The Pharma Innovation Journal 2023; 12(7): 3268-3275

www.ThePharmaJournal.com

The Pharma Innovation

ISSN (E): 2277-7695
ISSN (P): 2349-8242
NAAS Rating: 5.23
TPI 2023; 12(7): 3268-3275

www.thepharmajournal.com
Received: 19-04-2023
Accepted: 30-05-2023

S Sridevy
Department of Physical Sciences and Information Technology, TNAU, Coimbatore, Tamil Nadu, India

M Nivas Raj
Ph.D. Scholar, Department of Remote Sensing and GIS, TNAU, Coimbatore, Tamil Nadu, India

K Sivakumar
Associate Professor, Department of Remote Sensing and GIS, TNAU, Coimbatore, Tamil Nadu, India

R Ravikumar
Forest College and Research Institute Tamil Nadu Agricultural University, Coimbatore, Tamil Nadu, India

P Kumaresan
Centre for Water and Geospatial Studies, Tamil Nadu Agricultural University, Coimbatore, Tamil Nadu, India

N Balakrishnan
Centre for Water and Geospatial Studies, Tamil Nadu Agricultural University, Coimbatore, Tamil Nadu, India

M Tilak
Forest College and Research Institute Tamil Nadu Agricultural University, Coimbatore, Tamil Nadu, India

J Arockia Stephen Raj
Agricultural College and Research Institute, Killikulam Tamil Nadu Agricultural University, Coimbatore, Tamil Nadu, India

P Jona Innisai Rani
Department of Extension Education and Communication Management, Community Science College and Research Institute, Madurai Tamil Nadu Agricultural University, Coimbatore, Tamil Nadu, India

Corresponding Author:
S Sridevy
Department of Physical Sciences and Information Technology, TNAU, Coimbatore, Tamil Nadu, India

Remote sensing based estimation of soil organic matter: A review

S Sridevy, M Nivas Raj, K Sivakumar, R Ravikumar, P Kumaresan, N Balakrishnan, M Tilak, J Arockia Stephen Raj and P Jona Innisai Rani

Abstract
Investigating the SOM content is tedious and time-consuming. Each of the techniques employed for the assessment have its own constraints and advantages over the other techniques. These techniques range from the analytical experiments to long-term experiments with modelling procedures. (i.e.) employing geostatistical and ML algorithms. As the spatial variability of the soil is wide and complex besides the atmospheric conditions, achieving a homogeneous estimation of the SOM is difficult but considered indispensable in determining the soil health. Analytical means of precise SOM estimation are subjected to the sample SOM degradation and requirement of the high-cost reagents and equipment. Techniques for indirect measurement of SOM utilizing soil organic carbon through pedotransfer function also prevails. Spectroscopy based soil properties estimation are considered valuable for its heterogenous and site-specific assessment through training and validation. The assessment of SOM helps in site-specific and variable rate application of the remedials in the field and helps in maintaining the environmental sustainability and preventing fertilizer loads in the soil. This paper provides a comprehensive review on the selective analytical and spectroscopy means of SOM estimation and its influence over the other soil nutrient properties.

Keywords: SOM, soil organic carbon mapping, hybrid approaches; machine learning

Introduction

Soil is part of the basic skeletal framework performing key environmental functions of an ecosystem; from revitalizing carbon balance, water cycling, soil nutrient elements, habitats, climatic conditions, etc., About 80% of the carbon pool available are directly accounted from terrestrial ecosystem and surface of the soil contributes the largest carbon stock with the total of 1550 Gt of Organic Carbon. Hence, maintaining and stabilizing the carbon stocks are highly essential in enhancing soil fertility status and regulation of global carbon cycle (Duddigan *et al.*, 2019) [1].

The organic carbon stock in the soil is depends on the variability of the SOM content. Hence the measure of SOC helps in indirect qualitative assessment of SOM (Santoiemma *et al.*, 2018) [2]. The major aspect in assessing the organic matter involves managing the SOM from rapid decomposition and reformation. This can be achieved through several of the physio-chemical and biochemical management practices (i.e.) physical management by aggregate facilitation; chemical management by surface mineral organic matter binding; biochemical management through recalcitration (Alain F. Plante *et al.*, 2006) [3]. In general, three basic forms of carbon are available in soil (i.e.) organic, elemental and inorganic. Major loss in carbon content of the soil is associated with SOM mineralization, improper agricultural practices and deviated climatic conditions prevailing over the region, which are usually mitigated by recarbonization management practices. In most of the cases, SOC stored as SOM are lost quickly due to land management practices (i.e.) conversion of natural soils into agricultural lands (Kwiatkowska-Malina., 2018) [4]. Other mitigation strategies involved in reducing the loss of organic carbon subjected to decomposition and mineralization includes biochemical stabilization, physical stabilization and chemical stabilization (Duddigan *et al.*, 2019) [1]. SOM in soils and sediments are characterized from simple carbohydrates to complex fat substances and organic acids. This characterization helps SOM in forming water soluble and insoluble complexes; binding clay minerals; absorbing and releasing soil nutrients. SOM content in soils is classified into three major pools: labile, stable and humins. Further the SOM are classified into two fractions (i.e.) humic substances (HSs) and labile organic matter. (Strosser 2010) [5].

Humification is a key process mechanism in formation of humic substances (HSs) which in turn contributes to the SOM layer of the soil profile. Temporal analysis of the soil use analysis indicated the stability of the humus substances specific to a soil type over the years. Reduced SOM content in the soil is considered one of the eight main threats as presented in EU Soil Thematic Strategy, a fact stated by European Environment Agency, 2010.

