University of São Paulo
"Luiz de Queiroz" College of Agriculture

Land surface temperature and reflectance spectra integration obtained from Landsat on the soil attributes quantification

Veridiana Maria Sayão

Dissertation presented to obtain the degree of Master in Science. Area: Soil and Plant Nutrition

Piracicaba
2017

Veridiana Maria Sayão
Forest Engineer

Land surface temperature and reflectance spectra integration obtained from Landsat on the soil attributes quantification

Advisor:
Prof. Dr. **JOSÉ ALEXANDRE MELO DEMATTÊ**

Dissertation presented to obtain the degree of Master in Science. Area: Soil and Plant Nutrition

Piracicaba
2017

Dados Internacionais de Catalogação na Publicação
DIVISÃO DE BIBLIOTECA – DIBD/ESALQ/USP

Sayão, Veridiana Maria

Land surface temperature and reflectance spectra integration obtained from Landsat on the soil attributes quantification / Veridiana Maria Sayão. -- Piracicaba, 2017.

80 p.

Dissertação (Mestrado) - - USP / Escola Superior de Agricultura “Luiz de Queiroz”.

1. Temperatura de superfície terrestre 2. Mapeamento de atributos do solo 3. Sensoriamento remoto no infravermelho termal 4. Análise de imagens de satélite I. Título

I dedicate to my parents, Wagner and Valéria, to my grandparents, Bento and Áurea, and to my beloved Luis.

ACKNOWLEDGMENTS

First of all, I thank God for the blessings, the amazing people I had the chance to know during this period of my life and for giving me strength to carry on in the toughest moments.

My warmest thanks go to my family for all the love, support, trust and encouragement, and also to my beloved fiancé Luis, who always motivated me and helped me to face all challenges of becoming a Master of Science. Thank you for turning any regular day into the best I could have.

I would like to thank the “Luiz de Queiroz” College of Agriculture (ESALQ), within the University of São Paulo, for the granted opportunity of study. I am also very thankful to the Department of Soil Science, all its employees and the Soil and Plant Nutrition graduate program, for the infrastructure and assistance provided whenever it was needed.

Many thanks to the National Council for Scientific and Technological Development (CNPq) for the granted scholarship and for the research incentive in Brazil. I also thank the São Paulo Research Foundation (FAPESP) for all resources provided to the Remote Sensing laboratory applied to soil studies.

I thank my supervisor Alexandre Demattê for the instructions, knowledge, ideas and support, crucial for the development of this research.

I also thank all the Professors I had classes with, for the valuable time of learning I had with them.

Special thanks go to dear Professor Igo Lepsch for his immeasurable assistance and for sharing his knowledge.

I am deeply thankful to Fernanda Bacellar for her amazing work in the Center for Language Studies at ESALQ. Thank you for the opportunity of improving my English skills for a lifetime and above all, thank you for the friendship and support. Thanks to you and your crew Friday afternoons were more special than what was expected.

Thanks to my dear friends Clécia and Karina my workplace was such an enjoyable place to be. Thank you for helping me through all the way! I am also very thankful to my longtime friends Carla, Aline, Grace and Ágata, for the amazing time I had with you.

I am very grateful to all the current members of the Geotechnologies in Soil Science Group (GeoSS) – Vitor, Raul, Diego, Wanderson, Lucas, Caio, André, Ariane, Fellipe, Vinicio, Luiz Gonzaga, Arnaldo, Julia, Alisson, Natasha and Emmily – and to those who were part of the group before – Danilo, Bruna, Marcos, João Paulo and Yaser. Thank you all for the support in technical issues and for the friendship.

I would like to thank also Paulo Tavares, for the help in Geostatistics.

For all who directly or indirectly contributed to the development of my research, thank you very much.

EPIGRAPH

“The world breaks everyone and afterwards many are strong at the broken places.”
Ernest Hemingway

SUMMARY

RESUMO

Integração da temperatura de superfície terrestre e de espectros de reflectância obtidos do Landsat na quantificação de atributos do solo

Os atributos do solo influenciam diretamente na sua temperatura de superfície. Apesar de existir vários estudos utilizando espectros de solos obtidos de satélite, a avaliação do solo por meio da Temperatura de Superfície Terrestre (em inglês Land Surface Temperature, LST) ainda é escassa. A ampla disponibilidade de dados termais de satélite e o desenvolvimento de algoritmos para derivar a LST facilitou o seu uso em estudos de solos. O objetivo desse trabalho foi avaliar variações da LST do solo devidas à sua composição e verificar o potencial de uso da LST na quantificação de atributos do solo, também integrada com dados de espectros de reflectância e elevação. A área de estudo (198 ha) está localizada no estado de São Paulo, Brasil, e estava com solo exposto e arado na data de aquisição da imagem de satélite. Amostras de solo foram coletadas em um grid regular de 100 x 100 m (profundidades: 0.02 m e 0.8-1.0 m); a granulometria do solo, matéria orgânica (MO) e óxidos de ferro foram determinados via análises físicas e químicas laboratoriais. Neste estudo, uma imagem do Landsat 5 foi utilizada para extrair a temperatura de superfície usando a inversão da função da Lei de Planck na banda 6 (10.400 – 12.500 nm), e a emissividade de superfície foi estimada utilizando o método do limiar do Índice de Vegetação da Diferença Normalizada. Valores de reflectância das bandas 1, 2, 3, 4, 5 e 7 foram extraídos. Modelos para quantificação de atributos do solo foram feitos usando Regressão Linear (RL), com amostras de 62 pontos de tradagem distribuídos em 14 topossequências. A RL simples foi aplicada para gerar modelos de predição baseados na LST e também na elevação (extraída de um modelo digital de elevação). A RL múltipla foi aplicada para gerar modelos de predição usando os espectros de reflectância com correção atmosférica das bandas do Visível, Infravermelho próximo e Infravermelho de ondas curtas (Vis-NIR-SWIR) como preditores; também foi aplicada para predição de atributos do solo usando simultaneamente dados do Vis-NIR-SWIR, LST e elevação, e apenas variáveis significativas identificadas por teste T foram usadas. A performance preditiva dos modelos foi avaliada baseada no coeficiente de determinação ajustado (R^2_{adj}), raiz do erro quadrático médio (RMSE, g kg^{-1}) e razão de desempenho do intervalo interquartil (RPIQ) obtidos na validação. A krigagem ordinária também foi feita e as superfícies interpoladas resultantes foram comparadas com o melhor modelo de RL. Houve correlação significativa entre os atributos do solo e dados de reflectância, LST e elevação, e solos com textura argilosa foram diferenciados de solos arenosos com base em valores médios de LST. Para todos os atributos do solo, os modelos usando apenas elevação apresentaram a pior performance, modelos usando somente LST, performance moderada, e usando as bandas do Vis-NIR-SWIR, boa performance preditiva. Para argila, o melhor modelo obtido teve as bandas 4-7, LST e elevação como preditores; para areia e óxidos de ferro,

o melhor modelo teve as bandas 4-7 e LST; para MO, banda 4, banda 7 e LST. O uso da LST para estimar atributos do solo aumenta a performance preditiva de modelos de RL múltipla quando associada a outras variáveis obtidas via sensoriamento remoto (SR), particularmente dados de reflectância de superfície, melhorando a validação dos modelos atingindo altos valores de R^2_{adj} e RPIQ e baixos valores de RMSE. Os mapas para areia, MO e óxidos de ferro obtidos via krigagem ordinária superaram aqueles obtidos para os mesmos atributos usando modelos de RL baseados em co-variáveis obtidas via SR, e para argila, ambas abordagens atingiram o mesmo nível de acurácia. O mapeamento dos conteúdos de argila, areia, matéria orgânica e óxidos de ferro do solo via modelos de RL múltipla utilizando produtos do Landsat 5 é uma técnica simples e fácil de reproduzir, apropriada para o mapeamento de atributos do solo em áreas de agricultura com solo exposto.

Palavras-chave: Temperatura de superfície terrestre; Mapeamento de atributos do solo; Sensoriamento remoto no infravermelho termal; Análise de imagens de satélite

ABSTRACT

Land surface temperature and reflectance spectra integration obtained from Landsat on the soil attributes quantification

Soil attributes directly influence on its surface temperature. Although there are several studies using soil spectra obtained from satellites, soil evaluation through Land Surface Temperature (LST) is still scarce. The broad availability of satellite thermal data and the development of algorithms to retrieve LST facilitated its use in soil studies. The objective of this study was to evaluate soil LST variations due to its composition and verify the potential of using LST on soil attributes quantification, also integrated with reflectance spectra and elevation data. The study area (198 ha) is located in Sao Paulo state, Brazil, and had plowed bare soil during the satellite image acquisition date. Soil samples were collected in a regular grid of 100 x 100 m (depths: 0-0.2 m and 0.8-1.0 m); soil granulometry, organic matter (OM) and iron oxides were determined by wet chemistry analysis. In this study, an image of Landsat 5 was used for extracting LST using the inversion of Planck's function in band 6 (10,400 – 12,500 nm), and land surface emissivity was estimated using Normalized Difference Vegetation Index threshold method. Reflectance values were extracted from bands 1, 2, 3, 4, 5 and 7. Models for soil attributes quantification were performed using Linear Regression (LR), with samples from 62 auger points distributed in 14 toposequences. Simple LR was applied for generating prediction models based on LST and on elevation data (extracted from a Digital Elevation Model). Multiple LR was applied in order to generate prediction models using atmospherically corrected spectral reflectance from Visible, Near-Infrared and Shortwave infrared (Vis-NIR-SWIR) bands as predictors, and also for the prediction of soil attributes using simultaneously Vis-NIR-SWIR, LST and elevation data, and only significant variables identified by T-tests were used. Predictive performance of models was assessed based on adjusted coefficient of determination (R^2_{adj}), Root Mean Squared Error (RMSE, g kg^{-1}) and Ratio of Performance to Interquartile Range (RPIQ) obtained in validation. Ordinary kriging was also performed and the resulted interpolated surfaces were compared to the maps obtained from the best LR model. There was significant correlation between soil attributes and reflectance, LST and elevation data, and soils with clay texture were differentiated from sandy soils based on LST mean values. For all soil attributes, models using only elevation presented the worst performance; models using only LST, moderate performance; and using Vis-NIR-SWIR bands, good predictive performance. For clay, the best model obtained had bands 4-7, LST and elevation as predictors; for sand and iron oxides, the best model had bands 4-7 and LST; for OM, band 4, band 7 and LST. The use of LST for estimating soil attributes increases the predictive performance of multiple LR models when associated with other variables obtained through remote sensing, particularly surface reflectance data, improving the validation of models reaching high R^2_{adj}, high RPIQ and low RMSE values. Maps for sand, OM and iron oxides obtained through

ordinary kriging outperformed those obtained for the same attributes using LR models based on RS co-variables, and for clay, both approaches reached the same accuracy level. Mapping of soil clay, sand, OM and iron oxides contents through multiple LR models using Landsat 5 products is a simple and easy to reproduce technique, appropriate for soil attributes mapping in bare soil agricultural areas.

Keywords: Land surface temperature; Soil attributes mapping; Thermal infrared remote sensing; Satellite imagery analysis

LIST OF FIGURES

LIST OF TABLES

1. INTRODUCTION

Soil temperature is one of the most important factors for plant development, as it controls soil evaporation and aeration, besides the types and rates of chemical reactions occurring in the soil. Variations in soil temperature are related to changes in processes of energy exchange, which occur on soil surface. In order to understand the thermal regime in soils, it is essential to study thermal properties, such as specific heat capacity and the thermal conductivity. The former reflects the soil's capacity of acting as a heat reservoir while the last is related to the soil's capacity of transmitting heat (Prevedello, 2010; Hillel, 2004).

Processes of heat transfer in the soil occur by conduction and convection, and the surface temperature is actually a consequence of these processes and the heat exchange between the soil surface and the atmosphere, which can occur by radiation. In this last energy transfer process, the thermal energy is in the form of electromagnetic waves, while in the case of heat transfer by conduction, the energy exchange occurs in molecular and atomic levels, being the most important heat transfer process in dry soils (Prevedello, 2010).

The study of soil temperature based on Remote Sensing (RS) focuses on the electromagnetic radiation emitted by the soil surface primarily at a wavelength around 10,000 nm (Hillel, 2004). This wavelength is situated in the Thermal Infrared (TIR) domain, and sensors operating in this region capture this energy emittance, allowing the derivation of TIR RS products such as the Land Surface Temperature (LST) (Sabins, 1996; Kuenzer and Deck, 2013). LST data obtained from orbital sensors is useful for several environmental studies, including vegetation and fire monitoring, geological, sea and soil studies, among others (McMillin, 1975; Bonn and O'Neill, 1993; Kuenzer and Deck, 2013; Li et al., 2013).

Soil studies based on TIR RS are mainly directed to the estimation of soil moisture, as soil LST is highly influenced by its moisture content (Bonn and O'Neill, 1993). Water presents high thermal conductivity, which accelerates the transmission of heat along depth (Prevedello, 2010; Hillel, 2004). However, recent studies are revealing that it is possible to relate LST to soil attributes, such as texture (Osińka-Skotak, 2007; Wang et al., 2015a; Müller et al., 2016) and organic matter (OM) content (Zhao et al., 2014).

The estimation of soil attributes using proximal and orbital RS has been widely reported (Ben-Dor and Bannin, 1995; Viscarra Rossel et al., 2006; Shepherd and Walsh, 2002; Chen et al., 2008; Ben-Dor et al., 2009; Nanni et al., 2014), and its use in Digital Soil Mapping (DSM) associated with traditional soil identification has improved the availability of soil information at large scales. The spectral responses of soil in the regions of Visible (Vis), Near Infrared (NIR) and Shortwave Infrared (SWIR) have a strong relationship with soil attributes, such as clay, organic carbon and iron oxides (Chang et al., 2001) and their prediction through Vis-NIR-SWIR data is a consolidated technique. The use of remotely sensed Digital Elevation Models (DEM) also contributes to the prediction of soil attributes, considering that terrain features influence on the performance of models (McBratney et al., 2003; Moura-Bueno et al., 2016). However, the inclusion of TIR data in the form of LST for the prediction of soil attributes is a new approach, and very few works were performed in this research line.

There are differences regarding sand, clay and OM thermal properties, therefore LST differentiation can be enhanced in soils varying in texture and OM content. Osińka-Skotak (2007) found that topsoil texture has an important impact on LST, and the same type of soil can reach differences of up to 4°C in brightness temperatures, depending on its topsoil texture. These differences in thermal properties are closely related to the water content of soils, because thermal properties of water strongly influence on soil LST. Thus, the behavior of LST in response to soil moisture provides information regarding soil texture (Chang and Islam, 2000; Chang et al., 2003; Wang et al., 2012; Wang et al., 2015a; Müller et al., 2016).

In the case of soil OM, its presence in the topsoil is linked with clay content. Mechanisms of physical and chemical protection of soil OM from microbial mineralization take place in the soil due to the presence of clay particles (Hook and Burke, 2000; Konen et al., 2003). Therefore, soils with high clay content are also associated with high OM, and LST tends to follow a similar relationship with both (Zhao et al., 2014).

One aspect that can also be related to LST is soil clay mineralogy, but we do not found studies performed in this research line. Attention has been caught to the study of thermal properties of rocks and their respective minerals (Robertson, 1988;

Eppelbaum et al., 2014). Soil clay minerals such as iron oxides may contribute to the LST obtained from clayey soils.

Through soil texture, it is possible to infer soil hydraulic properties (Müller et al., 2016), such as the water holding capacity and the susceptibility to erosion (Wakindiki and Ben-Hur, 2002). In fact, knowledge of soil texture is crucial for proper soil management. The SOM also plays an important role in soil management, and its importance is linked with microbial activity, soil fertility, water-stable aggregation (Tisdall and Oades, 1982; Six et al., 2004), enhanced water holding capacity (Hudson, 1994) and carbon sequestration (Post and Kwon, 2000). The use of RS techniques including TIR products has the potential to improve soil attributes mapping at the farm scale, providing up-to-date reliable thematic maps of texture and OM.

Given this context, the hypotheses guiding this research are that (1) bare soil surface temperature differs in terms of soil texture, OM content and clay mineralogy and (2) LST estimated by RS allows the prediction of soil clay, sand, OM and iron oxides contents. Thus, the objectives of this research were to (1) estimate the soil surface temperature by orbital RS; (2) differentiate soil textural classes assessing LST values; (3) generate soil attributes prediction models based on different RS variables alone such as LST, Vis-NIR-SWIR reflectance data and elevation obtained from a DEM and associated with each other; (4) verify to what extent the LST can improve soil attributes mapping and (5) generate soil attributes maps based only on wet chemistry analysis using geostatitistics and compare them with the maps produced using RS variables.

Brief theoretical background on TIR RS

The term remote sensing (RS) is used in reference to methods that employ electromagnetic energy, such as light, in order to detect and measure target characteristics. One of its basic fundamentals are the interaction processes between electromagnetic radiation and matter. Once the electromagnetic energy reaches an object, it can be transmitted, absorbed, emitted, scattered or reflected. These interactions, determined by the physical properties of the matter and the energy wavelength, are registered on RS images, from which it is possible to interpret features of matter. In TIR RS, the emittance (or emission) is studied, which basically

represents the energy previously absorbed by an object that is converted to heat and released in longer wavelengths (Sabins, 1996).

All objects with temperature above absolute zero (0 K, which is equivalent to 273 ℃) emit electromagnetic radiation. The average Earth's temperature is approximately 300 k and its electromagnetic emittance peak is located in the TIR region, around 9,700 nm (Sabins, 1996). The Earth absorbs a great portion of the solar radiation and a corresponding amount is emitted in longer wavelengths. Sensors operating in the TIR domain are able to capture this radiation, allowing the derivation of thermal radiance images from objects situated on the Earth surface. These images can present the radiant temperature of targets, in the resolution of the sensor used. One of the most used products derived from TIR images is the LST (Kuenzer and Dech, 2013), providing temperature data at a global scale for applications in studies from several research areas.

There is not a consensus in the scientific community regarding the definition of the TIR domain. The definition of Sabins (1996) was adopted here, which states that TIR domain is situated between 3,000 and 14,000 nm. Atmospheric windows located in the ranges of 3,000-5,000 nm and 8,000-14,000 nm allow the thermal mapping of Earth's surface. Therefore, in the shorter range of 3,000-5,000 nm the thermal signal can be contaminated by the reflected sunlight, which needs to be corrected when using RS products from such wavelengths. While TIR RS records emitted radiation, multispectral RS in the Vis-NIR-SWIR domains records reflected radiation (Kuenzer and Dech, 2013).

In order to estimate LST from thermal images, the behavior of emitted thermal energy is studied through important laws of physics. The Planck's Law describes the amount of radiation emitted by a blackbody in thermal equilibrium at wavelength λ and temperature T (equation 1) (Li et al., 2013). The blackbody is a hypothetical ideal radiator: considering all energy incident upon it, it is completely absorbed and re-emitted (Kuenzer and Dech, 2013).

$$B_\lambda(T) = \frac{C_1}{\lambda^5 \left[\exp\left(\frac{C_2}{\lambda T}\right) - 1\right]} \qquad (eq. \quad 1)$$

Where: $B_\lambda(T)$ is the spectral radiance (W m^{-2} μm^{-1} sr^{-1}) of a blackbody at temperature T (K) and wavelength λ (μm); C_1 and C_2 are physical constants (C_1 = 1.191 × 10^8 W μm^4 sr^{-1} m^{-2}, C_2 = 1.439 × 10^4 μm K).

Considering that terrestrial targets usually do not present blackbody properties, it is necessary to account for the emissivity (ε) in order to correct these radiance differences, which is the ratio between the radiance of a target and that of a blackbody at the same temperature and in the same wavelength (Li et al., 2013; Ndossi and Avdan, 2016). If the atmosphere has no influence on the measured radiance, T can be obtained by putting temperature as the subject of equation 1, having the values of emitted radiance and emissivity known (Li et al., 2013).

Satellites with TIR sensors measure the radiation emitted from the Earth and its atmosphere along the line of sight. The Radiative Transfer Equation (RTE) is applied for calculating the channel infrared radiance received by a sensor at the top of the atmosphere (TOA), assuming there are no clouds in the atmosphere and it is under local thermodynamic equilibrium. The RTE considers basically three terms as described in details by Li et al. (2013): the surface outgoing radiation attenuated by the atmosphere; the atmospheric emission and the atmospheric scattering.

