

Remote sensing of the Earth's soil color in space and time



Rodnei Rizzo^{a,h}, Alexandre M.J.-C. Wadoux^b, José A.M. Dematté^{a,*}, Budiman Minasny^c, Vidal Barrón^d, Eyal Ben-Dor^e, Nicolas Francos^e, Igor Savin^{f,u}, Raul Poppiel^a, Nelida E.Q. Silvero^a, Fabrício da Silva Terra^g, Nícolas Augusto Rosin^a, Jorge Tadeu Fim Rosas^a, Lucas Tadeu Greschuk^a, Maria V.R. Ballester^h, Andrés Mauricio Rico Gómez^{a,h}, Henrique Belllinaso^w, José Lucas Safanelli^v, Sabine Chabrillat^{i,j}, Peterson R. Fiorio^k, Bhabani Sankar Das^l, Brendan P. Malone^m, George Zalidisⁿ, Nikolaos Tziolas^o, Nikolaos Tsakiridisⁿ, Konstantinos Karyotisⁿ, Nikiforos Samarinisⁿ, Eleni Kalopesaⁿ, Asa Gholizadeh^p, Keith D. Shepherd^q, Robert Milewski^j, Emmanuelle Vaudour^r, Changkun Wang^s, Elsayed Said Mohamed Salama^{t,u}

^a Department of Soil Science, Luiz de Queiroz College of Agriculture, University of São Paulo, Pádua Dias Av., 11, Piracicaba, Postal Box 09, São Paulo 13416-900, Brazil

^b LISAH, Université de Montpellier, INRAE, IRD, Institut Agro, Montpellier, France

^c Sydney Institute of Agriculture & School of Life and Environmental Sciences, The University of Sydney, Australia

^d Departamento de Agronomía, Universidad de Córdoba, Campus de Rabanales, 14071 Córdoba, Spain

^e Remote Sensing Laboratory, Geography Department, Porter School of the Environment and Earth Sciences, Faculty of Exact Sciences, Tel Aviv University, Tel Aviv 699780, Israel

^f V.V. Dokuchaev Soil Science Institute, 119017 Moscow, Russia

^g Institute of Agricultural Sciences, Federal University of Jequitinhonha and Mucuri Valleys, Av. Universitária 1000, Unaí, Minas Gerais 38610-000, Brazil

^h Center of Nuclear Energy in Agriculture (CENA), University of São Paulo, Piracicaba, São Paulo 13418-900, Brazil

ⁱ Helmholtz Center Potsdam GFZ German Research Centre for Geosciences, 14473 Potsdam, Germany

^j Institute of Soil Science, Leibniz University Hannover, Herrenhäuser Straße 2, 30419 Hannover, Germany

^k Department of Biosystems Engineering, ESALQ, USP, Ave. Pádua Dias 11, Cx. Postal 9, 13418-900 Piracicaba, SP, Brazil

^l Agricultural and Food Engineering Department, Indian Institute of Technology Kharagpur, West Bengal, India

^m CSIRO Agriculture and Food, Black Mountain, ACT, Australia

ⁿ Laboratory of Remote Sensing, Spectroscopy and GIS, Aristotle University of Thessaloniki, 54124 Thessaloniki, Greece

^o Southwest Florida Research and Education Center, Department of Soil, Water, and Ecosystem Sciences, Institute of Food and Agricultural Sciences, University of Florida, 2685 State Rd 29N, Immokalee, FL 34142, USA

^p Department of Soil Science and Soil Protection, Czech University of Life Sciences Prague, Suchdol, 16500 Prague, Czech Republic

^q Innovative Solutions for Decision Agriculture (ISDA), Rothamsted Campus, West Common, Harpenden, Hertfordshire, AL5 2JQ, United Kingdom

^r Université Paris-Saclay, INRAE, AgroParisTech, UMR ECOSYS, 91120 Palaiseau, France

^s State Key Laboratory of Soil and Sustainable Agriculture, Institute of Soil Science, Chinese Academy of Sciences, Nanjing 210008, China

^t National authority for remote sensing and space sciences, 23, Josef Tito St., Nasza El-Gedida, Cairo, Egypt

^u Department of Environmental Management, Institute of Environmental Engineering, People's Friendship University of Russia (RUDN University), 6 Miklukho-Maklaya Street, 117198 Moscow, Russia

^v Woodwell Climate Research Center, Falmouth, MA, USA

^w Coordination of Integrate Technical Assistance, Secretariat of Agriculture and Supply, CATI/SAA, Piracicaba Regional, Campos Salles Street, 507, Piracicaba, SP, Brazil

* Corresponding author at: Department of Soil Science, Luiz de Queiroz College of Agriculture (ESALQ), University of São Paulo (USP), Piracicaba, São Paulo 13418-900, Brazil.

E-mail addresses: rodnei.rizzo@gmail.com (R. Rizzo), alexandre.wadoux@inrae.fr (A.M.J.-C. Wadoux), jamdemat@usp.br (J.A.M. Dematté), budiman.minasny@sydney.edu.au (B. Minasny), vidal@uco.es (V. Barrón), bendor@tauex.tau.ac.il (E. Ben-Dor), nicolasf@mail.tau.ac.il (N. Francos), savin_iyu@esoil.ru (I. Savin), raulpoppiel@gmail.com (R. Poppiel), neli.silvero@usp.br (N.E.Q. Silvero), fabricio.terra@ufvjm.edu.br (F.S. Terra), narosin@usp.br (N.A. Rosin), jorge.fimrosas@usp.br (J.T.F. Rosas), lugasgreschuk@usp.br (L.T. Greschuk), vicky@cena.usp.br (M.V.R. Ballester), anmricogo@usp.br (A.M.R. Gómez), henrique.belllinaso@sp.gov.br (H. Belllinaso), jsafanelli@woodwellclimate.org (J.L. Safanelli), chabrillat@ifbk.uni-hannover.de, chabri@gfz-potsdam.de (S. Chabrillat), fiorio@usp.br (P.R. Fiorio), bsdas@agfe.iitkgp.ac.in (B.S. Das), brendan.malone@csiro.au (B.P. Malone), zalidis@auth.gr (G. Zalidis), ntziolas@ufl.edu (N. Tziolas), tsakirin@auth.gr (N. Tsakiridis), kbkaryotis@auth.gr (K. Karyotis), snnikiforos@auth.gr (N. Samarinis), kalopesa@auth.gr (E. Kalopesa), gholizadeh@af.czu.cz (A. Gholizadeh), keith.shepherd@isda-africa.com (K.D. Shepherd), milewski@gfz-potsdam.de (R. Milewski), emmanuelle.vaudour@inrae.fr (E. Vaudour), ckwang@issas.ac.cn (C. Wang), salama55@mail.ru (E.S.M. Salama).

ARTICLE INFO

Edited by Jing M. Chen

Keywords:

Soil spectroscopy
Spectral library
Color space models
Remote sensing
Landsat
Soil spatio-temporal monitoring

Abstract

Soil color is a key indicator of soil properties and conditions, exerting influence on both agronomic and environmental variables. Conventional methods for soil color determination have come under scrutiny due to their limited accuracy and reliability. In response to these concerns, we developed an innovative system that leverages 35 years of satellite imagery in conjunction with in-situ soil spectral measurements. This approach enables the creation of a global soil color map with a fine spatial resolution of 30 m x 30 m. The system initially identifies bare earth areas worldwide using reflectance bands acquired from Landsat 4 through Landsat 8 between 1985 and 2020. Soil color was quantified using the CIE-XYZ coordinates, utilizing 8005 soil spectral measurements within the visible range (380–780 nm) as ground truth data. We established transfer functions to convert Landsat reflectance bands to standardized XYZ color coordinates. These transfer functions were subsequently applied to images of bare surfaces, covering approximately 38.5% of the Earth's surface. We validated the resulting global soil color map using statistical indices derived from an independent set of ground-truth spectral data, demonstrating a high degree of agreement. By creating the world's first global soil color map, we have set a baseline for future spatial and temporal monitoring of soil conditions, thus enhancing our understanding and management of our planet's vital soil resources.

1. Introduction

Soil color is a fundamental indicator of soil characteristics and conditions, reflecting soil's chemical, physical, and biological properties and processes (Ibáñez-Asensio et al., 2013; Ketterings and Bigham, 2000; Schmidt and Ahn, 2021). Key factors influencing soil color include organic matter (which darkens the soil), calcium carbonates (resulting in a white hue), iron oxide (which can give a red or yellow color), and the size distribution of particles and aggregates (leading to a range of dark and bright colors). Additionally, water affects soil color both directly—by absorbing light energy (Ishida et al., 1991)—and indirectly, by influencing other soil properties, such as iron reduction (Schwertmann, 2015). Soil color is a useful indicator for understanding and managing various soil-related aspects. For instance, soil color can indicate the adsorption of nutrients and toxic compounds on iron oxy-hydroxides (Camargo et al., 2015). It could be used to indicate pesticide adsorption capacity due to organic matter, thereby informing herbicide application rates (Fernandez et al., 1988). Soil color can signify the presence of iron oxides that impact soil aggregation, a key indicator of erosion resilience (Cañasveras et al., 2010), or even act as a photocatalyst influencing the fixation of nitrogen gases (NO and NO₂) (Sánchez-Rodríguez et al., 2023). Moreover, the presence of iron oxides inferred from soil color can be used for paleoclimate reconstruction due to the pedogenic alterations of iron compounds (Jiang et al., 2022; Long et al., 2016). Soil color can also indicate fire severity's impact on biological, physico-chemical, and mineralogical soil properties (Guerrero et al., 2007; Ketterings and Bigham, 2000).

Soil color is measured quantitatively in standard color space systems. In soil science, color description is conventionally determined using the color plate notation from the Munsell color system (Munsell, 1907). Over the last century, the Munsell system was adopted by soil surveyors for soil classifications and is now an integral part of any field-level soil description. Several major classification systems, such as the USDA Soil Taxonomy, rely on the Munsell soil color to discriminate soil classes. However, the method of determining soil color by matching it with standard plates has drawn criticism due to concerns about the accuracy and repeatability of these measurements (Torrent and Barron, 1993). These concerns are especially pronounced when assessments are conducted by inexperienced examiners and under diverse soil moisture conditions.

The emergence of electro-optical sensors such as spectrophotometers and digital cameras has paved the way for a more objective determination of soil color. Many digital soil sensors (Barrett, 2002) and cameras (Adderley et al., 2002) are sensitive to the visible range (400–759 nm) of the electromagnetic spectrum. These devices allow quantitative estimates of soil color, addressing the subjectivity of human perception

(Mouazen et al., 2007; Stiglitz et al., 2016). It has been shown (Shields et al., 1966) that the three indices of the Munsell color system (i.e., the hue, value and chroma) correlate with spectral reflectance in the visible spectrum. Using this principle, several studies (Gholizadeh et al., 2020; Mouazen et al., 2007; Ramos et al., 2020) have used visible and infrared reflectance spectra to estimate soil color.

Almost all optical sensors mounted on satellite platforms record the data within the visible region of the electromagnetic spectrum. These spectral bands aid in image identification via the RGB composites (Escadafal, 1993). In fact, the visible bands of Landsat sensors have been used for estimating soil color since the 1990s (Escadafal et al., 1989; Mattikalli, 1997). For instance, Singh et al. (2004) mapped soil color using AVHRR data for modeling water erosion. Soil color maps obtained from NOAA/AVHRR data were also used as a proxy for estimating soil properties, such as organic matter content (Ben-Dor, 1997), iron content (Bartholomew et al., 2007), lichens over terrain (Bechtel et al., 2002) or surface soil moisture (Persson, 2005).

While remote sensing imagery has proven effective in mapping soil color, these endeavors often remain localized and offer limited spatial resolution. The primary obstacle in mapping soil color across large areas lies in securing satellite images unaffected by clouds, aerosols, vegetation, crop residues, soil wetness, and surface roughness. To address this challenge, recent advancements in digital soil mapping have introduced methods to generate per-pixel mosaics of bare surfaces from a collection of satellite images (Dematte et al., 2018; Diek et al., 2017; Roberts et al., 2019). This fusion of images opens up new avenues for soil assessment via remote sensing by offering time-averaged visuals of bare surfaces.

Remote sensing technology facilitates the extensive estimation of topsoil properties and provides the capability to monitor spatio-temporal changes (Ben-Dor et al., 2009; Padarian et al., 2022). Nevertheless, as of now, there is no comprehensive global map of topsoil color, despite some attempts to map soil color at a regional level (e.g., Poppel et al., 2020, among others). To the best of our knowledge, past studies had moderate success in mapping topsoil color. This was done either by interpolating point soil observations in Australia (Viscarra Rossel et al., 2010) and China (Liu et al., 2020), or by assigning a color to map units (polygons) corresponding to the dominant soil type in the case of USA (Soils, 2023). Both approaches have merits because they enable spatial coverage but have limitations. Maps of soil color made by point data interpolation rely on field data and reflect soil color at the time of observations (Liu et al., 2020) whereas polygon map units of soil colors are discrete surfaces and do not represent the natural continuous spatial variation of the soil. No study has mapped the soil (surface) color globally in space and time at a fine scale, which is essential for monitoring its conditions.

