Soil sampling and preparation for monitoring soil carbon

Dominique Arrouays[1], Nicolas P.A. Saby[1], Hakima Boukir[1], Claudy Jolivet[1], Céline Ratié[1],
Marion Schrumpf[2], Lutz Merbold[3,4], Bert Gielen[5], Sébastien Gogo[6], Nicolas Delpierre[7],
Gaëlle Vincent[7], Katja Klumpp[8], and Denis Loustau[9]

[1]INRA, InfoSol Unit, 45075 Orléans, France

[2]Max-Planck-Institute for Biogeochemistry, Hans-Knöll-Straße 10, 07745 Jena, Germany

[3]Institute of Agricultural Sciences, ETH Zürich, Universitätstrasse 2, 8092, Zürich, Switzerland

[4]Mazingira Centre, International Livestock Research Institute (ILRI), P.O. Box 30709, 00100, Nairobi, Kenya

[5]Antwerpen University, D.C.219, 2610 Wilrijk, Belgium

[6]Univ Orleans, CNRS, ISTO, UMR 7327, F-45071 Orleans, France

[7]UMR 8079 CNRS / Université Paris Sud, Laboratoire Ecologie, Systématique et Evolution. Avenue du Doyen André Guinier,
 91405 Orsay, France

[8]INRA, UMR IUREP, 63100 Clermont Ferrand, France

[9]INRA, UMR 1391 ISPA, 33140 Villenave d'Ornon, France

Abstract. There is an urgent need for standardized monitoring of existing soil organic carbon stocks in order to accurately quantify potential negative or positive feedbacks with climate change on carbon fluxes. Given the uncertainty of flux measurements at the ecosystem scale, obtaining precise estimates of changes in soil organic carbon stocks is essential to provide an independent assessment of long-term net ecosystem carbon exchange. Here we describe the standard procedure to monitor the soil organic carbon stocks within the footprint of an eddy covariance flux tower, as applied at ecosystem stations of the Integrated Carbon Observation System. The objectives are i) to ensure comparability between sites and to be able to draw general conclusions from the results obtained across many ecosystems and ii) to optimize the sampling design in order to be able to prove changes in time using a reduced number of samples. When sampling a given site at two periods, the objective is generally to assess if changes occurred in time. The changes that can be detected (*i.e.*, demonstrated as statistically significant) depend on several parameters such as the number of samples, the spatial sampling design, and the inherent within-site soil variability. Depending on these parameters, one can define the 'minimum detectable change' which is the minimum value of changed that can be statistically proved. Using simulation studies, we address the trade-off between increasing the number of samples and getting lower minimum detectable changes of soil organic carbon stocks.

Keywords: ICOS protocol, soil organic carbon stocks, sampling design, SOC measurements

INTRODUCTION

Soil organic carbon (SOC) is of high importance for the magnitudes of greenhouse gas fluxes (Lal, 2004; 2010). Small relative changes in SOC pool can significantly increase or decrease atmospheric carbon dioxide (CO_2) content. In many parts of the world, human-induced perturbations have strongly depleted SOC stocks (SOCS) (Bellamy *et al.*, 2005; Sanderman *et al.*, 2017). SOCS are expected to change not only following land use and land management changes (Post and Kwon, 2000; Guo and Gifford, 2002) but also as a result of climate change (temperature and moisture response), increased plant carbon (C) allocation due to elevated atmospheric CO_2 concentrations, long term drainage and N deposition (Gorham, 1991; Stockmann *et al.*, 2013; Dignac *et al.*, 2017). SOCS contain

twice to three times the amounts of C than that is stored in the vegetation or in the atmosphere (Batjes, 1996; Le Quéré *et al.*, 2016; Stockmann *et al.*, 2015). Therefore, a small relative change in this reservoir may have a large effect on CO_2 balance and climate change mitigation (Smith *et al.*, 2008; 2012; Lal, 2016; Paustian *et al.*, 2016; Minasny *et al.*, 2017). There is accordingly an urgent need for standardized monitoring of existing SOCS in order to accurately quantify potential negative or positive feedbacks with climate change on C fluxes. Moreover, SOC is probably one of the most important soil components as it affects soil physical, chemical and biological properties. Changes in SOCS are rather low compared to total SOCS, and most of the measurable changes occur in the upper soil layers (Arrouays and Pélissier, 1994a; Smith, 2004; Saby *et al.*, 2008a).

Soil organic carbon monitoring issues have been highlighted in many papers (*e.g.*, Smith, 2004; Bellamy *et al.*, 2005; Don *et al.*, 2007; Saby *et al.*, 2008a, 2008b; Schrumpf *et al.*, 2011; van Wesemael *et al.*, 2011; Orton *et al.*, 2012a, 2012b; Arrouays *et al.*, 2012). Arrouays *et al.* (2012) distinguished different approaches to the establishment of soil monitoring networks. The authors stressed that it is essential to establish an adequate sampling protocol that can be applied rigorously at each sampling location and time and that different statistical methods should be associated with the different types of sampling design. Among these issues one particular challenge is to get unbiased estimates of mean and total stock values and their uncertainties over specified areas.

Estimations of total SOCS and their changes over a defined area (*i.e.*, an ecosystem station to be monitored) are important variables in the portfolio of the pan-European long-term research infrastructure Integrated Carbon Observation System (ICOS). Given the uncertainty on CO_2 and methane flux measurements at the ecosystem scale (Loescher *et al.*, 2006), obtaining precise estimates of changes in SOCS, along with estimates of changes in biomass and lateral C fluxes, are essential to provide an independent assessment of long-term net ecosystem carbon exchange. All ICOS ecosystem stations should be sampled and treated in a harmonized way, in order to ensure comparability between sites and to be able to draw general conclusions from the results obtained across Europe. Here we describe a standard procedure to monitor the SOCS within the footprint of an Eddy Covariance flux tower, as applied at ecosystem stations of the Integrated Carbon Observation System (ICOS). The ultimate objective is to be able to draw general conclusions from the results obtained across many ecosystems. One other practical objective is to optimize the sampling design in order to be able to prove changes in time using a reduced number of samples. Indeed, the changes that can be detected (*i.e.*, demonstrated as being statistically significant) depend on several parameters such as the number of samples, the spatial sampling design, and the inherent within-site soil

variability. Depending on these parameters, one can define the 'minimum detectable change' which is the minimum value of changed that can be statistically proved. Our main hypothesis is that, using simulation studies, we will be able to address the trade-off between increasing the number of samples and getting lower minimum detectable changes of SOCS.

METHODOLOGY

We first need to consider if the rate of SOCS changes will allow to monitor them in time or if these changes will not be detectable, in which case a single estimate of SOCS at a given date will be enough to characterize the EC footprint.

Should we monitor SOCS or should estimate them only once?