The physical interpretation of LST values may lead to a few doubts. This temperature is called radiometric temperature, as the surface temperature is estimated from the radiance emitted by a surface. In fact, the radiometric temperature is derived from the emitted radiation from depths usually smaller than that of penetration considering the analyzed wavelength, which corresponds to a few millimeters. This radiometric temperature physically does not mean the same as the thermodynamic temperature measured by a thermometer for targets in thermal equilibrium. In the case of homogeneous and isothermal surfaces, these two types of temperatures are considered equivalent. Indeed, there are difficulties associated with thermodynamic temperature measurements, and the radiometric temperature is obtained in a more practical way (Prata et al., 1995; Becker and Li, 1995; Wan, 1999; Li et. al, 2013).

2. MATERIALS AND METHODS

2.1. Study Area

The research area (Figure 1) is located in the municipality of Rafard, southeast of São Paulo State, in a Paleozoic depression. Previous graduate studies were performed by Nanni (2000) and Bazaglia Filho (2012) in this same location. The total area is within the Tietê river watershed and has approximately 198 ha, with subtropical mesotermic climate (Cwa) according to Köppen classification, with dry winters and wet summers. The average temperature in the coldest month, July, is 18℃ and 22℃ in the warmest one, February. Annual rainfall is between 1100 and 1700 mm. The land use has been sugar cane cultivation for many years.

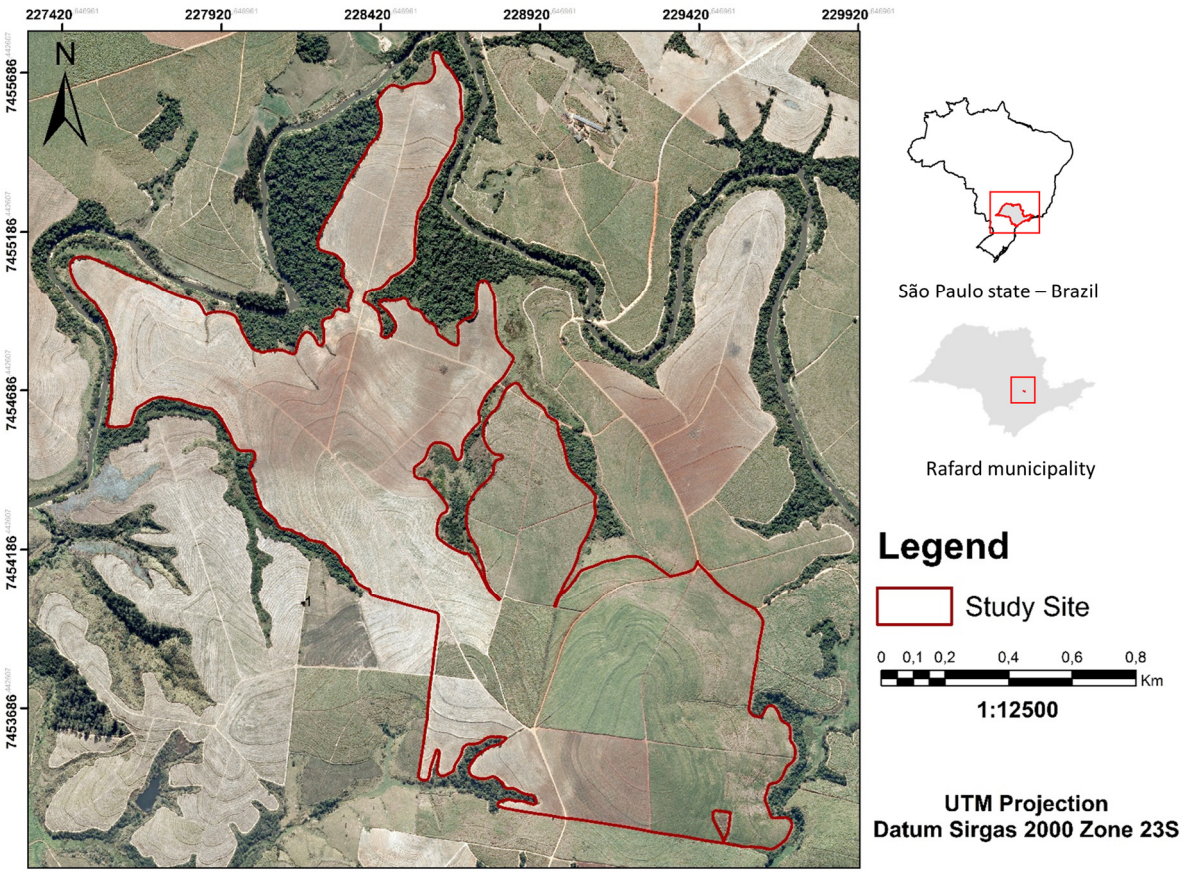

Figure 1. Study site location. Aerial photograph (2011) from Paulista Company of Metropolitan Planning S.A. (EMPLASA).

The Itararé Formation (Tubarão group) represents the geology of the area (Figure 2a) (Mezzalira, 1966), with siltstone as the predominant lithology. In addition, there are also eruptive diabase dike elements of the Serra Geral Formation (São Bento group) and fluvial old terrace sediments are found near the Capivari river (Nanni and Demattê, 2006). Altitude varies from 478 to 570 meters and the relief is rolling to gently rolling, with slope varying between 0 and 35%. Given the geology complexity and diversity, soils that occur in the study site are very diverse, comprising five groups from the nomenclature of World reference base for soil resources (IUSS Working Group WRB, 2015). Soils found in the area are classified as Lixisols, Nitisols, Cambisols, Leptosols, Gleysols and Chernozems (Figure 2b) (Bazaglia Filho, 2012).

2.2. Soil sampling and wet chemistry analyses

A regular sampling grid with 100 x 100 m² was established in the study site, having a sampling density of one sample per hectare (encompassing 182 ha from the total area, 198 ha), comprising auger points at two depths: 0-0.2 m for surface layer (182 points) and 0.8-1.0 m for subsurface layer (179 samples). For this study, only surface samples were used. The soil samples were oven-dried for 48 hours at 50°C, ground and sieved (2 mm mesh). Analysis of soil particle size distribution was performed using the densimeter method, in which sodium hydroxide (0.1 mol L^{-1}) and sodium hexametaphosphate (0.1 mol L^{-1}) were employed as dispersing agents (Camargo et al., 1986). For chemical analysis, soil OM content was determined based on the Walkley-Black method (Walkley and Black, 1934). This method determines the total organic C and the calculation of soil OM was made using a conversion factor of 1.724; total clay iron oxides (Fe_2O_3) was determined with a sulfuric acid based methodology (Camargo et al., 1986).

Percentages of soil particles sand, silt and clay were used in order to obtain the soil texture class, using the system proposed by the United States Department of Agriculture (USDA), based on the soil texture triangle. Soil textures were calculated using the *soiltexture* package (Moeys, 2016), in R environment (R Core Team, 2015).

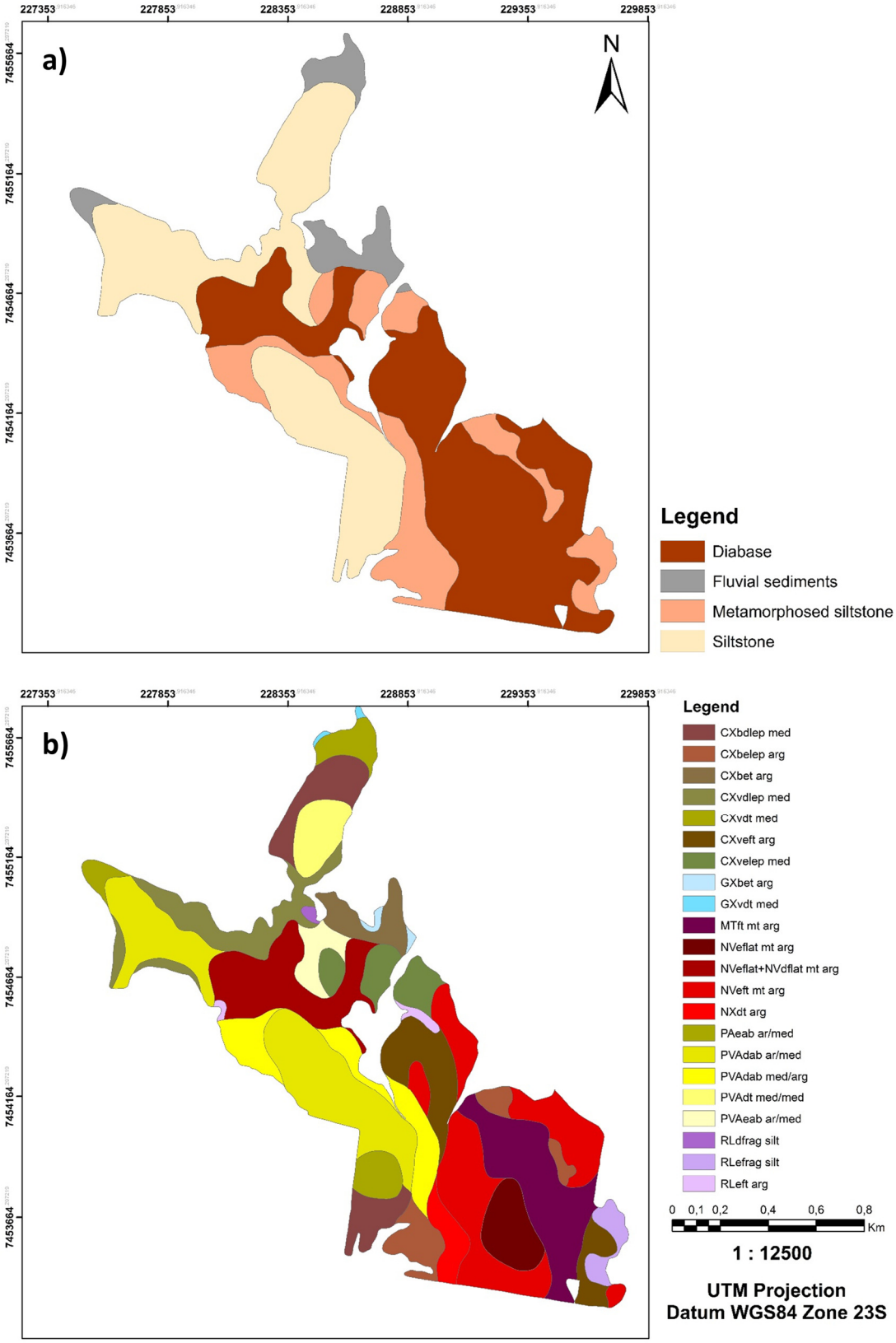

Figure 2. a) Geology of the study area; **b)** Soils from the study site (Bazaglia Filho, 2012). Nomenclature: CX = Cambisol; GX = Gleysol; MT = Chernosol; NV/NX = Nitisol; PA/PVA = Lixisol; RL = Leptosol; med = medium texture; arg = clayey texture; mt arg = very clayey texture; ar = sandy texture; silt = silty texture.

2.3. Remote sensing (RS) data

2.3.1. Data acquisition

Images from the satellite Landsat 5 can be obtained through the earth Explorer Platform, provided by the United States Geological Service (USGS). Landsat 5 has an imaging system called thematic mapper (TM) and the multispectral scanner (MSS). It has a ground resolution cell of 30 m in the regions of Vis and reflected infrared (NIR-SWIR) and 120 m in the TIR. However, band 6 is resampled and is acquired with a spatial resolution of 30 m as well (USGS, 2016a).

Landsat 5 has seven bands comprising the following wavelengths (nm): 450 to 520 (band 1 – B1); 520 to 600 (band 2 – B2); 630 to 690 (band 3 – B3); 760 to 900 (band 4 – B4); 1,550 to 1,750 (band 5 – B5); 10,400 to 12,500 (band 6 – B6) and 2,080 to 2,350 (band 7 – B7) (Sabins, 1996). An image from the year 1997, month August, day 27 was chosen, due to the availability of bare soil in almost all the study area, as reported by Nanni (2000). Besides this, August corresponds to the dry season, period in which there is low or null precipitation and, consequently, less moisture in the soil. This is important because water strongly influences in the reflectance and LST of soils.

Besides the Landsat image, the product SRTM (Shuttle Radar Topography Mission) was acquired in the same platform. The SRTM is a DEM of the Earth, has a ground resolution cell of 30 m and represents elevation data in meters, being useful in order to account for the influence of the landscape in the prediction of soil attributes. From the DEM it is possible to derive important information regarding the land surface, such as slope, aspect (slope orientation) and solar radiation. These products were obtained in ArcMap 10.3 (ESRI, 2011). Regarding the solar radiation, it was calculated for the same date and time of the selected Landsat 5 image, according to the algorithm developed by Fu and Rich (2000).

2.3.2. Atmospheric correction and bare soil mask

Atmospheric correction and the creation of a bare soil mask were performed in ENVI 5.1 software (Exelis Visual Information Solutions, Boulder, Colorado). For the atmospheric correction, the algorithm FLAASH (Line-of-sight Atmospheric Analysis of

Spectral Hypercubes) was used, which handles data from a variety of sensors, including Landsat 5 TM/MSS. This algorithm integrates the MODTRAN (Moderate Resolution Atmospheric Transmission) radiation transfer code and accounts for water vapor and aerosol retrieval (Cooley et al., 2002). Before applying this algorithm, digital numbers are converted to spectral radiance and only after going through radiometric calibration, image data from Vis-NIR-SWIR is converted to surface reflectance.

The bare soil mask was created in order to eliminate targets in the image that do not correspond to bare soils, such as water, clouds, vegetation, urban areas and residues from agriculture (straw). The methodology was adapted from Fongaro (2015) and Gallo (2015). For eliminating water and clouds, masks built by USGS acquired in ESPA (Science Processing Architecture) platform were used; for urban areas, polygons with the delineation of cities developed by the Brazilian Institute of Geography and Statistics (IBGE) were used.

In the case of vegetation and straw, spectral indexes were used in order to differentiate these targets from bare soil. The Normalized Difference Vegetation Index (NDVI) (Rouse et al., 1974) is one of the most used indexes to identify the presence of vegetation in multispectral images. Using a NDVI classification tool available in ENVI 5.1, it was possible to differentiate vegetation from all the other targets in the image. The threshold NDVI value that separates vegetation (dense, moderate and sparse) from all other elements was 0.31. NDVI is calculated according to equation 2, and in the case of Landsat 5, B4 represents reflectance values on near infrared band and B3, on red band. The index varies from -1 to 1.

$$\mathrm{NDVI} = \frac{\mathrm{NIR} - \mathrm{R}}{\mathrm{NIR} + \mathrm{R}} \qquad (\mathrm{eq.}\ \ 2)$$

Where: NIR – reflectance value on near infrared band (B4); R – reflectance value on red band (B3).

For the removal of straw, the MID Infrared Index (MIDII) developed by Madeira (1993) was applied, which uses reflectance values from B5 and B7. It also varies from -1 to 1, and a visual analysis of the spectral signature of straw pixels in the image was made so that it was possible to verify the threshold value of MIDII that

separates soil from straw. The interval within bare soil pixels were identified as -0.06 to 0.15. This index is calculated according to equation (3).

$$\mathrm{MIDII} = \frac{\mathrm{SWIR1} - \mathrm{SWIR2}}{\mathrm{SWIR1} + \mathrm{SWIR2}} \qquad (\mathrm{eq.}\ 3)$$

Where: SWIR1– reflectance value on shortwave infrared B5; SWIR2 – reflectance value on shortwave infrared B7.

The application of the built masks allowed the distinction of all elements from an image that do not represent soils. The creation of a mask enables to clip a classification raster from the total study area. The mask was built to contrast only soil elements, and its application eliminates all other areas in an image, which have a null value after the mask is applied.

2.3.3. Land Surface Temperature (LST) estimation

LST can be determined by several methods and/or algorithms, which may consider the influence of atmospheric factors such as upwelling and downwelling radiance and atmospheric transmittance. The Mono-Window Algorithm (MWA) (Qin et al., 2001) and the Radiative Transfer Equation (RTE) (McMillin, 1975; Schmugge et al., 1998) consider such parameters in their equations and, consequently, promote an atmospheric attenuation. However, their estimation is often difficult. One known tool for the estimation of these parameters is the NASA (National Aeronautics and Space Administration) Atmospheric Correction Parameter Calculator (Barsi et al., 2003), but unfortunately, atmospheric data prior to 1998 is not available.

LST was estimated using the inversion of Planck's Function in B6, with a correction for land surface emissivity. This step was executed in QGIS (2016), using a Python QGIS plugin (Land Surface Temperature Estimation Plugin) (Ndossi and Avdan, 2016). After the conversion of raw digital numbers from B6 to TOA radiance, the Brightness Temperature (BT) was calculated. BT represents the temperature required by a blackbody so that it has the same radiation emittance per unit of surface area as the studied body (Kruse et al., 1962). Computation of brightness

temperature is performed following equation 4, using band specific thermal constants and the TOA radiance (USGS, 2016b).

$$BT = \frac{K_2}{\ln\left(\frac{K_1}{L_\lambda} + 1\right)} \qquad (eq.\ 4)$$

Where: BT – Brightness Temperature (K); K_1 (774.8853 W m^{-2} srad^{-1} μm^{-1}) and K_2 (1321.0789 K) are band specific thermal constants obtained from the Landsat metadata file; L_λ (W m^{-2} srad^{-1} μm^{-1}) – TOA spectral radiance.

2.3.4. Land Surface Emissivity

The determination of Land Surface Emissivity (LSE) was performed using the NDVI threshold estimation algorithm, developed by Sobrino and Raissouni (2000). This algorithm differentiates bare soil from vegetation pixels that can be separated in three classes: when NDVI < $NDVI_{soil}$, the pixel is composed of bare soil or rock and a LSE value of soil is attributed to such pixel; when NDVI > $NDVI_{soil}$, the pixel is composed of vegetation only, having a LSE value typical of vegetation; when $NDVI_{soil}$ ≤ NDVI ≤ $NDVI_{vegetation}$, the pixel is a mixture of vegetation and soil/rocks and is necessary to consider the proportion of vegetation in an image in order to calculate the LSE. The authors proposed NDVI values of 0.2 and 0.5 as thresholds for soil and vegetation, respectively. For the calculation of the emissivity value of mixed areas, equation 5 (Sobrino and Raissouni, 2000) was used.

$$\varepsilon_{\lambda=} \varepsilon_{v\lambda} P_v + \varepsilon_{s\lambda}(1 - P_v) + C_\lambda \qquad (eq.\ 5)$$

Where: ε_v – vegetation emissivity; ε_s – soil emissivity; P_v – proportion of vegetation, calculated through equation 6 (Carlson and Ripley, 1997); C – effect of the geometry of a surface, C = 0 for flat surfaces, and for a mixed pixel and near-nadir view it can be calculated according to equation 7 (Sobrino and Raissouni, 2000). Values of emissivity of 0.966 and 0.973 were used for soil and vegetation, respectively (Wang et al., 2015b).

$$P_v = \left[\frac{NDVI - NDVI_{min}}{NDVI_{max} - NDVI_{min}}\right]^2 \qquad (eq.\ 6)$$

Where: $NDVI$ – Normalized Difference Vegetation Index; $NDVI_{min}$ – minimum NDVI value in the analyzed image; $NDVI_{max}$ – maximum NDVI value in the analyzed image.

$$C_\lambda = (1 - \varepsilon_{s\lambda})\varepsilon_{v\lambda} F'(1 - P_v) \qquad (eq.\ 7)$$

Where: ε_v – vegetation emissivity; ε_s – soil emissivity; F' –geometrical factor ranging from 0 to 1, a mean value of 0.55 was adopted (Sobrino et al., 1990); P_v – proportion of vegetation, calculated from equation 6.

2.3.5. Inversion of Planck's Function

The inversion of Planck's Function (PF) is the simplest method in order to estimate LST, and it uses the estimated LSE to correct the emissivity of targets in relation to a blackbody (equation 8). Ndossi and Avdan (2016) reported that this method produced accurate results of LST obtained from Landsat 5, as it not depends on estimated atmospheric parameters, which may be associated with errors. The result from equation 8 is in K, and it was converted to Celsius scale by subtracting 273.15.

$$T_s = \frac{BT}{\{1 + \left[\frac{\lambda * BT}{\rho}\right] * ln\varepsilon} \qquad (eq.\ 8)$$

Where: T_s – Land surface temperature (K); BT – at-sensor brightness temperature (K); λ – wavelength of the emitted radiance (10.8 µm for band 6); ρ – is a constant obtained from h * c/σ (h = Planck's constant, 6.626 * 10^{-34} J s; c = velocity of light, 2.998 * 10^8 m s^{-1}; σ = Boltzman constant, 1.38 * 10^{-23} J K^{-1}) equals to 1.438 * 10^{-2} mK; ε – land surface emissivity obtained through NDVI threshold method. Figure 3 illustrates all described steps since image acquisition until LST estimation.

