Scale-dependent variability of soil organic carbon coupled to land use and land cover

X. Xiong\textsuperscript{a}, S. Grunwald\textsuperscript{a,b}, R. Corstanje\textsuperscript{b}, C. Yu\textsuperscript{c}, N. Bliznyuk\textsuperscript{d}

\textsuperscript{a}Department of Soil and Water Science, University of Florida, Gainesville, FL 32611, USA
\textsuperscript{b}National Soil Resources Institute, Cranfield University, Cranfield, Bedfordshire, MK43 0AL, UK
\textsuperscript{c}College of Hydrology and Water Resource, Hohai University, Xikang Road, Nanjing, Jiangsu 210098, China
\textsuperscript{d}Department of Agricultural and Biological Engineering, University of Florida, Gainesville, FL 32611, USA

\textbf{A R T I C L E   I N F O}

Article history:
Received 12 May 2015
Received in revised form 1 March 2016
Accepted 3 March 2016
Available online 11 March 2016

\textbf{Keywords:}
Soil organic carbon (SOC)
Field-scale spatial variability
Spatially nested sampling
Semivariogram
Hierarchical analysis of variance

\textbf{A B S T R A C T}

Understanding the field-scale spatial variability of soil organic carbon (SOC) is critical to assess its spatial distribution at very fine scale (several meters) which is valuable for precision agriculture and natural resource management. The aim of this study was to investigate the field-scale spatial variability of SOC under five prevalent land use and land cover (LULC) types in Florida, U.S. with a uniform sampling scheme. Five scales, 2, 7, 22, 67 and > 200 m, were targeted and 108 soil samples at 0−20 cm depth were collected and analyzed for SOC and bulk densities within each LULC type in 2012. Results indicate that SOC variability was scale dependent. Hardwood Hammock and Forest and Improved Pasture demonstrated large variation at both coarse scale (67 and > 200 m) and very fine scale (2 m). Sandhill, Pineland and Dry Prairie were dominated by variation at very fine scales (2 and 7 m). All five sites showed large variability at very fine scales, indicating the close coupling of SOC stock variation to structure and composition of vegetation. This study also identified that log-transformed SOC showed variance-invariant behavior, which had an approximately constant overall variance (sill) of 0.067 ± 0.012 (log (kg m\textsuperscript{-2}))/m\textsuperscript{2} at field scale (~500 m) irrespective of LULC. These findings serve to explain field-scale variability of SOC relevant for precision agriculture and land management, but also facilitate better understanding of the scale-dependent fine-scale variability of SOC across larger soilscapes.

1. Introduction

In the context of sustainable agriculture and global climate change there has been great interest in soil organic carbon (SOC) (Batjes and Sombroek, 1997; Bellamy et al., 2005; Franzluebbers, 2005; Spargo et al., 2008). Globally, great efforts have been dedicated to quantifying the spatial distribution and achieving accurate assessments of SOC (Grunwald, 2009; Minasny et al., 2013). However, the current soil spatial data around the globe are at rather coarse scale (Grunwald et al., 2011). For example, in the U.S. the major two soil databases providing readily available soil data are State Soil Geographic Database (STATSGO2) and Soil Survey Geographic Database (SSURGO via Soil Data Mart) at scale of 1:250,000 and 1:24,000, respectively. These soil databases cannot reflect SOC variation at field scale (i.e., within a few hundred meters) and provide little spatial information for precision agriculture and land management.

