they change topsoil color. On the field, soil color is commonly and widely measured using a Munsell soil color chart (Munsell Color Company, 1975) which is a system for categorical qualifications of soil color. This chart was designed to reflect our perception of color and its variations, and is widely used by pedologist as a determinant of soil type. The topsoil color can also be studied by Visible-Near Infrared and Short Wave Infrared (VNIR-SWIR, 400 to 2500 nm) remote sensing as various soil components exhibit spectral response in the visible range of the electromagnetic spectrum, between wavelengths 400 and 700 nm. The topsoil color study by VNIR-SWIR remote sensing may allow to discriminate eroded and non-eroded soils (e.g., Pickup and Nelson, 1984), identify surface efflorescence and salt crust (e.g., De Jong, 1992; Mougenot et al., 1993) and estimate soil organic carbon (e.g., Viscarra Rossel et al., 2006).

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 ten days from 2015 to 2017 and five days since 2017 (European Space Agency, 2015; https://sentinel.esa.int/web/sentinel/missions/sentinel-2). Sentinel-2 images are acquired with a spatial resolution of 10 m to 60 m on 13 spectral bands in the VNIR-SWIR spectral domain. Several studies have proven the relevance of individual Sentinel-2 images for soil studies, including primary soil properties mapping such as soil organic mapping (e.g., Gholizadeh et al., 2018; Vaudour et al., 2019) and soil texture (e.g., Gomez et al., 2019), and soil salinity retrieval based on an Electrical Conductivity mapping (e.g., Taghadosi et al., 2019; Wang et al., 2019). Some studies have studied the multitemporal dimension of Sentinel-2 images for soil studies, and showed that it 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., Vaudour et al., 2019), ii) elaborate mosaic images of bare soil area (e.g., Loiseau et al., 2019) and iii) estimate uncertainties of permanent soil properties predictions (Gomez et al., 2019).

The objective of this work is to evaluate the utility of Sentinel-2 images for mapping tank silt application over red soils. This work is based on a changes analysis of soil color from a Sentinel-2 image to another one and was carried out in an Indian cultivated region (Berambadi catchment, Karnataka state). It used two Sentinel-2 images acquired on February and April 2017 and 217 soil surface samples, collected over the study area and characterized in terms of Munsell color.

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 km^2. 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 (AMBHAS, BVET, Sekhar et al., 2016; Tomer et al., 2015), which is part of the OZCAR network (Gaillardet et al., 2018). 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). Red soils (Ferralsols and Chromic Luvisols, Figure 2a) cover the uplands and hillslopes, while black soils (Vertisols and Vertic intergrades, Figure 2b) are found mostly in the valley bottoms (Barbiero et al., 2010). Uplands and hillslopes are mainly characterized by coarse soil texture (sandy loam) due to erosion processes, whereas valleys bottoms are mainly characterized by finer soil texture (clay) due to deposition processes (Gomez et al., 2019). Finally, every year some farmers apply tank silt over the Berambadi subcatchment during the summer season (Figure 2c).

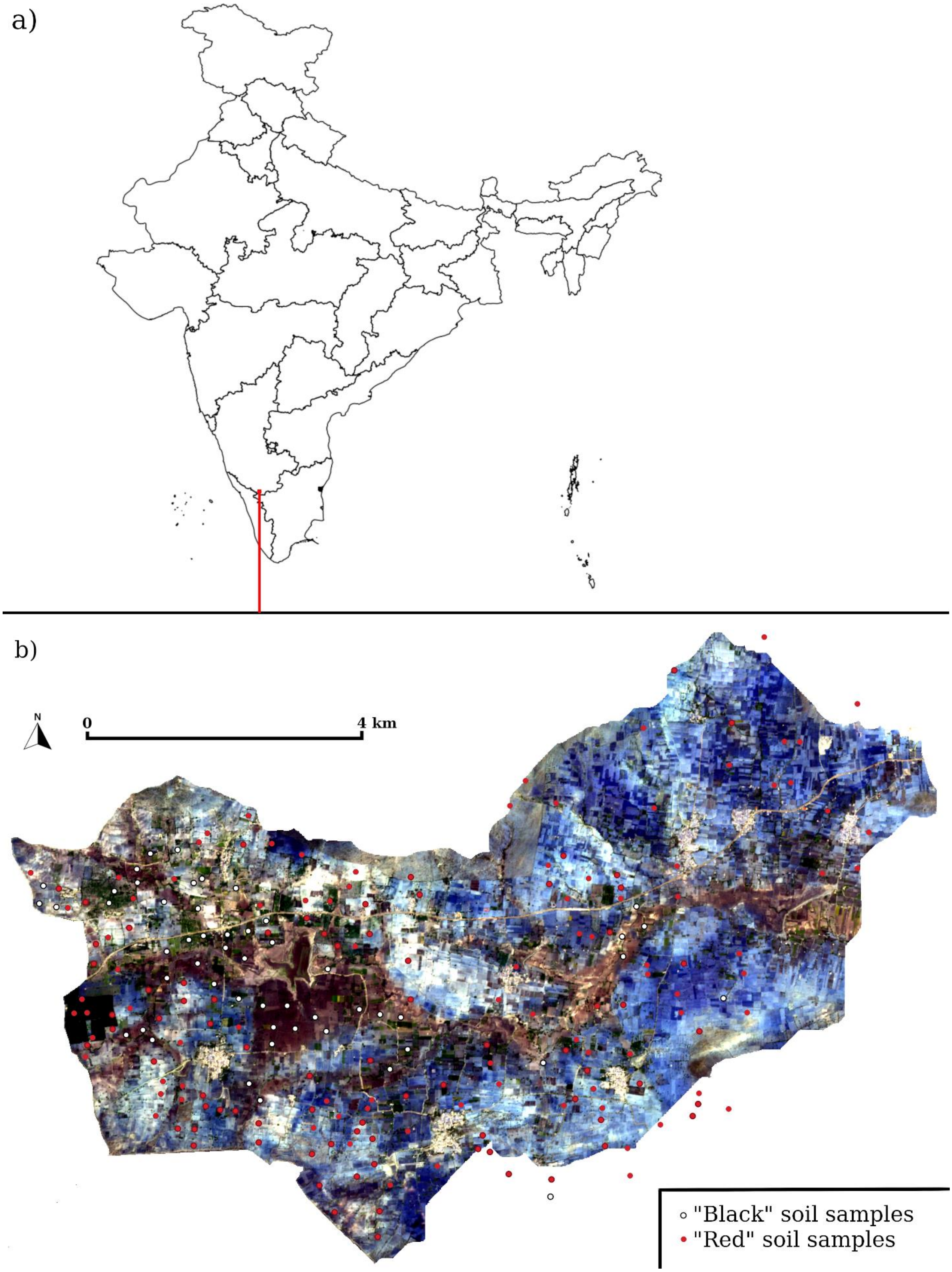

Figure 1: a) Location of the Berambadi catchment in India (Karnataka State) and b) 217 soil surface samples collected on the field and plotted on the Sentinel 2 image (in false color) acquired on 24th of April 2017**.**

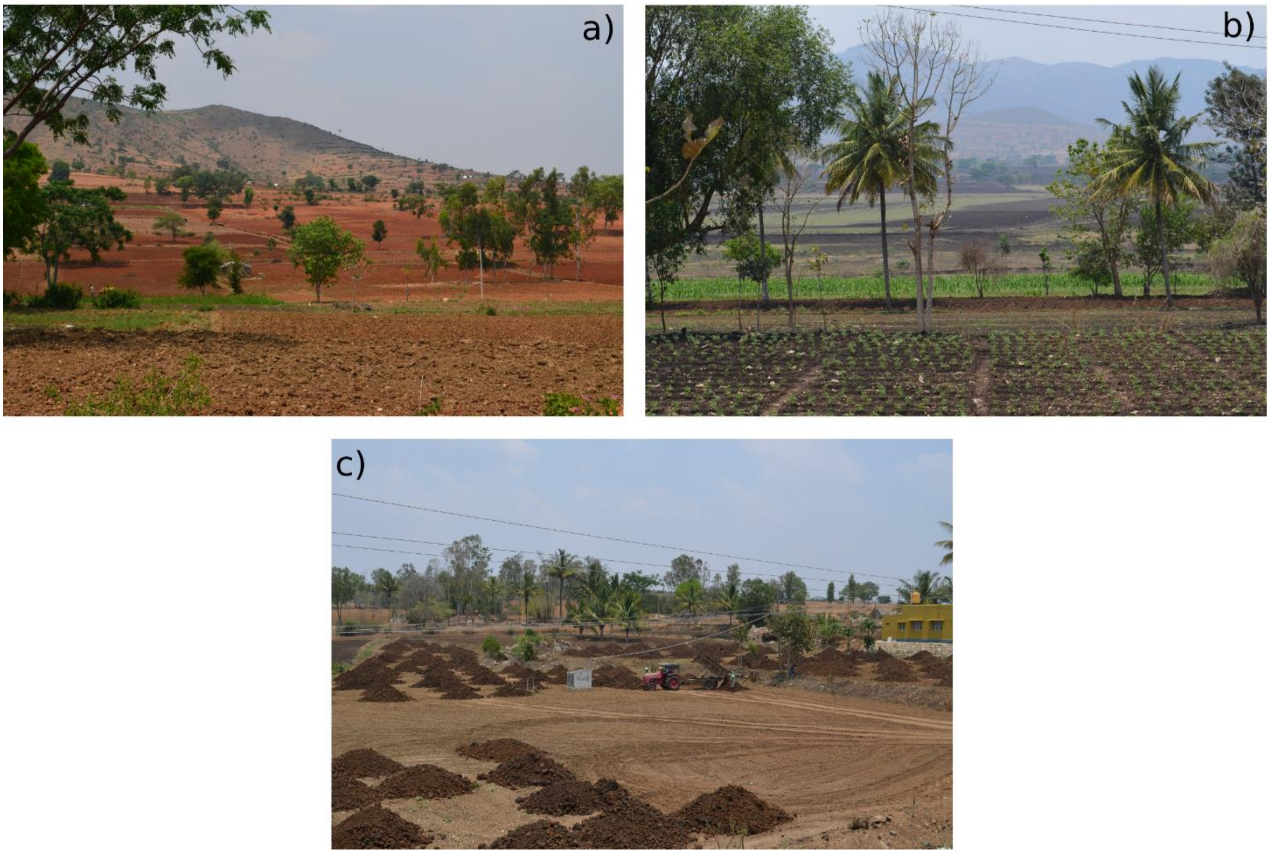

Figure 2: Example a) “Red” soils, b) “Black” soils and c) Tank silt application over “Red” soils**.**

2.2 Sentinel-2 images

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 (Table 1) (European Space Agency, 2015; https://sentinel.esa.int/web/sentinel/missions/sentinel-2). 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 directly used to perform atmospheric corrections and cloud detection (Table 1). Here, atmospheric and topographic corrections were performed with Sen2Cor software (Louis et al., 2016), and the level-2A images (bottom of atmosphere terrain corrected reflectance) were used to perform soil color classification and new anthroposol identification. The six spectral bands initially acquired with 20 m spatial resolution were resampled to 10 m.