The availability of 40 years of satellite imagery in the visible region is

highly valuable for earth monitoring, as many terrestrial phenomena have substantially impacted topsoil properties over time. It is well-established that color is a reliable indicator of these temporal topsoil transformations. Consequently, leveraging this treasure trove of satellite data to track the evolution of topsoil color becomes a critical step in understanding both local and global shifts in surface soil conditions.

In this study, we capitalize on the integration of remote sensing data and in-situ observations of soil reflectance spectra in the visible range for topsoil color mapping. We execute this both spatially and temporally at a fine scale with a grid spacing of approximately 30 m by 30 m across the world's bare surface. We illustrate the process of extracting soil color from satellite sensors, and demonstrate how these data can elucidate the spatial variability of soil. This integrative approach aims to advance our understanding of global soil characteristics significantly. The paper is structured as follows: i) we define the ground-truth spectral data to estimate soil color, then ii) we build a mathematical transfer function to link the ground-truth color values with the Landsat data, iii) we use temporal bare surface images obtained from various Landsat sensors to map soil color, and, finally, iv) we validate the soil color map.

2. Material and methods

Estimates of the bare earth surface soil color over space and time were executed in four main steps: i) creating a bare surface image of the world between 1985 and 2020; ii) estimating color with soil spectral measurements in the visible range; iii) using these estimates as ground-truth and fitting a mathematical transfer function to convert Landsat's reflectance bands into standardized CIE-XYZ color coordinates; iv) applying this transfer function to bare surfaces images of the world. This procedure enable mapping of topsoil color over time. The workflow is summarized in Fig. 1. In the next sections, we describe each of these steps in detail.

We acquired spectra from soils of all continents to build a global soil spectral library. Approximately 30,000 soil sample spectra were

compiled and evaluated. Soil spectra were provided by many researchers from different regions, and consequently, measurement conditions were slightly different (Supplementary 1). However, we considered only the samples from locations with available information in the bare surface image. We obtained a set of 8005 spectra that met these conditions, and these are hereafter referred to as the global soil library (GSL).

2.1. Color space models on the soil spectral library

Conventional soil color measurements are done by experienced users under natural day light condition using the Munsell color chart. To simulate this condition, we used a mathematical representation of the average human chromatic response based on color matching functions $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, $\bar{z}(\lambda)$ defined by the *Commission Internationale de l'Éclairage* (CIE). Illumination was standardized by the D65 spectral power distribution (Ohta and Robertson, 2006), which corresponds to the average midday light. These matching functions were applied to the GSL in the visible range (380–780 nm) of the electromagnetic spectrum to obtain the tristimulus values in the CIE 1931 XYZ color space (Standard CIE Commission internationale de l'éclairage, 2019). These matching functions were applied to our GSL as follows:

$$X = K \int_{380}^{780} S(\lambda) R(\lambda) \bar{x}(\lambda) d\lambda \quad (1)$$

$$Y = K \int_{380}^{780} S(\lambda) R(\lambda) \bar{y}(\lambda) d\lambda \quad (2)$$

$$Z = K \int_{380}^{780} S(\lambda) R(\lambda) \bar{z}(\lambda) d\lambda \quad (3)$$

$$K = \frac{100}{\int_{380}^{780} S(\lambda) \bar{y}(\lambda) d\lambda} \quad (4)$$

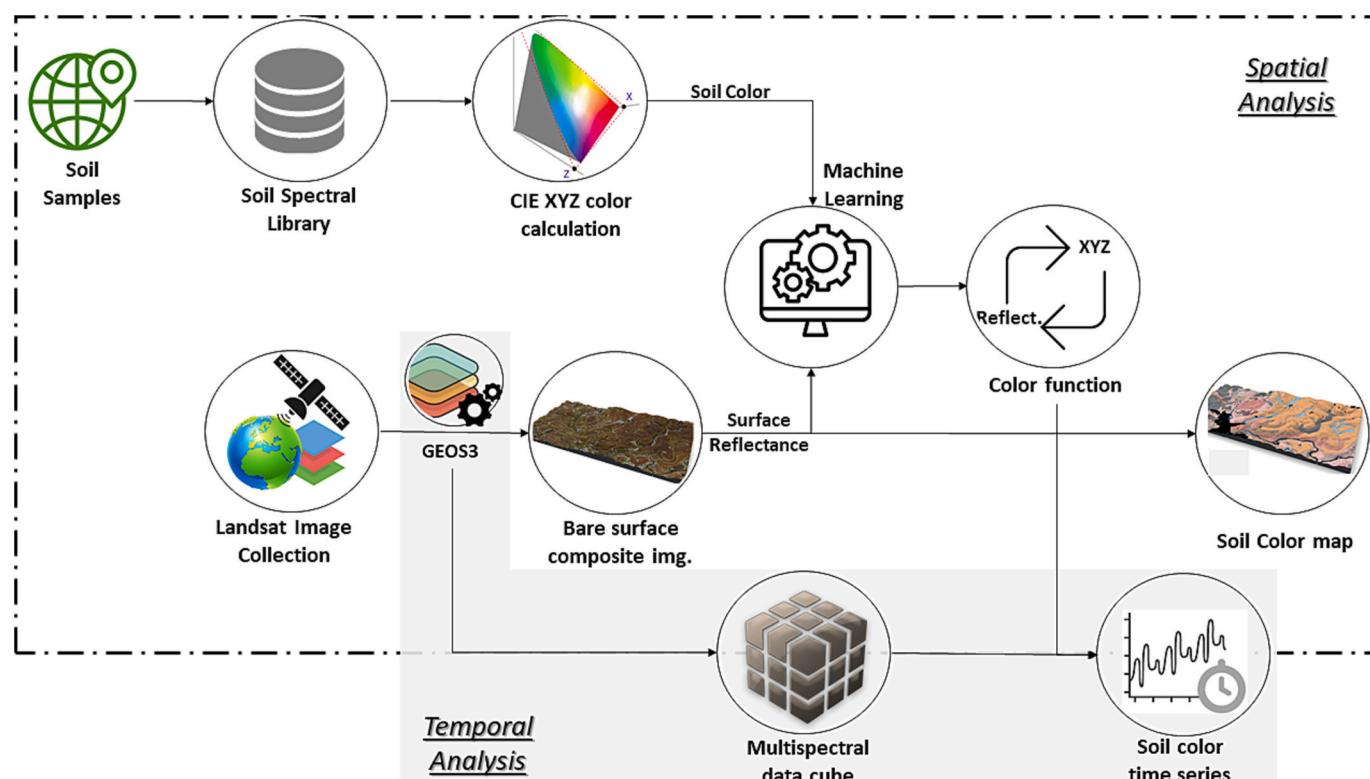


Fig. 1. Summary workflow of the procedure to obtain space and time estimates of the bare earth surface soil color. The gray shaded area relates to the temporal analysis of the remote sensing data.

where λ is the radiation wavelengths (in nm), $S(\lambda)$ is the spectral distribution of the illuminant relative power, in this case, the D65 illuminant standardized by CIE, $R(\lambda)$ is the spectral reflectance of samples, and $\bar{x}(\lambda)$, $\bar{y}(\lambda)$, $\bar{z}(\lambda)$ are the matching functions, for a standard observer with 2° field of view. The location of the spectra used in the GSL with their associated values of hue, value and chroma are shown in Fig. 2.

2.2. Spatio-temporal analysis of Earth's bare surface

We analyzed the Tier 1 Collection 1 Surface Reflectance data processed by the LEDAPS and LaSRC algorithms (USGS, 2020a, 2020b) from Landsat 4 Thematic Mapper (TM), Landsat 5 TM, Landsat 7 Enhanced Thematic Mapper Plus (ETM+), and Landsat 8 Operational Land Imager (OLI) sensors. These datasets are available through the Google Earth Engine (GEE) cloud platform (Gorelick et al., 2017; Padarian et al., 2015). Due to OLI sensors' different spectral configurations, Landsat 8 images were transformed into ETM+ following the harmonization process described by Roy et al. (2016).

The Geospatial Soil Sensing System (GEOS3) developed by Dematté et al. (2018) was used to flag bare surfaces (BS) in each available Landsat images from 1984 to 2020. These images have a 16-day temporal resolution, but bare surface reflectance cannot be retrieved regularly due to atmospheric conditions and vegetation coverage. The bare surfaces were flagged based on the greenness index (GREEN), normalized difference vegetation index (NDVI), normalized burn ratio index (NBR), and normalized burn ratio 2 index (NBR2). Bare surfaces were assigned when the following conditions were met: GREEN < 0.65, and $-0.05 < \text{NDVI} < 0.25$, and $\text{NBR} > -0.23$, and $-0.05 < \text{NBR2} < 0.15$. Clouds,

shadows, and other non-bare surface pixels were removed using the quality assessment band (QA band). Each pixel in the QA band contains a decimal value representing the surface, atmosphere, and sensor conditions that can affect the data's overall usefulness. The band is in unsigned 16-bit format, whose values are bit-packed and provide information pertaining to the following pixel conditions: fill, clear, water, cloud, cloud shadow, and snow. We created a mask to remove pixels that were not flagged as "clear" in the quality band and consequently did not represent the land surface reflectance. Further information about QA band, pixel quality indexes and values can be found elsewhere (USGS, 2020a, 2020b). The mask was then created and applied to the satellite image collection, preserving only surface reflectance data (e.g., images not influenced by clouds, snow, among others). The resulting product corresponded to a space-time cube comprising the masked images with bare surface pixels, at 30 m spatial and 16-day temporal resolutions. Based on this data cube, time series of surface reflectance (Landsat bands 1–5, 7 and 8) were retrieved and used to evaluate the temporal variability of soil color (see section 2.5).

The images within the space-time cube were then aggregated using the sparse spatiotemporal occurrences into a single composite product representing the median value of surface reflectance. This product is described herein as a Bare Surface Composite Image (BaSCI). In this case, BaSCI corresponded to natural abiotic surfaces (e.g., bare soil, sand and rock outcrops), where vegetation has a minor or no influence on the reflected signal (Dematté et al., 2020). Areas without spectral measurements were usually associated with natural vegetation fragments or soil surfaces with frequent crop residues.

We evaluated BaSCI by comparing the composite image with spectra

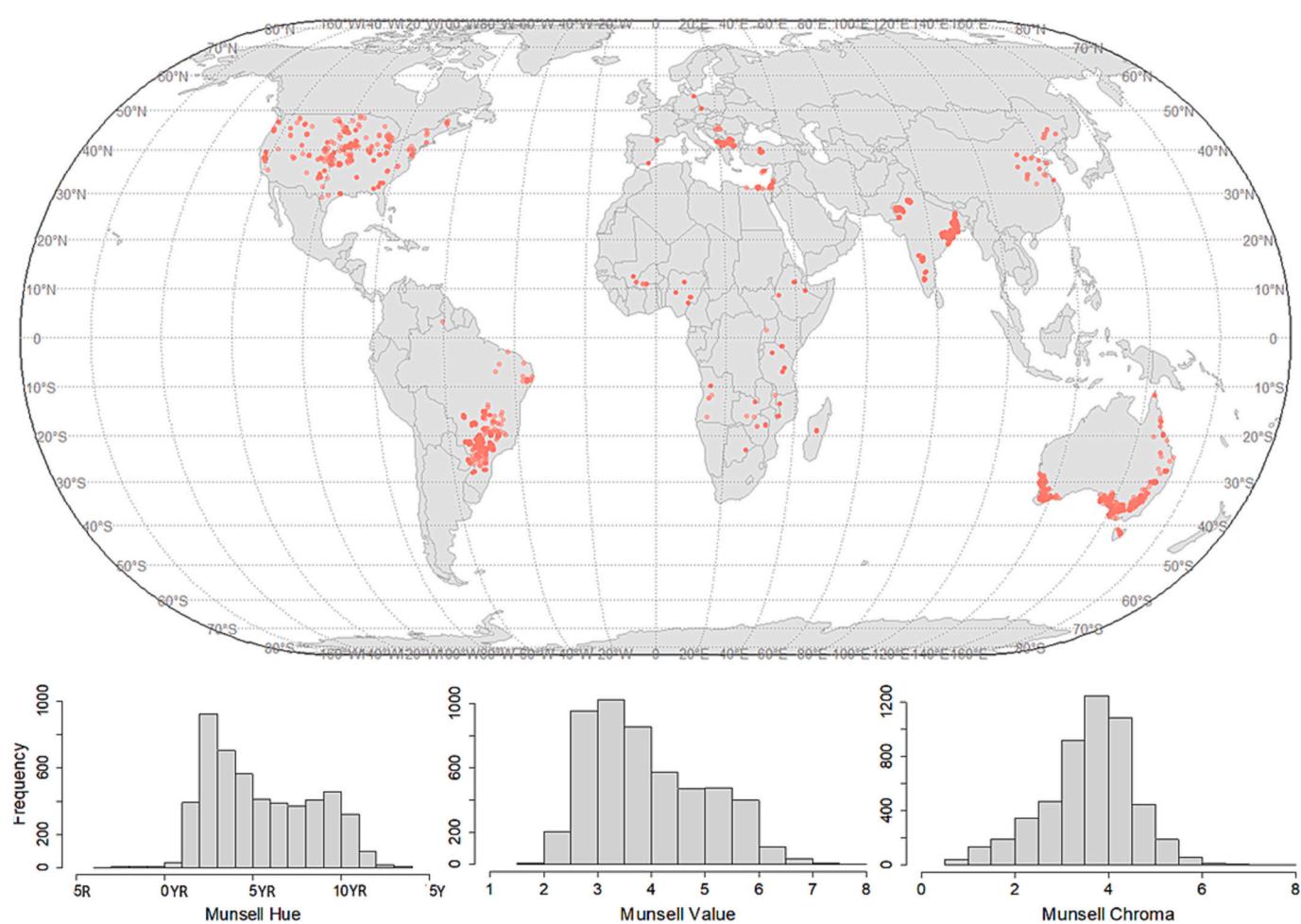


Fig. 2. Location of soil samples from the global soil library (GSL) and histograms describing the distribution of Munsell color coordinates.

from our GSL. Spectra in the visible range from the GSL were converted into XYZ color coordinates and later correlated to BaSCI data by Spearman correlation analysis. We also calculated the Euclidean distance between GSL and BaSCI spectra at each sampling location, analyzed the spatial distribution of possible outliers, and defined regions where BaSCI had higher similarity with the spectral library.