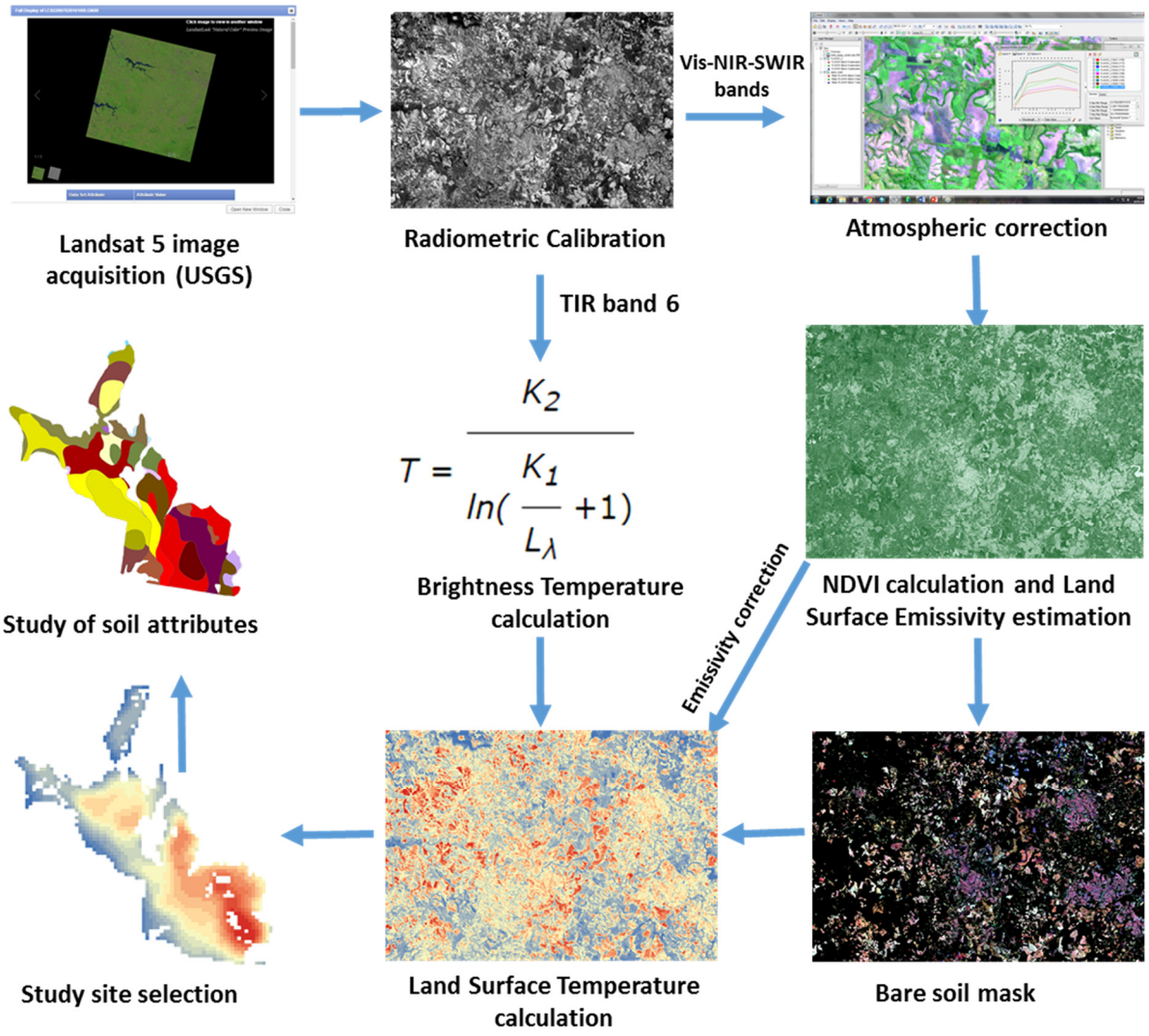

Figure 3. Flowchart with the methodology adopted for processing RS data

2.4. Statistical Analyses

2.4.1. Exploratory data analysis

Descriptive statistics such as mean, median, minimum, maximum, standard deviation (SD), coefficient of variation (CV), among others, were calculated for the soil attributes texture (clay, sand and silt), OM and iron oxides.

In order to verify how response and predictor variables are related, the Pearson correlation coefficient (r) was determined between and within them. Predictor variables include band's reflectance values, LST and elevation, while response variables include sand, silt, clay, OM and iron oxides (Fe_2O_3) from the topsoil (0-0.2m). In this step, 62 soil samples distributed in 14 toposequences (Figure

4a) were considered, as they were used for subsequent topics. The correlation analysis was performed in R environment (R Core Team, 2015), using the *corrplot* package (Wei & Simko, 2016).

2.4.2. Analysis of variance

One-way analysis of variance (ANOVA) was performed to verify if there are differences between LST mean vectors of each soil texture class, and the assumptions of normality of residuals and homogeneity of group variances were tested, using Shapiro-Wilk (Shapiro and Wilk, 1965) and Levene (Levene, 1960) tests, respectively. Tukey's honestly significant difference (HSD) post-hoc test for multiple comparisons of means (Tukey, 1949) was applied for identifying groups that actually differ from each other.

2.4.3. Linear regression models

Modeling of soil attributes (clay, sand, OM and Fe_2O_3) based on orbital RS data as predictors was performed using simple and multiple Linear Regressions (LR), fitted by ordinary least squares method. Models were calibrated using a subset of total samples, and their selection was based on the toposequence method. Using the DEM of the study site (Figure 4b), auger soil sampling points located in different positions of the relief were selected in such a way that it was possible to select soil samples in the main altitude variations, in all directions. Assuming that toposequences are sufficiently representative of the main soil features from the study site, a total of 70 auger points distributed in 14 toposequences were selected for the calibration of models. The remaining samples were used for the validation of models (Figure 4a).

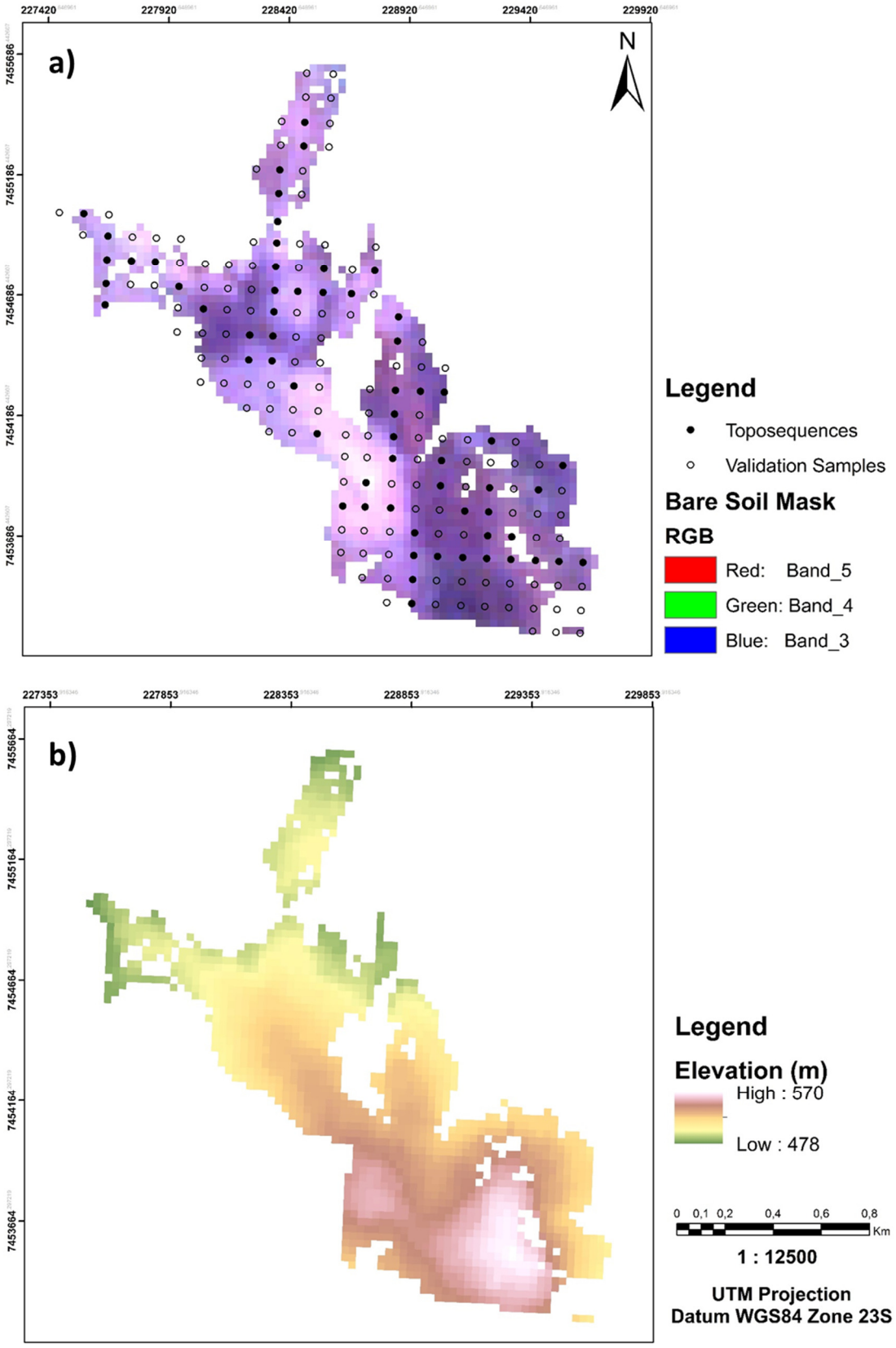

Figure 4. a) Samples used for calibration (toposequences) and validation of models over the bare soil area, obtained after the mask application described in section 2.1. Samples placed on blank areas in the map were not used; **b)** Elevation map obtained from SRTM digital elevation model.

Considering that after the application of the bare soil mask there were 152 auger points located in pixels corresponding to bare soil, only these points were used in statistical analyses, with 62 points in the toposequences and the remaining 90 samples in the validation subset. Both datasets present homogeneous variance, as identified by Levene's Test (Levene, 1960) (p-value > 0.05 for each soil attribute). Before running models, outliers were identified using Tukey's method (Tukey, 1977), based on the data distribution illustrated in boxplots, and eliminated in both calibration and validation sets.

The variables Vis-NIR-SWIR bands (B1-B5 and B7), LST and elevation were used as predictors in LR models. The approach adopted was to model first each soil attribute using the three predictors separately, applying simple LR in the case of LST and elevation and multiple LR (equation 9) in the case of Vis-NIR-SWIR bands. In this way, it was possible to assess how these different variables can model soil attributes.

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \ldots + \beta_p x_p + \varepsilon \quad (eq.\ 9)$$

Where: Y = response (dependent) variable to be modeled; x_1 through x_p are the predictor (independent) variables; β_1 through β_p are the estimated regression coefficients; β_0 = intercept term, which represents the value of Y when all independent variables are zero; ε = random error.

The second approach involved the selection of the best variables for prediction, considering the use of all predictor variables together. It was assumed that if more explanatory variables were used, the best predictive performance of models could be obtained. However, it is necessary to check if all considered predictors actually contribute in modeling the response variable.

In order to check for the importance of adding new predictors, backward selection was adopted, which consists in adding all potential predictors in multiple LR and verifying which of them are statistically significant to the model. Significant x_i variables with a β_i coefficient value statistically different from zero were selected, indicated by the p-value generated through t-statistic. The adopted $\alpha_{criteria}$ for

selection of variables was 0.1. In this matter, all variables with p-value above 0.1 were kept in the model.

The idea behind this methodology was to check the predictive performance of different explanatory variables when they are simultaneously used. In fact, it is expected that the larger the number of significant explanatory variables used, the more accurate the predictions would be. After identifying the significant predictors, a new multiple LR is performed and this procedure can be repeated until no more predictors can be withdrawn from the model (Faraway, 2015).

The assumptions underlying linear models such as normality of residuals, independency of residuals and constant variance (homoscedasticity) were all tested using Shapiro-Wilk (Shapiro and Wilk, 1965), Durbin-Watson (Durbin and Watson, 1950) and Breusch-Pagan tests (Breusch and Pagan, 1979), respectively. The p-value indicates the acceptance of the null hypothesis for each test (residuals normally distributed, independent and with constant variance, respectively), and the adopted significance level threshold was 0.05. In cases where the assumptions were violated in all proposed models, appropriate treatments were applied to the data so that it was possible to satisfy the necessary requirements for linear models.

In order to verify how well a model meets the purpose for which it is intended, a criterion for comparison between other models is necessary. The Root Mean-Squared Error (RMSE) (equation 10) and the adjusted R^2 (R^2_{adj}) (equation 11) obtained from calibration were used. The R^2_{adj} is recommended when comparing models with different number of predictors, as it makes a correction for the degrees of freedom and its value increases only when significant variables are included in the model (Faraway, 2015). For R^2_{adj}, high values indicate high explanatory effect of the predictors; for RMSE, low values indicate good performance of models.

In the case of backward selection, models were compared using the partial F-test, which compares the full model and the model based on a subset of variables from the former, testing if the removed coefficients are statistically equal to zero. All analyses described in this section were performed in R environment (R Core Team, 2015).

The validation was performed for all models of each soil attribute, and the figures of merit RMSE (equation 10), $\mathrm{R}^2_{\mathrm{adj}}$ (equation 11) (Williams, 1987) and Ratio of Performance to Interquartile Range (RPIQ) (Bellon-Maurel et al., 2010) (equation 13) were used. The RPIQ was used in substitution to the Ratio of Percent Deviation

(RPD), as Minasny and Mcbratney (2013) emphasized RPD is redundant in relation to R^2 (equation 12). Considering that there is not one classification system proposed for RPIQ assessment, the classification proposed for RPD by Viscarra-Rossel (2007) was adapted for RPIQ here, in which values lower than 1.5 are very bad; between 1.5 and 2.0 are poor; between 2.0 and 2.5, good; and higher than 2.5, very good (excellent predictions). The best model for each soil attribute was applied to produce maps.

$$RMSE = \sqrt{\frac{1}{N}\sum_{i=1}^{n}(\hat{y}_i - \bar{y}_i)^2} \quad (eq.\ 10)$$

$$R^2_{adj} = 1 - \left[\frac{(1-R^2)(N-1)}{N-K-1}\right] \quad (eq.\ 11)$$

$$R^2 = \frac{\sum_{i=1}^{n}(\hat{y}_i - \bar{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y}_i)^2} \quad (eq.\ 12)$$

$$RPIQ = \frac{IQ}{RMSE_{validation}} \quad (eq.\ 13)$$

Where: $\hat{y}_i$ – values estimated by the model; $\bar{y}_i$ – observed values; N – number of observations of the variable to be modeled; K = number of independent (predictive) variables; SD – Standard Deviation; IQ = inter-quartile range, representing the difference between quartile 3 (75% of values) and quartile 2 (25% of values).

In the case of soil texture, in order to visualize the results from all models, sand and clay predicted values were summed and the silt content was obtained by subtracting this sum from 1000 g kg^{-1}. Having these values for all available sample points, they were used for obtaining the soil texture classified by USDA texture triangle, for each model, using the package *soiltexture* (Moeys, 2016) in R environment (R Core Team, 2015).

In addition, an external validation was performed using a soil database encompassing agricultural areas outside the study site, in a radius of 40 km. This step aimed to verify if the best models generated for the small test site can perform well in areas spread outside the modeled site. The external validation was performed

only for clay, sand and OM, as there was not a databank with iron oxides data available. Figure 5 illustrates steps followed for modeling soil attributes.

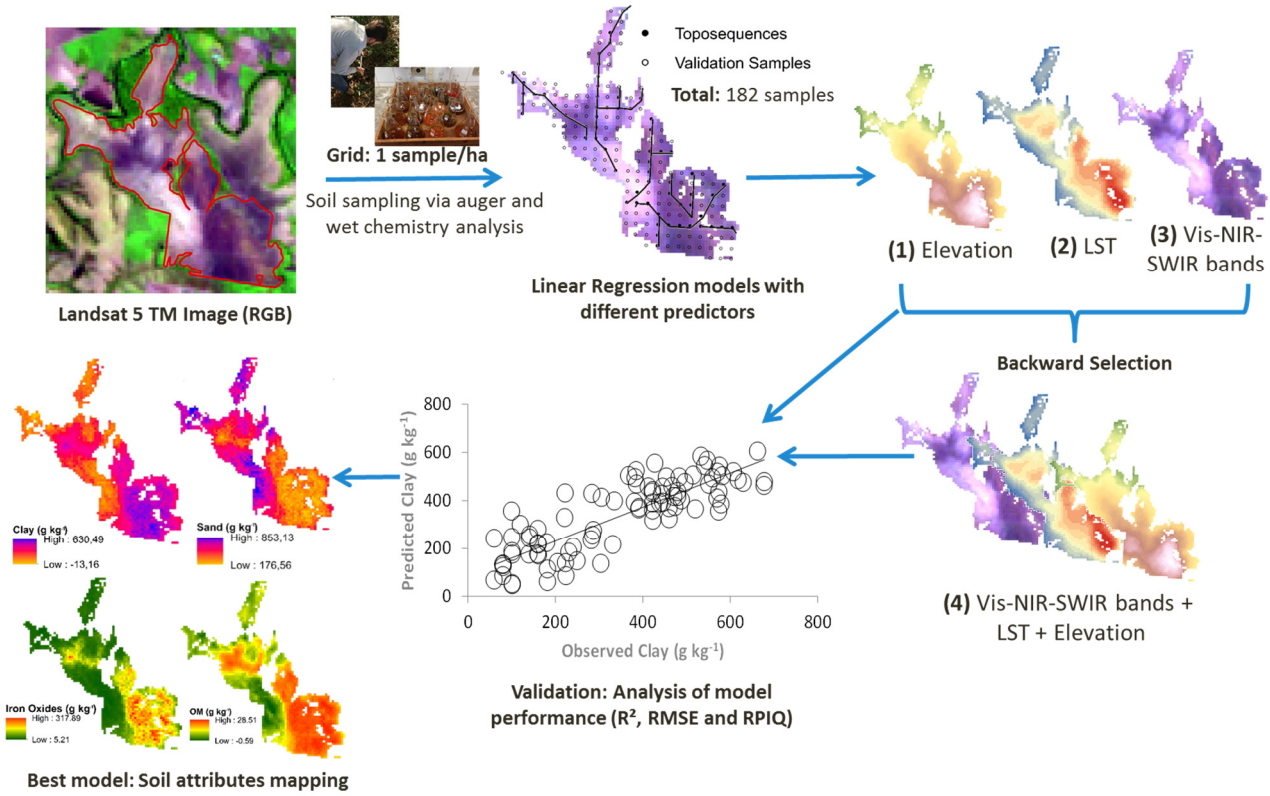

Figure 5. Flowchart with the methodology adopted in modeling of soil attributes.

2.4.4. Geostatistical Analysis

Soil attributes maps were performed using geostatistical methods, based only on soil data obtained through wet chemistry analysis. Soil attribute point values without outliers (previously identified in section 2.4.3) were spatialized using Ordinary Kriging. A semivariogram function was performed for verifying data spatial dependency and theoretical models were adjusted to the semivariograms, which were all isotropic. Exponential, spherical and Gaussian functions were tested for modeling of soil attributes spatial dependency. The adjust of models was compared using R^2 (equation 12) and RMSE (equation 10), and the model with best adjust was selected for interpolation. Data was interpolated using adjusted semivariogram parameters (Nugget (C_0), Sill ($C + C_0$) and Range (a)) and a leave-one-out cross-validation was performed to indicate the interpolation accuracy, from where RMSE (equation 10), R^2 (equation 12) and Average Standard Error (ASE, equation 14) were

calculated. These analyses were performed in ArcMap 10.3 (ESRI, 2011), using the Geostatistical Wizard.

$$\sqrt{\frac{\sum_{i=1}^{n} \hat{\sigma}^2(x_i)}{N}} \quad (eq.\ 14)$$

Where: x_i – observed values; $\hat{\sigma}^2$ – variance of the predicted values; N – number of observations.

The spatial dependence of soil attributes was assessed following the methodology of Cambardella et al. (1994) that classified spatial dependence using the nugget semivariance (C_0) in relation to the total semivariance ($C + C_0$). Values below 0.25 indicate strong spatial dependence; between 0.25 and 0.75, moderate and when above 0.75, weak spatial dependence. The degree of randomness of the data was assessed using the methodology of Guerra (1998), calculating the ratio (E) between the nugget semivariance (C_0) and the partial sill (C). When E is lower than 0.15, the random component is small, when E is between 0.15 and 0.30, is significant and when E is higher than 0.30, the random component is considered very significant.

