ESMC
Ecosystem Services Market Consortium

Working Group Project 1: Final Report

Phase 1a/b

Ecosystem Services Market Consortium/Ecosystem Services Market Research Consortium, January 2021

Table of Contents

EXECUTIVE SUMMARY **3**
OVERVIEW: TOTAL UNCERTAINTY FROM SOC STOCKS CHANGE AT THE FIELD SCALE **4**
CALCULATING STOCKS 4
SOURCES OF ERROR 4
CALCULATING STOCK CHANGES 5
1. SOC CONCENTRATION (%) **6**
1.1. REMOTE SENSING 6
1.2. GIS (SPATIAL) MODELING 10
1.3. PROXIMAL SENSING 10
1.4. DIRECT MEASUREMENT 13
2. BULK DENSITY **16**
2.1. PEDOTRANSFER FUNCTIONS 16
2.2. EXTENSIVE DIRECT SAMPLING 17
2.3. COMPARISON OF BD APPROACHES 17
3. DEPTH **18**
4. ROCK FRAGMENTS **18**
COST-UNCERTAINTY ANALYSIS 19
OVERALL UNCERTAINTY FOR STOCKS - RESULTS 22
DISCUSSION AND RECOMMENDATIONS **25**
APPENDIX A: INFLUENCE OF BULK DENSITY SOURCES IN THE FINAL ESTIMATION OF STOCKS WITH SATELLITE RS AT THE FIELD LEVEL: AN AD-HOC ANALYSIS **29**
DESCRIPTION 29
RESULTS 30
APPENDIX B: DETECTION PROBABILITIES AND COSTS FOR DIFFERENT SAMPLING INTENSITIES **32**

Funding: Research in this publication was supported by the Foundation for Food and Agriculture Research under award number – Grant ID: DFs-18-0000000012. The content of this publication is solely the responsibility of the authors and does not necessarily represent the official views of the Foundation for Food and Agriculture Research.

Attribution: This report was commissioned by ESMC/ESMRC and authored by Dr. Patrick Lawrence[1], Dr. Gisel Booman[2], Alexandra Thornton[1] & Dan Kane under contract to ESMC, January 2021.

[1]*Repliculture LLC*
[2]*Regen Network Development, INC*

Submission Contact: Patrick Lawrence; patrick@repliculture.com

Executive Summary

In the three interim reports preceding this final report, we reviewed in detail tools and technologies designed to improve in-field estimates of soil organic carbon (SOC) stock changes. From broad-scale approaches based on remote sensing to fine-scale sampling methods using penetrometers, all possible options were considered for improving accuracy and reducing costs. Each possible technology comes with its own unique trade-offs; there are no comprehensive solutions to quantification that eliminate the need to balance benefits and costs.

To help guide ESMC's approach to navigating this challenging landscape, this report provides a summary of our previous findings and an integrated analysis of all the trade-offs associated with estimating SOC stock changes. Beginning with a review of the inputs to the stock equation and sources of uncertainty, we describe the categories of tools available to estimate SOC concentrations, bulk density (BD), depth of sampling (d), and rock fragments (RF). Next, we perform a monte-carlo analysis of the impact of accumulated uncertainties on the probability of detecting stock changes and weigh those probabilities explicitly against costs. Finally, we synthesize the analysis with general observations on the most promising tools and technologies that may be available in the near future.

The process to cost-effectively quantify changes in soil carbon must be straightforward in order for carbon markets to become a reality and for land managers to receive fair compensation for climate-positive practices. Such a reality may, at times, seem distant and unachievable given the complexity of spatial variability in soils, the slow pace of carbon sequestration, and the difficulty of performing simple measurements of even small areas of land. However, as illustrated in this report and previous documents, incremental technological progress continues and several potential quantification breakthroughs may be available in the next several years. By pursuing a strategy of continuous engagement with emerging technology providers, testing of different tools and methods in pilot projects, and quantitative comparisons of different approaches, we are confident that ESMC will be well situated to take advantage of these developments in its nascent carbon marketplace.

Overview: Total Uncertainty from SOC Stocks Change at The Field Scale

Calculating Stocks

Uncertainty in SOC stock estimates can arise from four primary sources including the soil sampling depth, SOC concentration, bulk density, and the mass proportion of rock fragment content. These sources are related in the following equation,

$$stock = \frac{d \times C \times BD \times [1 - RF]}{100} \qquad [1]$$

where *stock* is the SOC stock (t C ha-1), d is the sampling depth (m), C is the soil organic carbon concentration (g C kg^{-1}), BD is the bulk density (kg m^{-3}) and RF is the mass proportion of rock fragment content (dimensionless). Alternatively, an equivalent soil mass (ESM) approach[1] may be used if significant changes in BD are expected between the two dates of sampling required for measuring stock changes. Soil mass layers ($M_{SOIL(DL)}$, Mg ha^{-1}) can be estimated via:

$$M_{SOIL(DL)} = \frac{mass}{area} = \frac{M_{SAMPLE(DL)}}{\pi\left(\frac{D}{2}\right)^2 \times N} \times 10000 \qquad [2]$$

where DL is the depth layer, $M_{sample(DL)}$ is the dry sample mass in g, and $\pi(D/2)^2$ is the cross-sectional area of the diameter inside the probe in mm. If multiple cores are combined, then N is equal to the number of soil cores that were sampled. From there, organic carbon mass can be calculated by:

$$M_{OC(DL)} = M_{SOIL(DL)} \times C_{OC(DL)} \qquad [3]$$

Where $M_{OC(DL)}$ is the OC mass in the depth layer (kg ha^{-1}) and $C_{OC(DL)}$ is the SOC concentration. These two soil mass equations can be used at any soil depth, with cumulative soil and OC masses calculated by summing the depth layer masses.

Sources of Error

With unlimited resources, one would ideally measure each of the stock equation inputs at every spatial location in the farm boundary with perfect analytical precision, completely eliminating uncertainty. In reality, practitioners are economically constrained to a small number of points for sampling, and those

[1] Wendt and Hauser, 2013

points will not perfectly represent stocks (even if the strata used to derive their locations are well designed). Furthermore, the mathematical method of combining the collected data will not be foolproof, and measurement techniques will never perfectly quantify the actual level of carbon in the soil. Therefore, uncertainty in the stock equation, and by extension each of the inputs, is inevitable. It can be characterized into three primary sources: sampling uncertainty (from imperfect stratification and sample densities), estimator uncertainty (generated by the population parameter calculation procedure), and analytical uncertainty (measurement and lab error).

In this report, we do not attempt to quantify the contribution of these three sources of uncertainty to overall SOC stock estimates despite their importance (this is the focus of the forthcoming ESMC RFP work). Instead, we assume that the accuracy metrics derived from academic and commercial sources are inclusive of errors from these three sources (even though they are not separately defined) especially where independent validation of the tools or technologies has occurred. Predictions in a new location will be influenced by all three sources of error and can be objectively compared against ground-truth verification samples.

Calculating Stock Changes

To translate SOC stocks at two periods in time into stock changes, a simple difference may be used:

$$\Delta\text{SOC stock(t/ha)} = \text{SOC stock}_{t1} - \text{SOC stock}_{t0} \qquad [4]$$

In some cases, the measured values of SOC stock_{t1} and SOC stock_{t0} may not be very accurate. This lack of accuracy is important if the intention of a survey is to understand overall C stored in a volume of soil. However, if the *precision* of both SOC stock_{t1} and SOC stock_{t0} is high and the goal of a survey is to quantify ΔSOC stock, then the accuracy will be immaterial.

For this report it was not possible to fully separate out accuracy versus precision. Measures of error referenced in academic literature and cited by commercial sources are most commonly reported in terms of root-mean-squared-error (RMSE) or similar metrics, which quantify deviation of predicted from actual values. Since actual values serve as the baseline for RMSE, it can be thought of as a basic measure of accuracy. However, in many cases, the statistical methods used to calibrate the tools and technologies are based on derivatives of linear regression, therefore errors would be expected to be normally distributed about the true values ($\sim N(0, \sigma)$). In this sense, RMSE can also serve as a measure of precision (clustering around one value). However, not all methods for estimating SOC exhibit normally distributed errors. For example, error may be higher for soils that are also higher in true carbon concentration. In these cases, RMSE may not serve as an adequate measure of precision and future research on such tools should better document error structure.

The primary directive for this evaluation was to quantify the impact of individual tool uncertainties on measured changes in carbon stocks, and to balance those uncertainties against cost. Given this target, it was determined that RMSE would serve as an adequate measure of precision for the inputs to the stock equation and for carbon stocks themselves.

1. SOC CONCENTRATION (%)

Bulk density, sampling depth, and rock fragments all impact estimation of SOC stocks, yet the SOC concentration itself is the key factor driving carbon accumulation or loss. As such, it has been the primary focal point for efforts to improve accuracy of SOC stock change estimates.