2.3. Transfer functions with Landsat data

Conversions between different color space models are well established in the literature (e.g., [Viscarra Rossel et al., 2006a, 2006b](#)), including the ones related to RGB - XYZ - Munsell systems. The visible bands from Landsat roughly correspond to the blue, green, and red spectral ranges, and if they were directly converted to other space models, the results would be overly saturated or even unrealistic ([Sovdat et al., 2019](#)). Therefore, we built a new transfer function to convert Landsat surface reflectance data (all bands) into CIE 1976 XYZ color coordinates. To fit the transfer functions, we used XYZ (from our GSL, see [Figs. 1 and 2](#)) as dependent variables and the corresponding reflectance data (from the median composite image called BaSCI) as independent variables. We also considered near infrared (NIR) and shortwave infrared (SWIR) spectral bands in the transfer function due to good results observed in previous studies ([Horvath et al., 1984; Poppie et al., 2020; Post et al., 1994](#)).

We used the random forest (RF) algorithm ([Breiman, 2001](#)) to build transfer functions for each color variable (X, Y, and Z). Tuning of the RF parameters consisted of a grid search examining a range of values of the hyperparameters: the number of independent variables randomly selected at each node (called mTry) was tested at 2, 4, 6; the minimum number of samples the terminal nodes for further splitting of a tree (called the minimum node size) was tested as 10, 30, and 50. We fixed the number of trees at 500 to obtain stable estimates ([Probst et al., 2019](#)). During the process, we validated the models with a 10-fold cross-validation, which has been described as an adequate method when dealing with spatial datasets ([Wadoux et al., 2022](#)). We calculated the mean absolute error (MAE), the root mean squared error (RMSE), the coefficient of determination (R^2) and the concordance correlation coefficient (CCC). Both MAE and RMSE have an optimal value as close to zero as possible. The MAE and RMSE represent the magnitude of the error. The R^2 and CCC evaluate the deviation from the line of equality and have an optimal value of 1.

2.4. Global soil color maps

The transfer functions were applied to each pixel in BaSCI, resulting in color maps. The resulting XYZ tristimulus values of color were further converted to the CIELab and Munsell parameters. These systems are commonly used in soil science. We did not perform direct conversion and mapping from Landsat bands to CIELab and Munsell because the process is time-consuming and would require a large space to store all information. Instead, all calculations were implemented on a cloud computing system (i.e., Google Earth Engine), so conversion between XYZ and other color systems could be done in real-time, according to users' specifications.

2.5. Time-series of soil color

Reflectance time series of bare surfaces were retrieved from the space-time cube (obtained by the GEOS3 system) at a specific location in Brazil. Then, we applied the transfer functions to these time series to estimate the temporal variation of topsoil color. The color time series were then converted to CIELab and Munsell systems and used in a temporal analysis. We calculated the CIELab distances (CIE76) between the temporal median (calculated by GEOS3) and each color prediction in our time series. The CIE76 allowed us to evaluate differences between single-date predictions and the multitemporal median. Because this

color space is constructed in a cartesian coordinate system, the distance between colors is calculated as:

$$\Delta E_{ab}^* (c_1, c_2) = \sqrt{(L_2^* - L_1^*)^2 + (a_2^* - a_1^*)^2 + (b_2^* - b_1^*)^2}, \quad (5)$$

where ΔE_{ab}^* is the measure of change in visual perception between color c_1 and c_2 ; L^* , a^* and b^* correspond to the color coordinates from CIELab. The reported just-noticeable difference (JND) for this metric is 2.3 ([Sharma et al., 2002](#)), which means that the human eye does not detect differences lower than 2.3 and, consequently, cannot be considered an error.

3. Results and discussion

3.1. World bare surface composite image vs. GSL

We created the BaSCI by evaluating each pixel, aggregating the barest soil conditions from 1985 to 2020 and calculating its pixel-wise median to obtain a soil color map. [Fig. 3](#) shows the global BaSCI related to arable lands (exposed by tillage practices), deserts, and rocks. Gaps in the image were mostly associated with natural vegetation or no-till agriculture. We could cover 38.5% of the earth's surface and 82.2% (1.53 billion ha) of the croplands mapped by [Thenkabail et al. \(2021\)](#).

Correlation coefficients between BaSCI and XYZ color coordinates derived from our GSL ([Table 1](#)) were all significant ($p < 0.0001$) and indicated moderate to strong linear relationships between the two datasets. Higher values of correlation coefficients were observed between XYZ color and BaSCI visible bands (B1, B2 and B3), where most r values range from 0.7 to 0.8. The NIR band (B4) were correlated with color coordinates, with values ranging between 0.64 and 0.7. SWIR bands (B5 and B7) had a moderate correlation with r values varying between 0.50 and 0.55. Such independent comparison indicates that the composite image adequately represents the variability of soil color at a relatively fine scale (e.g., 30 m resolution pixels).

In past years, some studies have derived bare surface images and evaluated their agreement with vis-NIR-SWIR reflectance spectra (350–2500 nm) measured in laboratory conditions. These studies were performed for smaller areas in Brazil or European countries, and used standardized libraries (e.g., [Dematté et al. \(2018\); Safanelli et al. \(2020\)](#) and [Tziolas et al. \(2020\)](#)). They found correlation coefficients between 0.49 and 0.88, and the highest values were reached in a local-scale study ([Dematté et al., 2018](#)) with highly controlled soil sampling and laboratory measurements.

The correlation analysis provided insights into the relationship between datasets, but the method is not very sensible to outliers and does not describe their geographical distribution. Therefore, we evaluated the Euclidean distance (dissimilarity) between GSL and BaSCI at each sampling location. Samples with higher dissimilarity are mostly likely related to outliers in GSL or poor performance of BaSCI at the local scale ([Fig. 4](#)). Euclidean distances were between 0.005 and 0.11 for 87% of sampling points. Samples with higher dissimilarities (> 0.11) had a sparse occurrence in the globe but were most frequently observed in Brazil, India and Australia ([Fig. 4b,c,d](#)).

Dissimilarity had a clear spatial pattern, particularly in India, which may be related to BaSCI limited potential to detect bare pixels in some specific environmental conditions. Samples in a semi-arid northern India region (Rajasthan state; [Santra et al., 2015](#)) ([Fig. 4c](#)) had the lowest dissimilarity between GSL and BaSCI. Higher dissimilarity values occurred in eastern India (West Bengal and Northern Odisha), where the monsoon climate favors the surface coverage by crops, residues or grass. [Roberts et al. \(2019\)](#) compared a bare surface image of Australia and a national spectral library and found better results in sparsely vegetated areas from semi-arid regions.

Different sampling protocols and equipment may also influence dissimilarities in Indian data. This study's soil samples originated from

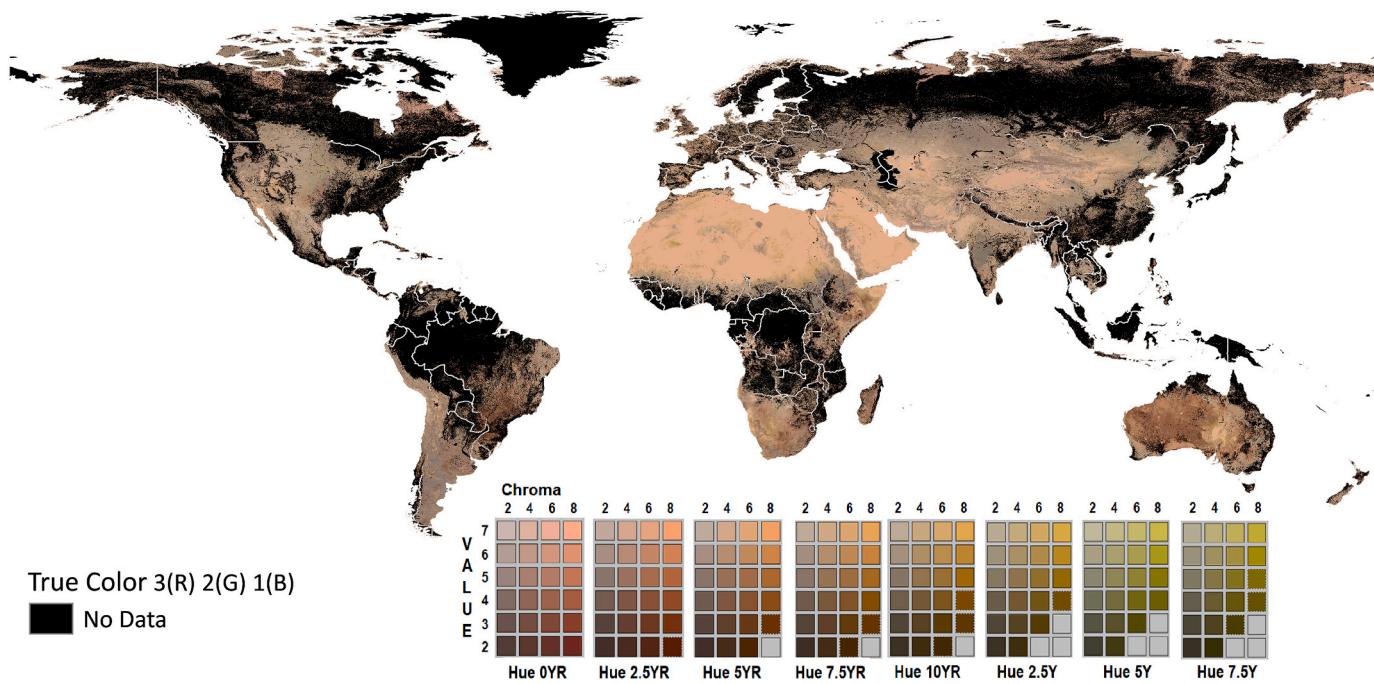


Fig. 3. BaSCI true color composite ($R = \text{Band } 3$; $G = \text{Band } 2$; $B = \text{Band } 1$) of the world's bare surfaces, and the Munsell color chips.

Table 1

Spearman correlation coefficients between soil reflectance measured in the laboratory and by Landsat sensors (spectral ranges according to TM). Laboratory spectra are represented by XYZ color parameters.

Landsat bands	Spectral range (nm)	Soil color coordinates (laboratory)		
		X	Y	Z
B1 (Blue)	450–520	0.79	0.80	0.81
B2 (Green)	520–600	0.78	0.79	0.77
B3 (Red)	630–690	0.69	0.69	0.64
B4 (NIR)	770–900	0.71	0.71	0.67
B5 (SWIR1)	1550–1750	0.65	0.65	0.62
B7 (SWIR2)	2090–2350	0.55	0.54	0.51

three datasets and multiple field surveys (Sarathjith et al., 2014; Santra et al., 2015). Errors in the measurements could be related to different sampling tools and users. This problem has been extensively discussed in the literature (van Leeuwen et al., 2022; Batjes et al., 2017) and it is an issue that affects all samples from our GSL, as well as other works that used data from multiple sources.

3.2. Monitoring Earth's topsoil color

While BaSCI spectral bands were related to the CIE XYZ data, there were still mismatches. Thus, we created a model to convert reflectance bands into CIE XYZ color coordinates using the RF machine learning model. Using a 10-fold cross-validation, we achieved a good agreement between predicted and observed values, as shown by a R^2 close to 0.66 and a CCC close to 0.8 (Fig. 5). This is similar to or better than soil color models derived at continental scales in past studies (e.g., Viscarra Rossel et al., 2010; Liu et al., 2020). These models were then applied to BaSCI to obtain the topsoil XYZ maps. We also derived maps using the Munsell color system, which is standard in soil science.

The Munsell color maps for hue, value and chroma are shown in Fig. 6. Low values are usually due to the presence of organic matter or Mn oxides. Reddish, yellowish, and orange hues (in some cases with high chroma) are typically caused by oxyhydroxides as hematite, goethite, lepidocrocite or ferrihydrite (Schwertmann, 1993). Grayish colors (low chroma) are related to the absence of the pigments

mentioned above or mixture with light or whitish components (high value), such as carbonates (calcite, dolomite) and soluble salts (gypsum, tenardite, epsomite, halite). Consequently, most key processes in the genesis include melanization, clay illuviation, rubefaction, braunification, redox, pseudogleyization, carbonation or salinization (Bockheim and Gennadiyev, 2000) which are reflected in soil color.

