Article

Soil Clay Content Mapping Using a Time Series of Landsat TM Data in Semi-Arid Lands

Marouen Shabou [1,2,*], Bernard Mougenot [1,†], Zohra Lili Chabaane [2,†] Christian Walter [3,†], Gilles Boulet [1,†], Nadhira Ben Aissa [2,†] and Mehrez Zribi [1,†]

[1] CESBIO/UMR 5126, 18 av.Edouard Belin,bpi 2801, F-31401 Toulouse Cedex 9, France;
E-Mails: bernard.mougenot@ird.fr (B.M.); gilles.boulet@ird.fr (G.B.); mehrez.zribi@ird.fr (M.Z.)

[2] Rural engineering, water and forest department, INAT/University of Carthage, 43,
Avenue Charles Nicolle 1082 Tunis-Mahrajène, Tunisia;
E-Mails: zohra.lili.chabaane@gmail.com (Z.L.C.); benaissa.nadhira@planet.tn (N.B.A.)

[3] Agrocampus Ouest, INRA, UMR 1069 SAS, 65 rue de st Brieuc- CS 84215,
F-35042 Rennes Cedex, France; E-Mail: christian.walter@agrocampus-ouest.fr

[†] These authors contributed equally to this work.

[*] Author to whom correspondence should be addressed; E-Mail: marouen.shabou@gmail.com;
Tel.: +216-9768-1630; Fax: +216-7184-0270.

Academic Editors: Clement Atzberger and Prasad S. Thenkabail

Abstract: Clay content (fraction < 2 µm) is one of the most important soil properties. It controls soil hydraulic properties like wilting point, field capacity and saturated hydraulic conductivity, which in turn control the various fluxes of water in the unsaturated zone. In our study site, the Kairouan plain in central Tunisia, existing soil maps are neither exhaustive nor sufficiently precise for water balance modeling or thematic mapping. The aim of this work was to produce a clay-content map at fine spatial resolution over the Kairouan plain using a time series of Landsat Thematic Mapper images and to validate the produced map using independent soil samples, existing soil map and clay content produced by TerraSAR-X radar data. Our study was based on 100 soil samples and on a dataset of four Landsat TM data acquired during the summer season. Relationships between textural indices (MID-Infrared) and topsoil clay content were studied for each selected image and were used to produce clay content maps at a spatial resolution of 30 m. Cokriging was used to fill in the gaps created by green vegetation and crop residues masks and to predict clay content of each pixel of the image at 100 m grid spatial resolution. Results showed that mapping clay content using a

time series of Landsat TM data is possible and that the produced clay content map presents a reasonable accuracy (R^2 = 0.65, RMSE = 100 g/kg). The produced clay content map is consistent with existing soil map of the studied region. Comparison with clay content map generated from TerraSAR-X radar data on a small area with no calibration point revealed similarities in topsoil clay content over the largest part of this extract, but significant differences for several areas. *In-situ* observations at those locations showed that the Landsat TM mapping was more consistent with observations than the TerraSAR-X mapping.

1. Introduction

Integrated water resource management requires an improved understanding of soil at fine scales. Soil texture is a key parameter allowing the establishment of pedotransfer functions to derive hydrodynamics soil parameters such as wilting point, field capacity or saturated hydraulic conductivity. Clay content (soil fraction <2 µm) is considered the most important parameter to be estimated. It is related to the cation exchange capacity of soils [1] and allow a tight control of soil hydraulic properties like water storage and availability to crop plants, field capacity and wilting point. It is the first statistical factor to be taken into account when building pedotransfer functions [2,3].

Soil texture maps are difficult to derive. Existing soil maps are neither exhaustive nor precise enough, and often describe soil types rather than texture. Point measurements through soil analysis can be performed to complete those maps but are expensive and dense sampling is required to characterize adequately the spatial variability, making broad-scale quantitative evaluation difficult [4]. Moreover, laboratory analysis performed by chemical treatments are expensive and time-consuming [5,6]. New and quick methods to measure soil properties and variability are required for the development of soil interpretations [6,7].

In recent decades, the soil science community has proposed several methodologies to better estimate topsoil texture using remote sensing [8–12]. Remote sensing data are an important component of predictive soil mapping. It provides a spatially contiguous, quantitative measure of surface reflectance, which is related to some soil properties [13]. This tool facilitates mapping inaccessible areas by reducing costly field surveys [14]. It also helps to produce more accurate and updated soil properties' maps through large scale studies by relating field-measured variables such as clay content to surface reflectance through linear regression models to supply continuous estimates for environmental variables [15]. Mapping of continuous variables from high resolution imagery such as Landsat Thematic Mapper (TM) has largely depended on modeling empirical relationships [15]. Landsat TM sensor has been providing near-continuous multispectral coverage, with wavelengths covering the visible (0.45–0.69 µm), near-infrared (0.76–0.9 µm) and shortwave infrared (1.55–2.35 µm) domains, supplying adequate data set for studies of improved management of natural resources [16].

Some studies have shown that physical factors (particle size and surface roughness) and components (surface mineralogy, organic matter content and moisture) control soil spectral reflectance [17]. Mineral

composite maps, including ferrous minerals, iron oxide and clay minerals, could be produced using remotely sensed data such as Landsat-5 Thematic Mapper (TM), Landsat-7 Enhanced Thematic Mapper Plus (ETM+), Landsat-8 Operational Land Imager (OLI), Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) and Moderate Resolution Imaging Spectroradiometer (MODIS) through composite indices [18–21]. Other researchers [8,22] have used hyper spectral data for clay content prediction, calcium carbonate (CaCO3), iron and cation-exchange capacity (CEC). Models for soil attribute prediction from spectral data, also named spectral transfer functions, have been developed recently [23–26]. Spectral transfer functions relate soil properties estimated from samples collected in the field and spectral reflectances measured above those samples, usually in the laboratory. Surface mineralogy can be derived by wavelength-specific charge transfer and crystal field absorptions associated with the presence of iron and iron-oxides (Fe^{2+} and Fe^{3+}), and vibrational absorptions associated with hydroxyl bonds in clay, absorbed water and the carbonate ion [17,27]. The middle infrared index (MID-infrared), which is a normalized difference index between TM5 and TM7 shortwave infrared bands, allows discrimination between sandy and clayey soils [20,21]. These bands are near the two main water absorption bands (1.4 µm–1.9 µm). The presence of clay in the soil composition enhances the absorption in TM7 band (2.2 µm). The use of such bands requires dry bare soil, which makes this method more suited to mapping semi-arid lands. All these studies indicate that the electromagnetic energy interacts with some soil properties in specific wavelengths. An analysis of the behavior of these wavelengths can be utilized in the modeling process to determine and map some soil attributes such as clay content. In our case, we used the MID-infrared index to produce clay content maps over semi-arid lands using time series of Landsat TM data. A similar index as the STI (Soil Tillage Index) was used to detect and quantify crop residue out of the soil baseline [28–30].

The aim of this research was to map topsoil clay content at fine spatial resolution over the Kairouan plain. This study used time series of Landsat Thematic Mapper (TM5) images and a dataset of 100 soil samples collected over our study area. To produce clay content map, we proceeded in three steps. The first step consisted in studying relationships between the MID-infrared index and observed clay content from calibration data. The second step was to combine multi-temporal Landsat TM data through cokriging to predict topsoil clay content within the Kairouan plain. The final step was to validate the produced clay content map by independent soil samples and by comparison with existing soil map and with topsoil clay content map produced from TerraSAR-X radar data.

2. Study Area and Data

2.1. Study Area: The Kairouan Plain

The Kairouan plain [31] is located in central Tunisia (9°30′E–10°15′E, 35′N, 35°45′N) (Figure 1). The climate in this region is semi-arid, with an average annual rainfall of approximately 300 mm/year, characterized by a rainy season lasting from October to May, with the two rainiest months being October and March. As is generally the case in semi-arid areas, the rainfall patterns in this area are highly variable in time and space. The landscape is mainly flat, and the vegetation is dominated by agricultural production (cereals, olive groves, fruit trees and winter and summer market gardening). Soil units reflect high spatial and vertical variability of alluvial deposits, except along wadi. The last flood of 1969 changed some soil