Efficient SOCS monitoring depends on the detectability of changes. Most studies on soil carbon monitoring have shown that changes cannot be significantly proved over time steps less than 10 years (*e.g.*, Smith, 2004; Bellamy *et al.*, 2005; Don *et al.*, 2007; Saby *et al.*, 2008a, 2008b; Schrumpf *et al.*, 2011; van Wesemael *et al.*, 2011; Orton *et al.*, 2012a, 2012b; Arrouays *et al.*, 2012). However, there is an obvious need to detect changes as earlier as possible. This is why the mandatory minimum time step for sampling SOC in ICOS sites has been fixed to 10 years. Samplings in-between may be decided depending on expected changes and available resources. However, temporal changes of the total SOCS over a period of 20 years are unlikely to be detected if they are less than 20-25% of its spatial standard deviation (Smith, 2004; Schrumpf *et al.*, 2011). This is mainly the case for soils where the SOCS is huge as compared with its expected changes over 20 years, as typical for peat soils. This may apply also for the soils having a large spatial heterogeneity due to stoniness (*e.g.* some forests soils on glacial deposits). In such cases, it is important to describe the soil in order to evaluate the feasibility of detecting a temporal change. Subsequently this will also allow to decide whether the standard instructions can be applied or an alternative approach needs to be considered. In any case, a baseline sampling respecting the protocol we detail hereafter, should be conducted during the first campaign to be able to estimate the initial mean value and variance of SOCS variance of an ecosystem under observation.

Spatial sampling strategy

The location of sampling points to monitor the SOCS at the ICOS ecosystem station requires particular attention. The flux tower footprint delineation issue is a crosscutting issue for all observations carried out at ICOS ecosystem monitoring stations and is therefore outside the scope of this paper. Each station should provide the delineation of the area to monitor and the location of places where any soil

or vegetation disturbance should be avoided. The standard procedure proposed here is, in theory, scale independent and can be adopted to any size and shape of the monitored area, on the condition that this area is large enough so as the disturbance created by coring or digging does not impact carbon stock change estimates. Indeed, the choice of a method for monitoring SOC changes should always be a trade-off between sampling effort, minimizing soil disturbance, and statistical power. In other words, and from a practical point of view, this protocol is recommended for sites having an area of several hectares.

Theory of sampling for monitoring

A major decision in designing a monitoring scheme is the selection of the statistical sampling approach (de Gruijter *et al.*, 2006). Two fundamentally different sampling approaches can be distinguished, the design-based and the model-based approach (Brus and de Gruijter, 2011). In a model-based approach, the statistical inference is based on a model of the spatial variation. Randomness is introduced *via* the model, which is a stochastic model, *i.e.* it contains a random error term. As a consequence, selection of the sampling units by probability sampling (*e.g.* by partitioning the area into stratas and sampling random points in each strata) is not strictly needed. This option needs to sample a very large set of locations in order to calibrate a model and is therefore not recommended for our purpose. In a design-based sampling approach, the estimation of the statistical parameter of interest, for instance the mean or the total, and its standard error is based on the inclusion probabilities of the sampling units. These probabilities are determined by the sampling design. This requires that the sampling units are selected by probability sampling, so that these inclusion probabilities are known.

Although it is not always possible to produce a map using probability sampling, it allows estimating the statistical parameters in an efficient way using a small set of samples (Brus and de Gruijter, 2011; de Gruijter *et al.*, 2006). Therefore, because of limited funding for sampling and analyses, and because our main objective is to estimate total SOCS over an area, stratified simple random sampling was chosen as the design-type.

As subjectivity in the spatial stratification used for probability sampling should be avoided, we adopted a stratified random sampling, using geographically compact sub-areas as strata. By using these compact sub-areas as strata, spatial clustering of the sample locations can be avoided and leads subsequently to better accuracy of the estimated spatial mean.

Implementation simulation study

The performance of the selected sampling design and statistical estimates of the parameters are illustrated here by summarizing the results of a simulation study. We have based our study on the results of the first campaign of nine CarboEurope sites, which differ in land use, climate and soil types (Schrumpf *et al.*, 2011). Within this framework, a first extensive soil inventory was conducted at all sites as a baseline for long-term soil carbon monitoring. Concentrations of organic carbon (OC), bulk density (BD), and fine earth fraction were determined from 0 to 60 cm depth at 100 sampling points per site (the depth of sampling resulted from a trade-off between costs and expected future changes). The 100 sampling points per site formed a regular grid at distances of 10 to 15 m at the grassland and cropland sites, and 30 m at the forest sites (Schrumpf *et al.*, 2011). Using these data and geostatistical technics, we simulated SOCS maps at the ICOS ecosystem station for two dates. Using these results, we realized 100 virtual samplings of increasing number of cores. Using this method, we investigated the minimum detectable changes in SOCS, according to compact strata probability sampling, assuming future re-sampling of a various set of cores and a constant spatial variance in time (*e.g.* same variogram for the two sampling rounds). As a correlation between sampling times is expected, but unknown, we did the simulations for two extreme values of this correlation, 0.2 and 0.8.

Protocol for spatial sampling

The target area to be monitored is partitioned into 10 compact strata of equal area and two first order sparse measurement points named SP-I are "randomly" selected per stratum.

From a practical point of view, purely theoretical random sampling is not always possible, as we need to avoid that points are too close each other. Therefore we use a default random sampling strategy and define a circle with a 10 m radius around each point. In practice this means that first one point is randomly selected per stratum. Then a second point is randomly selected in each stratum but being at least 30 m apart from the first one and a circle of 10 m is also defined around this second point. These circles are defined as subplots. The radius of the circles will be adjusted in case the stratum is not large enough to identify two subplots without overlap. In any case it should not be less than 5 m. The area for soil sampling is then defined by these 20 sub-plots (two circles in each of the 10 strata). In each of these subplots, five second order sparse measurement points are selected "randomly" for soil sampling (see figures in Saunders *et al.*, 2018). Samples from these five second order points will be then bulked. The advantage of this kind of bulk (or 'aggregate', or 'composite') sampling strategy has been demonstrated and discussed by Lark (2012). Interestingly, Lark (2012) also showed that the benefits of increasing the number of aliquots within such a small area beyond five were negligible. As sampling provokes some perturbations, it is therefore good practice to limit the number of subsamples to 5 in case of coring and to 3 in case of digging pits.

As for SP-I definition, the SP-II location within circular subplots is not strictly random, *i.e.*, SP-II should be at least located 2 m apart from each other in order to make sampling possible without disturbing adjacent points. In practice this means that first one SP-II is randomly selected per subplot. Then a second point is randomly selected in each subplot but being at least 2 m apart from the first one, etc. Supplementary reserve locations are provided in case of impossibility of sampling (*e.g.*, point under a trunk). Core samples of known volume are taken at the same fixed depths on these five points (coring protocol is described in the section 'Field sampling'). If digging soil pits is necessary because coring is impossible (*e.g.*, stony soils), then the number of points is reduced to 3 in order to minimise site disturbance and the minimum distance between SP-II is increased to 3 m.