Characterizing SOC concentration at the field level has typically required a high density of soil sampling regardless of whether the intent is to generate whole-field summary statistics or to generate spatial maps of SOC. Below, we list some of the alternative approaches to quantifying the SOC concentration based on our previous reports and provide ranges of uncertainty and associated costs that could be expected in each case.

1.1. Remote Sensing

Remote sensing (RS) is an intuitively appealing approach to quantifying SOC due to its inherent ability to predict SOC measurements over large landscapes at a minimal cost. Details on the mechanisms that enable these predictions are described in detail in the previous remote sensing report for this project[1]. The general approach for RS methodologies is to regress remotely sensed imagery (typically acquired from satellites) against ground-collected measurements of SOC where pixels and measurements overlap, then to extrapolate to the whole image or multiple images using the regression coefficients or algorithms. Numerous satellites currently provide imagery of the entire globe, albeit with differing return intervals, spectral resolutions, spatial resolutions, signal qualities, and costs (from free to expensive). Myriad statistical and machine learning algorithms are available to associate the reflectance values of the pixels to the actual SOC measurements, some of which are simple to implement and others which require much larger volumes of data for algorithm training.

In the previous report on remote sensing[2], we reviewed the different approaches to estimating SOC with RS and surveyed the offerings of several of the vendors. Since then, we have had the opportunity to meet with additional vendors to understand their products and their timelines for commercial availability. A list of these vendors is provided in Table 1.

[2] WG Project 1: Remote Sensing Technology Assessment Phase 1a

Table 1. Current vendors for remote sensing SOC quantification tools including approach and commercial readiness.

Vendor	Overview
Vultus	Estimates SOC w/ Sentinel-2 data with a cloud-based processing pipeline based on 0-10cm soil depth measurements and a 10m resolution, with upwards of 4-year historical data. Data must be taken before planting, with bare ground conditions. Satellite data available every 2-3 days within Europe and every 5 days elsewhere. Field surveys are not necessary. Commercial readiness: 10/10
UpToEarth	Estimates SOC w/Sentinel-2 data, bulk density with synthetic aperture radar. Has analysis pipeline partially automated but needs ground-truth data. Predicts 0-30 cm. Commercial readiness: 5/10
Platfarm	Startup app using hyperspectral data and field surveys to measure bulk density. Still in the development stage, with limited accuracy for bulk density measurements. Commercial readiness: 3/10
CQuest	Still developing RS approaches using USDA, EU, ISRIC, and proprietary ground-truth data. No work on BD yet. Predicts 0-30 cm. Commercial readiness: 2/10
Cloud Agronomics	Well-funded startup using hyperspectral imagery to estimate SOC. Moderately extensive set of ground-truth data in US, but still refining RS approach. Expected cost = $6/ac. Predicts 0-30 cm. Commercial readiness: 3/10
SmartCloudFarming	Have done some testing of RS approaches (w/Sentinel-1, Sentinel-2, and LANDSAT 8) in several fields but need partners w/ground-truth data to fully develop the approach. Predicts 0-30 cm. Commercial readiness: 1/10

Uncertainty

As RS has received increased attention as a potential method to quantify SOC at a reduced cost, the number of academic research articles addressing this topic has rapidly expanded[1]. However, consistency between studies is minimal, inhibiting the ability to objectively compare RS data sources, post-processing techniques, and algorithms. Over time, it is anticipated that consistency will improve, but, as of 2020, confounding factors such as study design, scale, land management, and bioclimatic conditions make it difficult to draw general conclusions. These factors strongly limit the ability to generalize any one RS approach and to apply it in new locations.

Methodological complexity, sensor complexity, and high costs are not inherently associated with a higher degree of accuracy. For example, the RS review[3] found that hyperspectral sensors did not appear

[3] WG Project 1: Remote Sensing Technology Assessment Phase 1a

to perform noticeably better than multispectral sensors on the basis of either RMSE or normalized RMSE metrics, and there was no discernible difference between aerial versus satellite-based sensors (Figure 1).

In our analysis of the distribution of published RMSE values, categorized by remote sensing platform, we found that most of the studies using satellite remote sensing to quantify SOC concentrations had RMSE values below 0.5% SOC. Similarly, most of the studies using airplane-mounted sensors reported RMSE values below 0.8%. The RMSE values reported from unmanned aircraft sensors were even better, below 0.2%. These metrics were better than anticipated, rivaling those derived from proximal sensing tools.

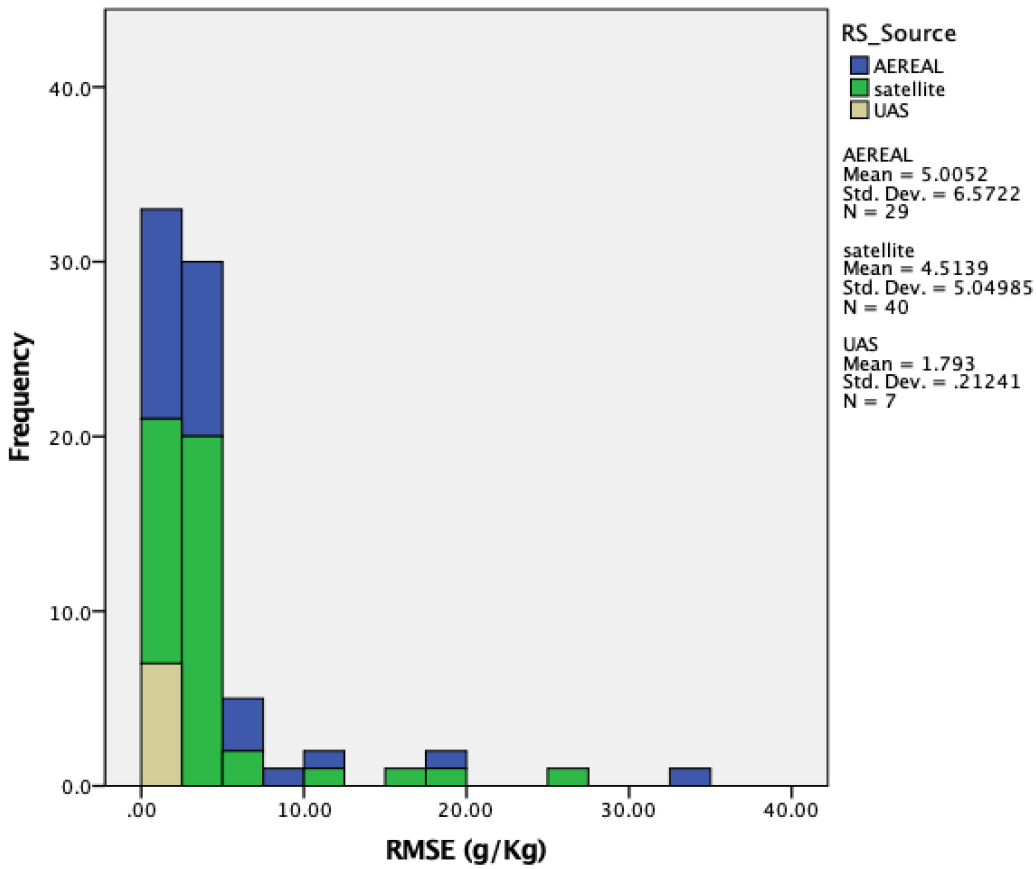

Figure 1. Frequencies of the RMSE values (in g/kg) from scientific publications where either satellite, aerial or unmanned aircrafts were used to quantify SOC.

Automation and standardization will enable more robust testing of RS tools across systems at a low cost. Achieving this will require the availability of standardized ground-truth data, but with efforts such as EMSC's pilot projects, these foundational systems will start to generate improved RS tools and knowledge of their true capabilities.

Costs

Each of the three RS sensor platforms have drastically different operating costs per unit area, with satellites as the least expensive, drones as moderately expensive, and aircraft as the most expensive.

Plane-mounted sensors have more recently fallen out of favor as the use and availability of drones has become more ubiquitous.

Conventional aircraft require pilots, fuel, airports, fuel trucks, and hangar fees, at a minimum, all of which quickly escalate costs. Even small LiDAR or spectral mapping jobs can cost upwards of $20,000 – $50,000 for a plane or helicopter operation. Given the logistics required[4], this price range may still be conservative.

Most companies that offer drone mapping services charge on an hourly basis. The number of hours billed will also include any post-processing needs, such as map preparation and data analysis. More than 30% of commercial drone pilots that offer drone mapping service charge an hourly rate of $150, although a range of $100 to $200 is generally acceptable. Prices also vary greatly depending on the industry being serviced: mapping services for oil and gas firms can cost up to $195 per hour, while agriculture and construction firms pay only $160 to $170 per hour. Some companies choose to offer drone mapping surveys on a per acre basis. However, they are in the minority and there are currently no established per acre rates for drone-based mapping.

