A Review: Soil properties mapping estimation using remote and proximal sensing data

D. Koumoulidis*[a], A. Efthimiadou[b], N. Katsenios[b], D. Hadjimitsis[a,c]
[a] Eratosthenes Centre of Excellence, Environment & Climate Department, Franklin Roosevelt 82, Limassol, 3012, Cyprus;
[b] Elgo - Demetra, Department of Soil, S. Venizelou 1, Lykovrisi, Athens, 14123, Greece;
[c] Technological University of Cyprus, Civil Engineering & Geomatics Department, 30 Arch. Kyprianos Str., 3036 Limassol, Cyprus
*dimitrios.koumoulidis@eratosthenes.org.cy; phone +35796334451; www.eratosthenes.org.cy

ABSTRACT

Soil is the life support system of terrestrial ecosystems, making comprehending its processes and functions vital for global food security, global climate change regulation, and achieving sustainability in agriculture. Knowledge of soil processes and up-to-date soil status is a prerequisite for sustainable environmental management and reducing decision-making risks.

Over the past decades, the notable introduction and adoption of digital technologies in Remote Sensing (RS), the improvement of spatial data applications, the development of quantitative techniques to understand soil patterns, and the detailed visualization of soils through new applications have increased our capacity to predict, assess, and explain soil and its patterns.

The current review paper, apart from accessing techniques/applications of RS and Proximal Sensing (PS) utilized towards soil mapping predictions, extends its interest in demonstrating the increasingly crucial role of RS and PS utilized to determine soil attributes such as texture, soil moisture, soil organic carbon (SOC) and iron content. Limitations and the difficulties of remote and proximal sensing are overviewed, while additionally, the crucial issue of accuracy of the classification of the thematic maps derived is addressed. Furthermore, this review paper aims to review the status of current mapping methods and provide more profound and more detailed insight into techniques in contemporary systematic soil mapping.

Keywords: Soil mapping methods, Remote & Proximal Sensing, Soil variables prediction, Multivariate statistical techniques.

INTRODUCTION

Over the last decades, implementing advanced geo-information techniques has offered remarkable results towards an accurate and efficient improvement in soil mapping. However, a challenging but promising effort on existing global soil maps is to document and calculate their accuracy precision to any usability issues when combined and examined with other global environmental layers [1]The first attempt to establish relationships between soil formation and soil properties dates to the 1990s [2]. Nowadays, the rapid advancement and availability of Geographic Information Systems (G.I.S.), high spatial and temporal resolution Digital Elevation Models (DEMs), predictive statistical models, and numerous applications for data analysis have advanced the potential within scientific soil surveying.

This paper outlines the fundamental principles of Remote and Proximal Sensing technologies for mapping and predicting soil attributes. It also seeks to assess the current mapping methodologies and provide more profound and detailed insight into contemporary soil mapping techniques. Finally, this review delves into the methods used to determine soil properties such as soil moisture, iron content, soil texture, and SOC.

Tenth International Conference on Remote Sensing and Geoinformation of the Environment (RSCy2024), edited by A. Christofe, S. Michaelides, D. Hadjimitsis, C. Danezis, K. Themistocleous, N. Kyriakides, G. Schreier,
Proc. of SPIE Vol. 13212, 1321207 · © 2024 SPIE · 0277-786X · doi: 10.1117/12.3034339

2. MATERIALS AND METHODS

Various techniques can be employed, including direct, indirect and correlative methods to measure soil attributes. For instance, Ion-Sensitive Field Effect Transistors (ISFET) can directly measure soil pH. At the same time, the indirect approach involves the utilization of Visible-Near Infrared (VIS-NIR) spectroscopy [3]. Spanning from 2006 to 2023, this review paper identified sixteen studies (Table 1) that specifically focus on Digital Soil Mapping Methods (DSMM). The preliminary examination of articles commenced with the amalgamation of keywords like "Digital Soil Mapping", "Remote Sensing", and "Soil Attributes" with "Imaging Spectroscopy" and "Soil Classification" in the Google Scholar search engine. A total of over twenty thousand articles were retrieved during this phase. Subsequent refinement of the article list involved a focus on terms such as "Accuracy-Uncertainty in Digital Soil Mapping" and specific soil attributes like "SOC", "Iron Content", "Soil Moisture", and "Soil Texture". In the final screening process, articles about study areas spanning the globe were chosen for further review. Finally, the following criteria to select studies were appointed:

- Predicted soil attributes
- Sampling numbers of soil samples (N), soil profiles/number of observation sites (SP), number of soil layers/horizons (SL), and frequency/period (F)
- Soil geographic region
- Spatial scale
- Digital mapping methods and multivariate methods employed
- Final R^2 coefficient results
- DSMM assessment mode with model development only (M), cross-validation (CV), and validation (V).

Table 1. DSMM papers reviewed.

References	Soil attribute	Sampling	Spatial Scale	Region	Digital /Multivariate Prediction Models	Coefficient R^2	Assessment mode[1]
Grinand et al (2012) [4]	SOC	N-2086	16*16km grid cells	France	MIR -PLSR	0.89	M, V
Zhang et al (2019) [5]	Soil Organic Matter (SOM)	N-787	Random area sampling	China	NDVI index - SLR, PLSR, ANN, OK	SLR:0.281 PLSR: 0.283 OK: 0.372 ANN: 0.391	M, V
Gholizadeh et al (2018) [6]	Moisture Content (MC), Clay, Silt, Sand	N-40	Random area sampling from 142ha	Malaysia	VIS/NIRS - SMLR	VIS/ Db:0.99, MC:0.82, Clay:0.99 Silt:0.97, Sand:0.96	M, V
Odebiri et al (2022) [7]	SOC	N-1936 SL-0.30cm	1,221,037km^2	South Africa	DNN, (Random Forest) RF, ANN, SVM Sentinel series data	*Deep Neural Networks (DNN)* R^2:67.3, RMSE:10.35(t/ha) *RF* R^2:64.7, RMSE:11.2(t/ha) *ANN* R^2:63.4, RMSE:11.6(t/ha) *SVM* R^2:58, RMSE:13.6(t/ha)	CV
Ben-Dor et al (2009) [8]	SOM, PH, Soil saturated moisture (SM), Electrical Conductivity (EC).	N-62	Sample per 30*30m	Israel	VIS/NIR-SWIR	SOM:0.837, PH:0.883 SM:0.81, EC:0.874	M, V
Vaudour et al (2019) [9]	SOC, CaCO3, Clay, Silt, Sand, Iron, PH, Cation Exchange Capacity (CEC)	N-215 SL (1st site) 0-8cm SL (2d site) 0-5cm	221km^2	France (2 sites)	PLSR	*Indicatively*: 1st Site *SOC* R^2: 0.56, RMSE: 1.23 *CaCO3* R^2: 0.48, RMSE: 20.3 *Clay* R^2: 0.39 RMSE: 1.23 *Silt* R^2: 0.14 RMSE: 103.0	CV

