SOIL PHYSICS NOTE

Improved Estimation of Soil Organic Carbon Storage Uncertainty Using First-Order Taylor Series Approximation

Dileep K. Panda*
R. Singh
D. K. Kundu
H. Chakraborty
A. Kumar
Water Technology Centre for Eastern Region (ICAR)
Chandrasekharpur
Bhubaneswar 751023
Orissa
India

Assessment of soil organic C (SOC) stocks is important for monitoring the effect of land use change in the C cycle and for formulation of C sequestration strategies in the context of global climate change. Discrepancies among the recent global SOC estimates by different researchers underscore the importance of precise estimation of the uncertainty associated with the SOC stocks. A method was recently proposed to estimate the SOC storage uncertainty using the Taylor series of approximations. Here we show that the accuracy of SOC storage uncertainty can be improved by incorporating the covariance among the input variables. Measurement of input variables from independent samples or use of an incomplete model leads to either over- or underestimation of the SOC storage uncertainty. The application of the method to an experimental data set indicated that ignoring covariance would lead to a substantial overestimate of the uncertainty.

Abbreviations: SOC, soil organic carbon.

The SOC storage is estimated indirectly using a function of input variables: organic C concentration, bulk density, percentage of fragments >2 mm, and thickness of the soil layer. Schwager and Mikhailova (2002) estimated the approximate SOC storage and its variability using the Taylor series, with the assumption that the input variables are measured from separate samples of the same soil layer; however, measurement of the input variables from separate soil samples leads to the propagation of uncertainty in the SOC storage estimate (Tiessen et al., 1982). The input variables for SOC estimation are interrelated in the sense that bulk density influences organic C, soil thickness influences bulk density, etc. (Homann et al., 1995; Akala and Lal, 2001). Hence, our objective was to improve the estimation of the SOC storage uncertainty using the first-order Taylor series of approximations by incorporating the covariance among input variables. One numerical example has been demonstrated to examine the precision of the proposed method.

Soil Sci. Soc. Am. J. 72:1708-1710
doi:10.2136/sssaj2007.0242N
Received 3 July 2007.

*Corresponding author (dileeppanda@rediffmail.com).

THEORY

Soil Organic Carbon Storage Uncertainty Estimation by Taylor Series of Approximations

The function to estimate the SOC storage in a particular soil layer is defined as (Homann et al., 1995; Batjes, 1996)

$$y = f(x) = x_1 x_2 x_3 (1 - b x_4) \quad [1]$$

where y is SOC storage (kg m^{-2}), x_1 is the SOC concentration in the ≤2-mm material (kg Mg^{-1}), x_2 is the soil bulk density (Mg m^{-3}), x_3 is the thickness of the soil layer (m), x_4 is the percentage of fragments >2 mm in diameter, and $b = 0.01$. For the soil profile as a whole, SOC can be estimated by $y_p = \sum_{l=1}^{p} x_{1l} x_{2l} x_{3l}(1 - b x_{4l})$ where $l = 1, 2, \ldots, p$ soil layers in the profile.

The Taylor series is used to obtain the state (dependent) variable when it exists as a nonlinear function of the input variables based on the series expansion and approximation to the function (Kendall and Stuart, 1969; Dudewicz and Mishra, 1988; Bakr and Butler, 2005). To estimate the SOC storage and the associated uncertainty, we define here the vector of input variables $\mathbf{x} = [x_1\ x_2\ x_3\ x_4]^T$, with the expected values $E(\mathbf{x}) = \boldsymbol{\mu} = [\mu_1\ \mu_2\ \mu_3\ \mu_4]^T$ and $E(\mathbf{x} - \boldsymbol{\mu}) = [x_1 - \mu_1\ \ x_2 - \mu_2\ \ x_3 - \mu_3\ \ x_4 - \mu_4]^T$, respectively, where the superscript T indicates the transpose of the vector. Furthermore, let us define $E(\mathbf{x} - \boldsymbol{\mu})(\mathbf{x} - \boldsymbol{\mu})^T = \boldsymbol{\Omega}_{(4\times4)}$, where $\boldsymbol{\Omega}_{(4\times4)}$ is the (4 × 4) order variance–covariance matrix of the input variables. The matrix $\boldsymbol{\Omega}$ can then be expressed as

$$\Omega = \sigma_{ij} = \begin{pmatrix} \sigma_1^2 & \sigma_{12} & \sigma_{13} & \sigma_{14} \\ \sigma_{21} & \sigma_2^2 & \sigma_{23} & \sigma_{21} \\ \sigma_{31} & \sigma_{32} & \sigma_3^2 & \sigma_{34} \\ \sigma_{41} & \sigma_{42} & \sigma_{43} & \sigma_4^2 \end{pmatrix} \quad [2]$$