OM supplements with increased carbon content and decreased nitrogen are proven to be stable half-life supplements with relatively slower mineralization processes. Hence, careful selection of the fertilizers with exogeneous OM for organic amendments must be made. SOM degradation due to the influence of increased tillage practices modelled by the compartment soil organic carbon and long-term experiments indicated the positive effect of SOC stocks due to no till practices and crop addition factor (Bayer *et al*., 2006) [18]. The long-term experiments on the SOC stocks by varying the influential parameters (i.e.) tillage operations for assessment of the SOM requires processing of the large quantity of the data which are subjected to errors due to environmental variables and manual errors. Sustainable soil management can be achieved through qualitative understanding and quantitative modelling of SOM.

As soil survey procedures provides possibilities for a finite assessment of the properties with decreased sampling points, quantitative assessment of properties majorly lies in the prediction. The analytical and long-term experiment procedures for estimation are generally time consuming and requires the need of high-cost laboratory reagents when the assessment and mapping is done at regional or national level. Other conventional mapping procedure involves spatial characterization based on the discrete soil classes, where the abrupt changes in the soil class boundary is assessed. These procedures were subjected to certain advantages (i.e.) spatial uncorrelation within classes, poor correlation between properties and mapped classes and misrepresentation of the abrupt change.

The extractable spectral features of the soil provide profound possibilities in mapping spatial variability of SOM and other soil nutrient properties. Though the spatial variability can be eliminated through the Digital Soil Mapping (DSM) procedures, within-site variability constrains exists. The variability is accounted by the several natural processes influenced by the factors such as climate, soil type, land use etc., The within-site variability can be excluded by the use of variable rate technology which facilitates the precision agriculture. Precision agriculture tends to the site-specific needs of the soil and the crop through the remote sensing and geospatial techniques, DEM and other climatic variables. (Kingsley John *et al*., 2019) [16].

In this context, the use of soil spectral information obtained through the spectroscopy measures for the model calibration involves the application of the analytical measures for training through Machine Learning (ML) procedures or calibration through employing statistical measures besides the other influential data (i.e.) Climatic data, environmental factors etc., The limitations in calibration and validation through statistical measures involves the selection of the spectral variable from Vis-NIR or MIR spectral bands or determination of the other influential soil parameter.

Earlier, the use of geostatistical framework was prevalent in the spatial prediction of the soil information, which is a linear combination of the environmental covariates and spatial autocorrelated residuals and the prediction at unobserved location estimated through interpolation technique. The geostatistical models are considered for its assumptions on spatial variations and the uncertainty associated with the prediction measures. Conversely, the geostatistical models have several limitations which affects the model fit and the prediction accuracy. The limitations include the stationarity of the residuals, increase in the parameter estimated and the increased computational load due to the increased sample size. As an alternative, Machine learning approaches are employed for their increased efficiency when compared to the geostatistical models. "Machine learning techniques refer to a large class of non-linear data-driven algorithms employed primarily for data mining and pattern recognition purposes, and now frequently used for regression and classification tasks in all fields of science." (Alexandre M.J-C. Wadoux *et al*., 2020) [7].

Unlike geostatistical models, machine learning techniques are void of assumptions and can process a large number of parameters. As conventional models (Geostatistical and statistical) are model-oriented and the predictive accuracy depends on the assumptions that makeup the model whereases machine learning techniques are data-driven and the predictions are made from the predictive model calibrated using an error-minimization process. This makes the model calibrated through the machine learning techniques more accurate than conventional models.

Table 1: Effect of soil properties in the formation of the SOM

Soil Property	Soil Processes
Soil structure	Aggregation, organic matter turnover, retention, and transportation of water and chemicals
Porosity	Plant available water capacity, soil crusting, aeration, water entry
Infiltration	Soil water availability and movement, leaching of nutrients, erosion
Bulk density	Soil structural conditions, compaction
Available water	Field capacity, permanent wilting point, water flow
pH	Soil acidification, salinization, soil structural stability, biological and chemical activity thresholds
Electrical conductivity	Plant and microbial activity thresholds, leaching of salts, soil structure decline, salinization
Plant available N, P, and K	Availability of nutrients for plant uptake, losses from the soil–plant system
Soil organic matter	Organic matter storage and quality, plant residue decomposition, metabolic activity of soil organisms, mineralization–immobilization turnover, microbial activity, nutrient supply
Total soil C and N	C and N mass and balance, soil structure, nutrient supply.

***Source*:** Adapted from Jat, Mangi L., Clare M. Stirling, Hanuman S. Jat, Jagdish P. Tetarwal, Raj K. Jat, Rajbir Singh, Santiago Lopez-Ridaura, and Paresh B. Shirsath. "Soil processes and wheat cropping under emerging climate change scenarios in South Asia." *Advances in Agronomy* 148 (2018): 111-171.

The spatial variability of the soil is generally accounted by its influential parameters. The parameters include soil physical, chemical and microbial properties besides the climatic condition prevailing (Table 1). The effect of the influential parameters will have a characteristic impact on the SOM content due to its limited viability and lower lifetime when compared to that of the other properties. Hence accounting the effects of influential parameters and its characteristic study is required to employ long term experiments and pedotranfer functionalities if possible and required. This chapter review the effect of inverse effects of SOM and other properties.