There have been field-scale studies of SOC variability, however they are difficult to compare because of differences in sampling design, protocols, density of observations, sample support, carbon (C) measurement techniques, soil depth, and other factors. Even though these studies were conducted in different climatic zones, land use and land cover (LULC), soil classes, topography, and intensity of human interferences, the maximum spatial autocorrelation ranges rarely exceeded 500 m (Cambardella et al., 1994; Cerri et al., 2004; McBratney and Pringle, 1999; Rossi et al., 2009; Terra et al., 2004; Tranmar et al., 1987). These studies provide knowledge about the field-scale variability of SOC under specific environmental and geographic conditions. However, they lack the ability to generalize because differences in the field-scale variation of SOC may be due to methodology, SOC variation, or scale-dependent relationships between SOC and other prominent environmental conditions, such as topography, climate, or net ecosystem productivity (Corstanje et al., 2007; Vasques et al., 2012). Florida soils hold a large amount of C and have great potential to accumulate more carbon. According to the estimates of SOC based on the State Soil Geographic database (STATSGO) by Vasques et al. (2010), Florida has the highest SOC stock value per
unit area among the conterminous U.S. states with minimum, median, and maximum values of 12.4, 35.3, and 64.0 kg m\(^{-2}\), respectively. In addition, Florida enjoys a wide variety of LULC classes unique to the Southeastern U.S. and SOC stocks can vary dramatically among ecosystems (Vasques et al., 2012; Xiong et al., 2014). However, little effort has been dedicated to the understanding of fine-scale spatial variability of SOC within and across different LULC types in this region.

Various sampling designs to investigate scale dependent variation of SOC that minimizes sampling effort have been suggested. Youden and Mehlich (1937) introduced a spatially nested sampling design to study soil spatial variation over multiple magnitudes of scales cost-effectively compared with systematic or random sampling techniques. Webster et al. (2006) formulated the theory of this technique. They successfully applied both balanced and unbalanced spatial nested sampling schemes to study various soil properties and reported the unbalanced design had comparable performance to the balanced one with much fewer samples and better distribution of degrees of freedom across scales. Lark (2011) then explored the scope for optimization of the unbalanced spatially nested design using simulated annealing and claimed that the optimization can theoretically yield better estimation of variance components (i.e., smaller variance of estimation) given a fixed number of samples and scales.

The major objective of this study was to investigate the field-scale variability of SOC under five common LULC classes in Florida with a uniform spatially nested sampling scheme. The specific objectives were to (i) investigate and compare the spatial structures of SOC in five LULC classes at field-scale (<500 m); (ii) identify the scale-dependent behavior and the controlling factors of SOC variation, and (iii) obtain an average semivariogram to capture the general characteristics of the SOC spatial variability at field scale.

2. Materials and methods

2.1. Description of study areas

The study areas are located in Florida, a state in Southeastern U.S., with latitudes from 24°27' N to 31°00' N and longitudes from 80°02' W to 87°38' W. The climate of Northern and Central Florida is humid, subtropical and Southern Florida has tropical climate. The mean annual precipitation of Florida is 1373 mm and the mean annual temperature is 22.3 °C (National Climatic Data Center, 2008). Overall, soils in Florida are sandy in texture. Dominant soil orders of Florida are: Spodosols (Podzols in World Reference Base for Soil Resources, 29%), Entisols (Fluvisol, 20%), Ultisols (Acrisol, 17%), Alfisols (Luvisols, 12%) and Histisols (Histisols, 10%) as shown in Fig. 1 (Natural Resources Conservation Service, 2009). Florida’s topography is muted with gentle slopes varying from 0 to 5% in almost the whole state (United States Geological Survey, 1999).

To investigate the field-scale variation of SOC under different LULC, five of the most prevalent and contrasting LULC classes in Florida were selected for sampling, namely Pineland (accounting

---

1 The definition of Dry Prairie adopted in this study is a large native grass and shrubland occurring on very flat terrain interspersed with scattered cypress domes and strands, bayheads, isolated freshwater marshes, and hardwood hammocks, according to Florida Fish and Wildlife Conservation Commission (2003).
for 15.5% of the total area of Florida), Improved Pasture (7.0%), Dry Prairie (2.9%), Hardwood Hammock and Forest (2.3%), and Sandhill (1.8%) (Florida Fish and Wildlife Conservation Commission, 2003). In total, the selected LULC classes account for approximately 36% of the Florida land area (excluding open water areas). Spatially, the five sites are spread out across Florida as shown in Fig. 1. The soil properties and environmental conditions of the five sites are summarized in Table 1.

