

Exploring bare soil digital mapping: identifying alternative variables to replace ECa via remote sensing, a case study on two Italian fields at different latitude

Matteo Petito ^a, Emanuele Barca ^{b,*} , Antonio Berti ^a, Silvia Cantalamessa ^a, Giancarlo Pagnani ^{c,**} , Michele Pisante ^c

^a Department of Agronomy, Food, Natural Resources, Animals and Environment, Agripolis, University of Padova, Viale Dell'Università 16, Legnaro Padova, 35020, Italy

^b Water Research Institute, National Research Council of Italy, 70132 Bari, Italy

^c Department of Bioscience and Technologies for Food, Agriculture and Environment, University of Teramo, Via Balzarini, 1, Teramo, 64100, Italy

ARTICLE INFO

Keywords:

Site-specific management
Satellite bare soil images
Sentinel-2
ECa
Spatial field variability
Multivariate statistical analysis

ABSTRACT

Site-specific management in agriculture, which accounts for variability within a field, is a cornerstone of sustainable agronomic practices. However, despite the availability of numerous methods to measure spatial variability, their limitations hinder large-scale adoption, posing challenges to the broader implementation of precision agriculture. This study aims to identify spectral indices derived from bare-soil analysis as potential substitutes for electrical conductivity (ECa) in mapping spatial variability. The approach aligns with the need for cost-effective, scalable, and less labor-intensive solutions to manage field variability. Using multi-temporal bare-soil imagery spanning five years across two fields under intermittent cultivation in Italy, we applied principal component analysis to evaluate correlations between spectral indices and ECa. Both fields demonstrated strong correlations between ECa and the first principal component (PC1). Key variables identified as highly correlated with ECa included the Brightness Index (0.66), Near-Infrared (0.53), and Red reflectance (0.58). The percentage variance explained by PC1 was 75.4 % for Field 1 and 79.0 % for Field 2. Finally, PC1 is correlated with ECa in the two areas in the measure of 0.73 and 0.53, respectively. This work addresses the problem of substituting ECa with covariates derived from bare-soil analysis from a purely statistical perspective as a first necessary step aiming at identifying the most promising covariates. A subsequent study will address this issue from a pedological standpoint. These findings highlight the potential of remote sensing data and spectral indices from multi-temporal imagery to replace direct ECa measurements, enabling rapid and accurate mapping of spatial variability in annual croplands.

1. Introduction

In the future, European agricultural production will need to boost both productivity and sustainability simultaneously [1]. Addressing this challenge involves embracing site-specific management (SSM) as a pathway to bring sustainable agriculture to a new level. SSM concentrates on comprehending how spatial variations in field properties affect crop yield [2]. Its assumptions propose that there is substantial spatial variability within fields affecting crop yield factors, and that these factors can be recognized, quantified, and utilized to improve both

economic and environmental sustainability [3]. Nevertheless, the effectiveness of SSM in optimizing production greatly depends on accurately assessing field variability. Several techniques gauge spatial field variability, with the most notable ones being Yield Monitoring, Topographic Factor Assessment, and mapping apparent soil Electrical Conductivity (ECa) [4].

Since its initial use in agriculture for measuring soil salinity, ECa has become widely accepted for assessing the spatial variability of various soil physical and chemical properties that directly or indirectly affect ECa measurements [5]. ECa measures the overall conductivity of the

* Corresponding author at: Construction Technologies Institute, National Research Council of Italy, Bari, Italy.

** Corresponding author at: Department of Bioscience and Technologies for Food, Agriculture and Environment, University of Teramo, Via Balzarini, 1, Teramo, 64100, Italy.

E-mail addresses: emanuele.barca@ba.irsa.cnr.it (E. Barca), gpagnani@unite.it (G. Pagnani).



Fig. 1. View of field no.1.

soil, encompassing everything conductive within the measurement volume. It is influenced, directly or indirectly, by any soil property that impacts bulk soil conductance [6]. Measuring apparent EC_a provides valuable insights into the spatial diversity of soil physical and chemical properties [7]. In non-saline soils, EC_a is correlated with clay content, organic matter, cation exchange capacity, and water content [8]. Methods such as electrical resistivity and electromagnetic induction gauge EC_a and have been linked to soil characteristics such as texture, salinity, nutrient levels, and organic matter [9,10]. However, proper measurement protocols are essential due to varied results in correlating EC_a with yield maps [10–12].

While EC_a mapping has proven effective in precision farming, its complexity, need for expertise, specific equipment, and susceptibility to

secondary influences pose challenges. Factors such as variations in soil texture, moisture content, temperature, salinity, mineral composition, compactness, buried materials (e.g., rocks or anthropogenic structures), and surface conditions like vegetation or organic residues can mask or distort the true variability of soil properties [8]. The swift growth of Precision Agriculture (PA) demands quicker and more dependable methods for mapping spatial field variations, prompting satellite bare-soil images to emerge as a feasible alternative. These images gather indirect soil property information economically and over time [13]. Innovative methodologies, such as the Geospatial Soil Sensing System (GEOS3), have been introduced in recent studies. These methodologies utilize satellite imagery such as Landsat and Sentinel-2 to generate Synthetic Soil Images (SYSI) for mapping topsoil properties [14,15].

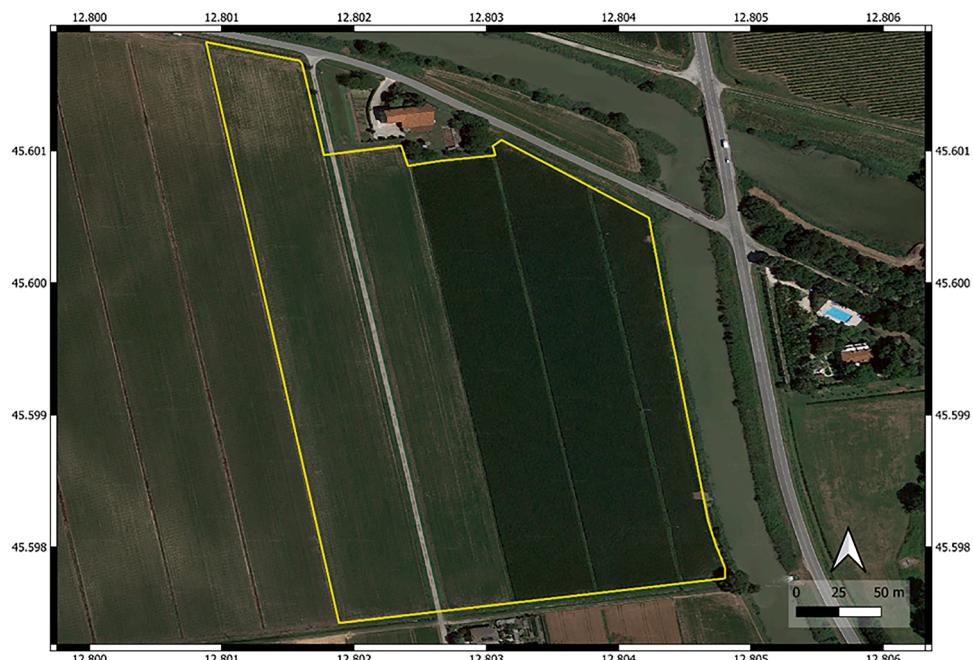


Fig. 2. View of field no.2.