The two advantages of using airborne sensors instead of satellite imagery would be to achieve a higher spatial resolution (submeter) and to reduce atmospheric distortions. However, satellite imagery often offers better spectral resolution (i.e., more spectral bands) plus frequent global coverage, and can be downloaded for free from several open data sources. One good example of this is the European Space Agency (ESA) Sentinel program, which offers super-spectral and radar imagery with 10 to 20m resolution and a revisit frequency of 5 days on average for the entire globe. Those characteristics are suitable for either the parcel or farm scale.

Moreover, the use of satellite imagery for estimating SOC concentrations at the field level could reduce sampling effort by a factor of 4 to a factor of 10, resulting in sampling costs declining from $1,440 to $140-360 for an average 100 ha field. By utilizing free satellite data sources, these costs would be further reduced. Therefore, the total cost for an assessment using remote sensing equals the sampling and analysis cost for the relatively small amount of samples, plus the GIS/RS expert that would execute the spatial analysis (until this process is automated). These costs are further broken down as follows:

1. The cost of sampling and laboratory analysis depends on the size of the assessed area. For reference, Regen Network estimates that for a farm of a size of 100 ha it would be necessary to extract around 5-12 samples for satellite calibration in each round of sampling. This would result in a total sample analysis cost between $35 and $330 depending on which laboratory is used. Note that the sample size required per unit area decreases as the project area increases, so a 1,000ha farm may require only around 20 samples, resulting in a total cost of between $140 and $550.
2. The cost of sample collection and analysis, and of RS could be further reduced by aggregating farms that use similar practices if they are located within a small distance of each other. For example, if 5 farms are aggregated, each farm would need to sample between one third and one fifth of the total number of samples for calibration of the satellite images, depending on the

[4] UAV LiDAR SERVICES

variability in the vegetation cover, soils, and other factors between farms. The cost to contract the GIS analyst could be also reduced as some parts of the analysis could be unified.

3. The costs of the GIS expert who would execute the analysis and reporting. We estimate the cost for the whole monitoring service and the corresponding provision of data and reports would be between 6,000 and $11,500 USD, although this cost could be spread over many farms (up to 40) for a cost of between $150 and $288 per farm.

1.2. GIS (Spatial) Modeling

Remote sensing imagery is but one of the many geospatial inputs that can be used to quantify SOC in agricultural systems. A multitude of other data layers are available that may be predictive of where SOC accumulates in a landscape. Based on first principles, these layers will often be associated with primary drivers of SOC accumulation, including soil clay content, soil taxonomy, NDVI, topography (and its derivatives), and soil pH.

In our report on modeling, we evaluated methodologies that did not overlap with approaches strictly based on imagery (covered in the RS section). A diversity of literature was reviewed, among which there was little overlap in methodologies, study locations, resolutions, or covariates. As a result, it was not possible to objectively compare accuracy metrics as in the RS report. Instead, we identified key classes of covariates that consistently explained SOC variability (in order of decreasing explanatory value):

- Topographic Covariates: Elevation, Slope , Plan Curvature (Plcu), Profile Curvature and Aspect
- LULC Covariates: spectral bands of satellites (e.g., Landsat), and NDVI.
- Hydrologic Covariates: Topographic Wetness Index (TWI) and Stream Power Index (SPI)
- Climatic Covariates: Temperature and Precipitation

These covariates were combined using many different statistical and machine learning algorithms, some of which performed well in some locations and not in others. Each unique study compared different sets of non-overlapping algorithms; therefore, comparison of results was of limited value. In terms of raw numeric frequency, artificial neural networks, boosted regression trees, multiple linear regression, ordinary kriging, random forest, and regression kriging were the algorithms that tended to perform better than others when compared in a single study.

The results and discussion provided in the modeling report[5] suggest possible layers and methodologies that may have value for predicting SOC in a variety of settings. Unfortunately, none of the methods outlined in the literature are known to be under development for commercialization, therefore speculation about possible costs and deeper interrogation of accuracy metrics is of limited value.

1.3. Proximal Sensing

Our evaluation of proximal sensing technologies identified ten potential tools for quantifying SOC, BD, or carbon stocks. For most of these tools, complete accuracy/uncertainty metrics were not available,

[5] WG Project 1: Modeling and Stratification Technology Assessment

either because the tool was proprietary or because it was still under development and accuracy data had not yet been collected. As a result, in our original report, accuracy ratings were necessarily subjective and only reported on a relative basis. For the total uncertainty analysis presented in the following section, we report accuracy metrics in terms of the root mean squared error (RMSE), but substantial caution is warranted since these metrics correspond to different sample sizes, locations, and sampling regimes. The RMSE values were derived by scaling the subjective ratings between a value of .052 g kg^{-1} for direct measurement, representing category 5, and 5.2 for Veris penetrometers.

Table 2. Available in-field proximal sensing tools and their associated operating details and estimated costs.

Tool	Underlying Tech	Field Method	BD	Analysis location	Date Commer-cially Ready	Accuracy Rating (1=low,5 =high)	Cost (E = estimated, A = actual)
Agro/soilcares	vis-NIR	In-field soil cores	No	in-field	Now	3	$3500 + $1900/yr (A)
Our-Sci reflectometer	vis-NIR	In-field soil cores	No	in-field	Now	2	~$400 (A)
Veris penetrometer	vis-NIR	penetration	No, possible	in-field	Now	2	$45,000 (A)
Texas A&M penetrometer	vis-NIR	penetration	No, possible	in-field	Unknown - maybe 2021	3	$63,000 (E)
X-Centric	PXRF	In-field soil cores	No	in-field	End of 2021	2	~$100-200 (E)
SoilReader	vis-NIR	On-the-go in-situ sensor	Yes	in-field	Now	3	$25,000/yr + $5k/yr subscription
LaserAg	LIBS	In-field soil cores	No	lab	2021	4	$400,000 (E)
INS	INS	On-the-go sensor	Yes	in-field	End of 2021	3-4	<$75,000 (E) or $10/ha
SCANS	vis-NIR, gamma-ray	In-field soil cores	Yes	in-field	Unknown	3-4	$100,000 - $130,000 (E)
Agricarbon	DC, BD	In-field soil cores	Yes	in-field	2021	5	1 GBP/sample (E)

Uncertainty

For individual locations, the lowest uncertainty/highest accuracy for any proximal sensing technology is likely to be achieved by Agricarbon, as it uses dry combustion and BD measurement methods that are typically used within off-site laboratories. However, this conjecture is not yet supported by data and

therefore requires validation. The SCANS system[6] provides the second-lowest uncertainty for estimating SOC stocks at the point level by combining vis-NIR sensors calibrated with the EPO algorithm[7] (for SOC) and gamma-radiometric sensors for estimating BD.

INS is also a system that can provide a relatively high level of accuracy; however, it is not entirely clear how changes in bulk density would impact the volume of soil analyzed by the system, and its uncertainty remains around its ability to interrogate soil to a consistent depth. Nevertheless, INS is capable of generating a spatially explicit map of SOC stocks using an on-the-go sensor, which would provide superior information about stocks over a large area even if it is slightly less accurate at individual locations.

Costs

The costs associated with each tool are described in Table 2.

1.4. Direct Measurement

Direct measurement is typically regarded as the "gold standard" against which other SOC quantification methodologies are compared. Using one of several laboratory techniques, samples are collected in-field, transported to a laboratory, pre-processed via drying, grinding, and sieving, and then analyzed for SOC concentration. These processing and analysis steps can introduce errors in themselves, in addition to the errors that are introduced during the process of sampling, sample site selection, and statistical averaging. At the point scale, these errors tend to be relatively small compared to errors associated with use of other quantification tools, but when extrapolating to fields or landscapes this relative advantage may dissipate. Nevertheless, direct measurement techniques, and especially dry combustion, are likely to serve as the "ground-truth" standard for other tools and technologies at small scales.

Uncertainty

The results from our report[8] indicate that direct measurement of SOC stocks and SOC stock change can provide a range of accuracy, depending on which SOC extraction method is used. Out of the three most common methods (Walkey-Black (WB), Loss-on-ignition (LOI), and elemental analyzers), combustion via en elemental analyzers provide the most accurate SOC content analysis. The North American Proficiency Testing Program (NAPT) determined that combustion has the most consistent measurement analysis. While WB and LOI are more cost effective, they have higher margins of error and suffer from error propagation. Furthermore, due to its high probability for error and safety hazards from handling the chromium waste and acidic reagents, most labs in the United States, but not all, have started to move away from the WB method and instead use LOI or an elemental analyzer.

According to NAPT, the LOI method can have upwards of +/-20% error, resulting from the need of individual conversion factors, assumptions based on water content and consistency of carbon content

---

[6] Viscarra Rossel *et al.*, (2017)

[7] Minasny *et al.*, (2011)

[8] WG Project 1: Direct Measurement and Proximal Sensing Technology Assessment Phase 1a

within SOC/SOM, and differing combustion temperature and duration of OC removal. Additionally, conversion factors, temperature, and time of combustion are all dependent on soil type, area samples, soil horizons, and types of organic content present within the samples. The WB method has even higher error margins, with some studies[9,10] estimating as high as 59% - 86% variability within results. This variability is attributed to the oxidation of only active OC within the samples, manual titration errors, and site-specific correction factors needed to estimate SOC.

