Data driven approach on in-situ soil carbon measurement

Umesh Acharya, Rattan Lal & Ranveer Chandra

To cite this article: Umesh Acharya, Rattan Lal & Ranveer Chandra (2022) Data driven approach on in-situ soil carbon measurement, Carbon Management, 13:1, 401-419, DOI: 10.1080/17583004.2022.2106310

To link to this article: https://doi.org/10.1080/17583004.2022.2106310

© 2022 The Author(s). Published by Informa UK Limited, trading as Taylor & Francis Group

Published online: 08 Aug 2022.

Submit your article to this journal

View related articles

View Crossmark data
ABSTRACT
Soil carbon (C) plays a key role in mitigating and adapting to global climate change. In-situ soil C measurement has faced many challenges including those related to aerial coverage, economics, accuracy, and availability. The concept of paying for C credits to farmers and ranchers who sequester C has necessitated availability of improved methods for in-situ measurement of soil C at large scale. The objective of this review is to (i) synthesize the existing knowledge on methods of soil C measurement, (ii) discuss their pros and cons (iii) review key factors affecting soil C measurement, and (iv) propose integrated data driven method of soil C measurement using Machine Learning (ML)/Artificial Intelligence (AI) approach. Lab and in-situ techniques of soil C determination are expensive, time consuming and lack scale. Although, remote sensing (RS) technique is used to predict soil C maps at large scale, it also lacks accuracy and requires high technical knowledge of image processing. Soil C measurements are affected by key soil physical properties such as color, texture, moisture content, bulk density etc. Thus, these factors must be considered while developing innovative methods for soil C determination. A prototype handheld device is proposed to measure these four properties along with Near Infrared (NIR) reflectance of soil that store data in cloud using Wi-Fi signals. A data driven model is proposed that can use the data from handheld devices and integrate with drone imagery to create soil C map of the entire field and satellite imagery for the entire region. This model uses data from in-situ soil C measurement technique in integrated form and soil C map can be updated every time the handheld device is used at different locations of the field.

KEYWORDS
dry combustion; wet oxidation; soil moisture content; soil bulk density; soil strength; satellite image; machine learning

Introduction
Globally, soil holds total carbon (C) stock of \( \approx 2500 \text{ Pg} \) (1Pg = \( \approx 10^{15} \)g) to 1-m depth, which is approximately three times of that in the atmosphere (800 Pg) [1–4]. Of the 2500 Pg of total C stock; 1550 Pg is soil organic carbon (SOC) and 950 Pg is soil inorganic carbon (SIC) [1, 5]. In 2019, about 4.8 billion hectares (B ha), 38% of global land area, is agricultural land, of which one-third (1.6 B ha) is cropland and two-thirds (3.2 B ha) is meadows and pasture for grazing livestock [6,7]. Since 1990, the cropland area has increased by 5% but permanent meadows and pastures decreased by 4%, with an overall decrease of agricultural land by 1% [7]. The human population was \( \approx 250 \) million (M) in circa 1000 AD, increased to 6.1 B by 2000 and is projected to reach 9.8 B by the year 2050 [8,9]. This growing trend of population will create greater demand for food and put pressure on finite land and water resources [7]. Agriculture, forestry, and other land uses (AFOLU) contribute \( \approx 23\% \) of total anthropogenic greenhouse gas (GHG) emission (2007-2016), of which 12-14% is contributed by the agriculture land use [10]. The conversion of natural vegetation to agricultural land increases atmospheric CO2 concentration due to increase in mineralization of SOC, accelerated loss of erosion and reduction of input of biomass -C [11–13]. Several studies [14–16] have documented the loss of SOC stock upon conversion of natural to agricultural ecosystems. However, the depleted SOC stock can be re-sequestered through adoption of improved rotations with deeper rooting cultivars and species, use of organic amendments, agroforestry [17–19]; along with conversion to reduced and no-till (NT) practices [20,21] and incorporation of soil organic matter (SOM) into subsoil [22]. In contrast, high inputs of fertilizers and use of conventional tillage can release soil C stock into the atmosphere [23]. Strategies of SOC protection and sequestration contribute 47% of...
total of total potential mitigation (2.3 PgCO$_2$e yr$^{-1}$) from grassland and agriculture, while 20% involves other GHGs related with improved soil management practices [24]. UN-Sustainable Development Goals (SDGs) can also be advanced by increasing or protecting soil C stock by long-term increase of soil fertility, maintaining resilience to climate change, reducing soil erosion, and improving habitat conversion [25]. The United Nations Framework Convention on Climate Change (UNFCCC) at Congress of Parties (COP) 21 Lima-Paris Plan of Action adopted new program of the Global Climate Action Agenda (GCAA), known as “4 per 1000” initiative (https://www.4p1000.org/). It aspires to increase SOM and C sequestration through implementation of agriculture practices such as agro-ecology, regenerative agriculture, agroforestry, conservation agriculture (CA) or landscape management at the annual rate of 0.4% to 40 cm depth [26,27]. Presently, the attention is also directed to land-based efforts to reduce C emission, remove CO$_2$ from the atmosphere and provide monetary credits to landowners [28,29].

The concept of C credit was introduced during the UNFCCC in Kyoto, Japan in 1997 [30]. A C credit is a tradable certificate or permit representing the right to emit one metric ton (megagram or Mg) of CO$_2$ and has been used by many companies to sell C credits to commercial and individual customers [31,32]. This mechanism also provides monetary value to the farmers or landowners who are encouraged to protect or improve soil C stock following sustainable agriculture methods [33]. The U.S. 117th congress passed the Growing Climate Solutions Act of 2021 that authorizes USDA to establish a voluntary GHG technical assistance provider and third-party verifier certification program to help reduce entry barriers into voluntary environmental credit markets for farmers, ranchers, and private forest landowners [34]. This plan is proposed to convert diverse regions across the U. S. into C sinks, aimed at offsetting the nation’s 7Pg CO$_2$e of GHG emission each year [35].