where $\sigma_{ii} = \sigma_i^2 = E(x_i - \mu_i)^2$ is the variance of the random variable i, and $\sigma_{ij} = E(x_i - \mu_i)(x_j - \mu_j)$, the covariance between random variables i and j ($i, j = 1, 2, 3, 4$). Furthermore, we define the vector of the partial derivation of the function y evaluated at μ as

$$\frac{\partial f(x)}{\partial \mathbf{x}} = \left[\frac{\partial f(x)}{\partial x_1}\ \ \frac{\partial f(x)}{\partial x_2}\ \ \frac{\partial f(x)}{\partial x_3}\ \ \frac{\partial f(x)}{\partial x_4}\right]^T \quad [3]$$
$$= \left[\mu_2\mu_3(1 - b\mu_4)\ \ \mu_1\mu_3(1 - b\mu_4)\ \ \mu_1\mu_2(1 - b\mu_4)\ \ -b\mu_1\mu_2\mu_3\right]^T$$

The first-order Taylor series expansion about the mean is defined as

$$y = f(x) \cong f(\mu) + (x - \mu)^T \frac{\partial f(x)}{\partial \mathbf{x}} \quad [4]$$

The expectation and variance of $f(x)$ in Eq. [1] can be derived as

$$E[f(x)] = f(\mu) = \mu_1\mu_2\mu_3(1 - b\mu_4) \quad [5]$$

and the variance (i.e., uncertainty)

$$\mathrm{Var}\left[f(x)\right]=\mathrm{E}\left[f(x)^2\right]-\left\{\mathrm{E}\left[f(x)\right]\right\}^2\cong\frac{\partial f(x)}{\partial \mathbf{x}}^{\mathrm{T}}\Omega\frac{\partial f(x)}{\partial \mathbf{x}} \quad [6]$$

As can be seen from Eq. [6], the computational requirement to estimate the uncertainty associated with a state variable is the calculation of the partial derivatives of $f(x)$ with respect to $\mathbf{x}$ (Eq. [3]) and the variance–covariance matrix $\mathbf{\Omega}$. The SOC storage estimate and the associated uncertainty for four experimental situations have been derived using Eq. [5] and [6], respectively (Table 1). The ideal situation (Function I) for precise estimation of SOC storage is ensured when all four input variables are measured on the same sample so that all the entries of matrix $\mathbf{\Omega}$ and vector $\partial f(x)/\partial \mathbf{x}$ can be obtained. It is not always possible, however, to measure all the input variables on the same sample. Sometimes a variable is assumed to be constant based on past studies or is estimated using an empirical formula (Post and Kwon, 2000). In some situations, a conversion coefficient is used for the percentage of fragments >2 mm to estimate the SOC storage (Wang et al., 2003). Often the bulk density and percentage of fragments >2 mm are measured on independent samples for simplification of the procedure (Tiessen et al., 1982). These situations can lead to over- or underestimation of the uncertainty, depending on the completeness of the function and the covariance among the input variables.

Function II (Table 1) represents an experimental situation where the input variables are measured on independent samples of different soil pedons or the data are collected from different sources. Consequently, the covariance (i.e., the off-diagonal elements of $\mathbf{\Omega}$) cannot be estimated. By substituting 0 for the covariance terms σ_{ij} ($i \neq j$ = 1, 2, 3, 4) in the uncertainty estimate of Function I or by using Eq. [6], the associated uncertainty can be estimated. This result equals Eq. D of Schwager and Mikhailova (2002). Under the assumption of independence among the input variables, the uncertainty of all the selected SOC storage functions of Schwager and Mikhailova (2002) can be obtained either from Eq. [6] or from the uncertainty estimate of Function I by substituting the terms as per the defined function and sampling scheme.

Function III represents an experimental situation where the soil samples have been obtained from a fixed thickness of the soil layer a, and SOC concentration, bulk density, percentage of fragments >2 mm are the sources of variability. The uncertainty of SOC storage can be obtained by substituting a for μ_3 and 0 for σ_{3i} (i = 1, 2, 3, 4) in the uncertainty of Function I. The SOC storage uncertainty for a soil column consisting of incremental soil layers can also be estimated by adding the individual soil layer uncertainties of the profile. Furthermore, if it is assumed that the percentage of fragments >2 mm does not contribute significantly to SOC storage (Function IV), then the uncertainty can be obtained by substituting a for μ_3 and 0 for μ_4, σ_{3i}, and σ_{4i} (i = 1, 2, 3, 4) in the uncertainty result of Function I. All terms for estimation of SOC storage and its uncertainty can be obtained by using the sample estimates m_i, s_i^2, and s_{ik} for μ_i, σ_i^2, and σ_{ij}, respectively.

