Progress report

Progress in Physical Geography
2019, Vol. 43(5) 707–719

DOI: 10.1177/0309133319873309
journals.sagepub.com/home/ppg
SAGE

The landscape of soil carbon data: Emerging questions, synergies and databases

Avni Malhotra
Stanford University, USA

Katherine Todd-Brown
Wilfrid Laurier University, Canada

Lucas E Nave
University of Michigan, USA

Niels H Batjes
ISRIC – World Soil Information, The Netherlands

James R Holmquist
Smithsonian Environmental Research Center, USA

Alison M Hoyt
Max Planck Institute for Biogeochemistry, Germany; Lawrence Berkeley National Laboratory, USA

Colleen M Iversen
Environmental Sciences Division, Climate Change Science Institute, Oak Ridge National Laboratory, USA

Robert B Jackson
Stanford University, USA

Kate Lajtha
Oregon State University, USA

Corey Lawrence
US Geological Survey, Geologic and Environmental Change Science Center, USA

Olga Vindušková
Stanford University, USA

William Wieder
National Center for Atmospheric Research, USA; University of Colorado, USA

Mathew Williams
University of Edinburgh, UK

Gustaf Hugelius
Stockholm University, Sweden; Bolin Centre for Climate Research, Sweden

Jennifer Harden
Stanford University, USA

Abstract

Soil carbon has been measured for over a century in applications ranging from understanding biogeochemical processes in natural ecosystems to quantifying the productivity and health of managed systems. Consolidating diverse soil carbon datasets is increasingly important to maximize their value, particularly with growing anthropogenic and climate change pressures. In this progress report, we describe recent advances in soil carbon data led by the International Soil Carbon Network and other networks. We highlight priority areas of research requiring soil carbon data, including (a) quantifying boreal, arctic and wetland carbon stocks, (b) understanding the timescales of soil carbon persistence using radiocarbon and chronosequence studies, (c) synthesizing long-term and experimental data to inform carbon stock vulnerability to global change, (d) quantifying root influences on soil carbon and (e) identifying gaps in model–data integration. We also describe

Corresponding author:
Avni Malhotra, Department of Earth System Science, Stanford University, Stanford, CA, USA.
Email: avnim@stanford.edu

the landscape of soil datasets currently available, highlighting their strengths, weaknesses and synergies. Now more than ever, integrated soil data are needed to inform climate mitigation, land management and agricultural practices. This report will aid new data users in navigating various soil databases and encourage scientists to make their measurements publicly available and to join forces to find soil-related solutions.

Keywords
Soil carbon data, soil database, wetland carbon, soil chronosequence, soil radiocarbon, long-term ecological research, root traits, soil carbon stabilization, model–data integration

1 Introduction

Soil carbon is a key component in our understanding of the biosphere's response to global change. There is a long history of soil carbon measurements that, together with other types of soil and ecosystem data, contribute to our understanding of the health and functioning of natural and managed ecosystems (Harden et al., 2018). To better utilize this body of work, the International Soil Carbon Network (ISCN) was formed in 2012 to connect soil carbon researchers and their data. Here, we present recent international efforts consolidating soil carbon data to address urgent questions in soil carbon science. We highlight advances in soil databases, led by the ISCN or other organizations, to synthesize datasets from diverse sources. Examples include data from boreal, arctic and wetland soils, long-term soil experiments, chronosequences, soil radiocarbon observations and root–soil linkages. These new data will help us to understand soil carbon stocks, change and vulnerability via syntheses and model–data integration.

1.1 What is the ISCN?

The ISCN is a science-based network that provides: (a) scientific and logistical infrastructure for sharing knowledge, information and data; (b) opportunities for synthesis activities; (c) data products beneficial to stakeholders and scientists; and (d) a framework for common scientific protocols and collaborative decision support tools.

1.2 Why soil carbon?

Soil carbon storage and cycling are measures of soil health, where soil health is defined as the capacity of soil to maintain a range of functions such as food and fiber provision (Lal, 2004; Banwart et al., 2014). Soil carbon is also directly linked to exchanges of carbon dioxide and other trace gases between land–water and land–air systems and is therefore a key component in regulating the global climate system (Ciais et al., 2013). Because soils are a focal point of terrestrial carbon cycling, current research prioritizes quantifying global and ecosystem-specific carbon stocks. In addition to stocks, understanding the processes controlling soil carbon timescales and vulnerability to global change are also critical (Figure 1). These research priorities require diverse data types synthesized across broad scales.

1.3 Why now?

Land is increasingly under pressure to maintain healthy ecosystems while providing food and fiber to growing human populations. Over one-third of the global land surface is currently grazed, forested or cropped (Erb et al., 2007), rendering three-quarters or more of the soil carbon down to a meter depth under human management (Harden et al., 2018). Past land management has depleted soil carbon and organic matter (Sanderman et al., 2018). However, the re-establishment and build-up of this organic matter through best practices can improve soil productivity and resilience to extreme climate events