Resampling

The sampling strategy for the successive sampling campaigns will use the same strata and subplots, but new SP-II will be selected within the circular areas at each campaign, using the same rule and adding a new minimum distance constraint of 2 m from the coring points (or 3 m in case of digging pits) of the previous sampling campaigns. In practice, this means that the area to exclude from sampling increases with time.

Total sampling depth and sampling depth increments

Defining a mandatory soil thickness to sample, results from a trade-off between measuring all SOC, being able to detect changes, and minimizing the sampling effort. Here, SOCS are calculated for the first meter (0 to 100 cm) as this thickness enables to capture most of the SOC subject to detectable changes over decadal periods. The SOCS is expressed in gC per m² for a given depth and layer. These depths are: 0-5, 5-15, 15-30, 30-60, 60-100 cm from the surface of mineral soils. Using smaller sampling increments close to the soil surface is recommended because SOC often exhibit a strong gradient close to the surface (Arrouays and Pélissier, 1994b) and larger changes are expected in the upper part of the soil. There are many good reasons to focus on topsoil layers and to adopt fixed increments (Kibblewhite *et al.*, 2008; Morvan *et al.*, 2008) rather than pedogenic horizons,*e.g.* to avoid subjectivity and to allow comparison between sites. Moreover, there may be different successions of horizons within a single site, hampering the formulations of averages. Another rationale is that it is better to have fixed increments for data assimilation in models that will be run for all ICOS ecosystem stations. Finally, the depth increments we recommend are consistent with the increments adopted by the international GlobalSoilMap initiative (Arrouays *et al.*, 2014).

Organic surface layers (O horizons) are characterized by a much higher turn-over rate of C than mineral layers. Therefore it is mandatory to sample them separately.

Moreover, they are also characterized by a higher short-range heterogeneity than mineral layers. Therefore, when present, organic surface layers will be sampled first and separately using a frame (25x25 cm) and their depth and mass will be recorded. Deeper mineral layers will be then sampled as described above.

Field sampling
Sampling conditions

Soil sampling should be carried out at a time when the humidity is close to the field capacity and in a stabilized ground, not recently ploughed. The ground water table should be preferably below the soil depth to be sampled. In croplands, the sampling must be carried out at least four weeks after the last ploughing at a time when the upper soil layers are firm and wet and allow for using a soil auger or automated corer. In addition, periods following cultivation operations (tillage, disking, seeding, fertilizers application *etc.*) must be avoided.

Protocol for field sampling

SOCS are measured using simultaneous measurements of SOC content and bulk density (BD). This simultaneous measurement enables to avoid biases linked to the correlation between SOC content and BD. The International Standard ISO-11279:1998 describes three methods for the determination of dry BD of soils calculated from the mass and the volume of a soil sample. The methods involve drying and weighing of soil samples, the volume of which is either known (core method) or has to be determined (excavation method, and clod method). As the clod method does not integrate some voids in soil, it is not adapted to our aim of estimating a total stock on a given volume. Thus, for the presented standard procedure we recommend either the core method (non- stony soils) or the excavation method (stony soils):

(i) In non-stony coherent soils, samples have to be taken as vertical undisturbed cores that are extracted with a motor hammer corer or a root corer. All five sampling points within each sub-plot (n = 20 sub-plots × 5 sampling points) have to be sampled. The core samples of known volume will be extracted from the depths defined above.

(ii) In stony or incoherent soils, only the first three SP-II locations around each SP-I have to be sampled (n = 20 × 3 sampling points). At each SP-II location, a soil pit is dug and individual volumetric samples are extracted in the pit from the layer depths defined above. Because volumetric samples thicker than 10 cm are practically difficult to extract in one piece, the layers 30-60 and 60-100 cm are sampled by means of two superposed samples of equal volume. The individual volumetric samples are taken either by the cylinder method or the excavation method according to the International Standard ISO 11272:1998. For the cylinder method, the volume of the cylinders should preferably be close to 500 cm³ and never less than 100 cm³.

For sampling organic soil horizons the upper surface of the soil at the interface between "O" horizons and the upper mineral horizon must be carefully identified. Separating "O" horizons from mineral horizons is not trivial and needs the expertise of a well trained soil scientist (Federer, 1983). Then, at each individual sampling location, the organic horizons (O) must be collected using a frame (25x25 cm²) and separated between the Oi, Oe, and Oa layers. Their thickness will be measured in order to allow volume calculation. Then, soil cores will be taken from sampling depths 0-5, 5-15, 15-30, 30-60, 60-100 cm for organo-mineral and mineral horizons. Whatever technique used, it is essential to obtain for each sample collected i) the sample volume, V in cm³, ii) the sample vertical thickness, LT in cm, and iii) the sample horizontal cross sectional area, A in cm². Cylinder volumes should be comprised between 100 and 500 cm³.

Cores should be about 10 cm diameter. If digging pits is necessary, their size has to be adapted to local conditions. In case of impossibility to achieve core sampling (*i.e.*, randomly distributed stones in depth), it is recommended to try to reach at least 30 cm by resampling one meter apart from the theoretical point. In case 30 cm cannot be reached again, the deepest sampling point will be retained, and the soil depth will be considered to have been reached. For cores deeper than 30 cm, any impossibility to dig down to a depth less than 1 m will define the soil depth. In case the soil depth is in between the standard depth values (*e.g.* 55 cm) the layer between the upper standard depth and the soil depth will be taken (*e.g.*, 30 - 55 cm). Pits should be refilled after digging, respecting the vertical order of layers.

Specifications for peat soils

Acknowledging that (i) the assessment of temporal changes in the total SOCS in virgin mires and peats using repeated measurements based on cores is not achievable and (ii) in some cases, the C export by lateral fluxes is a major component of C balance of mires, it is accepted not to apply the sampling scheme and SOCS determination as proposed in the standard procedure. An alternative method is proposed, especially in actively drained peatlands and shallow peat soils where temporal changes might be significant and detectable (Krüger *et al.*, 2015). Since the main objective of the soil sampling strategy is to close the ecosystem mass balance of C, the alternative method proposed should consider also to determine the discharge of C.

Nevertheless, it is compulsory (ecosystem stations with peat soils) to sample peat soils down to 1 m depth at a total of five to ten points split over ten plots in order to provide an initial estimate of the total SOCS.

An auger, that prevents compaction has to be used for peat sampling. The top of the capitula of the *Sphagnum* mosses, when present, are used as the reference (z = 0). To be able to highlight a change in SOC over time, it may be better to focus on detailed profiles with depth (*e.g.* one sample every 0.1 m), at the expense of replicate spatial sampling (*e.g.*, 1 peat core per plot).