[1] M: model, V: Validation, CV: Cross Validation.

						Sand R^2: 0.22 RMSE: 81.7 *Iron Content* R^2: 0.05 RMSE: 0.09 *PH* R^2: 0.51 RMSE: 0.51 *CEC* R^2: 0.75 RMSE: 1.23	
Bilgili et al (2010) [10]	CaCO3, SOM, PH, EC, Clay, Silt, Sand	N-512 SL-up to 30cm	25*25m grid from 32ha	Turkey	VNIR-PLSR&MARS	*PLSR/* CaCO3:0.65, SOM:0.75, PH:0.30, EC:0.30 Clay:0.84, Silt:0.37, Sand:0.80 *MARS/* CaCO3:0.79, SOM:0.79, PH:0.35 Clay:0.89, Silt:049, Sand:0.84	M, CV
Asfaw et al (2018) [11]	Salinity index (SI) through EC quantifying	-	6.539ha	Ethiopia	Landsat TM data 2012, computation of six RS indices - RMSE	0.78	M, V
Ali Aldabaa et al (2015) [12]	SI through EC quantifying	N-165 from 2 areas SL- 0-5cm	Two sites	USA	VNIR DRS - PLS, SVR PXRF, RS - PLSR	*VNIR DRS/* PLS:0.91, SVR:0.89 *PXRF/* 0.72 *RS (Landsat)/* 0.48	M, CV
Richter et al (2009) [13]	Fe oxide content	N-50 SL-0-2cm	6.200ha	Spain	VNIR DRS – RMSE, rel. RMSE	*RMSE/* (texture-dependent model-sand group):0.87 *rel. RMSE/* (texture-dependent model-sand group):13.9%	M, CV
Dvorakova et al (2023) [14]	SOC	N-124 SL-0-10cm	16.900km^2	Belgium - Netherlands	PLSR-Bootstrapping calibration	*Model Efficient Coefficient (MEC): 0.48 RMSE: 3.5*	M, CV
Swain et al (2021) [15]	Soil texture fractions	N-295 SL-0-10cm	2.712km^2	India	RF – SVR Sentinel series data	RF (sand and silt) *Highest measured value for sand:*	CV

						R^2:0.50 with RMSE 11.6% - *for silt:* R^2:0.46 with RMSE 5.8% SVR (clay) *Highest measured value* R^2: 0.44 with RMSE 6.9%	
Viscarra Rossel et al (2006) [16]	Organic carbon, CEC, Lime requirement, Clay, Silt, Sand, Exchangeable aluminium (Al), Available P, K, EC.	SL-0-20cm	17.5ha	Australia	VIS, NIR, MIR, Combination VIS-NIR-MIR - RMSE	*Quoting the combination method/* OC:0.72, CEC:0.09, LR:0.74 Clay:0.67, Silt: 0.52, Sand: 0.75 Al:0.37, P:0.07, K:0.46, EC:0.29	M, V
Lu et al (2008) [17]	24 soil attributes	N-200 SL-0-15cm	Two crop fields	USA	PMF	-	M
Paul et al (2022) [18]	Total organic carbon (TOC), Permanganate Oxidizable Carbon (POXc)	N-270 SL-0-15cm	3.700ha	Nova Scotia	Landsat 8 Level 2 surface reflectance images (Path 8/Row 28 or Path 7/Row 29), LiDAR-derived DEM.	Quantile Regression Forest, Gradient Boosting Model/ TOC:0.60, POXc:0.40	CV
Kaya et al (2022) [19]	SOC and available P	SL- 0-20cm 280 locations	-	Northeast Iran	Landsat 8 OLI (2018–2021) and Sentinel 2A MSI Level	Random Forest, Cubist packages of the R software/ Available Phosphorus Coefficient of variation (%) 85.74. SOC (%)— Coefficient of variation (%) 45.83	CV

3. RESULTS

3.1. Digital Soil Mapping (DSM)

The orology DSM was initially conceived as the act of generating and fusing spatial soil information through the application of field and laboratory observational methods alongside spatial and non-spatial soil inference systems [20]. The DSM method gained increased popularity for providing detailed and up-to-date data about soils [21];[22]; [23]. DSM presents itself as a flexible, quantitative, and precise method; nevertheless, it is essential to underscore specific significant points. The applicability of DSM models to diverse regions with distinct characteristics and profiles is primarily complex, as the prediction models are developed for each particular area. Furthermore, unlike conventional soil mapping products, DSM maps are designed with specific objectives rather than for general use. A crucial point to consider is the uncertainty parameter in the DSM products since the created models are approximations of reality. Some sources of uncertainty include the covariate accuracy and the model structure [24]. Random and systematic errors found in observations and covariates also represent a significant factor in the uncertainty [25]. The quantification of uncertainty in several models applied within DSM, notably kriging models such as regular kriging and kriging with external drift, has attracted substantial attention [26]. There is a notable uptick in the utilization of machine learning algorithms (ML) within the DSM, which are employed to forecast conditional quantiles [27]. Within uncertainty quantification, DSM utilizes prediction intervals to estimate the likely range of values for an actual measurement. A commonly used threshold in this context is the 90% prediction interval (PI) [21], which denotes the range of values where one can expect to find a new measurement nine times out of ten [28]. The frameworks proposed for assessing uncertainty in DSM are organized into three main classifications: spatial prediction frameworks, probabilistic frameworks, and frameworks that synthesize both spatial characteristics and probabilistic aspects to yield spatially explicit realizations [29].

