Method Article

Computation of total soil organic carbon stock and its standard deviation from layered soils

Tommaso Tadiello [a,*], Alessia Perego [a], Elena Valkama [b], Calogero Schillaci [c], Marco Acutis [a]

[a] *DiSAA, Department of Agricultural and Environmental Sciences, University of Milan, Italy*
[b] *Natural Resources Institute Finland (Luke), Bioeconomy and Environment, Sustainability Science and Indicators, Tietotie 4, 31600 Jokioinen, Finland*
[c] *European Commission, Joint Research Centre (JRC), Ispra, Italy*

ABSTRACT

To assess carbon sequestration in the agricultural and natural systems, it is usually required to report soil organic carbon (SOC) as mass per unit area (Mg ha-1) for a single soil layer (e.g., the 0-0.3 m ploughing layer). However, if the SOC data are reported as relative concentration (g kg-1 or %), it is required to compute the SOC stock and its standard deviation (SD) for a given layer as the product of SOC concentration and bulk density (BD). For a proper computation, it is required to consider that these two variables are correlated. Moreover, if the data are already reported as SOC stock for multiple sub-layers (e.g., 0-0.15 m, 0.15-0.3 m) it is necessary to compute the SOC stock and its SD for a single soil layer (e.g., 0-0.3 m). The correlation between stocks values from adjacent and non-adjacent soil sub-layers must be accounted to compute the SD of the single soil layer.

The present work illustrates the methodology to compute SOC stock and its SD for a single soil layer based on SOC concentration and BD also from multiple sub-layers. An Excel workbook automatically computes the means of stocks and SD saving the results in a ready-to-use database.

- Computation of a carbon (SOC) stock and its standard deviation (SD) from the product between SOC concentration and bulk density (BD), being correlated variables.
- Computation of a SOC stock and its SD from the sum of SOC stocks of multiple correlated sub-layers.
- An Excel workbook automatically computes the means of SOC stocks and SD and saves the results in a ready-to-use database.

---

* Corresponding author.
E-mail address: tommaso.tadiello@unimi.it (T. Tadiello).

Specifications Table

Subject Area;	Agricultural and Biological Sciences
More specific subject area;	*Agricultural soils*
Method name;	*Computation of total soil organic carbon stock and its standard deviation from layered soils*
Name and reference of original method;	*Bohrnstedt, G., Goldberger, A., 1969. On the Exact Covariance of Products of Random Variables. Journal of the American Statistical Association, 64:328, 1439-1442.* http://dx.doi.org/10.1080/01621459.1969.10501069 *Kenney, J.F. and Keeping, E. S., 1951. Mathematics of Statistics, Pt. 2, 2nd ed. Princeton, NJ: Van Nostrand*
Resource availability	https://doi.org/10.6084/m9.figshare.15066402.v3

Method details

Background

Soil organic carbon (SOC) is regarded as the second most extensive carbon reservoir of the planet [4]. SOC stock (i.e., carbon stock expressed as Mg ha^{-1} or kg m^{-2}) can be increased by enhancing biomass production and retaining crop residues as an effective mitigation action against climate change, as stated by the 4 per 1000 international initiative [6].

An exponential growth in numbers of experiments about SOC took place in the last decades [8]: a literature search carried out on Scopus in 2020 revealed 5586 articles under the keywords searching "soil organic carbon" and 100 articles with "soil organic carbon meta-analysis". The significant attention around this topic is mainly due to the need for the definition of the best practices to enhance carbon sequestration for a specific agroecosystem.

Most of the articles yielded by the literature search report SOC as relative concentration (e.g., g kg^{-1}) and frequently for multiple soil layers or sampling depths. However, when SOC is the dependent variable in meta-analysis it is required to treat SOC as stock for a certain depth, for example, for ploughing soil layers (0-0.3 m).