Soil physical properties

The increase of the SOM is highly correlated with the management practices (i.e.) tillage practices and system adopted and are highly regarded for the acidic soils. The effect of SOM on other soil physical properties are considered to in order to implicate the pedotransfer functions. SOM is a composition of plant and animal debris including other compounds containing carbon with fractions of active (unprotected C) and stable (protected C) portions. The active portions are readily decomposable to enrich the fertility content and the stable portions remains stable in order to maintain the integrity of the soil. Rapid decomposition and reformation can lead to nutrient leaching and hence a recurrent and stable cycle must be facilitated. This process is influenced by that of the physical properties. Important physical properties that influence the SOM are detailed in this chapter. Most of studies defined the soil temperature as a key factor that influences the decomposition rates of the organic matter present. The decomposition rates are higher in tropical region when compared to that of the temperate zones. Hence maintaining labile organic matter is much more difficult and requires additional management practices. Literatures stated that the decomposition rates are doubled for each increase of 8-9 ^{o}c in the mean annual air temperatures.

Soil texture is an essential soil physical property that influences almost every other soil property in conjunction. The increase in the SOM is subjected to the increase in the clay textural proportions. The process is implicated via two different mechanisms. The first mechanics involves formation of the bonds that limits the degradation and the other mechanics involves influence of the aggregate particles or occludes the organic matter and. The former involves formation of the stable complex mineral structure (unprotected C) and mineral association (protected C) between the clay minerals and OM and the latter involves increase in the aggregate stability or aggregation which is tended to increase with increase in the clay content. Alain F. Plante Under similar climatic conditions, the clay soils have organic matter three to four time that of the sandy soils. The particle size of each of the textural classes are depicted in the table 2.

Table 2: Particle size of the respective textural classes

Textural class	Particle size (Diameter)
Sand	2 to 0.2 mm
Slit	0.2 to 0.002 mm
Clay	< 0.002 mm

***Source*:** Adapted from Abd-Elmabod, Sameh K., Antonio Jordán, Luuk Fleskens, Jonathan D. Phillips, Miriam Muñoz-Rojas, Martine van der Ploeg, María Anaya-Romero, Soad El-Ashry, and Diego de la Rosa. "Modeling agricultural suitability along soil transects under current conditions and improved scenario of soil factors." In Soil mapping and process modeling for sustainable land use management, pp. 193-219.

As Kaolinite has lower smaller specific surface and nutrient exchange capacity, the clay-humus complexes formation is limited. This results in the leaching and degradation of the organic matter in kaolinite soils under wet tropic conditions as the degradation is favored through the prevailing soil and climatic condition. Alain F. Plante *et al*., 2006 [3] studied the effect of different textural classes on the C in the physically and chemically defined pools of the organic matter based on the measurable fractions. The study resulted in no considerable relationship between the soil textural classes and the unprotected organic C and stated that the clay soils or the clay and silt soils may not be always a best indicator of the soil organic matter has the organic C are subjected to changes through several direct and indirect mechanisms.

Soil Structure is related to the stability of the soil and has largely influenced by the organic matter. The structure involves formation of the micro-aggregates, resulted from that of the interaction between the organic matter and the mineral particles. This holds the soil particles as a firm structure based on the prevalence of the organic matter as it generates organic compounds (i.e.) polysaccharides that as a binding agent. The macro-aggregates are formed from linking of the micro-aggregates, facilitated by the fungal hyphae and the fine root zone of the plants. Decrease in the soil organic content affects the structure of the soil and results in the low infiltration rate, increased run off and soil erosion.

Soil moisture content is usually associated with the increased biomass production, which results in the increased organic residues and the increased organic matter. In this case, the increase is usually derived from the increase in the mean annual precipitation. Biological activity of the soil usually requires air and moisture. Though increased moisture is key benefactor parameter, the increased moisture content results in filling up the pores and deprives microbes of their oxygen. This affects mineralization process of the plants and slower decomposition rates of the plant residues which may lead to the reduced productivity. Similarly, dry soils considerably also have a slower decomposition rate, with other parameters being constant.

The soil-water relationship is also facilitated by the increased organic matter as it is promoted by the increase aggregate formation, which in turn increases the macropore formation and improves water retention and infiltration capacity of the soil. The dry and wet seasons of the humid, sub-humid and semi-arid tropics are subjected to the increased nitrate formation during their first rainy season. The nitrogen mineralization invoked by the first few rains results in the formation of the labile organic matter (Fao.org.in).

Soil chemical properties

The study of soil chemical properties and their interactions on the organic matter as a quantified variable is essential in modelling their formation and decomposition rates. Considering the spatial variability of the chemical properties and their heterogeneity in composition over the field conditions, maintaining an optimum condition is required for the growth of soil microbial population. Though the microbial population has the much more direct influence on the degradation and the breakdown of the plant residues, chemical properties of soil have its own interaction in the mineralization of the soil organic matter. This chapter reviews the effect of the chemical properties in the mineralisation of the organic matter.

Despite that the acidity of the soil is key parameter in influencing the mineralization of the soil organic matter, quantification of its effects in the agronomic system are difficult to access. Many of the modelling procedures for measuring the soil organic matter (i.e.) century and Rothamsted model, excluded the effect of the pH while calibrating the model. Studies revealed a progressive decrease in the organic matter when the pH below about 6. The rate of the mineralization respective of the pH change also depends upon the size of the organic matter pool. Denis curtin *et al*., 1997 [8] studied the effect of pH on the mineralization of the C and N with $Ca(OH)_2$ treatment to increase the pH of the soil samples. The treatment composition has been altered based on the acidity exchangeable proportions for each of the soil types. The increase in the pH resulted in the loosening of the bond and showed an evident increase in the mineralization process. Salinity of the soil also has a characteristic or negative effect on the Organic matter of the soil. Increase in the sodium ion content leads to the leaching and depletion of the organic matter.

