ESMC
Ecosystem Services Market Consortium

Determining Optimal Sample Size Using Power Analysis

Ecosystem Services Market Consortium/
Ecosystem Services Market Research Consortium
January 2023

Table of Contents

Funding: Research in this publication was supported by the Foundation for Food and Agriculture Research under award number – Grant ID: DFs-18-0000000012

Attribution: This report was commissioned by ESMC/ESMRC and authored by Dr. Dan TerAvest[1], Joel McClure[1], Dr. Dan Kane under contract to ESMC, January 2023.

[1]Our Sci, LLC

Submission Contact: Dan TerAvest, dan@our-sci.net

Introduction

The soil stratification process developed with ESMC uses ancillary data that co-varies with SOC stocks – NDVI, slope, organic matter, and clay content. Using this ancillary data allows us to take advantage of patterns of spatial variability within pilot project fields to lower sampling densities and reduce cost. Currently, however, the stratification process uses a default sampling density instead of a deterministic process related to the spatial variability of each field.

This study, part of Our Sci, LLC's "Pilot Project Soil Sampling Protocol Evaluation" project, seeks to evaluate a power analysis method for determining the "optimal" sampling size ex-ante. If appropriate, the power analysis could then be integrated into the stratification workflow such that the outcomes were both where to collect samples and how many samples to collect.

We initially considered a traditional power analysis approach, which determines the necessary number of samples to detect a specific magnitude of change in a sample population (i.e., collection of SOC samples in a field) given its background variability. However, we decided against using this type of power analysis for a couple of reasons. First, it relies on an arbitrary setting of the level of change it seeks to detect. Second, ESMC is conducting baseline sampling and will not be re-sampling to determine change for numerous years, so MDD isn't the most appropriate approach for the current stage of the project. Instead, we sought a power analysis method that recommends sampling densities based on their ability to explain the variance observed in the covariate ancillary data layers.

Method Evaluation

We randomly selected 104 fields from 2021 and 2022 ESMC pilot projects for post hoc analysis of the "optimal" sampling density. These fields were selected from 14 different projects, with between one and ten fields selected per project, depending on the number of fields in the project. The field size distribution of the 104 fields selected for this analysis were representative of the field size distribution of the 1019 ESMC pilot fields stratified in 2021-2022 (Figure 1 A, B).

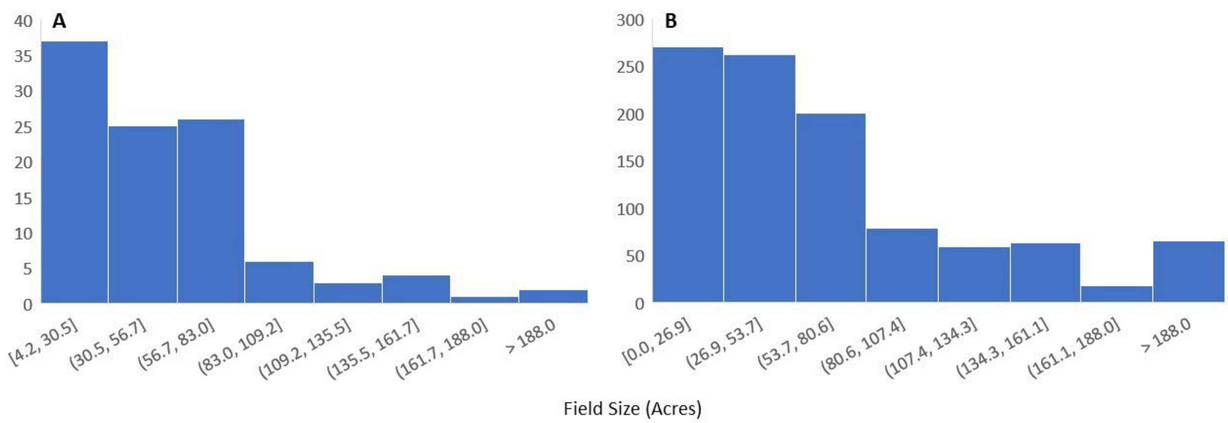

Figure 1. Distribution of field sizes in 104 fields selected for this report (A) and all 1019 fields stratified for ESMC during the 2021-2022 seasons (B).