horizons with sandy deposits of up to 100 cm and existing soil maps may become outdated. The dominant major soil group is classified as Fluvisols and saline soils [32]. Soil texture presents a large spatial variability from the south to the north of our region [33]. The clay content is ranging from 15.3 g/kg to 684.58 g/kg. Mainly four soil units exist in the study area. The first unit includes Vertic saline soils. The second unit is Fluvisols characterized by a loamy soil texture located near the wadi bed of Merguellil. Another soil unit includes Fluvisols with coarse soil texture on the downstream part of the Merguellil River. The last soil unit, located in the north of Kairouan, includes clayey soil. Vertic soil is an indicator of the presence of montmorillonite clay. The organic matter rate is very low, near 0.5% from the surface to 10 cm, which corresponds to an isohumic diagnostic property of all soil units.
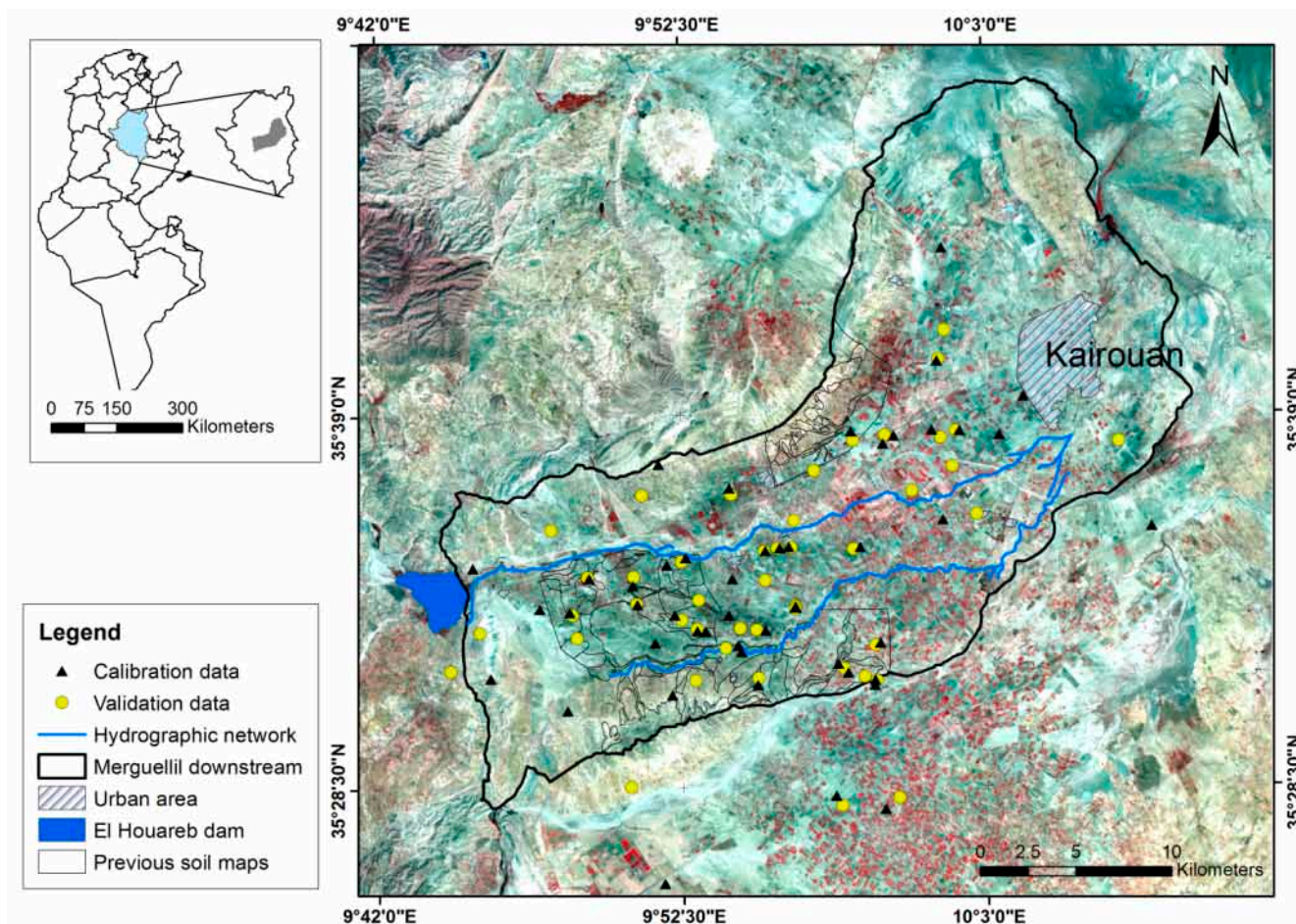

Figure 1. Study area and location of calibration and validation samples.

2.2. Data Collection and Processing

2.2.1. Ground Measurements

The dataset used in this study includes 100 samples. A first dataset of 30 samples was collected in March 2010 [12] over a limited area within the plain of Kairouan. An additional campaign was conducted during the summer season of 2013 after cereal harvesting over a larger part of the plain within a 670 km^2 area (Figure 1). During this campaign, 70 samples with a large spatial range of soil textures were selected along main and agricultural roads of the study area. Sample points were selected based on visual assessment of topsoil textural changes. These samples were located in bare soil areas and cereals

fields [12] corresponding to homogeneous areas extracted from MID-Infrared index maps, and over some olive fields with a vegetation fraction coverage lower than 10%. These plots are of sufficiently representative size (3 × 3 pixels) and easily accessible. The samples were collected from the surface to a depth of 5 cm. About 20 g were extracted for topsoil texture analysis. The initial samples were air-dried and sieved up to 2 mm. Topsoil texture measurements were carried out over the selected fields to measure clay content (soil fraction <2 µm), fine silt content (between 2 and 20 µm), coarse silt content (between 20 and 50 µm) and sand content (between 0.05 and 2 mm) by means of a classical laboratory analysis (Robinson pipette). To divide the dataset into calibration set and validation set, we kept one sample out of two. We considered 50 samples for calibration and 50 samples to validate the clay content map produced in this study. For clay content, the mean and standard deviation were 325.74 g/kg and 179.04 g/kg, respectively, and the measured values ranged between 15.3 and 684.6 g/kg. The coefficient of variation was high, about 55%. For sand content and silt content, the coefficients of variation were, respectively, about 85% and 60%. It confirms the high variability in the distribution of sand content over the Kairouan plain. Analyzed fields are plotted into texture triangle according to the USDA classification as illustrated by Figure 2. We can distinguish three dominant soil classes: clayey soil, silty soil and sandy soil as a continuous distribution between sandy and silty clay classes.
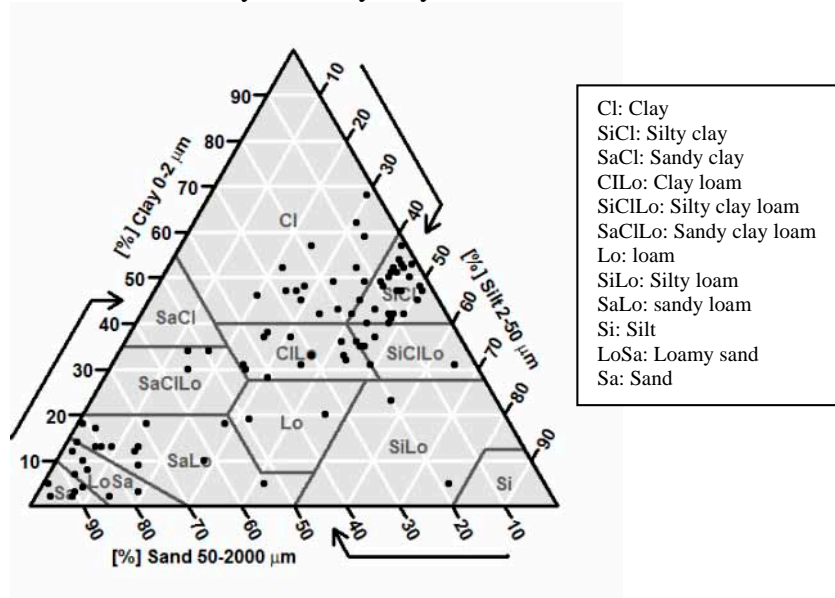
Figure 2. Distribution of 100 analyzed soil samples over the Kairouan plain across the USDA soil textural triangle.

2.2.2. Remote Sensing Data

To produce the topsoil texture map, time series of Landsat TM5 (2009) archival images were used. The Landsat TM5 was launched on 1 March 1984 and decommissioned on 5 June 2013. These images were downloaded from http://earthexplorer.usgs.gov/. Landsat TM5 data are acquired at 30 m spatial resolution. These images were geometrically and atmospherically corrected. All images are projected in the Universal Transverse Mercator system (UTM/Zone 32/North). The root mean square error of ortho-rectification is less than 1/2 pixel for each image. Visible, near- and mid-infrared bands calibration were carried out as developed in [34]. The method is based on a multi-temporal algorithm for detecting clouds, cloud shadows, free water and for estimating aerosol optical thickness (AOT). Subsequent preprocessing included

conversion of digital numbers to radiance, radiance to top-of-atmosphere reflectance and top of atmosphere reflectance to surface reflectance. The detailed method for correcting these images is illustrated in [35,36]. From the time series of Landsat TM data, we selected images during the seasons that minimize the influence of green vegetation (cereals mainly) and soil moisture due to rain. The latter is observed at the meteorological station. Calibration data were located on Landsat TM data. A polygon with 3×3 pixels was digitized to extract statistical data from Landsat TM images around each measurement. Polygons were represented within the sample plots in order to avoid selecting pixels at the edge.

2.2.3. Legacy Data

The 1:50,000 previous soil map used in our study was produced in 1975 by the Ministry of Agriculture, whose extent is shown in Figure 1. This map covers an area of about 13 km^2. It includes four soil type units: Fluvisols, Vertisols or Vertic soils, Calcisols and Isohumic soils.

We also used an existing topsoil clay content map derived from two TerraSAR-X radar images acquired over the Kairouan plain with a ground pixel spacing of 1 m [12]. The first TerraSAR-X radar image was acquired on 8 March 2010, just after a period of a heavy rain. The second one was acquired on 30 March 2010, when the soil was dry. The method used to produce clay content map from radar data was based on analyzing the difference in soil moisture content between the two images. The largest decrease in topsoil moisture is statistically related to the coarsest soil material based on a soil sample dataset of 15 samples [12,37].