Table 1: Specifications of the Sentinel-2 images. In italic characters are highlighted the spectral bands used in this work.

Spectral band	**Central wavelength (nm)**	**Band width (nm)**	**Original spatial resolution** (m)
B1	443	20	60
B2	*490*	*65*	*10*
B3	*560*	*35*	*10*
B4	*665*	*30*	*10*
B5	*705*	*15*	*20*
B6	*740*	*15*	*20*
B7	*783*	*20*	*20*
B8	*842*	*115*	*10*
B8A	*865*	*20*	*20*
B9	945	20	60
B10	1380	30	60
B11	*1610*	*90*	*20*
B12	*2190*	*180*	*20*

Two Sentinel-2 images covering the Berambadi catchment during the dry season of 2017 were selected. The study year, 2017, was selected because the tanks were dry over the study area during the dry season, so the period was suitable for tank silt applications. Both Sentinel-2 data were acquired under clear sky conditions. A first image was acquired on the 3rd of February 2017 after 6 days without any rainfall (purple vertical line, Figure 3) and a second image on the 24th of April 2017 after 3 days without any rainfall (red vertical line, Figure 3). Both images were previously described and used in Gomez et al. (2019).

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 865 nm and 665 nm (Rouse et al. 1973; Frampton et al., 2013). In April 2017, a field campaign was conducted over the Berambadi catchment, allowing to localize some cultivated fields (covered by green vegetation) and ploughed fields (covered by bare soil). Based on these observations and linked to the NDVI calculated over the S2 image acquired on 24th of April 2017, a threshold of NDVI has been determined to separate the bare soil pixels to the vegetated pixels. This threshold of NDVI was fixed at 0.3 and used for both images. Urban areas and lakes were also masked using a land use map available for the study area (e.g., AMBHAS Team, 2015). We selected the pixels which were both common to the two images and in bare soil condition (65 % of the total) to analyze the changes of soil color from a date to another one.

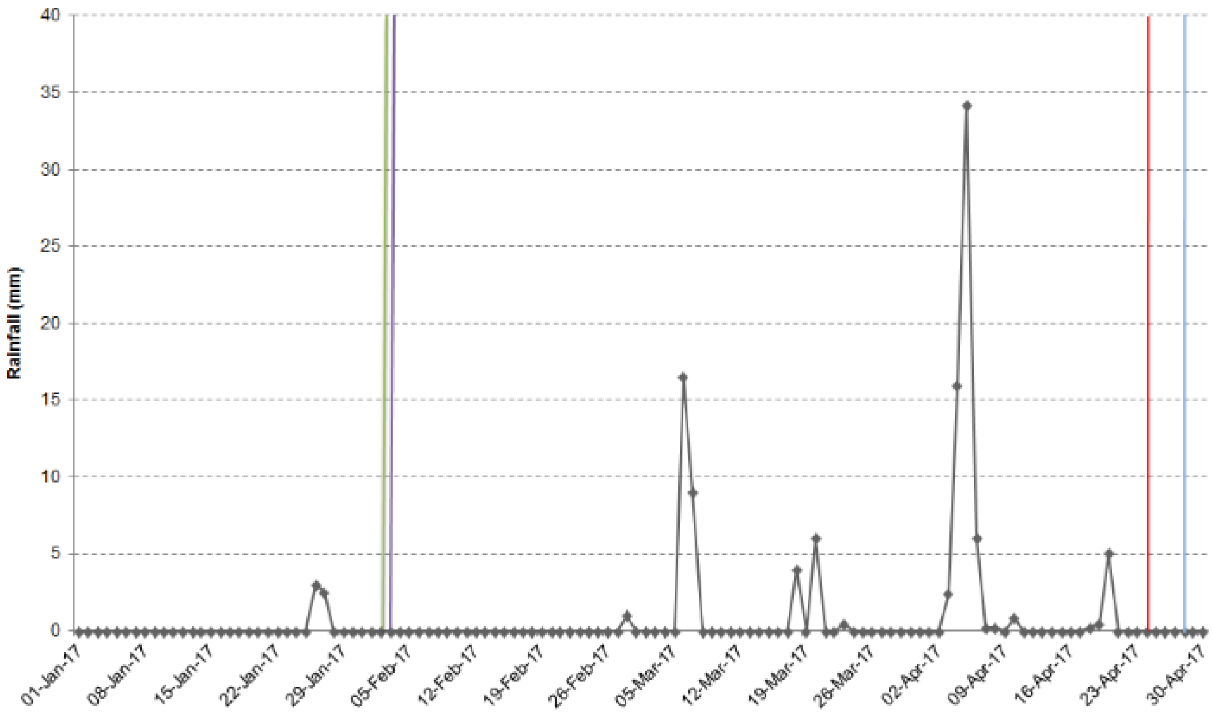

Figure 3: Rainfall (in mm) measured every day at 6 am, over Maddur village (Berambadi). The green, purple, red and blue vertical lines localize in time the acquisition date of the first SMAP image (2nd of February 2017), both first Sentinel-2 and Sentinel-1 images (3rd of February 2017),

the second Sentinel-2 (24th of April 2017) and the second Sentinel-1 image (28th of April 2017), respectively.

2.3 Soil data

217 soil profiles were studied over the Berambadi catchment (Figure 1b), comprising 177 soil profiles studied from March to July 2015, and 40 soil profiles studied from the 25th March to the 7th April 2017. The profiles sampling was based on the soil physiographic relationship (relationship between soil and landform) over the study area. First, several transects from ridge to valley were selected over the Berambadi catchment. Secondly, the soil profiles locations were selected along each transect, taking into account that these soil profiles are located at closely spaced intervals and represent the changes in land use, slope, erosion features, stones and gravels along the transect.

This study was based on 217 soil surface samples where each soil surface sample corresponds to the surface layer of a soil profile. As the surface layers had a depth from 10-30 cm (plough layer or top horizon), our soil surface samples had a depth from 10-30 cm. Each soil profile was located with a GPS and the soil surface sample was associated to a Munsell soil color, following the Munsell Chart. From the Munsell soil colors determined on the field, the 217 soil samples were firstly post-classified in four soil color classes (Table 2): Yellow, Brown, Red and Black soils. As Yellow and Brown are shades of Red color developed from granitic parent material, these three classes were then grouped as “Red” soil in this paper, and the following work was based only on two soil types: “Red” and “Black” soil (Table 2, Figures 2a and b).

Among the 217 soil surface samples associated with Munsell soil color, 130 were located over the common bare soil pixels between both Sentinel-2 images. Among these 130 soil surface samples, 25 correspond to “Black” soils and 105 correspond to “Red” soils, and so were used in this work to analyze the changes of soil color from a date to another one. As the number of “Black” soil samples is much less important than the number of “Red” soil samples, the dataset is imbalanced, which reflects the imbalanced repartition of soils types over the study area, as Black soils cover a minority of the area (mostly in the valley bottoms) (Barbiero et al., 2010).

Table 2: Correspondence between the observed Munsell soil color, the soil color classes and the soil type.

Munsell soil color	Soil Color Class	Soil type
10YR4/3, 10YR4/4, 10YR4/6, 10YR4/1, 10YR4/2, 10YR5/3, 10YR5/4, 10YR6/4	**Yellow soil**	**“Red” soil**
2.5YR3/2, 2.5YR3/3, 2.5YR3/4, 2.5YR3/6, 2.5YR4/4, 2.5YR4/6, 2.5YR5/4, 5YR3/2, 5YR3/3,5YR3/4, 5YR4/3, 5YR4/4, 5YR4/6, 5YR5/4, 5YR5/6	**Red soil**	
all the 7.5YR	**Brown soil**	
10YR3/1, 10YR3/2, 10YR3/4, 10YR4/1, 10YR4/2	**Black soil**	**“Black” soil**

2.4 Agricultural Parcel boundaries

A parcel delineation map was created by Sharma et al. (2018) based on the following methodology. First, individual parcels have been identified using a 2 m spatial resolution google earth image acquired in 2012. Secondly, all individual parcels were re-examined using successively three LISS IV satellite images acquired in 2013, 2014 and 2015 with 5 m of spatial resolution and one SPOT-6 satellite image acquired in 2016 with 1.5 m of spatial resolution. This parcel delineation work allowed identifying a total of 18178 polygons over the Berambadi catchment, with an average size of fields of 0.27 ha (Sharma et al., 2018).

2.5 Surface soil moisture map

Two soil moisture maps were created following the procedure presented in Tomer et al. (2016) based on Sentinel-1 and Soil Moisture Active Passive (SMAP) data, with a 20 m spatial resolution. This process used the closest Sentinel-1 and SMAP data from Sentinel-2 data to estimate surface soil moisture. A first surface soil moisture map was created based on a Sentinel-1 data acquired on 3rd of February, so simultaneously with our first Sentinel-2 data and after 6 days without any rainfall (red vertical line, Figure 3), and a SMAP data acquired on 2nd of February, so one day before our first Sentinel-2 data (green vertical line, Figure 3). A second surface soil moisture map was created based a Sentinel-1 data and a SMAP data both acquired on 28th of April, so four days after our second Sentinel-2 data and after 7 days without any rainfall (blue vertical line, Figure 3). As no rainfall have been recorded during 6 and 7 days before the Sentinel-1 data acquired on 3rd of February and 28th of April, respectively, variations in surface soil moisture can be supposed to be due only to irrigation. The estimated soil moisture content varies from 0 to 1, where a value of 0 correspond to the lowest soil moisture content and a value of 1 corresponds to the highest soil moisture content.

3. Methods

To map the tank silt application from February to April 2017, the proposed approach followed 3 main steps. In a first step, *N* Calibration and Validation databases were built with a bootstrap approach from both Sentinel-2 image and the corresponding observed Munsell soil colors (Figure 4b). A classification model was built from each of the *N* Calibrations dataset and then validated from the associated Validation dataset (Figure 4b). So *N* maps of “Black” soils and “Red” soils were provided (Figure 4a) from both selected Sentinel-2 image. The classification of “Black” and “Red” soils using the S2 image acquired on the 3rd of February 2017 and the n^{th} Calibration and Validation database, was called *Map_1_n*. The classification of ‘”Black” soils and “Red” soils using the Sentinel-2 image acquired on the 24th of April 2017 and the n^{th} Calibration and Validation database, was called *Map_2_n*. In a second step, *N* maps of changes between beginning (3rd of February

2017) and end (24th of April 2017) of the period were provided based on differences between *Map_1_n* and *Map_2_n* (Figure 4c). In a third and last step, a map of the frequency of changes was provided (Figure 4d), based on the *N* maps of changes (Figure 4c).

