



Federal Soil Carbon Data Sources

Analysis of long term carbon monitoring for the USDA Soil Carbon Monitoring Network

Background: USDA is faced with the monumental task of establishing a soil carbon monitoring network for agricultural lands. Measuring soil carbon is not a new endeavor for USDA or other federal agencies – currently many disparate efforts exist focused on long term inventorying or monitoring of soil carbon alongside other ecosystem services across federal agencies. USDA may be able to leverage some of these ongoing efforts to augment the soil carbon monitoring network (SCMN), but not all of these efforts are created equally: they differ in scientific rigor (i.e. standardization of methods, length of monitoring), and in purpose (i.e. monitoring of natural lands vs. working lands). Therefore, Carbon180 has compiled these datasets across the federal government and analyzed them for their potential to augment the USDA SCMN. We have ranked these datasets (Table 1) to inform USDA's work on coordinating existing soil carbon data as a foundation for the SCMN to be able to monitor changes in soil carbon in response to land management changes. For information on methods underpinning our analysis, please see [Appendix](#).

Overall recommendations and findings:

- **Move from inventorying to monitoring:** Current datasets on agricultural lands tend to focus on soil carbon inventorying more than monitoring soil carbon in response to management changes (Table 1; See methods for more information). To realize its full potential, the new SCMN should prioritize (1) long term monitoring of changes in soil carbon due to land management changes (or practice adoption) to help inform producer decisions, and (2) maintenance of inventorying soil carbon to ensure a robust baseline for which to compare monitoring data.
- **Need for a dedicated data management team:** It is essential USDA establish a dedicated team for the SCMN to focus on continual surveying and synthesis of data across sites and time. Ideally, a SCMN will incorporate monitoring data at least on a yearly basis from copious numbers of physical samples across soil types, geographies, and operation types. The NSF Long Term Ecological Research (LTER) Network [Office](#) can serve as a model for establishing a data coordination office at USDA.
- **Lay the foundation for better predictive modeling:** USDA could spur the development of a diversity of models (Table 2) by (1) scientific assessment of each to ensure a specific level of maturity, and (2) equipping the COMET platform and tools with the ability to use different models in mature stages of development. All current soil carbon models reviewed still require robust long term, wide-spread soil carbon data for model development and validation to accurately predict soil carbon changes linked to management changes. With



long term soil carbon data, we may find that certain models work better for different regions, practices, or crop systems. USDA should also provide more robust guidance on the appropriate use for different models, preventing the current situation of models being used for unintended purposes (i.e., calculating carbon credits with COMET).

Top datasets for SCMN. Based on our analysis (Table 1), we recommend USDA use the following datasets to establish the SCMN:

- *Primary datasets (Tier 1):* We identified short term USDA datasets and long term soil carbon data from other federal agencies that could augment the foundational, Tier 1 data: we recommend USDA rely on the **Long Term Agroecosystem Research (LTAR) Network, Greenhouse gas Reduction through Agricultural Carbon Enhancement network (GRACEnet), and Rapid Carbon Assessment (RaCA)**; and/or data from the Kellogg Soil Survey Lab). These datasets are top candidates for the foundation of the SCMN because they are overseen by USDA, and provide both monitoring (LTAR, GRACEnet) and baseline inventorying (RaCA) of soil carbon. The LTAR network has documented soil carbon in response to management practices in a research setting. Coordinating data across the LTAR sites will provide a robust monitoring dataset to base the SCMN on. In addition to these existing datasets, NRCS should consider collaborating with NIFA to fund a soil carbon monitoring data collection through the land grant universities, similar to [NSF's Research Coordination Network Program](#).
- *Secondary datasets (Tier 2 and 3):* We recommend USDA integrate Tier 2 data from USDA programs to augment baseline inventory or monitoring data. However, these datasets are limited in time and cannot be relied upon for long term input. In addition, we recommend USDA work with NSF and USGS to identify appropriate data from long term datasets supported by these departments. Although a majority of this data might be collected from natural ecosystems, some may be collected from working lands, for example, the Kellogg Biological Station in the LTER network– **USDA should consider data from managed land LTER and the National Ecological Observatory Network (NEON) sites as Tier 1 data.** USDA could work to directly integrate these data relating to managed lands and consider incorporating data from natural lands to better understand how general climate and weather variability affects soil carbon.
- *Other sources of soil carbon data (Tier 4):* In addition to long term datasets, the federal government funds a variety of research projects on soil carbon across departments (Tier 4). Some federally funded projects may be able to voluntarily contribute data into the SCMN by following a standardized methodology for reporting soil carbon data. It will be imperative to ensure that any data pulled from these projects has been produced with robust methodology and can be coordinated with existing SCMN data.
- *Supplementary data sources:* As the SCMN matures, there may be a need to incorporate environmental and climate related variables. For example, incorporating remote sensing data and above ground carbon flux data will expand the current scope of the SCMN and