Maps obtained through this step were adjusted to have a cell size of 30 m so that they could be compared to maps obtained from satellite RS data. This comparison was made by subtracting them, resulting in a map of residuals. By analyzing the latter, it was possible to identify contrasting areas, such as the ones that were over or underestimated in relation to each model. It is expected that the maps obtained from conventional analysis using geostatistical techniques represent well the distribution and variability of soil attributes and the closer they are from those obtained from RS data, the more reliable the last ones are considered.

3. RESULTS AND DISCUSSION

3.1. Soil attributes characteristics

Soil attributes have a wide variation in the study site, mainly due to their lithological diversity. For soil granulometry, the clay, sand and silt coefficient of variation (CV %) varies 56, 42 and 46 % around the mean, respectively. This aspect is important because it is linked with a great diversity of soil textural classes (Figure 6a). Silt content presents a high range, with a maximum value of 710 g kg^{-1} (Table 1). However, its mean is 194.63 g kg^{-1} and its median is 183.49 g kg^{-1}, showing that such high values are potential outliers and are causing both high positive kurtosis (9.46) and skewness (2.28). In the texture triangle (Figure 6a), there are only few samples with silt loam texture, which are apart from the main cloud of points, including mainly clay, clay loam, sandy clay loam, sandy loam and loamy sand textures. Despite the fact siltstone is the main soil parent material, soils with silt loam texture are not expressive and are represented by a very small area of Leptsols derived from a metamorphosed siltstone (Figure 2b).

Table 1. Descriptive statistics for soil attributes

	Fe_2O_3[1]	OM[1]	CLAY[1]	SAND[1]	SILT[1]
Mean	85.54	17.80	319.11	486.26	194.63
Standard Error	5.84	0.79	13.20	15.02	6.66
Median	48.50	15.00	286.16	440.97	183.49
Mode	8.00	10.00	182.19	739.22	101.21
Standard Deviation	78.75	10.70	178.01	202.69	89.80
CV (%)	92.05	60.12	55.79	41.68	46.14
Kurtosis	-1.20	0.63	-1.23	-1.24	9.46
Skewness	0.62	0.88	0.27	0.23	2.28
Range	242.00	52.00	617.91	738.57	669.76
Minimum	5.00	0.00	60.06	140.71	40.24
Maximum	247.00	52.00	677.97	879.28	710.00
Confidence Level (95.0%)	11.52	1.57	26.04	29.65	13.13

[1] g kg^{-1}

All soil attributes presented high standard deviations, which shows that their values expressively vary around the mean. This is crucial considering the first hypothesis of this research. Strong relationships between LST patterns and soils

determined by differences in soil attributes are expected to occur, as soil granulometry, clay mineralogy and organic matter content widely vary.

For OM, the CV (Table 1) was around 60%, indicating high variation as well. Very low and very high values occur, and both the mean (17.8 g kg^{-1}) and median (15 g kg^{-1}) indicate that very high values, such as those close to 50, are not likely to occur. Regarding skewness, the positive value (0.88) indicates that the variation in the right side of the distribution is higher and most of the values lie on the mean left side. The kurtosis value (0.63) is relatively low, which implies in a leptokurtic distribution, with high peaks in the data (Kim, 2013).

Lepsch et al. (1982) studied the relationship between OM content and soil texture in tropical soils with uniform conditions (such as good drainage and under the same climate domain, Cwa) cultivated with sugarcane in the São Paulo state and found that most soils present OM content between 0.6 and 3.8% (or from 6 to 38 g kg^{-1}), and the lowest values occur in soils with sandy texture, while the highest ones in soils with very clayey texture. For the dataset used here, there are few soil samples with OM content between 38 and 52 g kg^{-1}, which may be associated with hydric soils that enhance OM accumulation, as we have small areas of Gleysols near the riparian forests. In addition, the area with Chernozem (Figure 2b) also presents high values of OM (between 30 and 40 g kg^{-1}).

Regarding iron oxides, they presented the highest CV (92%) with a very high range; hipoferric (Fe_2O_3 < 80 g kg^{-1}), mesoferric (Fe_2O_3 between 80 and 180 g kg^{-1}) and ferric (Fe_2O_3 between 180 and 360 g kg^{-1}) soils occur in the study site. The mean value (85.5 g kg^{-1}) is not similar to the median (48.5 g kg^{-1}), suggesting that few very high values are influencing the mean. The negative value for kurtosis (-1.2) indicates data's platykurtic distribution, presenting a flat-topped curve (Kim, 2013). Occurrence of high values of iron oxides is common in highly weathered soils, particularly those with clayey texture. Parent material with high amounts of Fe bearing minerals also enhances the formation of iron oxides, such as the diabase occurring in a significant portion of the study site (Figure 2a).

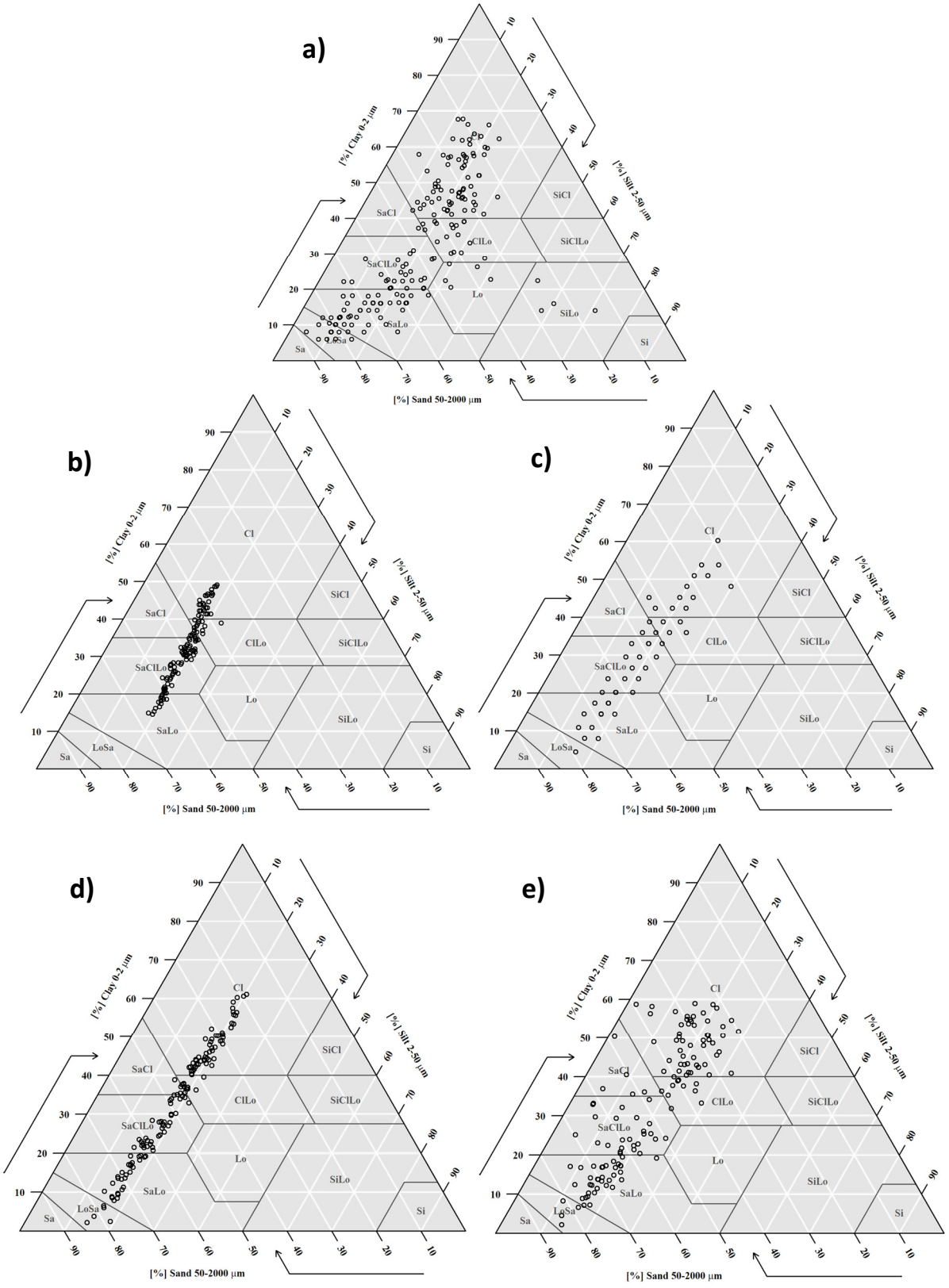

Figure 6. Texture triangles obtained from **a)** wet chemistry analysis; predictions of clay and sand contents from models using: **b)** elevation, **c)** LST, **d)** Vis-NIR-SWIR bands and **e)** bands 4-7 and LST for sand and bands 4-7, LST and elevation for clay as predictors.

3.2. Relationship between soil attributes and RS products: elevation, Landsat 5 Vis-NIR-SWIR bands and LST

Regarding the clay and sand soil attributes, they were negatively correlated (-0.93) (Figure 7), as with the increase in clay content, sand content tends do decrease. Sand presented significant negative correlation with silt (-0.49), indicating that soil samples with high silt content present low sand content. Conversely, clay and silt did not present statistically significant correlation, with clay content not being affected by the presence of silt fraction. OM content was positively correlated with clay and iron oxides (0.78 and 0.84, respectively), as OM is physically protected by clay particles, including oxides. Iron oxides (represented by Fe in figure 7) were positively correlated with clay and negatively correlated with sand, as clayey soils present high content of iron oxides, which is not likely to occur in sandy soils.

Elevation (Figure 4b) was positively correlated with clay, OM and iron oxides and negatively correlated with sand, as clayey soils occur in the highest portion of the study site and contents of OM and iron oxides both increase with clay content. Elevation varies up to 100 m considering the highest and lowest points in the study site, and soils with very clayey texture located in the highest peaks are mainly Nitisols and Chernozem (Figure 2b). Lixisols, which present textural gradient and clay enrichment in deeper layers, are located in general in the middle third part of a slope, and Cambisols (medium texture) are in the lowest portions of the area. Leptsols with silt loam texture cover a very small area, and there was no significant correlation between silt and elevation, as well as with Vis-NIR-SWIR bands.

Landsat 5 Vis-NIR-SWIR bands were all positively correlated with each other, as all of them represent reflectance values; in relation to LST, all of them were negatively correlated, as high temperatures tend to occur in clayey soils that present low reflectance in the regions of the spectrum represented by Landsat bands. Regarding elevation, significant negative correlations were obtained only for bands 4, 5 and 7. The negative correlation between elevation and these bands may be explained by the fact that sandy soils presented higher reflectance intensity in NIR and SWIR wavelengths comparing to clayey soils, as sand was negatively correlated with elevation.

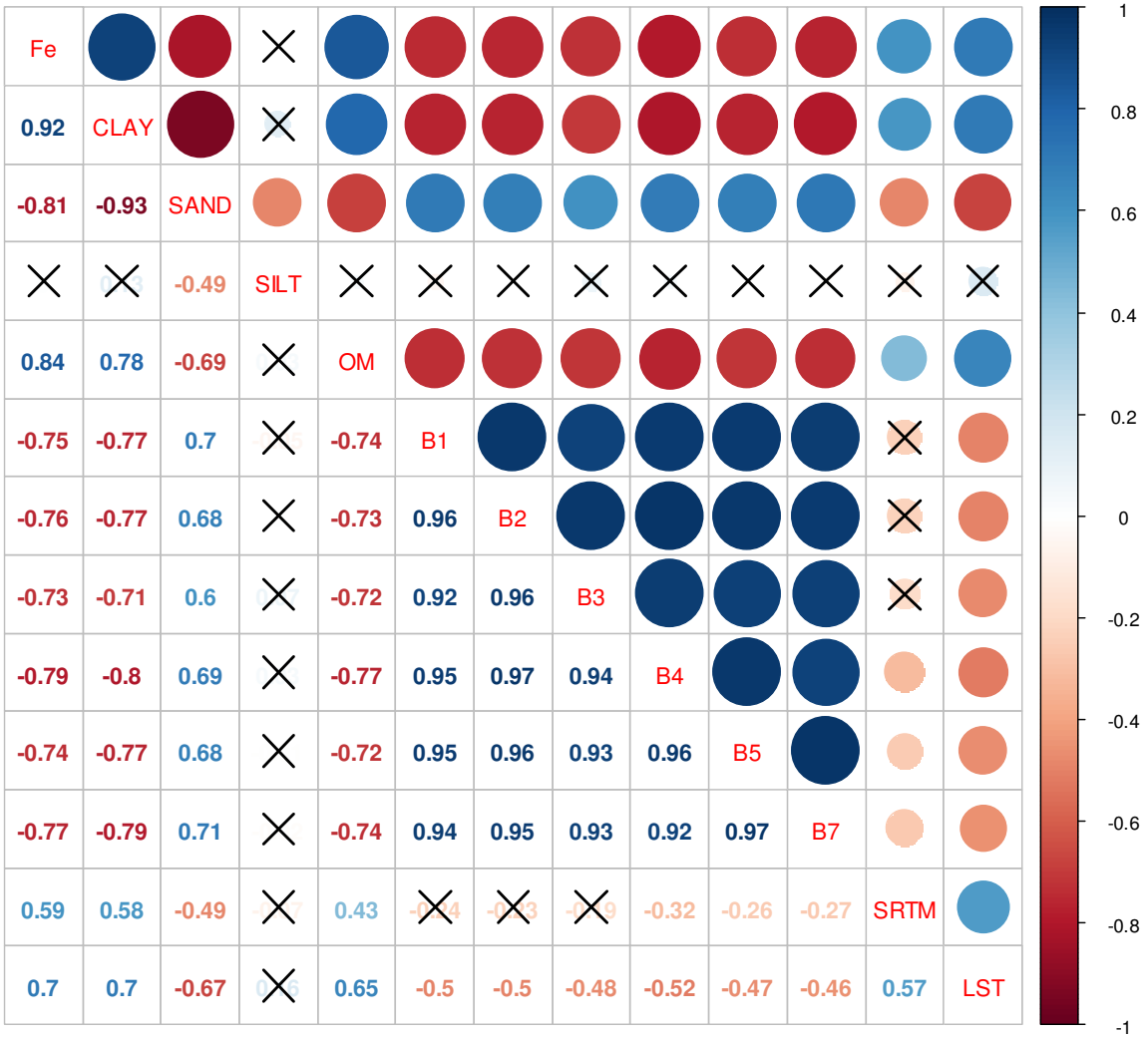

Figure 7. Pearson's correlation coefficient for toposequences data. Numeric correlation values are in the lower portion and a visual representation of them is in the upper portion of the figure; variables are indicated in the diagonal. Values overlapped with an "X" are not statistically significant at the 0.05 level.

All Vis-NIR-SWIR Landsat bands were negatively correlated with clay, OM and iron oxides, and positively correlated with sand. As mentioned before, sandy soils present higher reflectance intensity than clayey ones, due to the high quartz content of the first one. In addition, the presence of OM and iron oxides also leads to a more intense electromagnetic energy absorption, thus decreasing reflectance patterns of soils (Demattê and Garcia, 1999; Dalmolin et al., 2005).

Analyzing the relationship between LST (Figure 10a) and soil attributes, clay, OM and iron oxides were all positively correlated with LST, and only sand was negatively correlated. Clay, OM and iron oxides are closely related considering the clays are coated by Fe-oxides and the formation of organo-mineral colloids that

enhance soil aggregation stability. Therefore, LST followed the same pattern with these attributes. Thermal properties of soil constituents indicate how energy in the form of heat is retained in the soil surface, which directly affects LST values. Clay minerals present thermal conductivity (2.9 W m^{-1} k^{-1}, measured at 10 °C) lower than quartz (8.8 W m^{-1} k^{-1}, measured at 10 °C) (Hillel, 2004), the predominant mineral occurring in sand fraction, leading to more retention of heat in the soil surface. OM has an even lower thermal conductivity (0.6 W m^{-1} k^{-1}, measured at 10 °C) (Hillel, 2004), which enhances soil heat retention.

Wang et al. (2015a) found significant negative correlations between clay fraction and diurnal LST and significant positive correlations between sand fraction and diurnal LST, which was the opposite of the results reported so far. These authors studied LST from several scenes from satellite MODIS, with both day and nighttime LST products. The images were selected based on previous rainfall events, so that it was possible to assess differences in LST influenced by similar soil moisture conditions, obtained during daylight, nighttime and also from the difference between day and night (called diurnal temperature range, DTR). The presence of water highly influences thermal properties such as thermal inertia, which is basically the resistance of a material to heating (Kuenzer and Dech, 2013). In this case, clayey soils presented lower LST values comparing to sandy ones, because the former have a higher water storage capacity and thus, higher thermal inertia.

Given these results for Pearson's correlation coefficient, it is clear that there is a trend in the behavior of the studied soil attributes in relation to RS products: clay, OM and iron oxides behave the same regarding LST, Vis-NIR-SWIR reflectance and elevation. On the other hand, sand remains opposite to them. The high values of r obtained between soil attributes from toposequences samples and RS products show there are strong relationships among them, indicating that prediction models such as linear regression may be appropriate for explaining soil attributes patterns occurring in the study site.

3.3. Soil texture and LST

Before running ANOVA, the distribution and variance of the residuals from the simple LR of LST in function of soil texture as factor (eight levels, corresponding to eight classes) was checked. The null hypothesis of normality of residuals from

Shapiro-Wilk test was accepted (p-value = 0.48) and the null hypothesis of homogeneity of group's variance from Levene's test was accepted as well (p-value = 0.77), which enables the execution of parametric tests such as ANOVA. Results from ANOVA (Figure 8a) show that the null hypothesis that population's means are equal is rejected (F value = 17.54, p-value < $2.2^{*}\ 10^{-16}$), therefore not all of textural classes means are equal.

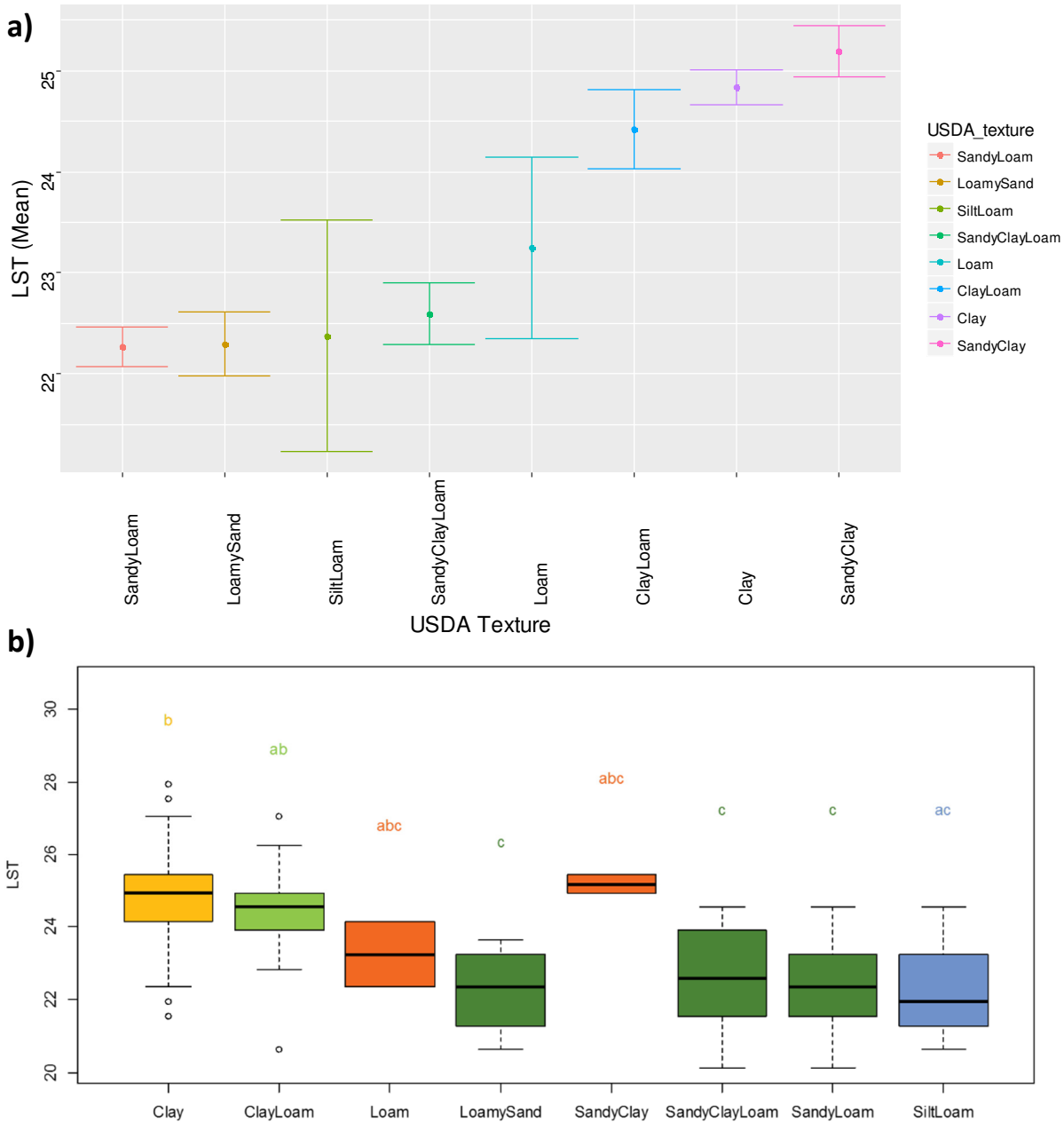

Figure 8. a) Results from ANOVA of LST in function of soil texture. Means of textural classes are in ascending order; **b)** results from Tukey HSD multiple comparison of means.