Soil description

The first core or pit of each subplot should be described and recorded by a soil scientist or an adequately trained staff. The following data must be reported:

− Soil type following WRB classification (FAO, 2014; IUSS Working Group WRB., 2015): (*e.g.*, Albic Stagnic Luvisol (Cutanic, Endoclayic)) (http://www.fao.org/3/i3794en/I3794en.pdf)

− Horizons description: each horizon of the profile must be described using the following variables:

- Upper and lower vertical limits of the horizon
- Color of the horizon using the Munsell color code chart
- Soil water state, estimated according to 5 classes (dry, moist, wet, wet-satiated, saturated; *e.g.* NRCS-USDA, 2012)
- Texture of the horizon using field finger test
- The presence of stones and their nature
- Diagnostic horizon name (if applicable) according to the FAO WRB (FAO, 2014, http://www.fao.org/3/i3794en/I3794en.pdf)
- Diagnostic parent material name (if applicable) according to the FAO WRB (FAO, 2014, http://www.fao.org/3/i3794en/I3794en.pdf)
- Diagnostic property name (if applicable) according to the FAO WRB (FAO, 2014, http://www.fao.org/3/i3794en/I3794en.pdf).

Sample processing

All sample processing steps included in this section will be done by each team at the ICOS station level. The sample processing describes the operating modes to obtain the composite samples that will be sent to chemical analysis and archiving; it is including the measurements for bulk density determination as well. The successive steps are summarized in Fig. 1. The following operating modes are adapted from the International Standard ISO-11074-2:1998. One should refer to this document for further details.

Mineral soil samples

The mineral soil samples preparation and processing involve three main steps:

(i) Drying and sieving each individual sample to obtain fine earth, stones and root mass values.

(ii) Combining fine earth of individual samples to obtain a composite sample: for each subplot, the fine earth of each depth interval of the five cores or three pits will be bulked in proportions corresponding to their respective mass of fine earth from the corresponding layers. For core samples, this bulking is simply done by mixing the samples. However, for samples obtained by excavation it is
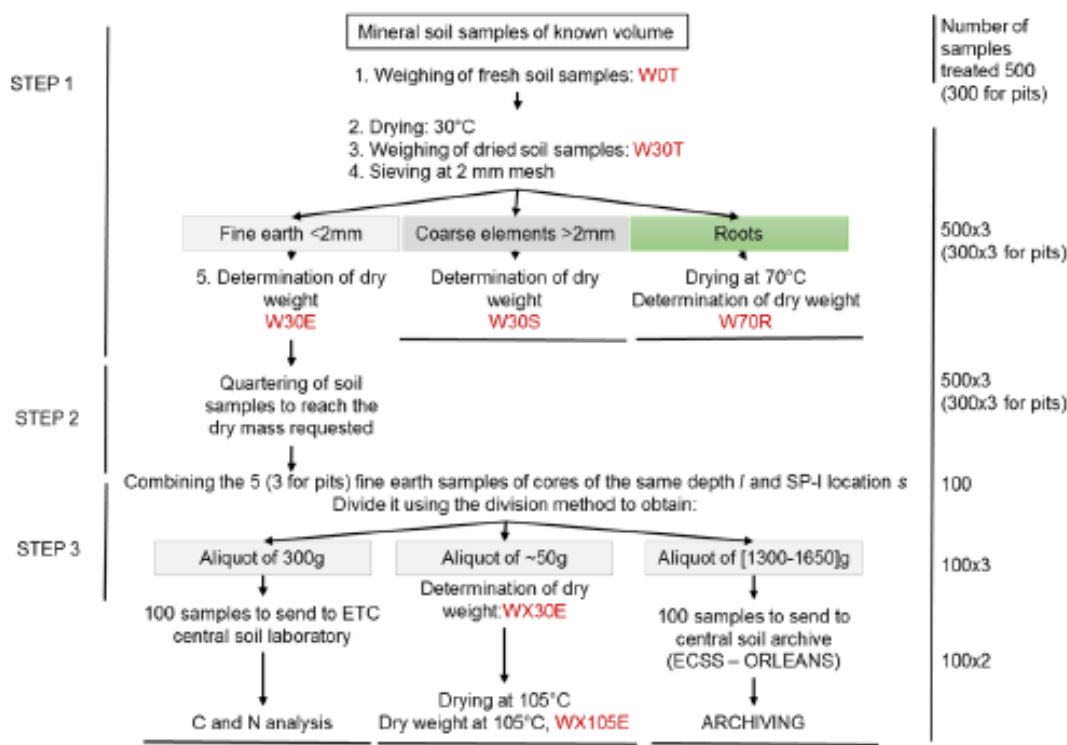
Fig. 1. Workflow for sample preparation and analyses- mineral soil samples. Steps should be done according to the numbering order. Red codes correspond to the codes to be used when reporting.

necessary to use bulk density values to calculate the mass of each SP-II sample to be bulked. An aliquot will be sent to the Ecosystem Thematic Centre (ETC) central laboratory.

(iii) Extracting three sub-samples from the composite samples. The obtained composite samples (one for each subplot location and layer) must be split in three aliquots: 1) a 50±10 g aliquot for determination of the residual humidity and the fine earth mass; 2) a 300±30 g aliquot that must be send to the ETC central laboratory for C and residual water content analyses; 3) the remaining sample for long-term archiving.

Organic samples

The organic samples preparation and processing can be divided in two main steps that are summarized in Fig. 2.

(i) Drying and crushing samples

(ii) Composite sample following the same method described for organo-mineral and mineral samples and extract sub-samples: 1) a 300 g aliquot that must be send to the ETC Central Laboratory for C, N and residual water content analyses, 2) The remaining sample to be archived.

Analytical measurements

The analytical measurements will be performed by the ETC central laboratory. Although various analytical methods can be used to estimate SOC content, including loss on ignition, wet oxidation and dry combustion, for ICOS

we recommend the commonly accepted standard that is dry combustion using an automated C-N analyzer. Total C content in soil will be determined by dry combustion (ISO, 1995). If $CaCO_3 < 699$ g kg^{-1}, SOC will be deduced by subtraction of carbonate to total C content using the following equation.

$$SOC = total\ carbon - 0.12\ CaCO_3, \qquad (1)$$

where *SOC* is expressed in g kg^{-1}.

The carbonate content will be measured according to ISO 10693 (ISO, 2014): an aliquot of the soil sample is treated with hydrochloric acid. The volume of CO_2 produced is measured using a volume Scheibler aparatus and converted into $CaCO_3$ equivalent concentration.

If $CaCO_3 > 699$ g kg^{-1}, SOC will be determined directly using an automated C-N analyzer on a test portion of the soil sample pretreated by HCl (removal of carbonates). The threshold of 699 g kg^{-1} corresponds to the value of $CaCO_3$ where the error on $CaCO_3$ propagates significant errors on the estimation of SOC using Eq. (1).