Costs

Unless lab-grade equipment is readily available, the most cost-effective way to analyze SOC content is to send soil samples to an independent lab for analysis. Table X outlines labs within the United States, their analysis method(s), and cost per soil sample. It should be noted that some labs only analyze for total C, regardless if inorganic carbonates are present. For these labs, we recommend contacting the lab directly to determine if further analysis is possible to remove the inorganic carbonates from the soil sample.

Table 3. Soil sampling labs and costs for SOC/SOM analysis.

Lab	Analysis Method for SOC	Analysis Cost Per Sample
University of Massachusetts at Amherst Soil Testing Lab	LOI	$6 for OM
Ward Laboratories	Elemental analyzer; LOI	$16 for NPK, OM, CEC, and S via an elemental analyzer $5.50 for a single OM analysis via LOI method
Atlantic Microlabs	Elemental analyzer	$26 for SOC analysis
Cornell University Lab	LOI	$20 for total C & N*
Pennsylvania State College Agricultural Analytical Services Lab	LOI; dry combustion	$5 for organic matter via LOI $15 for total C via combustion*
Oregon State University Central Analytical Lab	Elemental analyzer; LOI	$14 for TOC via an elemental analyzer for 1 - 16 samples $7 for SOC analysis via LOI for 1 - 16 samples
Oklahoma State Soil Testing Lab	Elemental analyzer	$8 for OM $8 for %OC and total N

[9] Bisutti *et al.*, (2004)
[10] Nayak *et al.*, (2019)

University of Idaho, Analytical Sciences	WB; elemental analyzer	$27 for OM via WB for 1-5 samples $23 for total C & N via elemental analyzer for 1-5 samples if samples *do not* need to have inorganic carbonates removed $27 for total C and N via elemental analyzer for 1-5 samples if samples need to have inorganic carbonates removed
Kansas State Soil Testing Lab	LOI	$2.50 for OM via LOI
Colorado State University	WB; LOI	$35 for soil analysis package that includes OM via WB if OM content is below 8% $35 for soil analysis package that includes OM via LOI if OM content is above 8%
Utah State University Analytical Laboratories	WB, LOI	$80 for soil analysis package that includes OM via WB if OM content is below 6% $80 for soil analysis package that includes OM via LOI if OM content is above 6%
University of Kentucky	Elemental analyzer	$5 for OM
Louisiana State University Soil Testing Lab	WB	$4 for OM

Notes: Adapted from Soil Sampling Resource Guide by Regen Network *Analysis for total carbon in sample, including inorganic carbonates. Does not specify if further analysis can be done to remove inorganic carbonates.

2. Bulk Density

Although secondary to SOC concentration, BD can have an outsized impact on SOC stock and SOC stock change calculations. As described in detail in the third report in this series, it is possible for BD to undergo significant changes over time, and these changes must be accounted for when performing successive measurements of SOC stocks. Direct measurement methods for quantifying BD are typically performed in conjunction with SOC measurements.

The most widely used approaches to estimate bulk density at the field-scale are:

- Pedotransfer functions
- Extensive direct sampling for calculation of field-scale spatial means via averaging or interpolation

Pedotransfer functions enable spatial granularity and may be usable even with limited sampling, but they are less accurate than extensive sampling.

Several less well-researched methods exist for quantifying BD that may become available in the next several years. These include BD quantification by penetrometer insertion force calibration, the use of a gamma-ray densitometer in conjunction with a vis-NIR sensor as part of the SCANS system[5], and inelastic neutron scattering (INS), which uses neutron attenuation at depth to estimate BD. Currently only the SCANS system has published accuracy metrics associated with BD estimation, which indicate that it can quantify BD with a high degree of accuracy[11] (RMSE = 0.055 g cm^{-3}). INS may be able to achieve this level of accuracy but only measures SOC by mass down to 30 cm rather than separately quantifying BD. The costs of these technologies are documented earlier in this report; only pedotransfer functions and extensive sampling are described in this section.

2.1. Pedotransfer Functions

Uncertainty

Pedotransfer functions (PTFs) have been developed over many years and are useful for estimating specific soil properties when data are only available for correlated soil properties. While we are unaware of literature that extensively tests the ability of PTFs to estimate bulk density, Regen Network has completed an unpublished analysis of their performance within the context of rangelands in New South Wales (NSW), Australia[12].

For the NSW analysis, total uncertainty ranged between 24 and 25% for the years 2017, 2018 and 2019. These metrics were derived from Sentinel-2 imagery calibrated to a minimum soil sample size, in conjunction with local pedotransfer functions to get bulk density maps. In this case, bulk density was the highest source of uncertainty, contributing to around **21%** to the total uncertainty.

[11] Lobsey *et al.*, (2017)

[12] Regen Network- Unpublished report and data

Costs

The cost of using pedotransfer functions will mainly depend on two factors:

1. The costs associated with sampling and laboratory analysis, which are determined by the number of samples required to validate the relationship between estimated and observed bulk density values. The cost of obtaining bulk density values depends on whether the assessment is carried out by a laboratory V0Cor by using a "DIY" bulk density method (which only incurs the time cost of sampling). Laboratory analysis for bulk density varies in cost as it is sometimes included in the traditional soil analysis package, or it can be added on for an additional cost of around $5.
2. The costs of the expert modeler that would assess the bulk density within the field boundaries based on the pedotransfer functions.

Once PTFs have been developed for a local region, it may be possible to omit site-specific sampling altogether. In addition, the process of applying PTFs to a specific farm or field boundary could likely be automated, further reducing costs.

2.2. Extensive Direct Sampling

Direct sampling includes the use of a soil coring implement to extract a specific volume of soil for bulk density calculations. These core extractions will typically occur alongside SOC sampling. Depending on the relative variability of SOC and BD, one or the other may be prioritized for determining core extraction locations and densities.

Uncertainty

The uncertainty associated with direct sampling for BD will be primarily determined by the degree to which the sample locations are representative of the area of interest. This uncertainty is unique to each location and therefore it is not possible to generalize its likely magnitude. However, if estimates of within-field variability are available from nearby locations or soil maps, these estimates may be used to roughly estimate required sampling densities to achieve an acceptable level of uncertainty.

Costs

The costs associated with direct sampling for BD will be dictated by specific labs and are expected to be approximately $5 USD per sample. Lower-cost methods for estimating BD may be available using DIY methods as previously mentioned.

2.3. Comparison of BD Approaches

In order to test the impact of bulk density measurement approaches on estimates of SOC stocks at the farm level, Regen Network carried out a simple *ad-hoc* analysis using data from a rangeland managed with prescribed grazing in NSW, Australia (see more details in Appendix A).

The mean differences between a) using PTFs for creating maps of bulk density, b) using values from samples, extrapolated to the whole strata, and c) using a homogeneous value for the whole project area, based on a regional database, only ranged between 2% and 10%. There was no statistical

difference between using a PTF or extrapolating sampled values. Slight overestimation was observed when comparing the measured to the PTF-estimated values, but there was no significant difference between extrapolated and measured values.

These preliminary results suggest that expensive and extensive sampling campaigns for gathering bulk density data may not be justified, at least in the context of the management regime and ecorregion endemic to NSW, Australia.

3. Depth

Information on uncertainty related to depth of sampling or rock fragments is extremely limited. In our review, only Goidts *et al*. (2009)[13] have explored the uncertainty associated with measuring these variables. For non-stony soils, depth was estimated to have an RMSE of 2.5 cm at the landscape scale and 6.5 cm at the field scale and was the primary (51%) contributor to uncertainty in SOC stocks for grasslands[12]. For non-stony croplands, depth contributed less than 30% to uncertainty in SOC stocks in croplands. For stony soils at the field scale, sampling depth contributed a much lower amount to overall stock uncertainty (approximately 10-15%).

It is important to note that the uncertainty generated by inconsistent depth measurements reported by Goidts *et al*. (2009)[12] was calculated with the use of augers rather than truck-mounted hydraulic soil probes. We hypothesize that a mechanically driven soil probe would reduce, but not eliminate depth-of-sampling related errors. Dry, clayey soils, rock fragments, wet soils, and other conditions will continue to inhibit sampling to consistent depths regardless of the sampling mechanism.

4. Rock Fragments

At the plot or field scale, rock fragments can significantly impact BD estimates and SOC calculations. Although there is some dispute about how to account for rock fragments and the degree to which they affect BD and SOC measurements, completely ignoring their impact is likely to bias SOC stock estimates where RF are present.

According to Goidts *et al.* (2009)[12], in stony soils RF have an outsized influence on calculated SOC stocks in grassland (35% of variability) and in cropland (23% of variability). The influence of RF in non-stony soils was not quantified. Associated RMSE values for stony soils were not provided, but standard deviations and CVs for RF were approximately 10% and 50%, respectively.