Figure 8a indicates what types of texture differ from others, but results from ANOVA only tell that there are statistically significant differences among means of groups, and not specifically, which groups differ. That is why it is necessary to run ANOVA post-hoc tests, such as Tukey HSD multiple comparison of means (Figure 8b).

Soils samples with clay (b) and clay loam (ab) textures were statistically different from loamy sand, sandy clay loam and sandy loam samples. These three last ones were grouped together (c), as their means were statistically equal (Figure 8b). This was the strongest difference among soil textures regarding LST. Sandy soils presented lower LST mean values, while clayey soils present higher mean values.

There are few samples with loam and sandy clay texture (only two samples for each texture class), then results found for these classes may not be representative. Regardless of this sample size issue, these two classes were not differentiated from all the others (were grouped as abc), which is reasonable considering that both of them present considerable clay and sand contents, hindering any differences in LST that may occur when sand is present in more expressive percentages rather than clay, and vice-versa.

Analyzing soil samples with silt loam texture (ac), they were differentiated only from clay-textured samples (b). However, there are only three samples from this texture class, and issues regarding the sampling representativeness are raised again. Even though there may be not enough samples for drawing conclusions for this soil texture class, it can be inferred that the behavior of LST of silty soils is very similar to that of sandy ones.

Considering all these differences, soil texture is one important factor that influences soil LST. However, a new issue is raised: why clayey soils were warmer than sandy ones? Lehnert (2014) highlighted the main factors affecting soil temperature besides soil properties, which can be divided between those regarding the soil environment, such as land relief and cover, and those relating to the state of the atmosphere, such as pressure gradients, wind and atmospheric transmittance. Land relief can be assessed using DEM and its derivative products, being a factor easier to explore rather than the atmosphere and climate factors, which are not available for the exact location of the study site and for old dates.

In the case of soil properties, grain size, mineralogy and OM content highly influence on soil thermal properties, as it was mentioned before. The water content has a high influence on the thermal properties as well. However, the dry season (winter) from which the Landsat 5 image was obtained (August 27th) may considerably decrease the influence of water on soil LST. Unfortunately, we did not found meteorological data from Rafard in this period, but there is information regarding the municipality of Piracicaba (series of climatological data from the "Luiz de Queiroz" College of Agriculture, Piracicaba, SP), which is approximately 40 km distant. We assumed that precipitation values registered for Piracicaba were the same in Rafard.

Considering a period of 60 days prior to the image date, there were only two significant rain events (precipitation higher than 0.5 mm): on July 21st (18.2 mm) and August 24th (15.5 mm). This amount of rain probably was not enough to cause water retention in the soil, particularly considering the last rain event, which happened in the conditions of bare soil, favoring water evaporation. In fact, the conditions of the surface during the image date are closer to that of dry soils. In this case, the surface overheating of clay soils with high OM and iron oxides contents is more likely to occur, as these soils absorb more radiation given their dark color and low albedo comparing to sandy soils.

Pramanik and Aggarwal (2013) compared thermal properties of soils with different texture: red loamy sand, black clay soil and alluvial sandy loam. The authors verified that thermal properties mainly vary with soil texture, water content, mineralogical composition and organic matter content, and the most evidenced effect on temperature was soil color. Even though the authors recorded surface temperature values in the field at saturation and a few days later, the black clay soil presented both the highest water content and temperatures. The effect of the black color as determined by mineralogical composition and OM had stronger influence on the increase of temperature because of more heat absorption rather than the effect of high moisture. The latter was not prominent on soil temperature. In our research, clay soils presented the highest temperatures, mainly Chernosol and Red Nitisol, which present high OM and iron oxides contents and, consequently, dark colors.

Another factor that may also explain why soils with surface sandy texture were cooler than those with clayey is their location in the terrain, perhaps being the most influential one. Kubiak and Stach (2012) pointed that the inclination and aspect

of slopes can expressively modify the distribution of LST values. The inclination of slopes (aspect) and solar radiation reaching the surface are conditioned by the relief, and there is a clear relationship between areas with low LST (Figure 10a) and the low amount of solar radiation reaching the surface in the exact time that the image was generated (9:42 a.m.) (Figure 9c), with slope generally south-oriented (Figure 9b). In contrast, the highest positions of the relief (Figure 4b), located in the steepest slopes (Figure 9a), received higher amounts of solar radiation (Figure 9c), and are mostly north-oriented (Figure 9b).

The soil is a very complex system, and many factors may be interacting within each other, which makes it difficult to list all influential ones on soil temperature. The influence of climatic factors is always connected with the relief, and thus, altered by local climatic effects. In addition, the behavior of soil temperature is specific for each area, and generalizations are not very likely to work (Lehnert, 2014).

3.4. Soil attributes modeling using RS data

3.4.1. Clay and sand prediction

Linear models were performed for prediction of clay and sand contents including different predictors (Table 2). For simple LR, using elevation as a predictor, the obtained value of R^2_{adj} was 0.33 and RMSE was 143.90 g kg^{-1}, in calibration. In the case of using only LST as a predictor, R^2_{adj} was 0.48 and RMSE was 126.67 g kg^{-1}. Considering Vis-NIR-SWIR bands as predictors, R^2_{adj} was 0.70 and RMSE was 92.51 g kg^{-1} in multiple LR. The model using elevation explained only 33% of the variability of the response data (clay), LST explained 48% and Vis-NIR-SWIR bands, 70%. The model using Landsat bands presented the best predictive capacity, as it also presented the lower value of RMSE. Results from validation followed the same trend, and RPIQ indicated that the first two models were considered poor while the last one was excellent.

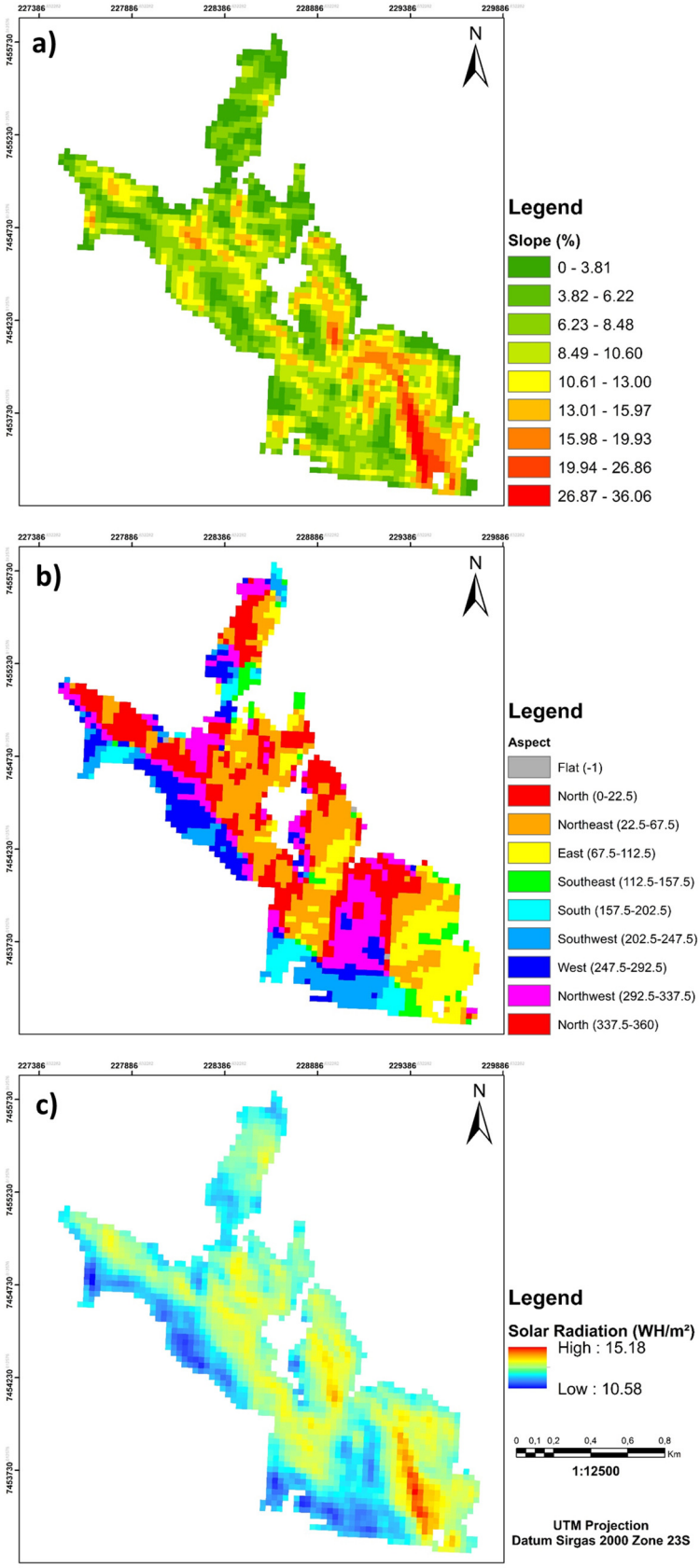

Figure 9. a) Slope, **b)** aspect and **c)** incoming solar radiation, obtained from the DEM (SRTM).

The use of only elevation data resulted in a model with very poor predictive capacity, as this variable is usually used in association with others for increasing the inference power of statistical models. In addition, the effect of relief variations in the study site may not be strongly expressed as the variations of reflectance and LST determined by inherent variations in the soil constituents.

In the case of soil surface reflectance, it is mainly resulted from the interaction of electromagnetic radiation with soil particles, and physical, chemical and mineralogical soil characteristics influence on how energy is absorbed, transmitted or reflected by the soil. Soil constituents directly influence on energy exchange processes, which impacts on LST. Differences in thermal properties and emissivity spectra of soil constituents are likely to influence on soil LST values, and many authors highlighted the influence of soil moisture as well (Bonn and O'Neill, 1993; Sun and Pinker, 2004; Zhao and Li, 2013).

Table 2. Calibration and Validation results from linear regression models for prediction of soil attributes clay, sand, organic matter and iron oxides.

Strategy	Calibration		Validation		
	R^2_{adj}	RMSE[1]	R^2_{adj}	RMSE[1]	RPIQ
	Clay				
Elevation	0.33	143.90	0.14	169.11	1.83
LST	0.48	126.67	0.35	156.83	1.97
TM Bands	0.70	92.51	0.67	100.86	3.06
TM Bands + LST + Elevation	0.83	69.41	0.75	89.84	3.44
	Sand				
Elevation	0.22	174.46	0.15	187.03	1.95
LST	0.45	147.12	0.36	172.51	2.12
TM Bands	0.65	112.60	0.55	131.20	2.79
TM Bands + LST	0.75	95.08	0.64	121.72	3.00
	Organic Matter				
Elevation	0.18	7.94	0.15	10.62	1.60
LST	0.44	6.58	0.36	9.83	1.73
TM Bands	0.63	5.08	0.42	7.37	2.24
TM Bands + LST	0.72	4.53	0.55	6.72	2.46
	Iron Oxides				
Elevation	0.21	0.45[2]	0.16	339.81	0.47
LST	0.51	0.36[2]	0.35	74.22	2.16
TM Bands	0.78	0.23[2]	0.55	54.15	2.96
TM Bands + LST	0.88	0.17[2]	0.70	43.69	3.67

[1] g kg^{-1}, [2] in logarithmic scale (base 10).

The use of Landsat 5 surface reflectance data for prediction of soil clay content is a common approach, but few studies were performed using only LST data. Müller et al. (2016) obtained an average RMSE value of approximately 7.22 % in the validation of multiple LR models for estimating clay content using time series of LST. The authors used ASTER images and worked with many scenes in order to model the variability of soil LST from different dates through Principal Component Analysis (PCA), and performed cross-validation. Using LST from a unique Landsat scene here, R^2_{adj} for validation was 0.35 and RMSE = 156.83 g kg^{-1} (15.68 %). These results indicate the use of LST data from only one scene may be not sufficient in order to satisfactorily model the variability of soil clay content.

Wang et al. (2015a) also developed prediction models for soil texture based on LST, but they used the DTR obtained from day and night LST MODIS scenes. Results of R^2 obtained from validation for clay vary from 0.35 to 0.10 depending on the scene date, being similar to the value found here using only LST as a predictor (R^2 = 0.35, Table 2). Regarding RMSE, the authors found values between 4.97 % and 6.53 %, which were much lower than the one found here (approximately 16 %). The authors suggested that the accuracy of predictions is influenced by the spatial heterogeneity and the complexity of landscape conditions, and considering that the region they developed their research in China was located in a plain area formed by alluvial deposits, it is not as heterogeneous as the one described here, which has four different parent materials.

At least one of the assumptions underlying linear models was violated in all presented models (Table 3), showing that they are not adequate for making predictions. The model using only elevation had all the assumptions violated (p-values for Shapiro-Wilk, Durbin-Watson and Breusch-Pagan tests < 0.05), and the other models did not presented independent residuals (p-value for Durbin-Watson test < 0.05).

The multiple LR using all presented predictors so far was executed and results of T-test showed that B1 and B2 were not statistically significant (p-values > 0.1). A new model excluding these two bands was executed, and F-test indicated that the reduced model is appropriate, as coefficients for B1 and B2 are statistically equal to zero (F = 0.12 and p-value = 0.89). Bands 1 and 2 from Landsat 5 are usually applied for studying vegetation and water bodies, which explains why they were not significant for soil studies. B2 represents the green portion of the visible

spectrum, region that is reflected by plants, and B1 represents the blue portion. Their applications have been related to the absorption of vegetation pigments (Billingsley, 1984).

Table 3. P-values obtained from Normality (Shapiro-Wilk), Independency (Durbin-Watson) and Homoscedasticity (Breusch-Pagan) tests in linear regression models.

Strategy	Normality	Independency	Homoscedasticity
		Clay	
Elevation	0.03	0.06	0.00
LST	0.38	0.00	0.14
Bands	0.98	0.02	0.80
Bands + LST + Elevation	0.88	0.19	0.34
		Sand	
Elevation	0.01	0.02	0.03
LST	0.14	0.03	0.91
Bands	0.98	0.08	0.85
Bands + LST	0.17	0.55	0.19
		Organic Matter	
Elevation	0.46	0.17	0.01
LST	0.14	0.46	0.38
Bands	0.13	0.90	0.09
Bands + LST	0.25	0.35	0.13
		Iron Oxides	
Elevation	0.00	0.02	0.03
LST	0.10	0.11	0.96
Bands	0.94	0.00	0.42
Bands + LST	0.70	0.50	0.36

During calibration, this new model presented R^2_{adj} = 0.83 and RMSE = 69.41 g kg^{-1}, and for validation, R^2_{adj} = 0.75 and RMSE = 89.84 g kg^{-1} (Table 2), presenting the best predictive performance comparing to models discussed before. In addition, residuals were normally distributed (p-value = 0.88), independent (p-value = 0.19) and presented constant variance (p-value = 0.34) (Table 3), as indicated by Shapiro-Wilk, Durbin-Watson and Breusch-Pagan tests, respectively. In validation, it presented the highest RPIQ value, and was considered excellent. Given these results, the model considering Vis-NIR-SWIR bands 3 to 7, LST and elevation was adequate for making predictions and a map for the whole study area was performed using the obtained coefficients from each predictor plus the intercept value (β_0) generated during calibration (Figure 10b).

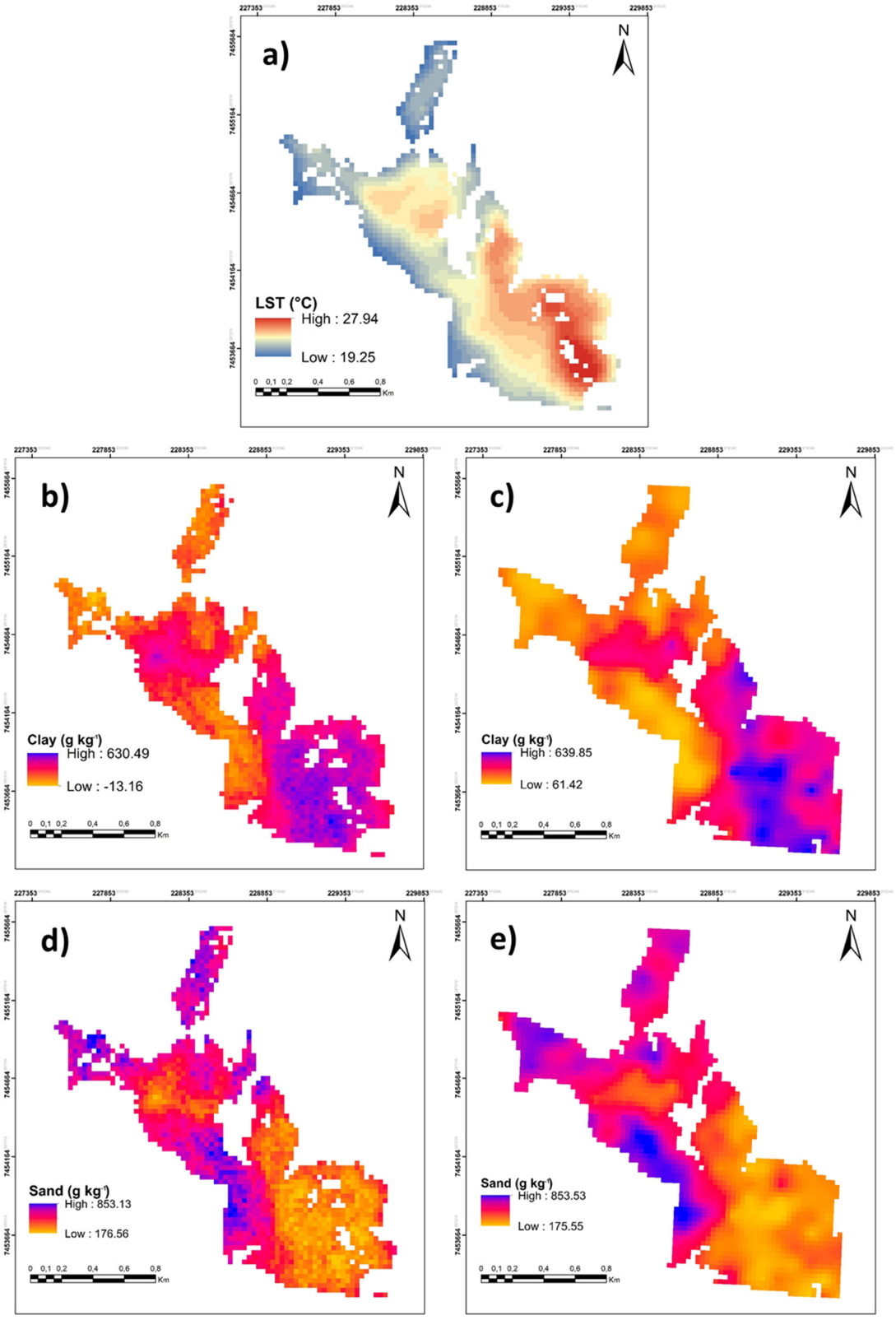

Figure 10. a) Land Surface Temperature obtained from Landsat 5 band 6; **b)** clay map obtained from the best linear regression model; **c)** clay map obtained from ordinary kriging; **d)** sand map obtained from the best linear regression model; **e)** sand map obtained from ordinary kriging.