MATERIALS AND METHODS

Data Set

The experimental site is farmland covering an area of ~4 ha, belonging to the Coastal Plains and Islands physiographic region on the northwestern side of India (22°26′ N, 72°49′ E to 22°55′ N, 73°23′ E). At 60 randomly located pits, the soil samples were collected to a random depth to 1.5 m from the surface. Samples were collected by core samplers of 0.1-m diameter for determining the bulk density of undisturbed soils of a pit of desired depth. Each pit was represented by a single average value of bulk density. To reduce the cost of sampling, a single sample representing different soil layers was obtained from each pit to measure organic C. Our interest was to estimate the SOC storage uncertainty at random depth, accounting more spatial and vertical representation of the study area. The mean and standard deviation of the input parameters of the 60 pits are: organic C, 3.16 ± 1.24 kg m^{-2}; bulk density, 1.50 ± 0.05 kg Mg^{-1}; and depth (i.e., thickness) 0.85 ± 0.48 m. The percentage of fragments >2 mm was assumed to be constant (i.e., mean 1.5%, standard deviation 0). Details of other soil characteristics can be obtained from Singh and Nayak (1999).

RESULTS AND DISCUSSION

The conventional method directly calculates the SOC storage for each sample and then takes the average and variance of the calculated values to obtain the overall SOC storage and associated uncertainty. Here, the average and the variance of SOC storage calculated from 60 samples (pits) was found to be 3.35 kg m^{-2} and 2.47 (kg $m^{-2})^2$, respectively. For the Taylor method of estimation, the experimental situation indicates that four input variables contribute to the SOC storage estimate. The SOC concentration, bulk density, and soil depth are the sources of uncertainty, however, as the gravel percentage has been assumed to be constant. The function would then be $f(x) = cx_1x_2x_3$, and SOC storage can be estimated as $c\mu_1\mu_2\mu_3$, where $c = (1 - b\mu_4)$. Substituting c for $(1 - b\mu_4)$ and 0 for σ_{4i} (i = 1, 2, 3, 4) in the uncertainty estimate of Function I, the SOC storage uncertainty of this experimental situation can be obtained as: $\mathrm{Var}[f(x)] = c^2[\mu_1^2\mu_2^2\sigma_3^2 + \mu_1^2\mu_3^2\sigma_2^2 + \mu_2^2\mu_3^2\sigma_1^2 + 2\mu_1\mu_2\mu_3(\mu_1\sigma_{23} + \mu_2\sigma_{13} + \mu_3\sigma_{12})]$. This

Table 1. The soil organic C (SOC) storage and associated approximate variance or uncertainty of selected experimental situations, where x_1 is SOC concentration in the ≤2-mm material, x_2 is soil bulk density, x_3 is thickness of the soil layer, x_4 is percentage of fragments >2 mm in diameter, and b = 0.01.

Function	SOC storage	Variance or uncertainty
I. $f(x) = x_1x_2x_3(1 - bx_4)$	$\mathrm{E}[f(x)] = \mu_1\mu_2\mu_3(1 - b\mu_4)$	$\mathrm{Var}[f(x)] = (1 - b\mu_4)^2[\mu_1^2\mu_2^2\sigma_3^2 + \mu_1^2\mu_3^2\sigma_2^2 + \mu_2^2\mu_3^2\sigma_1^2 + 2\mu_1\mu_2\mu_3(\mu_1\sigma_{23} + \mu_2\sigma_{13} + \mu_3\sigma_{12})] - 2b\mu_1\mu_2\mu_3(1 - b\mu_4)(\mu_1\mu_2\sigma_{34} + \mu_1\mu_3\sigma_{24} + \mu_2\mu_3\sigma_{14}) + b^2\mu_1^2\mu_2^2\mu_3^2\sigma_4^2$
II. $f(x) = x_1x_2x_3(1 - bx_4)$ independence assumption	$\mathrm{E}[f(x)] = \mu_1\mu_2\mu_3(1 - b\mu_4)$	$\mathrm{Var}[f(x)] = (1 - b\mu_4)^2(\mu_1^2\mu_2^2\sigma_3^2 + \mu_1^2\mu_3^2\sigma_2^2 + \mu_2^2\mu_3^2\sigma_1^2) + b^2\mu_1^2\mu_2^2\mu_3^2\sigma_4^2$
III. $f(x) = ax_1x_2(1 - bx_4)$	$\mathrm{E}[f(x)] = a\mu_1\mu_2(1 - b\mu_4)$	$\mathrm{Var}[f(x)] = (1 - b\mu_4)^2(a^2\mu_1^2\sigma_2^2 + a^2\mu_2^2\sigma_1^2 + 2a^2\mu_1\mu_2\sigma_{12}) - 2ab\mu_1\mu_2(1 - b\mu_4)(a\mu_1\sigma_{24} + a\mu_2\sigma_{14}) + a^2b^2\mu_1^2\mu_2^2\sigma_4^2$
IV. $f(x) = ax_1x_2$	$\mathrm{E}[f(x)] = a\mu_1\mu_2$	$\mathrm{Var}[f(x)] = a^2(\mu_1^2\sigma_2^2 + \mu_2^2\sigma_1^2 + 2\mu_1\mu_2\sigma_{12})$