3. Methods

3.1. Bare Soil Detection

Estimation of soil properties such as clay content is hampered if the pixels have a vegetation cover over 20% and also if the pixels are covered by crop residues [38]. We need rapid, accurate, and objective methods to detect residue cover and green vegetation in individual fields. Crop residue and soil are often spectrally similar and differ only by the magnitude of the reflectance in visible and near infrared wavelengths [39–42]. In fact, to extract bare soil extent over the Kairouan plain, we have to select Landsat TM data, which minimize the influence of soil moisture due to rain or irrigation. We also need to avoid crop residue after crop harvesting. To do so, we selected, in a first step, nine Landsat TM images with little influence of soil moisture due to rain on the basis of rainfall observations at meteorological stations. High soil moisture content decreases soil reflectance. In the second step, we retained only four images (17 July 2009, 1 August 2009, 2 September 2009 and 18 September 2009) in the summer season to avoid soil moisture due to local rain and irrigation. These images contain three types of pixels: bare soil pixels, plowed bare soil pixels and bare soil with crop residue. To extract only bare soil and to identify the agricultural practices, we analyzed simultaneously the TM3 profile, the NDVI profile and MID-infrared profile over all test fields as illustrated in Figure 3. In fact, using TM3 band, we can detect, approximately, the date of some agricultural practices such as crop harvesting and plowing. The plowing date can be identified by the decrease in red band reflectance values.

For each selected image, we masked urban areas, using a GIS shape file. The Normalized Difference Vegetation Index (NDVI) was used to mask green vegetation from each image by applying a sill of 0.2.

The image acquired on 18 September 2009 is contaminated by clouds. In this case, we used the provided cloud mask, generated during the image pre-processing, to mask them. The next step consists in surface changes detection (crop residues). We selected two MID-infrared images. One acquired at the beginning of the summer season, just after cereals harvesting (26 June 2009). This image contains a maximum of pixels affected by crop residue. The second image is acquired in 18 September 2009, when farmers are starting to plow their lands. The image of difference between these two MID-infrared dates was computed and analyzed. We observe that the highest values on the difference image correspond to fields with laying straw (crop residues) and not yet plowed. A value ranging from −0.08 to −0.04 extracted from the difference image was selected as a criterion for masking crop residues areas. Once we have no more crop residue, the MID-Infrared index becomes more stable. The image acquired on 2 September 2009 allows the best prediction of topsoil clay content. However, for the illustrated example in Figure 3, on this date the field was still covered by crop residues. Based on the MID-Infrared index profile extracted within the studied field, it becomes stable from the date of 18 September 2009. This date can be considered the best date for topsoil clay prediction within this field.
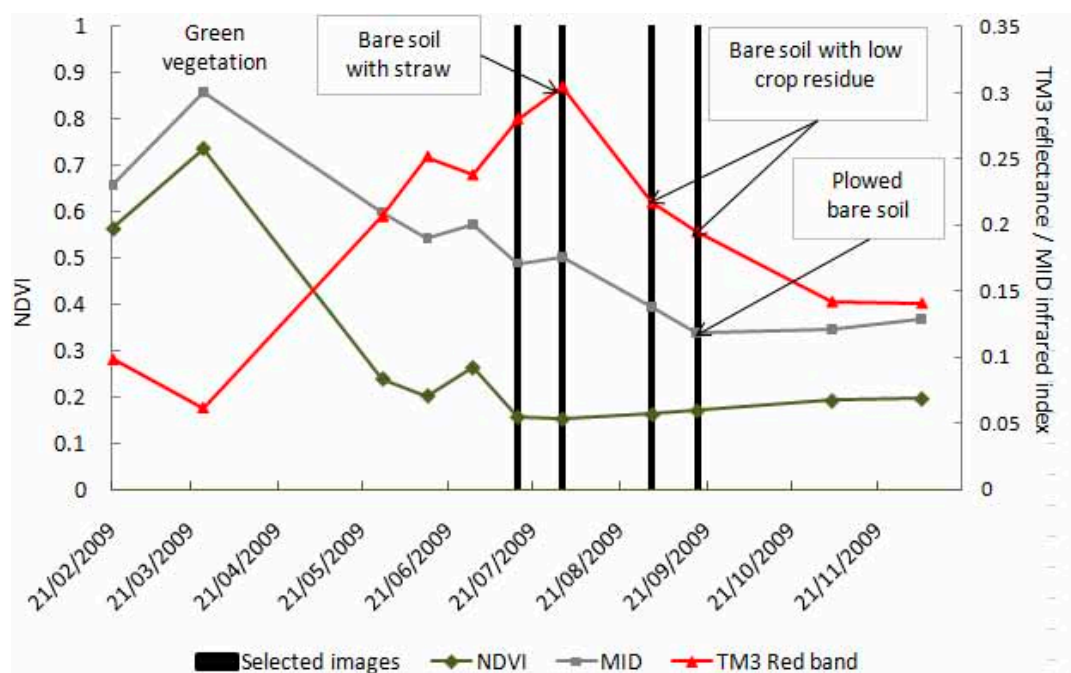
Figure 3. Example of Normalized Difference Vegetation index (NDVI), Landsat Red band (TM3) and MID-Infrared index profiles extracted from selected Landsat TM images, over a cereal test plot of 8 hectares (clayey soil texture) during the period ranging from 21 February 2009 until 21 November 2009.

3.2. Prediction of Clay Content

3.2.1. Calibrating MID Infrared Index and Topsoil Clay Content

The MID-infrared index is defined in Equation (1) and represents a normalized index between TM band 5 and TM band 7. Applying this index over the driest bare soils allows separating sandy soils from clayey ones.

This index was applied over bare soil for each selected image in order to produce mid-infrared index maps at 30 m spatial resolution.

$$MID - Infrared\ index = \frac{TM\ 5 - TM\ 7}{TM\ 5 + TM\ 7} \qquad (1)$$

One out of two samples from the 100 samples was selected in order to build the calibration dataset. We applied green vegetation and crop residue mask for each selected image: 20 samples corresponding to fields with either green vegetation or crop residue were thus masked and not available for calibration. The remaining 30 samples were used for calibration with topsoil clay content values ranging from 50 g/kg to 684 g/kg with a mean value of about 323 g/kg. These values span the whole range of observed topsoil clay content with a satisfying distribution amongst dominant texture classes. The use of these 30 samples seems therefore sufficient to represent topsoil clay content variability within the study area. They were distributed in the different textural classes in our study area. A mean value of MID-infrared index was extracted from each test field and used with clay content. This step allowed establishing a statistical relationship (one for each selected image) between the mid-infrared index and the soil clay content using a simple linear regression model. These relationships were used to produce clay content maps at 30 m spatial resolution over our study area and for each selected image. The resulting maps contain two types of pixels: pixels with clay content estimates and others with no data over the masked areas that differ from one image to another, except for orchards. To get an exhaustive coverage of clay content at 100 m spatial resolution over the masked area, we combined the four clay content images using ordinary cokriging (OCK).

3.2.2. Ordinary Cokriging

Ordinary cokriging is a multivariate variant of ordinary kriging and computes the best linear unbiased estimator with minimum error variance [43,44]. Ordinary cokriging takes advantages of the correlation between variables and allows for local variability of the means by restricting the stationarity of both primary and secondary variables to a local neighborhood *W(u)* centered on the location *u* being estimated. Ordinary cokriging examines the spatial relations among two or more variables by determining their coregionalization. If coregionalization exists, then it is feasible to use the spatial information of co-variables to improve the predictions of the target variable through cokriging. For each variable a variogram describing spatial dependence of a spatial random field or a stochastic process can be estimated by Equation (2):

$$\gamma(u) = \frac{1}{2\,n_i(u)} \sum_{\alpha_i=1}^{n_i(u)} \left[Z_i\left(u_{\alpha_i}\right) - Z_i\left(u_{\alpha_i} + u\right) \right]^2 \qquad (2)$$

where *u* is the spatial lag distance between two locations; $n_i(u)$ is the number of observed data pairs with the lag *u*; and $Z_i\left(u_{\alpha_i}\right)$ and $Z_i\left(u_{\alpha_i} + u\right)$ are two measured values at locations, u_{α_i} and $u_{\alpha_i} + u$, respectively.

The spatial interdependence between any two variables can be examined by the cross variogram $\gamma_{ij}(u)$, which is calculated according to Equation (3):

$$\gamma_{ij}(u) = \frac{1}{2\,n_i(u)} \sum_{\alpha_i=1}^{n_i(u)} \left[Z_i\left(u_{\alpha_i}\right) - Z_i\left(u_{\alpha_i} + u\right) \right] \left[Z_i\left(u_{\alpha_i}\right) - Z_j\left(u_{\alpha_i} + u\right) \right] \qquad (3)$$

In order to describe the experimental variograms and cross variograms for use in cokriging, a model of coregionalization has to be fitted to these variograms, which consider the spatial dependence of two or more sets of variables and their interdependence simultaneously. By incorporating related secondary information, cokriging takes into consideration spatial cross correlation between primary and secondary variables to improve interpolation accuracy.