Table 1
Survey period in the fields.

Field	ECa survey
Field no 1	October 2019
Field no 2	June 2019

Over the years, several authors have also tried to estimate ECa through remote sensing with the implementation of parameters such as geomorphology and terrain parameters with bare soil [4,16–19].

Given the range of contemporary methods used to map within-field variability, no attempt has been presented to use multitemporal bare-soil image analysis as a proxy for assessing field-scale soil variability in arable cropland. Therefore, this study investigates apparent ECa variation in two different areas in Italy through a methodology designed to select the most relevant ECa-related variables from direct field measurements. From a purely statistical perspective, this work represents a first necessary step toward identifying promising covariates that could substitute ECa in precision agriculture. A subsequent study will explore this issue from a pedological standpoint, providing a more comprehensive understanding of the implications. The motivation stems from the considerable time required and high economic cost of ECa surveys, uncertainties in ECa measurements, and the potential to extract numerous variables using remote sensing techniques.

2. Material and methods

2.1. Study areas

Field no 1

Field no 1 (Fig. 1) covers an area of 12 hectares and is situated within the administrative region of Emilia-Romagna. From a pedological standpoint, as outlined in the Soil Map of Italy [20], it falls within the soil region of the Po plain and its surrounding hills. The pedological province of the area encompasses Haplic Calcisol (Endogleyic) and (Hypercalcic), Calcaric and Calcaric Fluvic Cambisol, and Calcaric Fluvisol. In terms of climate, the field is characterized by a temperate subcontinental climate of northern Italy. This type of climate is present in the average and high Po plain, as well as in the western and locally eastern morainic plains, with a supratemperate/mesotemperate humid-subhumid climate. Regarding micromorphology, the field has a maximum elevation of 1.70 m above sea level and a minimum elevation of −1 m, with a range of 2.69 m. The average slope of the area is 0.24 %.

Field no 2

The second field (Fig. 2) spans 9 hectares and is located in the administrative region of Veneto, specifically in the city of Eraclea. This area is characterized by a climate classification, according to the Phytoclimatic Map of Italy, as a Temperate climate of northern Italy, typical of the eastern alluvial plains and the central morainic plains and valleys (Mesotemperate/Supratemperate humid). From a pedological perspective, the field falls within the region of soils of the Po plain and associated hills and in the Calcaric Fluvic Gleyic Cambisol province; Thapthohistic Thionic Fluvic Gleysol; Calcaric Gleyic Arenosol; Hyposalic; Calcaric Gleyic Fluvisols; Thionic Histosols [20].

In terms of micromorphology, the field ranges from a maximum elevation of −2.64 m above sea level to a minimum elevation of −4.34 m, resulting in a range of 1.70 m. The average slope of the area is 0.77 %.

2.2. ECa survey

An apparent soil electrical conductivity survey (Table 1) was performed utilizing the Veris I-scan soil electrical conductivity sensor (Veris Technologies Inc., Salina, KS, USA). The Veris cart was affixed to a tractor and pulled across the fields. The sensor operates based on the principle that smaller soil particles (such as clay-sized particles) conduct more current compared to larger particles (such as loamy or sandy

Table 2
Bare soil index.

Index	Formula	Soil properties (reference)
Hue Index (HI)	(2 B4-B3-B2)/(B3-B2)	Primary colours [22]
Brightness Index (BI)		Average reflectance [22]
Saturation Index (SI)	(B4-B2)/(B4+B2)	Magnitude [22]
Coloration Index (CI)	(B4-B3)/(B4+B3)	Spectra slope [22]
SOCI (Soil Organic Carbon Index)	(B2/(B4 B3))	Soil Colour [22]
Soil Background Line (SBL)	(B8 -(B4 1.24))	Organic Carbon [22]
Ferric Iron 3+ Index	(B4/B3)	Iron 3+ soil content [23]
Grain Size Index (GSI)	(B4-B2)/(/B4+B3+B2)	Soil Texture [24]
Iron Oxide Ratio (IOR)	(B4/B2)	Iron Oxide Soil content [25]
Salt Index (SIndx)		Salt Soil content [26]
Normalize Salt Index (NSI)	(B4-B8)/(B4+B8)	Salt Soil content [26]
Salinity Index 2		Salt Soil content [26]

particles). To assess this, a pair of colter electrodes was employed to gauge the voltage drop. The Veris I-scan produces two sets of data: topsoil data (EC30), which is weighted for depths of 0–30 cm, and depth data (EC90) covering depths of 0–90 cm. For this study, only soil data (EC30) were considered, as EC90 data, derived through more complex methodologies, are subject to greater uncertainty, which compromises their reliability. The monitoring was conducted by carrying out transects spaced 10 m apart, and within each transect, the sensor recorded the ECa value every 1.5 m

2.2.1. $ECa_{(10m)}$

The ECa values were resampled at a spatial resolution of 10 m. To do this, a pipeline was created in QGIS using the Model Designer tool. This pipeline consists of several steps and related algorithms. (a) Initially, a grid of polygons with 10×10 meters dimensions was created, corresponding to the Sentinel 2 pixels. Each polygon was assigned a unique ID. (b) ECa points falling within the polygons were subsequently intersected. Each point was assigned the ID of the corresponding polygon. (c) The ECa average was then calculated by aggregating the points with the same ID. (d) Finally, a point corresponding to the polygon's centroid was created, and the average ECa value was assigned to that point.