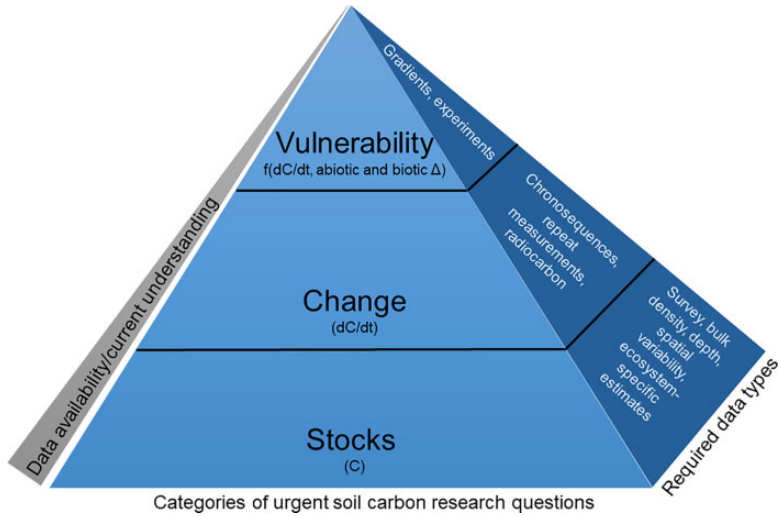

Figure 1. Key questions in soil carbon science and corresponding data requirements. Emergent questions about soil carbon vulnerability (e.g., the response to abiotic/climate or biotic/land cover change, management or disturbance) must be underpinned by questions of soil carbon change (timescales, persistence and stability, factors controlling microbial access, quality and fraction), which, in turn, are rooted in questions of carbon stocks (e.g., spatial variability, ecosystem-specific storage or depth variation). The syntheses described in this paper represent a range of data/efforts directed toward addressing each of these knowledge gaps. For example, research questions on soil carbon vulnerability may utilize data from experimental manipulations and plant trait databases; carbon change questions from databases such as the ISCN3 database, chronosequence and radiocarbon syntheses; and carbon stock questions from global or ecosystem-specific survey data.

while removing carbon dioxide from the atmosphere (Minasny et al., 2017; Batjes, 2019).

Although scientific research on soil carbon has led to numerous sources of data and information, such information is disparate and difficult to access (Harden et al., 2018). Communities interested in making carbon cycle projections or improving agricultural land management need synthesized data to evaluate soil carbon persistence and vulnerabilities to environmental change (Blankinship et al., 2018). With emerging technological advances in data, computing and instrumentation, we see an opportunity to inform and empower land managers with timely, relevant data and information for decision support.

1.4 ISCN data holdings

The ISCN database (latest version ISCN3; Nave et al., 2017) contains data from >70,000 soil profiles from a range of data sources, including the United States Department of Agriculture and the Northern Circumpolar Soil Carbon Database. More than 200 soil variables are present in the database, including the percentage of organic carbon, particle size distribution, pH and the percentage of nitrogen. Details of the data types and their calculations can be found at https://iscn.fluxdata.org/data/data-information/. The inclusion of a range of supporting measurements (describing, for example, the geography, soil properties and landform type) in the ISCN database makes it possible to investigate soil carbon as part of a dynamic cycle in addition to quantifying stocks. The strengths of the ISCN database include extensive coverage for soil profiles, horizons and depth internationally (with particularly strong representation of the US from United States Department of

Agriculture data) making ISCN one of the largest, most wide-ranging and diverse repositories of measured soil data.

II Recent advances in the ISCN database

2.1 Shift from template-only to script-based data ingestion

Historically, the ISCN has taken a template-based approach to data harmonization, in which data providers and curators manually input data into the ISCN database. Given that this approach can be both labor intensive and prone to errors, the ISCN is adding a scripted option for data users and providers. The SOC-DRaHR (Soil Organic Carbon Data Rescue and Harmonization Repository; https://github.com/ISCN/SOC-DRaHR) is a script repository with an associated R package designed to aid data ingestion and downloading. The SOC-DRaHR also provides a community platform to develop an R library to access and harmonize different data collections.

The SOC-DRaHR identifies and downloads soil carbon datasets that are publicly available, provides data harmonization scripts to integrate these datasets into R and provides output scripts for a harmonized data product. In short, these scripts match the variable names of the dataset to be ingested with those contained in the ISCN template. The SOC-DRaHR is not a data repository or archive, but is an open source software project that facilitates access to data and harmonizes units and naming conventions across data collections. One limitation of a script-based approach is that it may decrease data user/provider accessibility if they do not have experience with R or other programming languages. To address this, we plan to keep the template option for users that prefer it.

Led by Katherine Todd-Brown, the ISCN hosted two data hackathons (New Orleans, LA, USA in 2016; College Station, TX, USA in 2017) to train potential contributors and users of ISCN data on our scripted approach. We also provided guidance and expertise to other science communities building soil or ecological databases (Table 1).

2.2 Shift toward open data

The ISCN3 database contains data from sources with varying data-use policies (http://iscn.fluxdata.org/data/data-information/data-policy-and-use/). In the future, the ISCN4 database and subsequent versions will only contain data that are open source under a Creative Commons Attribution (CC-BY) License. The key update under this license will be that the requirements of data-provider involvement will be removed, but data attribution will be required as before. Previous versions of data bound to sharing restrictions will be retained, but will only be available through the ISCN3 database. The ISCN4 database will include the open source data from ISCN3 plus new datasets (Table 1). We consider this open source shift an important step in making the ISCN data easily accessible and usable.