Most soil classification systems are based on the morphogenetic of soils. We, therefore, compared our topsoil color map with the soil map of the world (FAO-UNESCO, 1974) (Fig. 7). Similarities between both maps indicated a consistent spatial pattern of color predictions and demonstrated the capacity of such a product to support the development of thematic maps at finer resolution, such as soil types, carbon pools or even lithology (Demattè et al., 2020). The visual comparison over South America, Africa, India, Australia and the European continent showed consistencies of soil color variation with soil types (Fig. 7).

In South America, the predicted chroma allowed the discrimination between Yermosols and Regosols in the Argentinian temperate grassland region, where areas with low chroma agree with the regional light-colored topsoil (Fig. 7a). In Southern Africa, Arenosols were correctly discriminated from chromic Luvisols and Fluvisols by Munsell hue and chroma, presenting a relatively abrupt transition in the map (Fig. 7b). A large area in India was accurately described by hue, value and chroma (Fig. 7c). In this case, the soils corresponded to Vertisols and Inceptisols derived from basaltic parent material, which can be detected by the low value and chroma.

In Eastern to Central Australia, Yermosols (soils under arid regions, Arenosols, Technosols and Kandosols in the Australian soil classification system) were highlighted by higher Munsell value (Fig. 7d), which is in agreement with the color characteristics according to the WRB classification criteria (IUSS Working Group WRB, 2015).

In Brazil, soils with red hue were observed in the Southern part of the country, while in the Northern region, yellow hue soils were detected (Fig. 7e). Originating from volcanic rocks, areas in Southeastern Brazil are closely related to Ferralsols, with a darker red-colored topsoil. On the other hand, Regosols and Yermosols in arid regions had higher albedo and brighter colors.

Another interesting pattern detected by our maps is the soil color transition from Africa to the European continents. Areas with lower

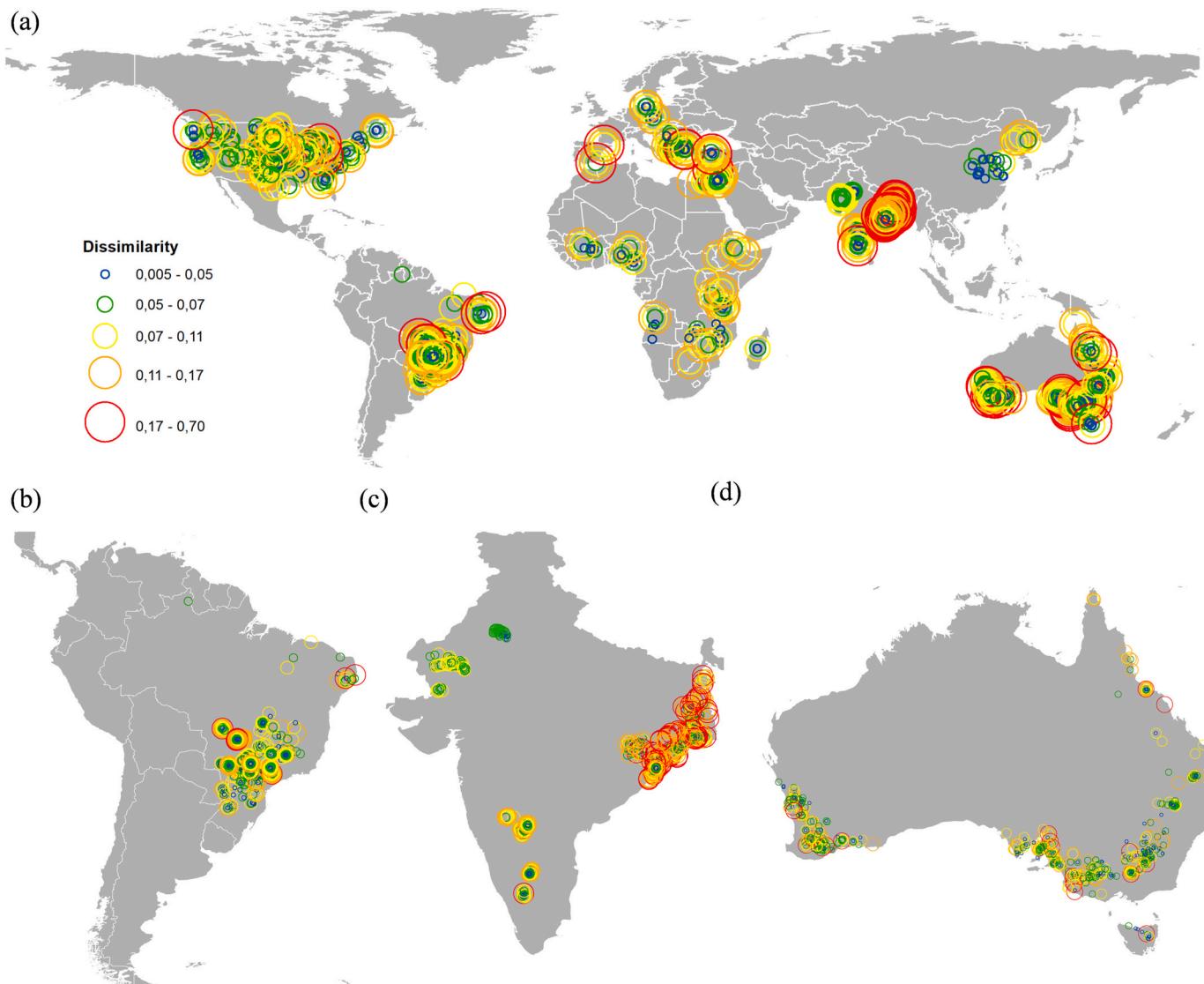


Fig. 4. Euclidean distance (dissimilarity) between GSL and BaSCI at each sampling location.

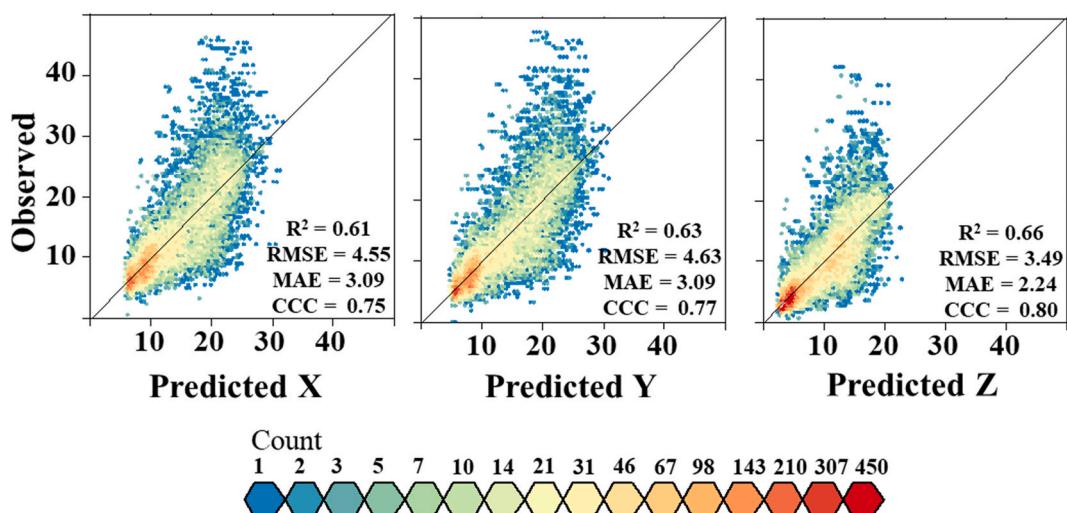


Fig. 5. Scatterplot of the results from the 10-fold cross validation of the X, Y and Z color coordinates prediction models using locations (≈ 8000) that matched the GSL and the BaSCI.

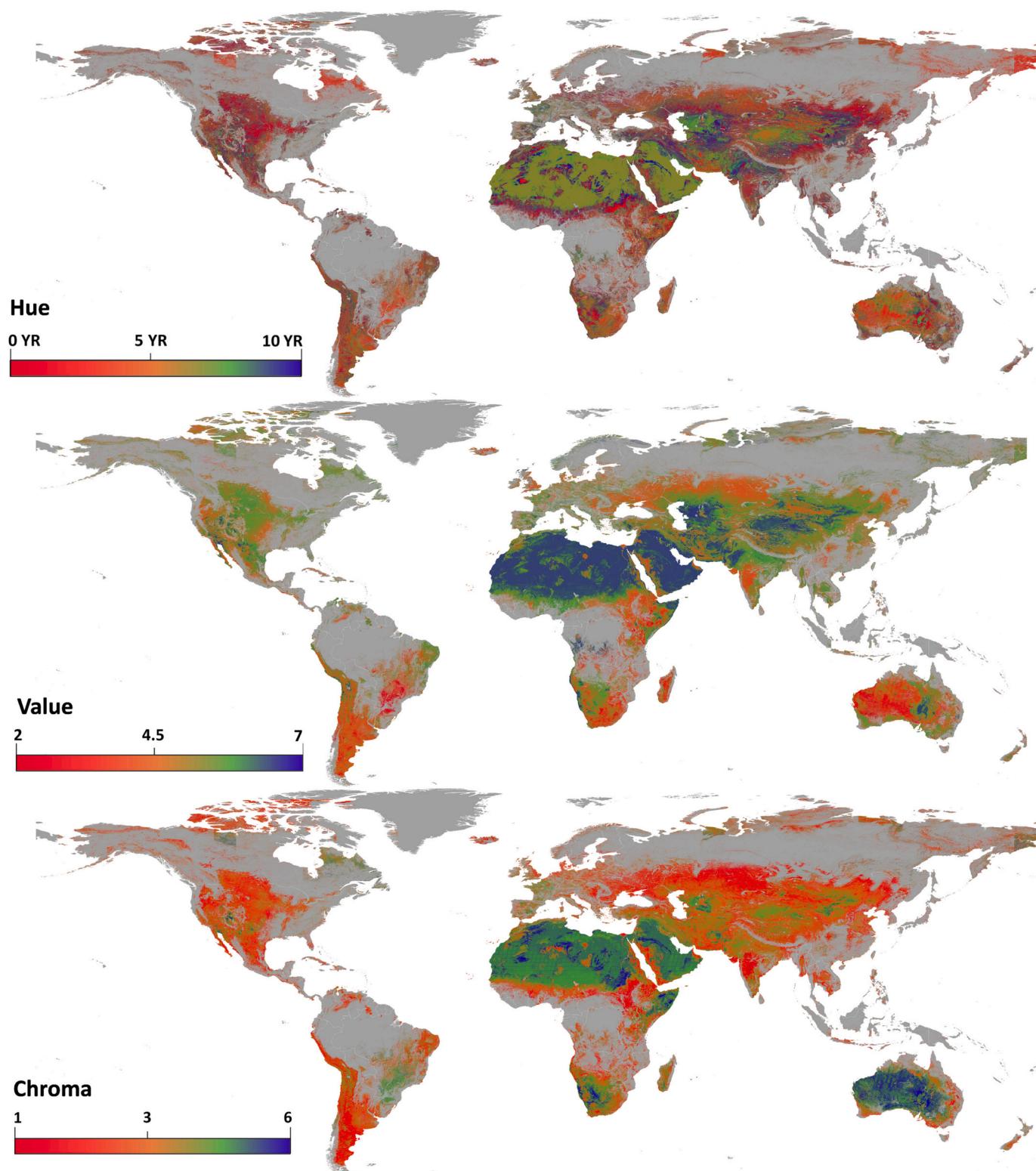


Fig. 6. Maps of the Munsell color coordinates.

Munsell value were observed in soils with higher organic carbon (e.g., Chernozems), while sandy soils (Yermosols) from dry regions (North of Africa) had higher value.

The similarity between soil color and soil types demonstrates the cogency of the global map. Remote sensing measurements only interact with a few millimetres of soil surface, while soil class maps are based on our conceptual understanding of the whole soil profile. This

corroborates past studies, which found that it is possible to obtain valuable information on subsurface variability from surface measurements (Lagacherie et al., 2013). This is especially true if remotely sensed spectra are combined with variables related to relief or other drivers (e.g., terrain attributes, climate, land use, and water table regime), that have been influencing pedogenetic processes and resulting in variability on deeper layers (Mendes et al., 2019).