2.3. Synthetic soil images

In Google Earth Engine (GEE), for each study area, SYSI were generated using multitemporal image collections from the Sentinel-2 Multispectral Instrument satellite (S2-MSI) and the Geospatial Soil Sensing System (GEOS 3) technique proposed by Demattè, et al. [14]. The system, initially designed for Landsat, was implemented using satellite images from the Sentinel-2 satellite [15]. The system can be outlined in the following steps: (a) Acquiring and compiling a series of satellite images spanning multiple time periods. (b) Identifying pixels classified as bare soil within each image. (c) Extracting VIS-NIR reflectance from each bare soil pixel, forming the Temporal Soil Spectral Reflectance (TESS). (d) Computing the median of all TESS values to generate the Synthetic Bare Soil Image. A database was established using Level-2A products from the Sentinel-2 satellite, comprising atmospherically corrected surface reflectance bands processed by the Sen2Cor algorithm [21]. In this study, we utilized bands from S2-MSI images with a spatial resolution ranging from 10 to 60 m, focusing specifically on bands with 10-pixel sizes. For each study area, a temporal interval was considered for image acquisition, ranging from 28/3/2017, the earliest available date for Level 2A products in GEE, to the most recent date on which the ECa survey was conducted. The images were selected based on a criterion of cloud coverage (<20 %). Once obtained, spectral indices such as NDVI (Normalized Difference Vegetation Index) and

Table 3

Summary statistics for ECa related the two study areas.

	n	mean	sd	Median	min	max	range	skew	kurtosis	se
Field no 1	1071	46.42	10.29	46.86	15.57	77.58	62.01	-0.09	-0.08	0.31
Field no 2	703	49.3	12.6	49.84	9.84	88.07	78.24	-0.18	0.42	0.48

n: number of measurements from Veris I scan; sd: standard deviation; skew: skewness; se: standard error.

NBR2 (Normalized Burn Ratio 2) were calculated and utilized to mask the pixel images. Following the GEOS3 method, NDVI values falling within the range of 0 to 0.25 were categorized as bare soil, while values exceeding this threshold were marked as NA. NBR2 values > 0.15 were also marked as NA, indicating areas covered with straw or affected by burning. Subsequently, the masked images were arranged chronologically, and the median reflectance was computed for each satellite.

2.3.1. Bare soil cube dataset

After generating the synthetic soil images, reflectance of the visible bands (458–680 nm): Blue (B2), Green (B3), Red (B4), and Near-Infrared (785–900 nm) band (B8 and NIR) were extracted. Subsequently, the following bare soil indices were computed (Table 2). Finally, the individual VIS-NIR bands of the SYSI and the bare soil indices were aggregated into a single multidimensional raster (Bare Soil Cube - BSC).

2.3.2. Bare soil cube dataset reduction

To evaluate the Bare Soil Covariates (BSC) that show the strongest correlation with ECa (10 m), a Pearson correlation analysis was conducted for each study area, this step identified the six covariates with the highest absolute correlation coefficients. The computations were performed using the cor function from the R stats library. These were then subjected to further reduction through Principal Components Analysis (PCA), a key multivariate statistical tool. PCA is instrumental in dimensionality reduction, transforming a potentially large set of correlated variables into a smaller set of uncorrelated variables called principal components (PCs). These PCs are ordered based on the variance they explain, offering a concise yet comprehensive representation of the original dataset while preserving its critical patterns and structures. This process not only facilitates data interpretation but also enhances subsequent modelling efforts. Before performing PCA, outlier detection was essential, as PCA is highly sensitive to their influence. Outliers can skew the results, distorting the variance distribution among PCs and potentially leading to misleading conclusions. Given the relatively symmetric empirical distributions of the ECa data (see Table 3), the 3-sigma rule of thumb [27] was deemed appropriate for this task. Following outlier

handling, the dataset's suitability for PCA was assessed using two diagnostic tests: the Kaiser-Meyer-Olkin (KMO) test and Bartlett's test of sphericity. The KMO index measures sampling adequacy, with values closer to 1.0 indicating high suitability, while values below 0.5 are considered unacceptable [28,29]. Bartlett's test, on the other hand, evaluates whether the variables are sufficiently inter-correlated to justify the application of PCA (Bartlett, 1951). Passing these tests ensures the robustness of the analysis and enhances the reliability of PCA-derived conclusions. The dataset size, comprising a medium range of 100–1000 observations and 20–200 variables, was also considered appropriate for PCA, striking a balance between capturing meaningful variance and maintaining computational efficiency. Computations were carried out using the R functions: bart_spher (REdaS library), KMO (psych library), and prcomp (stats library). After PCA, the final step was to determine the optimal number of principal components (PCs) to maximize the predictive power of ECa.

2.3.3. Spatial correlogram

The spatial correlogram stands as a fundamental tool in spatial statistics, offering insights into the spatial autocorrelation present within georeferenced data. This technique examines the degree of similarity or dissimilarity between observations at different spatial locations, elucidating the presence and magnitude of spatial dependence. By quantifying the correlation structure across varying lag classes, the spatial correlogram provides crucial information about the spatial patterns, helping to discern the scale at which spatial processes operate. It enables better understanding and modeling of spatial relationships, resource allocation, and the diffusion of phenomena across space. The spatial correlogram's ability to unveil spatial dependencies plays a crucial role in comparing the spatial maps of two different variables (bivariate correlogram or cross-correlogram) in order to understand if they could be considered spatially equivalent and, in the decision, if replacing one with the other [30]. The computations were carried out by means of the function spline.correlog of the R library(ncf).