2.3 ISCN-led community activities

We held our most recent all-hands meeting at the American Geophysical Union (AGU) fall meeting in December 2017. The meeting included updates from the ISCN as well as breakout groups on root–soil linkages, wetland soil carbon, the turnover times of soil carbon and reconciling multi-scale data (http://iscn.fluxdata.org/summary-of-pre-agu-2017-activities/). We also organized oral and poster sessions at AGU 2017 and 2018. In February 2017, we organized a workshop (Loisel et al., 2017) to discuss and define research and data priorities for soil carbon science and for the ISCN. We drafted an article highlighting the converging needs of the soil carbon science and soil health communities and the way forward for the ISCN (Harden et al., 2018). The ISCN plans to continue to coordinate and host workshops,

Table 1. New data and research synergies within the ISCN and across several working groups.

Dataset name	Dominant data type	Number of observations	Attributes	Geographical coverage	URL
ISCN 3	Survey	70k+ profiles	>200	Global, but US-heavy	https://iscn.fluxdata.org/data/data-information/
Shaw et al. (2018)	Canadian forest soils	3000 profiles	60	Canada	Included in the ISCN database
Treat et al. (2015)	Peatlands	500 cores	30	Global	Included in the ISCN database
C-PEAT	Peatlands	82 cores	10	Global	Included in the ISCN database
Coastal Carbon RCN	Wetland carbon	3000 profiles	131	Global, but US-heavy	https://github.com/Smithsonian/CCRCN-Data-Library
ISRaD	Radiocarbon	1700 profiles	>250	Global	www.soilradiocarbon.org
LTER SOM	Experimental, repeat measurements	140 locations	170	Global, but US-heavy	https://lter.github.io/som-website/
FRED	Below-ground trait data	105,000 root trait observations	300	Global	https://roots.ornl.gov/
ISRIC (WoSIS)	Survey	150,000 profiles	24	Global	www.isric.org/explore/wosis/accessing-wosis-derived-datasets

data hackathons and scientific sessions at international meetings (e.g., the AGU and the European Geophysical Union).

III New datasets and emerging ISCN partnerships

3.1 Advances in northern and wetland soil carbon data

Northern peatlands and permafrost soils are rich in carbon that is vulnerable to increased rates of warming and other feedbacks with climate change (Gorham, 1991; Oechel et al., 1993; Frolking et al., 2006; Tarnocai et al., 2009; Schuur et al., 2015). The drivers of soil carbon storage in organic soils can vary considerably relative to those in mineral soils (Limpens et al., 2008; Loranty et al., 2018; Malhotra et al., 2018a; Schuur and Mack, 2018). To better place these soils in a global context, the ISCN is including more data from peatlands (Treat et al., 2016, data from C-PEAT; https://github.com/ISCN/soilDataR/blob/master/R/readCPEAT.R) in the next version of the database (ISCN4; Table 1). We plan to also include Canadian forest soil surveys representing a decade of data (Shaw et al., 2018). Although not always organic soils, these northern forest soils are also expected to undergo warming (Meehl et al., 2007) and provide opportunities for contrasting studies of mineral and organic soils across climate gradients.

3.2 Bridging gaps in soil data types

The strengths of the ISCN3 database lie in global survey data that are reported with a range of supporting measurements and are best suited to investigating the mechanisms of soil carbon change (Figure 1). However, other types of data are necessary for carbon stock and vulnerability

questions (Figure 1)—for example, data from coastal systems, radiocarbon measurements, soil chronosequences, experiments (field manipulations), long-term repeat measurements or root–soil linkages. In an effort to increase our representation of diverse data types, we have built informal (e.g., sharing best practices and data harmonization scripts) or formal (memoranda of understanding) synergies with the various groups discussed in the following sections.

3.2.1 Coastal wetland carbon. Coastal wetlands are highly productive and, because they form soil as a dynamic response to sea-level rise (Kirwan and Megonigal, 2013), they act as long-term carbon sinks. "Blue carbon" syntheses have been used to support local greenhouse gas mitigation efforts (Kroeger et al., 2017) to include coastal wetlands in national-scale greenhouse gas inventories (Crooks et al., 2018; Holmquist et al., 2018) and to complete terrestrial–aquatic interface carbon budgets (Najjar et al., 2018). There is a tremendous need for a transparent, well-sourced and living synthesis of coastal carbon stocks. The Coastal Carbon Research Coordination Network (CC-RCN) is currently building such a dataset iteratively, producing standards for data formatting, assisting researchers in creating citable open data releases (Reichman et al., 2011; Wilson et al., 2017) and compiling public data releases into a central data clearinghouse. CC-RCN personnel are available (until at least 2021) to help providers prepare datasets for submission. To date, the CC-RCN has synthesized data from 3117 cores from salt marshes, mangroves and tidal freshwater wetlands of the contiguous USA (Holmquist et al., 2018) and from around the world. The ISCN and CC-RCN share lessons-learned on database best practices through workshops and hackathons. In the future, we aspire to formally link our databases through the SOC-DRaRH.