Calculations

Calculating SOCS for the subplot

The mass of fine soil in g m^{-2} will be calculated for each SP-II layer, using BD and the thickness of the standard layer or of the observed thickness in case it is less than the standard. The total mass of fine soil in each layer in g m^{-2}
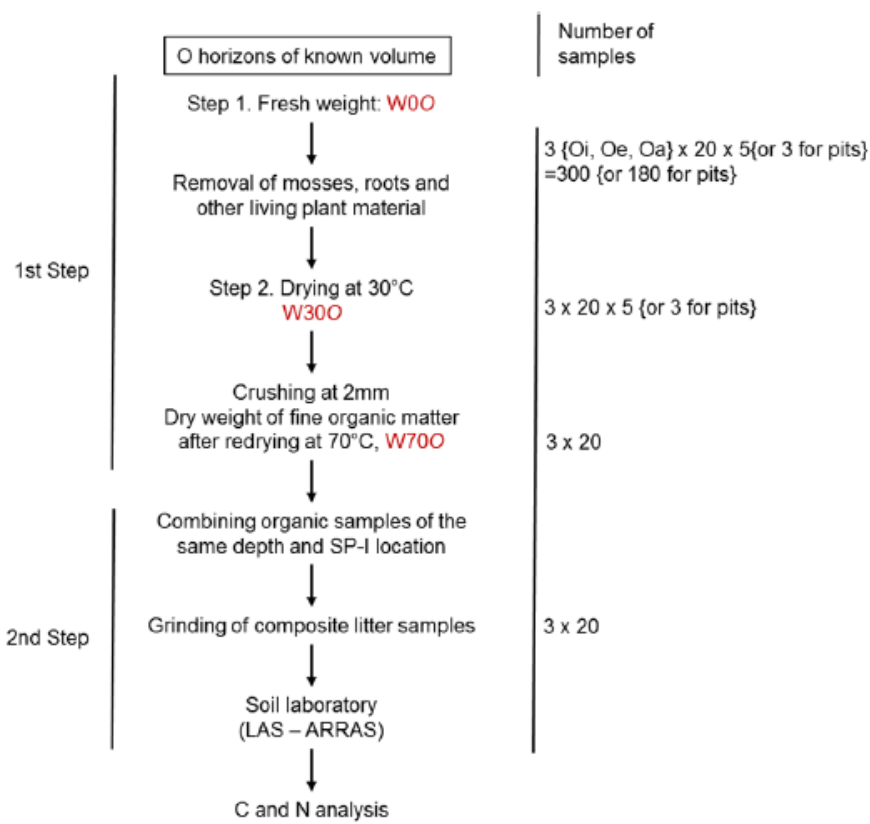
Fig. 2. Workflow for sample preparation and analyses – organic soil samples. Steps should be done according to the numbering order. Red codes correspond to the codes to be used when reporting. Three distinct organic layers may be present: one of leaves, pine needles and twigs (Oi); underlain by a partially decomposed layer (Oe); and then a very dark layer of well decomposed humus (Oa).

of each subplot will then be calculated by averaging the masses measured on the 5 sampling points (or 3 in case of pit excavation).

The estimations of the total and the mean of the SOCS of each layer of each subplot will be then calculated when the C content of the dried bulked samples is determined.

Estimation of spatial means and sampling variance

The estimation method uses the well-known formula of the stratified random sampling. The total stock z_i of the layer i (g m^{-2}) is estimated by:

$$\hat{z}_i = \sum_{h=1}^{H} a_h \hat{z}_{ih}, \tag{2}$$

where: H is the number of strata (10 in this case), a_h are the relative areas of the strata (stratum weights equal to 1/10 in this case), and $\hat{z}_{ih}$ is the estimated total stock of stratum h estimated by the following equation:

$$\hat{z}_{ih} = \sum_{j=1}^{n_h} z_{ihj}, \tag{3}$$

where: n_h equals to the number of sampling locations in stratum h (2 in this case), z_{ihj} is the observation j of the layer i of the stratum h.

The mean stock $\hat{\bar{z}}_i$ of the layer i is estimated by:

$$\hat{\bar{z}}_i = \sum_{h=1}^{H} a_h \hat{\bar{z}}_{ih}, \tag{4}$$

where: $\hat{\bar{z}}_{ih}$ is the estimated mean of stratum h estimated by the following equation:

$$\hat{\bar{z}}_{ih} = \frac{1}{n_h} \sum_{j=1}^{n_h} z_{ihj}. \tag{5}$$

Provided all numbers of sampling locations are larger than 1 in the strata, the sampling variance of can be estimated by:

$$\hat{V}(\hat{\bar{z}}_i) = \sum_{h=1}^{H} a_h^2 \hat{V}(\hat{\bar{z}}_{ih}), \tag{6}$$

where: $\hat{V}(\hat{\bar{z}}_{ih})$ is the estimated sampling variance of $\hat{z}_{ih}$ (the estimated mean of the stratum h):

$$\hat{V}(\hat{\bar{z}}_{ih}) = \frac{S_{ih}^2(z_i)}{n_h}, \tag{7}$$

with $S_{ih}^2(z)$ the spatial variance of z_i within stratum h, that is estimated by:

$$\widehat{S^2}_h(z_i) = \frac{1}{(n_h-1)} \sum_{j=1}^{n_h} (z_{ihj} - \hat{\bar{z}}_{ih})^2. \tag{8}$$

RESULTS AND DISCUSSION

The performance of the selected sampling design and statistical estimates of the parameters are illustrated here by summarizing the results of the simulation study based on the results of the first campaign of nine CarboEurope sites

(Schrumpf *et al.*, 2011). The results from our simulation study showed that increasing the number of plots decreased on average the mean detectable changes. As a correlation between sampling times is expected, but unknown, we did the simulations for two extreme values of this expected correlation, 0.2 and 0.8 (Figs 3 and 4, respectively).

Note that the correlation between sampling times has a strong effect on the MDC. The mean values of MDC roughly double when the correlation between sampling times is changed from 0.8 to 0.2.

For both cases, when the number of plots increases the minimum detectable change decreases sharply whereas this decrease slows down when the number of plots reaches 20 and more. This result confirms the choice of 20 plots as an acceptable trade-off between the possibility to detect a change and the sampling effort. Indeed, depending on the number of plots, the soil thickness investigated, and the correlation between sampling times the mean MDC range $1.5 - 0.05$ kg m^{-2}.

The results also show a large variability of the minimum detectable changes, which shows that, depending on the within-site variance and the rates of change, the same sampling effort may result in very different conclusions among

the ICOS ecosystem stations monitored. Conversely, this protocol enables to estimate the within-site variance for all ICOS ecosystem stations monitored. Therefore, assuming possible rates of changes it will be possible to estimate how the sampling design should be densified (by increasing the number of SP-I points) to reach a desired minimum detectable change. In other words, applying this protocol will help to decide if the sampling effort should be modulated. However, the choices we made (bulking 5 samples for cores and 3 for pits) are consistent with previous studies that showed that bulking a small number of samples may be as efficient as measuring SOC on individual cores (*e.g.*, Lark, 2012). Moreover, using this strategy reduces the analytical cost and avoids large soil perturbations.