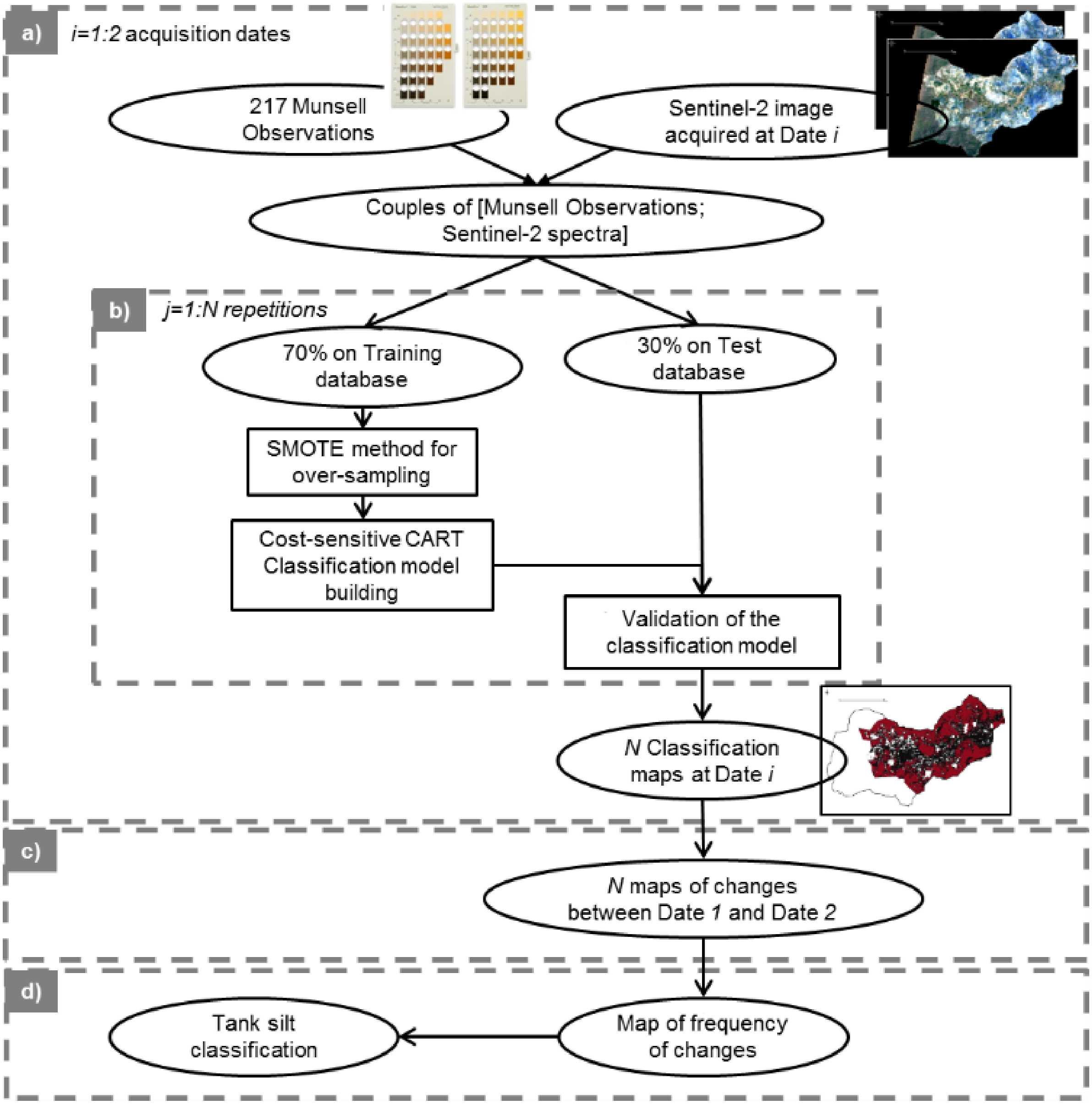

Figure 4: Workflow of the tank silt application mapping.

3.1. Classification Models

A standard normal variate (SNV) was applied to the sentinel-2 spectra consisting of substracting the mean of this spectrum and dividing the standard deviation of this spectrum to every spectral value along the spectrum. The SNV is an effective pretreatment method for baseline correction of diffuse reflection Vis-NIR spectra (Barnes et al.,1989).

Class-imbalanced datasets occur in many real-world applications. With imbalanced data sets, data mining learning algorithms produce degenerated models that do not take into account the minority class as most data mining algorithms assume balanced data set (e.g., Sun et al., 2009; Ali et al., 2015, Taghizadeh-Mehrjar et al., 2019). To deal with the class-imbalance problem, some solutions were proposed at the algorithmic level (Chawla et al. 2004) which include adjusting the costs of the various classes so as to counter the class imbalance, adjusting the probabilistic estimate at the tree leaf (when working with decision trees) or adjusting the decision threshold. Cost-sensitive learning, which takes the misclassification costs (and possibly other types of cost) into consideration, is often used to deal with datasets with very imbalanced class distribution (Japkowicz and Stephen, 2002).

Here, we used the Cost-Sensitive Classification And Regression Trees (CART) algorithm, initially proposed by Breiman et al. (1984), a nonparametric prediction method based on a binary decision tree algorithm dealing with imbalanced Calibration data. It processes continuous and categorical attributes and target, using the Gini splitting rule to search the best possible variable to split the node into two child nodes and grow the trees to their maximum size until no splits are possible. The Cost-Sensitive CART algorithm is the adaptation of the CART algorithm where the misclassification cost plays an important role in the learning process. The algorithm targets the cases where the cost of incorrectly classifying a minority (positive) class will have a higher cost than the cost of incorrectly classifying a majority (negative) class (Breiman et al., 1984). Details about the cost-sensitive CART algorithm can be found in Steinberg B. (2009). In our classification, the validation accuracy was performed using 10-fold cross-validation, selecting the model with the lowest cost-complexity metric. A value of 1 was affected to pixels classified as “Black” soil (so considered as the positive class) and a value of -1 was affected to pixels classified as “Red” soil (so considered as the negative class). The classification models were built using the *train* function provided in caret package (Kuhn et al., 2016).

3.2. Calibration and Validation Database preparation

The spectral information from the 130 pixels for which soil color observations are available was extracted for each Sentinel-2 image (Figure 4a). For each classification, these couples of spectral information and corresponding soil color were split into Calibration (70%) and Validation (30%) datasets using a stratified random sampling for Calibration and Validation data to follow a similar distribution of soil color classes (Figure 4b).

To deal with the class-imbalance problem, some solutions were previously proposed at the data level in addition to the ones proposed at the algorithmic level (Chawla et al. 2002). At the data level, these solutions include many different forms of re-sampling such as random oversampling with replacement, random undersampling, directed oversampling or creation of synthetic data. The simplest method to increase the size of the minority class corresponds to random over-sampling, that is, a non-heuristic method that balances the class distribution through the random replication of positive examples. Nevertheless, since this method replicates existing examples in the minority class, overfitting is more likely to occur. In this work, we used the Synthetic Minority Over-sampling Technique (SMOTE), proposed by Chawla et al. (2002) to create balanced Calibration data from the initial Imbalanced Calibration dataset build by splitting the 130 couples of spectral information and corresponding soil color (Figure 4b). This SMOTE preprocessing algorithm was used as an over-sampling approach in which the minority class is over-sampled by creating synthetic data rather than by over-sampling with replacement and is considered as standard in the framework of learning from imbalanced data (Douzas et al., 2019b). The perc.over value was selected at 100, to guarantee that, after applying SMOTE, the minority class would have the exact same number of training samples as the majority class (50% of samples associated to “Black” color and 50% of samples associated to “Red” color). The k value was set to 5 as the number of minority class samples was larger than 5. The balanced-datasets were built using the *smote* function provided in caret package (Kuhn et al., 2016).

3.3. Classification performance measures

The overall accuracy was calculated for each classification model and from each Validation dataset to measure the performance of the classifications (Cohen et al., 1960) (Figure 4b). The overall accuracy is commonly measured as the percentage of pixels correctly classified in the Validation dataset. The kappa coefficient was calculated on observed dataset to compare the observed accuracy with the expected accuracy resulting from randomness. According to Landis and Koch (1977), a value < 0 is indicating no agreement: 0–0.20 as slight, 0.21–0.40 as fair, 0.41–0.60 as moderate, 0.61–0.80 as substantial, and 0.81–1 as almost perfect agreement. The Sensitivity and Specificity were calculated for each classification model and from each Validation dataset. Sensitivity is the metric that evaluates the model's ability to predict true positives, which is the class of "Black" soil in our case, and is calculated as:

Sensitivity = True Positives / (True Positives + False Negatives) [1]

Where the *True Positives* corresponds to the number of pixels observed as "Black" soil and correctly classified in the "Black" soil class and *False Negatives* corresponds to the number of pixels observed as "Black" soil but misclassified in the "Red" soil class.

The Specificity, also called the true negative rate, which is the rate of "Red" soil correctly classified in our case, was calculated as following:

Specificity = True Negatives / (True Negatives + False Positives) [2]

Where the *True Negatives* corresponds to the number of pixels observed as "Red" soil and correctly classified in the "Red" soil class and *False Positives* corresponds to the number of pixels observed as "Red" soil but misclassified in the "Black" soil class. These performance measures were calculated using the *confusionMatrix* function provided in caret package (Kuhn et al., 2016).

3.4. Bootstrap

A bootstrap procedure was applied to the original dataset to define *N* sets of Calibration and Validation subsets for each Sentinel-2 image, where *N* is equal to 50 (Figure 4b) (Efron et al., 1993). The bootstrapping involved repeated stratified random sampling for Calibration and Validation data to follow a similar distribution of both classes, with replacement of the available data and the SMOTE over-sampling approach. For each bootstrap iteration, a classification model was fitted and applied to generate a soil classification (Figure 4b). So 50 maps of “Black” soils and “Red” soils were provided (Figure 4a) from each selected Sentinel-2 image. This bootstrap procedure follows the common way in soil properties mapping initiated by Brodsky et al. (2013) and pursued by Gomez et al. (2015), Viscarra Rossel and Hicks (2015) and Gomez et al. (2019).