3.2. Remote Sensing (RS)

A plethora of research studies have been conducted to explore the potential of RS in soil mapping surveys. These investigations have consistently indicated a significant increase in the effectiveness of conventional methods when RS data is incorporated [20].

Spatial models rely on accurate data regarding the continuous variations in soil properties. To ensure the utmost precision, these models require as much detail as possible when utilizing such data as input [30]. One of RS's major drawbacks arises at this point, as the predictions are deemed valid only for the very top layer of soils. It is essential to underscore an additional condition when analyzing RS data, which pertains to making corrections for atmospheric, geometric, radiometric, and topographic effects. An example that is commonly encountered is the interference caused by vegetation, which poses challenges to investigations. Once these corrections have been adequately addressed, geostatistical techniques can be utilized for spatial interpolation, enabling the mapping of spatial patterns in regions where soil data is limited.

The reliability of thematic mapping classification achieved via RS techniques is a crucial determinant of its relevance and precision for multiple applications. No universally established standard for accuracy in thematic mapping is produced through RS methods. Nonetheless, at least 85% accuracy in correct classifications is frequently recognized as the aspirational objective target in thematic mapping [31]. The integration of the error matrix technique, commonly referred to as the confusion matrix, is employed to address the challenge of accuracy in computer-generated mapping [32]. Functioning as a critical, descriptive statistic, the confusion matrix allows for evaluating classification results against established ground truth data. However, constructing and interpreting the confusion matrix is fraught with difficulties. These difficulties are typically rooted in assumptions that are either unattainable or biased [31].

3.3. Proximal Sensing (PS)

PS alludes to a technological advancement which enables the measurement of soil characteristics from a distance of less than 2 meters above the soil surface [33]. PS is predominantly associated with high-resolution soil mapping, usually at resolutions less than 20m. Recognizing the concept of applying PS in soil science is of utmost significance. While measurements' precision may be lower than conventional mapping techniques, PS facilitates the collection of larger volumes of spatial data at a lower cost and through less intricate and more straightforward methods [33]. Table 2 presents the suitability of each technology, while below is a description of three primary PS techniques used.

Table 2. Suitable PSS for acquiring soil information.

Sensing technology	Depth (m)	Dependent property	Inferred properties
Electromagnetic induction (EMI)	<1m to 6m	Resistivity, magnetic, permeability, permittivity	Conductivity, soil water and solutes, texture, temperature, salinity
Electrical resistivity (ER)	<1m to <30m	Resistivity	Conductivity, soil water and solutes, texture, temperature, salinity
Induced polarization (IP)	<1m to <50m	Resistivity, capacitance	Conductivity and polarization, soil water, hydraulic properties, lithology
Magnetic Sensors	<1m to <10km	Magnetization	Magnetic minerals Fe-oxides, structure, lithology
Gravity	>1m to<100km	Density	Lithology, hydraulic properties
Seismic Reflection (SR)	<1m to <500km	Elastic modulus, acoustic impedance, density	Soil layering, soil structure, soil depth, soil density, lithology
Ground penetrating radar (GPR)	<1m to <10m	Resistivity, magnetic permeability, permittivity	Soil water, texture, soil depth
Magnetic resonance sounding	<1m to <1.5km	Proton density	Soil water, porosity, hydraulic properties
γ-radiometrics	<1m	Radioisotopes of Cs, K, U, Th	Total K, mineralogy, clay content, soil type

3.3.1. The Ground-penetrating radar (GPR*)*

GPR is a near-surface geophysical non-destructive technique which creates high-resolution images of the Earth's dielectric properties. The fundamental concept behind this principle is that high-frequency electromagnetic waves have the potential to reflect or refract when they encounter subsurface features that exhibit changes in their electrical properties [34]. GPR makes transmitting and reflecting high-frequency electromagnetic waves within the Earth's subsurface possible. The accuracy and level of detail attained through GPR are determined by selecting different antenna frequencies. Higher frequencies contribute to enhancing the resolution of the images generated. GPR measures the time it takes for a direct ground wave to move from the source to the receiver antenna across the uppermost layer of the ground [35]. Figure 1 presents a simplified schematic design of a GPR.

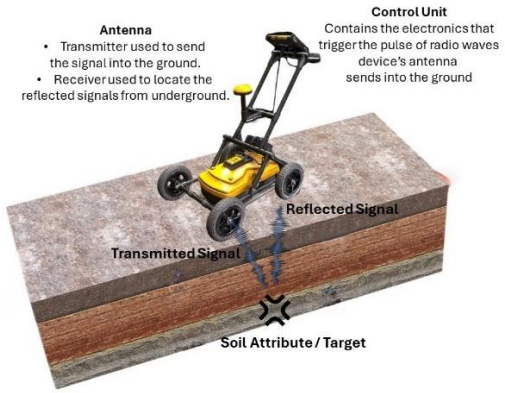

Figure 1. Schematic design of a GPR.

3.3.2. Electrical resistivity (ER)

Soil ER refers to the capacity of soil to resist the flow of electrical current. To determine the apparent ER, various methods utilize Ohm's law in conjunction with the measured injected current, the potential difference, and a geometric factor. A total of four electrodes are required to perform resistance measurements. Two of these electrodes are designated as current

electrodes (A and B) and are responsible for injecting the supply current. The other two electrodes, known as potential electrodes (M and N), record the potential difference [36]. The calculation of the potential difference in the following formula determines ΔV:

$$\Delta V = \frac{\rho l}{2\pi}\left[\frac{1}{AM} - \frac{1}{BM} - \frac{1}{AN} + \frac{1}{BN}\right] \quad (1)$$

AM, BM, AN, and BN refer to the geometrical distance between the electrodes A and M, B and M, A and N, and B and N, respectively. The formula then expresses the ER:

$$\rho = \left[\frac{2\pi}{\left(\frac{1}{AM}\right) - \left(\frac{1}{BM}\right) - \left(\frac{1}{AN}\right) + \left(\frac{1}{BN}\right)}\right]\frac{\Delta V}{I} \quad (2)$$

$$\rho = K\frac{\Delta V}{l} \quad (3)$$

The geometric coefficient K depends on the arrangement of electrodes A, B, M, and N.
ER varies widely, from 1Ω for saline soils to 105Ω per meter for dry soil overlying crystalline rocks. Figure 2 illustrates the typical ranges of resistivity values.