### 2.2. Optimized unbalanced spatially nested sampling

At each of the five LULC sites, the same sampling scheme and protocol were used. The sampling followed an optimized unbalanced spatially nested design. The theory of the spatially nested design can be found in Webster et al. (2006) and the optimization of the design in Lark (2011). Five levels of hierarchy were studied (Table 2) and a detailed account of the design is described as follows. Nine main stations gridded at 200 m intervals constitute Level 1 (Fig. 2a). At each main station, two additional sampling points (subnodes) were chosen 67 m away and form an equilateral triangle with the main station sampling point. The equilateral triangle is placed in a random direction. Then in similar fashion the 3rd, 4th and 5th hierarchical samples were sampled 22 m, 7 m and 2 m away from their parent nodes, respectively. The approximate three-fold hierarchy has proven effective in capturing as much variation as possible and avoiding overlaps among different branches (Webster et al., 2006). Fig. 2b shows the sample deployment of one main station. A number of 108 (9 × 12) samples were collected within each LULC resulting in a total number of 540 samples for the five sites.

### 2.3. Soil sampling and laboratory analysis

Predetermined sample locations were identified with a sub-meter accuracy global positioning system (Trimble GeoXT 2005 series). Soil samples were collected with a 5.7-cm-diameter-steel core at 0–20 cm depth. Soil samples were put into sample bags and transported to the laboratory in coolers at 0 °C. Then samples were air dried, weighed, sieved through a 2-mm sieve, and ball milled. Total carbon (TC) was analyzed by combustion catalytic oxidation at 900 °C (Shimadzu SS/M-5000A). Considering that the topsoil in Florida contains little inorganic carbon, a pedotransfer function developed for Florida soils was used to derive SOC (Vasques et al., 2009). The pedotransfer function is: SOC (% mass) = 0.9977 × TC (% mass) − 0.0389 (n = 1080, R² = 0.999). Laboratory SOC measurements in mass units (mg kg⁻¹) were converted to stock units (kg m⁻²) using the measured bulk density and soil depth (20 cm). The SOC data in the five LULC classes were positively skewed; hence they were log-transformed to approximate normal distribution and stabilize the variance for the statistical and geostatistical analyses. In this study, the soil bulk density was homogeneous within each site as indicated by the small standard deviation in Table 1 (range from 0.07 to 0.11 g cm⁻³). Therefore, its contribution to SOC stock uncertainty was negligible (Taalab et al., 2013).

### 2.4. Hierarchical analysis of variance

To estimate the variance components of SOC at the five scales, the linear mixed model was used (Eq. (1)).

\[
Z_{ijklmn} = \mu + A_i + B_i + C_{ijk} + D_{ijkl} + E_{ijkm} + e_{ijklmn}
\]

(1)

where \(Z_{ijklmn}\) is the log-transformed SOC value of the nth sampling point in the nth class at Level 5, in the ith class at Level 4, . . . in the jth class at Level 1. \(\mu\) is the overall mean of the log-transformed SOC at one site, \(A_i\) is the difference between mean of the ith main station and the overall mean \(\mu\), and \(A_i\) is an identically and independently distributed Gaussian random variable with mean zero and variance \(\sigma^2_{A}\). \(B_i\) is the difference between the mean of the jth class within the ith main station and the mean of the ith main station, and \(B_i\) is an identically and independently distributed Gaussian random variable with mean zero and variance \(\sigma^2_{B}\), and so on. The error term is represented by \(e_{ijklmn}\). In this case, there is only one fixed effect, which is the overall mean \(\mu\), while the five random effects represent the five scales. To estimate the variance components, \(\sigma^2_{A}, \sigma^2_{B}, \ldots, \sigma^2_{E}\) from the data with an unbalanced design, restricted maximum likelihood (REML) was used with the assumption of normally distributed random effects (Webster et al., 2006). The likelihood ratio test was used to test the significance of a variance component by comparing the full model (Eq. (1)) and the model with the random effect removed from the full model (Bates, 2010).

### 2.5. Geostatistical analysis

The experimental semivariogram was estimated with the Cressie-Hawkins estimator (Eq. (2)) that is robust to the impact of outliers (Cressie, 1993).

\[
y(h) = \frac{1}{2} \left( \frac{1}{N(h)} \sum_{i \neq j} (Z(s_i) - Z(s_j))^2 \right)^{1/2} \left( 0.457 + \frac{0.494}{N(h)} \right)
\]

(2)
where $N(h)$ is the number of data pairs separated by a distance of $h,$ $Z(s_i)$ and $Z(s_j)$ are the SOC observation at location $s_i$ and $s_j$ respectively.