Private companies (i.e. Bayer, CIBO, ESMC, Gradable, INDIGO, NORI, Soil and Water Outcomes Fund, TRUTERRA) are also connecting with farmers to follow sustainable agricultural practices (i.e. NT, cover cropping) for 5-10 years and paying them a US$7.5 to 24 ha$^{-1}$ yr$^{-1}$ [36,37]. These payments are based on baseline field information, history of agriculture farming and the future plan. Farmers have to sign a contract for 5-10 years and follow the specific guidelines to get paid for the C credit [36,38]. The major problems for the farmers lie in getting full benefits of C credit by using an easy and affordable soil C measurement method [37].

Smallholder farmers are defined as households who farm less than 2 hectares (ha) of land size and obtain annual economic revenue from the same farm [39]. Smallholder farming dominates the agricultural farming in sub-Saharan Africa, South Asia (i.e. India, China) where food security depends on how small holders use their limited resources and traditional knowledge to feed 80% of those populations [40,41]. The UN Food System Summit through its Action Track 3: Boosting Nature Positive Production suggested an opportunity to encourage small holder farmers in Africa and elsewhere to improve soil health and fertility by using the C credit program [42]. The farm size in Asia and Africa is decreasing and the number of smallholder farmers are in an increasing trend since 1950 and projected to increase through 2050 [43]. A study conducted by Shames [44] in Africa showed that out of many challenges (project management, cost, monitoring, unstable international policy), inaccessible and expensive methods of soil C measurement was the main challenge faced by smallholder farmers. An economical and accessible handheld device is instrumental in measuring soil C and that will boost the C credit for all farm sizes around the globe.

Researcher from the academic and industry sector are working to develop data-driven model that includes designing a handheld device, using cloud to store data and artificial intelligence to predict soil organic carbon or SOC [45,46]. Ewing et al. [45] proposed affordable, accessible handheld device (Our Sci Reflectometer; www.our-sci.net) an open-source hardware tool licensed under the GNU General Public License v3.0. This reflectometer using the co-variates field estimable textural class and slope class provides unbiased ($r^2=0.57; n=1155; p=0.06$) and actionable [area under curve (AUC)=0.88] data at field scale when compared with African Soil Information Services (AFSIS; www.soilgrids.org). Yard Stick (www.useyardstick.com) is also developing a handheld device that integrate Near Infrared (NIR) reflection, resistance sensor to measure bulk density and GPS locator to predict SOC using artificial intelligence. These initiatives show the future scope of data-driven soil carbon prediction at academic and industry sector.

Soil C sequestration is a potential climate solution [24], and its credible and economic measurement is key to addressing global climate change
Since the 1990s, several advanced analytical methods have been developed to estimate soil C stock \[47 – 50]\). Methods of soil C measurement can be categorized into three groups: laboratory methods, remote sensing techniques and in-situ procedures \[51]\). These measurement methods have their own merits, limitations, and challenges in terms of cost involved, laboratory equipment required, and accessibility to farmers \[52]\). Therefore, the objective of this article is to (1) synthesize the existing knowledge about the current methods of soil C measurement, (2) discuss pros and cons of those methods, (3) review key factors affecting soil C measurement, and (4) propose integrated methods to measure soil C using Machine Learning (ML)/Artificial Intelligence (AI) approach under in-situ conditions.

### Methodology

The systematic review of literature was done thorough Google Scholar, Web of Science, Scopus using key search terms: (soil carbon measurement) and (texture OR color OR moisture OR remote sensing OR UAVs OR SOC OR in-situ OR machine learning OR Artificial Intelligence OR laboratory OR bulk density OR carbon credit). Scopus identified 6111 search results, Web of Science 9136, and Google Scholar 36600 results of literature for string search. These results were refined by using soil C measurement in agricultural practices and fields resulting in 352 articles in Scopus, 223 in Web of science and 300 in Google Scholar. The bibliographic details were imported into endnote and duplicates were eliminated by applying exclusion-inclusion criteria and examining the title and keywords. Finally, 161 potentially usable articles were selected for this review.

These papers were categorized based on the objectives of this review. Literature related to soil C measurement (laboratory, remote sensing, and in-situ) were reviewed and critically analysed to prepare the pros and cons of each method under different field studies. Four key soil properties were identified, and tables developed on their effect on soil C measurement. Literature was also collated on use of machine learning and Artificial Intelligence (AI) in predicting soil C at different scales (field, farm and regional). Finally, with the understanding of the different aspects of soil C management a data driven model was suggested to predict soil C at different scales.

### Current methods of soil carbon measurement

Direct measurement of soil C being difficult, some methods are more direct, and some others are far from direct with assumptions and sources of error \[51, 53]\). Most available methods calculate soil C at point source (lab methods) \[47]\) or at large scale (remote sensing or high-tech methods) \[49, 54, 55]\). Laboratory analysis of soil C by the dry combustion method is considered as the Gold Standard \[56]\), but it is expensive and time consuming to map large areas \[52]\). Remote sensing techniques can be useful for large scale measurement and mapping of soil C \[57]\) but have numerous errors \[58]\). The following section discusses commonly used laboratory, remote sensing, and in-situ methods of soil C measurement.

### Laboratory methods

Dry combustion or elemental analysis and wet oxidation are commonly used laboratory methods for soil C measurement \[47, 59, 60]\). Dry combustion method is regarded as the standard method to conduct soil C analysis of field collected samples in a laboratory \[60]\). This method is considered to measure soil C with high precision and accuracy \[61, 62]\). It is quick and reliable with the use of an automatic elemental analyzer. In the dry combustion method, SOC is oxidized, and carbonate (CO$_3^-$) minerals are thermally decomposed in a medium temperature resistance furnace. The CO$_2$ produced is then trapped in a suitable reagent and determined titrimetrically or gravimetrically. The elemental analyzers to measure soil C by the dry combustion method are produced by the different companies \[60]\). However, the working principle is same for all, and uses a subsample of oven dried soil (40 °C) sieved through 250 microns ($\mu m = 1 \times 10^{-6}$ m) mesh. Some of the commonly used elemental analyzer for C analysis are LECO TruSpec CN, St. Joseph, MI \[63]\); Carlo-Erba Model NA 1500 analyzer, Milan, Italy \[64]\); Perkin-Elmer CHN2400, Northwalk, CT \[65]\); Thermo Scientific Flash 2000, Massachusetts, USA \[66]\); VELP Scientifica CN 802 CN elemental analyzer, Italy \[67]\).