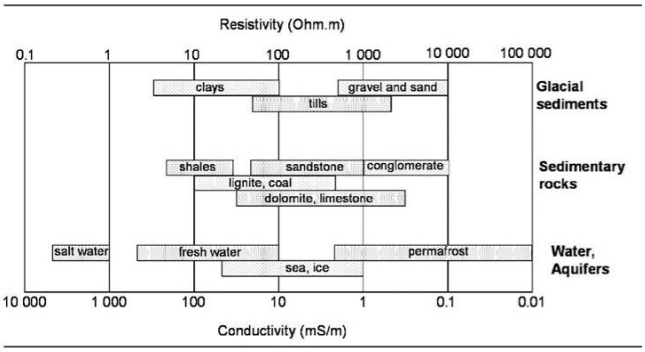

Figure 2. ER region of earth materials (after Palacky, 1987) [36].

3.3.3. Electromagnetic Induction (EMI)

Initially introduced in agriculture, an electro-magnetic sensor was a tool for quantifying soluble salts and the soil moisture index. Electro-magnetic applications in the agricultural field have expanded to include the determination of soil mapping units, measurements of topsoil depth in claypan soils, evaluation of herbicide degradation, and assessment of crop productivity[37].
By utilizing EMI, apparent electrical conductivity can be measured without invasive methods. This measurement is associated with various soil properties, including bulk density, soil structure, ionic composition, pH, SOC, and CaCO3 contents [38]. Such methods to measure ground conductivity involve using a transmitter coil and a single or multiple receiver coil positioned at predetermined intervals.

3.4. Soil Spectroscopy

The application of soil spectroscopy has become prevalent as a quantitative approach for estimating diverse soil properties. Spectroscopy is a quantitative method used in physical and analytical chemistry based on the principle that atoms and molecules absorb radiation in distinct wavelengths, resulting in their unique signature spectra [39]. Spectroscopic methods, such as mass spectroscopy, visible, near-infrared, and mid-infrared spectroscopy, are being acknowledged as potential substitutes for traditional laboratory and field approaches. They aim to enhance and, in some cases, even supplant our comprehension of soils.
Imaging spectroscopy employs airborne or satellite-based hyperspectral sensors to capture the pertinent spectral data spatially. Spectroscopy offers several advantages; however, there are certain drawbacks associated with this technique. One limitation is that it solely relies on reflectance measurements from bare surfaces to assess soil properties.

Consequently, it cannot estimate vertical gradients in SOC within the topsoil. The method is classified as less accurate than conventional methods, such as wet oxidation and dry combustion [9].

3.5. Soil Texture

Vaudour et al. (2019) emphasized the advantages presented by the Sentinel mission in evaluating key top-soil attributes. A multitude of methodologies advocates for the application of chemometrics techniques or the identification of specific absorption features to compute soil properties using the reflectance regions of visible and near-infrared (VIS-NIR, 400-1200nm) and shortwave infrared (SWIR, 1200-2500nm) [40]. Distinctive absorption features can differentiate soils rich in clay from those rich in quartz. Typically, clay minerals are identified by a hydroxyl absorption at 2200nm, whereas thermal bands displaying a reflectance peak corresponding to silica in the range of 8000nm-9500nm indicate the presence of quartz [41].
The correlation between surface soil properties and remotely sensed spectral data at various wavelengths has been explored using various laboratory statistical methods. The two most prominent techniques, multiple linear regression (MLR) and partial least-square regression (PLSR), are commonly used.

3.6. Soil Moisture

Measurement of soil moisture content on the near-surface can be achieved using optical and thermal infrared RS, as well as passive and active microwave RS techniques. Although imaging spectroscopy indices to calculate surface soil moisture content have been adopted using the reflectance in the SWIR region, the accuracy of the results was constrained, primarily due to shallow soil penetration and the interference caused by cloud cover.
The Microwave RS method is a supplementary technique for predicting soil moisture. It involves measuring the contrast in dielectric properties between dry soil and water, which is derived from the reflected data.
The Synthetic Aperture Radar (SAR) is the most extensively employed active microwave configuration for imaging purposes. Despite weather conditions, SAR systems offer high-resolution images suitable for day and night operations. These systems can achieve resolutions in the tens of meters range, covering swath widths up to 500 kilometres, with the best performance observed in space-based systems. SAR technology is widely applied in geoscience, climate change research, 2-D and 3-D mapping, and even 4-D mapping. The estimation of soil moisture has been facilitated by the utilization of various SAR satellites, including the European Remote Sensing (ERS) [42], the Environmental Satellite/Advanced SAR (ENVISAT/ASAR)[43], the RADARSAT [44], the Advanced Land Observing Satellite/Phased Array L-band SAR(ALOS/PALSAR) [45] and the Terra SAR-X platforms [46].