The Matérn theoretical semivariogram model (Eq. 3) was fitted the experimental semivariogram with REML. It generalizes a variety of semivariogram models and has a smoothness parameter that can be used to infer short-range variation (Minasny and McBratney, 2005).

$$\gamma(h) = \tau^2 + \sigma^2 \left( \frac{1}{2^{\nu-1} \Gamma(\nu)} \cdot \frac{h^\nu}{\phi^\nu} \cdot K_\nu \left( \frac{h}{\phi} \right) \right)$$

(3)

where $\tau^2$ is the nugget, $\sigma^2$ is the partial sill, $h$ is the lag distance, $K_\nu$ is a modified Bessel function of the second kind of order $\nu,$ $\Gamma$ is the gamma function and $\phi$ is the range parameter and $\nu$ is the smoothness parameter.

The Matérn model generalizes a variety of semivariogram models and has a smoothness parameter that can be used to infer the short-range variation (Minasny and McBratney, 2005).

### 2.6. Average semivariogram

The average semivariogram was developed to represent the common characteristics of the soil spatial variability regardless of the environmental conditions (e.g., LULC), which provides prior knowledge allowing to infer the soil variability at new sites (McBratney and Pringle, 1999). To derive the average SOC semivariogram, the spatial linear mixed model (Eq. (4)) was applied on the pooled five-site dataset (Table 3).

$$Y = X\beta + \varepsilon$$

(4)

where $Y$ is an $n \times 1$ vector of observed SOC values, $X$ is an $n \times 5$ dummy coded design matrix of five LULC classes (fixed effects), $\beta$ is the vector of fixed-effect coefficients, and $\varepsilon$ is an $n \times 1$ vector of errors.

Two methods were used to estimate the model parameters. The first method used ordinary least squares (OLS) to estimate $\beta$ by assuming $\varepsilon$ to be identically and independently distributed random variables of mean zero and variance $\sigma^2,$ i.e., $\varepsilon \sim N(0, \sigma^2 I),$ where $I$ is the identical matrix. The second method used generalized least squares (GLS) to estimate $\beta$ by taking the correlation among errors into account, i.e., $\varepsilon \sim N(0, V),$ where $V$ is the variance-covariance matrix of error variables that were estimated with REML (Lark and Cullis, 2004). The likelihood ratio test was used to test whether the GLS model is significantly better than the OLS model (Bates, 2010).

### 3. Results

#### 3.1. Descriptive statistics of soil organic carbon

The mean SOC varied greatly across the five LULC classes. Dry Prairie, Hardwood Hammock and Forest, and Improved Pasture had relatively high SOC stocks, while Sandhill and Pineland soil had relatively low SOC (Table 1). Linking the SOC means of the five sites with the edaphic and environmental conditions reveals the underlying controls of SOC levels across different LULC classes. Improved Pasture, although very high in soil available water capacity (AWC), is inclined to accumulate SOC. However, it ranked just as intermediate in the sequence of LULC classes in terms of accumulation of SOC due to the moderate net primary productivity (NPP) in the Improved Pasture system, which provided moderate C.