3.2.2 Soil radiocarbon data. The International Soil Radiocarbon Database (ISRaD) is an open source community-based project that brings together soil radiocarbon data and associated datasets (Lawrence et al., 2019). Radiocarbon data are an important tool for understanding the soil carbon cycle and can be used to constrain rates of carbon cycling in models (He et al., 2016) and to assess the timescales and persistence of soil carbon (Sierra et al., 2018). In particular, the application of radiocarbon methodology to improve our understanding of soil carbon dynamics has emphasized the need to conceptualize soils as a consortium of different carbon types, stabilized in soils via a variety of mechanisms. As such, there is a growing abundance of soil data collected from specific soil "fractions" that have been physically (e.g., density or particle size separation), chemically (e.g., chemical extraction) or biologically (e.g., soil incubation) partitioned from the bulk soil (Poeplau et al., 2018). Although these data may provide insights into the nature of a particular soil, it is often challenging to compare fractions across different soils because the fractional methods vary widely. The ISRaD also seeks to improve our ability to compare soil fractions and standardize fractionation methods, in addition to making soil radiocarbon data more accessible.

The data within ISRaD are structured hierarchically and include bulk soil radiocarbon data (about 500 sites and 1700 profiles), fractionation schemes (>3600 data points entered), flux measurements (>2100), incubations (>1900), interstitial gases and dissolved organics. Users can add data through a template, which is structured to reflect this hierarchy, or use a scripted approach for larger datasets. In addition to the dataset, the ISRaD also offers an associated R package, which includes quality-control checks and tools for exploring the data. Ongoing synthesis activities have compiled radiocarbon data from carbon fluxes in the arctic region to look at the potential release of old

permafrost carbon from soil incubations to assess rates of fast-cycling soil carbon and from different soil fractions. Although radiocarbon is the focus of the database, it is not a requirement, allowing the template, data structure and associated tools to be used for other synthesis efforts related to soil carbon. The ISRaD data template builds upon the ISCN template and profile-level soil data will be shared between the ISRaD and the ISCN databases.

3.2.3 Soil chronosequence data. Understanding long-term soil carbon dynamics is important for constraining the capacity of soils to store carbon and the spatiotemporal variations in soil carbon related to pedogenic mineralogy. The chronosequence approach has been traditionally used to study the role of time in pedogenesis (Stevens and Walker, 1970). As a result, many chronosequence studies have reported soil carbon data along with other soil and environmental variables. Comparisons of several chronosequences have been used to determine general patterns in soil and ecosystem development and to investigate the effects of other soil-forming factors on carbon, nutrients and mineralogy (Wardle et al., 2004). A recent effort synthesized data from soil chronosequences with the goal of determining the controls on long-term soil carbon dynamics during soil development. The structure of this dataset follows the hierarchical structure of the ISCN dataset and draws upon the ISRaD dataset in terms of the included variables and tools for data analysis. Upon completion, data from this synthesis will be ingested into the ISCN database.

3.2.4 Experimental and long-term data. Cross-site analysis is a central goal of the Long-term Ecological Research (LTER) program and significant advances have been made in synthesizing cross-site data in hydrology, vegetation dynamics, diversity and climate (Peters et al., 2013). Although soil carbon has been measured at almost all the LTER sites as well as at sites from other research networks, cross-network data have, to our knowledge, never been synthesized, compared, modeled or archived in standardized ways across sites (Weintraub et al., 2019). A new synthesis project is addressing this gap by synthesizing long-term soil carbon data not just from the LTER sites, but also from the National Ecological Observatory Network (NEON), the Critical Zone Observatory (CZO), the Detritus Input and Removal Treatment (DIRT) and the Nutrient Network (NutNet) sites. This project uses a scripted approach similar to the ISCN and ISRaD databases and involves researchers who developed soil models such as MIMICS and CORPSE (Sulman et al., 2014; Wieder et al., 2015), as well as the principal investigators who collected the soil carbon data. The model–data synthesis aims to answer questions such as the roles microbial and plant community composition have in transferring microbial byproducts to persistent soil organic matter and how nitrogen deposition affects soil organic matter composition across a range of climate types and mineralogies. Practical implications include outreach to land managers concerned with soil carbon consequences of specific practices.

3.2.5 Linking root traits to soil carbon. Plant root inputs are more likely to be stabilized as long-term soil carbon relative to above-ground plant inputs (Jackson et al., 2017; Sokol and Bradford, 2018; Sokol et al., 2018). Despite their recognized importance in soil carbon dynamics, data on root attributes or traits (e.g., root biomass or rooting depth) are severely lacking in soil databases (Harden et al., 2018). Root observations from across the globe have recently been compiled into the Fine-Root Ecology Database (FRED; Iversen et al., 2017). FRED version 2.0 includes more than 100,000 root trait observations and relevant ancillary data such as soil properties, providing an opportunity to harmonize soil and root data. The ISCN has

held breakout group discussions and a workshop to develop a framework linking root traits with soil carbon across the globe (Malhotra et al. 2018b). The root trait working group will continue their efforts, focusing on the three main stages of root–soil interactions—namely, rhizosphere engineering by living roots, root inputs to soil organic matter via turnover and the decay of root necromass throughout the soil profile.