Information on the cost associated with quantifying RF has not been described in any of the primary literature and has not been provided by any commercial laboratories, likely due to the cost being difficult to separate from the general process for BD analysis. Future research that addresses this topic would be valuable for understanding the cost of reducing uncertainty associated with RF, at least in stony soils.

[13] Goidts et al., 2009

Cost-uncertainty Analysis

The utility of any given tool or technology depends on its overall impact on estimated carbon stock change uncertainty and the expense required to achieve that specific level of accuracy. Simply knowing the ability of a tool to estimate one input for stock calculations will not provide insight into the other required inputs and the overall impact of a set of tools on uncertainty. For example, remote sensing technologies often appear to perform well for estimating surface-level SOC, but if this high accuracy is offset by low accuracy in estimated depths or estimated BD (which are often sometimes estimated using pedotransfer functions), then overall stock uncertainties may still be high.

To understand the impact of these joint uncertainties on carbon stock estimation errors and the subsequent trade-offs of uncertainty against cost, we performed a monte-carlo simulation of estimated carbon stock changes for many of the tools and technologies described in this report (excluding those without concrete accuracy metrics). Our fixed scenario for this analysis assumed a field size of 200 hectares, a SOC change from 1% to 1.5%, a fixed bulk density across time, and a fixed proportion of rock fragments, although each of these parameters was assumed to vary spatially within the 200 ha field. We did not analyze the impact of the spatial structure of these properties on uncertainty.

In addition to these baseline field conditions, we also assumed the following parameters to be true for the simulation (Table 4):

Table 4. Parameters used in monte-carlo simulation of soil carbon stocks.

Parameter	Value	Units
SOC - true within-field std. dev	3.27	%; based on average from Lawrence *et al.*, 2020[14]
Depth of sampling (d)	0.3	m
d - std. dev. direct sampling	0.05	m
Rock Fragments (RF)	0.01	dimensionless
RF - true within-field std. dev.	0.001	dimensionless
RF - std. dev. direct sampling	.0005	dimensionless
Bulk Density (BD)	1300	kg m^{-3}
BD - true within-field std. dev.	130	kg m^{-3}

[14] Lawrence *et al.*, 2020

BD - std. dev. Direct sampling	130	kg m^{-3}
BD - std. dev. pedotransfer function	325	kg m^{-3}
Number of samples (n)	No less than 10	Samples per field
Fixed cost of setting up sampling in one field	$100	USD
Technician cost for 1 day of sampling	$400	USD
Technician cost for ½ day of sampling	$200	USD
Cost per sample for SOC (direct measurement)	$15	USD
Cost/sample for BD (direct measurement)	$4.50	USD
Cost/sample for RF (direct measurement)	$2	USD
GIS/RS analysis, per farm	$200	USD

The end goal of the simulation was to quantify the probability of detecting the pre-determined stock change within specified limits (+/- .2% SOC or 40% on a relative basis) for each possible tool or technology. To accomplish this, we first simulated (Figure 2) the range of true values for each stock equation input (SOC, BD, RF; depth is not a field property and is only influenced by sampling) by drawing 1000 samples from a normal distribution with mean and standard deviation derived from Table 4. We refer to these distributions as the *true* distributions. These values represented the actual variability of the inputs across the field.

Next, we approximated the uncertainty associated with measuring each of the input variables. Random draws were sampled from a normal distribution, with the mean centered on a single sample from the *true* distribution and the standard deviation derived from RMSE values presented in the literature. The number of random draws corresponded to the number of samples that were assumed to be collected from each field; a range of these sample numbers was tested throughout the analysis.

If RMSE values were not provided in literature or by specific tool vendors, then those tools were omitted from the analysis. For most tools and technologies, SOC was the target variable, and RMSE values were therefore only provided for SOC itself. In such situations, we either assumed that BD was derived from a pedotransfer function with a standard deviation of 25% from the mean[15], or it was derived via direct measurements with the typical standard deviation presented in Table 4. Alternatively, if a tool or technology estimated total carbon stocks and provided the associated RMSE values, then stock values were simply drawn from a random normal distribution.

[15] APPENDIX A (ancillary data)

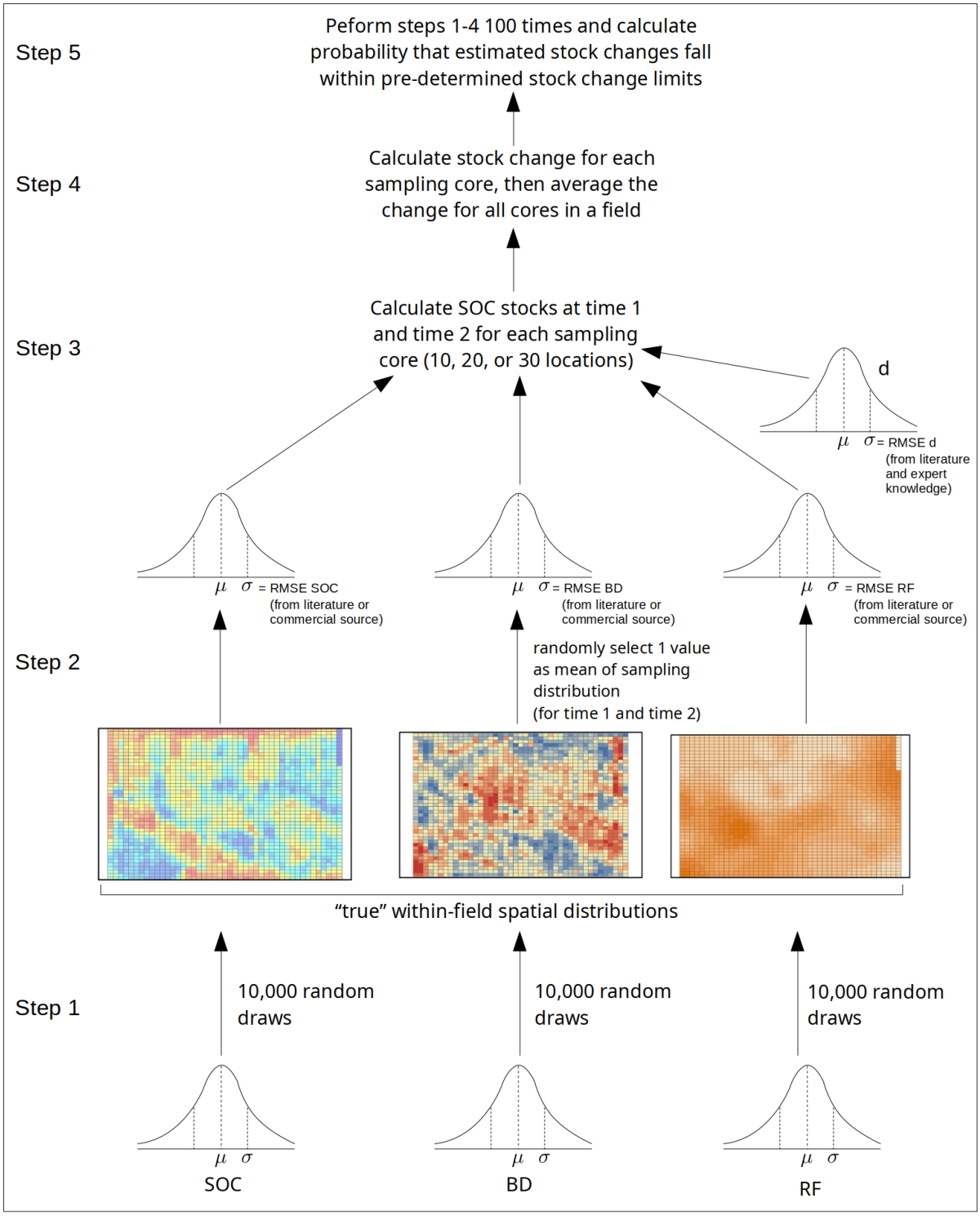

Figure 2. Monte-carlo analysis procedure for estimating probabilities of detecting stock changes using various tools and methodologies.

Regardless of the derivation method, carbon stocks were estimated at two points in time then subtracted at each sampling point to determine the stock change associated with each sampling location. All stock changes associated with the scenario (e.g., 20 changes for 20 sampling points) were averaged to get the stock change associated with the entire 200 ha field area. This process was repeated 100 times for each scenario, then these 100 average stock change values were used to calculate the probability of estimating stock changes that were within the threshold of stock changes associated with a change in SOC from 0.3% to 0.7%.

Finally, costs were calculated using the fixed and variable cost metrics outlined in Table 4. For remote sensing and modeling scenarios, we assumed that stock change analysis would be focused on a compact geographic area with similar bioclimatic and edaphic conditions. Therefore, only three samples were assumed to be required for the 200-ha field, since those three samples would be combined with samples from other fields in order to calibrate the remote sensing approach. These three samples were assumed to be acquired via direct measurement and therefore used the same variable costs as direct measurement techniques.