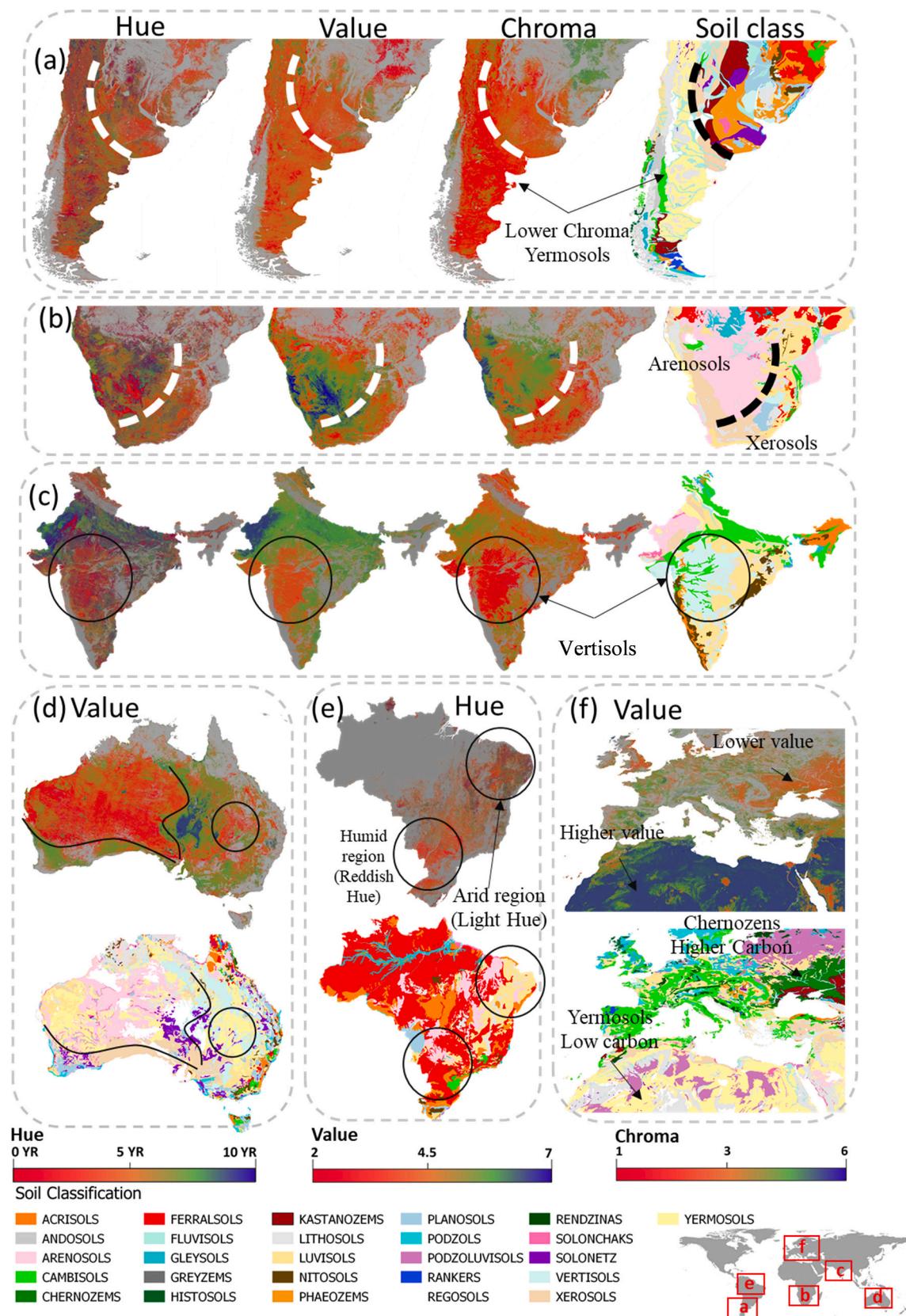


Fig. 7. Comparisons between predictions of soil color (hue, value, chroma) and the soil map in South of South America (a), South of Africa (b), India (c), Australia (d), Brazil (e) and Europe, North of Africa and Middle East (f).

Soil color maps proved to be not only suitable for global analysis, but they also delivered good results at a local scale (Figs. 8, 9 and 10). For example, in the Brazilian municipality of Luís Eduardo Magalhães (Bahia state) (Fig. 8), soil colors had high spatial variability, with topsoil colors transitioning from 2.5YR (in the western area) to 10YR in the eastern part of the site. Previous soil surveys of this region (Cooper et al., 2005) indicated that yellow hues are related to Arenosolos, while the reddish topsoil corresponded to Ferralsols.

Fig. 9 shows the topsoil color variability in an arid region on the border between Jordan and Syria. In this case, soils at the northern of the site were allocated to Vertic Cambisols, while in the south, there were Calcic Xerosols (FAO-UNESCO, 1974). An important characteristic of many Xerosols is the Ochric horizon, which resulted in a light-colored topsoil. Our color maps correctly discriminate such characteristics, indicating yellow Munsell hue (~10YR) and high values in areas dominated by the Xerosols.

3.3. Temporal variability of soil color

Topsoil variability in time is related to natural or human-induced changes on the earth's surface, and depending on the local conditions, it can greatly impact color predictions. This is exemplified in an area with two locations, 200 m apart, in Southeastern Brazil. The soils at the two sites are the same (i.e., Ferralsol) but had contrasting land uses (Fig. 10). The first site was always bare while the second site was under sugarcane, which was only occasionally bare. The greatest difference in

these time series was the number of available data. While soil in the sugarcane field had color estimated on 52 dates in the last 20 years, time series from the bare site had 244 observations. Color time series in crop fields are only available when the soil was tilled, which, in this case, corresponds to the period of transition between dry and wet seasons in Brazil (September–December). Besides, for both sites, cloud coverage is an important factor limiting data acquisition (Demattê et al., 2018). This further reduced the number of dates at which the bare surface image could be acquired.

The time series presented similarities in Munsell coordinates' temporal variability and median values (Fig. 10). Hue and value were one unit higher in the crop field than in the bare site, while the median chroma was 1.5 units higher in the bare area. Chroma on the bare site (Fig. 10f) also had the largest temporal variation, ranging from 3 to 10. Measurements on the bare site were acquired during the whole year and consequently at very different soil moisture conditions. Coelho and Audi (1964) evaluated the influence of water content on soils and found a variation of 2 units in chroma. Similarly, Post et al. (1994) indicated that soil wetting significantly influences chroma and variations can be positive or negative, depending on the soil types.

We calculated the CIE76 metric indicating noticeable differences in color (Fig. 11) to compare single-date color estimates and a long-term median. The results indicated that in nearly all cases, the variations were sufficient to be perceived by the human eye (Fig. 11). CIE76 distance in the dry (Fig. 11a, c) and wet (Fig. 11b, d) seasons did not have large differences, suggesting that soil moisture was not the only factor

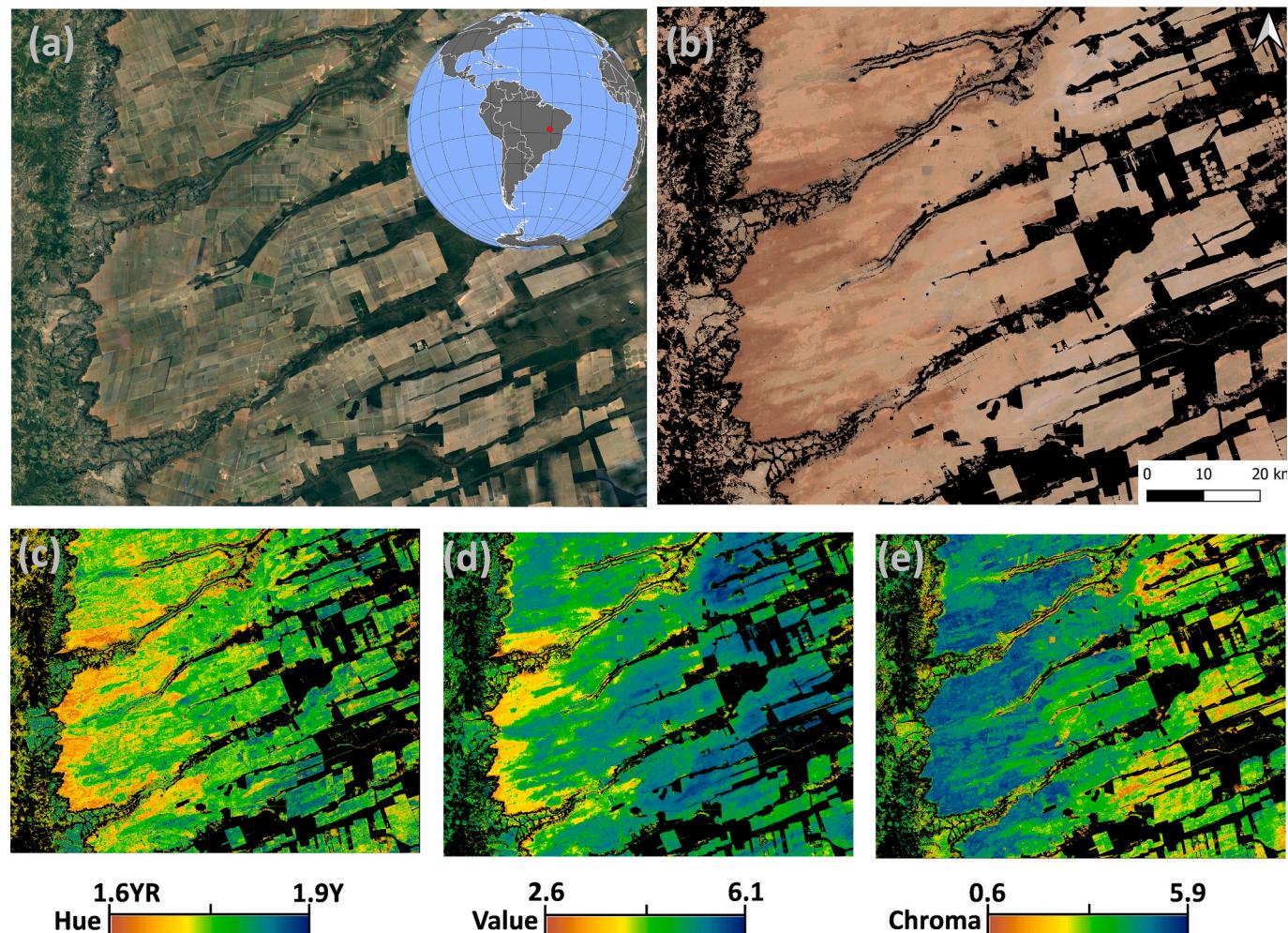


Fig. 8. High resolution satellite image (Google Earth) (a), bare surface image (true color composite) (b), and Munsell hue (c), value (d) and chroma (e) from the Municipality of Luís Eduardo Magalhães.

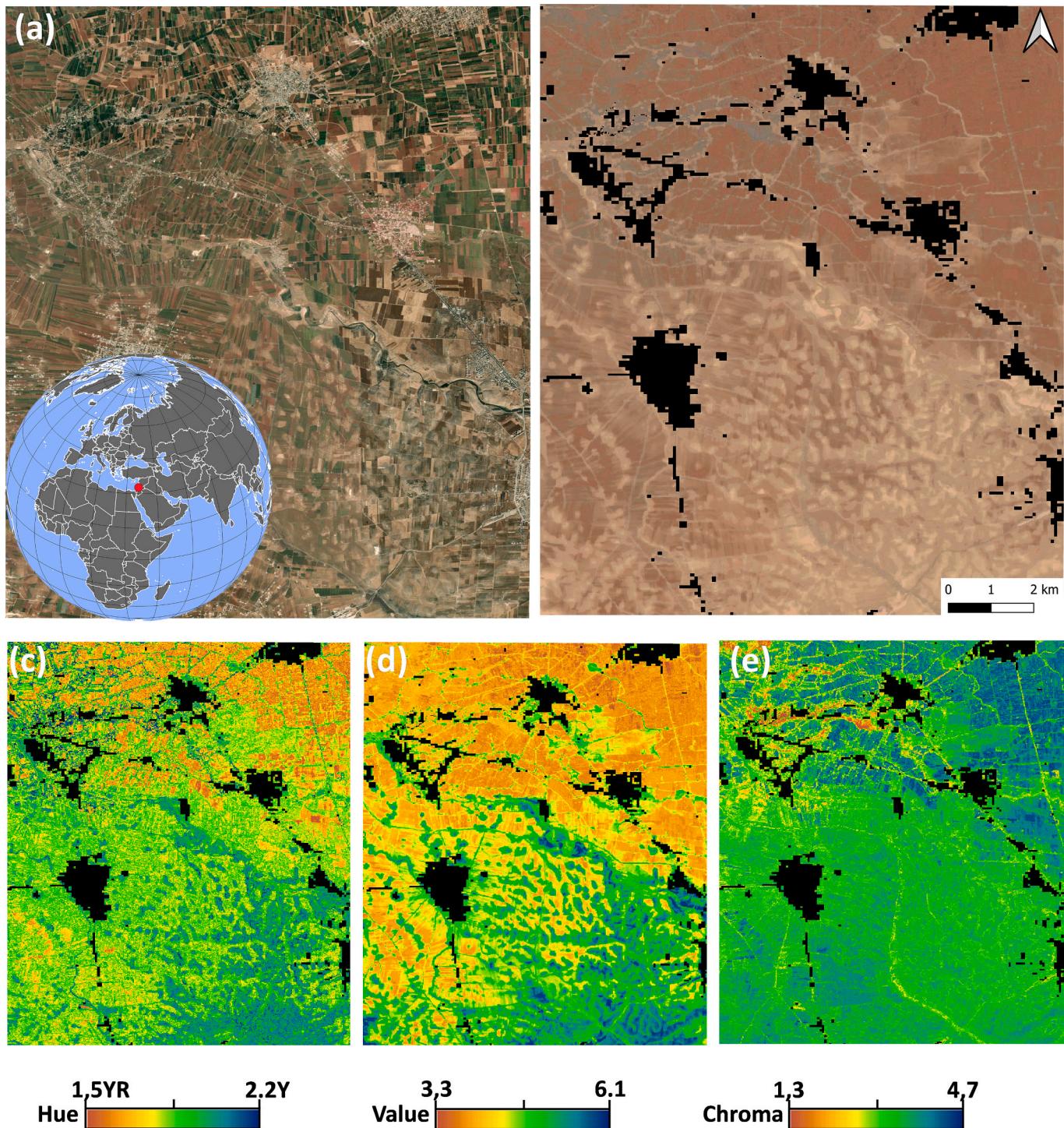


Fig. 9. High resolution satellite image (Google Earth) (a), bare surface image (true color composite) (b), and Munsell hue (c), value (d) and chroma (e) from a site at the border between Jordan and Syria.

influencing reflectance data in these soils. According to Diek et al. (2017), the temporal variability and deviation from the central measures might be associated not only with moisture, but also to differences between the Landsat sensors, atmospheric correction algorithms and changes in surface roughness (Diek et al., 2017). This finding indicates the importance of employing GEOS3, or similar methods to derive a central measure (e.g., median or mean) from multitemporal image collections instead of predicting soil attributes from unitemporal data. Averaging satellite images over a specific time frame provides a stable measurement (Safanelli et al., 2020), less affected by external factors,

such as surface conditions and uncertainties in sensors' measurements or image pre-processing algorithms.