3.2.6 Mechanisms of soil carbon storage and stability. In coordination with the ISCN, the U.S. Geological Survey (USGS) and the U.S. Department of Agriculture recently supported a series of USGS Powell Center workshops targeted toward improving our understanding of the mechanisms controlling soil carbon storage and stability. Several products were derived from these workshops, including an exploration of how soil measurements, models and theories are linked to better integrate rapidly expanding soil research efforts (Blankinship et al., 2018) and a re-evaluation of soil carbon controls using existing databases (Rasmussen et al., 2018). The results of these workshops highlight the crucial importance of including ancillary soil data in soil carbon syntheses and provide further opportunities to better coordinate future soil measurements with models and theory.

3.2.7 Model–data integration. Soil data synthesis efforts strive to inform model development and validation. Model evaluation is an important goal of the International Land Model Benchmarking Project (ILAMB; Collier et al., 2018). The ISCN participated in ILAMB's soil organic carbon working group (Oak Ridge National Laboratory, October 2018) to develop ISCN-derived data products that would be useful for model benchmarking.

Beyond benchmarking, there is a growing potential to use synthesized datasets for model–data integration to develop our understanding of soil carbon dynamics (Bloom et al., 2016; Luo et al., 2016). Model–data integration activities can help to determine model structures and parameterizations that are consistent with observations of carbon stocks, soil ages (radiocarbon data), above- and below-ground litter inputs and local conditions (soil texture, moisture and temperature), weighted according to measurement errors. Advances in computing power and algorithm development allow model calibration and evaluation across very large datasets, facilitating our capacity to simulate soil processes both regionally and globally. A key request from the model–data integration community is that soil databases include a clear quantification of all sources of measurement error (to allow for Bayesian modeling approaches). In addition, if point data have been converted to gridded products, the upscaling error is a key factor in model–data integration.

IV Navigating the landscape of soil data

The landscape of soil data is complicated and contains a range of databases representing different regions and variables (Figure 2). To a new data user (e.g., a graduate student), it may be daunting to select the right dataset to answer a research question or the best database to target for their data contributions. One of the ISCN's missions is to inform data users of the strengths and weaknesses of each database and to circumvent the issues related to multiple soil databases that are difficult to harmonize. Our recent synergies with the CC-RCN, ISRaD, chronosequences and the LTER/NEON/CZO data syntheses were therefore initiated with the intention of sharing information on best practices, standardizing controlled vocabularies and providing resources such as R scripts to ingest or harmonize data.

In addition, the ISCN and the International Soil Reference and Information Centre (ISRIC) developed a formal agreement to ensure that the ISCN soil profile data are fed

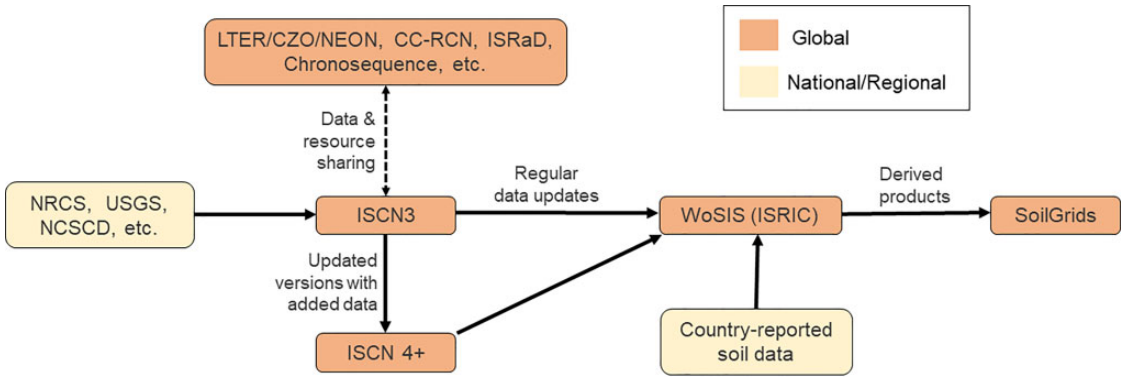

Figure 2. Navigating the landscape of soil data: the ISCN3 database and its current links to other large soil databases. The ISCN3 database comprises various independent data sources that are globally extensive, but with a strong US focus. The data sources include the Natural Resources Conservation Service (NRCS), the United States Geological Survey (USGS) and the Northern Circumpolar Soil Carbon Database (NCSCD). The ISCN publishes a new data version periodically (e.g., ISCN4 will contain new northern and peatland data). In turn, the ISCN data are regularly ingested into the World Soil Information Service database (WoSIS), a larger database focused on nationally reported profile data. Global gridded products such as SoilGrids are derived from profiles held in WoSIS, a set of environmental covariates and digital soil mapping. The ISCN database maintains synergies with various other data synthesis groups (e.g., the International Soil Radiocarbon Database (ISRaD), the Coastal Carbon Research Coordination Network (CC-RCN) and the Long-term Ecological Research (LTER) program; see Section 3.2), which encompass data types not well represented by the ISCN database (e.g., radiocarbon, coastal carbon and experimental manipulations).

into ISRIC's spatially extensive database on a regular basis, following a screening for possible duplicate profiles (Ribeiro et al., 2018). If users are interested in global carbon stocks, they may use the entire World Soil Information Service (WoSIS) database (Batjes et al., 2017) or its derived products (SoilGrids250 m; Hengl et al., 2017). Alternatively, if users are interested in abiotic or mechanistic controls of soil carbon, the ISCN database may be more appropriate because it provides more ancillary data on soil properties and ecology than the ISRIC database.