3.7. Soil Organic Carbon (SOC)

The colour of the soil serves as an initial indicator for estimating the levels of soil SOC. Soils with darker hues tend to contain higher quantities of SOM. As a result, the assessment of SOM using imaging spectroscopy proves to be an excellent approach, given its strong correlation with soil colour [8]. VIS-NIR spectroscopy is a valuable method that enables the estimation of soil's physical, chemical, and biological properties based on its reflectance characteristics. Notably, the SWIR (1300-2500nm) and NIR (700-1300nm) regions have been identified as particularly sensitive to SOC levels, with reflectance decreasing as SOC content increases [47].
The endeavor to map SOC has been facilitated by the implementation of sensors, such as the Sentinel series, which have strategically integrated bands that are capable of detecting SOC [6]. The European Space Agency's (ESA) Sentinel-3 Ocean and Land Colour Instrument (OLCI) is a cutting-edge technology that has recently joined the Sentinel series. With an extended sweep width of 1270km, the Sentinel-3 can collect data within the electromagnetic spectrum's VIS to NIR wavelength range, specifically from 400 to 1020nm. These VIS-NIR wavelengths are pivotal in offering reflectance information on SOC and are widely acknowledged as the most sensitive region for accurately assessing SOC concentrations [48].
A diverse set of multivariate statistical techniques and ML algorithms have been employed to map SOC. The Principal Component Analysis [49], Regression Trees [50], Support Vector Machines [51]; [52], and Partial Least Squares Regression [53], are commonly found in the relevant literature for SOC estimations.

3.8. Soil Salinity

Two methods can be used to identify soil salinity: direct detection on exposed soil using remotely sensed data or indirect detection through the analysis of vegetation type, which is influenced by the presence of salts in the soil.
Salt mineralogy determines concentrations of salt through identified absorption bands in the electromagnetic spectrum [54]. Spectral regions that identify soil salinity levels include the visible (0.55–0.77Am), near-infrared (0.9–1.3Am), and middle-infrared (1.94–2.15, 2.15–2.3, 2.33–2.4Am) bands. Six RS indices (Table 6) can be computed to assess salinity, which can then be used to map areas with salt-affected soils.

Table 6. Indices to analyze soil salinity and vegetation indices using Landsat 4-7 sensors [11].

Index name	Formula
Salinity Index (SI)	$\sqrt{Band3 * Band4}$
Brightness Index (BI)	$\sqrt{Band3^2 * Band4^2}$
Normalized Difference Salinity Index (NDSI)	$\frac{Band3 - Band4}{Band3 + Band4}$
Vegetation Soil Salinity Index (VSSI)	$2 * Band2 - 5 * (Band3 + Band4)$
Normalized Differential Vegetation Index (NDVI)	$\frac{(Band4 - Band3)}{(Band4 + Band3)}$
Soil Adjusted Vegetation Index (SAVI)	$\left(\frac{(Band4 - Band3)}{(Band4 + Band3 + 0.5)}\right) * 1.5$

The application of RS is not exempt from difficulties, as the ever-changing process of surface salinity can lead to inaccuracies in identification outcomes. Salt detection may be impeded by vegetation or other surface attributes, causing spectral confusion with salt reflectance. Using space-borne SAR systems represents an approach to mapping areas impacted by salinization. RADARSAT images [55] and Phased Array Type L–L-Band synthetic Aperture Radar (PALSAR) data [56] are commonly employed in the estimation of soil salinity. The justification for employing microwave RS to analyze the salinity factor is rooted in the dielectric characteristics of the soil, as salinity plays a crucial role in electric conductivity. Building upon this rationale, the ML regression model emerges as a primary method for addressing soil salinity, while the random forest (RF) models are frequently employed in ensemble learning approaches [57].

3.9. Iron Content

Employing multispectral or imaging spectrometer images, it becomes feasible to measure the absorption features related to iron oxide and iron hydroxides [8]. Iron exhibits three wide diagnostic absorption bands in the VIS/NIR in its trivalent state. These bands can be the main constituents (in iron oxides) or impurities (in iron smectites). In the case of hematite, these absorption bands are located at 550nm, 630nm, and 860nm, while for goethite, they occur at 480nm, 650nm, and 920nm. The reflectance spectra are predominantly influenced by the first absorption band (Fe-VIS:~500nm), and the third absorption band (Fe-NIR:~900nm) dominate the reflectance spectra [13].
A method to generate the redness index for Hematite content, derived from Thematic Mapper of Landsat, is the following formula: Redness Index (RI) = R2/(B*G3), while from Sentinel-2 general formulas to assess Ferric Oxides and Ferrous Iron is the B11(1600:1700)/ B08(760:860); and B12(2145:2185)/B08/ (760:860) +B03(520:600)/ B04(630:690) respectively.

4. CONCLUSIONS

The integration of technologies related to the collection and analysis of spatial data, such as G.I.S., has made it easier to access and use information and data from different origins and enhance the continuous investigations of spatial data towards a detailed understanding and creation of any thematic map. Despite the benefits obtained through DSM, conventional map strategies will continue to be vital as they gather highly detailed information about mapping units. Nonetheless, those conventional techniques are vulnerable to substantial temporal intervals between the fieldwork, sample retrieval, the determination of mapping limits, and ultimately, the disclosure of the conclusions [59].
RS and PS methods, among others, aim to overcome these obstacles and establish the groundwork for generating and delineating soil. Existing findings in academic literature demonstrate that combining G.I.S. technologies with RS/PS and ancillary data is a highly effective tool for soil science.
In a continuous long-term effort to improve our understanding of soil processes through RS and PS, advances are also deemed necessary to develop the geostatistical analysis using these data entirely. Historically, our knowledge of soil and

evaluation of its quality and functions has been significant. Further developing our analytical approaches and modelling techniques is crucial to achieving a detailed and accurate understanding of a complete system and a resource. Thus, we will achieve a more efficient use and preserve it sustainably for future generations.

ACKNOWLEDGEMENTS

1. The authors acknowledge the 'EXCELSIOR': ERATOSTHENES: Excellence Research Centre for Earth Surveillance and Space-Based Monitoring of the Environment H2020 Widespread Teaming project (www.excelsior2020.eu). The 'EXCELSIOR' project has received funding from the European Union's Horizon 2020 research and innovation programme under Grant Agreement No 857510 from the Government of the Republic of Cyprus through the Directorate General for the European Programmes, Coordination and Development and the Cyprus University of Technology.

2. The present work was carried out in the framework of the CARBONICA project, which received funding from the Horizon Europe HORIZON-WIDERA-2022-ACCESS-04-01—Excellence Hubs programme under Grant Agreement No 101087233.