<table>
<thead>
<tr>
<th>Sites</th>
<th>Mean</th>
<th>Median</th>
<th>SD</th>
<th>CV</th>
<th>Skew.</th>
<th>Mean</th>
<th>SD</th>
<th>CV</th>
<th>Skew.</th>
</tr>
</thead>
<tbody>
<tr>
<td>SH</td>
<td>0.86</td>
<td>1.49</td>
<td>1.58</td>
<td>2.87</td>
<td>0.42</td>
<td>26.6</td>
<td>0.92</td>
<td></td>
<td></td>
</tr>
<tr>
<td>HHF</td>
<td>2.16</td>
<td>3.90</td>
<td>4.13</td>
<td>7.02</td>
<td>1.18</td>
<td>28.6</td>
<td>0.60</td>
<td></td>
<td></td>
</tr>
<tr>
<td>PL</td>
<td>0.84</td>
<td>1.65</td>
<td>1.78</td>
<td>5.64</td>
<td>0.63</td>
<td>35.4</td>
<td>3.05</td>
<td></td>
<td></td>
</tr>
<tr>
<td>IP</td>
<td>1.79</td>
<td>3.13</td>
<td>3.28</td>
<td>6.97</td>
<td>0.85</td>
<td>25.9</td>
<td>1.56</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DP</td>
<td>2.31</td>
<td>4.06</td>
<td>4.17</td>
<td>8.54</td>
<td>0.98</td>
<td>23.5</td>
<td>1.13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Whole</td>
<td>0.84</td>
<td>2.88</td>
<td>2.99</td>
<td>8.54</td>
<td>1.41</td>
<td>47.2</td>
<td>0.65</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Abbreviations: SH = Sandhill, HHF = Hardwood Hammock and Forest, PL = Pineland, IP = Improved Pasture, DP = Dry Prairie, SD = standard deviation, CV = coefficient of variation, skew. = skewness.
input into soils. Sandhill had low AWC and low NPP and thus had very limited ability to sequester soil C. In contrast, Dry Prairie and Hardwood Hammock and Forest were high in both AWC and NPP and therefore tended to accumulate SOC.

The standard deviation of SOC also showed differences across the five LULC classes. Hardwood Hammock and Forest had the largest overall variation of SOC, followed by Dry Prairie and Improved Pasture, while Sandhill and Pineland showed modest variation. The coefficient of variation reflected that the overall variation generally increased with the average SOC stock, which results in a similar variation of SOC across the five sites. There was a significant linear correlation at 0.05 significance level (p-value = 0.021) between the variance and the squared mean of SOC (Fig. 3a). After natural log-transformation the variance was stabilized as indicated by Fig. 3b.

3.2. Hierarchical analysis of variance of soil organic carbon

Variations of SOC in the five LULC classes were scale dependent (Fig. 4). In Sandhill, the largest variance components were found at finest scales (7 and 2 m), whereas variations at coarser scale, 22, 67 and > 200 m, were negligible. Hardwood Hammock and Forest had the most variation at the coarsest scale (> 200 m) and the finest scale (2 m) also accounted for a large portion of the overall SOC variation. The variation at 7 m was also noticeable compared with the variations at 22 m and 67 m which were almost negligible. In Pineland, the finest scale variation was markedly greater than the other scales indicating most variation in Pineland occurred within 2 m. In Improved Pasture, the change of variation across scales was less prominent than those in other LULC classes. The relatively large variation stemmed from the coarsest scales (67 and > 200 m) followed by the finest scale variation. Similar to Sandhill and Pineland, Dry Prairie had the largest variance components at 2 and 7 m indicating that fine-scale variation dominated the variation under this LULC.

3.3. Semivariogram analysis of soil organic carbon

To complement the hierarchical analysis of variance, semivariogram analysis was conducted and the results are shown in Table 4 and Fig. 5. The sampled SOC of five sites showed distinct spatial structures. Improved Pasture showed strong spatial dependence (0.85), followed by Sandhill (0.37), Hardwood Hammock and Forest (0.37), and Dry Prairie (0.35) with moderate spatial dependence, whereas Pineland showed no clear spatial structure (Table 4). The effective spatial autocorrelation ranges of the five sites varied dramatically from site to site. Large ranges were observed in Hardwood Hammock and Forest (effective range: 391.2 m) and Improved Pasture (348.0 m) compared with small ranges in Sandhill (11.5 m) and Dry Prairie (23.4 m). These ranges were smaller than one-third of the sampling extent (~500 m) indicating stationarity in SOC within each of the fields. The smoothness parameter v is an important factor that not only measures the smoothness of the SOC processes but is also helpful in inferring short-range variation. The nugget variance is the combination of measurement errors and short-range spatial variation that occurs within the shortest sampling distance, i.e., 2 m in this study. The smoother the processes to accrete or deplete soil carbon, the less likely that short-range variation occurred. These processes of gaining or losing SOC under Hardwood Hammock and Forest were quite smooth (v = 99.0) approximating a Gaussian semivariogram model (Fig. 5). In contrast, Dry Prairie and Improved Pasture had relative rough processes of aggregating SOC stocks suggesting that short-range variation prevailed under these two LULC classes.