For the case of two or more secondary variables, the ordinary cokriging estimator is given by Equation (4):

$$Z_{OCK}^{(1)^*}(u) = \sum_{i=1}^{N_v} \sum_{\alpha_i=1}^{n_i(u)} \lambda_{\alpha_i}^{OCK}(u) Z_i(u_{\alpha_i}) \tag{4}$$

where N_v is the number of variables Z_i and $\lambda_{\alpha_i}^{OCK}$ represents data weights:

$$\sum_{\alpha_1=1}^{n_1(u)} \lambda_{\alpha_i}^{OCK}(u) = 1 \tag{5}$$

$$\sum_{\alpha_i=1}^{n_i(u)} \lambda_{\alpha_i}^{OCK}(u) = 0 \qquad i = 2,..., N_v \tag{6}$$

The cokriging variance can be estimated from the covariance C_{11} and the cross-covariances C_{i1} through Equation (7):

$$\sigma_{OCK}^2(u) = C_{11}(0) - \mu_1^{OCK}(u) - \sum_{i=1}^{N_v} \sum_{\alpha_i=1}^{n_i(u)} \lambda_{\alpha_i}^{OCK}(u) C_{i1}(u_{\alpha_i} - u) \tag{7}$$

where μ_1^{OCK} is the Lagrange multiplier of the cokriging system and $(u_{\alpha i}–u)$ denotes the distance between $u_{\alpha i}$ and u.

From each clay content image, a point vector layer at 100 m with clay content estimates was extracted and used for interpolating clay content over masked areas. For each date, an experimental semi-varioram was calculated. The Geostatistical operations were performed with ArcMap Software (version 10.1) using the Geostatistical wizard.

3.3. Validation Method

Clay content predicted by cokriging at each point of the validation data was extracted and compared to the measured clay content at the same location. The performance of the prediction models was evaluated with the following statistical criteria: the coefficient of determination (R^2) (Equation (8)), and the Root Mean Square Error (RMSE) (Equation (9)). The coefficient R^2 measures the effectiveness of a variable to predict another variable, whereas RMSE measures the average error of prediction. The indices were calculated according to [45]. To perform the calibration model estimation, we used only the R^2 and the RMSE. In addition, to independent soil samples, validation was performed using existing soil maps. The validation consists on comparing topsoil clay content and the soil typology extracted from existing soil maps.

$$R^2 = \frac{\sum_{i=1}^{n} \left(z^*(x_i) - \bar{z} \right)^2}{\sum_{i=1}^{n} \left(z(x_i) - \bar{z} \right)^2} \tag{8}$$

$$RMSE = \sqrt{\frac{1}{n} \times \sum_{i=1}^{n} \left(z(x_i) - z^*(x_i) \right)^2} \qquad (9)$$

where point observations were denoted by n, measured and predicted clay content at i^{th} location as $z(x_i)$ and $z^*(x_i)$, the horizontal bar represents the mean of the observed values.

4. Results

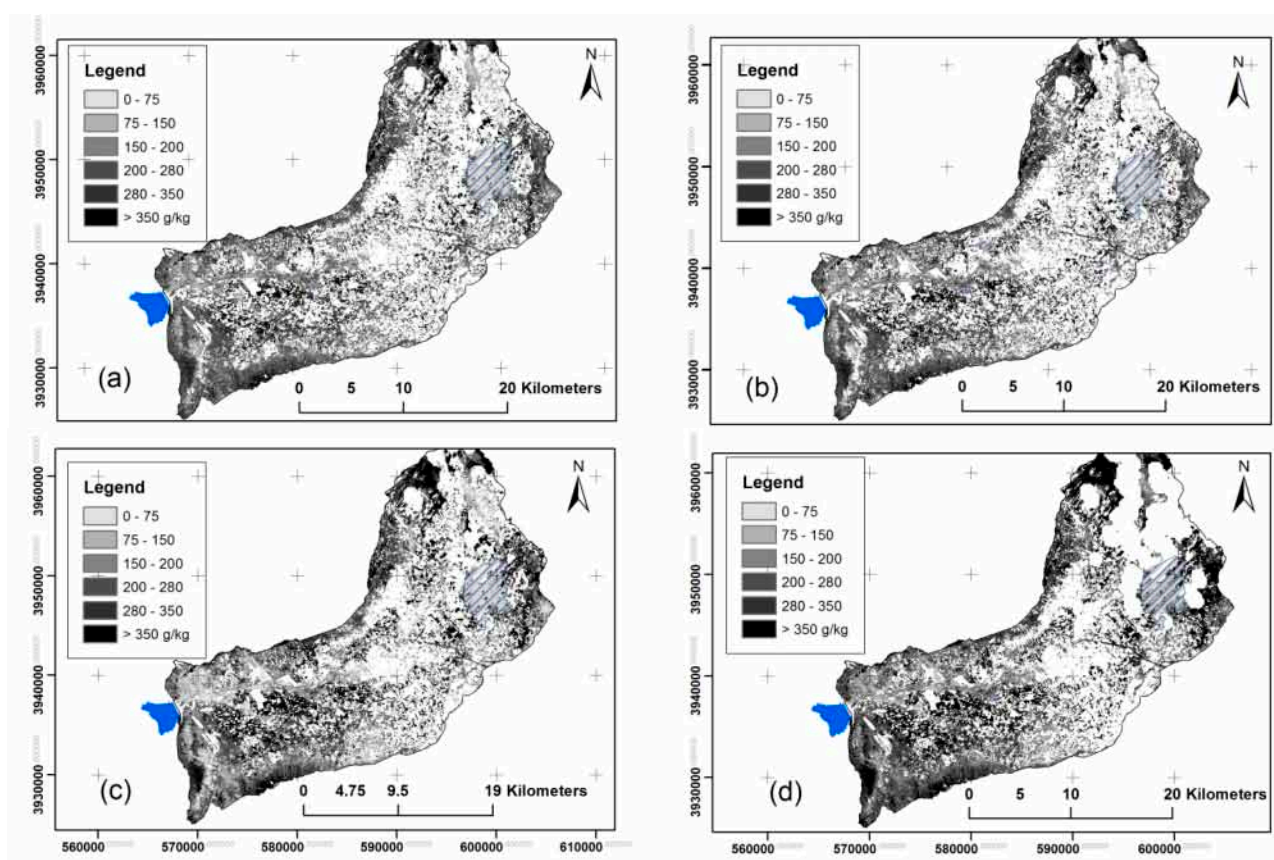### 4.1. Relationships between MID-Infrared Index and Soil Clay Content

To predict clay content over our study area, a linear regression model was built to establish the relationship between mid-infrared index and clay content corresponding to the 30 selected sites for each of the four selected dates: 16 July 2009, 1 August 2009, 2 September 2009 and 18 September 2009. Resulting relationships indicated that MID infrared index and clay content are significantly correlated (Figure 4). The coefficients of determination R^2 were, respectively, 0.48, 0.53, 0.73 and 0.63, while RMSE was, respectively, estimated at 13%, 12%, 10% and 11% of the mean clay content, which corresponds to about 130 g/kg.
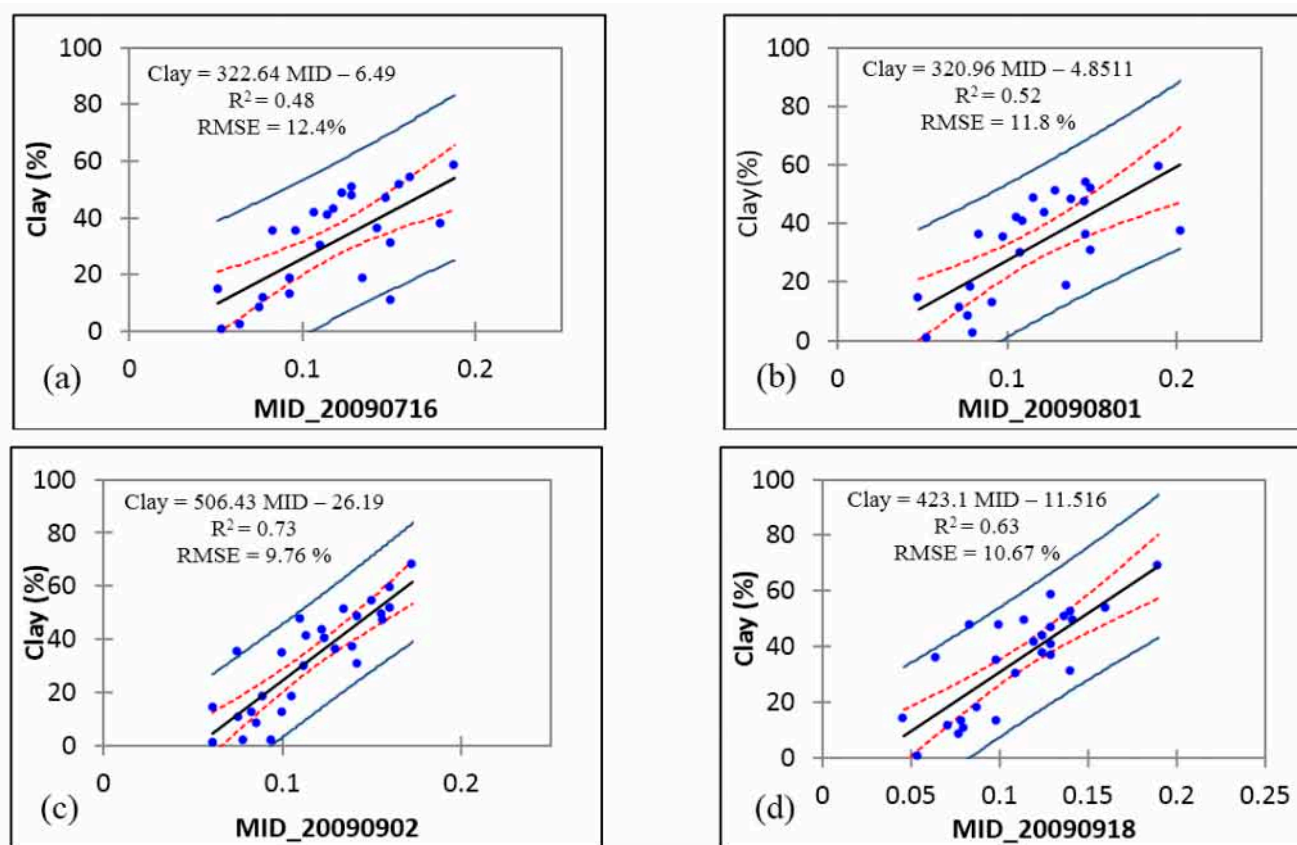
Figure 4. Linear relationship between MID-infrared index and observed clay content: (**a**) 16 July 2009; (**b**) 1 August 2009; (**c**) 2 September 2009; and (**d**) 18 September 2009. (Blue dots) observed clay content values. The red dotted line is the confidence interval of the mean. The blue line is the confidence interval for clay content observation. The continuous black line is the y-axis data *versus* x-axis data regression line.