We developed a power analysis test adapted from the technique described in Malone et al, 2019[1] to compare an "optimal" sampling density to ESMC's standard sampling density of one sample per four acres. The power analysis examines the variability of ancillary data layers and calculates the number of samples required to explain a pre-determined percent of the variability seen in the data layers using a statistic called the Kullback-Liebler (KL) divergence. In this case, we choose two levels, 60% and 80%. Therefore, we are calculating the number of soil sampling locations needed to explain 60% or 80% of the variability observed in the covariate data.

The power analysis used the same four ancillary data layers used in the ESMC stratification process (NDVI, slope, clay, organic matter) so it fits seamlessly into the existing stratification process. The following steps were used for a combination of all four ancillary data layers (matching the stratification workflow) and each data layer individually to evaluate the impact of each data layer on the "optimal" sampling size:

1. Repeated cLHS sampling was conducted using sequentially increasing samples sizes from 5 to 500 samples using increments of 5. This process was repeated 10 times at each sample size.
2. The KL divergence statistic was calculated for each simulated sample and normalized to a 0 to 1 scale.
3. The average KL divergence across all 10 iterations at each sample size was taken.
4. The "optimal" sample sizes were determined as the sample size number at which the KL statistic exceeded 0.6 (60% of variability explained) or 0.8 (80% of variability explained).

The optimal sampling densities at both levels using different data layers are presented in Table 1. Reaching the "optimal" sampling density would require 3.2 and 1.3 times more sampling locations than the ESMC sampling density at the 80% and 60% levels, respectively. Both the sampling densities and their standard deviations were much smaller when using all four ancillary layers combined. Conversely, using only the slope or organic matter layer individually would result in optimal sampling locations 8.3 and 5.7 times higher than the default sampling density, respectively. Using NDVI or clay content individually would result in lower optimal sampling locations, but the variability of those outcomes was much higher than using all layers together.

Table 1. Average "optimal" sampling locations using the ESMC default density and different combinations of power analysis and ancillary data layers.

	Sampling locations (#)	Standard deviation
ESMC Default sampling density (1 / 4 acres)	13.9	11.1
60% of variance explained (all ancillary data)	18.5	3.3
80% of variance explained (all ancillary data)	44.5	9.1
NDVI (80% of variance explained)	14.3	69.8
Slope (80% of variance explained)	114.8	72.0
Organic Matter (80% of variance explained)	79.0	260.8
Clay content (80% of variance explained)	18.1	32.5

As a function of field size, the ESMC density would not intersect with the 80% threshold until field sizes exceeded 175 acres (Figure 2). Even fields smaller than 20 acres would require between 20 and 40 sampling locations to explain 80% of the variability. If the threshold were lowered to 60%, then the number of sampling locations recommended by the power analysis and the ESMC density would intersect when field size reached approximately 75 acres. While this threshold is set quite low, it does suggest that, at a minimum, the sampling density of larger fields could be reduced to lower costs without reducing the modelling accuracy, compared to smaller fields.

At the 60% threshold the recommended sample size was largely independent of field size and intercepted the y-axis at 18.16 samples. At the 80% threshold the number of recommended samples started at 41.51 and increased slightly with field size (m = 0.0563). The ESMC sampling density is not related to the spatial variance of the ancillary layers and had the steepest slope (m = 0.2493).