If SOC data are not directly reported as stock, the following computation is needed:

$$SOC \ [Mg \ ha^{-1}] = OC \ \cdot LT \ \cdot BD \ \cdot (1 - RF) \ \cdot 10^4 \tag{1}$$

where OC is the organic carbon concentration (%), LT is the layer thickness (m), BD is the topsoil bulk density (Mg m^{-3}), and RF is the rock fragment content fraction (-) [1]. However, this formula does not take into account the correlation existing between OC and BD.

It is also essential to compute the variability (i.e., standard deviation, SD) associated with the estimated SOC stock mean. The correct SD computation plays a role in two different situations: (1) when SOC stock is computed as the product between the OC and BD; (2) when SOC stock of different sub-layers are summed to obtain a single value for a total soil layer. In both cases, the SD must be computed considering the correlations existing between SOC stock and BD, and between these values of different sub-layers. This computation is essential to analyse unbiased data: the mean of SOC stock derived from Eq. 1 and its SD cannot be merely computed as the arithmetic mean [9], without considering the measurement error and the fact that they are correlated [7].

This work aims to develop two different unbiased procedures to compute an SOC stock mean and its SD for a soil layer/horizon from the product of OC and BD (Case 1) or alternatively from the sum of SOC stocks of multiple sub-layers/horizons (Case 2). Both cases can occur when data came directly from field samplings or data extraction from scientific publications. An Excel© workbook was

developed to compute the two procedures and to save the results as a final database, which can be further used in meta-analysis.

Soil organic carbon stock computation

The methodology is conceived to compute the SD of SOC for a control and a treatment that are compared in a meta-analysis for both Case 1 and Case 2. This is crucial for the correct computation of the effect-size.

Researchers usually compare the study variable means of a control and a treatment in experimental fields that differ in terms of tillage operations, crop residues management and other factors that can possibly affect SOC stock. Therefore, it is required a different computation of SD because the extent of correlation between the variables among soil layers depends on the control and treatment characteristics: for instance, the studied variables measured in the different soil layers under no-till management have a different correlation coefficient than that of under soil ploughing.

Case 1: product between OC and BD

The computation is based on OC relative concentration and BD; the resulting SOC stock is then considered a composite variable. For this reason, methodology considers both the correlation and the SDs of OC and BD to compute SOC stock and its SD.

The corresponding computation of SOC (expressed as Mg ha^{-1}) for a soil layer follows the next formula:

$$SOC_{sub-layer} = (OC \cdot BD + Cov(OC, BD)) \cdot (D_U - D_L) \cdot 10^4 \qquad (2)$$

where C is the carbon concentration [%], BD is the bulk density [Mg m^{-3}], D_U and D_L are the upper and lower layer depths [m] of the sub-layer. The covariance is computed as:

$$Cov(C, BD) = \sigma_C \cdot \sigma_{BD} \cdot \rho_{C,BD} \qquad (3)$$

where σ is the SD. To calculate the correlation (ρ), the regression coefficient (-0.6) based on Manrique and Jones [5] has been utilized, since they considered the OC as a predictor of BD.

The following formula computes SD between two variables (namely 1, 2, which are OC and BD in our computation) that are normally distributed [2]:

$$SD = \sqrt{\mu_1^2\sigma_2^2 + \mu_2^2\sigma_1^2 + 2(\mu_1\mu_2 Cov(1, 2)) + \sigma_1^2\sigma_2^2 + Cov^2(1, 2)} \qquad (4)$$

where σ and μ are respectively the SD and the mean of the two variables (i.e., OC and BD), while the covariance is computed as:

$$Cov(1, 2) = \sigma_1 \cdot \sigma_2 \cdot \rho_{1,2} \qquad (5)$$

The result of Eq. (4) has to be transformed according to the unit of the variable, in this case with a multiplicative factor of [$(D_U - D_L) \cdot 10^4$] as reported in (2) Eq. 4. can be correctly applied to the SOC computation due to the normal distribution of OC and BD values [7].