The oldest method of soil C measurement, the wet oxidation method \[68]\), is still used in many countries \[5, 69]\). This method is simple, rapid and needs minimal equipment \[69, 70]\). Some major problems with the wet oxidation method include disposal of the waste produced during the procedure and being less accurate in soils with...
carbonates [71]. Wet oxidation method is based on oxidation of the C using potassium dichromate (K₂Cr₂O₇) in sulfuric acid (H₂SO₄). In the wet oxidation method, only the most active organic C is oxidized, leading to incomplete oxidation of organic compounds [60].

The detailed process of Walkley Black (wet oxidation) method for soil C measurement is described by Nelson and Sommers [60]. Oven dried (40 °C) and finely grind [250 μm sieved] 1 g soil is placed in an Erlenmeyer flask (125 ml), and 10 ml of 0.2 M potassium dichromate solution is added. Then 10 ml concentrated sulfuric acids slowly added to this solution. The soil sample with higher amount of SOC (>4%) requires a higher amount of potassium dichromate and sulfuric acid solution. After 30 min of oxidation under room temperature, 50 ml (cm³) of distilled water, 3 ml (cm³) of concentrated orthophosphoric acid (H₃PO₄) and four drops of the diphenylamine indicator are added. SOC content is determined by titration of the excess potassium dichromate using Mohr’s salt solution (0.1 M). Each set of soil samples require three blank reagents of Mohr’s salt solution to record its exact molarity. SOC content in the soil is determined using the following equation (i):

\[ gC\frac{kg^{-1}}{g} = \frac{(V_b - V_s) \times C_{Fe^{2+}} \times 0.003 \times 1000}{Wt.\ of\ sample\ (g)} \]

where, \(V_b\) and \(V_s\) are the volumes of Mohr’s salt solution used for the titration of the blank and the soil sample, respectively; \(C_{Fe^{2+}}\) is the molarity of the Mohr’s salt solution; 0.003 g mmol⁻¹ represents the ratio [(0.012)/4], where 0.012 is the molecular mass of the C (g mmol⁻¹), and 4 refers to the number of electrons involved in the oxidation of SOC, and wt. refers to the mass of soil sample (g).

This method generally underestimates the SOC content compared to that by the dry combustion method [71]. The SOC content measured by the wet oxidation method is adjusted with a C equivalent correction factor of 1.33 [68]; 1.63 for Russian Chernozemic soil [71]; 1.16 to 1.59 for a range of soils [60]; 1.05 for subtropical soil in southern Brazil [72]. The correction factor in the Walkley Black method is suggested because the temperature obtained by the H₂SO₄ dilution is not sufficient to totally oxidize all soil organic compounds that depends on organic C recovery under different soil and management systems [73,74]

**Remote sensing technique**

Conventional methods for soil C measurement are expensive and time consuming [52]. Scientific communities are looking for alternative methods to accurately predict soil C at large scale with limited resources in terms of cost and time [75–77]. Large scale soil C mapping is challenging due to variation of soil type, crops grown, tillage systems, and landscape characteristics [78]. Remote sensing techniques provide opportunity for cost-effective, non-destructive, rapid, and large-scale mapping of soil C [50]. These methods are based on the electromagnetic radiation that is radiated on soil surface and is reflected in distinct wavelength and energy providing specific spectral signature [79,80]. These specific reflectance on the spectrum, known as spectral signature, derive qualitative and quantitative information on the soil properties [81,82]. The near infrared (NIR) - short wave infrared (SWIR) spectroscopy ranging from 700-2500 nm (1 × 10⁻⁹m) is based on the characteristic’s vibrations of chemical bonds in a molecule [83]. NIR-SWIR have weak overtones and combinations of these vibrations occur due to stretching and bending of the N-H, O-H and C-H bonds [84,85]. Remote sensing (RS) techniques use different platforms to get the spectral signature such as spaceborne, airborne and unmanned aerial vehicles (UAVs) [80].

Spaceborne RS imageries, such as optical and multispectral satellites, have been used in SOC quantification from 1980 after the launch of Landsat [86]. Hyperspectral imagery started gaining popularity later after the Hyperion spaceborne system was operational [87–89]. Spaceborne imageries are popular because they capture the entire landscape and are available online to download and some are free (USGS), but they require atmospheric, geometric, and radiometric correction [90,91]. Airborne hyperspectral imaging offers accurate mapping of the SOC variability observed at the agricultural field level [92,93]. Airborne devices can produce information in single flight missions that cover large areas and are used to extend the existing dataset of soil properties to support the digital soil mapping [94]. UAVs have been a revolution since circa 2000 because they are novel and low cost and advancement in the sensor specifications [95]. The RS using UAVs provide high spatial resolution, temporal flexibility, and narrow-band spectral data from different wavelengths [96]. With the improvement in the analytical capabilities for data handling, UAVs can
be handy and useful in identifying soil properties [97].

The data derived from the remote sensing techniques can be used as auxiliary variables to develop high quality soil C maps [98,99]. Several studies have shown that the RS data (e.g. brightness, wetness, and vegetation indices) are strongly correlated with soil C as well as first and second derivative of digital elevation models [100]. The direct quantification of the soil C using spectral signature is challenging and requires multivariate statistical methods such as partial least square regression (PLSR) [101,102] or machine learning algorithms (random forest, support vector) [103,104]. Studies circa 2000 using different remote sensing platforms have shown huge potential in soil C measurement using various machine learning algorithms (Table 1).