3.4. Average semivariogram

The average semivariogram of the five sites had a smooth parameter close to 0.5, which indicates an approximate exponential semivariogram model. It had a relatively long range of ~100 m and moderate spatial dependence of 0.4.

Table 5 compares the results of two ANOVA models – one using OLS that renders samples as independent of each other and the other one using GLS that accounts for the autocorrelation. The result shows that accounting for the autocorrelation significantly improved the model (p-value < 0.0001), which justifies the two degrees of freedom used to estimate the two autocorrelation parameters (range and nugget). Estimates of SOC mean for each LULC from OLS are the arithmetic mean of corresponding samples, while GLS gave slightly larger estimates except for Improved Pasture. More importantly, the results of post hoc pair-wise comparisons from the two models are distinctly different. Ordinary least squares tended to reject the null hypothesis based on underestimated variance of SOC in each LULC, therefore ordered the mean SOC of the five LULC as: Dry Prairie > Hardwood Hammock and Forest > Improved Pasture > Pineland > Sandhill. In contrast, GLS detected no difference between Pineland and Sandhill, Pineland and Improved Pasture, and Dry Prairie and Improved Pasture. Because the samples are spatially correlated, it is necessary to account for the spatial autocorrelation among

---

**Fig. 3.** The relationship between the variance and the squared mean of soil organic carbon (SOC): (a) raw SOC data; (b) variance-stabilized (log-transformed) SOC data. SH = Sandhill, HHF = Hardwood Hammock and Forest, PL = Pineland, IP = Improved Pasture, DP = Dry Prairie.
samples when analyzing the fixed effect of LULC on SOC stock and estimating the mean SOC stock in the LULC.

4. Discussion

To the best of our knowledge, this is the first study to compare the field-scale variability of SOC of prevalent LULC classes in the southeastern landscape of the U.S. under a unified framework (i.e., standardized protocol, sampling design and models). This study demonstrates that the spatial heterogeneity of SOC in different LULC showed both similarities and differences which may be attributed to both natural processes and anthropogenic interference.

The heterogeneity in SOC is known to be primarily controlled by soil water content, vegetation, climate, and topography (Guo et al., 2006; Liu et al., 2006; Vasques et al., 2010; Xiong et al., 2014). Human activities such as land use, have been recognized as an important soil forming factor that exerts great influence on SOC variation (Guo and Gifford, 2002; Maia et al., 2010; Murty et al., 2002). The controls of these factors on SOC variability are known to operate at various scales corresponding to the scales of variability of the factors per se (Grunwald et al., 2011; Vasques et al., 2012). For example, annual precipitation and temperature show variation at regional scale rather than at field-scale. Therefore, climate is a confounding factor affecting the field-scale variability of SOC in this study. The topography in Florida is generally flat with pronounced interspersed micro-depressions. This micro-topography can dictate soil water content (e.g., toposequence) and result in dramatically different SOC accumulation (Kamara et al., 2007).

4.1. Scale-dependent variability of SOC and the controlling factors

In Sandhill, the vegetation was dominated by an overstory of scattered longleaf pines (Pinus palustris) (~10 m average spacing), along with an understory of oak trees (Quercus). The space between trees was filled by various herbs and grasses. This vegetation pattern coinciding at about 10 m scale explains the largest variance components of SOC that was found at 7 and 2 m scale (Fig. 4) and the short range of autocorrelation (Fig. 5). There were limited land management practices that had occurred at the Sandhill site, suggesting that the native spatial structure of SOC that was formed by natural ecosystem processes was not disturbed. This indicates that the vegetation was the only major controller of the spatial heterogeneity of SOC. In Pineland, the vegetation was exclusively longleaf pine trees (~2 m average spacing) that were planted in form of commercial stands converted from non-native slash pines (Pinus elliottii) approximately five years ago. The commercial pine plantation operation had altered greatly the native spatial structure of SOC under pine forest, which is evidenced by the lack of spatial autocorrelation at the Pineland site (Fig. 5). The

![Fig. 4](image-url) Components and accumulated components of variance values for soil organic carbon at 0–20 cm depth in five land use and land cover classes estimated by restricted maximum likelihood. Variance components with * under them are significant at level of 0.05.