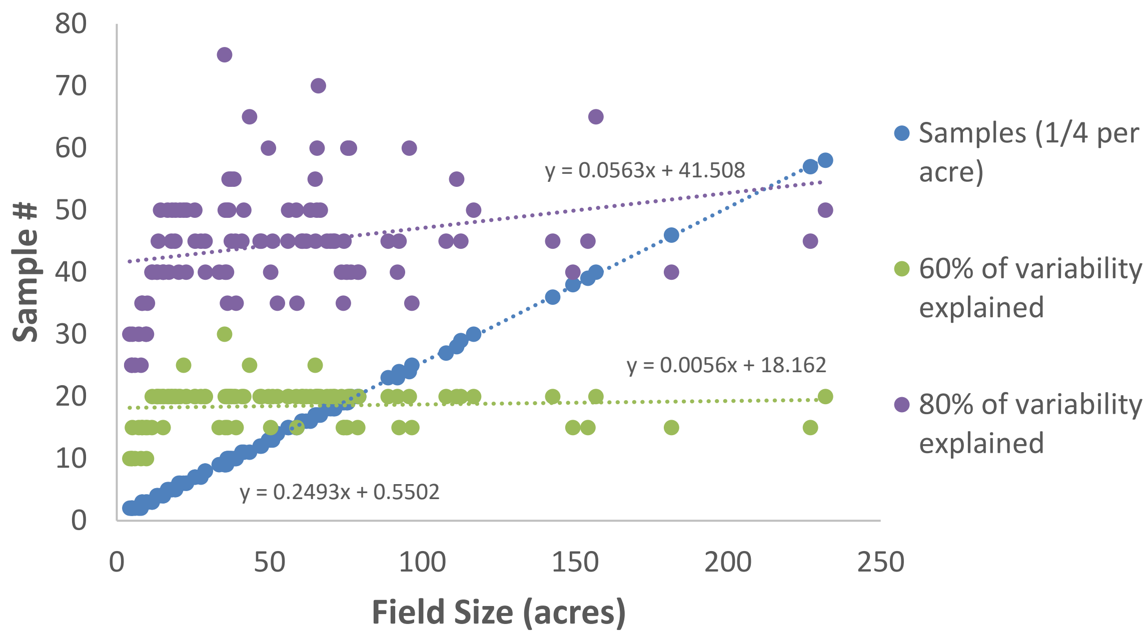

Figure 2. Comparison of ESMC standard sampling density (1 sample = 4 acres) to “optimum” sampling densities that explain 60% and 80% of the error in the ancillary data layers used.

Conclusions

The power analysis evaluated here is limited to a post hoc analysis of recommended sample sizes using only the ancillary data used in the stratification process. A more comprehensive analysis of sample densities would require the soil carbon and bulk density results for each sampling location and the results of SOC stock modelling along with the associated uncertainty estimates. However, these limited results do highlight a few questions that should be studied in greater detail as part of a more comprehensive assessment.

In this report, slope was the most variable of the ancillary data layers used and therefore resulted in the greatest recommended sampling size. Slope is included in the stratification because it is believed to be a covariate of SOC stocks. A more in-depth analysis of the impact of slope in soil stratification should be conducted.

The recommended sampling densities generated using power analysis were much less dependent on field size than the ESMC sample density and would have required far more sampling locations in small fields. These results suggest that the sampling density should be adjusted based on field size and that larger fields could be sampled at a lower sample density than small fields. Given that variability in soil carbon stocks is not expected to scale linearly with field size, this result is sensible. A more comprehensive analysis should compare uncertainty estimates by field size to evaluate if uncertainty follows the same pattern of the power analysis observed in this evaluation. If so, sampling densities could be reduced as field size increases without negatively affecting modelling outcomes.

Finally, while we were able to create a functional workflow for ex-ante power analysis to suggest sampling densities for soil carbon sampling, it’s unclear how reliable such methods may be in practical application. Importantly, they rely on readily available spatial data for various environmental covariates as an input, and while these covariates are broadly related to soil carbon stocks, they are unlikely to be perfectly representative of soil carbon stocks. Further testing is necessary to determine which of the available layers are most useful in such analyses to inform best methods for developing ex-ante sample planning tools.

References

[1]Malone BP, Minansy B, Brungard C. Some methods to improve the utility of conditioned Latin hypercube sampling. PeerJ. 2019 Feb 25;7. doi: 10.7717/peerj.6451