### In situ techniques

**In-situ** technique involves minimum soil disturbance operated in dynamic or static mode for scanning larger area, enable repetitive-sequenced measurement [49, 106]. Soil C measurement using classical lab methods are not only expensive but also time consuming and further complicated by the temporal and spatial variability inherent in soil horizons, bulk density, and soil C concentration [108]. **In-situ** soil C measurement techniques are important because they are rapid, potentially cost-effective and reduce sampling and laboratory errors [49]. Other benefits of **in-situ** techniques include minimum soil disturbance and ability to analyze large areas providing repetitive and sequential measurements to evaluate spatial and temporal variation [109,110]. Research attempts have been made to make economic, handy, and easily accessible devices to measure soil C at field level [45, 49]. Recent advances in **in-situ** instrument are based on Laser Induced Breakdown Spectroscopy (LIBS) [111], inelastic neutron scattering (INS) [106], near-infrared spectroscopy [112,113].

The LIBS method is based on the atomic emission spectroscopy, where laser pulse is focused on a sample and microplasma emission is recorded in time and spectrally resolved by a time-gated sensor to detect concentration of elements based on the unique spectral characteristics [114,115]. Some studies showed that C emission line at 247.8 nm is used for testing and calibration of LIBS and reported higher correlation ($r^2=0.96$) [115], whereas other studies [111] showed spectral interference between C and Fe at 247.8 nm line and suggested 193 nm line. Calibration curves are required for each sample set and soil specific calibration databases are required for the rapid analysis in the future [116,117].

**Inelastic neutron scattering (INS)** is a noninvasive method to determine soil C and was first proposed by Wielopolski *et al.* [118]. It is based on the inelastic scattering of 14 MeV neutrons from C nuclei present in the soil and measurement of the resulting 4.44 MeV gamma ray emission. The neutrons are produced by the Deuterium-Tritium generator and gamma ray emission is detected by NaI detectors. To measure soil C, calibration lines are derived using the mixture of sand and predetermined amounts of C power (0, 2%, 5%, and 10% by weight) [106, 118]. The INS method is advantageous because it does not need sample preparation and offers true sequential measurements at large area scanning. Wielopolski *et al.* [119] found higher correlation ($r^2=0.99$) between soil C estimated by INS and dry combustion in organic soil.

### Table 1. Soil organic carbon estimation using different remote sensing platforms.

<table>
<thead>
<tr>
<th>Sensors</th>
<th>Spectral range (nm)</th>
<th>$r^2$</th>
<th>RMSE (%)</th>
<th>RPD $%$</th>
<th>Soil depth (cm)</th>
<th>Algorithm</th>
<th>Country</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Hyperion</td>
<td>400-2500</td>
<td>0.66</td>
<td>0.61</td>
<td>1.69</td>
<td>0-10</td>
<td>PLSR</td>
<td>Australia</td>
<td>[105]</td>
</tr>
<tr>
<td>Agrispec Portable</td>
<td>350-2500</td>
<td>0.51</td>
<td>0.73</td>
<td>1.43</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hyperion + Agrispec</td>
<td></td>
<td>0.67</td>
<td>0.60</td>
<td>1.75</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Landsat 2 A</td>
<td>45B-2500</td>
<td>0.56</td>
<td>1.32</td>
<td>1.51</td>
<td>0-8</td>
<td>PLSR</td>
<td>Italy</td>
<td>[106]</td>
</tr>
<tr>
<td>Landsat (using SFC) Y</td>
<td>430-2400</td>
<td>0.43</td>
<td>1.01</td>
<td>–</td>
<td>0-5</td>
<td>Random Forest</td>
<td>Australia</td>
<td>[99]</td>
</tr>
<tr>
<td>without SFC</td>
<td></td>
<td>0.48</td>
<td>0.43</td>
<td>–</td>
<td>0-30</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EnMAP $\pm$</td>
<td>400-2500</td>
<td>0.65</td>
<td>0.22</td>
<td>1.78</td>
<td>0-8</td>
<td>PLSR</td>
<td>Italy</td>
<td>[107]</td>
</tr>
<tr>
<td>PRISM</td>
<td>400-2500</td>
<td>0.65</td>
<td>0.21</td>
<td>1.75</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>HypIRI</td>
<td>280-2510</td>
<td>0.60</td>
<td>0.22</td>
<td>1.65</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>EO-1 ALI</td>
<td>400-2500</td>
<td>0.49</td>
<td>0.25</td>
<td>1.43</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Hyperion</td>
<td>400-2500</td>
<td>0.62</td>
<td>0.22</td>
<td>1.70</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Landsat 8 OLI</td>
<td>430-2400</td>
<td>0.50</td>
<td>0.25</td>
<td>1.46</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Sentinel 2 MSI</td>
<td>45B-2500</td>
<td>0.56</td>
<td>0.23</td>
<td>1.55</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*Note: SFC-Seasonal Fraction Cover based on the four bands: b1-bare ground, b2-green vegetation fraction, b3-non-green vegetation fraction and b4-model fitting error derived from Landsat platform.*

*Castaldi *et al.* [103] used PONMAC spectral library for calculating SOC.

*$RPD$ stands for Ratio of Performance to Deviation.
success of measuring soil C [124], IR spectroscopy and statistical software [123]. After the laboratory
C due to improvements in the computer hardware
been predominantly used in measurement of soil
has been designed, fabricated, and tested as a
portable NIR spectrometer. The soil C estimation
has been recorded in the pastureland in Ohio [119].

Table 2. Pros and cons of the different methods of soil carbon measurement at field level.