The adoption of systematic depths is mandatory, in order to avoid subjectivity in sampling, harmonise sampling protocols, and facilitate comparisons between sites. However, assessing changes between two sampling dates, will require to take into account changes in bulk density which may happen. These changes in bulk density may happen because of natural causes, site management practices or even because of the sampling itself. The method we propose, including measurement of thickness, soil
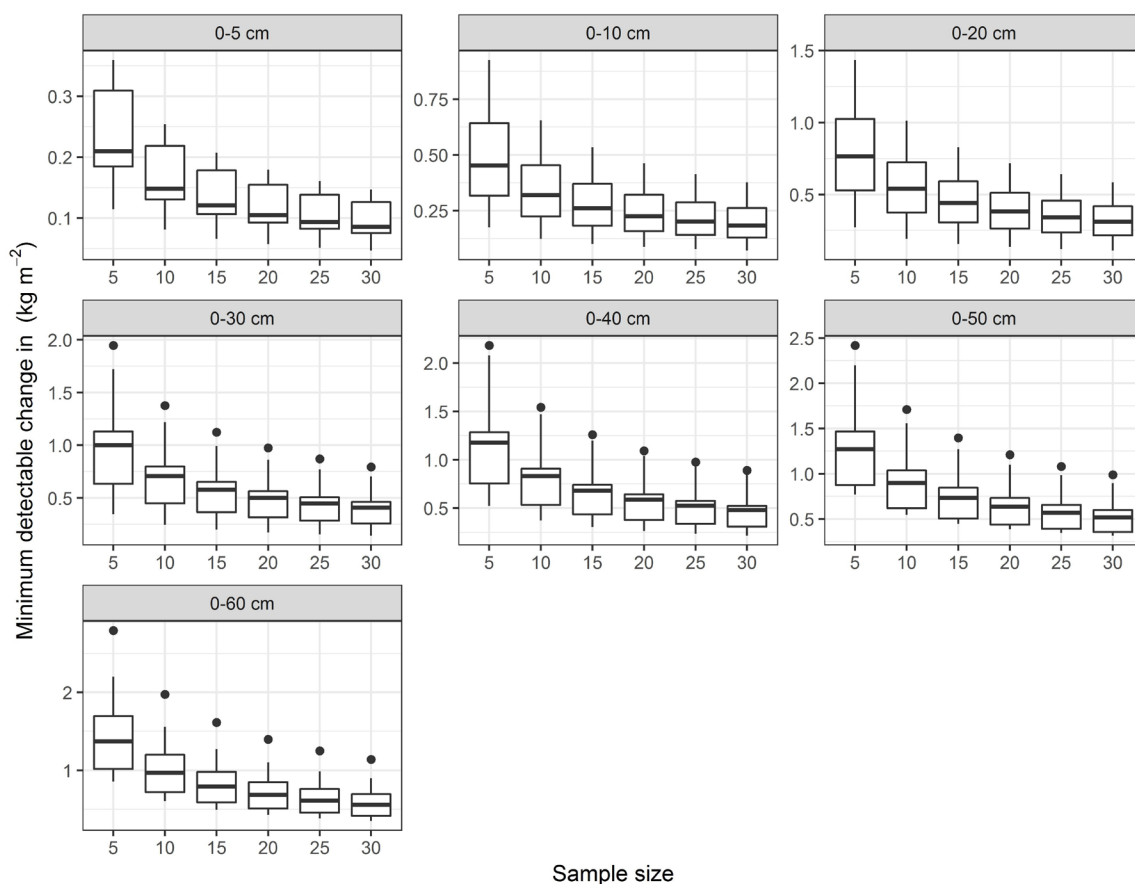

Fig. 3. Box-plots of the simulated minimum detectable changes over nine Carbo-Europe stations. One hundred simulations were run on each of the stations and each soil thickness (correlation between sampling times is fixed to 0.2). Dots are outlier values. The x-axis corresponds to the number of samples.

Fig. 4. Box-plots of the simulated minimum detectable changes over nine Carbo-Europe stations. One hundred simulations were run on each of the stations and each soil thickness (correlation between sampling times is fixed to 0.8). Dots are outlier values. The x-axis corresponds to the number of samples.
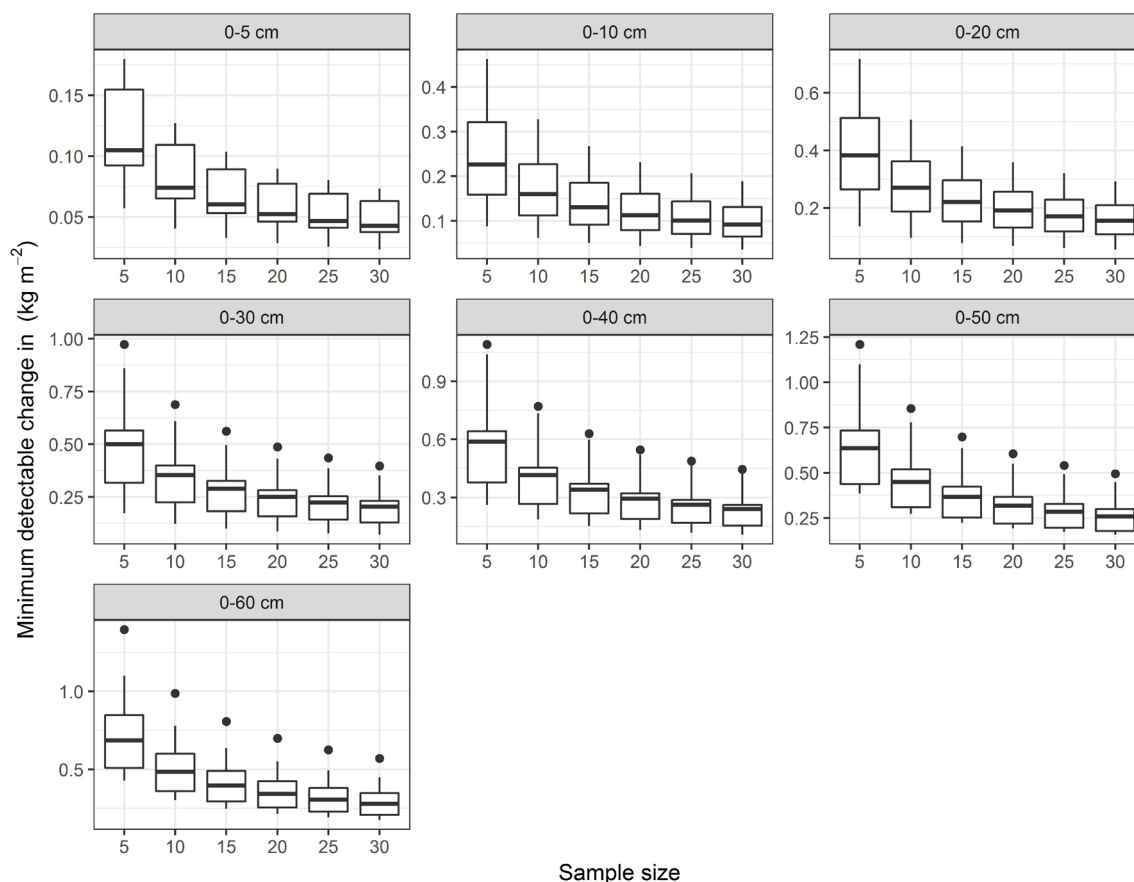
bulk density and bulking according to soil masses, enables to apply the 'equivalent soil mass' procedure (Ellert and Bettany, 1995; Wendt and Hauser, 2013).