Models for prediction of sand content (Table 2) presented similar results regarding clay in certain aspects: in calibration using elevation as a predictor, the lowest R^2_{adj} (0.22) and highest RMSE (174.46 g kg^{-1}) values were obtained; the predictive performance of the model using only LST was better than the former (R^2_{adj} = 0.45 and RMSE = 147.12 g kg^{-1}), and the model using all Vis-NIR-SWIR bands presented even better results (R^2_{adj} = 0.65 and RMSE = 112.60 g kg^{-1}); validation results also followed the same trend, but a higher RPIQ value for the model using only LST as a predictor was found (2.12), classifying this model as good.

Assumptions for linear models were violated in the same manner (Table 3): the model with only elevation as a predictor had all the assumptions violated (p-values for Shapiro-Wilk, Durbin-Watson and Breusch-Pagan tests < 0.05), and the other models did not presented independent residuals (p-values for Durbin-Watson test < 0.05). Linear models for sand prediction using MODIS DTR developed by Wang et al. (2015a) in plain areas from China reached R^2 values ranging from 0.32 to 0.45 and RMSE values between 10.69 % and 15.48 % in validation, being similar to the values reported here (R^2 = 0.36 and RMSE = 17.25 %, Table 2) using LST from one unique Landsat 5 scene. Müller et al. (2016) also found an average RMSE value of 16.66 % in the validation of multiple linear models for sand prediction using several ASTER LST products, represented by variables obtained from PCA.

Results for the T-test performed on the multiple LR using all predictors were different in one aspect from those obtained for the clay model. Coefficients for B1 and B2 were not significant, as well as elevation data (p-value > 0.1). In the case of elevation, this variable did not contribute in explaining the variance of sand content as it did for clay. In general, variations of clay content from the toposequences samples used for calibration of models followed variations in elevation, with most of clayey samples (clay content > 500 g kg^{-1}) located in the highest portion of the area (elevation > than 560 m). In the case of sand, the lowest values found in the toposequences samples (sand content < 300 g kg^{-1}) are not concentrated in the highest portion of the area, and the distribution of sand content does not follow a clear trend along elevation patterns.

The new model excluding bands 1-2 and elevation was executed, and F-test indicated that the coefficients for these predictors are statistically equal to zero (F = 0.29 and p-value = 0.83), and all the assumptions for linear regression models were met (residuals normally distributed (p-value = 0.17), independent (p-value = 0.55)

and with constant variance (p-value = 0.19), as indicated by Shapiro-Wilk, Durbin-Watson and Breusch-Pagan tests, respectively (Table 3). During calibration, it presented R^2_{adj} = 0.75 and RMSE = 95.08 g kg^{-1}, and for validation, R^2_{adj} = 0.64, RMSE = 121.72 g kg^{-1} and RPIQ = 3.0 (classified as excellent) (Table 2), having the best predictive performance in relation to models for sand prediction discussed before. A map representing sand content for the study area was performed using the coefficients obtained during calibration from each predictor (bands 4-7 and LST) plus the intercept value (β_0) (Figure 10d).

In order to visualize the predictive performance of all models presented for soil texture quantification, the texture triangle was calculated considering clay and sand predictions (silt content was obtained by subtraction) from each model (Figure 6b-e).

The use of only elevation as a predictor led to underestimation of values for both clay and sand contents, as the highest value predicted for the former was 499.22 g kg^{-1}, and for the latter, 678.56 g kg^{-1}, confirming that this predictor did not explain well the variance of the response variables in terms of continuous values. Considering the textural class accuracy, the models had good performance, as sample points from the full dataset of bare soil areas are situated in the predominant soil textures (Figure 6b), such as clay, clay loam, sandy clay loam and sandy loam. However, their distribution is very narrow, showing that the model did not capture precisely the variation occurring in both clay and sand contents, expressed in Figure 6a.

When using only LST as a predictor, the underestimation found in the case of the previous model using elevation was attenuated, as the maximum values obtained for clay and sand were 602.83 g kg^{-1} and 862.37 g kg^{-1}, respectively. These values were much closer to those resulting from wet chemistry analysis (677.97 g kg^{-1} for clay and 879.28 g kg^{-1} for sand, Table 1). In addition, models using LST represented better the variations in soil texture in the study area, although they resulted in many predictions with repeated values for clay and sand contents, as most of the sample points are overlapped in the texture triangle (Figure 6c). This indicates LST values are similar for soil samples pertaining to the same textural class, which contributes to a good accuracy in soil texture mapping.

The prediction models using all Vis-NIR-SWIR bands represented well the variations of soil texture in the study area (Figure 6d), embracing almost the same

range of soil clay and sand contents obtained from wet chemistry analysis. The maximum value for clay was 628.69 g kg^{-1} and for sand, 897.55 g kg^{-1}. Even though soil samples were placed quite correctly in their textural classes, variations in clay and sand contents encompassing each textural class are still not well represented. Sample points are not as spread in the texture triangle as they are in Figure 6a.

Only predictions obtained from the last executed models represented well both textural classes from sample points as analyzed by traditional methods and the variability of clay and sand contents around each textural class (Figure 6e). The maximum value obtained for clay was 630.49 g kg^{-1} and for sand, 853.13 g kg^{-1}. The minimum values were -13.16 g kg^{-1} for clay and 176.56 g kg^{-1} for sand. Traditional laboratory analysis obtained minimum values of clay and sand of 60.06 g kg^{-1} and 140.71 g kg^{-1}, respectively. There were only three pixels comprising negative values for clay content, which are errors associated with the model, as negative values for such soil attribute are inexistent. However, these values were estimated for areas with sample points having low clay content as determined by traditional analysis (around 80 g kg^{-1}), being within the RMSE value obtained in validation (89.84 g kg^{-1}, Table 2).

There were not soil samples having loam and silt loam texture in the texture triangles obtained from predictions of each model. Even though we can see a few soil samples having these textural classes in Figure 6a, almost none of them were included in the calibration dataset, due to two reasons: samples were located in pixels that did not have bare soil and were eliminated from the dataset; and in the case of samples having high silt content (> 500 g kg^{-1}), they were eliminated because they were identified as outliers in the boxplot analysis.

3.4.2. Organic Matter prediction

Results from modeling of soil OM content were similar to those obtained from sand. Using elevation as a predictor, the lowest R^2_{adj} (0.18) and highest RMSE (7.94 g kg^{-1}) values were obtained in calibration; the predictive performance of the model using only LST was better than the former (R^2_{adj} = 0.44 and RMSE = 6.58 g kg^{-1}), and the model using all Vis-NIR-SWIR bands presented even better predictive performance (R^2_{adj} = 0.63 and RMSE = 5.08 g kg^{-1}); validation results also followed the same order of improved predictive performance, and RPIQ indicated that the first

two models had poor predictions, while the one using Vis-NIR-SWIR bands had good predictions (Table 2).

Zhao et al. (2014) performed multiple LR using two MODIS DTR scenes for predicting topsoil OM in distinct areas and found R^2_{adj} values during calibration of 0.42 and 0.44, and when considering other variables such as a remotely sensed vegetation index and soil type, the authors also reached higher R^2_{adj} values, such as 0.73 and 0.53. This confirms the use of LST concomitantly with other co-variables improves the prediction performance of LR models when comparing to those using only LST.

The model using only elevation had residuals normally distributed and independent (p-value > 0.05), but the variance was non-constant (p-value < 0.05); the model using only LST had all assumptions underlying linear models met (p-values for Shapiro-Wilk, Durbin-Watson and Breusch-Pagan tests > 0.05), and the model using Vis-NIR-SWIR bands as predictors behaved the same (Table 3).

Results for the T-test performed on the multiple LR model using all predictors identified that only three were significant: B4, B7 and LST. Coefficients for B1, B2, B3, and B5 were not statistically significant, as well as elevation data (p-value > 0.1). As it was discussed in the case of sand, the variations of OM content in the study site are not strongly associated with patterns of elevation. The occurrence of the highest values of OM is associated with soils of different classes such as Gleysol, Chernozem and Nitisol, which are situated in different landscape domains (Figure 2b). Gleysols enhance OM accumulation due to their characteristic hydromorphic environment that potentially decreases OM decomposition rates. Chernozems are eutrophic soils with high organic carbon content in horizon A, and Nitisols are soils with clayey texture, which enhances OM physical protection against decomposition.

Regarding the relationship between OM and Vis-NIR-SWIR bands, even though the presence of OM decreases the reflectance of soils in this spectrum region, not all bands are affected in the same way. Demattê et al. (2003) found that depending on the soil type, the sensibility of reflectance values when comparing the original soil sample with one having OM removal does not occur in all bands.

F-test was performed considering the full model and the one excluding B1, B2, B3, B5 and elevation, and indicated that the coefficients for the eliminated predictors are statistically equal to zero (F = 1.05 and p-value = 0.40). In addition, all the assumptions for linear regression models were met (residuals normally

distributed, independent and with constant variance) (Table 3). During calibration, it presented R^2_{adj} = 0.72 and RMSE = 4.53 g kg^{-1}, and for validation, R^2_{adj} = 0.55, RMSE = 6.72 g kg^{-1} and RPIQ = 2.46 (classified as good), presenting the best predictive performance comparing to models for OM prediction discussed before. A map of OM content for the study area was performed using the coefficients obtained from calibration for the last model, using B4, B7 and LST (Figure 11a).

The highest value obtained for OM in the performed map (Figure 11a) was 28.51 g kg^{-1}, which is lower than that determined by traditional laboratory analysis (52 g kg^{-1}). The lowest value obtained in the map was -0.59 g kg^{-1}. As it happened to the lowest predictions for clay content, few negative values (only two pixels) were obtained in the case of OM. In fact, they are very close to the lowest value of OM in the area found by laboratory analysis (0 g kg^{-1}, Table 1).

The greatest limitation of OM predictions using B4, B7 and LST was the underestimation of areas with OM content higher than 30 g kg^{-1}. This happened because values higher than 45 g kg^{-1} were considered outliers and, therefore, were eliminated during calibration. The poor predictive performance of this model for OM values above 30 g kg^{-1} may be associated to the high variability of OM content occurring in a relatively small sampling area (182 ha), where abrupt changes in this soil attribute strongly affect the sampling effectiveness.

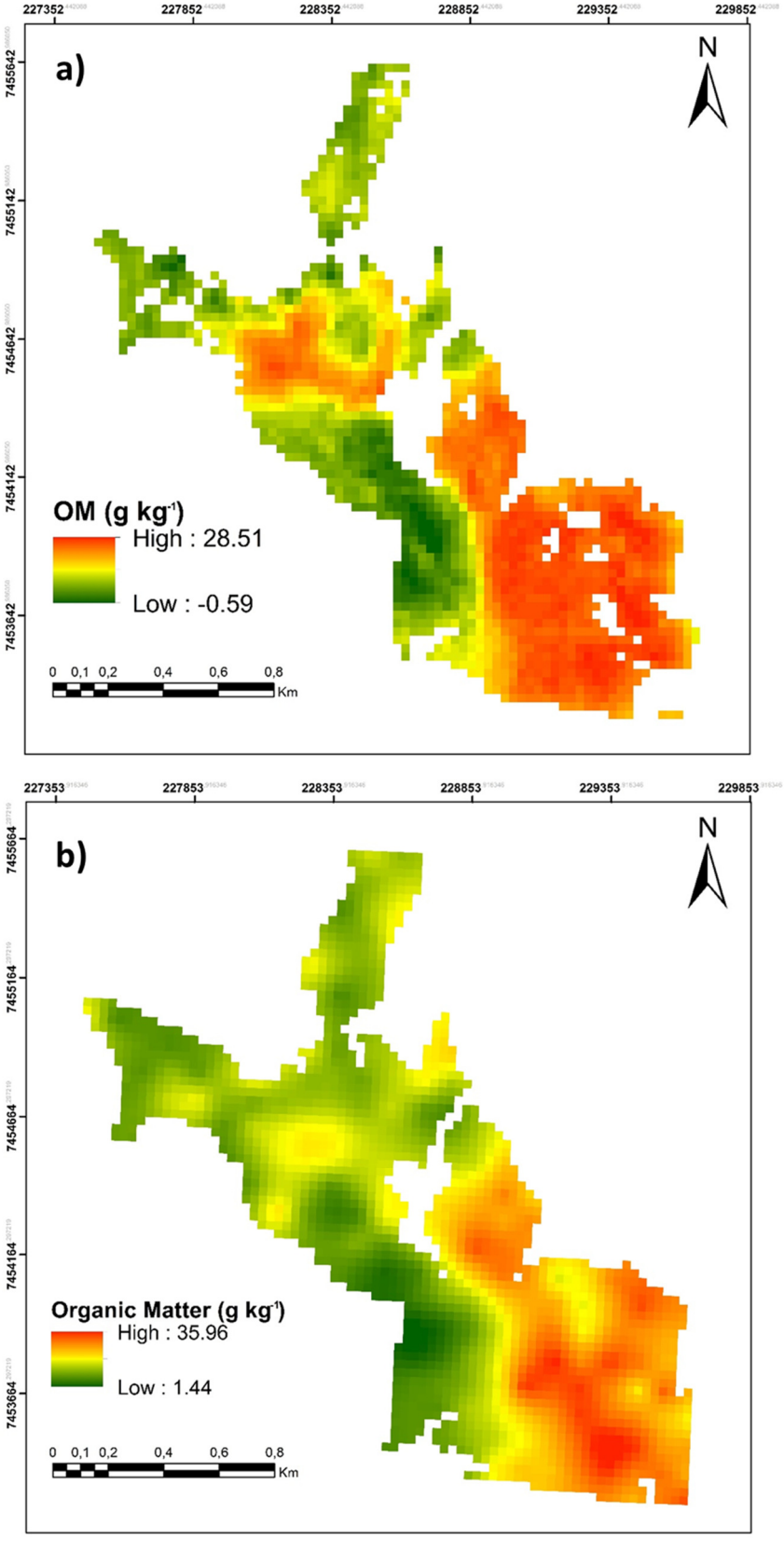

Figure 11. Organic Matter maps obtained from the **a)** best linear regression model and from **b)** ordinary kriging.

3.4.3. Iron Oxides prediction

All models performed for iron oxides prediction did not present residuals normally distributed, violating one of the assumptions underlying LR models. In order to mitigate this issue, a log transformation with base a = 10 was applied to the data, so that they were adjusted to a lognormal distribution. All models were executed again with log-transformed data. For validation purposes, predictions of iron oxides in the log scale were back-transformed to the data's original scale, applying the inverse form of logarithm.

The results obtained from models using different predictors were similar to those obtained for all analyzed soil attributes, with model predictive performance improvement moving from the models using only elevation (R^2_{adj} = 0.21 and RMSE in logarithmic scale = 0.45 g kg^{-1}), to the one using only LST (R^2_{adj} = 0.51 and RMSE in logarithmic scale = 0.36 g kg^{-1}) and finally, using all Vis-NIR-SWIR bands, presenting the best results (R^2_{adj} = 0.78 and RMSE in logarithmic scale = 0.23 g kg^{-1}). Validation results also followed the same order of predictive performance improvement, and regarding RPIQ values, the model using only elevation had very bad performance, that using only LST had good performance and the one using Vis-NIR-SWIR bands had very good performance (Table 2).

The model using only elevation had significant p-values (< 0.05) for Shapiro-Wilk, Durbin-Watson and Breusch-Pagan tests, while the model using only LST as a predictor did not present significant p-values for all tests, meeting all the assumptions of normal distribution, independence and constant variance of residuals. In the case of using Vis-NIR-SWIR bands, only Durbin-Watson test had significant p-values, and therefore residuals were not independent (Table 3).

Results for the T-test performed on the multiple LR using all predictors were similar to those obtained for sand model. Coefficients for B1 and B2 were not statistically significant, as well as elevation data (p-value > 0.1). Elevation did not contribute for modeling the variance of iron oxides in the study site. The occurrence of high values of this attribute is associated with areas that have diabase as the parent material (Figure 2a), as this rock has minerals with high iron content in its composition, such as magnetite. Soils in tropical humid regions originated from diabase through the action of weathering present mainly oxidic mineralogy. Nitisols are a good example, and they occur in a great portion of the study site (Figure 2b).

Indeed, the geology of the area may be influencing more expressively the variations of iron oxides rather than elevation.

F-test was performed considering the full model and the subset model excluding B1, B2 and elevation, and results showed the coefficients for the eliminated predictors were statistically equal to zero (F = 0.29 and p-value = 0.83). All the assumptions for linear regression models were met (Table 3). During calibration, it presented R^2_{adj} = 0.88 and RMSE = 0.17 (log-scale) g kg^{-1}, and for validation, R^2_{adj} = 0.70, RMSE = 43.69 g kg^{-1} and RPIQ = 3.67 (excellent performance) (Table 2). Considering that this model had the best predictive performance in relation to models for iron oxides prediction discussed before, a map of iron oxides content for the study area was performed using the coefficients obtained from calibration (Figure 12a).

The map obtained from the model using bands 4-7 and LST (Figure 12a) had 317.89 g kg^{-1} and 5.21 g kg^{-1} as maximum and minimum values, respectively. The maximum and minimum values reported by traditional analysis were 247 g kg^{-1} and 5 g kg^{-1}, respectively. The model overestimated the iron oxides content, particularly in the areas of Nitisols and Chernozem, which present the highest values of this attribute. However, the predictions for the remaining hipoferric and mesoferric areas presented good accuracy.

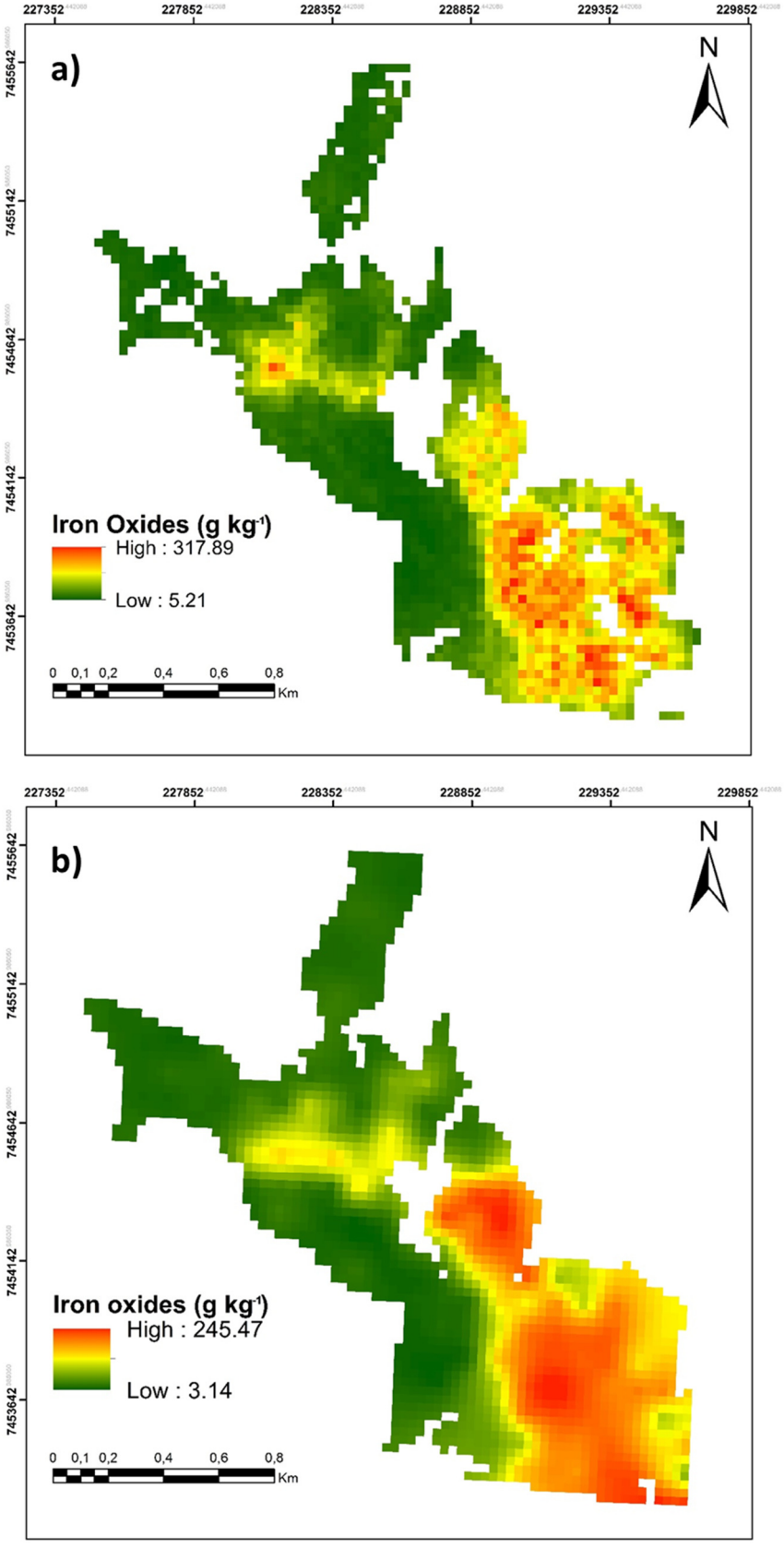

Figure 12. Iron oxides maps obtained from the **a)** best linear regression model and from **b)** ordinary kriging.