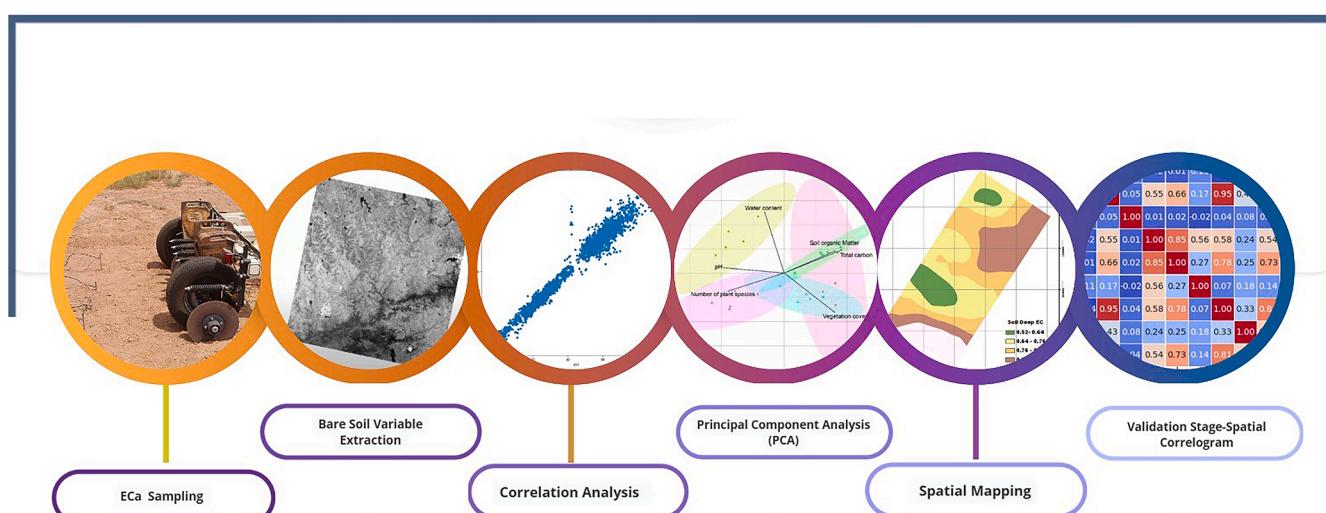
**Fig. 3.** Workflow of methodology.

Table 4

Selection of variables correlated with ECa in the two experimental area.

Field no 1	R	r abs
Eca-SOCl	0,69	0,69
ECa-SI	-0,66	0,66
ECa-GSI	-0,66	0,66
ECa-Red	-0,58	0,58
ECa-BI	-0,57	0,57
Eca-Nir	-0,57	0,57
Field no 2	r	r abs
ECa-Red	-0,57	0,57
ECa-GSI	-0,57	0,57
ECa-SI	-0,55	0,55
ECa-Nir	-0,53	0,53
ECa-BI	-0,53	0,53
ECa-NSI	-0,52	0,52

2.4. Workflow of methodology

With the objective of substituting ECa measurements with alternative variables derived from SYSI analysis, the proposed methodology (Fig. 3) is outlined as follows:

1. ECa Sampling: A comprehensive ECa sampling is conducted across the designated area of interest.
2. Bare Soil Variable Extraction: Relevant bare soil variables are meticulously extracted from the SYSI.
3. Correlation Analysis: The initial set of variables is subjected to a thorough correlation analysis to identify and retain the most significant and non-redundant factors correlated with ECa.
4. Principal Component Analysis (PCA): The reduced set of variables undergoes Principal Component Analysis (PCA) to capture the essential patterns and reduce dimensionality.
5. Spatial Mapping: Spatial maps for both ECa and the first principal component (PC1) are generated, facilitating a comparative analysis of spatial variability.
6. Validation Stage - Spatial Correlogram: The spatial correlogram is employed in the validation stage to assess and validate the spatial variability, comparing the spatial distribution of ECa with that of PC1.

This methodological framework aims to show that alternative variables derived from SYSI can provide the same spatial variability as ECa, ensuring a robust and validated approach to spatial variability analysis.

3. Results

3.1. ECa basic statistics

Table 3 provides statistical summaries of ECa values for two distinct areas: field no 1 and field no 2. Field no 1, with a mean ECa value of 46.42 and a standard deviation of 10.29, demonstrates moderate variability in soil ECa. The distribution appears approximately Gaussian, as indicated by a skewness close to zero (-0.09) and a kurtosis of -0.08, suggesting symmetric tails. The median ECa value (46.86) aligns closely with the mean, reflecting a balanced distribution. Field no 2, displays characteristics similar to field 1, with a mean ECa value of 49.3 and a standard deviation of 12.6. The distribution also appears Gaussian, with a skewness near zero (-0.18) and a kurtosis of 0.42, indicating relatively normal tails. The median ECa value (49.84) closely mirrors the mean, reinforcing the symmetry of the distribution. The computations were carried out by means of the function *describe* of the R library(psych).

3.2. Bare soil variable extraction - correlation matrix analysis

For the two study areas, the variables most strongly correlated with

Table 5

Illustrates the frequency of appearance of the selected indices in the list of indices most strongly correlated with ECa.

Indices	N	Field
BI	3	All
Nir	3	All
Red	3	All
GSI	2	Field no 1 and no 2
SI	2	Field no 1 and no 2
SOCI	2	Field no 1
NSI	1	Field no 2

ECa were identified and selected (Table 4, Table 5). This initial set of variables was further refined by analyzing the inverse of the correlation matrix and excluding columns (variables) that rendered the matrix singular. These variables were deemed redundant and could be safely excluded from further analysis. After this refinement process, the final set of variables considered in the study included: BI, NIR, Red, GSI, SI, SOCI, NSI, Salt Index, and Salt Index 2. Among these, the variables BI, NIR, and Red consistently ranked among the top six variables with the highest correlation to ECa across all plots, each demonstrating a negative correlation. Additionally:

- GSI and SI were selected for fields no 1 and no 2.
- SOCI, the only variable positively correlated with ECa, was included for fields no 1 and no 3.
- NSI appeared exclusively in field no 2.

3.3. PCA analysis

Firstly, the KMO and Bartlett tests were performed on the two datasets to check their suitability to PCA analysis. All the outcomes of KMO and Bartlett test are reported on Table 6 and Table 7. KMO outcomes in Table 6 are in majority larger of 0.8 that is the threshold testifying that datasets are well suited for PCA analysis. The outcomes of Bartlett tests are reported in Table 7 and show that PCA can be safely applied. Since KMO and Bartlett test indicated the suitability of the dataset to PCA/Factor Analysis, a last pre-processing step was carried out before applying Principal Component Analysis (PCA): the data standardization. This step ensures that each variable contributes equally to the analysis, which is particularly important when variables have different units or scales. Standardization rescales the data so that each variable has a mean of zero and a standard deviation of one, allowing PCA to accurately capture the underlying structure of the data without being biased by differences in measurement units or ranges. Afterwards, the above reported variables (indices) were compressed in their principal components by means of the PCA process. The variance percentage (var %) of the PC1 is the following for the three datasets corresponding to the two study areas: field no 1, PC1 var % 7.54e+01; field no 2, PC1 var % 7.90e+01; field 3, PC1 var % 5.87e+01. The variance percentage contained in the PC1 of the two areas are 1071 and 703, respectively. Finally, PC1 is correlated with ECa in the two areas in the measure of 0.73 and 0.53, respectively.