The classification performance measures were calculated for the 50 tests. The mean and standard deviation of the classification performance measures were analyzed to evaluate, respectively the models performances and the variabilities from each S2 image.

After the bootstrap procedure, a single-date map of the most frequent class obtained for each pixel of each Sentinel-2 image was produced.

3.5. Temporal color changes mapping method

The *N* classifications based on the Sentinel-2 image acquired on the 3rd of February 2017 (*Map_1_n*) may reflect the initial mapping of the “Black” and “Red” soils before tank silt applications. The *N* classifications based on the Sentinel-2 image acquired on the 24th of April 2017 (*Map_2_n*) may reflect the mapping of both “Black”, “Red” soils and tank silt applications.

To highlight the tank silt application over the study area, *N* maps of soil color changes between the 3rd of February 2017 and the 24th of April 2017 were provided based on differences between *Map_1_n* and *Map_2_n* (Figure 4c). A value of 0 was affected to pixels where no change of soil color is identified between the *Map_1_n* and the *Map_2_n*. A value of 1 was affected to pixels classified as “Red” soils on *Map_1_n* and then classified as “Black” soils on *Map_2_n*. A value of -1 was affected to pixels classified as “Black” soils on *Map_1_n* and then classified as “Red” soils on *Map_2_n*.

Then the mean of frequency of changes over the 50 iterations was calculated at the field scale (Figure 4d) using the agricultural field boundaries (section 2.4), based on the *N* maps of soil color changes (Figure 4c). Only fields covered by more than 30 % of bare soil pixels were analyzed. Changes over fields covered by less than 30 % of bare soil pixels were not studied and mapped. High and positive values (from 40 to 50) correspond to frequent changes of soil color from “Red” to “Black”, which may be considered as tank silt application. Null values correspond to none change of soil color between both dates along the bootstrap. Low and negative values (from -40 to -50) correspond to frequent changes of soil color from “Black” to “Red”.

3.6. Analysis of color changes regarding to surface soil moisture changes

The mean of relative surface soil moisture difference $Diff_{moisture}$ between the 3rd of February and the 28th of April was calculated, for each field *f* using the agricultural field boundaries (section 2.4), following:

$$Diff_{moisture_f} = \frac{1}{p}\sum_{i=1}^{p}(MoistureT1_i - MoistureT2_i) \quad [3]$$

Where $MoistureT1_i$ and $MoistureT2_i$ are the relative surface soil moistures estimated on the 3rd of February and on the 28th of April, respectively, both on pixel *i* over the field *f*. As for creating the map of mean of frequency of changes over the 50 iterations (section 3.5), only fields covered by more than 30 % of bare soil pixels were analyzed. Changes over fields covered by less than 30 % of bare soil pixels were not studied and mapped. When the surface soil moisture contents are higher on the 3rd of February 2017 ($MoistureT1$) than on the 24th of April 2017 ($MoistureT2$), the values of $Diff_{moistude}$ are positive. Inversely, when the surface soil moisture contents are lower on the 3rd of February 2017 ($MoistureT1$) than on the 24th of April 2017 ($MoistureT2$), the values of $Diff_{moistude}$ are negative.

4. Results

4.1Classifications performances

50 classification models were built from the Sentinel-2 image acquired on the 3rd of February 2017 and using balanced data obtained by the SMOTE method and the cost-sensitive CART classification method. These 50 models provided 50 classifications of “Black” and “Red” soils with accurate overall accuracies (mean overall accuracy = 0.81) and moderate agreement (mean kappa = 0.51) (Figure 5). The standard deviations of overall accuracies and kappa over the 50 models derived from the Sentinel-2 image acquired on the 3rd of February 2017 were low, with values of

0.05 and 0.12, respectively. The averaged sensitivity and specificity of these classification models were around 0.75 and 0.83 respectively.

50 classification models were built from the Sentinel-2 image acquired on the 24th of April 2017 and using balanced data obtained by the SMOTE method and the cost-sensitive CART classification method. These 50 classification models provided 50 classifications of “Black” and “Red” soils with also accurate overall accuracy (mean overall accuracy = 0.76) and moderate agreement (mean kappa = 0.40) (Figure 5). The standard deviations of overall accuracies and kappa over the 50 models derived from the Sentinel-2 image acquired on the 24th of April 2017 were low, with values of 0.12 and 0.18, respectively. The averaged sensitivity and specificity of these classification models were around 0.66 and 0.79 respectively.

Whatever the Sentinel-2 image, the sensitivity is lower than the specificity (Figure 5). So the proportion of actual positives that are correctly identified as such is lower than the proportion of actual negatives that are correctly identified as such. It means that the proportion of "Black" soils that are correctly identified as such is lower than the proportion of "Red" soils that are correctly identified as such, in spite of the calibration data balanced thanks to the SMOTE method.

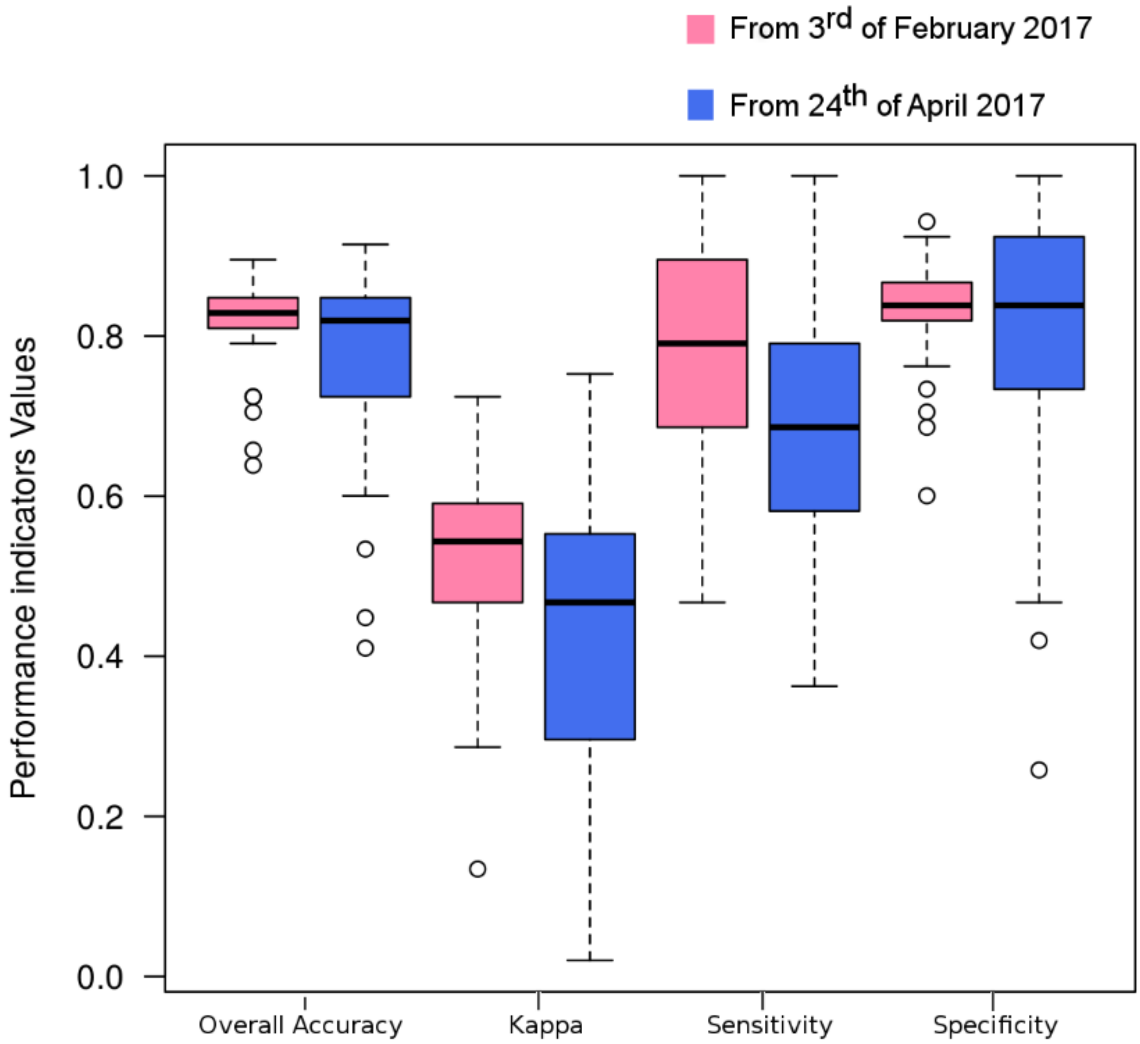

Figure 5. Boxplot of overall accuracy, kappa, sensitivity and specificity calculated from the Validation datasets and obtained for 50 classification models built using well-balanced data (using SMOTE method) and the Sentinel-2 image acquired on 3rd of February 2017 (pink) and on the 24th of April 2017 (blue).

The most frequent soil color classes among the 50 *Map_1_n* after the bootstrap process provided a classification of soil color for the Sentinel-2 data acquired on 3rd of February 2017 (Figure 6, left). As well, the most frequent soil color classes among the 50 *Map_2_n* after the bootstrap process provided a classification of soil color for the Sentinel-2 data acquired on the 24th of April 2017 (Figure 6, right). Whatever the Sentinel-2 data, the soil color over valley bottoms was mainly classified as "Black", while the soil color over hillslopes was mainly classified as "Red" (Figure 6).

The “Black” soils represent 45.7 % and 44.1 % of the bare soils pixels, from the Sentinel-2 data acquired on 3rd of February 2017 and 24th of April 2017, respectively.

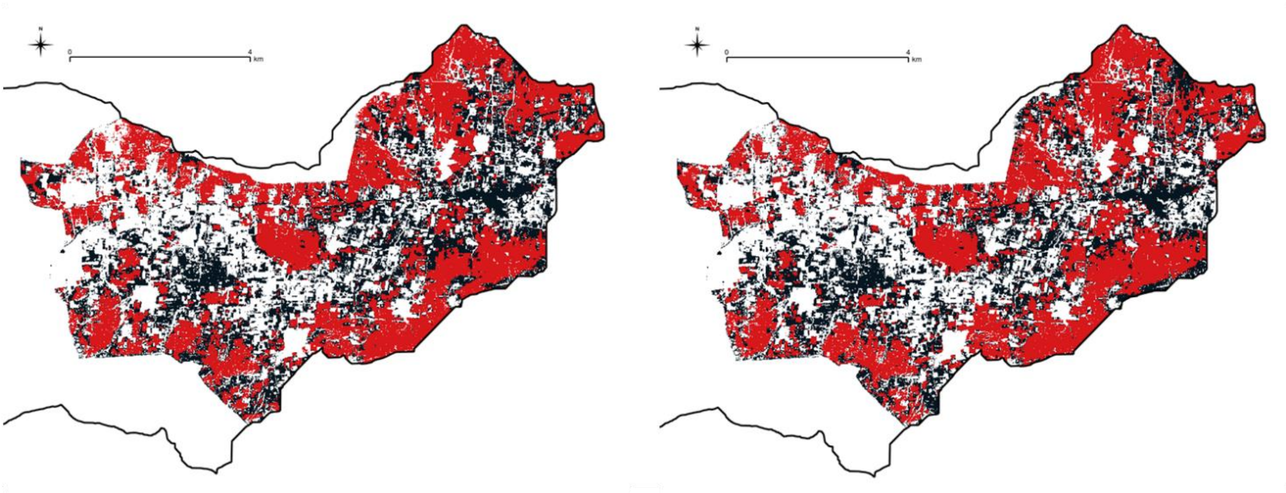

Figure 6. Map of “Black” and “Red” soils obtained from the Sentinel-2 image acquired on 3rd of February 2017 (left) and 24th of April 2017 (right).