V Future directions

In the short term, our goal is to provide data infrastructure that enables interoperability, not just between the ISCN's data sources, but also across the synthesis efforts mentioned here. This is a non-trivial task, but the community is ready and the need for harmonized soil datasets is clear.

In the long term, in addition to maintaining the aforementioned data and infrastructure, we would also like to consolidate new data sources and types. Most urgently, given that the extent of managed soils exceeds that of unmanaged soils globally (Harden et al., 2018), the ISCN would like to include more data from agricultural and other managed systems. We hope to continue our discussions with entities such as FarmOS (https://farmos.org/) and the Coordination of International Research Cooperation on soil CArbon Sequestration in Agriculture (CIRCASA; www.circasa-project.eu/) to consolidate agricultural data into a central repository. This first step is necessary to link management practices with the resulting soil properties.

Activities summarized in this report highlight emerging priorities within soil carbon science. We especially highlight recent advances in high-latitude soils and at the terrestrial–aquatic interface as well as in experimental, long-term, chronosequence or radiocarbon data. In a complex landscape of soil carbon data and applications, the ISCN and our partners strive to provide resources, data and opportunities for disparate soil carbon communities to exchange ideas and solutions. Promoting healthy soils and finding creative solutions for climate change mitigation and adaptation will require collaboration among land managers, policy-makers and scientists. Our report serves as a call for input from these diverse and often disconnected communities. In addition to data contributions, we welcome and encourage working groups that can synthesize existing data to bridge the gap between soil carbon science and soil management.

Acknowledgements

ISCN is self-chartered and has become a volunteer-run entity since its original funding by USDA. ISCN would not exist without the generous support, both of time and funding, from various organizations and individuals (see https://iscn.fluxdata.org/ for detailed contributions). USDA and USGS funding were critical for the development of ISCN. Data infrastructure continues to be supported by Lawrence Berkeley National Laboratory. We especially thank Deborah Agarwal, Marty Humphrey and Norman Beekwilder.

Declaration of conflicting interests

The author(s) declared no potential conflicts of interest with respect to the research, authorship, and/or publication of this article.

Funding

The author(s) disclosed receipt of the following financial support for the research, authorship, and/or publication of this article: A subset of the workshops and synthesis efforts mentioned in this report were funded by Bonanza Creek LTER, Stanford University's Cox Fellowship and The Bolin Centre for Climate Research at Stockholm University. G. Hugelius acknowledges the Swedish Research Council (2014-06417 and 2018-04516). J. Holmquist and the Coastal Carbon Research Coordination Network are funded by NSF Grant [DEB 1655622]. ISRaD and A. Hoyt are funded by the European Research Council (ERC) under the EU Horizon 2020 programme [Grant Agreement No. 695101 (14Constraint)]. The experimental and long-term data synthesis is being conducted as a part of the LTER-CZO-NEON soil organic matter research Synthesis Group funded by the National Science Foundation under Grant DEB#1545288, through the Long Term Ecological Research Network Communications Office (LNCO), National Center for Ecological Analysis and Synthesis (NCEAS), University of California-Santa Barbara. Funding for the root trait and soil carbon workshop came from the Climate Change Science Institute at Oak Ridge National Laboratory. C.M. Iversen and the Fine-Root Ecology Database are supported by the Department of Energy's Office of Science, Biological and Environmental Research Program. Mechanisms of soil carbon storage and stability working group was supported by the "What Lies Below? Improving quantification and prediction of soil carbon storage, stability and susceptibility to disturbance" by the John Wesley Powell Center for Analysis and Synthesis, funded by the U.S. Geological Survey.

ORCID iD

Avni Malhotra https://orcid.org/0000-0002-7850-6402

References

Banwart S, Black H, Cai Z, et al. (2014) Benefits of soil carbon: Report on the outcomes of an international scientific committee on problems of the environment rapid assessment workshop. *Carbon Management* 5(2): 185–192. DOI: 10.1080/17583004.2014.913380.

Batjes NH (2019) Technologically achievable soil organic carbon sequestration in world croplands and grasslands. *Land Degradation and Development* 30: 25–32.

Batjes NH, Ribeiro E, Van Oostrum A, et al. (2017) WoSIS: Providing standardised soil profile data for the world. *Earth System Science Data* 9(1): 1–14. DOI: 10.5194/essd-9-1-2017.

Blankinship JC, Berhe AA, Crow SE, et al. (2018) Improving understanding of soil organic matter

dynamics by triangulating theories, measurements, and models. *Biogeochemistry* 140(1): 1–13. DOI: 10.1007/s10533-018-0478-2.

Bloom AA, Exbrayat J-F, van der Velde IR, et al. (2016) The decadal state of the terrestrial carbon cycle: Global retrievals of terrestrial carbon allocation, pools, and residence times. *Proceedings of the National Academy of Sciences* 113(5): 1285–1290. DOI: 10.1073/pnas.1515160113.

Ciais P, Gasser T, Paris JD, et al. (2013) Attributing the increase in atmospheric CO_2 to emitters and absorbers. *Nature Climate Change* 3(10): 926–930. DOI: 10.1038/nclimate1942.

