YARD STICK Soil Carbon Revealed

Stratification Methods for OC Stock Quantification

Updated 04/06/2023

Overview

The purpose of this document is to describe Yard Stick's stratification and sample plan design techniques as of April 2023. Broadly speaking, we employ a stratified random sampling approach which minimizes the number of samples required to achieve known goals for confidence and margin of error (MOE).

Stratification is a widely-discussed and -debated topic within soil carbon market stakeholders. If employed correctly, it offers great potential to improve soil stock measurement, reporting, and verification (MRV) quality while reducing cost. We believe our stratification methodology represents a compelling solution to the question of how to stratify US soils for the purpose of quantifying soil carbon stocks at scale.[1]

Note some of these methods are protected by one or more patents.

What is our goal?

The goal of Yard Stick's stratification methodology is to reduce the cost and complexity of soil carbon stock quantification while preserving precision and overall scientific rigor. Soil sampling is expensive, therefore maximizing the stock quantification value of each individual sample is paramount. Stratification is one way to maximize this value.[2]

---

[1] Throughout this document we will use the phrase "soil carbon stocks" to generally refer to the mass of soil organic carbon (SOC) in a given field to a given depth (most often 30 cm in commodity croplands in the US). While there is an extensive and worthy conversation to be had regarding which fractions of soil carbon are most valuable to quantify, for simplicity this document presumes our customers are most interested in OC stocks and changes for the purposes of soil CDR markets and similar.

[2] Yard Stick is of course also working to reduce the cost of direct quantification of soil C stocks via our in situ spectral probe. This document exclusively concerns our stratification techniques which are a key complement to our overall stock quantification offering.

What is stratification?

Stratification is the process of dividing one or many fields[3] into smaller areas which can be expected to share certain characteristics. These areas are called strata (singular: stratum). As Yard Stick is focused on quantifying soil carbon stocks, our stratification techniques focus on creating strata of maximally homogenous expected soil carbon stocks. To eliminate bias, sample sites are placed randomly within each stratum (though note further details on sample site placement techniques below).

By dividing fields into more-homogeneous strata we are able to reduce the overall variation of target areas, resulting in a more precise estimate of the soil carbon stock for a given budget. Stratified random sampling nearly always results in lower variation within the estimated property (in this case soil carbon) than a simple random sample for the same number of samples taken.[4]

Our stratification methodology requires three inputs:

1. Field boundaries in appropriate GIS formats for the area to be stratified and sampled.
2. Target confidence level as a %. This describes with what confidence we expect the true mean of a stock to be within the calculated range we report after sampling. For example, if target confidence level = 95%, we will report that there is a 95% likelihood that the true mean of a stock is within the range *calculated mean ± MOE in tons*. Consequently, there is a 5% chance the true mean is **not** within our calculated range due to pure randomness (even though our sampling plan was ideally suited to capture the true mean).
3. One of either Target margin of error (MOE) or Total Number of Samples
    a. Option 1 is Target MOE as a % of the mean. This describes the estimated amount of random sampling error in the results of our sampling effort relative to the mean stock, and is used to determine the number of samples required. For example, if MOE = 10%, we estimate the number of samples required so that the width of our confidence interval is equal to *2 * 10% * calculated mean*.
    b. Option 2 is Total Number of Samples. This option is typically used when a specific budget target exists and precision is therefore limited by the total number samples possible for a project. We distribute this number of samples across strata to minimize resulting MOE.

Note that our stratification methodology produces sample plans with **target** confidence and MOE values. The actual confidence and MOE values of a stock may vary from a plan's target due to differences between expected and actual soil carbon and bulk density variability (see section "A priori estimate vulnerability" below for a more detailed explanation).

---

[3] In this document for simplicity we will refer to land areas as "fields," reflecting Yard Stick's focus on agricultural lands. Note however that these techniques are applicable to many other types of land management units including grazing paddocks, whole farms or ranches, multiple farms or ranches, grasslands, and beyond.