Soil microbial properties

The role of soil microbes in synthesizing SOM is very evident and several of the model depicting the quantification of the SOM included or supplemented the microbial activity as a variable for the SOM formation. Some of the stable SOM formed shown the characteristic feature of the microbial enzymes and cytoplasmic materials and excretions rather than the properties of the plant material. The conversion of the plant residues to that of the microbial residues followed by its stabilization determines the effect of the surface and the climatic conditions on the SOM. The SOM formation in the soils of low biological activity is subjected to the destabilization and leaching. The quantitative assessment of the SOM is typically hindered by the microbial biomarkers such as lipids and amino acids as they cannot be scaled to a whole-soil basis. In general, the plant residues are degraded by that of the microbial community to microbial residues of organo-mineral compounds includes ligand bonds with low temperature sensitivity.

Cynthia M. Kallenbach *et al*., 2016 [9] studied and provided results on the direct effect of the microbial community on the formation of the stable SOM besides the ecophysiological functions that affects the formation. The study provided information on the plant residue decomposition by the microbial community and its increased importance than that of the clay mineral content. The study involved modeling for a quantitative assessment of the microbial activity on the simple C low molecular weight substrate rather than the complex plant system. Gradient system at different levels of simple C substrate was established to hypothesize or depict the formation of the microbial colonies and their interactions on the C substrate with different treatments (i.e.) sugar; syringol; sterile sand and kaolinite; dissolved organic carbon and field soils. The model used resulted in defining the importance of the microbial communities on the SOM formation over the effect of clay mineralogy and other known control SOM dynamics.

Recent and advanced analytical measures of SOM

Space borne remotely sensed imagery has an immense potential as an enabling tool for the generation of spatial maps of the upper soil horizon, owing to the proven background in interlinkages among soil's specific chemical bonds and electromagnetic radiation. Optical satellite multispectral imagery started to be used extensively in quantitative SOC characterization with the launch of the first satellites in the 1980s (Frazier and Cheng, 1989) [19]. Applications based on hyperspectral data became popular several years later when the Hyperion spaceborne system became operationally available (Castaldi, *et al.,* 2014). Until now, their use was limited for soil observation due to (i) the required atmospheric, geometric and radiometric data corrections, (ii) simultaneous ground observations, (iii) the difficulty in finding large bare soil areas within a single image (Demattê, *et al.* 2018) [10] and (iv) obstacles related to vegetation cover (Barnes, *et al.* 2003) [20]. Consequently, there are few studies using satellite sensors for SOC estimation (Croft, *et al.* 2003) [21]. Currently, SOC estimation and mapping based on spaceborne data is undergoing a significant shift. The relevant USGS policy change, that enabled Landsat data to be distributed at no charge, can be considered a major milestone to that direction (Woodcock, *et al*., 2008) [22]. Furthermore, this is driven by the advent of the Big earth observation data era, spearheaded by Sentinel-2 free and open super spectral imagery, as well as by the emergence of large fleets of small satellites (e.g., Planet Cubesats, (www.planet.com). In addition, the forthcoming hyperspectral sensors, such as the Environmental Mapping and Analysis Program (EnMAP) (Stuffler, *et al*., 2007) [11], will soon provide unprecedented data streams (high spatial, spectral and temporal resolution) for the retrieval and hence monitoring of SOC, across the VNIR–SWIR spectral range.

Remote sensing based assessment of SOM

RS techniques vary depending on their spatial, spectral, temporal and radiometric resolution and the platforms that are mounted. Selecting the proper technique depends on the field of application, the measured property and the expected accuracy. These technologies have shown their great use for monitoring environmental parameters towards management of natural resources and their rapidly increasing use is due to the significant advancements in terms of sensors specifications. Sensors mounted on satellite platforms have improved from panchromatic to multispectral and the forthcoming hyperspectral, such as EnMAP, HyspIRI, and PRISMA. Hence the availability of these sophisticated hyperspectral sensors, could expedite RS applications in the field of agriculture, while contribute to an advancement of operational applications for environmental purposes. Subsequently, they could provide valuable information on soils' condition and SOC estimation either directly or by providing auxiliary data. Consequently, they could supply the necessary data for accurate and up-to-date soil maps to meet the current and future needs for soil monitoring.

The main advantages of RS applications can be summarized as follows: (i) they are a non-destructive way to gather information about soil properties, (ii) the provided data cover large geographical areas, (iii) they can provide information about inaccessible areas, (iv) they provide data that hold information for several attributes, (v) they have the ability to provide concise data and (vi) provide the means to reduce traditional and laborious soil sampling campaigns.

Soil organic matter (SOM) is the organic matter content of soil, consisting of plant and animal detritus at various stages of decomposition, cells and tissues of soil microbes, and substances that soil microbes synthesize (Brady & Weil, 2008) [12]. SOM provides numerous benefits to soil physicochemical properties (such as preserving soil moisture, increasing soil stability, and protecting soil structure) and its capacity to provide regulatory ecosystem services (Lehmann & Kleber, 2015) [13]. In addition, SOM can be a source or sink for atmospheric CO2 depending on land use, and soil management, vegetation and water resources (Lal *et al*., 2021) [14]. Increased carbon stocks in the soil increase soil fertility, workability, water holding capacity, and reduce erosion risk. Increasing soil organic matter can thus reduce the vulnerability of managed soils to future global warming (Schmidt *et al*., 2011; Smith, 2012) [15, 25]. Therefore, SOM is especially critical for maintaining soil functions, and its content is considered a reliable indicator for evaluating soil quality and degradation (Franzluebbers, 2002) [16].