<table>
<thead>
<tr>
<th>Methods</th>
<th>Pros</th>
<th>Cons</th>
<th>References</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Laboratory Methods</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Dry combustion method</td>
<td>• Gold standard with high precision and accuracy</td>
<td>• Costly and destructive method</td>
<td>[60]</td>
</tr>
<tr>
<td></td>
<td>• Quick and reliable</td>
<td>• Difficult to upscale at larger area</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Does not produce hazardous waste</td>
<td>• Expensive instrument and not</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>affordable by all soil labs</td>
<td></td>
</tr>
<tr>
<td>Wet oxidation method</td>
<td>• Low cost, minimum equipment requirement, simple</td>
<td>• Required high level of training</td>
<td>[68]</td>
</tr>
<tr>
<td></td>
<td>• Cheaper cost of apparatus and easy</td>
<td>• Old and traditional method,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>to assemble compared to dry combustion</td>
<td>required recovery factor due</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>to incomplete oxidation of soil</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>carbon</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Produce hazardous chemicals such</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>as dichromate, a carcinogen</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>• Not sufficiently sensitive to</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>identify small changes over time</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>in response to land use changes</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>and management practices.</td>
<td></td>
</tr>
<tr>
<td><strong>Remote sensing techniques</strong></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Spaceborne</td>
<td>• Non-destructive method, capture entire earth</td>
<td>• Need image correction (geometric,</td>
<td>[90]</td>
</tr>
<tr>
<td></td>
<td>and provide consistent viewing geometry</td>
<td>atmospheric, radiometric)</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Most of the time freely available (Landsat,</td>
<td>• Images are taken in specific dates,</td>
<td></td>
</tr>
<tr>
<td></td>
<td>Sentinel, etc.)</td>
<td>must coincide work with the</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Are easily available online and can be</td>
<td>passing day of satellite</td>
<td></td>
</tr>
<tr>
<td></td>
<td>downloaded any time for any years after they</td>
<td>• Difficult to process the image of</td>
<td></td>
</tr>
<tr>
<td></td>
<td>were launched</td>
<td>cloudy days</td>
<td></td>
</tr>
<tr>
<td>Airborne</td>
<td>• Cover larger areas of land in one flight</td>
<td>• More susceptible to geometry</td>
<td>[94]</td>
</tr>
<tr>
<td></td>
<td>mission</td>
<td>problems</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Easier to re-fly area on aircraft to study</td>
<td>• Radar system must use sophisticated</td>
<td></td>
</tr>
<tr>
<td></td>
<td>changes</td>
<td>navigation equipment</td>
<td></td>
</tr>
<tr>
<td>Unmanned aerial vehicles (UAVs)</td>
<td>• Non-destructive method, easily available and</td>
<td>• Expensive to conduct using aircraft</td>
<td>[97]</td>
</tr>
<tr>
<td></td>
<td>economical nowadays</td>
<td>• Poor performance under windy and</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Does not require high level of expertise to</td>
<td>cloudy conditions</td>
<td></td>
</tr>
<tr>
<td></td>
<td>fly UAVs</td>
<td>• Required high level of knowledge</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>during image processing</td>
<td></td>
</tr>
<tr>
<td>In-situ techniques</td>
<td>• Required small amount of sample</td>
<td>• Expensive and complex system;</td>
<td>[111]</td>
</tr>
<tr>
<td>Laser Induced Breakdown</td>
<td>therefore practically non-destructive</td>
<td>chances of eye damage</td>
<td></td>
</tr>
<tr>
<td>Spectroscopy</td>
<td>• Time and money saving potential</td>
<td>• Difficult to obtain suitable</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>standards</td>
<td></td>
</tr>
<tr>
<td>Inelastic neutron scattering</td>
<td>• Noninvasive method</td>
<td>• Calibration curves required for</td>
<td>[106]</td>
</tr>
<tr>
<td></td>
<td></td>
<td>each sample set</td>
<td></td>
</tr>
<tr>
<td>Infrared reflectance spectroscopy</td>
<td>• Most popular and non-destructive method</td>
<td>• Need calibration for each soil</td>
<td>[113]</td>
</tr>
<tr>
<td></td>
<td>• Possible to developed simplified version</td>
<td>type</td>
<td></td>
</tr>
<tr>
<td></td>
<td>• Easy to calculate due to</td>
<td>• Expensive and complex device</td>
<td></td>
</tr>
<tr>
<td></td>
<td>advancement in the hardware and</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>statistical software</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

However, large discrepancy was recorded in the pastureland in Ohio [119].

Infrared reflectance spectroscopy (IR spectroscopy) is a non-destructive and a widely used method for soil C measurement [120]. This reflectance spectroscopy is based on diffusely reflected radiation of illuminated soil [121]. The unique absorptive or reflective properties of SOC have a unique spectral signature at near infrared range (0.7-2.5µm) [122]. Near infrared (NIR) range has been predominantly used in measurement of soil C due to improvements in the computer hardware and statistical software [123]. After the laboratory success of measuring soil C [124], IR spectroscopy has been designed, fabricated, and tested as a portable NIR spectrometer. The soil C estimation has been affected by the physical properties of soil (moisture, texture, bulk density etc.) [125]. Some statistical techniques (i.e. partial least squares regression [126], random forest [127], support vector machines [128]) are used for calibration development and soil property estimation.

Based on the IR Spectroscopy two handheld devices are under field trial i.e. Our Sci Reflectometer [Our Sci, LLC; https://www.our-sci.net/; General Public License v3.0.] and Yard Stick [https://www.useyardstick.com/]. Our Sci Reflectometer has 10 light emitting diodes ranging from UV-B to near infrared wavelengths (365, 385, 450, 500, 530, 587, 632, 850, 880 and 940 nm.) These light sources are then pulsed allowing software to filter ambient light and reflection is measured using two pin photodiode detectors in the 300-700 and 700-1000 nm ranges. The area of
influence is $9.9 \times 22\text{mm} (1 \times 10^3 \mu\text{m})$ [45]. Our Sci Reflectometer uses air dried soil and measures reflectance in a 3.5-cm diameter petri dish. In addition to NIR reflectance values, soil physical properties (i.e., bulk density, soil texture and slope) are also considered in calibration using random forest regression method. Yard Stick is a second-hand-held probe which has a small camera on the tip and uses wavelengths to sense the presence of SOC. In addition to the small camera, it has a resistance sensor that measures bulk density that can be used to calculate the amount of C sequestered in a particular soil. It can be used in standing crops and instantly collects SOC and bulk density measurement and saves in the cloud. This probe is under field trial and developed in collaboration with Soil Health Institute, University of Nebraska-Lincoln, University of Sydney, and Yard Stick (https://www.useyardstick.com/solutions).