result can also be derived through the matrix operation in Eq. [6]: substituting 0 for σ_{4i} (i = 1, 2, 3, 4) in $\mathbf{\Omega}$ and deriving the sensitivity coefficients as $\partial f(x)/\partial \mathbf{x} = a[\mu_2\mu_3\ \ \mu_1\mu_3\ \ \mu_1\mu_2\ \ 0]^{\mathrm{T}}$. The corresponding sample estimate of SOC storage is $\hat{y} = cm_1m_2m_3,$ and the associated uncertainty is

$$\mathrm{Var}\left(\hat{y}\right)=c^2\begin{bmatrix} m_2^2m_3^2s_1^2+m_1^2m_3^2s_2^2+m_1^2m_2^2s_3^2+ \\ 2m_1m_2m_3\left(m_1s_{23}+m_2s_{13}+m_3s_{12}\right)\end{bmatrix} \qquad [7]$$

If the covariance among the input parameters is ignored, the uncertainty would be estimated by

$$c^2\left(m_2^2m_3^2s_1^2+m_1^2m_3^2s_2^2+m_1^2m_2^2s_3^2\right) \qquad [8]$$

Substituting the mean, variance, and covariance of the input parameters, the SOC storage was found to be 3.98 kg m^{-2} for the cultivated farmland soils of the west coast ecosystem. Using Eq. [7], the uncertainty associated with SOC storage was estimated to be 2.45 (kg m^{-2})^2. Our proposed method is precise as it yields a lesser uncertainty than the conventional method although the SOC storage is overestimated. The conventional method computes the variance among the calculated values (i.e., between estimates) of SOC storage. The Taylor method uses observed input variables for uncertainty estimation, however, which helps in tracking the relative change in uncertainty due to changes in input parameters. Ignoring the covariance between SOC concentration and bulk density (s_{21} = −0.018), SOC concentration and soil depth (s_{31} = −0.44), and soil depth and bulk density (s_{32} = 0.013) leads to a substantial overestimation of the uncertainty (7.43 (kg m^{-2})^2 using Eq. [8]) associated with the SOC storage. Thus, the sampling strategy should ensure the estimation of covariance for precise estimation of SOC storage uncertainty.

CONCLUSIONS

Ignorance of the covariance among input variables or use of an incomplete model could lead to either over- or underestimation of the SOC storage uncertainty. Application of this method to an experimental data set indicated that the ignorance of covariance led to a substantial overestimate of the uncertainty. The approximation derived in Eq. [6] can be used to obtain SOC storage uncertainty for a variety sampling campaigns.

REFERENCES

Akala, V.A., and R. Lal. 2001. Soil organic carbon pools and sequestration rates in reclaimed minesoils in Ohio. J. Environ. Qual. 30:2098–2104.

Bakr, M.I., and A.P. Butler. 2005. Nonstationary stochastic analysis in well capture zone design using first-order Taylor's series approximation. Water Resour. Res. 41:W01004, doi:10.1029/2004WR003648.

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

Dudewicz, E.J., and S.N. Mishra. 1988. Modern mathematical statistics. John Wiley & Sons, New York.

Homann, P.S., P. Sollins, H.N. Chappell, and A.G. Stangenberger. 1995. Soil organic carbon in a mountainous, forested region: Relations to site characteristics. Soil Sci. Soc. Am. J. 59:1468–1475.

Kendall, M., and A. Stuart. 1969. The advanced theory of statistics: Distribution theory. Vol.1. 3rd ed. Hafner Publ. Co., New York.

Post, W.M., and K.C. Kwon. 2000. Soil carbon sequestration and land-use change: Processes and potential. Global Change Biol. 6:317–328.

Schwager, S.J., and E.A. Mikhailova. 2002. Estimating variability in soil organic carbon storage using the method of statistical differentials. Soil Sci. 167:194–200.

Singh, R., and A.K. Nayak. 1999. Water retention and transmission characteristics of Mahi Right Bank Canal Command Area of Gujarat. J. Indian Soc. Soil Sci. 47:9–15.

Tiessen, H., J.W.B. Stewart, and J.R. Bettany. 1982. Cultivation effects on the amounts and concentrations of carbon, nitrogen, and phosphorus in grassland soils. Agron. J. 74:831–835.

Wang, S., H. Tian, J. Liu, and S. Pan. 2003. Pattern and change of soil organic carbon storage in China: 1960s–1980s. Tellus B 55:416–427.