The clay content maps at 30 m spatial resolution generated by the relationships between the mid-infrared index and the clay content are illustrated in Figure 5. We can distinguish between two types of pixels: white ones correspond to the masked areas and grey ones correspond to clay content values. The predicted clay content is ranging from 15.3 to 684 g/kg. The vegetation and crop residues masks differ from one image to another and the coverage of the remaining area is incomplete. The percentage of information contained in each image is 49%, 47%, 52% and 49%, respectively, for 16 July 2009 image, 1 August 2009, 2 September 2009 and 18 September 2009. The decrease in the percentage of information that the image contains for the 18 September 2009 image is due to cloud mask applied to mask cloud coverage.

From these maps, we extracted the clay content value of each 100 m grid and regular lattices at 100 m spatial resolution were prepared in order to interpolate the clay content over masked areas.

Figure 5. Clay content maps at 30 m spatial resolution obtained using relationships between observed clay content and mid-infrared index for each selected date: (**a**) 16 July 2009; (**b**) 1 August 2009; (**c**) 2 September 2009; and (**d**) 18 September 2009.

4.2. Ordinary Cokriging Parameters

The image acquired on 2 September 2009 was used as the primary variable for cokriging. Images acquired on 16 July 2009 and 1 August 2009 and 18 September 2009 were used as secondary variables. To fit the experimental variogram and the cross variogram of the four selected images, we used a stable model. Parameters of established variograms are listed in Table 1. The same range and lag size values were used in the fitting procedure, which are 3681 m and 102 m, respectively. The resulting variograms

show a nugget effect. It represents random variance often caused by measurement error or micro variability of topsoil clay content, which cannot be detected at the scale of sampling.

Table 1. Stable model variogram parameters for the image acquired on 16 July 2009, 1 August 2009, 2 September 2009 and 18 September 2009.

| Variables	C_0 (Nugget)	C (Partial Sill)	Total sill ($C_0 + C$)	Range (m)	$C_0/(C_0+C)$ (%)
16 July 2009	23.84	68.01	91.85	3681	25.9
1 August 2009	29.62	69.65	99.27	3681	29.8
2 September 2009	55.38	86.17	141.55	3681	39.1
18 September 2009	54.57	122.81	177.38	3681	30.7

The resulting variogam for topsoil clay content extracted from the image acquired on 2 September 2009 and cross-variograms between the primary variable and the secondary variable are illustrated in Figure 6.
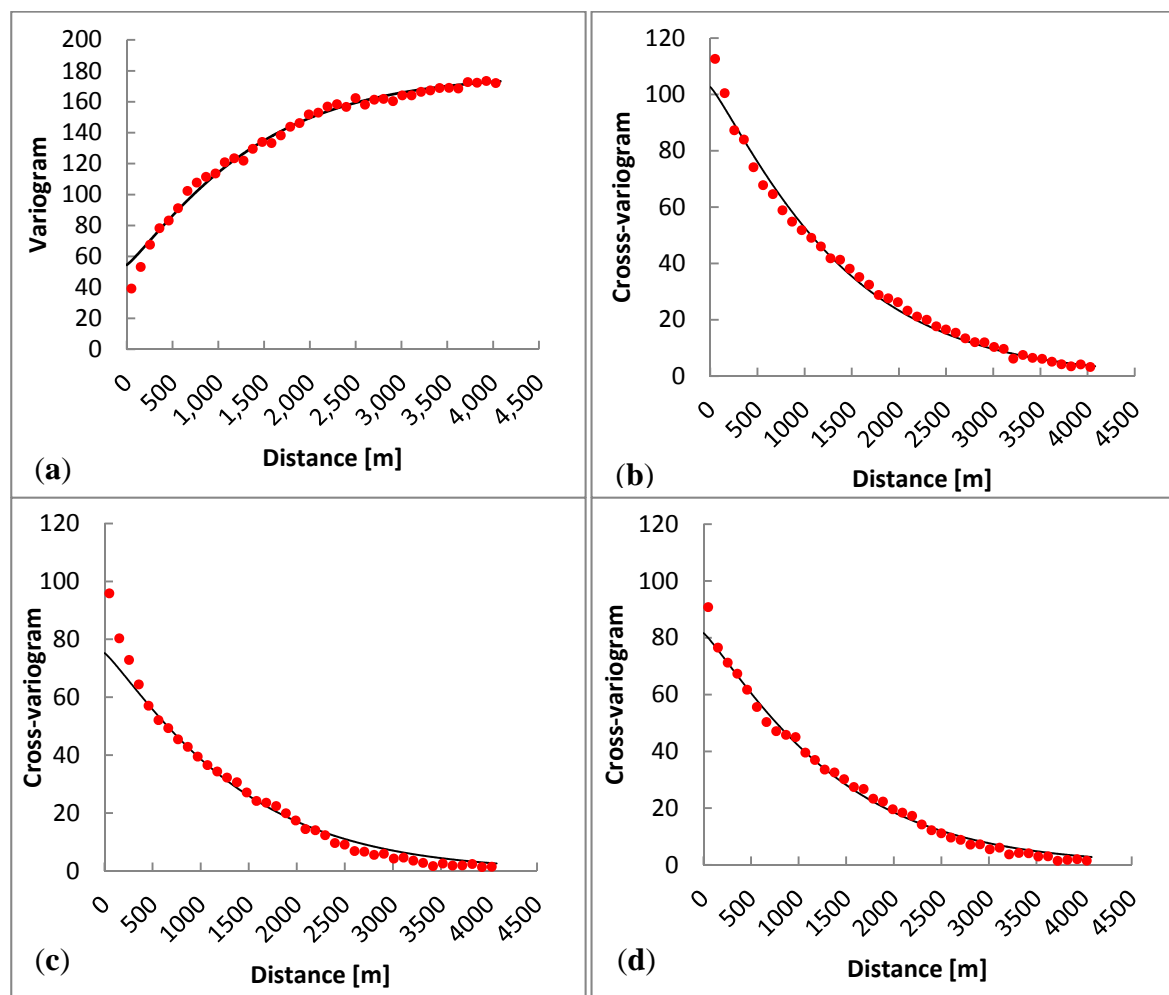
Figure 6. (**a**) Experimental semi-variograms (red dots) and model fits (black line) for topsoil clay content extracted from Landsat TM5 image acquired on 2 September 2009; (**b**) Cross-variogram between clay content from 2 September 2009 and 18 September 2009; (**c**) Cross-variogram between clay content from 2 September 2009 and 1 August 2009; (**d**) Cross-variogram between clay content from 2 September 2009 and 16 July 2009.