The OC and the BD are seldom measured on the same sample with independent replications. In this case, the best estimation of the SD of the carbon stock is the SD of the product (i.e., OC • BD) computed for each replication. When OC and BD are measured on samples separately taken, the Eq. 4 estimates the SD of the SOC stock.

Case 2: sum of stocks

This case applies when the SOC stock of a total soil layer is needed to be computed as the sum of SOC stocks of several sub-layers. In this case, SOC data from sampling or published studies are already expressed as stock (Mg ha^{-1} or kg m^{-2}). If data are reported as kg m^{-2}, they have to be expressed as Mg ha^{-1}. The SOC stock of the total soil layer is computed as:

$$SOC_{total\ soil\ layer} = \sum_{i=1}^{n} SOC_{stock} \qquad (6)$$

where SOC_{stock} is the carbon stock of the i_{th} sub-layer.

Fig. 1. Example of a single template for the setting of the data input. The user is allowed to fill the green cells with all the information related to the sample/study selected.

The SD of the total soil layer is obtained from the the formula [3] of the variance (σ_z^2) of the sum of two correlated variables:

$$\sigma_z^2 = \sigma_x^2 + 2r\sigma_x\sigma_y + \sigma_y^2 \tag{7}$$

where x and y are the two correlated variables, σ the SD, and r is their correlation coefficient.

The formula to compute the SD of the total soil layer is then:

$$SD = \sqrt{\sum_{i=1}^{n}\sigma^2(X_i) + 2(\sum_{i=1}^{n-1}Cov(X_i, X_{i+1}) + \sum_{i=1}^{n-2}Cov(X_i, X_{i+2}) + \sum_{i=1}^{n-3}Cov(X_i, X_{i+3}))} \tag{8}$$

where σ is the SD of the i_{th} sub-layer and n is the total number of layers analyzed. A single covariance term is computed as:

$$Cov(X_i, X_j) = \sigma_i \cdot \sigma_j \cdot \rho_{i,j} \tag{9}$$

where σ_i and σ_j are the SD of the X_i and X_j sub-layer, while ρ is the correlation between a pair of sub-layers according to their adjacent or non adjacent spatial disposition. Only layers that are close to each other (adjacent) or at a maximum of three layer far away (non-adjacent) are considered in the correlation matrix.

The Eq. 8 is to be used when different layers are taken from the same undisturbed soil core. The covariance terms in the Eq. 8 reflect the correlation patters existing between soil depth layers also when they are sampled in different locations.

The correlation between layers depends on the control and treatment. For instance, let the control be conventional agriculture and the treatment be conservation agriculture. In control, soil is perturbed, and soil layers are mixed due to ploughing; the correlation can be assumed equal to zero. In the treatment, the correlation coefficient in the Eq. 9 could be empirically computed from the initial carbon data of each one of the different soil layers. In particular, it is assumed that two contiguous soil layers are more correlated than distant layers. In the Excel© Workbook, described in the next paragraph, the user can let the workbook automatically retrieve the correlation coefficients from the data entered, or manually set different correlation coefficients.

Excel© workbook

The Excel© workbook was developed for the SOC stock and SD computation under *Case 1* and *Case 2*. The code was developed in VBA language.

The users are allowed to let the Excel© workbook automatically retrieve the correlation coefficients (only for the Case 2 computations) from the different carbon stock data entered or manually insert them. This difference defines the computation type. In fact, in the Excel© workbook, the "SELECT COMPUTATION TYPE" list allows choosing the computation type between "Automatic" and "Manual".

The automatic and the manual computations are described below separately.

Automatic computation type

The user has first to insert the carbon data following the template showed in Fig. 1. In each row of the template, and separately for control and treatment, the user can alternatively:

Fig. 2. Example of automatic setting of the correlation coefficients (product between OC and BD as Case 1, and sum of stocks as Case 2). In the Case 2, the correlation values are automatically retrieved separately for the control and treatment, and for the adjacent layers and non adjacent layers.