3.4. Assumptions and limitations

We hypothesized that discrepancies between the BaSCI and the GSL stemmed from surface-related attributes, such as surface roughness, soil crust, and moisture, which likely impact measurements from spaceborne sensors (Prudnikova and Savin, 2021; Savin and Vindeker, 2021). The GSL measurements were conducted in a laboratory under

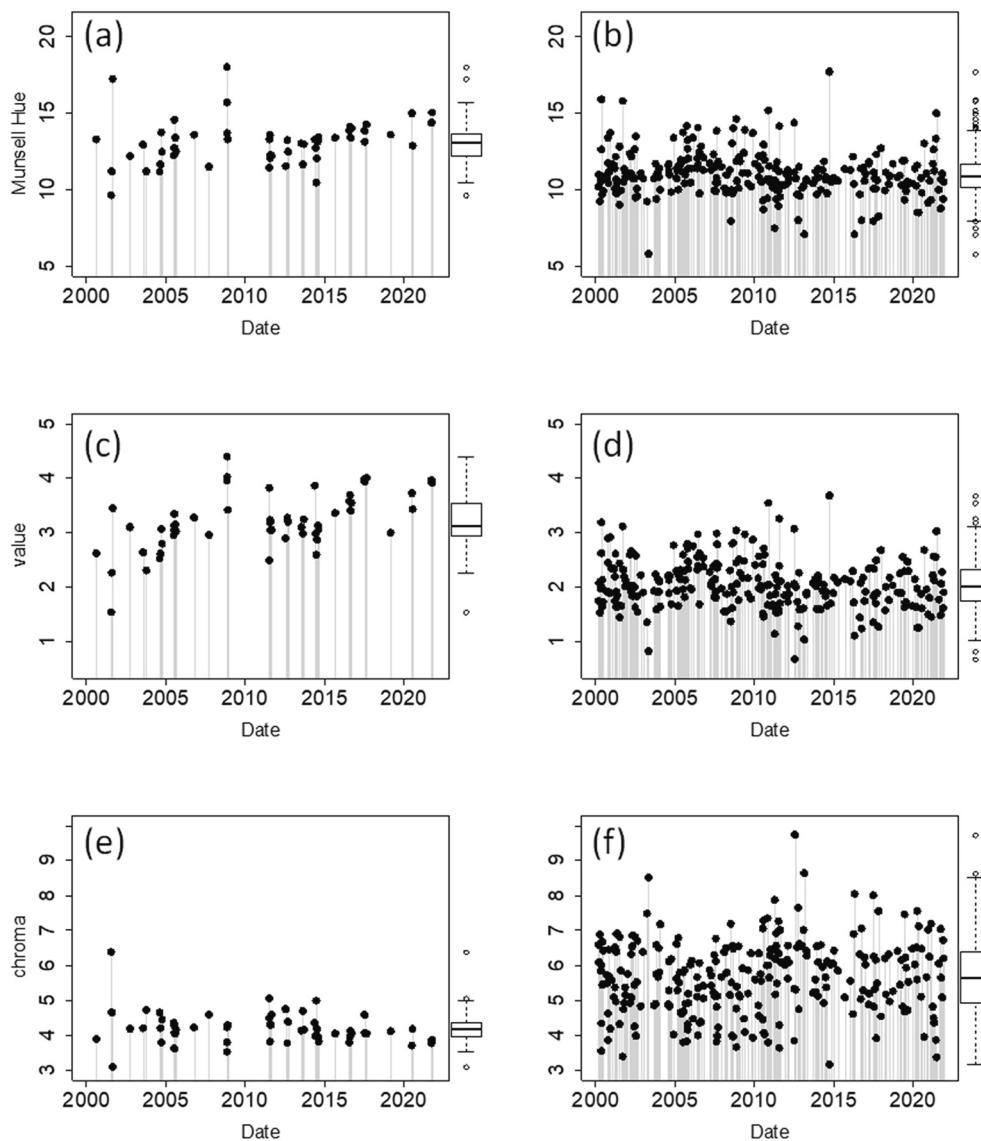


Fig. 10. Temporal comparison between the Munsell hue, value and chroma from an agricultural land occasionally bare (a,c,e) and from a constant bare surface (b,d,f). Locations are 200 m apart from each other and have the same soil type.

standardized conditions with dry soils, while space-borne sensors gathered soil data under field conditions (Francos and Ben-Dor, 2022). In addition, the remote sensing signal is also influenced by residual atmospheric conditions even after corrections (Anderson et al., 2005), bidirectional reflectance distribution, sensor spectral resolutions, geometric distortions, and the spectral mixture of features (Mulder et al., 2011; Richter and Schläpfer, 2002). Differences may also be related to the scale or support of measurements because satellite-based pixels averaged the signals over an area, while GSL measurements were performed on a soil aliquot in the order of a volume of a few cm³.

Another issue is related to BaSCI and the methodology used to flag bare surface pixels. The vegetation indexes and their thresholds have been tested and adapted to many countries and their different climate, land uses and soils. In our work, broad thresholds were used to produce a bare surface image at the global scale. Therefore BaSCI might not fully represent the soil spatial variability at regional or local scales. External factors influencing the quality of spectral data are primarily related to moisture, surface roughness and crop residues (Mulder et al., 2011). An increase in soil moisture results in higher light absorption due to the refraction index of water molecules (Nolet et al., 2014). Consequently, when soils are wet, they exhibit lower reflectance at all visible spectrum

wavelengths and, thus, lower Munsell values. The effect on the hue parameter is more complex, and the trend depends on the mineral composition (Bedidi et al., 1992). Roughness is related to irregularities in soil surface, which produce shadow areas, where light does not directly reach the surface. The reflectance spectrum of ploughed soils, characterized by a rough surface, exhibits lower values compared to the reflectance associated with sunlit fragments (Cierniewski and Guliński, 2010). In some situations, the soil surface might decrease by 25% in reflectance after being ploughed (Matthias et al., 2000).

It is generally challenging to find completely bare soils in the field. The soil surface is usually partially or completely covered by vegetation and organic and inorganic debris. Bartholomeus et al. (2007) indicated that a more than 20% vegetation cover prevents accurate soil properties estimation with remote sensing. In other cases, only a few percent fractional vegetation cover was sufficient to decrease the accuracy of predictions dramatically (Bartholomeus et al., 2007). Dry vegetation greatly influences the albedo of the vis-NIR region, due to the presence of cellulose and lignin (Ben-Dor et al., 2018). Murphy and Wadge (1994) found that crop residues greatly impact soil spectra.

The influence of external factors on BaSCI spectra was reduced by defining the median bare surface signal over long periods. In our case,

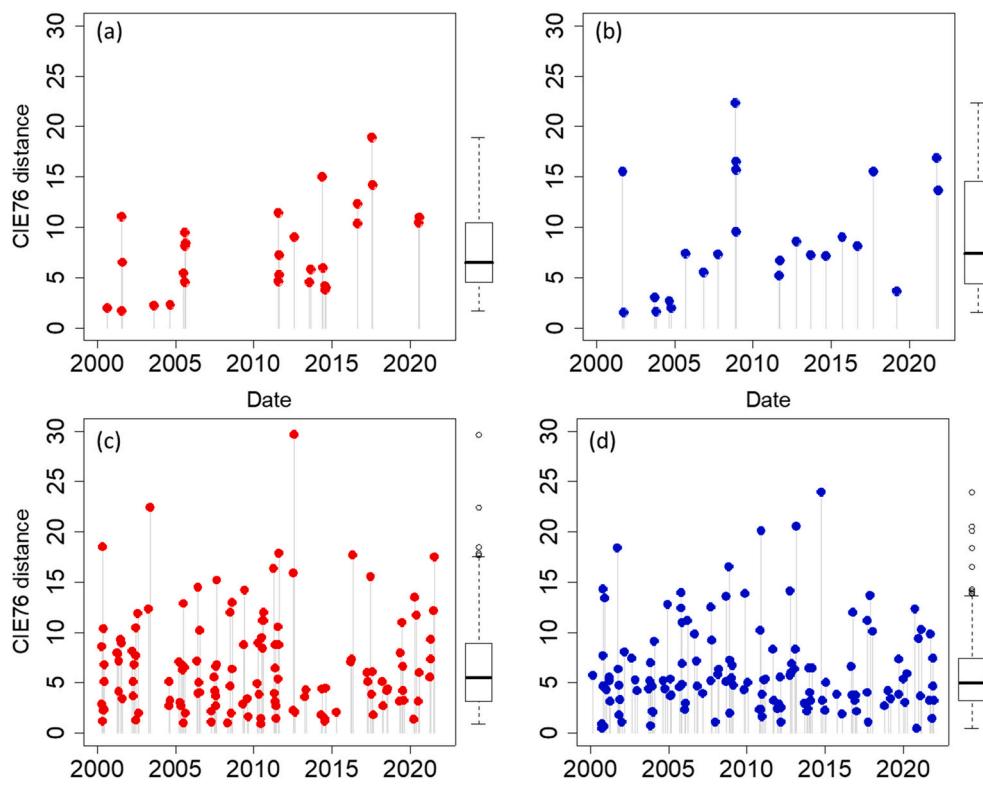


Fig. 11. CIE76 metric indicating the just-noticeable differences in color measurements from an agricultural land in wet (a) and dry (b) season and from a constantly bare surface in wet (c) and dry (d) season.

we used a Landsat collection with data from the last 35 years. Some studies have indicated that working with large image collections improves bare surface coverage, increasing the mapped area and the quality of topsoil reflectance (Shi et al., 2022; Zep et al., 2023). The problem with a long-term BaSCI is the temporal mismatch with the soil samples. GSL samples represent soil conditions at a specific date, while BaSCI reflectance is a median value that may or may not correspond to the same dates of field sampling campaigns. Several earth surface processes, e.g., soil erosion, can rapidly change topsoil conditions. While these changes could be easily observed in an early soil survey and sampling, BaSCI could not detect this change immediately.

While our approach facilitated a comprehensive characterization of soil's spatial variability, certain issues warrant consideration. One significant concern pertains to the primary database, the Global Spectral Library (GSL). The GSL library has a limited spatial coverage, and consequently there are topsoil conditions that are not represented by our data. In the American continent, soil samples are concentrated in two countries, the United States and Brazil. There are no samples in the southern Africa and Sahara desert, and only a few observations in the Middle Eastern. Asian soils are represented by samples from India and eastern China, while in the Australian continent there are no information in the northern and central regions. According to FAO's world soil map, 67% of GSL samples correspond to 4 soil classes, i.e., Ferralsols, Arenosols, Luvisols and Acrisols (Supplementary 2). A fundamental problem in global soil mapping is the lack of well distributed point observations within the soil property geographic and features space (Poggio et al., 2021). In general, least sampled areas tend to have highest prediction uncertainties. This can be explained by the models performing less accurately within areas that are under-represented by the calibration dataset. This is specially true in regions with large spatial variation, imposing difficulties for predictions. Besides that, the GSL consolidates data from a multitude of field surveys conducted over the years. Despite our efforts to include only samples (and spectra) that met specific criteria in the GSL, soil sampling protocols were not

standardized. Consequently, methods such as individual or composite sampling, depth, date and time of sampling, tools used, and georeferencing precision varied across locations. Given the considerable variation in the GSL samples, which seldom align with the spectral data in BaSCI, this inconsistency might have impacted the performance of our predictive models.

4. Conclusions

This study showcased the capability of integrating topsoil spectral data collected from both ground-based and orbital sensors with machine learning techniques to create high-resolution global digital soil maps. The validation of the models demonstrated strong agreement with ground-truth data, despite several limitations. These limitations included the lack of representativeness in the database, challenges associated with non-standardized soil sampling and spectral measurements, and, most notably, the considerable spatiotemporal variability of the soil measurements.

The intrinsic relationship between electromagnetic spectrum reflectance and soil chromophores underpins the positive outcomes obtained in this study. This connection ensures the reliability of our proxy, BaSCI, for digitally mapping topsoil color, along with other crucial environmental attributes. Moreover, the fine-scale, time-averaged soil color maps can also serve as covariates for digital soil mapping, offering comprehensive insights into the soil's chemical, physical, and biological properties.

Our methodology and analysis provide reliable knowledge that could benefit national and continental-level soil assessments. By performing temporal analysis and monitoring changes in soil color over time, one could identify areas experiencing rapid changes and raise alerts for potential soil degradation or other environmental risks. This understanding can help inform decisions about land management, agriculture, environmental protection, and urban planning.