Collier N, Hoffman FM, Lawrence DM, et al. (2018) The International Land Model Benchmarking (ILAMB) System: Design, theory, and implementation. *Journal of Advances in Modeling Earth Systems* 10(11): 2731–2754. DOI: 10.1029/2018MS001354.

Crooks S, Sutton-Grier AE, Troxler TG, et al. (2018) Coastal wetland management as a contribution to the US National Greenhouse Gas Inventory. *Nature Climate Change* 8: 1109–1112. DOI: 10.1038/s41558-018-0345-0.

Erb KH, Gaube V, Krausmann F, et al. (2007) A comprehensive global 5 min resolution land-use data set for the year 2000 consistent with national census data. *Journal of Land Use Science* 2(3): 191–224. DOI: 10.1080/17474230701622981.

Frolking S, Roulet N and Fuglestvedt J (2006) How northern peatlands influence the Earth's radiative budget: Sustained methane emission versus sustained carbon sequestration. *Journal of Geophysical Research: Biogeosciences* 111(1): 1–10.

Gorham E (1991) Northern peatlands: Role in the carbon cycle and probable responses to climatic warming. *Ecological Applications* 1(2): 182–195.

Harden JW, Hugelius G, Ahlström A, et al. (2018) Networking our science to characterize the state, vulnerabilities, and management opportunities of soil organic matter. *Global Change Biology* 24(2). DOI: 10.1111/gcb.13896.

He Y, Trumbore SE, Torn MS, et al. (2016) Radiocarbon constraints imply reduced carbon uptake by soils during the 21st century. *Science* 353(6306): 1419–1424.

Hengl T, Mendes de Jesus J, Heuvelink GBM, et al. (2017) SoilGrids250 m: Global gridded soil information based on machine learning. *PLoS One* 12(2): e0169748. DOI: 10.1371/journal.pone.0169748.

Holmquist JR, Windham-Myers L, Bliss N, et al. (2018) Accuracy and precision of tidal wetland soil carbon mapping in the conterminous United States. *Scientific Reports* 8(1): 9478. DOI: 10.1038/s41598-018-26948-7.

Iversen CM, Powell AS, McCormack ML, et al. (2017) Fine-Root Ecology Database (FRED): A global collection of root trait data with coincident site, vegetation, edaphic, and climatic data, version 1. DOI: 10.3334/cdiac/ornlsfa.005.

Jackson RB, Lajtha K, Crow SE, et al. (2017) The ecology of soil carbon: pools, vulnerabilities, and biotic and abiotic controls. *Annual Review of Ecology, Evolution, and Systematics* 48(1): 419–445. DOI: 10.1146/annurev-ecolsys-112414-054234.

Kirwan ML and Megonigal JP (2013) Tidal wetland stability in the face of human impacts and sea-level rise. *Nature* 504(7478): 53–60.

Kroeger KD, Crooks S, Moseman-Valtierra S, et al. (2017) Restoring tides to reduce methane emissions in impounded wetlands: A new and potent Blue Carbon climate change intervention. *Scientific Reports* 7(1): 1–12. DOI: 10.1038/s41598-017-12138-4.

Lal R (2004) Soil carbon sequestration impacts on global climate change and food security. *Science* 304(5677): 1623–1627.

Lawrence CR, Beem-Miller J, Hoyt AM, et al. (2019) An open source database for the synthesis of soil radiocarbon data: ISRaD version 1.0. *Earth System Science Data Discussions* 1–37. DOI: 10.5194/essd-2019-55.

Limpens J, Berendse F, Blodau C, et al. (2008) Peatlands and the carbon cycle: From local processes to global implications – a synthesis. *Biogeosciences* 5(5): 1475–1491. DOI: 10.5194/bg-5-1475-2008.

Loisel J, Malhotra A and Phillips C (2017) A new platform for managing soil carbon and soil health. *Eos* 98(11). DOI: 10.1029/2017EO080753.

Loranty MM, Abbott BW, Blok D, et al. (2018) Reviews and syntheses: Changing ecosystem influences on soil thermal regimes in northern high-latitude permafrost regions. *Biogeosciences* 15(17): 5287–5313. DOI: 10.5194/bg-15-5287-2018.

Luo Y, Ahlström A, Allison SD, et al. (2016) Toward more realistic projections of soil carbon dynamics by Earth system models. *Global Biogeochemical Cycles* 30(1): 40–56. DOI: 10.1002/2015GB005239.

Malhotra A, Moore TR, Limpens J, et al. (2018a) Post-thaw variability in litter decomposition best explained by microtopography at an ice-rich permafrost peatland.

Arctic, Antarctic, and Alpine Research 50(1): 1–10. DOI: 10.1080/15230430.2017.1415622.

Malhotra A, Sihi D and Iversen C (2018b) The fate of root carbon in soil: Data and model gaps. *Eos* 99. DOI: 10.1029/2018EO112593.

Meehl GA, Stocker TF, Collins WD, et al. (2007) IPCC, 2007: Climate Change 2007: the physical science basis. In: *Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge U.K.: Cambridge University Press, pp. 747–846. DOI: 1452cb7a-9f93-44ea-9ac4-fd9f6fd80a07.

Minasny B, Malone BP, McBratney AB, et al. (2017) Soil carbon 4 per mille. *Geoderma* 292: 59–86. DOI: 10.1016/j.geoderma.2017.01.002.