**Pros and cons of current methods**

Soil C has been considered as a key soil quality indicator since 1990s [129], but now it is also an important option to mitigate and adapt to global climate change [17, 130]. Sequestering C in soil by adopting recommended agricultural practices (RAPs) is a pertinent option to improve soil health and offset anthropogenic emissions [11]. Soil C measurement protocols and tools have been developed to measure soil health [131, 132]. With the increasing interest of farmers, private sector and policy makers towards credits and payment for C sequestration in soil; new easy, affordable, farmer’s friendly device to measure soil carbon is a high priority [38, 45]. The major challenge is to calculate soil C stock accurately at farm level under conservation agriculture (CA) and other RAPs to sequester C. Current methods that measure/estimate soil C for crediting farmers are expensive, highly sophisticated and needs technical expertise. However, these remote sensing techniques have been useful for large scale mapping of soil C, but it lacks the accuracy and needs technical expertise [58]. However, these remote sensing techniques have a large potential with the efficient machine learning and artificial intelligence for soil C estimation at a large scale [140, 141].

Soil C stock calculated for C credit must consider some important on-farm parameters that can improve the accuracy and address the variability. Soil C credits are based on the following specific cultural practices such as reduced tillage ($7.5 \text{ ha}^{-1}$), cover cropping ($15 \text{ ha}^{-1}$) and $22.5 \text{ ha}^{-1}$ for both (https://www.soyohio.org/council/carbon-markets/). This practice is not scientific, and farmers have to follow the guidelines of the private big agriculture companies that include up to ten years of contract. Soil C needs to be monitored by understanding different factors that affect soil C determination. Thus, the next section focuses on the key parameters that may be pertinent to soil C measurement. These parameters can be used to predict soil C by using machine learning and artificial intelligence.
Table 3. Literature showing relationship between soil color and soil carbon under different land use.

<table>
<thead>
<tr>
<th>Device used to measure color</th>
<th>Variable</th>
<th>Equation</th>
<th>Land use</th>
<th>$r^2$</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Konica Minolta chroma meter CR410</td>
<td>$y = \text{Soil lightness (L}^*\text{)}$</td>
<td>$y = -0.07x + 45.58$</td>
<td>Grassland</td>
<td>0.60</td>
<td>[156]</td>
</tr>
<tr>
<td></td>
<td>$x = \text{SOC (g kg}^{-1}\text{)}$</td>
<td>$y = 0.17x + 50.34$</td>
<td>Forest</td>
<td>0.52</td>
<td></td>
</tr>
<tr>
<td>Minolta CR-310 chroma meter</td>
<td>$y = % \text{ Reflectance}$</td>
<td>$y = -4.339 \ln(x) + 23.52$</td>
<td>Agriculture</td>
<td>0.74</td>
<td>[145]</td>
</tr>
<tr>
<td></td>
<td>$x = \text{SOC (g kg}^{-1}\text{)}$</td>
<td>$y = -0.7128 \ln(x) + 5.867$</td>
<td>Agriculture</td>
<td>0.74</td>
<td></td>
</tr>
<tr>
<td>SPAD-503</td>
<td>$y = \text{Total Carbon (g kg}^{-1}\text{)}$</td>
<td>$y = -2.32x + 159$</td>
<td>Agriculture</td>
<td>0.70</td>
<td>[147]</td>
</tr>
<tr>
<td>Minolta CR-310 chroma meter</td>
<td>$y = \text{Munsell value}$</td>
<td>$y = -1.35 \ln(x) + 6.55$</td>
<td>Hayden Prairie</td>
<td>0.74</td>
<td>[144]</td>
</tr>
<tr>
<td></td>
<td>$x = \text{SOC (g kg}^{-1}\text{)}$</td>
<td>$y = -0.9 \ln(x) + 4.78$</td>
<td>Agriculture</td>
<td>0.74</td>
<td></td>
</tr>
</tbody>
</table>

Key parameters in the context of soil carbon measurement

*In situ* measurement of soil C is important for farmers to get C credits and map soil C stock for larger areas. New device must take measurement of multiple variables which are pertinent to predicting accurate soil C content and relevant to mapping soil C at farm scale. The measurement of these variables can be used as covariates in the machine learning algorithm/Artificial Intelligence and get precise soil C values for an entire farm area. This process can be termed as integrated data-driven soil C mapping. This data driven soil C mapping can integrate laboratory, remote sensing, and other *in-situ* methods to precisely measure soil C that can be used for calculating soil C stock. The following discussion focuses on some important variables that affect soil C measurement at the field scale.

**Soil color**

Direct measurement of SOC is expensive and time consuming and is also complicated by the need of many samples to evaluate spatial heterogeneity [52, 142]. Soil color can be used as a cost-effective proxy to determine SOC content [143–145]. SOM is one of the primary pigmenting agents of soil that determines soil color [146–148] and is also an indicator of soil quality [149]. In the USA, there is a long history of relating soil color to SOM content even before the introduction of the Munsell color system by Brown and O’Neal [150] in Iowa and later Alexander [151] in Illinois using field color charts. Several studies have been conducted to estimate SOC content using the Field Munsell color chart [152,153]. Soil color is measured using Munsell color chart [154], Minolta CR-310 chroma meter (Minolta Crop., Ramsay, NJ) [144], Nix Pro Color Sensors and Konica Minolta CR-400 [155], soil color reader (SPAD-503) [146]. Research studies establishing the relationship between SOC and soil color are shown in Table 3. Wills et al. [143] reported that soil color was the best predictor for SOC by weight for agricultural fields ($r^2=0.79$) compared to prairie ($r^2=0.53$). The study conducted by Mueller and Pierce [156] in Central Michigan observed that the regression-derived maps closely resemble the spatial patterns of soil color and concluded that soil color maps can be used as a proxy of soil C maps. The lightness values ($L^*$) was negatively correlated with the total C content ($r^2=0.70$) in the Japanese agricultural soils [146]. These above studies showed that soil color can play an important role in estimating SOC content as a proxy variable.