provide more context to changes in soil carbon stocks. We recommend that following USDA's establishment of baseline and continuous monitoring methodologies and datasets for the SCMN, USDA work with other federal agencies, including NASA, EPA, and NSF to identify the most appropriate data to augment the SCMN.



Table 1: Prioritization of federal long term soil carbon data sets for establishment of SCMN

Data set/program	Tier	Best use for data	Agency lead (methods development)	Samples from working ag lands	Data synthesized across sites	Tie management practices to SOC	Data collection time	Data publicly available
Long Term Agroecosystem Research (LTAR) soil carbon data	1	Baseline + continuous monitoring	USDA Agricultural Research Service (ARS)	Yes	Variable- site dependent	Yes	Long term (decades)	In process
Greenhouse gas Reduction through Agricultural Carbon Enhancement Network (GRACEnet)	1	Baseline + continuous monitoring	USDA ARS	Yes	Yes	Yes	Long Term (decades)	Yes
Kellogg Soil Survey Laboratory	1	Baseline inventory	USDA Natural Resources Conservation Service (NRCS)	Yes	Yes	Maybe some sites	Long term (decades)	Yes
Rapid Carbon Assessment (RaCA)	1	Baseline inventory	USDA NRCS	Yes	Yes	No	Very limited (2 years)	Yes
Long Term Ecological Research (LTER) Network	2*	Baseline + continuous inventory (compare to managed sites)	National Science Foundation (NSF)	Few sites	Yes	Few sites	Long term (decades)	Yes
Conservation Effects Assessment Project (CEAP)	2	Baseline monitoring	USDA NRCS	Yes	Yes	Limited practices	Very limited (2x survey)	Possibly within USDA?
Conservation Reserve Program (CRP) soil carbon data	2	Baseline + continuous monitoring (comparison to	USDA reviewed-- Partner led	Maybe	Unclear	Limited practices	Limited (~9 year grants)	Not yet



		managed sites)						
National Ecological Observatory Network (NEON)	3*	Baseline + continuous inventory (compare to managed sites)	NSF funded, Battelle led	Maybe	Yes	Maybe few sites	Mid term (12 yrs of data)	Yes
US Geological Survey (USGS) Soil Carbon Measurements	3	Baseline + continuous inventory (compare to managed sites)	USGS	Not necessarily	Yes	No	Long term (decades)	Yes
Nutrient Network	3	Baseline + continuous inventory	NSF funded, University of Minnesota led	some sites (grazing lands)	Yes	Maybe some sites	Mid term (15 yrs of data)	Yes
Short term data collection & disparate efforts (examples)	Tier	Best use for data	Agency lead (methods development)	Samples from working ag lands	Data synthesized across sites	Tie management practices to SOC	Data collection time	Data publicly available
Conservation Evaluation and Monitoring Activity (CEMA)	4	Short term monitoring	USDA NSF	Yes	No	Needs methods update	Short term (depends on producer)	Possibly within USDA?
Partnerships for Climate Smart Commodities	4	Methods testing	USDA reviewed-- Partner led projects	Yes	Hopefully within each project	Yes	Limited ~5 year grants	Not yet
USDA externally funded research (ex. SARE, AFRI...)	4	Methods testing/short term monitoring	USDA NIFA reviewed-- Partner led projects	Yes	Within each project	Yes	Limited based on grant funding timeline	Yes, through publications
Conservation Innovation Grant (CIG)	4	Methods testing/short term monitoring	USDA NRCS reviewed-- Partners led projects	Yes	Within each project	Yes	Limited based on grant funding timeline	Yes, through reporting/publications