[4] Webster and Lark 2012

What data are fields stratified by?

Fields are stratified by expected soil carbon stock values.

We derive these stock values from the POLARIS[5] dataset which was created using existing data about soils (e.g. SSURGO) as well as other relevant environmental data. POLARIS contains predictions about an array of soil properties for different depth increments throughout the United States and is widely used within the soil science community. It is a comprehensive, robust set of data well-suited to our stratification goals.

We use predictions of organic matter and bulk density from POLARIS to derive an estimate of soil organic carbon stocks in each part of a field.[6] Our methodology provides these estimates at a pixel size of approximately 5 m square.[7]

Once these pixel-scale estimations are available, we use a statistical technique called *k*-means clustering to group, or "cluster," pixels of similar expected soil carbon stock. This *k*-means approach minimizes within-group variation, placing locations of similar properties into the same clusters.[8] These clusters become our strata. In this way, each stratum we create is more similar to itself, in terms of mean soil carbon stock and variation, than to other strata. That means the area of each stratum can be randomly sampled and each sample site within a given stratum can be expected to be as predictive of the stratum mean as any other site. From a statistical perspective, the soils within each stratum are considered "the same" for the purposes of quantifying soil carbon stocks.

How are the number of strata determined?

Using *k*-means (described above), our methodology creates a variety of scenarios which contain varying total number of strata: Some with few strata, and some with very many.[9] Next,

---

[5] Chaney et al. 2019

[6] POLARIS does not include estimations of organic carbon content, however organic matter and organic carbon have a very close relationship, the latter comprising ~60% of the former in most agricultural soils in the US.

[7] While our priority POLARIS data are only available at 30 m pixel resolution, we know soil is highly heterogeneous within a single pixel of that size. Hence to better simulate real-world conditions, we split each 30 m pixel into 36 5 m pixels and assign values to each 5 m pixel by randomly sampling from the 30 m pixel distribution. We derive a single 30 m pixel's distribution from the P5 and P95 values provided by POLARIS. By assigning a single 5 m pixel a single value, we are assuming it is homogenous. We chose 36 samples since it satisfies the Central Limit Therorem's "rule of thumb" of minimum number of samples required for the CLT to hold.

[8] Brus 2019

[9] Strata per field range from 1 to 11. If a field (or group of fields, if fields are combined for stratification) is less than 20 acres we only permit one stratum. Fields larger than 20 acres are only allowed to have max strata of (total_acres/10) or 11, whichever is larger.

for each scenario, we calculate the total number of samples required to achieve target precision and any other protocol requirements (e.g. CAR SEP requires a minimum of 3 samples per stratum). We select the scenario which minimizes the total number of samples required, therefore also minimizing cost. This scenario determines the total number of strata.

How is sampling density determined?

A key consideration in all soil carbon sampling efforts is the tradeoff between sampling density (and therefore cost) and stock precision. A project with too-low density will result in high stock uncertainty values leading to large deductions in tons and therefore weak carbon "yield." This in turn will negatively impact the economic attractiveness of many soil carbon projects. A project with too-high density, however, will "overshoot" confidence requirements, resulting in wasted budget on unnecessarily costly sampling.

Yard Stick's customers know that high-uncertainty stocks are of limited value to those wanting to be meaningfully rewarded for storing carbon in soils. As a consequence, Yard Stick is generally aligned to confidence and margin of error requirements specified by major market methodologies such as CAR Soil Enrichment Protocol, Verra VM0042, and other similar methodologies. We aim to "right size" sampling density: high enough to be confident (= low uncertainty and therefore acceptable to major market methodologies), but not so high that projects overspend on sampling.