3.5. External Validation

The external validation, including sample points from agricultural areas outside the study site, presented higher RMSE and lower R^2_{adj} and RPIQ values comparing to results obtained from validation within the study site. Assessing RPIQ, models for clay and OM had very bad predictions, while the model for sand, poor prediction. Graphs with predicted versus observed values (Figure 13) show that the cloud of points is very scattered along the trend line, which characterizes the poor predictive performance of models for the external validation. Both underestimation and overestimation were identified in the validation of all models, which led to high RMSE values, higher than 200 g kg^{-1} for clay and sand, and higher than 8 g kg^{-1} for OM.

Considering that the area where the model was calibrated was only 182 ha, the extrapolation for bare soil areas situated in a radius of 40 km from the core of the study site resulted in high errors because the variability occurring in such a wider area may be related to other factors that were not included in the model, such as differences in soil management and natural characteristics of the area including geology.

If possible, the use of only sample points located near the study site could minimize the high errors obtained in this validation, as the areas surrounding it tend to present similar features. However, there were not many sample points located nearby, and using only them would be insufficient to present consistent results. In addition, not all sample points were located in bare soil areas, and it was not possible to use all soil observations available in the dataset.

The landscape dynamics is variable from one place to another. This external validation was performed in the whole region but not all of it present the same dynamics, which hinders the models predictive performance. However, considering we are using satellite data located approximately 705 km away from terrestrial targets and with a spatial resolution of 30 m in the case of Vis-NIR-SWIR bands and 120 m for the thermal band as input variables, these results show this technique has good potential on the mapping of soil attributes in great bare soil agricultural regions.

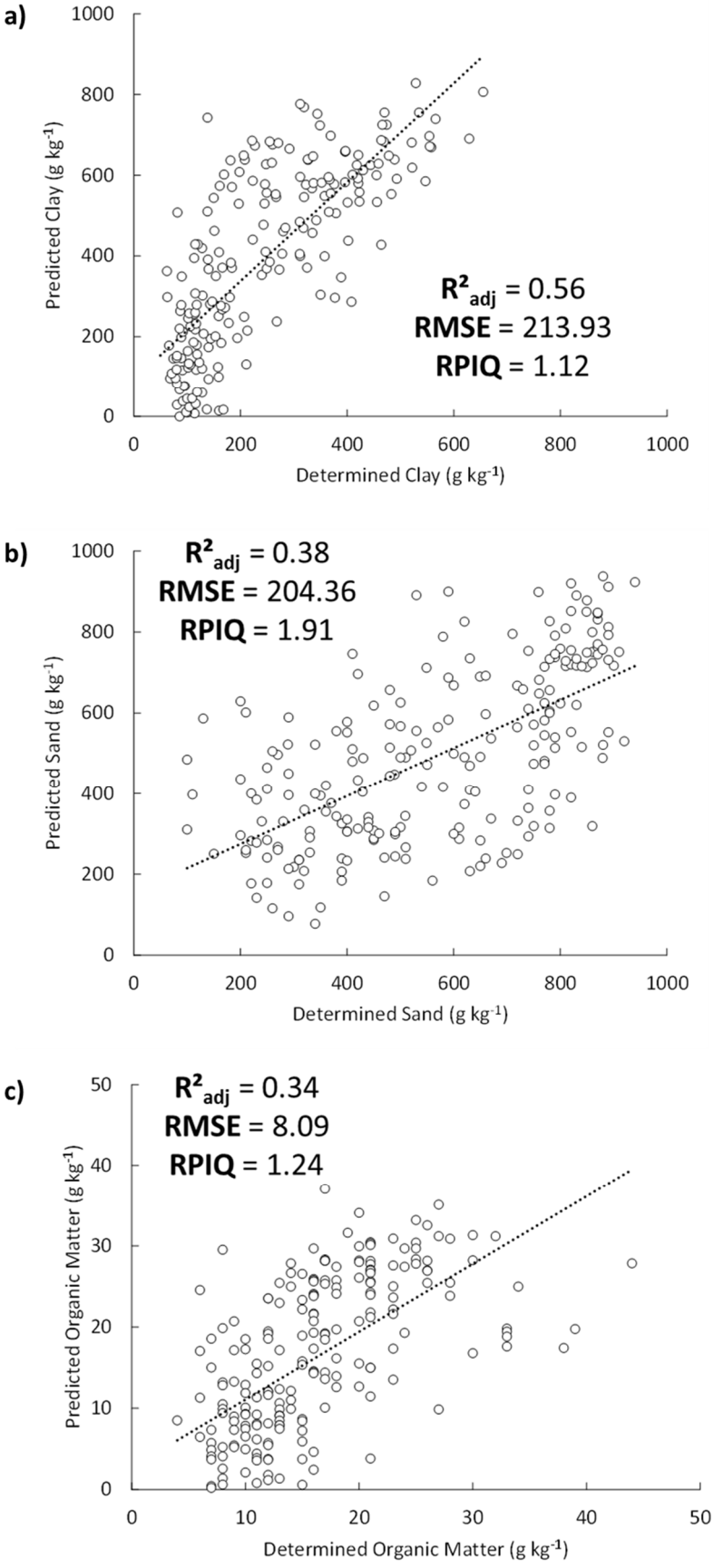

Figure 13. Graphs obtained from external validation for **a)** clay, **b)** sand and **c)** OM, using the best linear regression models.

3.6. Soil attributes modeling based on conventional analyses

The spatial variability of soil attributes illustrated in the empirical semivariograms was best fitted using the theoretical exponential function, which presented highest R^2 and lowest RMSE (Figure 14). Parameters obtained from semivariograms such as the range (a), the sill ($C + C_o$) and the nugget (C_o) indicate how soil attributes are spatially structured and related. Only those obtained from the exponential function fitted to the empirical semivariograms were assessed for each soil attribute (Table 4).

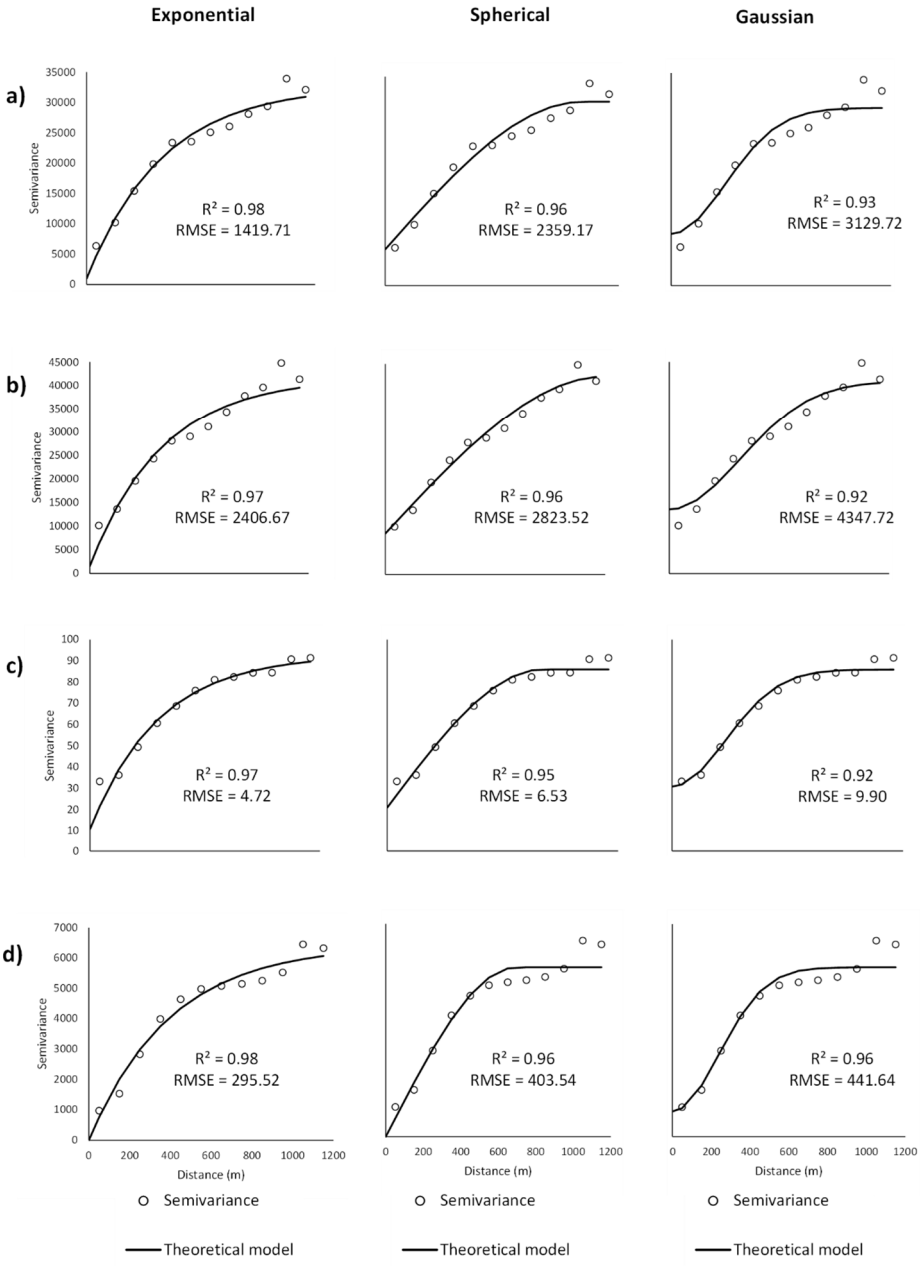

Figure 14. Semivariograms for **a)** clay, **b)** sand, **c)** organic matter and **d)** iron oxides, using Exponential, Spherical and Gaussian theoretical functions.

Table 4. Semivariogram parameters obtained from exponential models

Soil Attribute	Nugget (C_o)	Sill ($C + C_o$)	Range (m)
Clay	1004.10	32724.48	1200.00
Sand	1584.39	41833.01	1200.00
OM	10.75	92.91	1072.483
Fe_2O_3	0	6432.90	1200.00

For clay, sand and iron oxides, the same range value was obtained, equal to 1200 m, representing the distance in which samples are no longer correlated. In the case of OM, the obtained range value was 1072.48 m. Considering that the maximum distance between sample points located in the study site is around 2800 m, it means that there is a high variability for all soil attributes, which is confirmed by the analysis of the CV (Table 1). Wilding and Drees (1983) proposed a CV classification to indicate the degree of variability, and when CV is higher than 35%, the data is characterized as having high variability. All obtained CV values are above this threshold.

Regarding the degree of spatial dependence, it was classified as strong for all soil attributes, presenting values lower than 25% (Table 5). This strong dependence indicates that all variables are highly influenced by variations in intrinsic soil characteristics, such as parent material and weathering action (Cambardella et al., 1994; Parfitt et al., 2009).

Table 5. Results from cross validation, spatial dependence and degree of randomness obtained from geostatistical analyses

Soil Attribute	Cross Validation			Spatial dependence		Degree of randomness	
	R^2	RMSE[1]	ASE[1]	C_o / ($C + C_o$)	Level	C_o/C	Random component
Clay	0.73	92.10	89.45	0.03	Strong	0.03	Small
Sand	0.72	107.35	102.96	0.04	Strong	0.04	Small
OM	0.59	6.02	5.74	0.12	Strong	0.13	Small
Fe_2O_3	0.82	33.31	36.87	0	Strong	0	Small

[1] g kg^{-1}

While semivariogram ranges rely on the spatial interaction of soil processes affecting each soil attribute at the adopted sampling scale, the nugget variance (C_o), which is the variance at distance (h) = 0, reflects field and experimental variability, or

random variability that cannot be identified at the sampling's scale (Cambardella et al., 1994). The degree of randomness assessed by the ratio between the nugget variance and the partial sill (Guerra, 1998) was classified as small for all soil attributes, confirming that the random component associated with the nugget variance is not significant, and does not influence in the modeling of spatial dependence described by the semivariogram function.

Results from cross validation (Table 5) assessed the quality of the kriging technique, based on the spatial structure obtained from the semivariogram adjust. For clay, sand and iron oxides, there was a good adjust between measured and predicted values, which is reflected by the R^2 values higher than 0.7. In the case of OM, R^2 was 0.59, showing that there are greater differences between measured and predicted values than those obtained for the other soils attributes. Regarding clay and sand, RMSE values of 92.10 g kg^{-1} and 107.35 g kg^{-1} were obtained, respectively, which are relatively high. The kriging technique underpredicted large values and overpredicted small ones, which may be causing this high error. The interpolated surfaces obtained from kriging for clay and sand, in a resolution cell of 30 m, are illustrated in figures 10c and 10e, respectively.

The ASE values (Table 5) for both clay and sand are similar to RMSE values. For clay, there is variation of approximately 3 g kg^{-1} between these errors, and for sand, 4 g kg^{-1}. If these errors are close, there is an indicative that the data variability is correctly assessed in prediction. The RMSE value obtained for iron oxides was 33.31 g kg^{-1} and the difference between RMSE and ASE was about 4 g kg^{-1}, which was similar to the results obtained for sand and clay. Thus, considering the range occurring for these soil attributes as determined by traditional soil analysis (Table 1), these RMSE values represent approximately 15 % of it. For OM, the RMSE (6.02 g kg^{-1}) represents 13 %, considering that values above 45 g kg^{-1} were eliminated prior to kriging, as they were outliers identified by the boxplot analysis executed before. The difference between RMSE and ASE for OM was also small (0.28 g kg^{-1}). The interpolated surfaces obtained from kriging for OM and iron oxides, in a resolution cell of 30 m, are illustrated in figures 11b and 12b, respectively.

In general, results from cross validation indicated that the performance of kriging was good, but considering that there is a high variability in all soil attributes in the study site, a higher sampling density in problematic areas like those presenting

the highest values could improve the modeling of spatial dependence using geostatistical methods, thus decreasing the uncertainties reported here.

3.7. Comparison between thematic maps obtained from laboratory analysis and RS variables

For clay (Figure 15a) and sand (Figure 15b) contents, there were expressive differences between the maps obtained from laboratory analysis using geostatistical methods and from RS variables, using multiple LR. The areas in white inside the study site perimeter represent low residuals, between -50 and 50 g kg^{-1} for clay and -60 and 60 g kg^{-1} for sand. These areas were similar regarding the prediction performance of both methods.

In general, the presence of high residuals across most of the study area indicates that there is a high variation regarding the predictions performed by both methods, and the areas that the kriging technique overestimated clay contents in relation to multiple LR predictions are associated with underestimated sand contents by kriging when comparing to multiple LR, and vice-versa.

For OM (Figure 15c) and iron oxides (Figure 15d) contents, there were also high differences between the maps obtained from laboratory analysis using geostatistical methods and from RS variables, using multiple LR, but they were concentrated in a few regions and not spread all over the study site like for clay and sand. The areas in white inside the study site perimeter also represent low residuals, between -5 and 5 g kg^{-1} for OM and -30 and 30 g kg^{-1} for iron oxides, being similar regarding the prediction performance of both methods.

The higher differences for OM (between 7 and 15 g kg^{-1}) were detected for the areas presenting the highest contents of this attribute, and the kriging (Figure 11b) map presented overestimated results compared to those from multiple LR map (Figure 11a). Even though outliers were eliminated in both cases, the kriging map presented a higher range (34.52 g kg^{-1}) comparing to that from multiple LR, which was 29.1 g kg^{-1}. In fact, OM variability across the study site was more accurately represented utilizing geostatistical methods.

For iron oxides, the differences obtained between the adopted mapping approaches are concentrated in the southeast portion of the area. The areas with a

high underestimation of iron oxides content of kriging in relation to multiple LR results (-136.21 to -56.04 g kg^{-1}) are probably due to the higher range obtained in the LR map (312.68) (Figure 12a). The range obtained from the kriging map was much smaller (242.33 g kg^{-1}) and virtually the same as the one reported by soil wet chemistry analysis (242 g kg^{-1}).

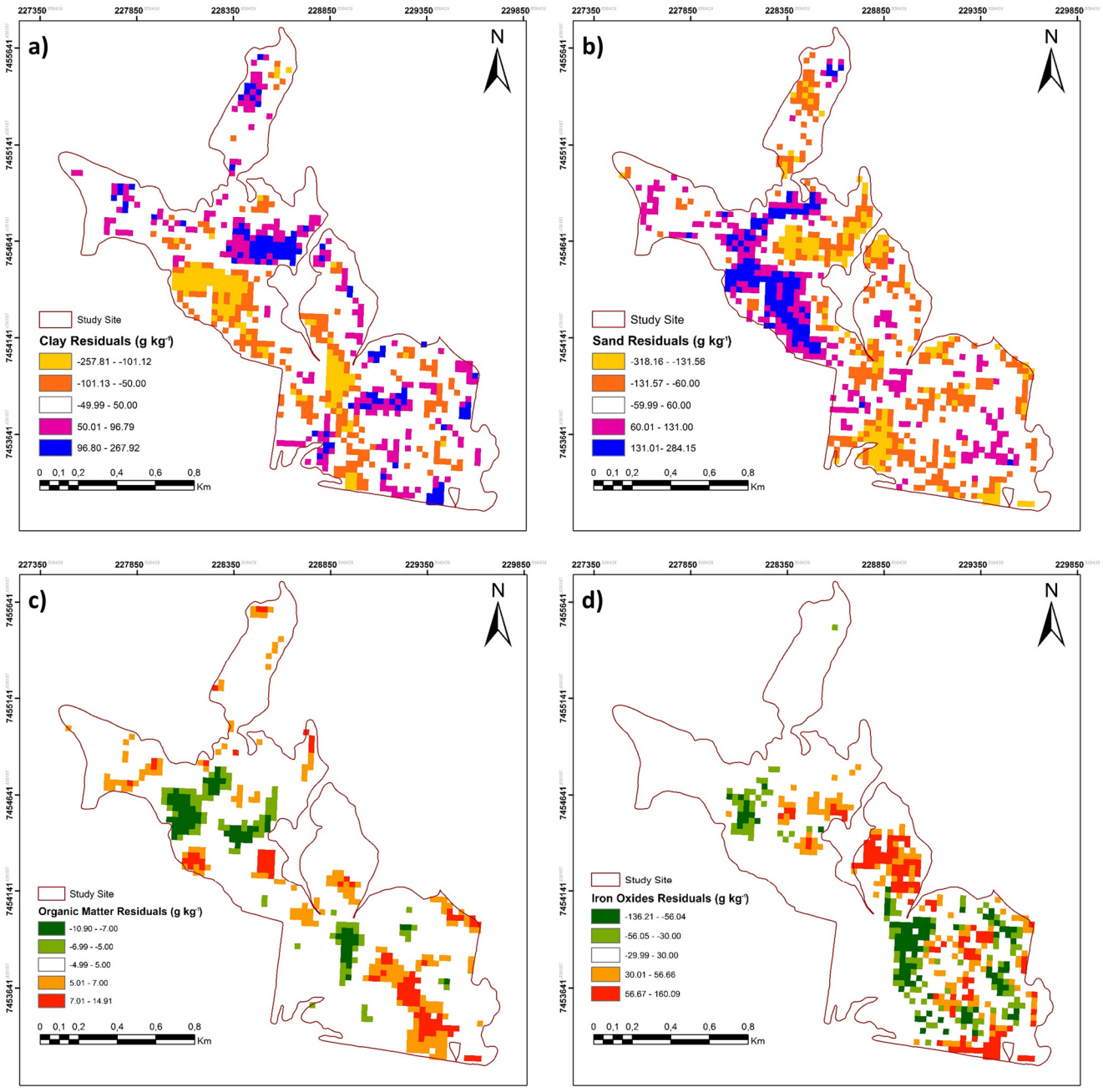

Figure 15. Residuals obtained between kriging and linear regression maps for **a)** clay; **b)** sand; **c)** organic matter and **d)** iron oxides.