There is a wide absorption region of SOM in the visible bands (Stenberg *et al*., 2010) [17]. In general, a higher SOM content of soil corresponds to darker colors. Spectral reflectance decreases with an increase in soil organic matter content, mainly affected by the soil chromophore and the dark humic substances (Shi, Wang, *et al*., 2014) [24]. In terms of the near infrared region, the overtones and combinations due to SOM are mainly attributed to the stretching and bending of NH, CH, and CO groups (Stenberg *et al*., 2010) [17]. Considering the explicit relations between SOM and spectral reflectance, qualitative detection and quantitative estimation of SOM based on RS and proximal sensing methods have always received much research attention

Many scholars have used multispectral RS images to estimate and map spatial patterns of SOM across different regions, scales, and soil types (Stenberg *et al*., 2010; Wang *et al*., 2010) [17, 27]. These studies have demonstrated that the optical multispectral RS data, including Moderate-resolution Imaging Spectroradiometer (MODIS), Landsat series, Advanced Space borne Thermal Emission and Reflection Radiometer (ASTER), China & Brazil Earth Resource Satellite (CBERS), Chinese Gaofen-1 (GF-1), Chinese Huanjing satellite constellation-1 (HJ-1), Systeme Probatoire D'Observation De La Terre (SPOT) series, Advanced Land Observation Satellite (ALOS), and IKONOS present great potential for the assessment and prediction of the SOM distributions, especially at large scale (Angelopoulou *et al*., 2019; Sullivan *et al*., 2005) [23]. However, coarse spectral resolution and cloud cover always hampered the employment of these data. In this case, researchers have introduced hyperspectral RS and portable VNIR spectrometers to increase the performance of SOM estimation (Mulder *et al*., 2011) [28]. Mallah Nowkandeh *et al*. (2018) [28] investigated the potential of Hyperion imagery for estimating SOM content and compared several regression technologies. Based on Hyperion spectra, Wang *et al*. (2010) [27] mapped the SOM in Mu Us desert based on land degradation spectral response units. For field/laboratory VNIR spectroscopy, Shi, Wang, *et al*. (2014) [24] developed a national VNIR soil spectral library (1581 soil samples) to predict SOM concentrations, providing an alternative for soil degradation monitoring. Rossel *et al*. (2016) [30] analyzed a global soil VNIR spectral library and explored the usefulness of the global spectra for predicting soil attributes such as soil organic carbon. The results suggest that the global VNIR library describes soil variation and that the spectra provide an integrative measure of the soil, which can be used for both qualitative and quantitative soil analyses. In fact, SOM is one of the typical soil properties with high spatial variability. However, the estimated SOM content is the individual value at a single point, and these data are actually discrete in space. Therefore, it is hard to characterize its spatial patterns only using field/laboratory VNIR spectroscopy (Angelopoulou *et al*., 2019) [23]. Furthermore, a specific SOM content (20 g/kg) seems to be a threshold, that is, if the SOM drops below 2% (commonly found in degraded lands), it became less accurate for the evaluation of the spectral response (Ben-Dor, 2002) [24]. In other words, some other compositions, such as salinity and iron oxide, affect soil spectral behaviors more significantly (Ben-Dor *et al*., 2002) [24]. Thus, how to remove these interferences has attracted much attention in recent years. For example, Liu *et al*. (2018) introduced the External Parameter Orthogonalisation (EPO) preprocessing algorithm to reduce the potential effects of iron oxide and moisture. Wang *et al*. (2014) [14] established a predictive model with better accuracy and stability based on fractional-order derivative pretreated spectra. Considering salt-affected soils, Zhang, *et al*. (2021) [46] developed some strategies for efficiently estimating SOM through NIR spectroscopy. These advanced approaches provide new solutions and reliable support for accurately estimating SOM.

Remote sensing methodologies provides rapid, non-destructive means of estimation of SOM at places that are inaccessible for chemical and analytical means of estimation. Though quantification of the SOM through chemical analytical means are found to be accurate, spatial variability of the SOM cannot be accounted via conventional means. Hence, remote sensing methodologies can be used to account the spatial variability, which in turn can be utilized for the site-specific management and formation of the prescription map. In general, the SOM estimation through remote sensing measures can be facilitated through geo-statistical modelling and through the Digital soil mapping procedures.

Digital soil mapping – SOM assessment and mapping

Why will we need more DSM?

- All plausible futures require more soil information and the utility promised through DSMA. A focus on alternative futures allows us to explore the breadth of that potential need.
- New options arise in the 'outlook vision' view. The breadth of land uses grows and the emphasis on matching those choices to a more nuanced understanding of soil capacity will be crucial. As corporate sustainability goals become increasingly common, consumer demands for sustainably produced food will increase (Thomson *et al*. 2020) [38]. The condition of the soil becomes a reported asset within the supply chain which will potentially grow sustainable producers' profit relative to those producing food in unsustainable ways. In both the slow decline and outlook vision scenarios, the focus on managing soil capacity and resilience, and thus the need for appropriate soil information, will need to increase but through different drivers, the former as a crisis or restoration need, the latter as part of the accent on information-led productivity and capacity optimization.