4.2 Soil color changes analysis

50 maps of changes between the 3rd of February 2017 and the 24th of April 2017 were provided based on differences between the *Map_1_n* and *Map_2_n* (Figure 4c). 79 % of the fields (i.e., 14 474 fields) were covered by more than 30% of bare soil pixels and so were analyzed in term of soil color change. A map of the frequency of changes was provided (Figures 4d and 8) for these 14 474 fields, based on these 50 maps of changes.

Among these 14 474 fields, 100 fields had been classified as “Red” soil on the 3rd of February 2017 and then “Black” soil on the 24th of April 2017, more than 45 times over the 50 iterations (Figures 7 and 8A). So a soil color change from “Red-to-Black” soil between February and April over these 100 fields was detected over more than 90 % of the iterations. So these 100 fields can be considered as affected by tank silt applications (example Figures 8B1, 8B2, 8C1 and 8C2). These 100 fields are characterized by an average size of 0.16 hectares and cover a total of 16.2 hectares. 102 fields had been classified as “Red” soil on the 3rd of February 2017 and then

“Black” soil on the 24^{th} of April 2017, more than 40 times over the 50 iterations but less than 45 times over the 50 iterations (Figures 7 and 8A). So a soil color change from “Red-to-Black” soil between February and April over these 102 fields was detected over more than 80 % of the iterations. These 102 fields are characterized by an average size of 0.25 hectares and cover a total of 25.6 hectares.

Among the 14 474 fields, 33 fields had been classified as “Black” soil on the 3^{rd} of February 2017 and then “Red” soil on the 24^{th} of April 2017, more than 45 times over the 50 iterations (Figures 7 and 8A). These 33 fields are characterized by an average size of 0.12 hectares and cover a total of 3.99 hectares. 98 fields had been classified as “Black” soil on the 3^{rd} of February 2017 and then “Red” soil on the 24^{th} of April 2017, more than 40 times over the 50 iterations but less than 45 times over the 50 iterations (Figures 7 and 8A). These 98 fields are characterized by an average size of 0.17 hectares and cover a total of 17.27 hectares.

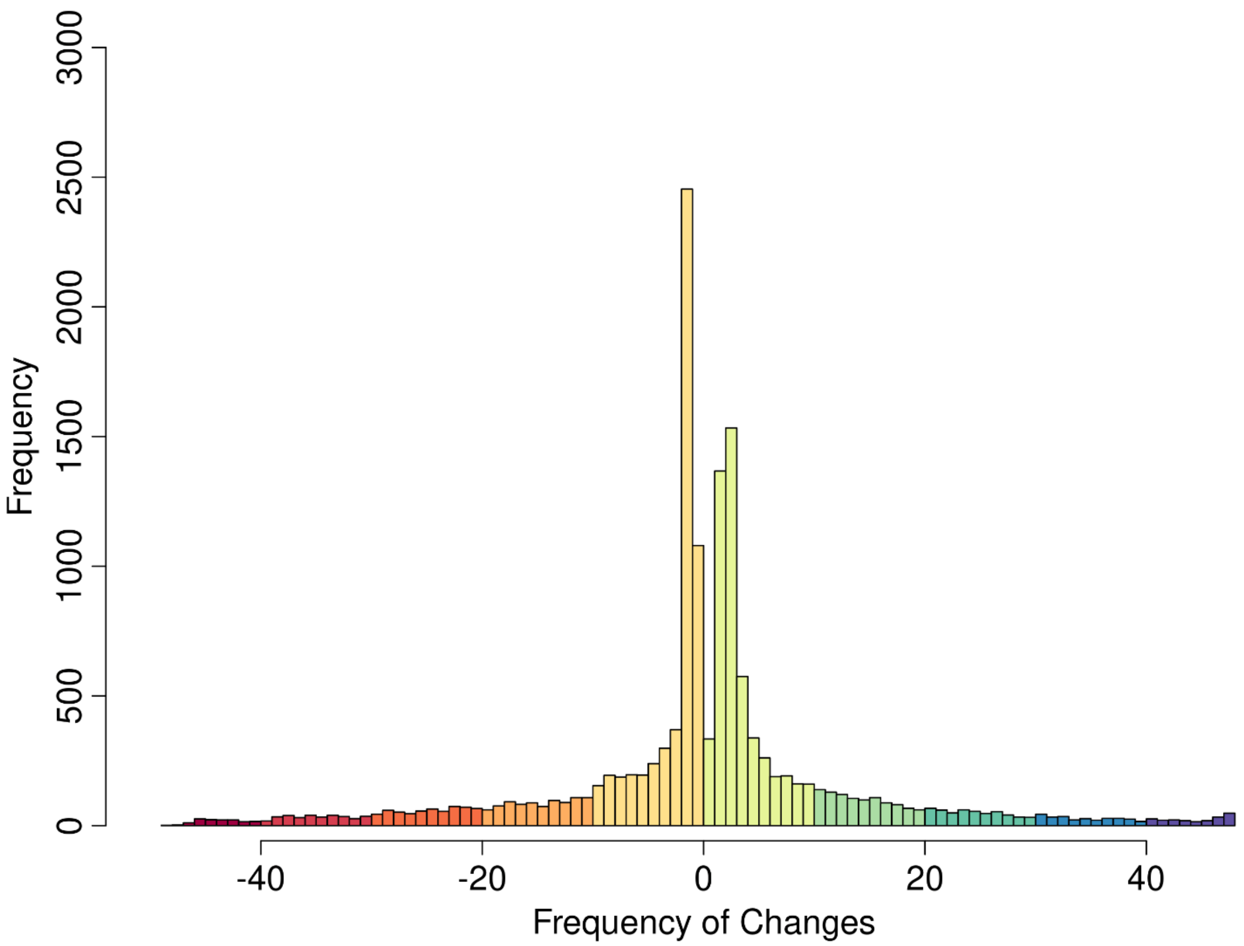

Figure 7. Density of the soil color changes frequency (over the 50 iterations) calculated over fields covered by more than 30 % of bare soil pixels. Positive values correspond to changes from "Red-to-Black" soil. Negative values correspond to changes from "Black-to-Red" soil.

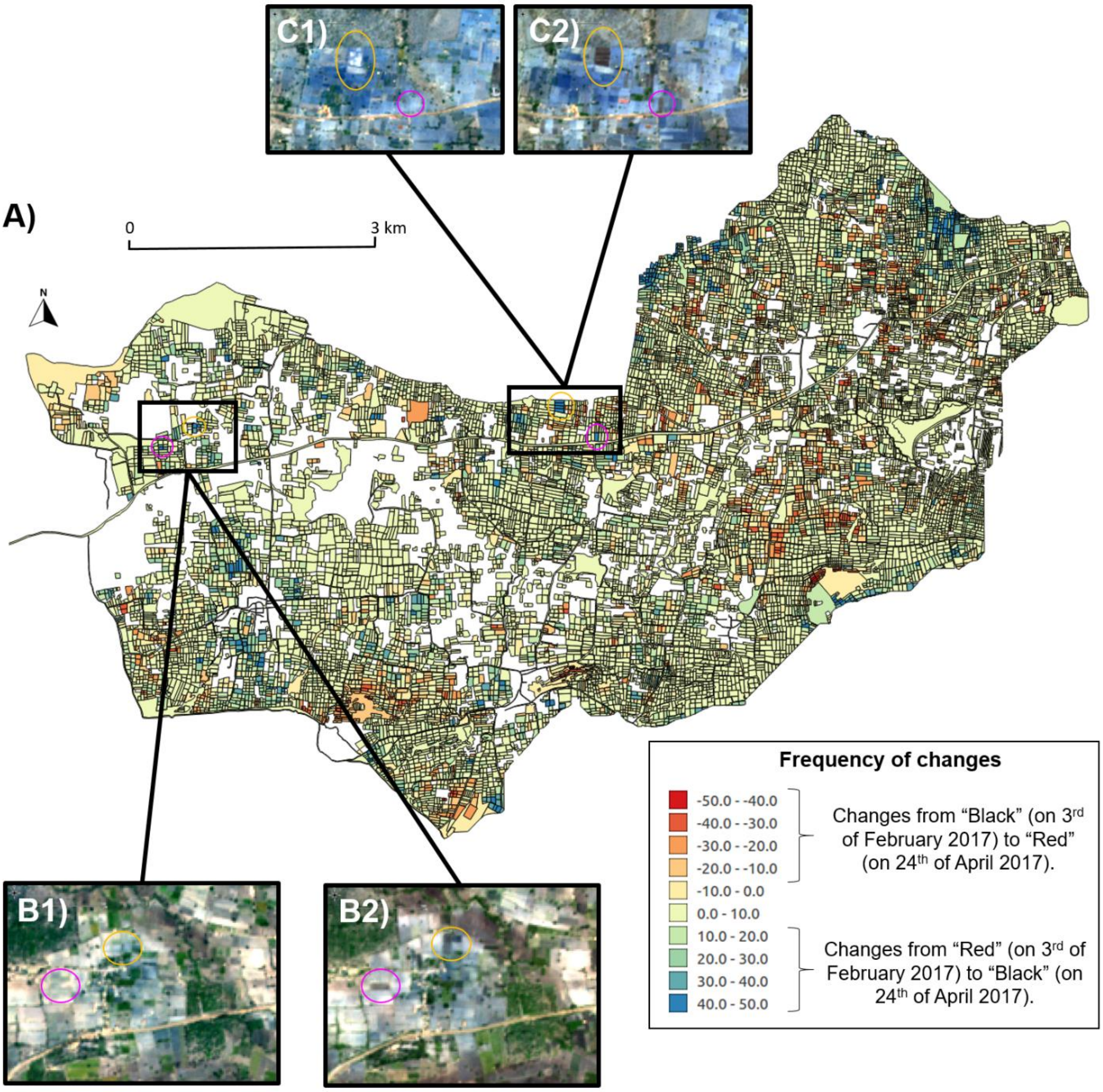

Figure 8. A) Number of changes per parcel from the 3rd of February 2017 to the 24th of April 2017, obtained after 50 classification models built using balanced data obtained by the SMOTE method. B1) and C1) corresponds to zooms on the Sentinel-2 image acquired on 3rd of February 2017. B2) and C2) corresponds to zooms on the Sentinel-2 image acquired on 24th of April 2017.