3.4. Spatial cross-correlation

Cross-correlation measures the degree of association between two variables at different space lags. Cross-correlation analysis can put on evidence similarities between a couple of maps and show the coincidence of homogeneous zones. In Table 8 are shown cross-correlation values at different lags between ECa and PC1 of the two study areas.

Field no 1 (Fig. 1) showed the best correlation and field no 2 the poorest (Fig. 2). Similar to field no 1, field no 2 demonstrates a positive correlation between ECa and PC1, with slightly lower correlation coefficients compared to field no 1 across all two lags. In general, the cross-correlation analysis indicates a reliable positive correlation between ECa

Table 6
KMO index estimation outcomes.

KMO index values									
Study Area	SOCI	red	BI	nir	salt_index2	salt_index	green	GSI	IOR
field no.1	0.878	0.813	0.813	0.849	0.848	0.836	0.804	0.736	0.733
field no.2	0.801	0.794	0.831	0.895	0.879	0.798	0.795	0.8	0.802

Table 7
Bartlett Sphericity tests outcomes.

Bartlett test outcomes		
X ² statistics	df	p-value
6429.508	171	< 2.22e-16
113,128.633	171	< 2.22e-16
171,284.754	171	< 2.22e-16

Table 8
Cross-correlation values at two lags between ECa and PC1.

Study area	0m	10m	20m	30m
Field no 1	0.69	0.64	0.55	0.45
Field no 2	0.57	0.50	0.45	0.40

and PC1 across all research sites, with field no 1 demonstrating the most robust correlations, trailed by field no 2. Additionally, the correlations tend to weaken as the lag between ECa and PC1 increases, indicating a decrease in the strength of the relationship over space.

Moreover, has been outlined the concluding step of the methodology used to compare the spatial variability of ECa with that of the synthetic variable PC1. Through examination of maps depicting these variables across two designated study areas, it becomes evident that they display similar patterns.

4. Discussion

This study outlines a statistical pathway for identifying variables that can potentially replace ECa as the primary descriptor of soil variability. The proposed approach leverages statistical methods to identify the most promising covariates derived from bare soil analysis. Fields no 1 and no 2 exhibit relatively similar distributions, characterized by moderate variability and an approximately Gaussian shape. To date, no experimental studies have explored the correlation between specific bare soil indices, extracted from Synthetic Soil Images (SYSI), and apparent electrical conductivity (ECa) values obtained through proximal sensing.

The selection of variables for bare soil analysis builds upon the established notion that in non-saline soils, apparent electrical conductivity correlates with various properties such as clay content, cation exchange capacity, and organic matter content [7,8,31]. Numerous studies highlight that bare soil reflectance, captured via remote sensing, alongside associated indices, can serve as reliable indicators for estimating soil texture and organic matter characteristics [32–34]. Among the indices analyzed, the Brightness Index (BI) and Saturation Index (SI) stand out as soil colorimetric indicators. BI quantifies soil reflectance, correlating with ground brightness, while SI measures the gradient across the visible spectrum (red to blue), indicating deviations from a neutral, grey-like spectrum [22]. Soil color, a qualitative indicator of soil properties, is strongly associated with characteristics such as organic matter and texture [35]. Furthermore, Silvero, et al. [15] demonstrated that reflectance in the Red and NIR bands, derived from Sentinel-2-based SYSI, negatively correlates with clay and organic matter content. In our analysis, NIR reflectance exhibited a distinct pattern: Field no 1 < Field no 2. Similarly, the Grain Size Index (GSI),

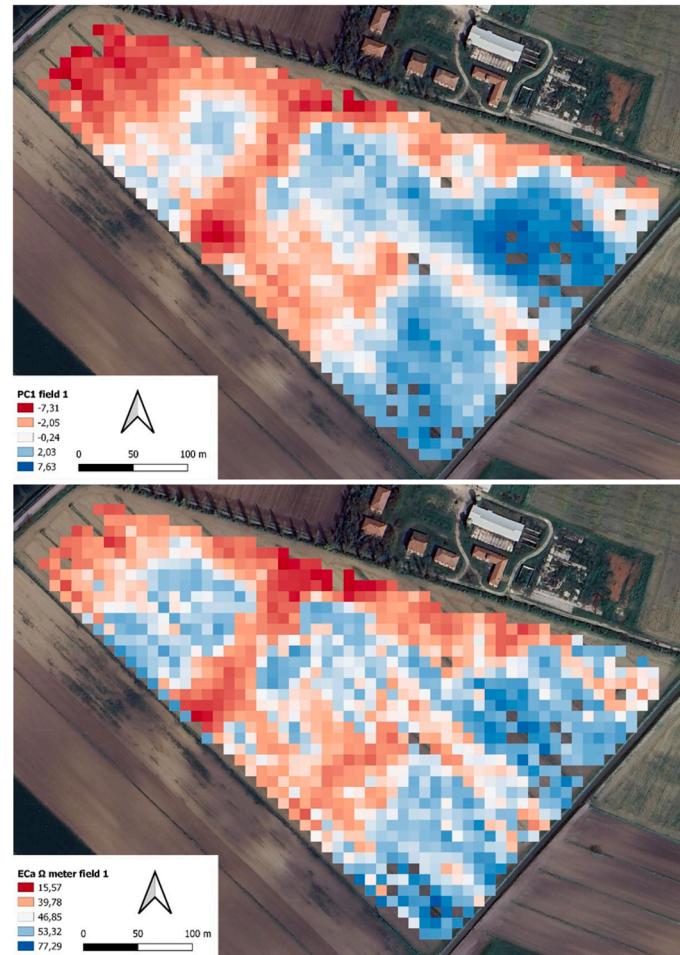


Fig. 4. Field no 1 maps of (top) synthetic variable PC1 and (bottom) spatial variability of ECa.

typically associated with fine sand content [24], showed a negative correlation with ECa in all plots, corroborating findings from Carroll and Oliver, [36] and Medeiros, et al. [37] about the inverse relationship between sand content and ECa.

The **Soil Organic Carbon Index (SOCI)**, designed to estimate soil organic carbon based on reflectance in visible bands [38], was the only bare soil index positively correlated with organic matter in our study, reinforcing the validity of the approach.