**Table 4**

<table>
<thead>
<tr>
<th>Sites</th>
<th>Nugget $c_0$ (log[kg m$^{-2}$])$^a$</th>
<th>Partial sill $c_1$ (log[kg m$^{-2}$])$^a$</th>
<th>Range parameter $r$ (m)</th>
<th>Smoothness $\nu$</th>
<th>Spatial dependence$^a$</th>
</tr>
</thead>
<tbody>
<tr>
<td>SH</td>
<td>0.043</td>
<td>0.025</td>
<td>1.5</td>
<td>4.3</td>
<td>0.37</td>
</tr>
<tr>
<td>HHF</td>
<td>0.039</td>
<td>0.023</td>
<td>11.3</td>
<td>99.0</td>
<td>0.37</td>
</tr>
<tr>
<td>PL</td>
<td>0.071</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>IP</td>
<td>0.010</td>
<td>0.057</td>
<td>134.7</td>
<td>0.35</td>
<td>0.85</td>
</tr>
<tr>
<td>DP</td>
<td>0.034</td>
<td>0.018</td>
<td>9.4</td>
<td>0.32</td>
<td>0.35</td>
</tr>
<tr>
<td>Average</td>
<td>0.049</td>
<td>0.033</td>
<td>103.2</td>
<td>0.48$^b$</td>
<td>0.40</td>
</tr>
</tbody>
</table>

Abbreviations: SH = Sandhill, HHF = Hardwood Hammock and Forest, PL = Pineland, IP = Improved Pasture, DP = Dry Prairie.

$^a$ Spatial dependence is defined as partial sill/sill.
The largest variance component at 2 m indicates that the restored longleaf pine exerted influence on the spatial heterogeneity of SOC within 2 m. In Dry Prairie, the native vegetation was dominated by saw palmettos (Serenoa repens) (~1 m average spacing) sparsely intermixed with various grasses, sedges, herbs, and shrubs (e.g., Leucothoe fontanesiana, Aristida beyrichiana, Axonopus fissifolius, Vaccinium formosum). This site had been maintaining the natural conditions with limited human interferences and was subject to lightning-induced fires and seasonal flooding that controlled the ever-invading woody plants. Therefore, the short range of spatial autocorrelation and the largest variance components found at the finest scales (2 and 7 m) were primarily in accordance to the spatial distribution pattern of saw palmettos. In Hardwood Hammock and Forest the vegetation was more diverse in species than the other LLUC and the spatial distribution of species coincided with the soil moisture regimes—hardwoods predominantly found in wetter areas and pine trees in well-drained areas. This spatial pattern of

Fig. 5. Cressie-Hawkins robust estimators of semivariogram (dots) and Matérn semivariogram (solid lines) with restricted maximum likelihood for soil organic carbon at 0–20 cm depth in five land use and land covers.
vegetation species accounted for the large variation at the coarsest scale (> 200 m) and the influence of individual trees explained the large variation at 2 m. In Improved Pasture, the soil was heavily managed. The site had been cleared, tilled, reseeded with forage grass and applied with fertilizer periodically. Besides, some areas were more heavily frequented by cattle flocks than other areas, which resulted in spatial variation of manure input and grass consumption. These practices may have determined the strong spatial structure of SOC observed in this study, similar to the findings by (Zhang et al., 2012).