For proximal sensing technologies, variable costs were proportional to the number of samples collected in each field (10-30) and varied by technology. For technologies using hand-held vis-NIR in-field spectrometers, costs per sample were relatively low ($1.5) due to the overall low cost of equipment and negligible costs of analysis. For in-field direct measurement (Agricarbon), costs (1 GBP) were assumed to be roughly 2X ($3) the cost expected by the manufacturer since the technology is not yet available. Variable costs associated with penetrometers (Veris, Texas A&M penetrometer) or in-field core analysis (SCANS) were assumed to be proportional to the analysis time. For INS, individual core samples are not collected. As a result, we used the expected costs as described by the vendor ($10/ha). Variable costs for all other technologies were assumed to follow vendor expectations.

To calculate total costs for any given technology, the final equation used was as follows:

$$Total\ Cost\ =\ Cost_{SS}\ +\ Cost_{sampler}\ +\ n * VC_{SOC} + n * VC_{BD} + N * VC_{RF} * 0.1 \qquad [5]$$

Where $Cost_{SS}$ is the cost of setting up the sampling scheme ($100) including determining sample locations, gathering equipment, and transportation costs, $Cost_{sampler}$ is the cost of the technician for one day of sampling, *n* is the number of samples collected, VC_{SOC}, VC_{BD}, and VC_{RF} are the variable costs of collecting and analyzing SOC, BD, and rock fragments. The cost associated with measuring RF was reduced by 90% under the assumption that only 10% of sampling locations would have a sufficient proportion of RF to warrant measurement.

Overall Uncertainty for Stocks - Results

Probabilities of Detecting Change

As expected, increased sample sizes improved the probability of detecting the stock change associated with a 0.5% increase in SOC, but simultaneously increased costs (Figure 3; Appendix B). Direct

measurement methods were most likely to detect the change within the +/-40% error threshold; at lower sampling intensities they easily outperformed most other methods. Proximal sensing tools achieved similar detection probabilities to remote sensing methods on average, although both had a relatively wide spread of outcomes.

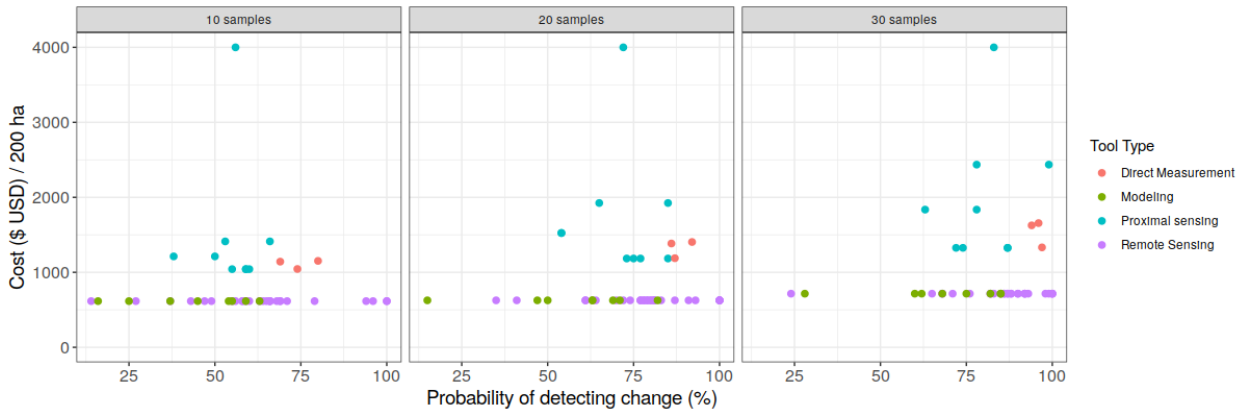

Figure 3. Probability of detecting a change in SOC stocks associated with a 0.5% change in SOC to within +/- 0.2% SOC (or within 40% of the actual change) when (a) 10, (b) 20, and (c) 30 samples have been collected at both the outset and completion of a soil carbon monitoring project.

For this analysis, detection probabilities within 10% of each other can likely be regarded as analogous in terms of accuracy, given the number of assumptions that were required. For example, all of the direct measurement methods are virtually indistinguishable in terms of detection probabilities, although this was expected given that two of the methods rely on the same measurement techniques (albeit performed in-lab vs. in the field).

Remote sensing/modeling methods and tools had a very large spread of probabilities of detection, ranging from worse than chance (50%) to 100%. As noted in the previous RS report, confidence in the broader applicability of reported accuracy metrics is low due to the lack of independent validation and the unique characteristics of each distinct study location. As such, the remote sensing probabilities should be interpreted as a whole rather than in isolation.

Proximal sensing tools covered a smaller, yet still broad spread of detection probabilities, with the SCANS system[16] performing the best and the Texas A&M penetrometer achieving the lowest probabilities.

If detecting stock changes is required at a lower error threshold of +/- 0.1% SOC or +/- 20% on an absolute basis, detection probabilities are drastically reduced, even when 30 samples are collected during each sampling event (Figure 4). The same patterns of relative probabilities between measurement methods are observed, but rarely exceed 75%.

[16] Viscarra Rossel *et al.*, 2017

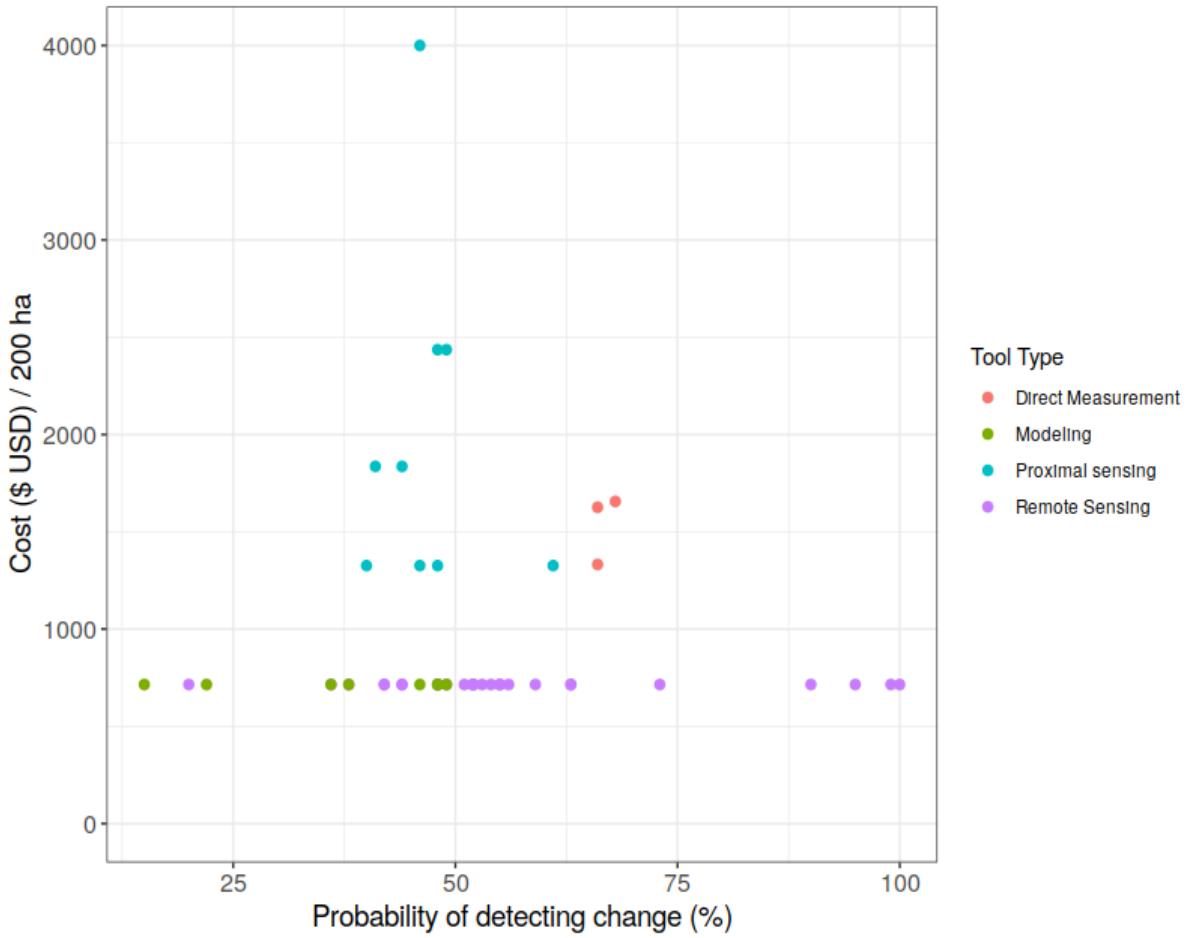

Figure 4. Probability of detecting a change in SOC stocks associated with a 0.5% change in SOC to within +/- 0.1% SOC (or within 20% of the actual change).