Non-USDA, Federally funded research (ex. Dept. of Energy (DOE) SMARTFARM, ARPA-E, NSF Division on Environmental Biology)	4	Methods testing/short term monitoring	Various (DOE, NSF) agency reviewed-- Partner led projects	Maybe	Within each project	Maybe	Limited based on length of grant	Yes, through publications & reporting
Supplementary Data for Modeling	Tier	Best use for data	Agency lead (methods development)	Samples from working ag lands	Data synthesized across sites	Tie management practices to SOC	Data collection time	Data publicly available
Earth Observation (ex. Harvest, Landsat)	2	Supplementing SOC models with above ground data and weather data (ex. yield)	National Aeronautics and Space Administration (NASA)	Yes-- remote sensing	Yes	Depending on project-- ex. Harvest SARA	Long term (decades)	Yes
GHG inventory	3	Supplementing SOC models with broader GHG data	Environmental Protection Agency (EPA)	Yes	Yes	No	Long term (decades)	Yes
Soils Data Harmonization Database (SoDaH)	3	Includes SOC data from LTER, NEON and other sites	NSF	Maybe	Yes	Maybe some	Incorporates different datasets of different lengths	Yes
CarbonTracker	3	Including carbon fluxes in models	National Oceanic and Atmospheric Administration (NOAA)	Yes-- global	Yes	No	Long term (decades)	Yes

**Data from LTER and NEON sites that focus on managed lands should be considered Tier 1 data*



Table 2: Existing modeling efforts to benefit from robust SCMN

Model	Tier	Model management	Validation method	Up-front <i>in situ</i> samples	Other methods incorporated	Public	Private sector champion
DayCent (Underpins COMET tools)	For most accurate estimation of soil carbon, all require more development/validation across a diversity of regions, operations, and practices	Colorado State University	Data from public lit & few <i>in situ</i> samples	Very few	Vegetation cover, climate data	Certain versions	Indigo
DNDC		University of New Hampshire	Remote sensing + data from public literature	Yes	<i>In situ</i> sampling, management practices, soil nutrient cycling (process based model)	Certain versions	Regrow
System of Systems		Idea championed by University of Illinois	Localized environmental, crop and management data	Yes	Remote sensing, carbon and nitrogen fluxes, machine learning	Unclear	Habiterre
Soil Mapper		EarthOptics	<i>In situ</i> samples + above ground sensors	Yes	Machine Learning, remote sensing, near surface sensing	Unclear	Earthoptics
Digital Soil Mapping		No individual lead	<i>In situ</i> samples	Yes	Machine Learning, remote sensing	Certain versions	Perennial
SALUS model		Michigan State University	Needs <i>in situ</i> samples-- based on modified Century Model (DayCent adjacent)	Very few	Artificial Intelligence, remote sensing, ecosystem modeling, geospatial data	Certain versions	CIBO



Appendix

Methods:

In this file, we have made recommendations for prioritizing certain datasets to build the foundation of the soil carbon monitoring network. The “Soil carbon in situ measurements” tab includes current federal soil carbon monitoring and inventory datasets as well as supplementary datasets (i.e. remote sensing, greenhouse gas fluxes) that can help build out the soil carbon monitoring network. The “Soil carbon modeling efforts” outlines existing modeling efforts for soil carbon stocks. Currently, we do not have enough information on the accuracy of each model to prioritize one over the other.

In carrying out this analysis we focused on the following criteria:

- *Soil carbon monitoring vs. inventorying*: Monitoring refers to measuring SOC changes in response to management practices, whereas inventorying refers to point measurements of SOC without taking into effect management changes. To establish which data sets focused on monitoring or inventorying, datasets focused on working lands, and tied soil carbon to management changes were prioritized higher than those on natural lands.
- *Baseline vs. continuous monitoring*: While inventory samples will help to establish a baseline for soil carbon stocks, the monitoring network will require a main focus on continuously monitoring changes in carbon stocks in response to management practices. Datasets focused on continuous monitoring were prioritized higher than those with a primary focus on establishing baseline data.
- *Length of the dataset*: Meaningful changes in soil carbon take years to measure post management change. Therefore, a top criteria for prioritizing datasets is the length of time they have been collecting this data- longer running datasets were prioritized higher than short term datasets.
- *Data coordination and management*: Data translation will be essential to ensure that external stakeholders can access the data from the soil carbon monitoring network and make meaningful inferences from it. Therefore, standardized methods across sites and data interoperability will be essential. We factored into our analysis which datasets currently have some level of coordination or synthesis.