The **Normalized Salt Index (NSI)**, along with **Salt Index** and **Salt Index 2**, has historically been used to predict soil electrical conductivity and identify salinity risks [26]. Interestingly, these indices were present only in Field no 3, which might suggest saline soil presence. However, this claim is contradicted by the low ECa values observed. Saline soils typically exhibit high reflectance due to their light color and surface crust formation, particularly in dry conditions [26]. In this case, the elevated Salt Index values could instead be attributed to sandy soil properties, which do not consistently align with high EC readings [39]. This interpretation also aligns with the observed higher NIR reflectance values in Field no 3 and the overall lower ECa levels compared to Fields

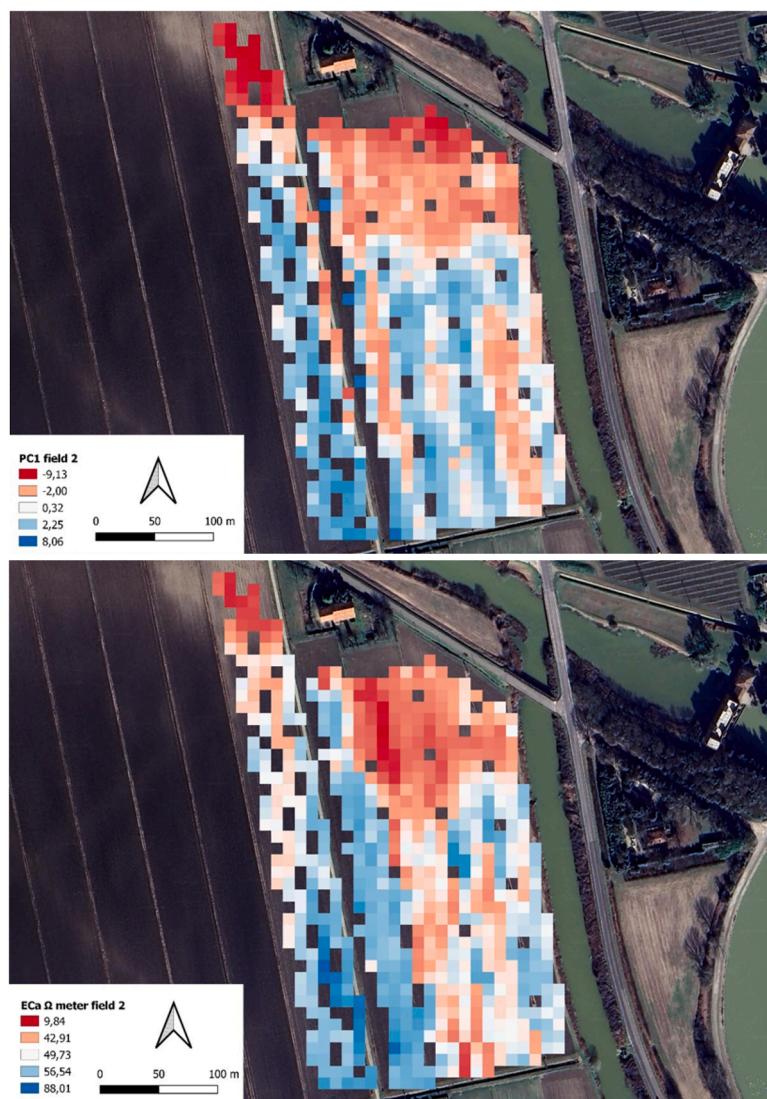


Fig. 5. Field no 2 maps of (top) synthetic variable PC1 and (bottom) spatial variability of ECa.

no 1 and no 2.

Spatial cross-correlation analysis further revealed the degree of association between ECa and the first principal component (PC1) across varying spatial lags. Field no 1 showed the strongest correlation, followed by Field no 2, likely influenced by differences in dataset size (Table 3). Despite variations in correlation strength, a consistent positive association between ECa and PC1 was evident across all study areas, with correlations decreasing as spatial lag increased (Table 7). This consistency underscores the potential of synthetic variables derived from remote sensing as proxies for ECa measurements, offering cost-effective solutions for assessing soil variability on larger spatial scales. Silvero et al. [15] emphasized that enhanced access to synthetic imagery significantly improves the coverage of bare soil pixels, boosting the accuracy of predictions for soil properties such as ECa, clay content, organic matter, and cation exchange capacity. However, not all bare soil pixels are equally effective for accurate soil property predictions. In agricultural contexts, irrigation practices and farm management techniques can heavily influence soil spectral attributes. Activities like salt leaching and organic fertilization tend to lower salt content in the root zone, particularly in arid irrigation areas where soil moisture is critical for vegetation growth [40].

Lastly, a comparison of ECa and PC1 maps revealed similar spatial patterns across the study areas (Fig. 4 and Fig. 5), further validating the

utility of synthetic variables for capturing soil variability. This study reaffirms the significant potential of remote sensing to extract meaningful information about soil properties, thereby facilitating more informed decisions in agricultural and environmental management. Moreover, the methodology proposed in this study entirely based on open-access data offers a scalable and cost-effective solution for monitoring large-scale agricultural areas. By reducing the reliance on traditional proximal sensing equipment and labor-intensive field surveys, this approach markedly enhances the operational feasibility of precision agriculture, especially in regions where data are scarce or resources are limited.

5. Conclusions

Traditionally, ECa has been relied upon as a primary indicator for assessing soil organic matter content, crucial for effective agricultural management. However, conducting ECa assessments is expensive and often yields spatial maps with significant uncertainty, diminishing their usefulness. To address this, the study proposes extracting a set of variables from the analysis of bare soil synthetic images to replace ECa measurements. Bare soil images are constructed by collecting over time a series of remotely sensed images and isolating pixels with specific spectral signatures related to bare soil. This approach allows for the

creation of bare soil images on a pixel-by-pixel basis, providing a cost-effective alternative to traditional EC_a assessments that also minimizes the need for extensive field data collection. Furthermore, analyzing variables derived from synthetic soil images can offer insights into the underlying processes influencing the spatial distribution of apparent electrical conductivity. By condensing and consolidating the valuable information contained within these variables into a single principal component using PCA technique, the methodology becomes more user-friendly for both scholarly and practical applications. The study utilized spatial statistics techniques, specifically cross-correlation at various lags, to examine the relationship between EC_a and PC1 respective spatial maps. Through this approach, similarities between the two variables were identified and quantified. PC1 serves as a summary of the SYSI variables considered in the study, facilitating comparative analysis through the examination of their respective maps.