Prior soil information can be used as a covariate in digital mapping of soils (McBratney *et al*., 2003) [29]. Soil classes and properties such as texture, bulk density and clay mineralogy can explain the variability in SOC. Some studies in digital mapping of SOC have successfully used legacy soil maps as the covariates for predicting SOC. For example, reference soil group was an influential predictor of SOC stocks in a semiarid steppe ecosystem in China (Wiesmeier *et al*., 2011) [31]. However, it had less influence than land use and more influence than geological units for predicting SOC in their study. Categorical soil map of 1:50,000 scale was taken by Adhikari *et al*. (2014) as an environmental variable to map soil organic carbon stocks and content in Denmark and the relative usage of the soil map in their prediction model was reported to be above 60% for predicting SOC. Various soil attributes have been demonstrated to be influential in determining SOC content. Soil properties, such as texture, mineralogy and bulk density, were reported to be strongly associated with SOC, especially below the depth of 20 cm (Badgery *et al*., 2013; Jobbagy and Jackson, 2000). Bulk density was found to be an important explanatory variable for SOC content in the study of Hobley *et al*. (2015). It was negatively correlated with SOC content. However, the relationship was not linear. Influence of site factors including bulk density was found to increase with depth. Total nitrogen content was highly correlated with the spatial distribution of SOC stocks (Were *et al*., 2015). However, using soil nitrogen content as a predictor may not be feasible due to unavailability of the data to cover large geographic extents. However, such data might be available at farm-scale mapping. Soil texture data is highly recommended as a covariate for DSM of SOC as far as available as it has been demonstrated to be highly correlated with SOC content (Zinn *et al*., 2005). While legacy soil maps may be useful covariates, the spatial coverage and mismatch in location in relation to recent covariates may limit their use. 5.5. Validation and the mapping of uncertainty Minasny *et al*. (2013) [32] reported that half of the SOC mapping studies using a digital mapping approach had not validated their work. In our review, all of the studies were found to have performed validation. Of 120 studies, external validation was carried out in 11 studies, which is more than in the review of Minasny *et al*. (2013) [32] that reported only 3 studies to have performed external validation for predictive mapping of SOC. Majority of the studies in our review used data-splitting technique for evaluating the results, followed by cross-validation. Regarding the mapping of uncertainty, 49 of 120 articles were found to have presented spatially explicit estimation of uncertainty for predicting SOC. Contrary to the earlier review by Minasny *et al*. (2013) [32], studies that used ML techniques have also performed spatially explicit assessment of uncertainty. This shows a clear progress in the context of validating the predictive mapping of SOC. However, most of the studies that used data splitting technique have claimed it to be an independent and external validation. Brus *et al*. (2011) [33] define data-splitting as a form of internal validation. Furthermore, according to Brus *et al*. (2011) [33], when the original point sample data set is not collected through probability sampling, data-splitting cannot yield random samples that could provide unbiased and valid estimates of the quality measures and associated estimation errors for validating the predictive mapping in DSM framework. Most of these studies were following the data-splitting or cross-validation method even for the soil sample datasets that were collected through purposive, haphazard or convenient sampling techniques. In addition to mapping of SOC concentration and stocks, there have been some trends in mapping SOC in other dimensions using DSM approach. Chen *et al*. (2018) [10] mapped carbon sequestration potential in France and found subsoils to have larger potential to sequester SOC compared to topsoils. Some recent DSM studies have mapped changes in the level of SOC concentration and stocks with respect to current and projected land use/land cover and climate change scenarios (Gray and Bishop, 2016; Yigini and Panagos, 2016; Zhou *et al*., 2019a) [30]. Mapping in such dimensions can assist in visualizing different probable scenarios and proactive planning of projects to target enhanced sequestration of SOC.

Conclusion

Following the systematic mapping approach, this paper reviewed various algorithms and environmental covariates used in the digital mapping of SOC concentration and stocks in the recent past and their suitability. It identified geographic clusters and gaps regarding the empirical knowledge in the field of digital mapping of SOC. There is an uneven spatial distribution of empirical studies aimed at mapping SOC using digital mapping approaches. Studies are clustered in several countries, namely China, Australia and the USA. Regarding the temporal trend of studies, from 2013 onward, the number of publications reached a maximum in 2016 and 2017, but decreased substantially after that time until 2018. Regarding the predictive models, there has been a shift from Linear to ML ones in comparison to the earlier review in 2013. Although RF was found to be better than other algorithms in most of the comparative studies, no single model was found to be the strongest in all circumstances. Regression Kriging or hybrid models combining the modelling of deterministic and stochastic errors were superior to the separate models that either dealt with deterministic parts or interpolate only using the spatial autocorrelation of SOC. Among various predictive models, there were a significant number of primary studies in relation to some promising algorithms, namely RF, Cubist, BRT, SVM, NN and GWR. Therefore, in order to achieve rigorous comparison of these models, a meta-analysis approach is recommended to assess the most competitive algorithms. However, for other algorithms, primary research is still required to fill existing knowledge gaps. The relationship of environmental covariates to soil carbon levels was found to depend principally upon environmental conditions, depth of soil, resolution of mapping, and the extent of the area under concern. For mapping at regional extents, climate was reported to be the most important factor in SOC levels, followed by parent materials, topography and land use. However, for mapping at a resolution that represents plots or small fields, variation in land use was claimed to be more influential in predicting SOC. Local variation in topography was also stated to be influential for determining SOC level. Minasny *et al*. (2013) [32], reported topographic variables as the most widely used covariates for predicting SOC. However, our review shows that variables representing ‘organisms’ factor are among the most frequent ones among top five important covariates, followed by the covariates representing ‘climate’ and then ‘topography’ factors. While better models and covariates are important for improving the prediction accuracy, other factors such as the size and the representativeness of the training samples are equally significant for predictive mapping of SOC. In comparison to an earlier review by Minasny *et al*. (2013) [32], it has become a more common practice to validate the SOC mapping tasks and estimate spatially explicit uncertainty of prediction using statistical methods in order to improve reliability and accuracy of SOC estimation.

