Influence of the spatial extent and resolution of input data on soil carbon models in Florida, USA

Gustavo M. Vasques,[1,2] S. Grunwald,[1] and D. Brenton Myers[1,3]

Received 7 February 2012; revised 19 June 2012; accepted 13 September 2012; published 23 October 2012.

[1] Understanding the causes of spatial variation of soil carbon (C) has important implications for regional and global C dynamics studies. Soil C predictive models can identify sources of C variation, but may be influenced by scale parameters, including the spatial extent and resolution of input data. Our objective was to investigate the influence of these scale parameters on soil C spatial predictive models in Florida, USA. We used data from three nested spatial extents (Florida, 150,000 km^2; Santa Fe River watershed, 3,585 km^2; and University of Florida Beef Cattle Station, 5.58 km^2) to derive stepwise linear models of soil C as a function of 24 environmental properties. Models were derived within the three extents and for seven resolutions (30–1920 m) of input environmental data in Florida and in the watershed, then cross-evaluated among extents and resolutions, respectively. The quality of soil C models increased with an increase in the spatial extent (R^2 from 0.10 in the cattle station to 0.61 in Florida) and with a decrease in the resolution of input data (R^2 from 0.33 at 1920-m resolution to 0.61 at 30-m resolution in Florida). Soil and hydrologic variables were the most important across the seven resolutions both in Florida and in the watershed. The spatial extent and resolution of environmental covariates modulate soil C variation and soil-landscape correlations influencing soil C predictive models. Our results provide scale boundaries to observe environmental data and assess soil C spatial patterns, supporting C sequestration, budgeting and monitoring programs.

Citation: Vasques, G. M., S. Grunwald, and D. B. Myers (2012), Influence of the spatial extent and resolution of input data on soil carbon models in Florida, USA, *J. Geophys. Res.*, *117*, G04004, doi:10.1029/2012JG001982.

1. Introduction

[2] The global soil carbon (C) pool including wetlands and permafrost (3,250 Pg C) is about five times the biotic pool (650 Pg C) and about four times the atmospheric pool (780 Pg C) [*Field et al.*, 2007], which highlights the importance of soils to store C and mitigate global warming. The geographic distribution of this global soil C resource varies according to numerous soil-forming factors, including other soil properties, climate, topography, parent material, vegetation, land use and human influence [*Jenny*, 1941; *Gower et al.*, 1997; *Houghton*, 2000; *Jobbágy and Jackson*, 2000; *Post and Kwon*, 2000; *Guo and Gifford*, 2002; *Florinsky et al.*, 2002; *McBratney et al.*, 2003; *Tan et al.*, 2004; *Masek and Collatz*, 2006; *Pärtel et al.*, 2008; *Desai et al.*, 2010; *Grunwald et al.*, 2011; *Liao et al.*, 2012]. Understanding how these factors account for soil C variation is critical for proposing land use and management strategies to mitigate soil C loss and avoid further depletion of this valuable resource.

[3] The assessment of soil C stocks depends on the spatial distribution of the aforementioned soil-forming factors and on how each factor contributes to soil C accretion or depletion. Thus, soil-landscape correlations can be derived between soil C and environmental properties (i.e., soil-forming factors) to estimate the amount and distribution of soil C. For example, *Chaplot et al.* [2010] observed correlations between soil organic C (SOC) and land use, annual rainfall and latitude, regionally, and distance to the stream and slope, at the hillslope level, in Laos (230,566 km^2), whereas *Martin et al.* [2011] found the most important variables to estimate SOC in France (541,060 km^2) to be soil clay content, land use and monthly precipitation, among other regional studies [e.g., *Arrouays et al.*, 2001; *Henderson et al.*, 2005; *Meersmans et al.*, 2008; *Phachomphon et al.*, 2010]. More locally, correlations have been found, for example, between: soil C and elevation, aspect, vegetation index and solar radiation, in a 2,800-ha watershed in southwestern Idaho, USA [*Kunkel et al.*, 2011]; SOC and slope, in a 6,000-ha area in southwestern France [*Arrouays et al.*, 1998]; SOC and slope, slope curvature and other topographic variables, in a 1,500-ha watershed in eastern Kentucky, USA [*Thompson and Kolka*, 2005]; and indirectly through a soil

[1]Soil and Water Science Department, University of Florida, Gainesville, Florida, USA.
[2]Now at the National Center for Soil Research, Brazilian Agricultural Research Corporation, Rio de Janeiro, Brazil.
[3]Now at the Agricultural Research Service, U.S. Department of Agriculture, Columbia, Missouri, USA.