Most leading methodologies prescribe minimum stock uncertainty criteria - we work backward from these values to determine density. Yard Stick's stratification methodology requires target confidence and margin of error (MOE) inputs, from which required sample density is calculated. Using POLARIS, we estimate local carbon variability in a given field, and use these estimates to establish required sampling density on a per-field or per-project basis using Neyman allocation.[10] By tailoring the sampling density to the expected variability of each stratum, we create a sampling plan which minimizes the number of samples required to meet a desired level of precision.

Imagine two fields which are identical in terms of shape, size, management practices, topography, and mean soil carbon stocks. The only difference is that Field A has higher carbon variability than Field B. As a consequence, Field A will require a larger number of samples to reach the same target confidence and MOE as Field B. This is the math which our stratification methodology uses to calculate required sample density.

Currently, all sample sites include testing for both carbon and BD. In the future, we will likely allow the creation of sample plans with differing densities of carbon and BD (e.g. 50% of sample sites are C only, 50% are C + BD).

---

[10] The result of this approach is that sampling density may vary substantially across even a single field. Variability is not evenly distributed; areas with high variability will have higher density and vice versa.

How are sample sites placed within strata/fields?

To eliminate bias, all sample sites are placed randomly within each stratum excluding a 25 foot setback edge of each field. This setback is used for two reasons:

1. Field boundaries may often depart from expected GPS values by reasonable amounts. A setback ensures sampling plans are confidently "in" each field despite inevitable boundary imprecision.
2. The edges of fields commonly receive somewhat different treatment from the interior (e.g. more traffic from equipment resulting in very high bulk density). A setback helps us avoid sampling head rows and similarly less-representative sections of a field.

Relatedly, no two sample sites will be placed within 5 m of each other, as in-field GPS precision can vary by device, location, and conditions.

How does Yard Stick use other types of data?

Management

Management is an important determinant of the level of soil carbon in soil. For example, if a field has been intensively tilled for many more years than a neighboring location, we would expect the soil carbon content to be diminished, as intensive tillage generally causes a loss of soil carbon over time.[11]

However, Yard Stick does not explicitly use detailed management data in our stratification, as its quality is often highly variable and its effect upon soil carbon can be strong or weak depending on a significant number of hard-to-evaluate variables.

Instead, we group fields according to their shared current and historical management; if two fields have a different management history, no part of them will be grouped together in the stratification. Similarly, if there are several fields that share identical management, then our stratification methodology allows them to be grouped, and stratified together. Grouping multiple fields with similar management often results in much lower sampling density. However, as the definition of "identical" is complex and hotly debated, we generally advise customers to stratify at a per-field level.

Environmental properties

POLARIS was created by combining a wide variety of high-quality soil data, including a number of environmental properties such as topography and wetness which are understood to predict soil carbon with some success.[12] [13] As a result, POLARIS can be considered to "include" these

---

[11] Sanford and Lin 2012

[12] McBratney, Mendonça Santos, and Minasny 2003

[13] Minasny et al. 2013

data in its estimation of OM and BD. Therefore to additionally include these data as direct covariates (e.g. via SSURGO, USGS National Elevation Dataset) would be “double counting” and likely result in overweighting of these data. Hence, we only stratify by stock and variability estimates derived from SSURGO and omit direct inclusion of other covariates such as slope, aspect, and wetness/TWI.

What are the limitations of these techniques?

Geography

As these techniques rely heavily on POLARIS data, which is only available for the United States, we cannot stratify soils outside the US “out of the box.” However similar datasets often exist in other countries, and alternative covariates are often available. We are actively exploring how our techniques would be adapted to international soils.

A priori estimate vulnerability

There is no way to precisely estimate expected soil carbon stock and variability values *a priori*. Soil carbon practitioners have two options:

1. Use the best data available for stratification. That is the approach we currently take by stratifying on soil carbon stock and variability estimates derived from POLARIS.
2. Conduct a pre-sampling campaign. This is often prohibitively expensive and operationally complex. And a pre-sampling campaign of course has its own stratification and sample plan design challenges!