The spatial structures of the topsoil SOC under five common LULC in Florida from this study show similarities compared to other studies. Rossi et al. (2009) applied a similar nested sampling design to investigate the spatial variation of topsoil SOC in tropical forests in Southeastern Tanzania. They also found that under pine plantation, there was no spatial autocorrelation due to human management activities, while the spatial structures of SOC in other natural forests were pronounced. Their results also suggest that in forests, where the vegetation composition and structure were more complex, the ranges of spatial autocorrelation were larger, which is confirmed by this present study (Hardwood Hammock and Forest). Cambardella et al. (1994) studied the field-scale variogram of SOC in two farms of Central Iowa and observed similar spatial structure to that of Improved Pasture from this study with observed ranges between 104 and 129 m with strong spatial dependence. This may be because the agricultural practices, such as tilling, weed control and fertilizing, are similar between the two studies and resulted in similar impact on the spatial structure of SOC. Cerri et al. (2004) used a systematic sampling design with a 25 m grid to study the spatial variation of soil total carbon in an Amazon pasture and also observed relatively large nugget effects accounting for 85 and 73% of the overall variances at 0–10 cm and 10–20 cm respectively. Their result suggested that large variability existed either within 25 m and the measurement error. These similarities of SOC spatial structures from distinct studies suggest that the spatial structure model obtained at the five sites may apply to other sites with the same or similar land use and management practices at field scale (~500 m).

4.2. Average semivariogram and its utilities

The average semivariogram (Fig. 5) derived by pooling the data of five sites together provides a general picture of the field-scale spatial structure of SOC. The model is comparable to the average SOC semivariogram reported by McBratney and Pringle (1999) who gathered a set of field-scale semivariograms from nine studies and obtained an average semivariogram for soil carbon. Despite the difference in the models used to fit the semivariogram, the spatial structure is fairly similar between the two studies. This similarity suggests the generality in the average semivariogram model of SOC at the scale of ~500 m regardless of the different edaphic and environmental conditions between the two studies. It has been well known that SOC follows a log-normal distribution and log-transformation is a common practice to normalize SOC data (Orton et al., 2014; Vasques et al., 2012). Furthermore, there exists proportionality between the variance and the squared mean of SOC (Fig. 3). The proportionality ensures that log-transformed SOC data has approximately constant variance (Carroll and Ruppert, 1988). This variance-invariant property suggests that SOC should display similar overall variance (sill) of 0.067 ± 0.012 (log[kg m⁻²])² at field scale (< 500 m) irrespective of LULC, which explains an important facet of the generality of the SOC average semivariogram derived in this study. The SOC average semivariogram can be readily used as prior knowledge of the soil carbon variability at field scale.

Another important utility of the average semivariogram is in modeling the effects of various environmental factors and forcings (fixed effects) on SOC. Any spatial sampling design can barely avoid the autocorrelation among samples, especially at field scale. Therefore, it is desirable to account for the spatial autocorrelation among samples when estimating the parameters of the fixed effects. The average semivariogram is of great value for more accurate modeling of soil processes.

5. Conclusions

This study investigated the field-scale variability (< 500 m) of SOC at 0–20 cm in five prevalent LULC classes in Florida, U.S. under a uniform framework — same sampling design, size, sample support, and data analysis methods — allowing an intra- and inter-comparison among fields. The field-scale variability of SOC was scale-dependent and tightly coupled to LULC types and the vegetation composition and structure. The average semivariogram depicted the overall spatial variability of SOC at field scale with the range parameter of ~100 m, moderate spatial dependence (0.4) and a smoothness parameter of ~0.5. The variance-invariant property of log-transformed SOC indicates that SOC displays similar overall variance (sill) of 0.067 ± 0.012 (log[kg m⁻²])² at field scale (< 500 m) irrespective of LULC. These findings do not only serve as prior knowledge of field-scale variability of SOC for precision agriculture and land management, but also facilitate better understanding of the scale-dependent fine-scale variability of SOC across larger soilscales.

Acknowledgements

This study was funded as part of a larger USDA-CSREES-NRI grant award 2007-35107-18368 ‘Rapid Assessment and Trajectory Modeling of Changes in Soil Carbon across a Southeastern Landscape’ (National Institute of Food and Agriculture (NIFA) — Agriculture and Food Research Initiative (AFRI)). This project is a Core Project of the North American Carbon Program. We are grateful to Dr. Murray Lark (British Geological Survey, UK) for his help in optimizing the spatially nested sampling scheme. We also thank Ce Yang, Christopher Wade Ross, Bajing Cao, Jongsung Kim, and Pascha Chaikaew (University of Florida) for their assistance in field sampling and laboratory analyses.