Ethical statement

This study did not involve human participants or animals. The research was conducted in accordance with ethical guidelines and best practices in scientific research. The data used in this study were obtained from publicly accessible sources and do not include any sensitive or personal information.

CRediT authorship contribution statement

Matteo Petito: Software. **Emanuele Barca:** Methodology. **Antonio Berti:** Project administration. **Silvia Cantalamessa:** Supervision. **Giancarlo Pagnani:** Formal analysis. **Michele Pisante:** Funding acquisition.

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.

Acknowledgements

The authors gratefully acknowledge Dr. Maria D'Etorre for her valuable contribution to the editing of the manuscript.

Data availability

The data that has been used is confidential.

References

- [1] J. MacPherson, A. Voglhuber-Slavinsky, M. Olbrisch, P. Schöbel, E. Dönitz, I. Mouratiadou, K. Helming, Future agricultural systems and the role of digitalization for achieving sustainability goals. A review, *Agron. Sustain. Dev.* 42 (2022) 70, <https://doi.org/10.1007/s13593-022-00792-6>.
- [2] D.L. Corwin, S.M. Lesch, Delineating site-specific management units with proximal sensors. *Geostatistical Applications for Precision Agriculture*, Springer, Netherlands, 2010, pp. 139–165, https://doi.org/10.1007/978-90-481-9133-8_6.
- [3] R.E. Plant, Site-specific management: the application of information technology to crop production, *Comput. Electron. Agric.* 30 (2001) 9–29, [https://doi.org/10.1016/S0168-1699\(00\)00152-6](https://doi.org/10.1016/S0168-1699(00)00152-6).
- [4] R. Taghizadeh-Mehrjardi, B. Minasny, F. Sarmadian, B.P. Malone, Digital mapping of soil salinity in Ardakan region, central Iran, *Geoderma* 213 (2014) 15–28, <https://doi.org/10.1016/j.geoderma.2013.07.020>.
- [5] J. Kühn, A. Brenning, M. Wehrhan, S. Koszinski, M. Sommer, Interpretation of electrical conductivity patterns by soil properties and geological maps for precision agriculture, *Precis. Agric.* 10 (2009) 490–507, <https://doi.org/10.1007/s11119-008-9103-z>.
- [6] Y. Fu, R. Horton, T. Ren, J.L. Heitman, A general form of Archie's model for estimating bulk soil electrical conductivity, *J. Hydrol. (Amst)* 597 (2021) 126160, <https://doi.org/10.1016/j.jhydrol.2021.126160>.
- [7] D.L. Corwin, S.M. Lesch, Characterizing soil spatial variability with apparent soil electrical conductivity, *Comput. Electron. Agric.* 46 (2005) 103–133, <https://doi.org/10.1016/j.compag.2004.11.002>.
- [8] D.L. Corwin, E. Scudiero, Field-scale apparent soil electrical conductivity, *Soil Sci. Soc. Am. J.* 84 (2020) 1405–1441, <https://doi.org/10.2106/sssaj20153>.
- [9] E. Barca, D. De Benedetto, A.M. Stellacci, Contribution of EMI and GPR proximal sensing data in soil water content assessment by using linear mixed effects models and geostatistical approaches, *Geoderma* 343 (2019) 280–293, <https://doi.org/10.1016/j.geoderma.2019.01.030>.
- [10] D.L. Corwin, S.M. Lesch, Apparent soil electrical conductivity measurements in agriculture, *Comput. Electron. Agric.* 46 (2005) 11–43, <https://doi.org/10.1016/j.compag.2004.10.005>.
- [11] S. Popolizio, A.M. Stellacci, L. Giglio, E. Barca, M. Spagnuolo, M. Castellini, Seasonal and soil use dependent variability of physical and hydraulic properties: an assessment under minimum tillage and No-tillage in a long-term experiment in Southern Italy, *Agronomy* 12 (2022) 3142, <https://doi.org/10.3390/agronomy12123142>.
- [12] S. Popolizio, E. Barca, M. Castellini, F.F. Montesano, A.M. Stellacci, Investigating the spatial structure of soil hydraulic properties in a long-term field experiment using the BEST methodology, *Agronomy* 12 (2022) 2873, <https://doi.org/10.3390/agronomy12112873>.
- [13] L. Poggio, A. Gimona, Assimilation of optical and radar remote sensing data in 3D mapping of soil properties over large areas, *Sci. Total Environ.* 579 (2017) 1094–1110, <https://doi.org/10.1016/j.scitotenv.2016.11.078>.
- [14] J.A.M. Demattè, C.T. Fongaro, R. Rizzo, J.L. Safanelli, Geospatial soil sensing System (GEOS3): a powerful data mining procedure to retrieve soil spectral reflectance from satellite images, *Remote Sens. Environ.* 212 (2018) 161–175, <https://doi.org/10.1016/j.rse.2018.04.047>.
- [15] N.E.Q. Silvero, J.A.M. Demattè, M.T.A. Amorim, N.V. dos Santos, R. Rizzo, J. L. Safanelli, R.R. Poppi, W. de Sousa Mendes, B.R. Bonfatti, Soil variability and quantification based on Sentinel-2 and Landsat-8 bare soil images: a comparison, *Remote Sens. Environ.* 252 (2021) 112117, <https://doi.org/10.1016/j.rse.2020.112117>.
- [16] T. Gorji, E. Sertel, A. Tanik, Monitoring soil salinity via remote sensing technology under data scarce conditions: a case study from Turkey, *Ecol. Indic.* 74 (2017) 384–391, <https://doi.org/10.1016/j.ecolind.2016.11.043>.
- [17] E. Asfaw, K.V. Suryabagavan, M. Argaw, Soil salinity modeling and mapping using remote sensing and GIS: the case of Wonji sugar cane irrigation farm, Ethiopia, *J. Saudi Soc. Agric. Sci.* 17 (2018) 250–258, <https://doi.org/10.1016/j.jssas.2016.05.003>.
- [18] M. Bouaziz, M.Y. Chtourou, I. Triki, S. Mezner, S. Bouaziz, Prediction of soil salinity using multivariate statistical techniques and remote sensing tools, *Adv. Remote Sens.* 07 (2018) 313–326, <https://doi.org/10.4236/ars.2018.74021>.
- [19] O.R. Neto, A. Teixeira, R. Leão, L. Moreira, L. Galvão, Hyperspectral remote sensing for detecting soil salinization using ProSpectIR-VS aerial imagery and sensor simulation, *Remote Sens.* 9 (2017) 42, <https://doi.org/10.3390/rs9010042>.
- [20] E. Costantini, R. Barbetti, M. Fantappiè, G. L'Abate, R. Lorenzetti, R. Napoli, A. Marchetti, R. Riveccio, The soil map of Italy, *GlobalSoilMap*, CRC Press, 2014, pp. 109–112, <https://doi.org/10.1201/b16500-23>.
- [21] J. Louis, B. Pflug, M. Main-Knorn, V. Debaecker, U. Mueller-Wilm, R.Q. Iannone, E. G. Cadau, V. Boccia, F. Gascon, Sentinel-2 Global surface reflectance level-2a product generated with Sen2Cor, in: *IGARSS 2019 - 2019 IEEE International Geoscience and Remote Sensing Symposium*, IEEE (2019) 8522–8525, <https://doi.org/10.1109/IGARSS.2019.8898540>.
- [22] R. Mathieu, M. Pouget, B. Cervelle, R. Escadafal, Relationships between satellite-based radiometric indices simulated using laboratory reflectance data and typic soil color of an arid environment, *Remote Sens. Environ.* 66 (1998) 17–28, [https://doi.org/10.1016/S0034-4257\(98\)00030-3](https://doi.org/10.1016/S0034-4257(98)00030-3).
- [23] L.C. Rowan, J.C. Mars, Lithologic mapping in the Mountain Pass, California area using Advanced Spaceborne Thermal Emission and Reflection Radiometer (ASTER) data, *Remote Sens. Environ.* 84 (2003) 350–366, [https://doi.org/10.1016/S0034-4257\(02\)00127-X](https://doi.org/10.1016/S0034-4257(02)00127-X).
- [24] J. Xiao, Y. Shen, R. Tateishi, W. Bayaer, Development of topsoil grain size index for monitoring desertification in arid land using remote sensing, *Int. J. Remote Sens.* 27 (2006) 2411–2422, <https://doi.org/10.1080/0143160600554363>.
- [25] D. S., Theoretical basis for differentiation of ferric-iron bearing minerals using Landsat MSS data, in: *Proceedings of Symposium for Remote Sensing of Environment, 2nd Thematic Conference on Remote Sensing for Exploratory Geology*, 1982, p. 951.
- [26] A.E.K. Douaoui, H. Nicolas, C. Walter, Detecting salinity hazards within a semiarid context by means of combining soil and remote-sensing data, *Geoderma* 134 (2006) 217–230, <https://doi.org/10.1016/j.geoderma.2005.10.009>.
- [27] J.W. Tukey, *Exploratory Data Analysis*, Reading/Addison-Wesley, 1977.
- [28] E.E. Cureton, R.B. D'Agostino, *Factor Analysis*, Psychology Press, 2013, <https://doi.org/10.4324/9781315799476>.
- [29] N.H. G.D. & Sofroniou, the multivariate social scientist: introductory statistics using generalized linear models., (n.d.).
- [30] M. Castellini, A.M. Stellacci, M. Tomaiuolo, E. Barca, Spatial variability of soil physical and hydraulic properties in a durum wheat field: an assessment by the BEST-procedure, *Water. (Basel)* 11 (2019) 1434, <https://doi.org/10.3390/w11071434>.
- [31] S.P. Friedman, Soil properties influencing apparent electrical conductivity: a review, *Comput. Electron. Agric.* 46 (2005) 45–70, <https://doi.org/10.1016/j.compag.2004.11.001>.
- [32] E. Ben-Dor, S. Chabillat, J.A.M. Demattè, Characterization of soil properties using reflectance spectroscopy. Fundamentals, Sensor Systems, Spectral Libraries, and Data Mining For Vegetation, CRC Press, 2018, pp. 187–247, <https://doi.org/10.1201/9781315164151-8>.