In the validation of both methods, the uncertainty was assessed, and the RMSE values obtained in kriging cross validation (Table 5) were lower than those obtained from validation of best multiple LR models (Table 2), except for clay content, when RMSE values were very close. Considering that all sample points were used in the modeling of the semivariogram function and subsequent kriging interpolation, as well as in the cross validation, it is expected that interpolated maps present higher accuracy comparing to multiple LR maps when assessing RMSE. Therefore, two different techniques are being compared, and the calibration sample size is also different. In multiple LR, models were calibrated using only 62 sample points distributed in toposequences, while for ordinary kriging approximately 180 sample points were used. For the last one, the use of a large sample dataset is essential for proper modeling of spatial dependence, while for multiple LR, the choice of good predictors is crucial for improving the predictive capacity of models.

The use of interpolation techniques for mapping soil attributes variability in agricultural fields is a common approach, and the high sampling density required involves high costs in wet chemistry analysis, despite the environmental impact of using chemical reagents. On the other hand, prediction models based on RS data, such as Landsat LST and surface reflectance bands, require a smaller sample set in order to model soil attributes, and reliable results can be obtained. Furthermore, Landsat images can be acquired free of costs and there is an extensive bibliography in the applications of such images in soil studies.

Mapping of soil attributes using RS contributes to the development of soil digital maps, which in turn, are essential for proper field management. The association between the conventional approach and the digital one can improve mapping of soil attributes in large scales, leading to a sustainable land use planning and management.

4. CONCLUSIONS

1. Soils with clayey texture, high OM and iron oxides contents had higher LST values and were statistically different from the other soils from the study site;
2. The highest LST values obtained for clayey soils may be associated with their position in the landscape, their dark color and low albedo;
3. LST was significantly correlated with all assessed soil attributes, but its use as a predictor in simple LR models for estimating soil attributes presented moderate predictive performance;
4. The use of elevation as a co-predictor was only significant for prediction models of clay;
5. The use of LST for estimating soil clay, sand, OM and iron oxides contents increases the predictive performance of multiple LR models when associated with other RS variables, particularly surface reflectance data, improving the validation of models reaching high R^2_{adj}, high RPIQ and low RMSE values;
6. The prediction of soil attributes in a wide neighborhood using the model calibrated in a small area (182 ha) presented poor accuracy;
7. Soil attributes presented strong spatial dependence and the use of ordinary kriging for surface interpolation represented well their spatial variability;
8. Maps obtained from kriging presented good accuracy and outperformed those obtained from multiple LR models for sand, OM and iron oxides;
9. The use of a small calibration sample set using toposequences sampling technique in association with Landsat 5 products allowed the mapping of soil clay, sand, OM and iron oxides contents through multiple LR models, which is a simple and easy to reproduce technique, appropriate for soil attributes mapping in bare soil agricultural areas.

REFERENCES

Barsi, J. A., Barker, J. L., Schott, J. R., 2003. An Atmospheric Correction Parameter Calculator for a Single Thermal Band Earth-Sensing Instrument. In Proceedings of the 2003 IEEE International Geoscience and Remote Sensing Symposium, Toulouse, France, 21–25 July 2003.

Bazaglia Filho, O., 2012. Comparação entre mapas de solos obtidos pelos métodos convencional e digital numa área complexa. Piracicaba, University of São Paulo, "Luiz de Queiroz" College of Agriculture, 190 p. (Master Dissertation in Soil and Plant Nutrition)

Becker, F., Li, Z.-L., 1995. Surface temperature and emissivity at various scales: Definition, measurement and related problems. Remote Sensing Reviews. 12, 225–253.

Bellon-Maurel, V., Fernandez-Ahumada, E., Palagos, P., Roger, J-M., McBratney, A.B., 2010. Critical review of chemometric indicators commonly used for assessing the quality of the prediction of soil attributes by NIR spectroscopy. TrAC Trends in Analytical Chemistry. 29, 1073–1081.

Ben-Dor, E., Banin, A., 1995. Near-infrared analysis as a rapid method to simultaneously evaluate several soil properties. Soil Science Society of America Journal. 59, 364–372.

Ben-Dor, E., Chabrillat, S., Demattê, J. A. M., Taylor, G. R., Hill, J., Whiting, M. L., Sommer, S., 2009. Using Imaging Spectroscopy to study soil properties. Remote Sensing of Environment. 113, S38–S55.

Billingsley, F. C., 1984. Remote sensing for monitoring vegetation: an emphasis on satellites. In: The Role of Terrestrial Vegetation in the Global Carbon Cycle. p. 161-180. Edited by Woodwell, G.M. John Wiley and Sons, New York.

Bonn, F. J., O'Neill, N. T., 1993. Thermal infrared remote sensing of soils: evolution, trends and perspectives. Remote Sensing Reviews 7, 281-302.

Breusch, T. S., Pagan, A. R., 1979. A Simple Test for Heteroscedasticity and Random Coefficient Variation. Econometrica. 47, 1287–1294.

Camargo, O. A., Moniz, A. C., Jorge, J. A., Valadares, J. M., 1986. Métodos de análise química, mineralógica e física de solos do IAC. Campinas: Instituto Agronômico, 94p. (IAC. Boletim Técnico, 106)

Cambardella, C. A., Moorman, T. B., Novak, J. M., Parkin, T. B., Karlen, D. L., Turco, R. F. and Konopka, A. E., 1994. Field-scale variability of soil properties in central Iowa soils. Soil Science Society of America Journal. 58, 1501-1511.

Carlson, T. N., Ripley, D. A., 1997. On the relation between NDVI, fractional vegetation cover, and leaf area index. Remote Sensing of Environment. 62, 241–252.

Chang, D. H., Islam, S., 2000. Estimation of soil physical properties using remote sensing and artificial neural network. Remote Sensing of Environment. 74, 534–544.

Chang, C. W., Laird, D. A., Mausbach, M. J., Hurburgh Jr, C. R., 2001. Near infrared reflectance spectroscopy – Principal components regression analysis of soil properties. Soil Science Society of America Journal. 65, 480-490.

Chang, D. H., Kothari, R., Islam, S., 2003. Classification of soil texture using remotely sensed brightness temperature over the southern great plains. IEEE Transactions on Geoscience and Remote Sensing. 41, 664–674.

Chen, F., Kissel, D. E., West, L. T., Adkins, W., Rickman, D., Luvall, J. C., 2008. Mapping soil organic carbon concentration for multiple fields with image similarity analysis. Soil Science Society of America Journal. 72, 186–193.

Cooley, T., Anderson, G. P., Felde, G. W., Hoke, M. L., Ratkowski, A. J., Chetwynd, J. H., Gardner, J. A., Adler-Golden, S. M., Matthew, M. W., Berk, A., Bernstein, L. S., Acharya, P. K., Miller, D., Lewis, P., 2002. FLAASH, A Modtran4-based atmospheric correction algorithm: its application and validation. IEEE Transactions on Geoscience and Remote Sensing. 3, 1414–1418.

Dalmolin, R. S. D., Gonçalves, C. N., Klamt, E., Dick, D. P., 2005. Relação entre os constituintes do solo e seu comportamento espectral. Ciência Rural, 35(2), 481-489.

Demattê, J. A. M.; Garcia, G. J., 1999. Alteration of soil properties through a weathering sequence as evaluated by spectral reflectance. Soil Science Society of America Journal. 63, 327–342.

Demattê, J. A. M., Epiphanio, J. C, N., Formaggio, A. R., 2003. Influência da matéria orgânica e de formas de ferro na reflectância de solos tropicais. Bragantia. 62(3), 451–464.

Durbin, J., Watson, G. S., 1950. Testing for Serial Correlation in Least Squares Regression I. Biometrika. 37, 409–428.

Eppelbaum, L., Kutasov, I., Pilchin, A., 2014. Applied Geothermics, Springer, 751 p.

ESRI, 2011. ArcGIS Desktop: Release 10. Redlands, CA: Environmental Systems Research Institute.

Faraway, J. J., 2015. Linear Models in R. Chapman & Hall/CRC, Boca Raton, Flórida.

Fongaro, C. T., 2015. Mapeamento granulométrico do solo via imagens de satélite e atributos de relevo. Piracicaba, University of São Paulo, "Luiz de Queiroz" College of Agriculture, 106 p. (Master Dissertation in Soil and Plant Nutrition)

Fu, P., Rich, P.M., 2000. A geometric solar radiation model and its applications in agriculture and forestry. Proceedings of the Second International Conference on Geospatial Information in Agriculture and Forestry. p. 357-364.

Gallo, B. C., 2015. Imagens multitemporais do Landsat TM como estratégia no apoio ao levantamento pedológico. Piracicaba, University of São Paulo, "Luiz de Queiroz" College of Agriculture, 78 p. (Master Dissertation in Soil and Plant Nutrition)

Guerra, P. A. G., 1988. Geoestatística operacional. Brasília: Ministério das Minas e Energia, Departamento de Produção Mineral, 145p.

Hillel, D., 2004. Introduction to Environmental Soil Physics. Elsevier.

Hook, P. B., Burke, I. C., 2000. Biogeochemistry in a short grass landscape: control by topography, soil texture, and microclimate. Ecology. 81, 2686–2703.

Hudson, B. D., 1994. Soil organic matter and available water capacity. Journal of Soil and Water Conservation. 49, 189–94.

IUSS Working Group WRB, 2015. World Reference Base for Soil Resources 2014, update 2015: International soil classification system for naming soils and creating legends for soil maps. World Soil Resources Reports n. 106. Rome: FAO, 203 p.

Kim, H.-Y., 2013. Statistical notes for clinical researchers: assessing normal distribution (2) using skewness and kurtosis. Restorative Dentistry & Endodontics. 38(1), 52–54.

Konen, M. E., Burras, C. L., Sandor, J. A., 2003. Organic carbon, texture, and quantitative color measurement relationships for cultivated soils in north central Iowa. Soil Science Society of America Journal. 67, 1823–1830.

Kruse, P. W., McGlauchlin, L. D., McQuistan, R. B., 1962. Elements of Infrared Technology: Generation, Transmission and Detection; Wiley: New York, NY, USA, Volume 1962.

Kubiak, M., Stach, A., 2013. Influence of relief and land cover on the distribution of values of the land surface temperature in upper Parsęta river area. Quaestiones Geographicae. 32(3), 39–51.

Kuenzer, C.; Dech, S., 2013. Theoretical Background of Thermal Infrared Remote Sensing. In: Kuenzer, C.; Dech, S. (eds.), Thermal infrared remote sensing: sensors, methods, applications. Springer, Dordrecht, 2013, 736 p.

Lehnert, M., 2014. Factors affecting soil temperature as limits of spatial interpretation and simulation of soil temperature. Acta Universitatis Palackianae Olomucensis, Facultas Rerum Naturalium, Geographica. 45(1), 5–21.

Levene, H., 1960. Robust tests for equality of variances. In: Olkin, I. (ed.), Contributions to Probability and Statistics. . Stanford Univ. Press, Palo Alto, CA., 1960, 278–292.

Lepsch, I. F., Silva, N. M. da, Espironelo, A., 1982. Relação entre matéria orgânica e textura de solos sob cultivo de algodão e cana-de-açúcar, no estado de São Paulo. Bragantia. 41(1), 231–236.

Li, Z.; Tang, B.; Wu, H.; Ren, H.; Yan, G.; Wan, Z.; Trigo, I. F.; Sobrino, J. A., 2013. Satellite-Derived Land Surface Temperature: Current Status and Perspectives. Remote Sensing of Environment. 131, 14–37.

Madeira, N. J., 1993. 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, 250 p.

McBratney, A. B., Santos, M. L. M., Minasny, B., 2003. On digital soil mapping. Geoderma. 117, 3–52.

McMillin, L. M., 1975. Estimation of sea surface temperature from two infrared window measurements with different absorptions. Journal of Geophysical Research. 80, 5113–5117.

Mezzalira, S., 1966. Folha Geológica de Piracicaba: SF 23-M 300. São Paulo, Instituto Geográfico e Geológico do Estado de São Paulo. 1 mapa. Escala 1:100.000.

Minasny, B., McBratney, A., 2013. Why you don’t need to use RPD. Pedometron. 33, 14–15.

Moeys, J., 2016. soiltexture: Functions for Soil Texture Plot, Classification and Transformation. R package version 1.4.1. https://CRAN.R-project.org/package=soiltexture

Moura-Bueno, J. M., Dalmolin, R. S. D., ten Caten, A., Ruiz, L. F. C., Ramos, P. V., Dotto, A. C., 2016. Assessment of Digital Elevation Model for Digital Soil Mapping in a Watershed with Gently Undulating Topography. Revista Brasileira de Ciência do Solo v. 40, 15 p.

Müller, B., Bernhardt, M., Jackisch, C., Schulz, K., 2016. Estimating spatially distributed soil texture using time series of thermal remote sensing – a case study in central Europe. Hydrology and Earth System Sciences. 20, 3765–3775.

Nanni, M. R., 2000. Dados radiométricos obtidos em laboratório e no nível orbital na caracterização e mapeamento de solos. Piracicaba, University of São Paulo, “Luiz de Queiroz” College of Agriculture, 365p. (PhD thesis in Soil and Plant Nutrition)

Nanni, M. R., Demattê, J. A. M., 2006. Spectral reflectance methodology in comparison to traditional soil analysis. Soil Science Society of America Journal. 70, 393–407.

Nanni, M. R., Demattê, J. A. M., Silva Jr., C. A., Romagnoli, F., Silva, A. A., Cezar, E., Gasparotto, A. C., 2014. Soil Mapping by Laboratory and Orbital Spectral Sensing Compared with a Traditional Method in a Detailed Level. Journal of Agronomy. 13(3), 100–109.

Ndossi, M. I., Avdan, U., 2016. Application of open source coding technologies in the production of Land Surface Temperature (LST) maps from Landsat: A PyQGIS Plugin. Remote Sensing. 8, 31 p.

Osińka-Skotak, K., 2007. Studies of soil temperature on the basis of satellite data. International Agrophysics. 21, 10 p.

Parfitt, J. M. B., Timm, L. C., Pauletto, E. A., Sousa, R. O. de, Castilhos, D. D., Ávila, C. L. de, Reckziegel, N. L., 2009. Spatial variability of the chemical, physical and biological properties in lowland cultivated with irrigated rice. Revista Brasileira de Ciência do Solo. 33(4), 819–830.

Post, W. M., Kwon, K. C., 2000. Soil carbon sequestration and land use change: processes and potential. Global Change Biology. 6, 317–27.

Pramanik, P., Aggarwal, P., 2013. Comparison of thermal properties of three texturally different soils under two compaction levels. African Journal of Agricultural Research. 8(28), 3679–3687.

Prata, A. J., Caselles, V., Coll, C., Sobrino, J. A., Ottlé, C., 1995. Thermal remote sensing of land surface temperature from satellites: Current status and future prospects. Remote Sensing Reviews. 12, 175–224.

Prevedello, C. L., 2010. Energia Térmica do Solo. In: Jong van Lier, Q., ed. Física do solo. Viçosa, MG, Sociedade Brasileira de Ciência do Solo, 2010. p.241-281.

QGIS Development Team, 2016. QGIS Geographic Information System. Open Source Geospatial Foundation Project. http://www.qgis.org/

Qin, Z.-H., Karnieli, A., Berliner, P., 2001. A mono-window algorithm for retrieving land surface temperature from Landsat TM data and its application to the Israel-Egypt border region. International Journal of Remote Sensing. 22, 3719–3746.

R Core Team, 2015. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL https://www.R-project.org/.

Robertson, E. C., 1988. Thermal properties of rocks. U.S.G.S. Open file report 88-441, 106 p., U.S. Geol. Survey, Reston, Va.

Rouse, J. W., Haas, R. H., Schell, J. A., Deering, D. W., 1974. Monitoring vegetation systems in the Great Plains with ERTS. Third Earth Resources Technology Sattelite-1 Symposium, Greenbelt: NASA SP-351 I, p. 309–317.

Sabins, F. F., 1996. Remote sensing: Principles and Interpretation, 3rd ed. W. H. Freeman and Company, New York, 494p.

Schmugge, T., Hook, S. J., Coll., C., 1998. Recovering surface temperature and emissivity from thermal infrared multispectral data. Remote Sensing of Environment. 65, 121– 131.

Shapiro, S. S., Wilk, M. B., 1965. An analysis of variance test for normality (complete samples). Biometrika. 52, 591–611.

Shepherd, K. D., Walsh, M. G., 2002. Development of reflectance spectral libraries for characterization of soil properties. Soil Science Society of America Journal. 66, 988–998.

Six, J., Bossuyt, H., Degryze, S., Denef, K., 2004. A history of research on the link between (micro)aggregates, soil biota, and soil organic matter dynamics. Soil and Tillage Research. 79, 7–31.

Sobrino, J., Caselles, V., Becker, F., 1990. Significance of the remotely sensed thermal infrared measurements obtained over a citrus orchard. ISPRS Journal of Photogrammetry and Remote Sensing. 44, 343–354.

Sobrino, J. A.; Raissouni, N., 2000. Toward remote sensing methods for land cover dynamic monitoring: Application to Morocco. International Journal of Remote Sensing. 21(2), 353–366.

Sun, D., Pinker, R. T., 2004. Case Study of Soil Moisture Effect on Land Surface Temperature Retrieval. IEEE Geoscience and Remote Sensing Letters. 1(2), 127 – 130.

Tisdall, J. M., Oades, J. M., 1982. Organic matter and water-stable aggregates in soils. European Journal of Soil Science. 33, 141–63.

Tukey, J. W., 1977. Exploratory data analysis. Addison-Wesely, 688 p.

Tukey, J. W., 1949. Comparing Individual Means in the Analysis of Variance. Biometrics. 5 (2), 99–114.

USGS, 2016a. What are the band designations for the Landsat satellites? Available from <https://landsat.usgs.gov/what-are-band-designations-landsat-satellites>. Accessed on March 27, 2017.

USGS, 2016b. Using the USGS Landsat 8 product. Available from < https://landsat.usgs.gov/using-usgs-landsat-8-product>. Accessed on April 3rd, 2017.

Viscarra Rossel, R. A., Walvoort, D. J. J., McBratney, A. B., Janik, L. J., Skjemstad, J. O., 2006. Visible, near infrared, mid infrared or combined diffuse reflectance spectroscopy for simultaneous assessment of various soil properties. Geoderma. 131(1), 59–75.

Viscarra Rossel, R. A., 2007. Robust modelling of soil diffuse reflectance spectra by “bagging-partial least squares regression”. Journal of Near Infrared Spectroscopy. 15, 39–47.

Wakindiki, I. I. C., Ben-Hur, M., 2002. Soil mineralogy and texture effects on crust micromorphology, infiltration and erosion. Soil Science Society of America Journal. 66, 897–905.

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

Wan, Z., 1999. MODIS Land-Surface Temperature Algorithm Theoretical Basis Document. In: Greenbelt MD, USA: NASA/GSFC.

Wang, D.-C., Zhang, G. L., Pan, X. Z., Zhao, Y. G., Zhao, M. S., Wang, G. F., 2012. Mapping soil texture of a plain area using fuzzy-c-means clustering method based on land surface diurnal temperature difference. Pedosphere. 22, 394–403.

Wang, D.-C., Zhang, G.-L., Zhao, M.-S., Pan, X.-Z., Zhao, Y.-G., Li, D.-C., Macmillan, B., 2015a. Retrieval and Mapping of Soil Texture Based on Land Surface Diurnal Temperature Range Data from MODIS. PLoS ONE. 10(6), 14 p.

Wang, F., Qin, Z., Song, C., Tu, L., Karnieli, A., Zhao, S., 2015b. An improved mono-window algorithm for land surface temperature retrieval from landsat 8 thermal infrared sensor data. Remote Sensing. 7, 4268–4289.

Wei, T., Simko, V., 2016. corrplot: Visualization of a Correlation Matrix. R package version 0.77. https://CRAN.R-project.org/package=corrplot.

Wilding, L. P., Drees, L. R., 1983. Spatial variability and pedology. In: Wilding, L. P., Smeck, N. E., Hall, G. F., eds. Pedogenesis and soil taxonomy: Concepts and interactions. New York, Elsevier, 1983. p.83-116.

Williams, P. C. Variables affecting near-infrared reflectance spectroscopy analysis. In: Williams, P. C., Norris, K. (Ed.), 1987. Near-infrared technology in the agricultural and food industries. Saint Paul, MN: American Association of Cereal Chemists, p. 143–167.

Zhao, W., Li, Z.-L., 2013. Sensitivity study of soil moisture on the temporal evolution of surface temperature over bare surfaces. International Journal of Remote Sensing. 34, 9–10.

Zhao, M.-S., Rossiter, D. G., Li, D.-C., Zhao, Y.-G., Liu, F., Zhang, G.-L., 2014. Mapping soil organic matter in low-relief areas based on land surface diurnal temperature difference and a vegetation index. Ecological Indicators. 39, 120–133.