**Soil moisture content**

Soil moisture content (SMC) plays a key role in transportation of nutrients and minerals and is a life-giving element for plant and soil microorganisms [157]. While SMC does not directly influence the SOC content but conversely soil C influences soil moisture reserves. Parajuli and Duffy [158] observed a weak relationship between SOC and SMC and high degree of variability within soil moisture data in two watersheds of south-eastern U.S. region of Mississippi. In addition, SMC exerted a stronger control on the decomposition rate under different temperatures in the tea-bags decomposition study conducted in Italy [159]. SMC has a strong effect on runoff, land surface energy dynamics, biomass yield, and root zone productivity. It also has far-reaching implications for irrigation, agricultural management, water quality and climate change [160,161]. Kerr and Ochsner [162] argued that soil-climate variables (soil moisture and temperature) play crucial roles in statistical models to predict soil C in the study conducted in temperate grassland sites across Oklahoma, USA. They reported that SMC is a single most influential predictor variable using the least absolute shrinkage and selection operator (LASSO) regression analysis and improved SOC estimates from $r^2 = 0.42$ to 0.52 at 5cm soil depth. Similarly, Fantappie et al. [163] showed that soil moisture regime is a
key predictor of SOC variations across Italy. Hursh et al. [164] reported strong correlation between SMC and SOC in a global study. Soil moisture plays a vital role in soil respiration affecting soil C stock and creating uncertainty in climate-carbon cycle [165].

SMC is measured using the ruggedized components that measure the resistance, capacitance or conductivity change of sensors. These commercial soil moisture sensors are expensive with the lowest cost of ~100 dollars, but economic sensors are not reliable and degrade quickly. Ding and Chandra [166] proposed a new technique called "strobe" for estimating SMC and electrical conductivity (EC) using Wi-Fi signals. Strobe estimates EC and soil moisture by measuring the relative time flight of Wi-Fi signals between multiple antennas and the ratios of the amplitude of the signals. Ding and Chandra [166] stated that strobe technique can be useful in soil sensing to small holder farmers using their own smartphone at economical value.

Soil strength

Soil strength is an important dynamic soil mechanical property which is influenced by land use and management practices [167]. Soil strength affects seedling emergence, soil erodibility, bearing capacity of soils and traction required to pull farm implements [168]. High soil strength reduces hydraulic conductivity and water infiltration rate, increases runoff and soil losses, affects root growth, and reduces crop production [105, 169]. Several management practices (i.e. tillage, cover cropping, crop residue retention) affect soil C stock as well as soil strength [170]. Soil strength increases because of compaction caused by a high mechanical load, less crop diversification, intensive grazing, and irrigation methods [171]. The effect of SOC on soil shear strength is moderated through changes in soil aggregate stability [168]. The distribution of SOC concentration and soil strength behavior are also affected by landscape position [172]. Blanco-Canqui et al. [167] reported an inverse relationship between soil strength and SOC content in forest soils compared to pasture and cultivated soils. Soil strength is measured as shear strength and cone index and its relationship with SOC content under different land uses are shown in Figure 1.

Figure 1. Relationship between soil strength parameter (Cone index and shear strength) and SOC for different management practices at Northern Appalachian Experimental Watershed in Ohio. (Adapted from [168]); DiskM, disk with beef cattle manure; NTm, no-till with beef cattle manure; NTnm, no-till with no beef cattle manure.

Soil bulk density

Soil bulk density, defined as mass per unit bulk volume, determines pore space available for air and water movement in soil. Bulk density is influenced by SOM content, soil texture, mineral composition, and porosity, which in turn are influenced by soil management practices [173]. Bulk density is strongly correlated with SOC content and depends on compaction, consolidation, and amount of SOC present [174–176]. Chaudhari et al. [173] observed a strong negative relationship \( r^2 = 0.79 \) between SOM and bulk density of soil samples. Similar opposite relationship was also observed by Curtis and Post [177] and Sakin [178]. Thus, bulk density is also useful to calculate the SOC stock.

The discussion presented herein shows the use of four soil physical parameters in determining SOC stock. These four variables include soil color, soil moisture, soil strength and bulk density, and they affect soil C content and soil functions. The data of these parameters can be used to estimate soil C using artificial intelligence/machine learning.
Data driven soil carbon estimation using artificial intelligence

A range of methods of C measurement (lab, in situ and remote sensing) are used globally. The choice of method is based on time constraints, equipment and reagent availability, software to analyze imagery, lack of resources (drone, aircraft), and expensive equipment (LIBS) [47, 49, 77, 99]. The accuracy of soil C measurement is higher with elemental analysis (dry combustion) [56] and slowly decreases while following other methods (in situ and remote sensing) [101]. Soil C measurement is affected by soil type and texture, and strength [167], tillage methods [179], crop grown [99], etc. Further, long-term changes in soil C stock depend on the rainfall, temperature, slope, etc. [180]. Highly efficient computing power is being used in predicting different soil properties using machine learning algorithms or artificial intelligence (AI) [181]. The use of AI needs a lot of data to improve the accuracy in predicting different soil properties such as soil moisture, EC, nitrogen (N) content etc. [182,183]. Data-driven agriculture has a bright future, and it can help in planning farming to meet the needs of growing population [184]. AI can be used in soil C measurement by integrating important soil physical characteristics. AI not only can predict soil properties at specific point but also create map for entire farm and region using remote sensing techniques [185]. This can be done at three phases; (i) field data collection using handheld device (ii) process the image data collected from drone, and satellite, and (iii) build model to predict soil C for entire farm. These three phases are described below.