Yard Stick believes strongly that option #1 is a superior approach in most cases, however we acknowledge we may not be “right” about expected variability and, therefore, required sample density. Once a sample plan is created and sampling is complete, Yard Stick can provide an “expected vs. actual” analysis which evaluates the estimated variability as derived from SSURGO to the real variability established by the real-world sampling campaign. From this we can determine if a given project accomplished its intended confidence/MOE, and, if not, what additional sampling may be required to further reduce stock uncertainty.

Considering the known imperfections of POLARIS data, we have implemented a number of compensating techniques to increase required sampling density and therefore provide a conservative sample plan more likely to hit a required confidence level:

- When the POLARIS-estimated variation is unreasonably low, we adjust it upwards to avoid inappropriately low sampling density.[14]
- If a pre-sampling campaign has been completed, we will adjust the estimated variance calculated with POLARIS by a factor computed by the variance observed in the field only if the field variance is larger than POLARIS estimates.

---

[14] We scale the standard deviation to mean ratio by 1/0.3 when variation is low. 0.3 is the lower of two R^2 between bulk density and organic matter in the POLARIS dataset.

- We sample from the 30 m pixel distribution for each 5 m by 5 m pixel to increase within-pixel variability. The 30 m pixel distribution is derived from the P5 and P95 values provided by POLARIS. A large range between the P5 and P95 value in the POLARIS dataset will create a wide sampling distribution and therefore a larger computed variance among the 5 m pixels.
- We round up when partial samples are required.

These techniques maximize the utility of POLARIS and result in more conservative sample plans most likely to hit target precision values. This further reinforces that option #1 above is likely to result in better outcomes.

“Grouping” fields by identical management

As described above, fields which are truly “identical” may be stratified together. This can often reduce estimated required sampling densities considerably - an attractive proposition to projects always working to maximize the value of each dollar!

While the post-sampling analysis described above will reveal whether these lower-density sampling efforts accomplished their stock precision goals or not, we acknowledge there is unanswered science regarding the stratification “unit” of a project and how that should inform field grouping. We broadly recommend customers choose the most conservative option which fits their budget, and being too precise is often less costly than being not precise enough.

References

Brus, D. J. 2019. “Sampling for Digital Soil Mapping: A Tutorial Supported by r Scripts.” *Geoderma* 338: 464–80. https://doi.org/10.1016/j.geoderma.2018.07.036.

Chaney, N. W., Minasny, B., Herman, J. D., Nauman, T. W., Brungard, C. W., Morgan, C. L. S., McBratney, A. B., Wood, E. F., & Yimam, Y. (2019). POLARIS soil properties: 30‐m probabilistic maps of soil properties over the contiguous United States. *Water Resources Research*, *55*(4), 2916–2938. https://doi.org/10.1029/2018wr022797

McBratney, A. B, M. L Mendonça Santos, and B Minasny. 2003. “On Digital Soil Mapping.” *Geoderma* 117 (1): 3–52. https://doi.org/10.1016/S0016-7061(03)00223-4.

Minasny, Budiman, Alex B. McBratney, Brendan P. Malone, and Ichsani Wheeler. 2013. “Chapter One - Digital Mapping of Soil Carbon.” In *Advances in Agronomy*, edited by Donald L. Sparks, 118:1–47. Advances in Agronomy. Academic Press. https://doi.org/10.1016/B978-0-12-405942-9.00001-3.

Sanford, Posner, G. R., and T. Lin. 2012. “Soil Carbon Lost from Mollisols of the North Central U.S.A. With 20 Years of Agricultural Best Management Practices.” *Agriculture Ecosystems and Environment* 162: 68–76. https://doi.org/10.1016/j.agee.2012.08.011.

Webster, R., and M. Lark. 2012. *Field Sampling for Environmental Science and Management*. Vol. 1. 1. Routledge. https://doi.org/10.4324/9780203128640.