The surface soil moisture difference between both dates over fields covered by more than 30 % of bare soil pixels, follow a normal distribution centered around -0.5 (Figure 9A). So a majority of fields has a negative moisture difference showing an increase of moisture from the 3rd of February 2017 to the 24th of April 2017 (Figure 9A), which can be explained by irrigation after seeding as it is usually time in April.

Finally, no relation exists between the soil color changes frequency and the surface soil moisture difference between both dates (Figure 9B). The Pearson correlation coefficient between the changes from “Black-to-Red” soil and the surface soil moisture difference between both dates is around 0.008 with a p-value of 0.9 (Figure 9C). And the Pearson correlation coefficient between the changes from “Red-to-Black” soil and the surface soil moisture difference between both dates is around -0.03 with a p-value of 0.75 (Figure 9D).

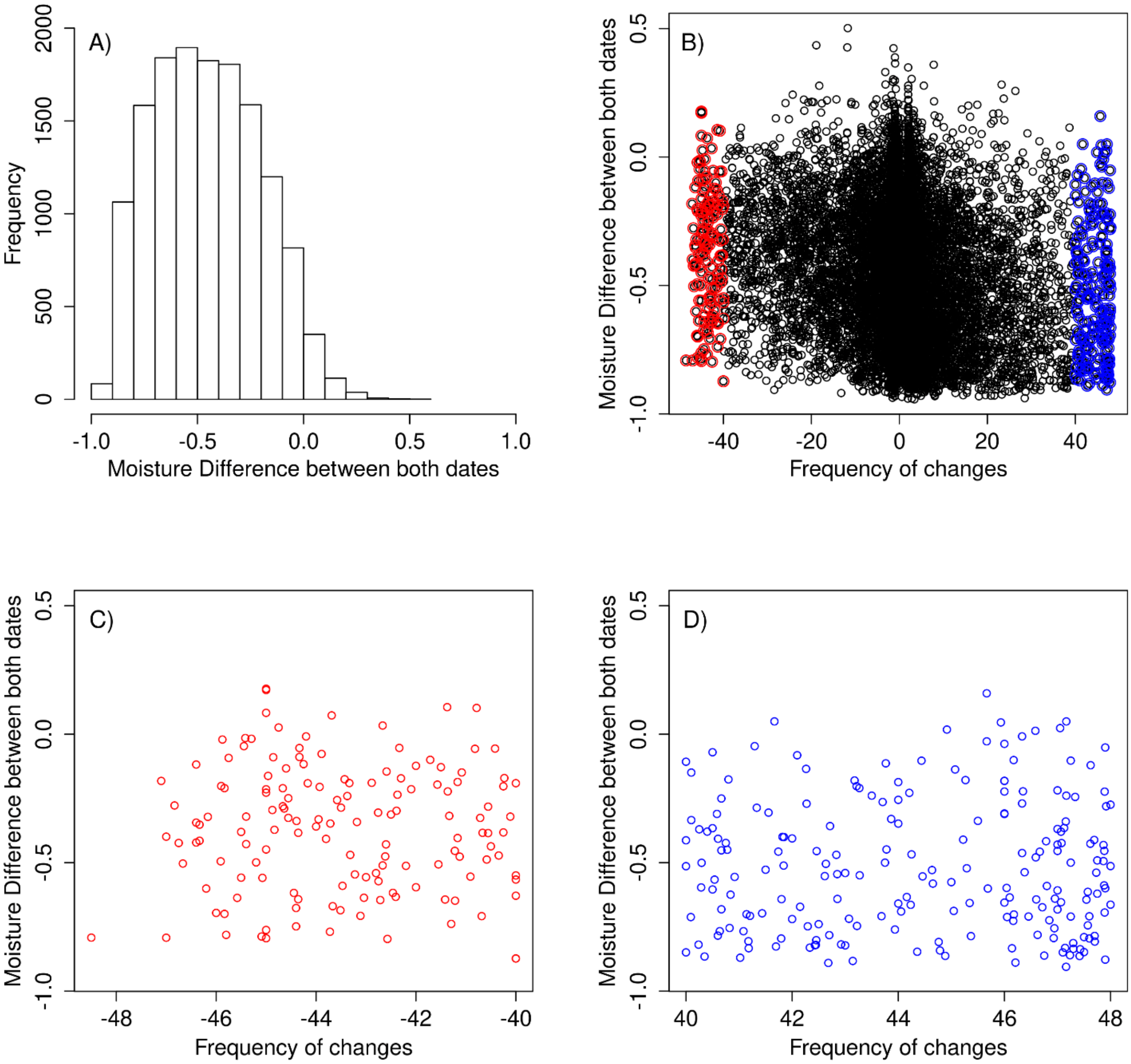

Figure 9. A) Frequency of surface soil moisture difference over the Berambadi catchment. Scatters plots of soil color changes frequency versus the surface soil moisture difference for B) all fields covered by more than 30 % of bare soil pixels (red dots correspond to fields identified as changed from “Black-to-Red” soil; blue dots correspond to fields identified as changed from “Red-to-Black” soil), C) fields identified as changed from “Black-to-Red” soil with medium to high confident (between 40 to 50 changes over the 50 iterations), and D) fields identified as changed from “Red-to-Black” soil with medium to high confident (between 40 to 50 changes over the 50 iterations).

5. Discussion

Mapping of tank silt applications

Based on the proposed approach, 100 fields corresponding to 16.2 hectares and 102 fields corresponding to 25.6 hectares were associated to a silt tank application with more than 45 changes over 50 iterations and between 40 changes to 45 iterations, respectively. So this approach is able to identify this old-age practice in our context in spite of the small size of fields (0.27 ha in average, Sharma et al., 2018) and allows highlighting that this practice of tank silt application concerns a minority of farmers each year. Moreover this methodology allows highlighting that the fields receiving tank silt are geographically far apart from each other and are distributed throughout the area (Figure 8). So this practice of tank silt application is not restricted to a small part of the catchment.

Mapping of "Red-to-Black" soil

Based on the proposed approach, 33 fields were identified as changed from "Black-to-Red" soil with more than 45 changes over the 50 iterations and 98 fields were identified as changed from "Black-to-Red" soil with between 40 to 45 changes over the 50 iterations. As no relation was highlighted between soil moisture changes and frequencies of color changes between both dates, these changes of "Black-to-Red" soil cannot be attributed to a soil moisture change. Based on our field knowledge, these changes of "Black-to-Red" soil may be attributed to an application of tank silt before the acquisition of our first Sentinel-2 image which may explain the black color identified on our first Sentinel-2 image over these fields and then a ploughing of the tank silt and original red soil between our both Setinel-2 images which may explain the red color identified on our second Sentinel-2 image over these fields. These changes of "Black-to-Red" soil may also be attributed to a bad correspondence between the observed Munsell soil color "5YR5/4" and "10YR3/4" and the soil color class "Red" and "Black", respectively (Table 2). Indeed the observed Munsell soil colors "5YR5/4" and "10YR3/4" correspond both to very dark reddish soils and are associated to different soil color classes (Table 2).

Classifications dealing with imbalanced calibration data

In soil science, Sharififar et al. (2019) highlighted that uncertain maps of soil classes result in imbalanced distribution of class observations, with a loss of minority classes and relatively poor classification accuracy. As final produced soil classes maps produced using imbalanced classes could be misleading for users or decision makers, a data pre-treatment, with either an over- or under-sampling of the Calibration data, coupled by a decision trees model showed significant improvement in keeping the minority soil classes (Sharififar et al., 2019). Sharififar et al., 2019 also highlighted that, in case of a use of data resampling, decision trees algorithms perform better than random forest and multinomial logistic regression algorithms. Our methodology based on a well-balanced Calibration data created from the initial imbalanced Calibration dataset thanks to the SMOTE methodology, coupled to the CART algorithm, which is a binary decision tree algorithm, follows this line, avoiding uncertain maps, misclassifications, and loss of the minority class (the “black” soil in our case). In a future methodological work dedicated to soil classes mapping, the performance of the SMOTE algorithm could be compared to the Geometric-SMOTE developed by Douzas and Bacao (2019a), as the Geometric-SMOTE algorithm has been recently demonstrated to outperform every other oversampling technique, for land cover classification (Douzas et al., 2019b).

Two-class classifications

Our methodology provided correct accuracies of the two-class classifications with a mean overall accuracy around 0.81 and 0.77 from the Sentinel-2 image acquired on February and April 2017, respectively (Figure 5). These performances are higher than those obtained for a soil texture classification using Landsat-5 TM images (Dematte et al., 2016) and Sentinel-2 images (Gomez et al., 2019), and for soil type discrimination using Landsat-5 TM images (Nanni et al., 2012). Both soil color classification maps showed good discrimination ability for hillslopes and uplands which correspond to red soils and for valleys which correspond to black soils (Figure 6) as previously described over this study area by Barbiero et al. (2010).

Limitation to bare soil pixels

As the number of mapped fields receiving tank silt is low, the success of our proposed approach dealing with two Sentinel-2 images depends on the area covered by bare soil pixels. The larger the bare soil surface area, the more accurate the spatial representation of tank silt will be. An increase in vegetation cover changes completely the shape of soil spectral reflectance (Bartholomeus et al., 2011; Ouerghemmi et al., 2011). So a characterisation of the topsoil color has to be done only over bare soils, which limits the mapped surface. Moreover as the study is based on two images, only the bare soil pixels common to both images can be studied. In our case, after masking the vegetation areas on both images in addition to the urban areas and lakes, only 65% of the catchment remained to be map in term of topsoil color.

Futures studies

As we showed that a couple of Sentinel-2 images encompassing the tank silt applications period allows localizing the tank silt application, an extended study of this farmer practice could be addressed using couples of Sentinel-2 images acquired each year from 2015 (start of the Sentinel-2A satellite) till today. This extended study would provide a multi-year time series of tank silt application maps and so would allow to monitor this practice. Such multi-year time series of tank silt application maps would bring new knowledge about both spatial repartition and frequency of this practice over the Berambadi catchment.