We should also note that the MDC is increasing with the total depth taken into account, therefore we recommend that calculations about changes should be done for all thicknesses (*e.g.* 0-5, 0-15, 0-30 cm, *etc.*) because calculating changes on the whole first meter only may mask changes occurring at the topsoil layers.

Finally, the method we propose is adapted to ICOS sites that have generally an area of several hectares. One should keep in mind that sampling smaller areas will need a very carefull trade-off between the target minimum detectable change and a minimum site perturbation.

CONCLUSIONS

1. We provide standardised guidelines for sampling and measuring SOCS to ensure comparability between sites and to be able to draw general conclusions from the results obtained across many ecosystems.

2. Using simulation studies, we addressed the trade-off between increasing the number of samples and getting lower mean detectable changes of SOCS. We show that using design based sampling, a significant reduction of uncertainty on SOCS can be achieved when gathering 20 composite samples taken on the basis of 10 compact strata.

3. The protocol enables a first estimate of the within-site variance which can be further used to adapt the sampling effort to required minimum detectable changes.

4. The standardisation of methodologies to characterise SOCS in ecosystem stations are essential to develop a coherent pan-European flux network.

Conflict of interest: The Authors declare no conflict of interest.

REFERENCES

Arrouays D., Grundy M.G., Hartemink A.E., Hempel J.W., Heuvelink G.B.M., Hong S.Y., Lagacherie P., Lelyk G., McBratney A.B., McKenzie N.J., Mendonça-Santos M.D., Minasny B., Montanarella L., Odeh I.O.A., Sanchez P.A., Thompson J.A., and Zhang G.-L., 2014. GlobalSoilMap: towards a fine-resolution global grid of soil properties. Advances in Agronomy, 125, 93-134.

Arrouays D., Marchant B.P., Saby N.P.A., Meersmans J., Orton T.G., Martin M.P., Bellamy P.H., Lark R.M., and Kibblewhite M., 2012. Generic issues on broad scale soil monitoring schemes: A review. Pedosphere, 22, 456-469.

Arrouays D. and Pélissier P., 1994a. Changes in carbon storage in temperate humic loamy soils after forest clearing and continuous corn cropping in France. Plant Soil, 160, 215-223.

Arrouays D. and Pélissier P., 1994b. Modeling carbon storage profiles in temperate forest humic loamy soils of France. Soil Science, 157, 185-192.

Batjes N.H., 1996. Total carbon and nitrogen in the soils of the world. Eur. J. Soil Sci., 47, 151-163.

Bellamy P.H., Loveland P.J., Bradley R.I., Lark R.M., and Kirk G.J.D., 2005. Carbon losses from all soils across England and Wales 1978-2003. Nature, 437, 245-248.

Brus D.J. and de Gruijter J.J., 2011. Design-based Generalized Least Squares estimation of status and trend of soil properties from monitoring data. Geoderma, 164, 172-180.

de Gruijter J., Brus D.J., Bierkens M.F.P., and Knotters M., 2006. Sampling for Natural Resource Monitoring. Springer-Verlag.

Dignac M.F., Derrien D., Barré P., Barot S., Cécillon L., Chenu C., Chevallier T., Freschet G.T., Garnier P., Guenet B., and Hedde M., 2017. Increasing soil carbon storage: mechanisms, effects of agricultural practices and proxies. A review. Agronomy for Sustainable Development, 37, 14.

Don A., Schumacher J., Scherer-Lorenzen M., Scholten T., and Schultze E-D., 2007. Spatial and vertical variation of soil carbon at two grassland sites – Implications for measuring soil carbon stocks. Geoderma, 141, 272-282.

Ellert B.H. and Bettany J.R., 1995. Calculation of organic matter and nutrients stored in soils under contrasting management regimes. Can. J. Soil Sci., 75, 529-38.

FAO, **2014.** World reference base for soil resources. International soil classification system for naming soils and creating legends for soil maps. FAO, Roma, Italy.

Federer C.A., 1983. Nitrogen mineralization and nitrification: Depth variation in four New England forest soils. Soil Sci. Soc. Am. J., 47, 1008-1014.

Gorham E., 1991. Northern peatlands: role in the carbon cycle and probable response to global warming. Ecological Appl., 1(2), 182-195.

Guo L.B. and Gifford R., 2002. Soil carbon stocks and land use change: A meta analysis. Global Change Biology, 8, 345-360.

International Standard ISO-11074-2:**1998** (Soil quality – Vocabulary – Part 2: terms and definitions relating to sampling).

ISO, **1995.** Soil quality - Determination of organic and total carbon after dry combustion (elementary analysis).

ISO, **1998.** ISO 11279. Soil quality – Determination of dry bulk density.

ISO, **1998.** ISO 11272. Soil quality – Determination of dry bulk density.

ISO, **2014.** ISO 10693. Soil quality – Determination of carbonate content – Volumetric method.

IUSS Working Group WRB, **2015.** World Reference Base for Soil Resources 2014, update 2015, International soil classification system for naming soils and creating legends for soil maps. World Soil Resources Reports No. 106. FAO, Rome.

Kibblewhite M.G., Jones R.J.A., Montanarella L., Baritz R., Huber S., Arrouays D., Micheli E., and Stephens M., (Eds**). 2008.** Environmental Assessment of Soil for Monitoring: Volume VI Soil Monitoring System for Europe. EUR 23490 EN/6, Office for the Official Publications of the European Communities, Luxembourg, 188pp.

Krüger J.-P., Alewell C., Minkkinen K., Szidat S., and Leifeld J., 2015. Calculating carbon changes in peat soils drained for forestry with four different profile-based methods. For. Ecol. Manag., 381, 29-36, 10.1016/j.foreco.2016.09.006.

Lal R., 2004. Soil carbon sequestration impacts on global climate change and food security. Science, 304, 1623-1627.

Lal R., 2010. Managing soils and ecosystems for mitigating anthropogenic carbon emissions and advancing global food security. BioScience, 60, 708-721.

Lal R., 2016. Beyond COP 21: potential and challenges of the "4 per Thousand" initiative. J. Soil Water Conservation, 71, 20A-25A.

Lark R.M., 2012. Some considerations on aggregate sample supports for soil inventory and monitoring. European J. Soil Science, 63, 86-95.