3. The authors acknowledge the 'GreenCarbonCY': Transitioning to Green agriculture by assessing and mitigating Carbon emissions from agricultural soils in Cyprus. The GreenCarbonCy project has received funding from the European Union - Next Generation, the Recovery and Resilience Plan "Cyprus_tomorrow", and the Research & Innovation Foundation of Cyprus under the Restart 2016-2020 Program with contract number CODEVELOP-GT/0322/0023.

REFERENCES

[1] T. Hengl *et al.*, “SoilGrids250m: Global gridded soil information based on machine learning,” *PLOS ONE*, vol. 12, no. 2, p. e0169748, Feb. 2017, doi: 10.1371/journal.pone.0169748.

[2] I. D. Moore, P. Gessler, G. A. E. Nielsen, and G. Peterson, “Soil Attribute Prediction Using Terrain Analysis,” *Soil Science Society of America Journal - SSSAJ*, vol. 57, Mar. 1993, doi: 10.2136/sssaj1993.572NPb.

[3] R. A. V. Rossel, N. J. McKenzie, and M. J. Grundy, “Using Proximal Soil Sensors for Digital Soil Mapping,” in *Digital Soil Mapping: Bridging Research, Environmental Application, and Operation*, J. L. Boettinger, D. W. Howell, A. C. Moore, A. E. Hartemink, and S. Kienast-Brown, Eds., Dordrecht: Springer Netherlands, 2010, pp. 79–92. doi: 10.1007/978-90-481-8863-5_7.

[4] C. Grinand *et al.*, “Prediction of soil organic and inorganic carbon contents at a national scale (France) using mid-infrared reflectance spectroscopy (MIRS),” *European Journal of Soil Science*, vol. 63, no. 2, pp. 141–151, 2012, doi: 10.1111/j.1365-2389.2012.01429.x.

[5] J. Zhang, Y. Huang, K. Reddy, and B. Wang, “Assessing crop damage from dicamba on non-dicamba-tolerant soybean by hyperspectral imaging through machine learning,” *Pest Management Science*, vol. 75, May 2019, doi: 10.1002/ps.5448.

[6] A. Gholizadeh, “Detecting Vegetation Stress as a Soil Contamination Proxy: A Review of Optical Proximal and Remote Sensing Techniques,” *International Journal of Environmental Science and Technology*, vol. 16, pp. 2511–2524, Apr. 2019, doi: 10.1007/s13762-019-02310-w.

[7] O. Odebiri, O. Mutanga, and J. Odindi, “Deep learning-based national scale soil organic carbon mapping with Sentinel-3 data,” *Geoderma*, vol. 411, p. 115695, Apr. 2022, doi: 10.1016/j.geoderma.2022.115695.

[8] E. Ben-Dor *et al.*, “Using Imaging Spectroscopy to study soil properties,” *Remote Sensing of Environment*, vol. 113, pp. S38–S55, Sep. 2009, doi: 10.1016/j.rse.2008.09.019.

[9] E. Vaudour, C. Gomez, Y. Fouad, and P. Lagacherie, “Sentinel-2 image capacities to predict common topsoil properties of temperate and Mediterranean agroecosystems,” *Remote Sensing of Environment*, vol. 223, pp. 21–33, Mar. 2019, doi: 10.1016/j.rse.2019.01.006.