These maps of tank silt application might also provide inputs for agronomic modelling as tank silt application brought out an increase in the productivity of different cropping systems (Sharma et al., 2015), and organic carbon and available nutrient status (Patil et al., 2017). It might also provide inputs for hydrological modelling as tank silt application has been demonstrated to increase available water capacity (Deshmukh et al., 2019). Moreover, a field survey could be done to estimate the depth of tank silt applied over fields. This would allow calculating the volume of silt extracted from the tanks each year and so studying the erosion processes on the study area.

6. Conclusion

The objective of this study was to evaluate the utility of Sentinel-2 images for mapping tank silt application characterized by a black color as they come from Vertisols soils, over red soils using a couple of Sentinel-2 images encompassing the tank silt applications period. In this study a two-class classification at each Sentinel-2 date was produced with correct accuracies and then the changes of soil color from the first image to the second one were produced. This approach took care of having a well-balanced Calibration data to avoid some misclassification of “soil changes due to over representation of the “Red” soil in the Calibration data. This approach was able to identify this old-age practice in our context in spite of the small size of fields and allows highlighting that i) this practice of tank silt application concerns a minority of farmers each year and ii) the fields receiving tank silt are geographically far apart from each other and are distributed throughout the area. This work was also able to demonstrate the potential of Sentinel-2 images for monitoring farmer practices. Finally, the study of this farmer practice is a large issue that needs to be address and the current multi temporal remote sensing data might help for improving its spatio-temporal characterization.

Acknowledgments

The authors are indebted to NBSS for soil samples collection. The Kabini Critical Zone Observatory (AMBHAS, BVET, Sekhar et al., 2016; Tomer et al., 2015, www.ambhas.com; https://mtropics.obs-mip.fr/) which is part of the OZCAR network (Gaillardet et al., 2018, http://www.ozcar-ri.org/ozcar/), are also acknowledged.

References

AMBHAS Team (2015) A manual for agro-hydrological monitoring in Pilot Experimental Watersheds. Indian Institute of Sciences and AMBHAS. August 2015. 65 pages. https://watershed.karnataka.gov.in/storage/pdf-files/Sujala%20Docs/Hydrology%20Manual.pdf

Ali A., Shamsuddin S.M. and Ralescu A.L. (2015). Classification with class imbalance problem: A Review. *Int. J. Advance Soft Compu. Appl*, Vol. 7, No. 3, November 2015.

Baran A., Tarnawski M. and Urbaniak M. (2019). An assessment of bottom sediment as a source of plant nutrients and an agent for improving soil properties. *Environmental Engineering and Management Journal.* 18, No. 8, pp. 1647-1656. *DOI: 10.30638/eemj.2019.155*

Barbiero L., Kumar M.S.M., Violette A., Oliva P., Braun J.J., Kumar C., Furian S., Babic M., Riotte J. and Valles V. (2010). Ferrolysis induced soil transformation by natural drainage in vertisols of sub-humid South India. *Geoderma* 156, pp. 173–188. *https://doi.org/10.1016/j.geoderma.2010.02.014*

Barnes R.J., Dhanoa M.S. and Lister S.J. (1989). Standard Normal Variate Transformation and De-trending of Near-Infrared Diffuse Reflectance Spectra. *Applied Spectroscopy*, Vol. 43, Issue 5, pp. 772-777.

Bartholomeus H, Kooistraa L, Stevens A, van Leeuwen M, van Wesemael B, Ben-Dor E and Tychon B (2011) Soil organic carbon mapping of partially vegetated agricultural fields with imaging spectroscopy. *Int J Appl Earth Obs Geoinf* 13: pp. 81–88. *https://doi.org/10.1016/j.jag.2010.06.009*

Breiman L., J.H. Friedman, R.A. and Olsen, C.J., (1984). Stone Classification and Regression Trees Belmont, Wadsworth, CA.

Brodský L., Vašát R., Klement A., Zádorová T., Jakšík O. (2013). Uncertainty propagation in VNIR reflectance spectroscopy soil organic carbon mapping, *Geoderma*, 199, pp. 54-63, *https://doi.org/10.1016/j.geoderma.2012.11.006.*

Chawla N.V., Bowyer K.W., Hall L.O. and Kegelmeyer W.P. (2002). Smote: synthetic minority over-sampling technique, *Journal of Artificial Intelligence Research*, 16, pp. 321-357. *DOI: 10.1613/jair.953*

Chawla N. V., Japkowicz N. and Kolcz A. (2004). Editorial: Special Issue on Learning from Imbalanced Data Sets. ACM SIGKDD Explorations Newsletters. Volume 6, Issue 1. *DOI: 10.1145/1007730.1007733*

Cohen J., 1960 “A coefficient of agreement for nominal scales,” *Educational and Psychological Measurement,* vol. 20, pp. 37-46.

CWC (2010). Water and Related Statistics, Information System Organisation, Water Planning and Projects Wing, Central Water Commission, Government of India, New Delhi, December.

Dazzi C. and Lo Papa G. (2015). Anthropogenic soils: general aspects and features. Ecocycles 1(1), pp. 3-8. *DOI: 10.19040/ecocycles.v1i1.23*

De Jong S. (1992) The analysis of spectroscopical data to map soil types and soil crusts of Mediterranean eroded soils, *Soil Techniques*, 5 (1992), pp. 199-211. *https://doi.org/10.1016/0933-3630(92)90022-S*

Dematte J. A. M., Alves M. R., Terra F. da S., Bosquilia R. W. D., Fongaro C. T. and Barros P. P. da S. (2016). Is It Possible to Classify Topsoil Texture Using a Sensor Located 800 km Away from the Surface? *Revista Brasileira de Ciência do Solo*, 40. *DOI: 10.1590/18069657rbcs20150335*

Deshmukh DP, PW Deshmukh, SM Bhoyar and SP Gawai (2019). Characterization of physical attributes of tank silt of natural reservoirs in western Melghat region of Dharni. *International Journal of Chemical Studies.* 7(3), pp. 2782-2785.

DHAN Foundation. Tank silt Application for Agricultural Production Enhancement - Scope, Issues and Challengers. Centre of policy and planning, *Policy Brief report 12*.

Douzas, G. and Bacao, F. (2019a) Geometric SMOTE a geometrically enhanced drop-in replacement for SMOTE. *Information Science,* 501, pp. 118–135. *https://doi.org/10.1016/j.ins.2019.06.007*

Douzas G., Bacao F., Fonseca J. and Khudinyan M. (2019b). Imbalanced Learning in Land Cover Classification: Improving Minority Classes’ Prediction Accuracy Using the Geometric SMOTE Algorithm, *Remote Sens.* 11, 3040; *doi:10.3390/rs11243040.*

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

European Space Agency (ESA) (2015). Sentinel-2 User Handbook. 24 July 2015. 64 pages.

Frampton W.J., Dash J., Watmough G. and Milton E.J. (2013) Evaluating the capabilities of Sentinel-2 for quantitative estimation of biophysical variables in vegetation. *ISPRS Journal of Photogrammetry and Remote Sensing,* 82, pp. 83-92 *https://doi.org/10.1016/j.isprsjprs.2013.04.007*

Gaillardet, J., I. Braud, F. Hankard, S. Anquetin, O. Bour, N. Dorfliger, J.R. de Dreuzy, S. Galle, C. Galy, S. Gogo, L. Gourcy, F. Habets, F. Laggoun, L. Longuevergne, T. Le Borgne, F. Naaim-Bouvet, G. Nord, V. Simonneaux, D. Six, T. Tallec, C. Valentin, G. Abril, P. Allemand, A. Arènes, B. Arfib, L. Arnaud, N. Arnaud, P. Arnaud, S. Audry, V. B. Comte, C. Batiot, A. Battais,

H. Bellot, E. Bernard, C. Bertrand, H. Bessière, S. Binet, J. Bodin, X. Bodin, L. Boithias, J. Bouchez, B. Boudevillain, I. B. Moussa, F. Branger, J. J. Braun, P. Brunet, B. Caceres, D. Calmels, B. Cappelaere, H. Celle-Jeanton, F. Chabaux, K. Chalikakis, C. Champollion, Y. Copard, C. Cotel, P. Davy, P. Deline, G. Delrieu, J. Demarty, C. Dessert, M. Dumont, C. Emblanch, J. Ezzahar, M. Estèves, V. Favier, M. Faucheux, N. Filizola, P. Flammarion, P. Floury, O. Fovet, M. Fournier, A. J. Francez, L. Gandois, C. Gascuel, E. Gayer, C. Genthon, M. F. Gérard, D. Gilbert, I. Gouttevin, M. Grippa, G. Gruau, A. Jardani, L. Jeanneau, J. L. Join, H. Jourde, F. Karbou, D. Labat, Y. Lagadeuc, E. Lajeunesse, R. Lastennet, W. Lavado, E. Lawin, T. Lebel, C. Le Bouteiller, C. Legout, Y. Lejeune, E. Le Meur, N. Le Moigne, J. Lions, A. Lucas, J. P. Malet, C. Marais-Sicre, J. C. Maréchal, C. Marlin, P. Martin, J. Martins, J. M. Martinez, N. Massei, A. Mauclerc, N. Mazzilli, J. Molénat, P. Moreira-Turcq, E. Mougin, S. Morin, J. N. Ngoupayou, G. Panthou, C. Peugeot, G. Picard, M. C. Pierret, G. Porel, A. Probst, J. L. Probst, A. Rabatel, D. Raclot, L. Ravanel, F. Rejiba, P. René, O. Ribolzi, J. Riotte, A. Rivière, H. Robain, L. Ruiz, J. M. Sanchez-Perez, W. Santini, S. Sauvage, P. Schoeneich, J. L. Seidel, M. Sekhar, O. Sengtaheuanghoung, N. Silvera, M. Steinmann, A. Soruco, G. Tallec, E. Thibert, D. V. Lao, C. Vincent, D. Viville, P. Wagnon, and R. Zitouna. (2018). OZCAR: The French Network of Critical Zone Observatories. *Vadose Zone Journal*, 17:180067. *https://doi.org/10.2136/vzj2018.04.0067*

Gholizadeh A., Žižala D., Saberioon M. and Borůvka L. (2018) Soil organic carbon and texture retrieving and mapping using proximal, airborne and Sentinel-2 spectral imaging, *Remote Sensing of Environment*, 218, pp. 89-103. *https://doi.org/10.1016/j.rse.2018.09.015*

Gomez C., Drost A.P.A. and Roger J.M. (2015). Analysis of the uncertainties affecting predictions of clay contents from VNIR/SWIR hyperspectral data, *Remote Sensing of Environment,* 156, pp. 58–70. *https://doi.org/10.1016/j.rse.2014.09.032*

Gomez C., Dharumarajan S., Féret J.-B., Lagacherie P., Ruiz L. and Sekhar M. (2019) Use of Sentinel-2 Time-Series Images for Classification and Uncertainty Analysis of Inherent Biophysical Property: Case of Soil Texture Mapping. *Remote Sensing,* 11, 565. *doi:10.3390/rs11050565*

Gunnell, Y. and Krishnamurthy, A. (2003). Past and present status of runoff harvesting systems in dryland peninsular India: A critical review. AMBIO: A Journal of the Human Environment, 32(4), 320-324.