4.3. Topsoil Clay Content Map Produced by Ordinary Cokriging

Cokriging was used to combine the 4 images and to produce a clay content map at 100 m spatial resolution, which is illustrated in Figure 7. The obtained clay content map is classified into six classes of clay content according to the USDA texture. Defined textural classes boundaries allow mapping homogeneous areas. The two first classes include areas with clay content values ranging from 10 to 150 g/kg. These areas correspond to the Merguellil wadi bed represented by a linear structure in the north of the western half of the plain. A medium texture class includes clay content values between 200 and 280 g/kg. Within the red areas, the clay content values are higher than 280 g/kg. These areas are cultivated essentially by irrigated cereals and winter and summer vegetables. In the north of Kairouan city, we obtained a fine topsoil texture with clay content ranging from 280 g/kg to more than 350 g/kg. It corresponds to the last soil unit with clayey salt-affected soils described in Section 2.1.
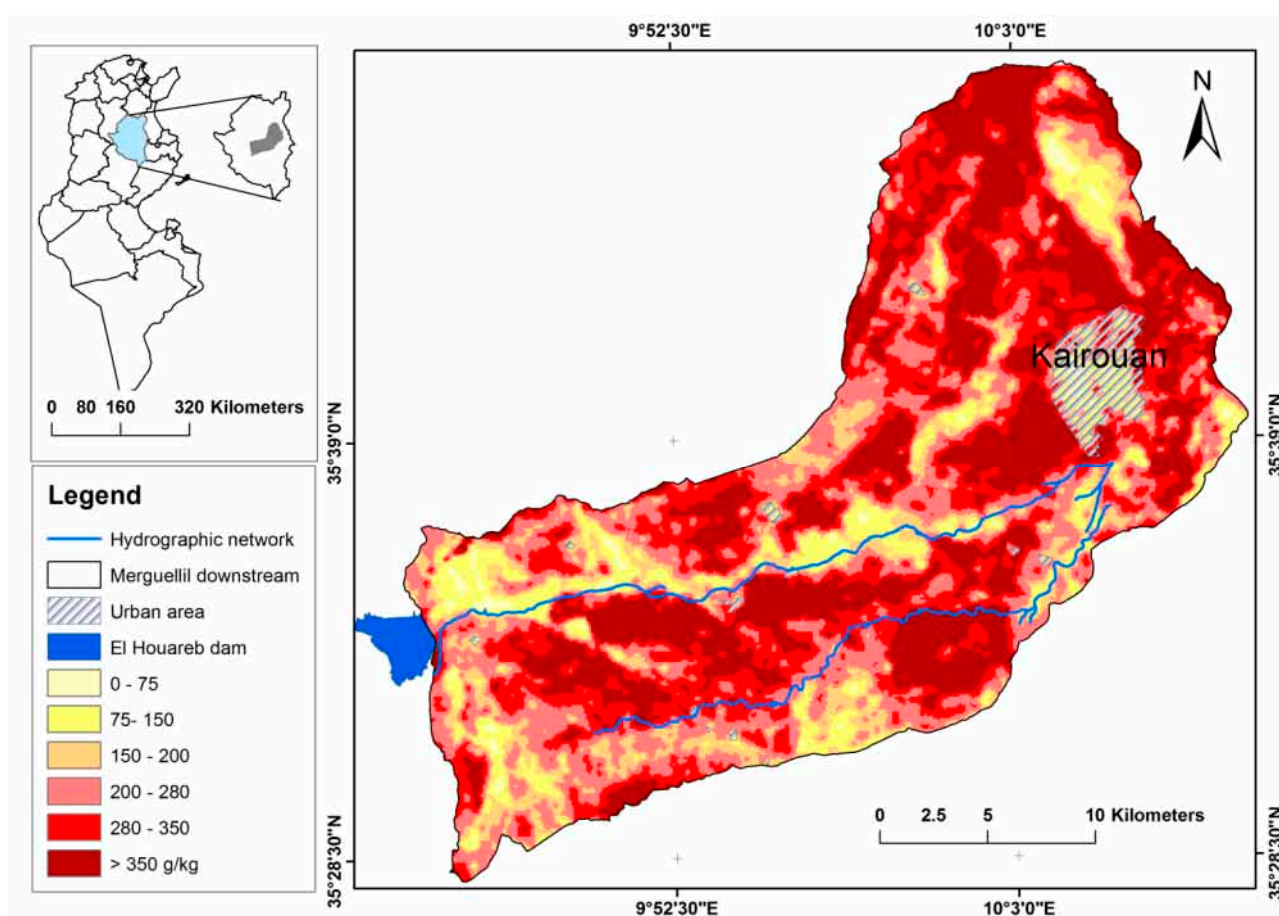

Figure 7. Clay content map at 100 m spatial resolution produced by combining four Landsat TM images: 16 July 2009, 1 August 2009, 2 September 2009 and 18 September 2009. (Blue line) corresponds to hydrographic network.

4.4. Validation of Clay Content Map from Landsat TM Data Using Independent Soil Samples and Existing Soil Maps

The resulting clay content map was validated by comparison with 50 samples over the study area (Figure 1). RMSE is about 10%, corresponding to 100 g/kg of clay. The R^2 between predicted and observed values is close to 0.65. Comparison between validation data set and clay content values

predicted by cokriging is illustrated in Figure 8. The model used to predict topsoil clay content overestimates low observed clay content.
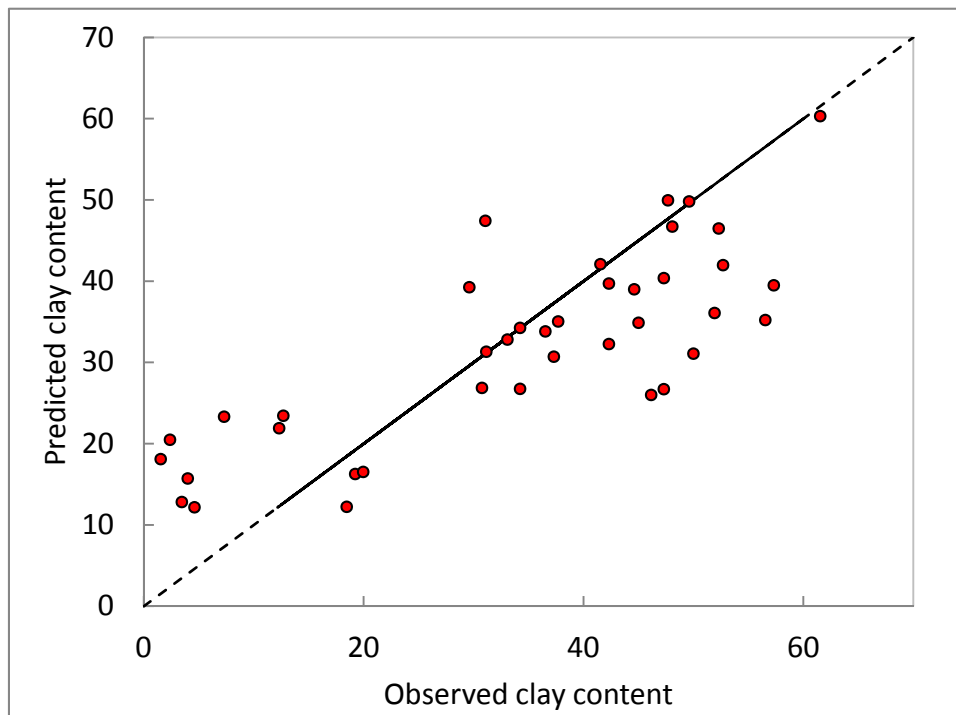
Figure 8. Comparison of observed topsoil clay content and cokriging predicted topsoil clay content at 50 independent sampling points. The dashed line is 1:1 line.

Table 2. Area percentage of each combination between soil units and produced clay content map from Landsat TM data over an area of 13 km^2.

			Clay Content Classes					
			1	**2**	**3**	**4**	**5**	**6**
			0–75	**75–150**	**150–200**	**200–280**	**280–350**	**>350 g/kg**
Soil units from existing soil map	1	Fluvisols	0.19	3.95	9.18	<u>15.44</u>	9.00	8.02
	2	Vertisols	0.09	0.73	0.99	4.14	<u>10.44</u>	<u>16.45</u>
	3	Isohumic	0	0.97	4.31	6.94	2.79	1.28
	4	Calcisols	0	0.33	1.20	1.52	1.04	1.00

Another validation of the clay content map was achieved by comparison with existing soil maps, produced in 1975 by the Ministry of Agriculture at 1:50,000 scale. To validate our map, we compared those existing soil type maps and the soil clay map derived from Landsat TM data. The percentage of clay content classes is calculated for each soil map unit in Table 2. The first soil unit (Fluvisols) is marked by a high variability in clay content values in agreement with known characteristics of these soils. The dominant class represents 15% of the area and the clay content value is ranging from 200 to 280 g/kg. It corresponds to a medium topsoil texture class. The distribution of clay content classes higher than 350 g/kg is representative of Vertisols (16% of the area) and in agreement with clay content values observed by [46] (clay content value ranging from 350 to 520 g/kg). The isohumic soil is present in the Kairouan plain with different clay content values as identified with Landsat TM data. In addition, the Merguellil wadi bed is well identified as a linear structure in the map obtained from Landsat TM. The

results show that the distribution of clay content produced by Landsat TM data is consistent with the soils units. Thus, Landsat TM type data can be used to update existing soil maps, to improve complex soil units according to clay criterion and to produce soil texture information where not available.

5. Discussion

The model developed for clay content prediction was built from time series of Landsat TM data. Its main advantage is to provide an acceptable clay content estimation over fields with a limited number of samples (only 30 samples were used for calibration), which can be applied over large regions.

5.1. Calibration Models

For each selected image, results showed that the MID infrared index is correlated with topsoil clay content. Results were approximately in the range of error estimations derived from other remote sensing techniques to estimate clay content. With hyper-spectral Hymap images, authors obtained an R^2 of 0.73 [47]. Using TerraSAR-X radar data, authors obtained an R^2 of 0.6 [12]. The relative scattering of the plots can be due to sites with a heterogeneous surface: incomplete plowing, mixed pixels and the related MID infrared sill used for bare soil mapping. The effect of the atmosphere on the signal attenuation decreases in the MID-infrared domain.

5.2. Comparison between Clay Content Maps Produced from Landsat TM and TerraSAR-X Radar Data

The accuracy of our clay content prediction model ($R^2 = 0.64$ and a RMSE of about 10% corresponding to 100 g/kg) is in agreement with an existing study [12] who obtained an $R^2 = 0.6$ and an RMSE about 12% (120g/kg). A comparison was therefore carried out with this study based on TerraSAR-X radar data on a smaller area (Figure 9a,b) where neither calibration nor validation points were present. Before comparing these two maps, the same mask of green vegetation is applied over the clay content map extracted from Landsat TM data. The clay map produced from TerraSAR-X radar data displays sharp contrasts over short distances compared to the smoother fields of Landsat TM. This is due to the effect of residual speckle. Cokriging and the rough resolution of Landsat TM lead to more continuous fields and therefore a smoother variation from one spatial element of similar texture to the other. Both maps show similar features such as the wadi bed of Merguellil with low clay content values. The other regions show similarities and differences in clay content values. Additional field observations were made without analysis to differentiate coarse and fine textural classes over some areas. Within region (I), we obtained high clay content value of about 60% when using Landsat TM data and low clay content value of about 15% using TerraSAR-X radar data. Field observation shows that the fields were not cultivated and this situation corresponds to bare soil areas with fine topsoil texture. Landsat TM, therefore, estimates clay content better than TerraSAR-X radar data within this area.