Field data collection using handheld device

Soil sample collection from the field and analyzing at the lab is tedious and expensive for the entire farm. In this paper, four important soil physical properties are suggested that affect soil C viz., soil color, density, moisture, and strength. New handheld devices to measure soil C content must have sensors to measure these additional properties of soil and store them in the cloud (such as Microsoft Azure, Google cloud, IBM cloud services). The prototype of the handheld device is shown in Figure 2. This handheld device can measure bulk density, soil moisture, soil color and NIR reflectance. These data are then stored in the cloud for each GPS location (Geographical Position System).

Soil samples from some of these GPS locations are randomly selected to measure soil C using dry combustion method. Model can be developed using machine learning algorithms (random forest, boosted regression trees, support vector machine, artificial neural network). SOC content can be predicted using the handheld device using this model for the entire field. The predictor variables are soil moisture, bulk density, NIR reflectance, soil strength and soil color to predict soil C at each point (Figure 3). The data driven SOC measurement will address the effect of those soil physical properties and can be applied to diverse soil types.

Image processing collected from drone and satellite

Soil C mapping for the entire farm and region is challenging due to labor and time requirements. Remote sensing techniques have used drone and satellite imagery to map soil properties at larger scale [50]. However, the accuracy of the prediction of soil C has not been excellent (lower r^2 values and higher RMSE) (Table 1). This problem with accuracy can be addressed by identifying key soil properties that affect soil C content [186]. Once soil C content is predicted using a machine
learning algorithm from key soil physical properties (i.e., color, moisture, strength, and bulk density) and NIR spectroscopy. These properties can be used as key variables in predicting soil C content using the reflectance values or vegetative indices processed using drone or satellite image. Drone has the time and coverage limitation and covers a couple of fields at one flight [187]. However, satellites can cover a larger area, but it is essential to follow the time frame of the satellite revolving around the earth [188]. Drone can collect many images at the field and farm levels, and these can be used in mapping soil C at farm level. Accuracy at large regions can be improved by using drone imagery processing.

Drone and/or satellite imagery have been used not only to detect the type of vegetation but also tillage method and residue management which are key factors affecting the SOC content. These parameters are valuable in predicting SOC and can

Figure 3. Schematic diagram of soil carbon mapping using handheld device, drone and satellite imagery at field, farm and regional scale respectively.

Figure 4. Flow chart showing the soil carbon mapping from field to regional scale using artificial intelligence.
be used as predictor variables in machine learning/AI.

**Building model to predict soil carbon for entire region**

The goal of developing soil C maps at larger regions can be started at the farm field level. The schematic diagram (Figure 3) and flow chart (Figure 4) show how the randomly selected field data can be used to map soil C for the entire region considering key soil parameters and AI/Machine learning. The upsampling of the soil C measurements from a field using handheld devices goes through the series of machine learning/AI processes that use drone image to map soil C at farm level and satellite image at region scale. In remote sensing, researchers have either used different vegetative indices or specific wavelength reflectance to predict soil C content. Therefore, a model is proposed that can be effectively used to create soil C map and this can be updated every year to get more precise results. Farmers can select a few fields of their farm, take measurements using the proposed handheld device and link these data with the drone imagery. Different farms can be selected in a region and satellite images are used to create a soil C map for the entire region. This approach can help build a data store in the cloud for each time and AI can be used to improve accuracy and precision every year.

**Conclusion**

Soil C stock being the indicator of soil health and has monetary value due to C credits is global concern. The measurement of soil C at field level has faced challenges of time, money, accuracy following different methods from lab, in-situ to remote sensing. Laboratory methods provide higher accuracy but are expensive and give value at one point of field, whereas in-situ methods do provide soil C measurement for larger areas but are expensive and have lower accuracy. Remote sensing techniques showed a lot of promise but lacked accuracy. The pros and cons of these methods suggested four soil physical properties have a huge impact on soil C measurement. In this review article, the data driven soil C measurement is proposed that can be used at different scales (field, farm, and region). A prototype handheld device is suggested that can measure soil color, bulk density, moisture, strength and NIR spectroscopy, record the data and store in cloud (ex. Microsoft Azure). Soil C can be predicted using the machine learning/AI for that field using these above covariates. Next step is integrating data from drone imagery using ML/AI that will give soil C maps for the entire farm. For the larger scale soil C mapping, we suggested the use of satellite images (Landsat 8, Sentinel 2 A) adding it with the soil C predicted using drone imagery. This process integrates all the methods from field level (laboratory method) to farm level (in-situ) and finally regional level (remote sensing). It not only helps to create a soil C map but also facilitates calculating change in soil C stock over the time and location. This conceptual framework has a huge potential of soil C mapping and future study in applying this method in the field.

**Disclosure statement**

No potential conflict of interest was reported by the author.

**Funding**

This research was funded by the Microsoft (GF317072) and CFAES Rattan Lal Center for Carbon Management and Sequestration program support fund (C-MASC).

**Data availability statement**

The data used to support the findings of this study are available from the corresponding author upon request.

**References**


29. Crespi JM, Tidgren K. The first legal step for an agricultural carbon market is in the Growing Climate Solutions Act of 2021. Ames, IA: Center for Agricultural and Rural Development (CARD) at Iowa State University; 2021.


62. Matejevic I. Determination of carbon and nitrogen in samples of various soils by the dry combustion.


Ou L. Comparison of soil organic carbon methods and carbon sequestration and soil characteristics in Louisiana crawfish ponds (Master Thesis) Department of Agronomy and Soils, Auburn University, Auburn, AL. 2013.


Moritsuka N, Matsuoka K, Katsura K, et al. Soil color analysis for statistically estimating total carbon, total