Author contributions

Conceptualization, R. Rizzo and J.A.M. Dematté; development and coding of the method, R. Rizzo, A.M.R. Gómez, N A Rosin, J.T.F. Rosas, L.T. Greschuk; data curation, J.A.M. Dematté, R. Rizzo, B.S. Das, B. Malone, N. Tziolas, N. Tsakiridis, K. Karyotis, N. Samarinis, E. Kalopesa, A. Wadoux, A. Gholizadeh, K. Shepherd, S. Chabrilat, R. Milewski, E. Vaudour, C. Wang, El Sayed Said M., E. Ben-Dor, N. Francos, I. Savin; writing-original draft preparation, R. Rizzo, A. Wadoux, J.A.M. Dematté, R Poppiel, J. L. Safaneli, H. Belllinaso, F. Terra, N.E.Q. Silvero, M.V. Ballester, P.R. Fiorio; writing-review and editing, A. Wadoux, R. Rizzo, J.A.M Dematté, R.R. Poppiel, V. Barrón, E. Ben-Dor, N. Francos, B. Minasny, N.A. Rosin; visualization, R. Rizzo; funding acquisition, J.A.M. Dematté. All authors have read and agreed to the published version of the manuscript.

Funding

This research was funded by the following Brazilian institutions: São Paulo Research Foundation (FAPESP), grant number 2014/22262-0, 2016/26176-6, 2018/23760-4; National Council for Scientific and Technological Development (CNPQ), grant number 150319/2021-5. MHES RF (075-15-2022-321). For the purpose of Open Access, a CC-BY public copyright licence has been applied by the authors to the present document and will be applied to all subsequent versions up to the Author Accepted Manuscript arising from this submission.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Data availability

The soil color maps generated during the current study are available in raster format at the following link: <https://esalqgeocis.wixsite.com/english/thematic-products>. The code used to generate the final soil maps is available from the first author upon reasonable request.

Acknowledgements

The authors are grateful to the Soil Science Department in Luiz de Queiroz College of Agriculture, University of São Paulo for funding and providing the facilities for the development of this research. The authors are also grateful to the Geotechnologies in Soil Science group for providing the soil database.

Appendix A. Supplementary data

Supplementary data to this article can be found online at <https://doi.org/10.1016/j.rse.2023.113845>.

References

- Adderley, W.P., Simpson, I.A., Davidson, D.A., 2002. Colour description and quantification in mosaic images of soil thin sections. *Geoderma* 108, 181–195. [https://doi.org/10.1016/S0016-7061\(02\)00123-4](https://doi.org/10.1016/S0016-7061(02)00123-4).
- Anderson, M.C., Norman, J.M., Kustas, W.P., Li, F., Prueger, J.H., Mecikalski, J.R., 2005. Effects of vegetation clumping on two-source model estimates of surface energy fluxes from an agricultural landscape during SMACEX. *J. Hydrometeorol.* 6 (6), 892–909.
- Barrett, L.R., 2002. Spectrophotometric color measurement in situ in well drained sandy soils. *Geoderma* 108, 49–77. [https://doi.org/10.1016/S0016-7061\(02\)00121-0](https://doi.org/10.1016/S0016-7061(02)00121-0).
- Bartholomaeus, H., Epema, G., Schaepman, M., 2007. Determining iron content in Mediterranean soils in partly vegetated areas, using spectral reflectance and imaging spectroscopy. *Int. J. Appl. Earth Obs. Geoinf.* 9, 194–203. <https://doi.org/10.1016/j.jag.2006.09.001>.
- Batjes, N.H., Ribeiro, E., Van Oostrum, A., Leenaars, J., Hengl, T., Mendes de Jesus, J., 2017. WoSIS: providing standardised soil profile data for the world. *Earth Syst. Sci. Data* 9 (1), 1–14.
- Bechtel, R., Rivard, B., Sánchez-Azofeifa, A., 2002. Spectral properties of foliose and crustose lichens based on laboratory experiments. *Remote Sens. Environ.* 82, 389–396. [https://doi.org/10.1016/S0034-4257\(02\)00055-X](https://doi.org/10.1016/S0034-4257(02)00055-X).
- Bedidi, A., Cercelle, B., Madeira, J., Pouget, M., 1992. Moisture effects on visible spectral characteristics of lateric soils. *Soil Sci.* 153, 129–141.
- Ben-Dor, E., 1997. The reflectance spectra of organic matter in the visible near-infrared and short wave infrared region (400–2500 nm) during a controlled decomposition process. *Remote Sens. Environ.* 61, 1–15. [https://doi.org/10.1016/S0034-4257\(96\)00120-4](https://doi.org/10.1016/S0034-4257(96)00120-4).
- Ben-Dor, E., Chabrilat, S., Dematté, J.A., 2018. Characterization of soil properties using reflectance spectroscopy. In: *Fundamentals, Sensor Systems, Spectral Libraries, and Data Mining for Vegetation*. CRC Press, pp. 187–247.
- Ben-Dor, E., Chabrilat, S., Dematté, J.A.M., Taylor, G.R., Hill, J., Whiting, M.L., Sommer, S., 2009. Using imaging spectroscopy to study soil properties. *Remote Sens. Environ.* 113, S38–S55. <https://doi.org/10.1016/j.rse.2008.09.019>.
- Bockheim, J.G., Gennadiyev, A.N., 2000. The role of soil-forming processes in the definition of taxa in soil taxonomy and the world soil Reference Base. *Geoderma* 95, 53–72. [https://doi.org/10.1016/S0016-7061\(99\)00083-X](https://doi.org/10.1016/S0016-7061(99)00083-X).
- Breiman, L., 2001. Random forests. *Mach. Learn.* 5–32 <https://doi.org/10.1023/A:1010933404324>.
- Camargo, L.A., Marques, J., Barrón, V., Alleoni, L.R.F., Barbosa, R.S., Pereira, G.T., 2015. Mapping of clay, iron oxide and adsorbed phosphate in oxisols using diffuse reflectance spectroscopy. *Geoderma* 251–252, 124–132. <https://doi.org/10.1016/j.geoderma.2015.03.027>.
- Cañasveras, J.C., Barrón, V., del Campillo, M.C., Torrent, J., Gómez, J.A., 2010. Estimation of aggregate stability indices in Mediterranean soils by diffuse reflectance spectroscopy. *Geoderma* 158, 78–84. <https://doi.org/10.1016/j.geoderma.2009.09.004>.
- Cierniewski, J., Guliński, M., 2010. Furrow microrelief influence on the directional hyperspectral reflectance of soil at various illumination and observation conditions. *IEEE Trans. Geosci. Remote Sens.* 48 (11), 4143–4148.
- Coelho, A.G.S., Audi, R., 1964. A côn em relação ao teor de umidade de algumas séries de solos. *Bragantia* 23, 117–128. <https://doi.org/10.1590/S0006-87051964000100012>.
- Cooper, M., Mendes, L.M.S., Silva, W.L.C., Sparovek, G., 2005. A national soil profile database for Brazil available to international scientists. *Soil Sci. Soc. Am. J.* 69 (3), 649–652.
- Dematté, J.A.M., Fongaro, C.T., Rizzo, R., Safanelli, J.L., 2018. Geospatial soil sensing system (GEOS3): a powerful data mining procedure to retrieve soil spectral reflectance from satellite images. *Remote Sens. Environ.* 212, 161–175. <https://doi.org/10.1016/j.rse.2018.04.047>.
- Dematté, J.A.M., Safanelli, J.L., Poppiel, R.R., Rizzo, R., Silvero, N.E.Q., Mendes, W.S., Bonfatti, B.R., Dotto, A.C., Salazar, D.F.U., Mello, F.A.O., Paiva, A.F.S., Souza, A.B., Santos, N.V., Maria Nascimento, C., Mello, D.C., Bellinaso, H., Gonzaga Neto, L., Amorim, M.T.A., Resende, M.E.B., Vieira, J.S., Queiroz, L.G., Gallo, B.C., Sayão, V., Lisboa, C.J.S., 2020. Bare Earth's surface spectra as a proxy for soil resource monitoring. *Sci. Rep.* 10, 1–11. <https://doi.org/10.1038/s41598-020-61408-1>.
- Diek, S., Fornallaz, F., Schaepman, M.E., 2017. Barest pixel composite for agricultural areas using landsat time series. *Remote Sens.* 9, 1245. <https://doi.org/10.3390/rs9121245>.
- Escadafal, R., 1993. Remote sensing of soil color: principles and applications. *Remote Sens. Rev.* 7, 261–279. <https://doi.org/10.1080/02757259309532181>.
- Escadafal, R., Girard, M.-C., Courault, D., 1989. Munsell soil color and soil reflectance in the visible spectral bands of landsat MSS and TM data. *Remote Sens. Environ.* 27, 37–46. [https://doi.org/10.1016/0034-4257\(89\)90035-7](https://doi.org/10.1016/0034-4257(89)90035-7).
- FAO-UNESCO, 1974. *FAO/UNESCO Soil Map of the World at 1:5 000 000 Scale, vol. I*.
- Fernandez, R.N., Schulze, D.G., Coffin, D.L., Van Scyoc, G.E., 1988. Color, organic matter, and pesticide adsorption relationships in a soil landscape. *Soil Sci. Soc. Am. J.* 52, 1023–1026. <https://doi.org/10.2136/sssaj1988.03615995005200040023x>.
- Francos, N., Ben-Dor, E., 2022. A transfer function to predict soil surface reflectance from laboratory soil spectral libraries. *Geoderma* 405, 115432.
- Gholizadeh, A., Saberioon, M., Viscarra Rossel, R.A., Boruvka, L., Klement, A., 2020. Spectroscopic measurements and imaging of soil colour for field scale estimation of soil organic carbon. *Geoderma* 357, 113972. <https://doi.org/10.1016/j.geoderma.2019.113972>.
- Gorelick, N., Hancher, M., Dixon, M., Ilyushchenko, S., Thau, D., Moore, R., 2017. Google earth engine: planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* 202, 18–27. <https://doi.org/10.1016/j.rse.2017.06.031>.
- Guerrero, C., Mataix-Solerá, J., Arcenegui, V., Mataix-Beneyto, J., Gómez, I., 2007. Near-infrared spectroscopy to estimate the maximum temperatures reached on burned soils. *Soil Sci. Soc. Am. J.* 71, 1029–1037. <https://doi.org/10.2136/sssaj2006.0187>.
- Horvath, E.H., Post, D.F., Kelsey, J.B., 1984. The relationships of landsat digital data to the properties of Arizona rangelands. *Soil Sci. Soc. Am. J.* 48, 1331–1334. <https://doi.org/10.2136/sssaj1984.03615995004800060026x>.
- Ibáñez-Asensio, S., Marqués-Mateu, A., Moreno-Ramón, H., Balasch, S., 2013. Statistical relationships between soil colour and soil attributes in semiarid areas. *Biosyst. Eng.* 116, 120–129. <https://doi.org/10.1016/j.biosystemseng.2013.07.013>.
- Ishida, T., Ando, H., Fukuhara, M., 1991. Estimation of complex refractive index of soil particles and its dependence on soil chemical properties. *Remote Sens. Environ.* 38, 173–182. [https://doi.org/10.1016/0034-4257\(91\)90087-M](https://doi.org/10.1016/0034-4257(91)90087-M).
- IUSS Working Group WRB, 2015. *World Reference Base for Soil Resources 2014, Update 2015*. FAO, Rome. *World Soil Resources Reports* 106.