Hardaker, J. (2004). In: Hardaker, C.J.B., et al. (Eds.), Coping with Risk in Agriculture. CABI, Wallingford.

Japkowicz, N. and Stephen, S. (2002). The Class Imbalance Problem: A Systematic Study. Intelligent Data Analysis, 6(5): pp. 429-450. *DOI: 10.3233/IDA-2002-6504*

Karanam P.V., Wani S.P., Sahrawat K.L. and Jangawad L.S. (2008). Economic evaluation of sediments as a source of plant nutrients. *Current Science*, 95, 8, 25 October 2008.

Kuhn, J. Wing, S. Weston, A. Williams, C. Keefer, A. Engelhardt, T. Cooper, Z. Mayer, B. Kenkel, R Core Team, *et al.* (2016). Caret: Classification and Regression Training; R Core Team: Vienna, Austria. [Online]. Available: http://topepo.github.io/caret/index.html

Landis JR and Koch GG. (1977). “The Measurement of Observer Agreement for Categorical Data” 1 (33). *Biometrics*: pp. 159–74.

Loiseau T., Chen S., Mulder V.L., Román Dobarco, M., Richer-de-Forges A.C., Lehmann S., Bourennane H., Saby N.P.A., Martin M.P., Vaudour E., Gomez C., Lagacherie P. and Arrouays D., (2019). Satellite data integration for soil clay content modelling at a national scale. *International Journal of Applied Earth Observation and Geoinformation*, 82, 101905. *https://doi.org/10.1016/j.jag.2019.101905*

Louis J., Debaecker V., Pflug B., Main-Knorn M., Bieniarz J., Mueller-Wilm U., Cadau E. and Gascon F. (2016) Sentinel-2 Sen2Cor: L2A Processor for Users, *Proceedings Living Planet* Symposium, pp. 1–8.

McBratney A.B., Mendonça-Santos M.L. and Minasny B. (Eds.), 2003. On Digital Soil Mapping. Geoderma, 117, pp. 3-52, *https://doi.org/10.1016/S0016-7061(03)00223-4*

McBratney A., Field D.J. and Koch A., (2014) The dimensions of soil security, *Geoderma*, 213, pp. 203-213, *https://doi.org/10.1016/j.geoderma.2013.08.013*

Mougenot B. and Pouget M. (1993). Remote sensing of salt affected soils. *Remote Sens Rev* 7: pp. 241-259.

Munsell Color Company (1975) Munsell Soil Color Charts. Munsell Color Company, Baltimore.

Nanni M.F., Demattê J.A.M., Chicati M.L., Fiorio P.R., Cézar E. and de Oliveira R.B. (2012). Soil surface spectral data from Landsat imagery for soil class discrimination, Acta Sci., Agron., ISSN 1807-8621. 34, n. 1, pp. 103-112. *Doi: 10.4025/actasciagron.v34i1.12204*

Narayanamoorthy A. (2007). Tank irrigation in India: a time series analysis. *Water Policy*, 9; pp. 193–216

Obi Reddy G.P., Patil N.G. and Chaturvedi A., (2017). Sustainable Management of Land Resources: An Indian Perspective. ISBN hard: 978-1-77188-517-1. Cat# N11824, ISBN ebook: 978-1-315-36556-5.

Ouerghemmi W., Gomez C., Naceur S. and Lagacherie P. (2011) Applying blind source separation on hyperspectral data for clay content estimation over partially vegetated surfaces. *Geoderma* 163(3–4): pp. 227–237. *https://doi.org/10.1016/j.geoderma.2011.04.019*

Osman M., Wani, S.P., Vineela, C and Murali, R. (2009). Quantification of Nutrients Recycled by Tank Silt and its Impact on Soil and Crop - A Pilot Study in Warangal District of Andhra Pradesh. Global Theme on Agroecosystems Report no. 52. Patancheru 502 324, Andhra Pradesh, India; International Crops Research Institute for Semi-Arid Tropics. 20 pages.

Patil P.D., Vaidya P.H., Dhawan A.S. and Kumbhar C.S. (2017). Impact of tank silt and FYM application on soil quality and yield of soybean (Glycine max L.) under inceptisol. *International Journal of Chemical Studies,* 5(5): pp. 962-966.

Pickup G. and Nelson D.J. (1984) Use of landsat radiance parameters to distinguish soil erosion, stability, and deposition in arid Central Australia, *Remote Sensing of Environment*, 16 (3), pp. 195-209. *https://doi.org/10.1016/0034-4257(84)90064-6*

Reddy V.R., Reddy M.S. and Palanisami K. (2018) Tank rehabilitation in India: Review of experiences and strategies. *Agricultural Water Management*, 209, 32–43 *https://doi.org/10.1016/j.agwat.2018.07.013*

Roberts JT. Global Inequality and Climate Change. Soc Nat Resour 2001;14: 501–9. *DOI: 10.1080/08941920118490*

Rouse J.W., Haas R.H., Schell J.A. and Deering W.D. (1973). Monitoring vegetation systems in the Great Plains with ERTS. *In: Third ERTS Symposium*, NASA SP-351, pp. 309–317.

Sekhar M., Riotte J., Ruiz L., Jouquet P. and Braun J.J. (2016). Influences of Climate and Agriculture on Water and Biogeochemical Cycles: Kabini Critical Zone Observatory. *Proceeding of the Indian National Science Academy*, 82 (3), pp. 833-846. *DOI: 10.16943/ptinsa/2016/48488*

Shah T., Roy A.D., Qureshi A.S. and Wang J. (2003) ‘Sustaining Asia’s groundwater boom: an overview of issues and evidence’, *Natural Resources Forum* 27: pp. 130–140. *https://doi.org/10.1111/1477-8947.00048*

Sharififar A., Sarmadian F., Malone B.P. and Minasny B. (2019). Addressing the issue of digital mapping of soil classes with imbalanced class observations. *Geoderma* 350, pp. 84–92. *https://doi.org/10.1016/j.geoderma.2019.05.016*

Sharma S.K., Sharma R.K., Kothari A.K., Osman M. and Chary G.R. (2015). Effect of Tank Silt Application on Productivity and Economics of Maize-Based Production System in Southern Rajasthan. *Indian Journal of Dryland Agricultural Research and Development,* 30(2), pp. 24-29. *DOI : 10.5958/2231-6701.2015.00021.4*

Sharma A.K., Hubert-Moy L., Buvaneshwari S, Sekhar M., Ruiz L., Bandyopadhyay S. and Corgne S. (2018). Irrigation History Estimation Using Multitemporal Landsat Satellite Images: Application to an Intensive Groundwater Irrigated Agricultural Watershed in India. *Remote Sensing*. 10, 893. *https://doi.org/10.3390/rs10060893*

Steinberg B. (2009) CART: Classification and Regression Trees, In The Top Ten Algorithms in Data Mining Book, Edited ByXindong Wu, Vipin Kumar, Chapter 10, pp. 193-216.

Sun Y., Wong A.K.C. and Kamel M.S. (2009). Classification of imbalanced data: a review. *International Journal of Pattern Recognition and Artificial Intelligence*, 23(04), pp. 687–719. *https://doi.org/10.1142/S0218001409007326*

Taghizadeh-Mehrjardi, R., Schmidt, K., Eftekhari, K., Behrens, T., Jamshidi, M., Davatgar, N., Toomanian, N. and Scholten, T., (2020). Synthetic resampling strategies and machine learning for digital soil mapping in Iran. *European Journal of Soil Science*, 71(3), pp.352-368. *https://doi.org/10.1111/ejss.12893.*

Taghadosi M.M., Hasanlou M. and Eftekhari K., (2019) Retrieval of soil salinity from Sentinel-2 multispectral imagery, *European Journal of Remote Sensing,* 52:1, pp. 138-154. *DOI: 10.1080/22797254.2019.1571870*

Tomer S.K., Al Bitar A., Sekhar M., Zribi M., Bandyopadhyay S., Sreelash K., Sharma A.K., Corgne S. and Kerr Y. (2015). Retrieval and multi-scale validation of soil moisture from multi-temporal SAR data in a semi-arid tropical region. *Remote Sensing*, 7, pp. 8128–8153. *DOI: 10.3390/rs70608128*

Vaudour E., Gomez C., Fouad Y. and Lagacherie P. (2019). Sentinel-2 image capacities to predict common topsoil properties of temperate and Mediterranean agroecosystems. *Remote Sensing of Environment*, 223, pp. 21-33. *https://doi.org/10.1016/j.rse.2019.01.006*

Viscarra Rossel, R. A., Minasny, B., Roudier, P., and McBratney, A. B. (2006). Colour space models for soil science. *Geoderma* 133, pp. 320–337. *https://doi.org/10.1016/j.geoderma.2005.07.017*

Viscarra Rossel R.A. and Hicks W.S. (2015). Soil organic carbon and its fractions estimated by visible–near infrared transfer functions. *European Journal of Soil Science*, 66 (3), pp. 438-450. *https://doi.org/10.1111/ejss.12237*

Walter K., Gunkel G. and Gamboa N. (2012). An assessment of sediment reuse for sediment management of Gallito Ciego Reservoir, Peru. *Lakes & Reservoirs: Research and Management* 2012 17: pp 301–314. *https://doi.org/10.1111/lre.12008*

Wang J., Ding J., Yu D., Ma X., Zhang Z., Ge X., Teng D., Li X., Liang J., Lizaga I., Chen X., Yuan L. and Guo Y., (2019) Capability of Sentinel-2 MSI data for monitoring and mapping of soil salinity in dry and wet seasons in the Ebinur Lake region, Xinjiang, China. *Geoderma*, 353, pp 172-187. *https://doi.org/10.1016/j.geoderma.2019.06.040*

Woodcock C.E., Allen R., Anderson M., Belward A., Bindschadler R., Cohen W.B., Gao F., Goward S.N., Helder D., Helmer E., Nemani R., Oreopoulos L., Schott J., Thenkabail P.S., Vermote E.F., Vogelmann J., Wulder M.A. and Wynne R. (2008). Free access to Landsat imagery. *Science*, 320, pp 1011. *DOI: 10.1126/science.320.5879.1011a*