However, additional probability sampling for evaluating the predictive performance was still not found in most studies reviewed, probably due to the additional resources and time needed for this practice. Most studies used data-splitting and claimed it to be an independent evaluation of the results. It is recommended to perform external validation using soil sample datasets collected through additional probability sampling approaches for an unbiased assessment of the prediction of SOC concentration and stocks.

References

1. Adhikari, Bikash, Sanjay N Khanal, Dhiraj Giri, Janardan Lamichhane. Seasonal variation of pH, BOD, COD and BOD/COD ratio in different ages of landfill leachate in Nepal. Journal of Biomolecule Reconstruction. 2014;11(2):89-99.
2. Angelopoulou T, Tziolas N, Balafoutis A, Zalidis G, Bochtis D. Remote sensing techniques for soil organic carbon estimation: A review. Remote Sensing. 2019 Mar 21;11(6):676.
3. Badgery, Warwick B, Aaron T Simmons, Brian M Murphy, Andrew Rawson, Karl O Andersson, *et al*. Relationship between environmental and land-use variables on soil carbon levels at the regional scale in central New South Wales, Australia. Soil Research. 2013;51(8):645-656.
4. Barnes WL, Dereux A, Ebbesen TW. Surface plasmon subwavelength optics. nature. 2003 Aug;424(6950):824-30.
5. Bayer JK, Sanson AV, Hemphill SA. Parent influences on early childhood internalizing difficulties. Journal of Applied Developmental Psychology. 2006 Nov 1;27(6):542-59.
6. Ben-Dor A, Chor B, Karp R, Yakhini Z. Discovering local structure in gene expression data: the order-preserving submatrix problem. In Proceedings of the sixth annual international conference on Computational biology; c2002 Apr 18 p. 49-57.
7. Brady, Nyle C, Ray R Weil, Ray R Weil. The nature and properties of soils. Upper Saddle River, NJ: Prentice Hall; c2008. p. 13.
8. Brus DJ, Kempen B, Heuvelink GB. Sampling for validation of digital soil maps. European Journal of Soil Science. 2011 Jun;62(3):394-407.
9. Castaldi, Peter J, Jennifer Dy, James Ross, Yale Chang, George R. Washko, *et al*. Cluster analysis in the COPD Gene study identifies subtypes of smokers with distinct patterns of airway disease and emphysema. Thorax. 2014;69(5):416-423.
10. Chen, Yuanjun, Shufang Ji, Shu Zhao, Wenxing Chen, Juncai Dong, Weng-Chon Cheong, *et al*. Enhanced oxygen reduction with single-atomic-site iron catalysts for a zinc-air battery and hydrogen-air fuel cell. Nature communications. 2018;9(1):5422.
11. Croft SL, Coombs GH. Leishmaniasis–current chemotherapy and recent advances in the search for novel drugs. Trends in parasitology. 2003 Nov 1;19(11):502-8.
12. Curtin Denis, Rostad HPW. Cation exchange and buffer potential of Saskatchewan soils estimated from texture, organic matter and pH. Canadian Journal of Soil Science. 1997;77(4):621-626.
13. Demattê, José Alexandre Melo, Caio Troula Fongaro, Rodnei Rizzo, José Lucas Safanelli. Geospatial Soil Sensing System (GEOS3): A powerful data mining procedure to retrieve soil spectral reflectance from satellite images. Remote Sensing of Environment. 2018;212:161-175.
14. Duddigan Sarah, Liz J Shaw, Paul D Alexander, Chris D, Collins. A comparison of physical soil organic matter fractionation methods for amended soils. Applied and Environmental Soil Science; c2019. p. 1-12.
15. Franzluebbers, Alan J. Water infiltration and soil structure related to organic matter and its stratification with depth. Soil and Tillage research. 2002;66(2):197-205.
16. Frazier BE, Cheng Y. Remote sensing of soils in the eastern Palouse region with Landsat Thematic Mapper. Remote Sensing of Environment. 1989 Apr 1;28:317-25.
17. Gray JM, Bishop TF, Wilford JR. Lithology and soil relationships for soil modelling and mapping. Catena. 2016 Dec 1;147:429-40.
18. Hobley, Eleanor, Brian Wilson, Arjan Wilkie, Jonathan Gray, and Terry Koen. Drivers of soil organic carbon storage and vertical distribution in Eastern Australia. Plant and Soil. 2015;390:111-127.
19. Jobbágy, Esteban G, Robert B Jackson. The vertical distribution of soil organic carbon and its relation to climate and vegetation. Ecological applications. 2000;10(2):423-436.
20. Kallenbach, Cynthia M, Serita D Frey, Stuart Grandy A. Direct evidence for microbial-derived soil organic matter formation and its ecophysiological controls. Nature communications. 2016;7(1):13630.
21. Kingsley, *et al*. Predictive mapping of soil properties for precision agriculture using geographic information system (GIS) based geostatistics models. Modern Applied Science. 2019;13(10):60-77.
22. Kwiatkowska-Malina, Jolanta. Qualitative and quantitative soil organic matter estimation for sustainable soil management. Journal of soils and sediments. 2018;18:2801-2812.
23. Lal, *et al*. Soils and sustainable development goals of the United Nations: An International Union of Soil Sciences perspective. Geoderma Regional. 2021;25:e00398.
24. Lehmann, Johannes, Markus Kleber. The contentious nature of soil organic matter. Nature. 2015;528:60-68.
25. McBratney AB, Santos MM, Minasny B. On digital soil mapping. Geoderma. 2003 Nov 1;117(1-2):3-52.
26. Minasny B, McBratney AB, Malone BP, Wheeler I. Digital mapping of soil carbon. Advances in agronomy. 2013 Jan 1;118:1-47.
27. Mulder VL, De Bruin S, Schaepman ME, Mayr TR. The use of remote sensing in soil and terrain mapping-A review. Geoderma. 2011 Apr 15;162(1-2):1-9.
28. Nowkandeh, Sina Mallah, Ali Akbar Noroozi, Mehdi Homaee. Estimating soil organic matter content from Hyperion reflectance images using PLSR, PCR, MinR and SWR models in semi-arid regions of Iran. Environmental development. 2018;25:23-32.
29. Plante, Alain F, Richard T. Conant, Catherine E. Stewart, Keith Paustian, and Johan Six. Impact of soil texture on the distribution of soil organic matter in physical and chemical fractions. Soil Science Society of America