[10] A. Bilgili, H. van Es, F. Akbas, A. Durak, and W. Hively, "Visible-near infrared reflectance spectroscopy for assessment of soil properties in a semi-arid area of Turkey," *Journal of Arid Environments*, vol. 74, pp. 229–238, Feb. 2010, doi: 10.1016/j.jaridenv.2009.08.011.
[11] E. Asfaw, K. V. Suryabhagavan, and M. Argaw, "Soil salinity modeling and mapping using remote sensing and GIS: The case of Wonji sugar cane irrigation farm, Ethiopia," *Journal of the Saudi Society of Agricultural Sciences*, vol. 17, no. 3, pp. 250–258, Jul. 2018, doi: 10.1016/j.jssas.2016.05.003.
[12] A. A. A. Aldabaa, D. C. Weindorf, S. Chakraborty, A. Sharma, and B. Li, "Combination of proximal and remote sensing methods for rapid soil salinity quantification," *Geoderma*, vol. 239–240, pp. 34–46, Feb. 2015, doi: 10.1016/j.geoderma.2014.09.011.
[13] N. Richter, T. Jarmer, S. Chabrillat, C. Oyonarte, P. Hostert, and H. Kaufmann, "Free Iron Oxide Determination in Mediterranean Soils using Diffuse Reflectance Spectroscopy," *Soil Science Society of America Journal*, vol. 73, no. 1, pp. 72–81, 2009, doi: 10.2136/sssaj2008.0025.
[14] K. Dvorakova, U. Heiden, K. Pepers, G. Staats, G. van Os, and B. van Wesemael, "Improving soil organic carbon predictions from a Sentinel–2 soil composite by assessing surface conditions and uncertainties," *Geoderma*, vol. 429, p. 116128, Jan. 2023, doi: 10.1016/j.geoderma.2022.116128.
[15] S. Swain *et al.*, "Estimation of soil texture using Sentinel-2 multispectral imaging data: An ensemble modeling approach," *Soil and Tillage Research*, vol. 213, Sep. 2021, doi: 10.1016/j.still.2021.105134.
[16] R. Viscarra Rossel, D. J. J. Walvoort, A. Mcbratney, L. Janik, and J. O. Skjemstad, "Visible, Near Infrared, Mid Infrared or Combined Diffuse Reflectance Spectroscopy for Simultaneous Assessment of Various Soil Properties," *Geoderma*, vol. 131, pp. 59–75, Mar. 2006, doi: 10.1016/j.geoderma.2005.03.007.
[17] J. Lu, P. Jiang, L. Wu, and A. C. Chang, "Assessing soil quality data by positive matrix factorization," *Geoderma*, vol. 145, no. 3, pp. 259–266, Jun. 2008, doi: 10.1016/j.geoderma.2008.03.021.
[18] D. Paul and H. Lade, "Plant-growth-promoting rhizobacteria to improve crop growth in saline soils: a review," *Agronomy for Sustainable Development*, vol. 34, no. 4, pp. 737–752, 2014, doi: 10.1007/s13593-014-0233-6.
[19] F. Kaya, A. Keshavarzi, R. Francaviglia, G. Kaplan, L. Başayiğit, and M. Dedeoğlu, "Assessing Machine Learning-Based Prediction under Different Agricultural Practices for Digital Mapping of Soil Organic Carbon and Available Phosphorus," *Agriculture*, vol. 12, no. 7, Art. no. 7, Jul. 2022, doi: 10.3390/agriculture12071062.
[20] H. Wulf, V. L. Mulder, M. E. Schaepman, A. Keller, and P. Joerg, *Remote Sensing of Soils*. 2015.
[21] S. Chen *et al.*, "Digital mapping of *GlobalSoilMap* soil properties at a broad scale: A review," *Geoderma*, vol. 409, p. 115567, Mar. 2022, doi: 10.1016/j.geoderma.2021.115567.
[22] F. Liu *et al.*, "Mapping high resolution National Soil Information Grids of China," *Science Bulletin*, vol. 67, no. 3, pp. 328–340, Feb. 2022, doi: 10.1016/j.scib.2021.10.013.
[23] X. Zhang *et al.*, "Improving model parsimony and accuracy by modified greedy feature selection in digital soil mapping," *Geoderma*, vol. 432, p. 116383, Apr. 2023, doi: 10.1016/j.geoderma.2023.116383.
[24] J. L. Boettinger, D. W. Howell, A. C. Moore, A. E. Hartemink, and S. Kienast-Brown, Eds., *Digital Soil Mapping: Bridging Research, Environmental Application, and Operation*. Dordrecht: Springer Netherlands, 2010. doi: 10.1007/978-90-481-8863-5.
[25] M. Nussbaum *et al.*, "Evaluation of digital soil mapping approaches with large sets of environmental covariates," *SOIL*, vol. 4, pp. 1–22, Jan. 2018, doi: 10.5194/soil-4-1-2018.
[26] A. Moreira, P. Prats-Iraola, M. Younis, G. Krieger, I. Hajnsek, and K. Papathanassiou, "A Tutorial on Synthetic Aperture Radar," *IEEE Geoscience and Remote Sensing Magazine (GRSM)*, vol. 1, pp. 6–43, Mar. 2013, doi: 10.1109/MGRS.2013.2248301.
[27] B. Kasraei, B. Heung, D. D. Saurette, M. G. Schmidt, C. E. Bulmer, and W. Bethel, "Quantile regression as a generic approach for estimating uncertainty of digital soil maps produced from machine-learning," *Environmental Modelling & Software*, vol. 144, p. 105139, Oct. 2021, doi: 10.1016/j.envsoft.2021.105139.
[28] D. Arrouays *et al.*, "Soil legacy data rescue via GlobalSoilMap and other international and national initiatives," *GeoResJ*, vol. 14, pp. 1–19, Dec. 2017, doi: 10.1016/j.grj.2017.06.001.
[29] S. Ma *et al.*, "Application of the water-related spectral reflectance indices: A review," *Ecological Indicators*, vol. 98, pp. 68–79, Oct. 2018, doi: 10.1016/j.ecolind.2018.10.049.
[30] T. Hengl *et al.*, "SoilGrids250m: Global gridded soil information based on machine learning," *PLOS ONE*, vol. 12, no. 2, p. e0169748, Feb. 2017, doi: 10.1371/journal.pone.0169748.
[31] G. M. Foody, "Harshness in image classification accuracy assessment," *International Journal of Remote Sensing*, vol. 29, no. 11, pp. 3137–3158, Jun. 2008, doi: 10.1080/01431160701442120.