Najjar RG, Herrmann M, Alexander R, et al. (2018) Carbon budget of tidal wetlands, estuaries, and shelf waters of eastern North America. *Global Biogeochemical Cycles* 32(3): 389–416. DOI: 10.1002/2017GB005790.

Nave L, Johnson K, van Ingen C, et al. (2017) International Soil Carbon Network (ISCN) Database v3-1. DOI: 10.17040/ISCN/1305039.

Oechel WC, Hastings SJ, Vourlrtis G, et al. (1993) Recent change of Arctic tundra ecosystems from a net carbon dioxide sink to a source. *Nature* 361(6412): 520–523. DOI: 10.1038/361520a0.

Peters D, Laney C, Lugo A, et al. (2013) *Long-term trends in ecological systems: A basis for understanding responses to global change*. U.S.D.A. Agricultural Research Service, Technical Bulletin No. 1931.

Poeplau C, Don A, Six J, et al. (2018) Isolating organic carbon fractions with varying turnover rates in temperate agricultural soils – A comprehensive method comparison. *Soil Biology and Biochemistry*. 125(April): 10–26. DOI: 10.1016/j.soilbio.2018.06.025.

Rasmussen C, Heckman K, Wieder WR, et al. (2018) Beyond clay: Towards an improved set of variables for predicting soil organic matter content. *Biogeochemistry* 137(3): 297–306. DOI: 10.1007/s10533-018-0424-3.

Reichman OJ, Jones MB and Schildhauer MP (2011) Challenges and opportunities of open data in ecology. *Science* 331(6018): pp. 703–705.

Ribeiro E, Batjes NH and Van Oostrum AJM (2018) *World Soil Information Service (WoSIS) – Towards the standardization and harmonization of world soil data. Procedures manual 2018*. ISRIC Report 2018/01. Wageningen: ISRIC – World Soil Information. DOI: 10.17027/isric-wdcsoils.20180001.

Sanderman J, Hengl T and Fiske GJ (2018) Soil carbon debt of 12,000 years of human land use. *Proceedings of the National Academy of Sciences of the USA* 115(7): 201800925.

Schuur EAG and Mack MC (2018) Ecological response to permafrost thaw and consequences for local and global ecosystem services. *Annual Review of Ecology, Evolution, and Systematics* 49(1): 279–301. DOI: 10.1145/1999030.1999036.

Schuur EAG, McGuire AD, Schädel C, et al. (2015) Climate change and the permafrost carbon feedback. *Nature* 520(7546): pp. 171–179. DOI: 10.1038/nature14338.

Shaw C, Hilger A, Filiatrault M, et al. (2018) A Canadian upland forest soil profile and carbon stocks database. *Ecology* 99(4): 989. DOI: 10.1002/ecy.2159.

Sierra CA, Hoyt AM, He Y, et al. (2018) Soil organic matter persistence as a stochastic process: Age and transit time distributions of carbon in soils. *Global Biogeochemical Cycles* 32(10): 1574–1588. DOI: 10.1029/2018GB005950.

Sokol NW and Bradford MA (2018) Microbial formation of stable soil carbon is more efficient from belowground than aboveground input. *Nature Geoscience* 12: 46–53.

Sokol NW, Kuebbing SE, Karlsen-Ayala E, et al. (2018) Evidence for the primacy of living root inputs, not root or shoot litter, in forming soil organic carbon. *New Phytologist* 221(1): 233–246. DOI: 10.1111/nph.15361.

Stevens PR and Walker TW (1970) The chronosequence concept and soil formation. *The Quarterly Review of Biology* 45(4): 333–350.

Sulman BN, Phillips RP, Oishi AC, et al. (2014) Microbe-driven turnover offsets mineral-mediated storage of soil carbon under elevated CO_2. *Nature Climate Change* 4(12): 1099–1102. DOI: 10.1038/nclimate2436.

Tarnocai C, Canadell JG, Schuur EAG, et al. (2009) Soil organic carbon pools in the northern circumpolar permafrost region. *Global Biogeochemical Cycles* 23(2): GB2023. DOI: 10.1029/2008GB003327.

Treat CC, Jones MC, Camill P, et al. (2016) Effects of permafrost aggradation on peat properties as determined from a pan-Arctic synthesis of plant macrofossils. *Journal of Geophysical Research: Biogeosciences* 121(1): 78–94. DOI: 10.1002/2015JG003061.

Wardle D, Walker L and Bardgett R (2004) Ecosystem properties and forest decline in contrasting long-term chronosequences. *Science* 305(5683): 509–513.

Weintraub SR, Flores AN, Wieder WR, et al. (2019) Leveraging environmental research and observation networks to advance soil carbon science. *Journal of Geophysical Research: Biogeosciences* 124(5): 1047–1055. DOI: 10.1029/2018JG004956.

Wieder WR, Grandy AS, Kallenbach CM, et al. (2015) Representing life in the Earth system with soil microbial functional traits in the MIMICS model. *Geoscientific Model Development* 8(6): 1789–1808. DOI: 10.5194/gmd-8-1789-2015.

Wilson G, Bryan J, Cranston K, et al. (2017) Good enough practices in scientific computing. *PLoS Computational Biology* 13(6): e1005510. DOI: 10.1371/journal.pcbi.1005510.