Within region (II), the use of TerraSAR-X data overestimates the clay content compared to Landsat TM data. It can be due to the presence of olive trees, which cover 5% of the soil surface. Concerning the third region, we predicted high clay content when using Landsat TM data. This is in agreement with our *in-situ* observations, but only in partial agreement according to radar data. In these areas, the dominant crops are cereals, beans, chilies and tomatoes; all of them being irrigated and over fine topsoil texture.

For region (IV), more consistent outputs were provided when using Landsat TM data than when using TerraSAR-X radar data. We predicted fine topsoil texture using TM data with clay content ranging from 35% to 45%. The difference between both approaches can be due to soil moisture variation and to green vegetation at the date of image acquisition and to the erroneously non-masked crop residue areas.
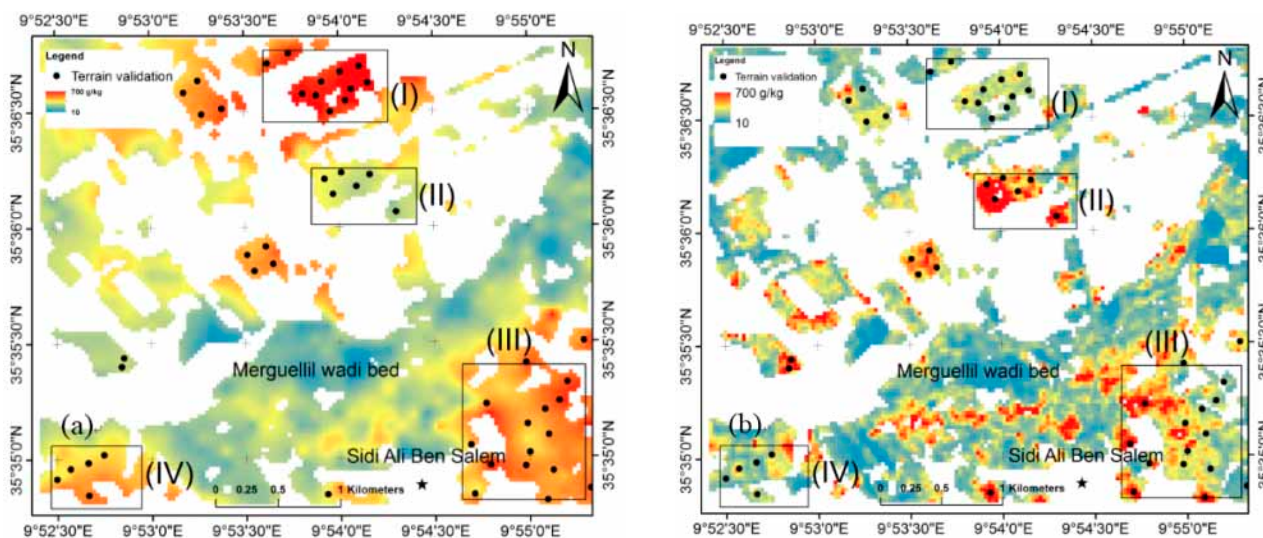
Figure 9. (**a**) Extract of the clay content map produced by Landsat TM data (2009). (**b**) Clay content map produced by TerraSAR-X data over only bare soil areas (2010). Black dots are the field observation in January 2015. Black frames are the compared sub sites.
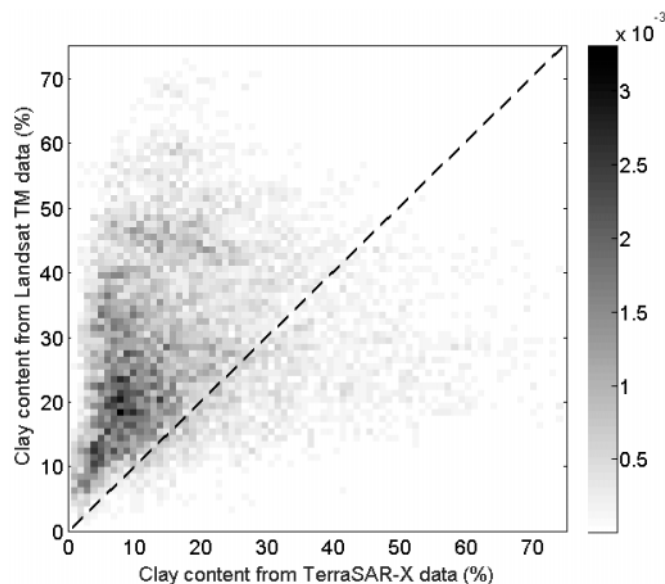
Figure 10. Density plot of clay content obtained from Landsat TM data and clay content value from TerraSAR-X radar data. The dashed line is the 1:1 line.

The difference between both approaches is validated by the density plot of clay content values obtained from Landsat TM data and TerraSAR-X radar data (Figure 10). In this area, the bias is about 5% for radar data. The central part of the scatter plot is occupied by clay content values ranging from 10% to 25% for Landsat TM data and from 5% to 10% for TerraSAR-X radar data. The average clay content value from TerraSAR-X radar data is significantly lower than the average clay content estimated

using Landsat TM data. Disagreement between Landsat TM and TerraSAR-X products could be particularly explained by the presence of irrigation over some agricultural fields during experimental campaigns, which induce large errors for radar methodology.

6. Conclusions

The purpose of this paper was to develop a multi-temporal approach using Landsat TM data in order to map topsoil clay content at a fine spatial resolution. To predict topsoil clay content spatial interpolation can be performed with ordinary cokriging using more than one variable. The results show that remote sensing is an efficient and low-priced tool that enables generating a topsoil clay content map. We can predict clay content for each pixel of the image, thus avoiding sampling a large number of fields. The resulting topsoil clay content map is at 100 m spatial resolution and is reasonably accurate with an R^2 about 0.64 and a RMSE of 10% (100 g/kg). Cokriging made it possible to fill in the information gaps created by the vegetation mask and to validate model predictions. Improvement of relationships between MID-infrared index and clay content need more sample sites. The method needs to be tested with other satellites e.g., the new Landsat-8 and the future Sentinel2 (ESA, launch scheduled in April 2015). The ground resolution of the latter one is more accurate (20 m) to detect homogenous surface. The revisit frequency of this sensor is less than five days. With this type of data, we can improve mapping of bare soils in dry conditions. Comparisons with radar show differences. Moreover, these two approaches are complementary if we consider that visible and middle infrared reflectances are limited to surface characteristics, and radar data are limited to the drying speed of the first centimeters. Such topsoil clay content maps will be used to calculate soil hydrodynamic properties with pedotransfer functions. The surfaces, detected in the visible and MID-infrared bands, represent the agronomic horizon. This horizon is regularly homogenized by plowing. This makes it difficult to estimate clay content of horizons beyond 30 cm. To do so, we need to combine remote sensing data with some auxiliary data, such as digital elevation model, geological and geomorphologic maps.

Author Contributions

Marouen Shabou, Bernard Mougenot and Zohra Lili Chabaane proposed to use Landsat TM data to produce clay content map over the Kairouan plain. Christian Walter suggested using cokriging to fill information gaps created by vegetation and crop residue masks. Gilles Boulet and Mehrez Zribi proposed to compare clay content map produced from Landsat TM data and TerraSAR-X radar data. Nadhira Ben Aissa participated to ground measurements and laboratory analysis.