Journal. 2006;70(1):287-296.
30. Rossel RA, Viscarra T, Behrens E, Ben-Dor DJ, Brown JAM, Demattê Keith D, *et al*. A global spectral library to characterize the world's soil. Earth-Science Reviews. 2016;155:198-230.
31. Santoiemma, Giacomo. Recent methodologies for studying the soil organic matter. Applied Soil Ecology. 2018;123:546-550.
32. Schmidt, *et al*. Persistence of soil organic matter as an ecosystem property, Nature. 2011;478:49-56.
33. Smith PK. Cyberbullying: Challenges and opportunities for a research program-A response to Olweus. European journal of developmental psychology. 2012 Sep 1;9(5):553-8.
34. Stenberg, Bo, Raphael A, Viscarra Rossel, Abdul Mounem Mouazen, Johanna Wetterlind. Visible and near infrared spectroscopy in soil science. Advances in agronomy. 2010;107:163-215.
35. Strosser, Eduard. Methods for determination of labile soil organic matter: an overview. Journal of Agrobiology. 2010;27(2):49.
36. Stuffler, *et al*. The En MAP hyperspectral imager-An advanced optical payload for future applications in Earth observation programmes. Acta Astronautica. 2007;61(1-6):115-120.
37. Sullivan, Dana G, Shaw JN, Doug Rickman. IKONOS imagery to estimate surface soil property variability in two Alabama physiographies. Soil Science Society of America Journal. 2005;69(6):1789-1798.
38. Thomson, Allison M, Erle C Ellis, HEctor Ricardo Grau, Tobias Kuemmerle, Patrick Meyfroidt, *et al*. Sustainable intensification in land systems: trade-offs, scales, and contexts. Current Opinion in Environmental Sustainability. 2019;38:37-43.
39. Wadoux, Alexandre MJC, Budiman Minasny, Alex B McBratney. Machine learning for digital soil mapping: Applications, challenges and suggested solutions. Earth-Science Reviews. 2020;210:103359.
40. Wang H, Bi N, Saito Y, Wang Y, Sun X, Zhang J, *et al*. Recent changes in sediment delivery by the Huanghe (Yellow River) to the sea: causes and environmental implications in its estuary. Journal of Hydrology. 2010 Sep 24;391(3-4):302-13.
41. Wang YP, Zhou LS, Zhao YZ, Wang SW, Chen LL, Liu LX, *et al*. Regulation of G 6 PD acetylation by SIRT2 and KAT9 modulates NADPH homeostasis and cell survival during oxidative stress. The EMBO journal. 2014 Jun 17;33(12):1304-20.
42. Were, Kennedy, Dieu Tien Bui, Øystein B Dick, Bal Ram Singh. A comparative assessment of support vector regression, artificial neural networks, and random forests for predicting and mapping soil organic carbon stocks across an Afromontane landscape. Ecological Indicators. 2015;52:394-403.
43. Wiesmeier M, Barthold F, Blank B, Kögel-Knabner I. Digital mapping of soil organic matter stocks using Random Forest modeling in a semi-arid steppe ecosystem. Plant and soil. 2011 Mar;340:7-24.
44. Woodcock CE, Allen R, Anderson M, Belward A, Bindschadler R, Cohen W, *et al*. Free access to Landsat imagery. Science. 2008 May 23;320(5879):1011-.
45. Yigini, Yusuf, Panos Panagos. Assessment of soil organic carbon stocks under future climate and land cover changes in Europe. Science of the Total Environment. 2016;557:838-850.
46. Zhang, Chiyuan, Samy Bengio, Moritz Hardt, Benjamin Recht, Oriol Vinyals. Understanding deep learning (still) requires rethinking generalization. Communications of the ACM. 2021;64(3):107-115.
47. Zinn, Yuri L, Rattan Lal, Dimas VS Resck. Changes in soil organic carbon stocks under agriculture in Brazil. Soil and Tillage Research. 2005;84(1):28-40.