- [33] S. Grunwald, G.M. Vasques, R.G. Rivero, Fusion of soil and remote sensing data to model soil properties, in: 2015: pp. 1–109. <https://doi.org/10.1016/bs.agron.2014.12.004>.
- [34] N.E.Q. Silvero, J.A.M. Dematté, B. Minasny, N.A. Rosin, J.G. Nascimento, H.S.R. Albarracín, H. Bellinaso, A.M.R. Gómez, Sensing technologies for characterizing and monitoring soil functions: a review, in: 2023: pp. 125–168. <https://doi.org/10.1016/bs.agron.2022.08.002>.
- [35] S. Ibáñez-Asensio, A. Marqués-Mateu, H. Moreno-Ramón, S. Balasch, Statistical relationships between soil colour and soil attributes in semiarid areas, *Biosyst. Eng.* 116 (2013) 120–129, <https://doi.org/10.1016/j.biosystemseng.2013.07.013>.
- [36] Z.L. Carroll, M.A. Oliver, Exploring the spatial relations between soil physical properties and apparent electrical conductivity, *Geoderma* 128 (2005) 354–374, <https://doi.org/10.1016/j.geoderma.2005.03.008>.
- [37] W.N. Medeiros, D.S.M. Valente, D.M. de Queiroz, F. de Assis de Carvalho Pinto, I. R. de Assis, Apparent soil electrical conductivity in two different soil types, *Revista ciência agronómica* (2018) 49, <https://doi.org/10.5935/1806-6690.20180005>.
- [38] E.A. Thaler, I.J. Larsen, Q. Yu, A new index for remote sensing of soil organic carbon based solely on visible wavelengths, *Soil Science Society of America Journal* 83 (2019) 1443–1450, <https://doi.org/10.2136/sssaj2018.09.0318>.
- [39] G.I. Metternicht, J.A. Zinck, Remote sensing of soil salinity: potentials and constraints, *Remote Sens. Environ.* 85 (2003) 1–20, [https://doi.org/10.1016/S0034-4257\(02\)00188-8](https://doi.org/10.1016/S0034-4257(02)00188-8).
- [40] E. Scudiero, T.H. Skaggs, D.L. Corwin, Regional-scale soil salinity assessment using Landsat ETM + canopy reflectance, *Remote Sens. Environ.* 169 (2015) 335–343, <https://doi.org/10.1016/j.rse.2015.08.026>.