Costs

Measurement costs were tightly spread when a low number of samples were collected but tended to diverge as sample sizes increased. INS was by far the costliest measurement technology, requiring $4,000 to measure the entire field. INS "scans" the soil rather than collecting physical soil samples, therefore its price is only determined by field size (its anticipated pricing is $10/ha).

Remote sensing was consistently the lowest-cost technology due to its need for a smaller number of samples per unit area than other technologies, especially if many fields in a region can be analyzed simultaneously. There was no way to distinguish the number of local calibration samples that would be required for one approach over another, therefore a singular cost structure was adopted for all of the methods. The same holds true for modeling-based methods. However, it should be noted that our assumptions about sampling requirements for both types of approach were based on minimal data, therefore the actual location of the cost intercepts may be substantially different than what was used for this analysis.

In the Regen example, the estimation of SOC by calibrating Sentinel-2 satellite imagery to soil samples produced accuracies between 3.19% and 4.12%, reaching a total uncertainty of around 24-25% in stocks (see bulk density section). This is consistent with the mid- to upper range of our detection probabilities at sampling densities of 20 and 30 cores per 200 ha, and with reported direct SOC stock errors as noted in the RS report[1], which varied between 1.6 and 33.2 %.

Discussion and Recommendations

In the series of reports generated for this project we explored in detail the costs and accuracies associated with a broad range of SOC quantification tools and methodologies. Less consideration was given to commercial readiness and robustness of testing that are equally important for tool viability. Here all these factors are considered together, with acknowledgement that the rapid pace of technological development may render our recommendations outdated in short order.

RS and spatial modeling technologies clearly possess the potential to achieve high accuracies for estimating SOC stock changes at much reduced costs over many farms. However, these higher accuracies are strongly dependent on improving PTF accuracy and increasing the number and breadth of calibration datasets to ensure that predictions in new locations are valid. These improvements must occur simultaneous to reducing the number of local ground-truth calibration samples required in any given location. Furthermore, any optimism in RS/spatial modeling technologies relies on accuracy metrics that are primarily derived from academic literature and have not been independently verified by vendors in a large number of locations. Therefore, the largest hurdle to implementation of RS and spatial modeling tools is the availability of sufficient ground-truth data for calibration and testing. In our previous recommendations, we suggested that ESMC partner with RS/spatial modeling tool vendors to simultaneously assist in model training and testing. We still believe that this approach has merit; it would enable ESMC to evaluate the tools at a minimal cost while also enhancing their capabilities.

Proximal sensing tools currently appear to convey few advantages over direct measurement except for slightly lower costs. Most tools require sampling campaigns and extraction of soil cores. Since the largest costs associated with direct measurement such as planning and executing a field campaign (choosing sampling locations, transportation, cost of a sampling technician, and measuring BD/RF) are also required for proximal sensing, minimal justification exists for choosing proximal sensing technologies over direct measurement. For example, if 20 samples are gathered from a field at two points in time, the total planning, fieldwork, and auxiliary measurement (BD and RF) costs amount to $1,264 whereas the laboratory analysis costs for SOC only total $140. Reducing the $140 even by half would only offer a minor cost advantage. In addition, many of the proximal sensing technologies either require significant up-front investments in equipment or are in the prototype phase of development and have not undergone rigorous testing. This reality does not suggest that proximal sensing tools should be entirely disregarded; it indicates that their opportunities lie in offering incremental improvements rather than drastic reductions in costs.

In consideration of these realities, we recommend that ESMC adopt a cautious approach towards all technologies which promotes collaborative development and testing of technologies but does not commit to endorsing any specific tools at this moment. As such, we continue to endorse a pragmatic approach which engages with different vendors in terms of their development timelines, level of

commercial availability, and cost structure. These tiers of engagement are divided into four categories (Table 5).

- Category one represents tools that are currently available and have well developed processes for easily estimating SOC, BD, or SOC stocks. Only two tools are in this category (the Agrocares scanner and SoilReader), and they are only able to estimate SOC. We recommend these tools for immediate testing within ESMC pilots, while recognizing that the Agrocares SOC soil spectral library (SSL) has incomplete coverage in the US, and SoilReader has not provided well-established accuracy metrics. The Agrocares tool is relatively inexpensive, so it may be feasible for ESMC to purchase for testing rather than entering into a collaborative development/testing agreement. The SoilReader tool is more costly ($25,000 plus a minimum $5,000/yr subscription), therefore we recommend engaging in a partnership with them to reduce or eliminate the capital costs of the equipment.
- Category two represents tools that are currently available but may require a higher level of testing prior to broader adoption. The Our-Sci reflectometer falls into this category even though it is similar to the Agrocares sensor because its back-end database and system for automated quantification of SOC is less robust. The rest of the tools in this category are remote-sensing tools for estimating SOC. Most of these tools are in the prototype stage and have tested their methods on a small subset of ground-truth samples. The tool developers would likely welcome a partnership with ESMC to test and refine their tools at no cost.
- Category three represents tools that have been prototyped (Agricarbon) or have been tested (INS) but are not commercially available. Due to the high potential for these tools for quantifying spatially explicit SOC stocks (both SOC and BD) at potentially low costs (the cost structure of INS services remains to be fully determined), we recommend that ESMC keep in close contact with these vendors and schedule future testing as soon as possible.
- Category four represents a diverse set of tools that warrant monitoring on a periodic basis. The X-Centric PXRF sensor seems promising but has only been tested on carefully curated samples (Ravansari, 2021) and may have a longer path towards being manufactured and commercially available. LaserAg has the potential to reduce lab-based SOC analysis costs significantly over the long-term but is not commercially available and will require significant financial resources for soil testing labs to procure. The Veris penetrometer is available right now for use, but it is not paired with a robust SSL and therefore does not currently offer significant advantages over other lab-based direct sampling. However, if the penetrometer can be calibrated to effectively measure BD, its utility could significantly increase. Finally, the Texas A&M penetrometer has achieved good results for estimating SOC and is under active development, but its availability for commercial use appears to be some time off in the future.

Table 5. Currently available commercial tools and quality of prior testing. Proximal sensing, remote sensing, and direct measurement tools highlighted in red, green, and orange, respectively.

	Suggested Testing Timeline									
	2020	**2021**				**2022**				
	Q4	**Q1**	**Q2**	**Q3**	**Q4**	**Q1**	**Q2**	**Q3**	**Q4**	**Robustness of prior testing**
Category 1: Testing										
Agro/soilcares										4
SoilReader										2
Category 2: Collaborative refinement										
Our-sci reflectometer										3
UpToEarth										1
CQuest										1
SmartCloudFarming										1
Vultus										2
Cloud Agronomics										2
Category 3: Close monitoring; future testing										
Agricarbon										3*
INS (CarbonAssetSolutions)										3
Category 4: Periodic Monitoring										
LaserAg										3

X-Centric										2
Texas A&M penetrometer										3
Veris Penetrometer										3

Despite recent enthusiasm for carbon sequestration and significant investment in SOC quantification capabilities, none of the currently available tools are silver bullets for cost-effectively quantifying SOC stock changes with high accuracy. Remote sensing tools may provide a solution if a rigorous, coordinated campaign of calibration and independent testing is implemented. Proximal sensing tools may incrementally reduce field sampling measurement costs and improve efficiency of sample analysis. Direct measurement costs may be marginally reduced with more efficient lab measurement, or significantly reduced if traditional measurement techniques can be implemented in the field. Combinations of all these tools (e.g., remote sensing and proximal measurement) may jointly reduce costs to a sufficient degree that they are viable when used as a complete package, although accuracy metrics are not sufficiently robust to support specific recommendations at this time.

Each of these potential advancements is based on a unique set of assumptions about accuracy, cost, and commercial availability that may or may not be consistent with the current or future reality. However, as a whole, it is likely that one or more of them may meet or exceed expectations, providing a plausible pathway for cost-effective SOC stock change quantification. By adopting a measured approach to evaluating and supporting the development of these tools, we believe that ESMC will be well positioned to take advantage of this shifting landscape as its carbon market initiative matures.

Appendix A: Influence of Bulk Density Sources in the Final Estimation of Stocks with Satellite RS at the Field Level: An Ad-Hoc Analysis

Description

In order to test the influence of the data source for bulk density when accounting for SOC stocks at the field level, Regen Network's science team carried out a simple *ad-hoc* analysis for the data from a 1,400ha rangeland located in New South Wales, Australia, under prescribed grazing. This rangeland has been previously stratified according to management history and environmental variables and is covered mostly by native grassland species. During 2019, nine (9) strata were sampled, the sampling points were geolocated, and the SOC content (%) along with the bulk density were determined for each composite sample by a certified laboratory.

Based on these data, Regen Network carried out a calibration of satellite spectral imagery (Sentinel-2) in order to map out the SOC % over the whole project area, and also account for the SOC stocks at the project area level. More details on that methodology can be found at Regen Network´s Methodology for GHG and Co-Benefits in Grazing Systems.