[32] A. E. Maxwell and T. A. Warner, “Thematic Classification Accuracy Assessment with Inherently Uncertain Boundaries: An Argument for Center-Weighted Accuracy Assessment Metrics,” *Remote Sensing*, vol. 12, no. 12, Art. no. 12, Jan. 2020, doi: 10.3390/rs12121905.
[33] R. Viscarra Rossel, S. Cattle, A. Ortega, and Y. Fouad, “In situ measurements of soil colour, mineral composition and clay content by vis–NIR spectroscopy,” *Geoderma*, vol. 150, pp. 253–266, May 2009, doi: 10.1016/j.geoderma.2009.01.025.
[34] R. Knight, “Ground Penetrating Radar for Environmental Applications,” *Annual Review of Earth and Planetary Sciences*, vol. 29, no. Volume 29, 2001, pp. 229–255, May 2001, doi: 10.1146/annurev.earth.29.1.229.
[35] D. Singh and A. Kumar, “Binary mixture of nanoparticles in sewage sludge: Impact on spinach growth,” *Chemosphere*, vol. 254, p. 126794, Sep. 2020, doi: 10.1016/j.chemosphere.2020.126794.
[36] A. Samouëlian, I. Cousin, A. Tabbagh, B. A., and G. Richard, “Electrical resistivity survey in soil science: A review,” *Soil and Tillage Research*, vol. 83, pp. 173–193, Sep. 2005, doi: 10.1016/j.still.2004.10.004.
[37] J. G. Davis, N. R. Kitchen, K. A. Sudduth, and S. T. Drummond, “Using Electromagnetic Induction to Characterize Soils,” vol. 81, no. 4, 1997.
[38] J. Doolittle and E. Brevik, “The use of electromagnetic induction techniques in soil studies,” *Geoderma*, vol. s 223–225, pp. 33–45, Jul. 2014, doi: 10.1016/j.geoderma.2014.01.027.
[39] C. Pasquini, “Near infrared spectroscopy: A mature analytical technique with new perspectives – A review,” *Analytica Chimica Acta*, vol. 1026, pp. 8–36, Oct. 2018, doi: 10.1016/j.aca.2018.04.004.
[40] D. Curcio, G. Ciraolo, F. D’Asaro, and M. Minacapilli, “Prediction of Soil Texture Distributions Using VNIR-SWIR Reflectance Spectroscopy,” *Procedia Environmental Sciences*, vol. 19, pp. 494–503, Jan. 2013, doi: 10.1016/j.proenv.2013.06.056.
[41] C. Cristina Barbosa Guimarães *et al.*, “Soil weathering behavior assessed by combined spectral ranges: Insights into aggregate analysis,” *Geoderma*, vol. 402, p. 115154, Nov. 2021, doi: 10.1016/j.geoderma.2021.115154.
[42] C. Wang, J. Qi, S. Moran, and R. Marsett, “Soil moisture estimation in a semiarid rangeland using ERS-2 and TM imagery,” *Remote Sensing of Environment*, vol. 90, no. 2, pp. 178–189, Mar. 2004, doi: 10.1016/j.rse.2003.12.001.
[43] F. Baup, E. Mougin, P. De Rosnay, F. Timouk, and I. Chênerie, “Surface soil moisture estimation over the AMMA Sahelian site in Mali using ENVISAT/ASAR data,” *Remote Sensing of Environment*, vol. 109, no. 4, pp. 473–481, Aug. 2007, doi: 10.1016/j.rse.2007.01.015.
[44] G. Bertoldi, S. Della Chiesa, C. Notarnicola, L. Pasolli, G. Niedrist, and U. Tappeiner, “Estimation of soil moisture patterns in mountain grasslands by means of SAR RADARSAT2 images and hydrological modeling,” *Journal of Hydrology*, vol. 516, pp. 245–257, Aug. 2014, doi: 10.1016/j.jhydrol.2014.02.018.
[45] Y. Izumi *et al.*, “Soil Moisture Retrieval by Means of Adaptive Polarimetric Two-Scale Two-Component Model with Fully Polarimetric ALOS-2 Data,” Jul. 2018, pp. 4619–4622. doi: 10.1109/IGARSS.2018.8518189.
[46] N. Baghdadi, M. Aubert, and M. Zribi, “Use of TerraSAR-X Data to Retrieve Soil Moisture Over Bare Soil Agricultural Fields,” *IEEE Geoscience and Remote Sensing Letters*, vol. 9, pp. 512–516, May 2012, doi: 10.1109/LGRS.2011.2173155.
[47] H. M. Bartholomeus, M. E. Schaepman, L. Kooistra, A. Stevens, W. B. Hoogmoed, and O. S. P. Spaargaren, “Spectral reflectance based indices for soil organic carbon quantification,” *Geoderma*, vol. 145, no. 1, pp. 28–36, May 2008, doi: 10.1016/j.geoderma.2008.01.010.
[48] Q. Lin, H. Huang, J. Wang, L. Chen, H. Du, and G. Zhou, “Early detection of pine shoot beetle attack using vertical profile of plant traits through UAV-based hyperspectral, thermal, and lidar data fusion,” *International Journal of Applied Earth Observation and Geoinformation*, vol. 125, p. 103549, Dec. 2023, doi: 10.1016/j.jag.2023.103549.
[49] A. Chang, G. Pan, A. Page, and T. Asano, “Developing Human Health-related Chemical Guidelines for Reclaimed Waster and Sewage Sludge Applications in Agriculture,” Jan. 2001.
[50] J. Peón, C. Recondo, S. Fernández, J. Fernandez Calleja, E. Miguel, and L. Carretero, “Prediction of Topsoil Organic Carbon Using Airborne and Satellite Hyperspectral Imagery,” *Remote Sensing*, vol. 9, p. 1211, Nov. 2017, doi: 10.3390/rs9121211.
[51] E. Aldana Jague, H. Goswin, A. Macdonald, B. Wesemael, and K. Oost, “UAS-based soil carbon mapping using VIS-NIR (480–1000 nm) multi-spectral imaging: Potential and limitations,” *Geoderma*, vol. 275, pp. 55–66, Aug. 2016, doi: 10.1016/j.geoderma.2016.04.012.
[52] R. Viscarra Rossel and T. Behrens, “Using data mining to model and interpret soil diffuse reflectance spectra,” *Geoderma*, vol. 158, pp. 46–54, Aug. 2010, doi: 10.1016/j.geoderma.2009.12.025.
[53] R. Viscarra Rossel *et al.*, “Diffuse reflectance spectroscopy for estimating soil properties: A technology for the 21st century,” *European Journal of Soil Science*, vol. 73, Jun. 2022, doi: 10.1111/ejss.13271.

[54] G. Metternicht and J. A. Zinck, “Remote sensing of soil salinity: Potentials and constraints,” *Remote Sensing of Environment*, vol. 85, pp. 1–20, Apr. 2003, doi: 10.1016/S0034-4257(02)00188-8.
[55] W. Zhao, B. Wang, and G. Yu, “Antibiotic resistance genes in China: occurrence, risk, and correlation among different parameters,” *Environ Sci Pollut Res Int*, vol. 25, no. 22, pp. 21467–21482, Aug. 2018, doi: 10.1007/s11356-018-2507-z.
[56] I. Nurmemet *et al.*, “Monitoring Soil Salinization in Keriya River Basin, Northwestern China Using Passive Reflective and Active Microwave Remote Sensing Data,” *Remote Sensing*, vol. 7, pp. 8803–8829, Jul. 2015, doi: 10.3390/rs70708803.
[57] A. Myles, R. Feudale, Y. Liu, N. Woody, and S. Brown, “An Introduction to Decision Tree Modeling,” *Journal of Chemometrics*, vol. 18, pp. 275–285, Jun. 2004, doi: 10.1002/cem.873.
[58] E. Ben-Dor *et al.*, “Using Imaging Spectroscopy to study soil properties,” *Remote Sensing of Environment*, vol. 113, Sep. 2009, doi: 10.1016/j.rse.2008.09.019.
[59] E. Weber, H. Hasenack, C. Flores, and P. Fasolo, “GIS as a Support to Soil Mapping in Southern Brazil,” 2008, pp. 103–112. doi: 10.1007/978-1-4020-8592-5_9.