Conflicts of Interest

The authors declare no conflict of interest

References

1. Russel, E.J. *Soil Conditions and Plant Growth*, 10th ed.; Longman: London, UK, 1973.
2. Bresler, E.; Dagan, G.; Wagenet, R.J.; Laufer, A. Statistical analysis of salinity and texture effects on spatial variability of soil hydraulic conductivity1. *Soil Sci. Soc. Am. J.* **1984**, *48*, 16–25.
3. Frenkel, H.; Goertzen, J.O.; Rhoades, J.D. Effects of clay type and content, exchangeable sodium percentage, and electrolyte concentration on clay dispersion and soil hydraulic conductivity. *Soil Sci. Soc. Am. J.* **1978**, *42*, 32–39.
4. Dent, D.; Young, A. *Soil Survey and Land Evaluation*; HarperCollins Publishers Ltd.: London, UK, 1981.
5. Demattê, J.A.M.; Galdos, M.V.; Guimarães, R.V.; Genú, A.M.; Nanni, M.R.; Zullo, J. Quantification of tropical soil attributes from etm+/landsat-7 data. *Int. J. Remote Sens.* **2007**, *28*, 3813–3829.
6. Shepherd, K.D.; Walsh, M.G. Development of reflectance spectral libraries for characterization of soil properties. *Soil Sci. Soc. Am. J.* **2002**, 988–998.
7. McKenzie, N.J.; Cresswell, H.P.; Ryan, P.J.; Grundy, M. Contemporary land resource survey requires improvements in direct soil measurement. *Commun. Soil Sci. Plant Anal.* **2000**, *31*, 1553–1569.
8. Gomez, C.; Lagacherie, P.; Coulouma, G. Regional predictions of eight common soil properties and their spatial structures from hyperspectral VIS-NIR data. *Geoderma* **2012**, *189–190*, 176–185.
9. Mulder, V.L.; de Bruin, S.; Schaepman, M.E.; Mayr, T.R. The use of remote sensing in soil and terrain mapping—A review. *Geoderma* **2011**, *162*, 1–19.
10. Odeh, I.O.A.; McBratney, A.B. Using AVHRR images for spatial prediction of clay content in the lower namoi valley of Eastern Australia. *Geoderma* **2000**, *97*, 237–254.
11. Sullivan, D.G.; Shaw, J.N.; Rickman, D. IKONOS imagery to estimate surface soil property variability in two alabama physiographies *Soil Sci. Soc. Am. J.* **2005**, *69*, 1789–1798.
12. Zribi, M.; Kotti, F.; Lili-Chabaane, Z.; Baghdadi, N.; Ben Issa, N.; Amri, R.; Duchemin, B.; Chehbouni, A. Soil texture estimation over a semiarid area using TerraSAR-X radar data. *IEEE Geosci. Remote Sens. Lett.* **2012**, *9*, 353–357.
13. Scull, P.; Franklin, J.; Chadwick, O.A.; McArthur, D. Predictive soil mapping: A review. *Progr. Phys. Geogr.* **2003**, *27*, 171–197.
14. Dogan, H.M.; Kilic, O.M. Modelling and mapping some soil surface properties of central Kelkit basin in Turkey by using Landsat-7 ETM+ images. *Int. J. Remote Sens.* **2013**, *34*, 5623–5640.

15. Cohen, W.B.; Maiersperger, T.K.; Gower, S.T.; Turner, D.P. An improved strategy for regression of biophysical variables and landsat ETM+ data. *Remote Sens. Environ.* **2003**, *84*, 561–571.

16. Moran, M.S.; Bryant, R.; Thome, K.; Ni, W.; Nouvellon, Y.; Gonzalez-Dugo, M.P.; Qi, J.; Clarke, T.R. A refined empirical line approach for reflectance factor retieval from Landsat-5 TM and Landsat-7 ETM +. *Remote Sens. Environ.* **2001**, *78*, 71–82.

17. Irons, J.R.; Weismiller, R.A.; Pterson, G.W. Soil reflectance. In *Theory and Applications of Optical Remote Sensing*; Asar, G., Ed.; John Wiley: New York, NY, USA, 1989; pp. 66–106.

18. Dogan, H.M. Applications of remote sensing and geographic information systems to assess ferrous minerals and iron oxide of Tokat Province in Turkey. *Int. J. Remote Sens.* **2008**, *29*, 221–233.

19. Dogan, H.M. Mineral composite assessment of kelkit river basin in Turkey by means of remote sensing. *J. Earth Syst. Sci.* **2009**, *118*, 701–710.

20. Madeira, N.J. *Etude Quantitative des Relations Constituants Minéralogiques-Réflectance Diffuse des Latosols Brésiliens: Application à l'utilisation Pédologique des Données Satellitaires TM (Région de Brasilia)*; Universite Pierre et Marie Curie: Paris, France, 1993.

21. Scull, P.; Franklin, J.; Chadwick, O.A.; McArthur, D. The application of classification tree analysis to soil type prediction in a desert landscape. *Ecol. Modell.* **2005**, *181*, 1–15.

22. 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.

23. Ben-Dor, E.; Taylor, R.G.; Hill, J.; Demattê, J.A.M.; Whiting, M.L.; Chabrillat, S.; Sommer, S. Immaging spectrometry for soil applications. In *Advances in Agronomy*; Donald, L.S., Ed.; Academic Press: Waltham, MA, USA, 2008; Volume 97, pp. 321–392.

24. Chang, C.W.; Laird, D.A.; Musbach, M.J.; Helder, D.L. Near-infrared reflectance spectroscopy–principal components regression analyses of soil properties. *Soil Sci. Soc. Am. J.* **2001**, *65*, 480–490.

25. COLEMAN, T.L.; MONTGOMERY, O.L. Soil moisture, organic matter and iron content effects on thespectral characteristics of selected vertisols and Alfisols of Alabama. *Photogramm. Eng. Remote Sens.* **1987**, *53*, 1695–1663.

26. Dunn, B.W.; Batten, G.D.; Beecher, H.G.; Ciavarella, S. The potential of near-infrared reflectance spectroscopy for soil analysis—A case study from the riverine plain of South-Eastern Australia. *Austr. J. Exp. Agric.* **2003**, *42*, 607–614.

27. Madeira, J.N. Spectral reflectance properties of soils. *Photointerpretation* **1996**, *34*, 59–70.

28. Daughtry, C.S.T.; Doraiswamy, P.C.; Hunt, E.R., Jr.; Stern, A.J.; McMurtrey, J.E., III.; Prueger, J.H. Remote sensing of crop residue cover and soil tillage intensity. *Soil Tillage Res.* **2006**, *91*, 101–108.

29. Jacques, D.C.; Kergoat, L.; Hiernaux, P.; Mougin, E.; Defourny, P. Monitoring dry vegetation masses in semi-Arid areas with MODIS SWIR bands. *Remote Sens. Environ.* **2014**, *153*, 40–49.

30. van Deventer, A.P.; Ward, A.D.; Gowda, P.H.; Lyon, J.G. Using thematic mapper data to identify contrasting soil plains and tillage practices *Photogramm. Eng. Remote Sens.* **1997**, *63*, 97–93.

31. Zribi, M.; Chahbi, A.; Shabou, M.; Lili-Chabaane, Z.; Duchemin, B.; Baghdadi, N.; Amri, R.; Chehbouni, A. Soil surface moisture estimation over a semi-arid region using ENVISAT ASAR radar data for soil evaporation evaluation. *Hydrol. Earth Syst. Sci.* **2011**, *15*, 345–358.

32. Batjes, N.H. *Soil Property Estimates for Tunisia Derived from Soter and Wise (Sotwis-Tunisia, Ver. 1.0)*; ISRIC-World Soil Information: Wageningen, The Netherlands, 2010; p. 41.

33. Barbery, J.; Mohdi, M. *Carte des Ressources en Sols de la Tunisie, Feuille de Kairouan*; Direction des ressources en eau et en sol: Division des sols, Tunisie, 1987; p 53.

34. Chander, G.; Markham, B.L.; Helder, D.L. Summary of current radiometric calibration coefficients for Landsat MSS, TM, ETM+, and EO-1 ALI sensors. *Remote Sens. Environ.* **2009**, *113*, 893–903.

35. Hagolle, O.; Dedieu, G.; Mougenot, B.; Debaecker, V.; Duchemin, B.; Meygret, A. Correction of aerosol effects on multi-temporal images acquired with constant viewing angles: Application to Formosat-2 images. *Remote Sens. Environ.* **2008**, *112*, 1689–1701.

36. Hagolle, O.; Huc, M.; Pascual, D.V.; Dedieu, G. A multi-temporal method for cloud detection, applied to Formosat-2, Venμs, Landsat and Sentinel-2 images. *Remote Sens. Environ.* **2010**, *114*, 1747–1755.

37. Paris Anguela, T.; Zribi, M.; Baghdadi, N.; Loumagne, C. Analysis of local variation of soil surface parameters with TerraSAR-x radar data over bare agricultural fields. *IEEE Trans. Geosci. Remote Sens.* **2010**, *48*, 874–881.

38. Bartholomeus, H.; Epema, G.; Schaepman, M. Determining iron content in Mediterranean soils in partly vegetated areas, using spectral reflectance and imaging spectroscopy. *Int. J. Appl. Earth. Obs. Geoinform.* **2007**, *9*, 194–203.

39. Aase, J.; Tanaka, D. Reflectances from four wheat residue cover densities as influenced by three soil backgrounds. *Agron. J.* **1991**, *83*, 753–757.

40. Baird, F.; Baret, F. Crop residue estimation using multiband reflectance. *Remote Sens. Environ.* **1997**, *59*, 530–536.

41. Nalger, P.; Daughtry, C.S.T.; Goward, S. Plant litter and soil reflectances *Remote Sens. Environ.* **2001**, *71*, 207–215.

42. Streck, N.A.; Rundquist, D.; Connot, J. Estimating residual wheat dry matter from remote sensing measurements. *Photogramm. Eng. Remote Sens.* **2002**, *68*, 1193–1201.

43. Goovaerts, P. *Geostatistics for Natural Resources Evaluation*; Oxford University Press: New York, NY, USA, 1997; p. 483.

44. Wackernagel, H. *Multivariate Geostatistics*; Spinger-Verlag: London, UK, 1995.

45. Williams, P.C.; Norris, K. Variables affecting near-infrared reflectance spectroscopy analysis. In *Near-Infrared Technology in the Agricultural and Food Industries*; American Association of Cereal Chemists Inc.: Saint Paul, MN, USA, 1987; pp. 143–166.

46. Belkhodja, K. *Etude Pédologique des Zones de Draa Affane nord Merguellil et Chebika Mellouche (Kairouan)*; Direction des ressources en eau et en sol: Division des sols, Tunisie, 1975; p. 16.

47. 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.