Le Quéré C., Andrew R.M., Canadell J.G., Sitch S., Ivar Korsbakken J., Peters G.P., Manning A.C., Boden T.A., Tans P.P., Houghton R.A., Keeling R.F., Alin S., Andrews O.D., Anthoni P., Barbero L., Bopp L., Chevallier F., Chini L.P., Ciais P., Currie K., Delire C., Doney S.C., Friedlingstein P., Gkritzalis T., Harris I., Hauck J., Haverd V., Hoppema M., Klein Goldewijk K., Jain A.K., Kato E., Körtzinger A., Landschützer P., Lefèvre N., Lenton A., Lienert S., Lombardozzi D., Melton J.R., Metzl N., Millero F., Monteiro P.M.S., Munro D.R., Nabel J.E.M.S., Nakaoka S.-I., O'Brien K., Olsen A., Omar A.M., Ono T., Pierrot D., Poulter B., Rödenbeck C., Salisbury J., Schuster U., Schwinger J., Séférian R., Skjelvan I., Stocker B.D., Sutton A.J., Takahashi T., Tian H., Tilbrook B., Van Der Laan-Luijkx I.T., Van Der Werf G.R., Viovy N., Walker A.P., Wiltshire A.J., and Zaehle S., 2016. Global Carbon Budget 2016. Earth System Science Data, 8, 605-649.

Loescher H.W., Law B.E., Mahrt L., Hollinger D.Y., Campbell J., and Wofsy S.C., 2006. Uncertainties in, and interpretation of, carbon flux estimates using the eddy covariance technique. J. Geophysical Res. Atmospheres, 11(D121), DOI: 10.1029/2005JD006932.

Minasny B., Malone B.P., McBratney A.B., Angers D.A., Arrouays D., Chambers A., Chaplot V., Chen Z.S., Cheng K., Das B.S., Field D.J., Gimona A., Hedley C.B., Hong S.Y., Mandal B., Marchant B.P., Martin M.P., McConkey B.G., Mulder V.L., O'Rourke S., Richer-de-Forges A.C., Odeh I., Padarian J., Paustian K., Pan G., Poggio L., Savin I., Stolbovoy V., Stockmann U., Sulaeman Y., Tsui C-C., Vågen T.-G., van Wesemael B., and Winowiecki L., 2017. Soil carbon 4 per mille. Geoderma, 292, 59-86.

Morvan X., Saby N.P.A., Arrouays D., Le Bas C., Jones R.J.A., Verheijen F.G.A., Bellamy P.H., Stephens M., and Kibblewhite M.G., 2008. Soil monitoring in Europe: A review of existing systems and requirements for harmonization. Sci. Total Environ., 391, 1-12.

NRCS-USDA, **2012.** Field Book for Describing and Sampling Soils. Version 3.0. National Soil Survey Center. Natural Resources Conservation Service. U.S. Department of Agriculture, September 2012.

Orton T., Saby N.P.A., Arrouays D., Walter C., Lemercier B., Schvartz C., and Lark R.M., 2012a. Spatial prediction of soil organic carbon from data on large and variable spatial supports. I. Inventory and mapping. Environmetrics, 23, 129-147.

Orton T., Saby N.P.A., Arrouays D., Walter C., Lemercier B., Schvartz C., and Lark R.M., 2012b. Spatial prediction of soil organic carbon from data on large and variable spatial supports. II. Mapping temporal change. Environmetrics, 23, 148-161.

Paustian K., Lehmann J., Ogle S., Reay D., Robertson G.P., and Smith P., 2016. Climate-smart soils. Nature, 532, 49-57.

Post W.M. and Kwon K.C., 2000. Soil carbon sequestration and land-use change: Processes and potential. Global Change Biology, 6, 317-327.

Saby N.P.A., Arrouays D., Antoni V., Lemercier B., Follain S., Walter C., and Schvartz C., 2008a. Changes in soil organic carbon in a mountainous French region, 1990-2004. Soil Use Manage., 24, 254-262.

Saby N.P.A., Bellamy P.H., Morvan X., Arrouays D., Jones R.J.A., Verheijen F.G.A., Kibblewhite M.G., Verdoodt A., Berényiüveges J., Freudenschuß A., and Simota C., 2008b. Will European soil-monitoring networks be able to detect changes in topsoil organic carbon content? Glob. Change Biol., 14, 2432-2442.

Sanderman J., Hengl T., and Fiske G.J., 2017. Soil carbon debt of 12,000 years of human land use. P.N.A.S., 114, 9575-9580.

Saunders M., Dengel S., Kolari P., *et al*., 2018. Assessing the impacts of site characteristics, management, disturbance and lateral fluxes on greenhouse gas dynamics. Int. Agrophys., 32, 457-469.

Schrumpf M., Schulze E.D., Kaiser K., and Schumacher J., 2011. How accurately can soil organic carbon stocks and stock changes be quantified by soil inventories? Biogeosciences, 8(5), 1193-1212.

Smith P., 2004. How long before a change in soil organic carbon can be detected? Global Change Biology, 10, 1878-1883.

Smith P., Davies C.A., Ogle S., Zanchi G., Bellarby J., Bird N., Boddey R.M., McNamara N.P., Powlson D., Cowie A., and Noordwijk M., 2012. Towards an integrated global framework to assess the impacts of land use and management change on soil carbon: current capability and future vision. Global Change Biology, 18, 2089-2101.

Smith P., Martino D., Cai Z., Gwary D., Janzen H., Kumar P., McCarl B., Ogle S., O'Mara F., Rice C., Scholes B., Sirotenko O., Howden M., McAlliste T., Pan, G., Romanenkov V., Schneider U., Towprayoon S., Wattenbach M., and Smith J., 2008. Greenhouse gas mitigation in agriculture. Philosophical Trans. Royal Society B, 363, 789-813.

Stockmann U., Adams M.A., Crawford J.W., Field D.J., Henakaarchchi N., Jenkins M., Minasny B., McBratney A.B., de Remy de Courcelles V., Singh K., Wheeler I., Abbott L., Angers D.A., Baldock J., Bird M., Brookes P.C., Chenu C., Jastrow J.D., and Zimmermann M., 2013. The knowns, known unknowns and unknowns of sequestration of soil organic carbon. Agric., Ecosys. Environ., 164, 80-99.

Stockmann U., Padarian J., McBratney A.B., Minasny B., de Brogniez D., Montanarella L., Hong S.Y., Rawlins B.G., and Field D.J., 2015. Global soil organic carbon assessment. Global Food Security, 6, 9-16.

van Wesemael B., Paustian K., Andrén O., Cerri C.E.P., Dodd M., Etchevers J., Goidts E., Grace P., Kätterer T., McConkey B.G., Ogle S., Pan G.X., and Siebner C., 2011. How can soil monitoring networks be used to improve predictions of organic carbon pool dynamics and CO_2 fluxes in agricultural soils? Plant Soil, 338, 247-259.

Wendt J.W. and Hauser S., 2013. An equivalent soil mass procedure for monitoring soil organic carbon in multiple soil layers. European J. Soil Sci., 64, 58-65.