In order to convert the estimated SOC% values to SOC stocks, the bulk density values are needed. For the purpose of comparing the final stocks at the field level when using different data sources of bulk density with different levels of accuracy, we gathered three different data sources of bulk density values at the pixel level:

1. Pedotransfer function (PTF): We created a raster map based on the SOC% map from the satellite calibration, by applying a local pedotransfer function proposed by Merry[17], which was validated by plotting the observed (from samples) against the estimated bulk density values and confirming a significant (p-value < 0.05) correlation.

The PTF equation follows:

$$BD = 1.608 - 0.0872 \times \text{Percent SOC} \qquad \text{Eq. 1}$$

2. Linear extrapolation of the sampled values: We simply extrapolated the BD values from the samples to the whole strata.

[17] In: Spouncer, L.R., Skjemstad, J.O., Merry, R.H. (2000) Soil Carbon Information for Major Soils in IBRA regions - South Australia. CSIRO Land and Water Consultancy Report. Pp 22

3. Empirical value: we used a unique bulk density value of 1.29 g/cm3 for the whole project area, based on a mean bulk density value for that area from a global map of Soil Bulk Density at a 0-5cm depth[18]

Results

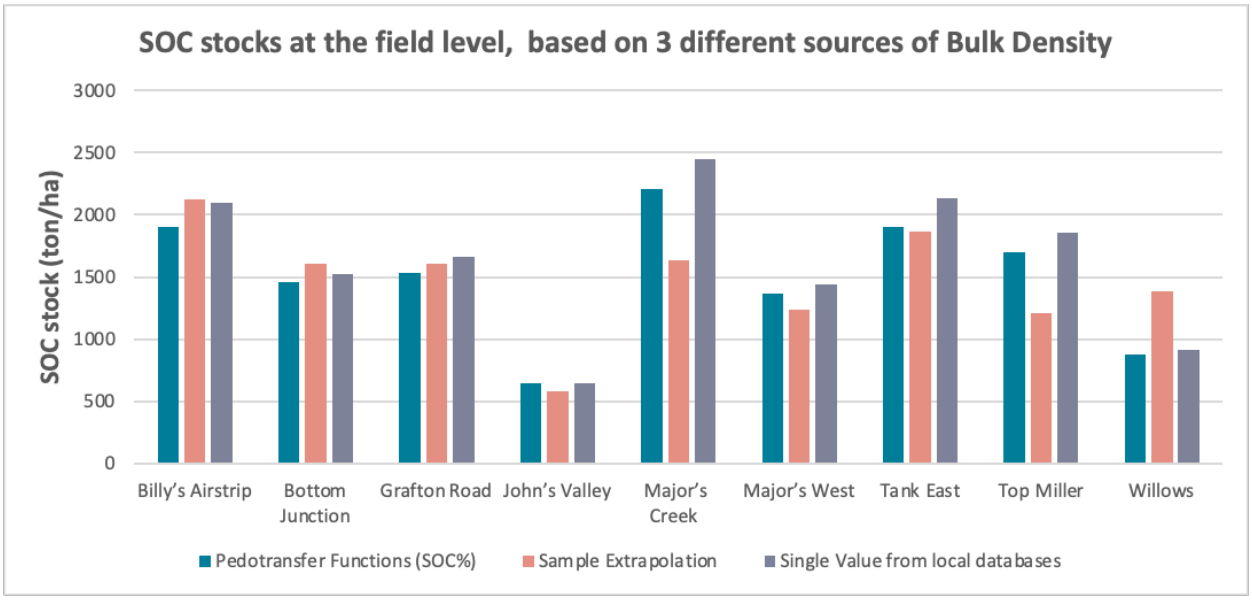

Figure A1. Statistical comparison of results at the field level, from applying three different sources of Bulk Density data at 8 strata in a farm in New South Wales, Australia.

Statistical Analysis Results: Paired Sample T-Test for Comparing Means

The mean differences between (a) using PTFs for creating maps of bulk density, (b) using values from samples, extrapolated to the whole strata, and (c) using a homogeneous value for the whole project area, based on a regional database, ranged between 2% and 10%. There was no statistical difference between using a PTF or extrapolating sample values. There was some overestimation from using the empirical value over the PTF. And there was no significant difference between using the extrapolated samples and using a measured value.

Although this is a very small experiment and only carried out over one property with a limited sample size, it highlights the need to do the whole analysis at the field level in order to analyze the impact of each source of uncertainty and thus evaluate if the final influence on the overall uncertainty counterbalances the costs in time and money to carry out exhaustive and complex bulk density

[18] Soil Bulk density for the World (global map) - Data Basin

samplings and analysis. If the final errors rise from 2% to 8 or 10%, like the results from this analysis show, it might be acceptable to just get regional empirical values of bulk density for estimating soc stocks at the farm level.

Appendix B: Detection Probabilities and Costs for Different Sampling Intensities

Table 1B. Detection probabilities and associated costs for different sampling Iitensities.

		10 samples		20 Samples		30 Samples	
Tool Type	Author / Tool	Detection Probability (%)	Cost ($)	Detection Probability (%)	Cost ($)	Detection Probability (%)	Cost ($)
Remote Sensing	Aldana-Jague *et al.*, 2016	65	616	83	626	90	715
Remote Sensing	Ben-Dor *et al*, 2002	64	616	78	626	93	715
Remote Sensing	Castaldi *et al.*, 2019	63	616	91	626	88	715
Remote Sensing	DeTar *et al.*, 2008	79	616	93	626	98	715
Remote Sensing	Devine *et ai.*, 2020	100	616	100	626	100	715
Remote Sensing	Forkuor *et al.*, 2017	69	616	81	626	92	715
Remote Sensing	Gomez *et al.*, 2008	27	616	41	626	65	715
Remote Sensing	Guo *et al.*, 2019	68	616	79	626	87	715
Remote Sensing	Huang *et al.*, 2007	71	616	80	626	86	715
Remote Sensing	Lu *et al.*, 2013	66	616	80	626	90	715
Remote Sensing	Mirzaee *et al.*, 2016	58	616	81	626	88	715
Remote Sensing	Munoz and Kravchenko, 2011	69	616	78	626	92	715
Remote Sensing	Page *et al.*, 2013	59	616	77	626	83	715
Remote Sensing	Patzold *et al.*, 2008	58	616	79	626	82	715
Remote Sensing	Samuel-Rosa *et al.*, 2015	14	616	35	626	24	715
Remote Sensing	Selige *et al.*, 2006	60	616	72	626	87	715
Remote Sensing	Simbahan *et al.*, 2006	94	616	100	626	99	715
Remote Sensing	Steinberg *et al.*, 2006	55	616	74	626	92	715
Remote Sensing	Stevens *et al.*, 2010	56	616	77	626	76	715
Remote Sensing	Stevens *et al.*, 2012	49	616	61	626	68	715
Remote Sensing	Thaler *et al*, 2019	37	616	61	626	71	715
Remote Sensing	Uno *et al.*, 2005	43	616	64	626	68	715
Remote Sensing	Vaudour *et al.*, 2019	66	616	87	626	85	715
Remote Sensing	Wang *et al.*, 2018	96	616	100	626	100	715
Remote Sensing	Wang *et al.*, 2012	100	616	100	626	100	715
Remote Sensing	Zhang *et al.*, 2019	47	616	70	626	82	715
Remote Sensing	Zizala *et al.*, 2019	66	616	81	626	86	715
Modeling	Van Apeldoorn et al., 2014	63	616	82	626	82	715
Modeling	Castro-Franco et al., 2015	59	616	71	626	85	715
Modeling	de Menezes et al., 2016	16	616	15	626	28	715
Modeling	Mosleh et al., 2016	25	616	47	626	62	715
Modeling	Ratnayake et al., 2016	55	616	63	626	85	715
Modeling	Zeng et al., 2016	54	616	69	626	75	715

Modeling	Laub *et al*., 2018	45	616	63	626	68	715
Modeling	Wang et al., 2018	37	616	50	626	60	715
Proximal sensing	Agrocares	60	1042	85	1184	87	1326
Proximal sensing	Our-Sci	59	1042	75	1184	87	1326
Proximal sensing	INS	56	4000	72	4000	83	4000
Proximal sensing	Christy, 2008	50	1212	54	1524	78	1836
Proximal sensing	Lobsey et al., 2017	66	1412	85	1924	99	2436
Proximal sensing	Morona et al., 2017	55	1042	73	1184	74	1326
Proximal sensing	Tang et al., 2015	59	1042	77	1184	72	1326
Proximal sensing	Viscarra Rossel et al., 2017	53	1412	65	1924	78	2436
Proximal sensing	Wijewardane et al., 2020	38	1212	54	1524	63	1836
Direct Measurement	Direct Measurement	80	1152	92	1404	96	1656
Direct Measurement	LaserAg	69	1142	86	1384	94	1626
Direct Measurement	Agricarbon	74	1044	87	1188	97	1332