1) insert the OC % (and its SD) and the BD (and its SD) and leave the Excel© workbook to compute the SOC stock and its SD for a soil layer (Case 1);
2) directly insert the SOC stock and its SD of a sub-layers and leave the Excel© workbook to compute SOC stock and its SD for the total soil layer (Case 2).

Within each template (Fig. 1), when the standard error is inserted, it is required to insert the sample size, being the number of replicates adopted in each sampling/study. In this way, the Excel© workbook can compute the standard deviation from the standard error. In this Excel© workbook it is assumed that the sample size does not vary within each template (i.e., each sub-layer has the same sample size). In addition, it is mandatory to insert the lower and the upper layer depth ("Lower depth [cm]", "Upper depth [cm]") values.

After all data have been inserted in one or more templates, press "COMPUTE AND SAVE" (Fig. 2). The Excel© workbook will automatically retrieve the correlation coefficients and compute SOC stock for the total soil layer with the new correlation values. Then, the results are automatically saved in the "Database" sheet and the correlation values adopted in the computation appear in the automatic settings section, as showed in Fig. 2.

In the Automatic computation type:

- before pressing "COMPUTE AND SAVE", all the values in each template will be temporarily computed assuming correlation equal to zero;
- the correlation coefficients are retrieved from all carbon data inserted in the "Data input" sheet separately for control and treatment;
- the correlation value for adjacent/non adjacent layers is retrieved only when at least three adjacent/non adjacent pair layers are inserted;
- the correlation value for non adjacent layers is computed as a unique value considering only layers that are two or three layers far from each other. For example, if a single sampling/study reports five layers (i.e., 1-2-3-4-5) from the top one (1) to the bottom (5), the Excel© workbook computes the non adjacent correlation coefficients, separately for the control and treatment, considering the 1-3, 1-4, 2-5, 3-5 non adjacent layers pairs;
- the SD value of the total soil layer (Case 2) is computed applying the "correlation between adjacent layers" coefficient to the layers that are close to each other, and the "correlation between non adjacent layers" to the layers that are two or three layers far away from each other (Fig. 2);
- the correlation between OC and BD is always automatically set to -0.6.

Manual computation type

The user has to first insert the correlation coefficients for the product between OC and BD (Case 1) and the sum of stock (Case 2) for adjacent and non adjacent layers separately (Fig. 3). Then, the user can insert the carbon data following the template shown in Fig. 1. In each row of the template, and separately for control and treatment, the user can alternatively:

1) insert the OC % (and its SD) and the BD (and its SD) and leave the Excel© workbook to compute SOC stock and its SD for a soil layer (Case 1);

Fig. 3. Example of manual setting of the correlation coefficients (product between OC and BD as Case 1, and sum of stocks as Case 2). In the Case 2, the correlation values have to be defined by the user separately for the control and treatment, and for the adjacent layers and non adjacent layers.

Fig. 4. Database sheet. For each sample/study, SOC stock and SD in the total soil layer are saved in a single row separately for control and treatment. In the first row of the example, SOC stock and SD of a total soil layer (0-0.3 m) are displayed for control and treatment. Ancillary data, such as study and sample number, data source, date, and sample size are shown in the same row.

2) directly insert the SOC stock and its SD of the sub-layers and leave the Excel© workbook to compute SOC stock and its SD for the total soil layer (Case 2).

As for the automatic computation type, when the standard error is inserted (in the template shown in Fig. 1), it is required to insert the sample size, being the number of replicates adopted in each sampling/study. In this way, the Excel© workbook can compute the standard deviation from the standard error. In this Excel© workbook it is assumed that the sample size does not vary within each template (i.e., each sub-layer has the same sample size). In addition, it is mandatory to insert the lower and the upper layer depth ("Lower depth [cm]", "Upper depth [cm]") values.