- Jiang, Z., Liu, Q., Roberts, A.P., Dekkers, M.J., Barrón, V., Torrent, J., Li, S., 2022. The magnetic and color reflectance properties of hematite: from earth to Mars. *Rev. Geophys.* 60 <https://doi.org/10.1029/2020RG000698>.
- Ketterings, Q.M., Bigham, J.M., 2000. Soil Color as an Indicator of Slash-and-Burn Fire Severity and Soil Fertility in Sumatra, Indonesia. *Soil Sci. Soc. Am. J.* 64, 1826–1833. <https://doi.org/10.2136/sssaj2000.6451826x>.
- Lagacherie, P., Sneep, A.-R., Gomez, C., Bacha, S., Coulouma, G., Hamrouni, M.H., Mekki, I., 2013. Combining Vis–NIR hyperspectral imagery and legacy measured soil profiles to map subsurface soil properties in a Mediterranean area (Cap-bon, Tunisia). *Geoderma* 209–210, 168–176. <https://doi.org/10.1016/j.geoderma.2013.06.005>.
- Liu, F., Rossiter, D.G., Zhang, G.-L., Li, D.-C., 2020. A soil colour map of China. *Geoderma* 379, 114556. <https://doi.org/10.1016/j.geoderma.2020.114556>.
- Long, X., Ji, J., Barrón, V., Torrent, J., 2016. Climatic thresholds for pedogenic iron oxides under aerobic conditions: processes and their significance in paleoclimate reconstruction. *Quat. Sci. Rev.* 150, 264–277. <https://doi.org/10.1016/j.quascirev.2016.08.031>.
- Matthias, A.D., Fimbres, A., Sano, E.E., Post, D.F., Accioly, L., Batchily, A.K., Ferreira, L.G., 2000. Surface roughness effects on soil albedo. *Soil Sci. Soc. Am. J.* 64 (3), 1035–1041.
- Mattikalli, N., 1997. Soil color modeling for the visible and near-infrared bands of landsat sensors using laboratory spectral measurements. *Remote Sens. Environ.* 59, 14–28. [https://doi.org/10.1016/S0034-4257\(96\)00075-2](https://doi.org/10.1016/S0034-4257(96)00075-2).
- Mendes, W.S., Medeiros Neto, L.G., Dematté, J.A.M., Gallo, B.C., Rizzo, R., Safanelli, J.L., Fongaro, C.T., 2019. Is it possible to map subsurface soil attributes by satellite spectral transfer models? *Geoderma* 343, 269–279. <https://doi.org/10.1016/j.geoderma.2019.01.025>.
- Mouazen, A.M., Maleki, M.R., De Baerdemaeker, J., Ramon, H., 2007. On-line measurement of some selected soil properties using a VIS–NIR sensor. *Soil Tillage Res.* 93, 13–27. <https://doi.org/10.1016/j.still.2006.03.009>.
- Mulder, V.L., de Bruin, S., Schaeppman, M.E., Mayr, T.R., 2011. The use of remote sensing in soil and terrain mapping — a review. *Geoderma* 162, 1–19. <https://doi.org/10.1016/j.geoderma.2010.12.018>.
- Munsell, A.H., 1907. A color notation. *Geo. H. Ellis Co, Boston*.
- Murphy, R.J., Wedge, G., 1994. The effects of vegetation on the ability to map soils using imaging spectrometer data. *Remote Sens.* 15 (1), 63–86.
- Nolet, C., Poortinga, A., Roosjen, P., Bartholomeus, H., Ruessink, G., 2014. Measuring and modeling the effect of surface moisture on the spectral reflectance of Coastal Beach sand. *PLoS One* 9, e112151. <https://doi.org/10.1371/journal.pone.0112151>.
- Soils, N.R.C.S., 2023. Soil Colors of the Continental United States [WWW Document]. URL (accessed 4.16.22). <https://www.nrcs.usda.gov/wps/portal/nrcs/detail/soils/survey/geo/?cid=nrcsepr1423827>.
- Ohta, N., Robertson, A.R., 2006. In: 3:9: Standard and Supplementary Illuminants. Wiley, pp. 92–96. <https://doi.org/10.1002/0470094745.ch3>. ISBN 978-0-470-09472-3.
- Colorimetry.
- Padarian, J., Minasny, B., McBratney, A.B., 2015. Using Google's cloud-based platform for digital soil mapping. *Comput. Geosci.* 83, 80–88. <https://doi.org/10.1016/j.cageo.2015.06.023>.
- Padarian, J., Stockmann, U., Minasny, B., McBratney, A.B., 2022. Monitoring changes in global soil organic carbon stocks from space. *Remote Sens. Environ.* 281, 113260. <https://doi.org/10.1016/j.rse.2022.113260>.
- Persson, M., 2005. Estimating surface soil moisture from soil color using image analysis. *Vadose Z.* 4, 1119–1122. <https://doi.org/10.2136/vzj2005.0023>.
- Poggio, L., De Sousa, L.M., Batjes, N.H., Heuvelink, G.B.M., Kempen, B., Ribeiro, E., Rossiter, D., 2021. SoilGrids 2.0: producing soil information for the globe with quantified spatial uncertainty. *Soil* 7 (1). <https://doi.org/10.5194/soil-7-217-2021>.
- Poppel, R.R., Lacerda, M.P.C., Rizzo, R., Safanelli, J.L., Bonfatti, B.R., Silvero, N.E.Q., Dematté, J.A.M., 2020. Soil color and mineralogy mapping using proximal and remote sensing in Midwest Brazil. *Remote Sens.* 12, 1197. <https://doi.org/10.3390/rs12071197>.
- Post, D.F., Lucas, W.M., White, S.A., Ehasz, M.J., Batchily, A.K., Horvath, E.H., 1994. Relations between soil color and landsat reflectance on semiarid rangelands. *Soil Sci. Soc. Am. J.* 58, 1809. <https://doi.org/10.2136/sssaj1994.03615995005800060033x>.
- Probst, P., Wright, M.N., Boulesteix, A., 2019. Hyperparameters and tuning strategies for random forest. *WIREs Data Min. Knowl. Discov.* 9 <https://doi.org/10.1002/widm.1301>.
- Prudnikova, E., Savin, I., 2021. Some peculiarities of arable soil organic matter detection using optical remote sensing data. *Remote Sens.* 13, 2313. <https://doi.org/10.3390/rs13122313>.
- Ramos, P.V., Inda, A.V., Barrón, V., Siqueira, D.S., Marques Júnior, J., Teixeira, D.D.B., 2020. Color in subtropical brazilian soils as determined with a munsell chart and by diffuse reflectance spectroscopy. *Catena* 193, 104609. <https://doi.org/10.1016/j.catena.2020.104609>.
- Richter, R., Schlüpfner, D., 2002. Geo-atmospheric processing of airborne imaging spectrometry data. Part 2: Atmospheric/topographic correction. *Int. J. Remote Sens.* 23, 2631–2649. <https://doi.org/10.1080/01431160110115834>.
- Roberts, D., Wilford, J., Ghattas, O., 2019. Exposed soil and mineral map of the australian continent revealing the land at its barest. *Nat. Commun.* 10, 5297. <https://doi.org/10.1038/s41467-019-13276-1>.
- Roy, D.P., Kovalskyy, V., Zhang, H.K., Vermote, E.F., Yan, L., Kumar, S.S., Egorov, A., 2016. Characterization of Landsat-7 to Landsat-8 reflective wavelength and normalized difference vegetation index continuity. *Remote Sens. Environ.* 185, 57–70. <https://doi.org/10.1016/j.rse.2015.12.024>.
- Safanelli, J.L., Chabrilat, S., Ben-Dor, E., Dematté, J.A.M., 2020. Multispectral models from bare soil composites for mapping topsoil properties over Europe. *Remote Sens.* 12, 1369. <https://doi.org/10.3390/rs12091369>.
- Sánchez-Rodríguez, A.R., Gómez-Alvarez, E., Méndez, J.M., Skiba, U.M., Jones, D.L., Chadwick, D.R., del Campillo, M.C., Fernandes, R.B., Kleffmann, J., Barrón, V., 2023. Photocatalytic fixation of NO_x in soils. *Chemosphere* 338, 139576. <https://doi.org/10.1016/J.CHEMOSPHERE.2023.139576>.
- Santra, P., Singh, R., Sarathjith, M.C., Panwar, N.R., Varghese, P., Das, B.S., 2015. Reflectance spectroscopic approach for estimation of soil properties in hot arid western Rajasthan, India. *Environ. Earth Sci.* 74, 4233–4245.
- Sarathjith, M.C., Das, B.S., Wani, S.P., Sahrawat, K.L., 2014. Dependency measures for assessing the covariation of spectrally active and inactive soil properties in diffuse reflectance spectroscopy. *Soil Sci. Soc. Am. J.* 78 (5), 1522–1530.
- Savin, I.Y., Vindeker, G.V., 2021. Some specifics in using optical properties of soil surface for moisture detection. *Eurasian Soil Sci.* 54, 1019–1027. <https://doi.org/10.1134/S1064229321070127>.
- Schmidt, S.A., Ahn, C., 2021. Analysis of soil color variables and their relationships between two field-based methods and its potential application for wetland soils. *Sci. Total Environ.* 783, 147005. <https://doi.org/10.1016/j.scitotenv.2021.147005>.
- Schwertmann, U., 1993. Relations between iron oxides, soil color, and soil formation. In: Bigham, J.M., Ciolkosz, E.J. (Eds.), *Soil Color*. SSSA, Madison, pp. 51–69.
- Schwertmann, U., 2015. In: *Relations Between Iron Oxides, Soil Color, and Soil Formation*, pp. 51–69. <https://doi.org/10.2136/sssaspecpub31.c4>.
- Sharma, G., Sharma, G., Bala, R., 2002. *Digital Color Imaging Handbook*. CRC Press.
- Shi, P., Six, J., Sila, A., Van Lanwe, B., Van Oost, K., 2022. Towards spatially continuous mapping of soil organic carbon in croplands using multitemporal Sentinel-2 remote sensing. *ISPRS J. Photogramm. Remote Sens.* 193, 187–199. <https://doi.org/10.1016/J.ISPRSJPRS.2022.09.013>.
- Shields, J.A., Arnaud, R.J., Paul, E.A., Clayton, J.S., 1966. Measurement of soil color. *Can. J. Soil Sci.* 46, 83–90. <https://doi.org/10.4141/cjss66-012>.
- Singh, D., Herlin, I., Berroir, J., Silva, E., Simoes Meirelles, M., 2004. An approach to correlate NDVI with soil colour for erosion process using NOAA/AVHRR data. *Adv. Sp. Res.* 33, 328–332. [https://doi.org/10.1016/S0273-1177\(03\)00468-X](https://doi.org/10.1016/S0273-1177(03)00468-X).
- Sovdat, B., Kadunc, M., Batić, M., Milčinski, G., 2019. Natural color representation of Sentinel-2 data. *Remote Sens. Environ.* 225, 392–402. <https://doi.org/10.1016/j.rse.2019.01.036>.
- Standard CIE Commission internationale de l'éclairage, 2019. ISO/CIE 11664-1:2019: Colorimétrie — partie 1: observateurs CIE de référence pour la colorimétrie colorimetry — part 1: CIE standard colorimetric observers [WWW Document]. <https://www.iso.org/obp/ui/#iso:std:iso:cie:11664-1:ed-1:v1:fr> accessed 4.1.22.
- Stiglitz, R., Mikhailova, E., Post, C., Schlaumann, M., Sharp, J., 2016. Evaluation of an inexpensive sensor to measure soil color. *Comput. Electron. Agric.* 121, 141–148. <https://doi.org/10.1016/j.compag.2015.11.014>.
- Thenkabail, P.S., Teluguntla, P.G., Xiong, J., Oliphant, A., Congalton, R.G., Ozdogan, M., Gumma, M.K., Tilton, J.C., Giri, C., Milesi, C., Phalke, A., Massey, R., Yadav, K., Sankey, T., Zhong, Y., Aneecie, I., Foley, D., 2021. Global cropland-extent product at 30-m resolution (GCEP30) derived from landsat satellite time-series data for the year 2015 using multiple machine-learning algorithms on Google earth engine cloud. USGS, Reston.
- Torrent, J., Barrón, V., 1993. Laboratory measurement and soil colour: Theory and practice. In: Bigham, J.M., Ciolkosz, E.J. (Eds.), *Soil Colour*. SSSA Spec, Madison, pp. 21–33.
- Tziolas, N., Tsakiridis, N., Ben-Dor, E., Theocharis, J., Zalidis, G., 2020. Employing a multi-input deep convolutional neural network to derive soil clay content from a synergy of multi-temporal optical and radar imagery data. *Remote Sens.* 12, 1389. <https://doi.org/10.3390/rs12091389>.
- USGS, 2020. Landsat 4-7 Collection 1 (C1) Surface Reflectance (LEDAPS) Product Guide [WWW Document]. URL (accessed 9.1.21). https://d9-wret.s3.us-west-2.amazonaws.com/assets/palladium/production/s3fs-public/atoms/files/LSDS-1370_L4-7_C1-SurfaceReflectance-LEDAPS_ProductGuide-v3.pdf.
- USGS, 2020. Landsat 8 Collection 1 (C1) Land Surface Reflectance (LaSRC) Product Guide [WWW Document]. URL (accessed 9.1.21). https://d9-wret.s3.us-west-2.amazonaws.com/assets/palladium/production/s3fs-public/atoms/files/LSDS-1368_L8_C1-LandSurfaceReflectanceCode-LaSRC_ProductGuide-v3.pdf.
- van Leeuwen, C.C., Mulder, V.L., Batjes, N.H., Heuvelink, G.B., 2022. Statistical modelling of measurement error in wet chemistry soil data. *European Journal of Soil Science* 73 (1), e13137.
- 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, 59–75. <https://doi.org/10.1016/j.geoderma.2005.03.007>.
- Viscarra Rossel, R.A., Bui, E.N., De Caritat, P., McKenzie, N.J., 2010. Mapping iron oxides and the color of australian soil using visible-near-infrared reflectance spectra. *J. Geophys. Res. Earth Surf.* 115, 1–13. <https://doi.org/10.1029/2009JF001645>.
- Viscarra Rossel, R.A., Minasny, B., Roudier, P., McBratney, A.B., 2006. Colour space models for soil science. *Geoderma* 133, 320–337. <https://doi.org/10.1016/j.geoderma.2005.07.017>.
- Wadoux, A.M.-C., Walvoort, D.J.J., Brus, D.J., 2022. An integrated approach for the evaluation of quantitative soil maps through Taylor and solar diagrams. *Geoderma* 405, 115332. <https://doi.org/10.1016/j.geoderma.2021.115332>.
- Zepp, S., Heiden, U., Bachmann, M., Möller, M., Wiesmeier, M., Wesemael, B., 2023. Optimized bare soil compositing for soil organic carbon prediction of topsoil croplands in Bavaria using landsat. *ISPRS J. Photogramm. Remote Sens.* 202, 287–302. <https://doi.org/10.1016/j.isprsjprs.2023.06.003>.