After all the data have been inserted, press "SAVE" to let the Excel© workbook save all the results in the "Database" sheet.

In the Manual computation type:

- For those correlation values leaved empty, the correlation is set to zero;
- the SD value of the total soil layer (Case 2) is computed applying the "correlation between adjacent layers" coefficient to the layers that are close to each other, and the "correlation between non adjacent layers" to the layers that are two or three layers far away from each other.

As reported above, the Excel© workbook gives the possibility to store the results about SOC stock and SD for a total soil layer, also saving the general information (i.e., study number, or/and sample number as well as the data source, date of sampling and sample size). All this information is reported and ready to be used in the "Database" sheet (Fig. 4), in which SOC stocks and SDs are given for all the total soil layers computed.

All the information required to compute the SD with the present Excel© workbook are summarized in a user manual directly in a workbook sheet (i.e., "Manual & Tips" sheet).

The Excel© workbook can be used to collect and elaborate data for meta-analysis or for re-elaborating data from field samples that have been measured as carbon concentration or in soil sub-layers.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgements

References

[1] N.H. Batjes, Total carbon and nitrogen in the soils of the world, Eur. J. Soil Sci. 47 (1996) 151–163, doi:10.1111/j.1365-2389.1996.tb01386.x.

[2] [2] G. Bohrnstedt, A. Goldberger, On the exact covariance of products of random variables, J. Am. Stat. Assoc. 64 (328) (1969) 1439–1442, doi:10.1080/01621459.1969.10501069.

[3] [3] J.F. Kenney, E.S. Keeping, Mathematics of Statistics, Pt. 2, 2nd ed., Van Nostrand, Princeton, NJ, 1951.

[4] [4] E. Lugato, J.M. Lavallee, M.L. Haddix, P. Panagos, M.F. Cotrufo, Different climate sensitivity of particulate and mineral-associated soil organic matter, Nat. Geosci. 14 (2021) 295–300, doi:10.1038/s41561-021-00744-x.

[5] [5] L. Manrique, C. Jones, Bulk density of soils in relation to soil physical and chemical properties, Soil Sci. Soc. Am. J. 55 (1991) 476–481, doi:10.2136/sssaj1991.03615995005500020030x.

[6] [6] B. Minasny, B.P. Malone, A.B. McBratney, D.A. Angers, D. Arrouays, A. Chambers, V. Chaplot, Z.-S. Chen, K. Cheng, B.S. Das, D.J. Field, A. Gimona, C.B. Hedley, S.Y. Hong, B. Mandal, B.P. Marchant, M. Martin, B.G. McConkey, V.L. Mulder, S. O'Rourke, A.C. Richer-de-Forges, I. Odeh, J. Padarian, K. Paustian, G. Pan, L. Poggio, I. Savin, V. Stolbovoy, U. Stockmann, Y. Sulaeman, C.-C. Tsui, T.-G. Vågen, B. van Wesemael, L. Winowiecki, Soil carbon 4 per mille, Geoderma 292 (2017) 59–86, doi:10.1016/j.geoderma.2017.01.002.

[7] [7] S.M. Jensen, J. Svensgaard, C.,R. Ritz, Estimation of the harvest index and the relative water content – two examples of composite variables in agronomy, Eur. J. Agron. 112 (2020) (2020) 125962, doi:10.1016/j.eja.2019.125962.

[8] [8] P. Smith, S. Lutfalla, W.J. Riley, M.S. Torn, M.W.I. Schmidt, J.F. Soussana, The changing faces of soil organic matter research, Eur. J. Soil Sci. 69 (2018) 23–30, doi:10.1111/ejss.12500.

[9] [9] A.W. Van der Vaart, in: Delta Method. Asymptotic Statistics, Cambridge University Press, Cambridge, 1998, pp. 25–34, doi:10.1017/CBO9780511802256.