Corresponding author: G. M. Vasques, National Center for Soil Research, Brazilian Agricultural Research Corporation, Rua Jardim Botânico 1024, Rio de Janeiro, RJ 22460-000, Brazil. (gustavo@cnps.embrapa.br)

0148-0227/12/2012JG001982

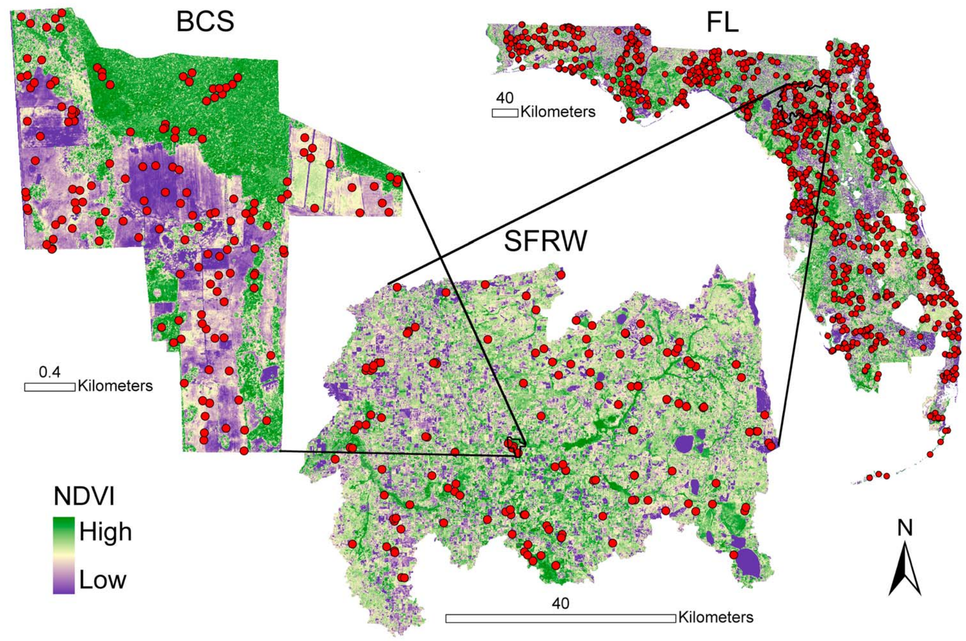

Figure 1. Sampling locations at the three nested study areas (spatial extents). BCS is the University of Florida Beef Cattle Station, FL is Florida, and SFRW is the Santa Fe River watershed. Normalized difference vegetation index (NDVI) is shown in the background.

morphologic index between SOC and elevation above the stream bank, slope and upslope contributing area, in a 2-ha agricultural area in western France [*Chaplot et al.*, 2001].

[4] Characterization of these soil-landscape relationships is influenced by the scale at which soil C and environmental properties are measured. Scale is determined by multiple parameters, including the spatial extent (i.e., size of the study area) and spatial resolution (i.e., pixel size) [*Meentemeyer and Box*, 1987; *Woodcock and Strahler*, 1987; *Turner et al.*, 1989], which can be associated to processes acting at multiple spatial and temporal levels. Soil properties are susceptible to spatial scale influences [*Burrough*, 1983; *Grunwald et al.*, 2011], which should be acknowledged when modeling over large regions, but are rarely incorporated in soil C models.

[5] For instance, it has been shown for select soil properties that their relationship with other soil and environmental properties changes as a function of scale [*Bourennane et al.*, 2003; *Corstanje et al.*, 2007; *Pringle and Lark*, 2007; *Martin and Bolstad*, 2009]. For example, *Corstanje et al.* [2007] observed scale-dependent correlations between SOC and urease activity of about 0.10–0.89 as a function of the distance between observations and land use. *Martin and Bolstad* [2009] observed differences in the variation of soil moisture, temperature and respiration among three scales ranging from intra-site (<1 m) to landscape (>1000 m), as indicated by their standard deviations. Thus, it can be expected that the capacity of environmental properties to explain soil variation changes as a function of scale. This would imply that the applicability of a predictive model derived for soil C (or another soil property) is restricted to the range of scale parameters that is specific to the study region and soil and environmental properties involved. Although some research has been presented on the effect of scale on environmental models [*Bian and Walsh*, 1993; *Mykrä et al.*, 2008; *Mackay et al.*, 2010], the influence of the extent and resolution of input data on predictive soil models has received little attention [*Chaplot et al.*, 2000; *Thompson et al.*, 2001].

[6] The aim of this study was to test the influence of scale parameters on soil total C (TC) models in Florida, USA. Specifically, our objectives were to evaluate: (a) the influence of the spatial extent on TC models within three nested extents; (b) the transferability of TC models among the three extents; (c) the influence of the spatial resolution of input data on TC models within two extents; and (d) the transferability of TC models among the seven resolutions of input data within the two extents.

[7] We hypothesized that the TC models were sensitive to the spatial extent and resolution of input data (objectives (a) and (c)), in other words, that the TC models were scale-dependent. In spite of the specificity of the models to the scale parameters, we also expected that similar soil-landscape relationships occur across scales and evaluated this by applying derived models to other scales (objectives (b) and (d)).

2. Material and Methods

2.1. Study Areas, Sampling Designs and Laboratory Methods

[8] Our study was conducted at three nested study areas in Florida, USA, including the whole state (FL; ~150,000 km^2), the Santa Fe River watershed (SFRW; ~3,585 km^2) in the north-central portion of FL, and the University of Florida Beef Cattle Station (BCS; ~5.58 km^2) in the central portion of the SFRW (Figure 1). We collected soil C data from these three extents to allow testing of hypotheses as outlined above.

2.1.1. Florida

[9] Florida is located in the subtropical climatic zone between latitudes 24.55 and 31.00 N, and longitudes 80.03

and 87.63 W. Mean annual precipitation and temperature are 1,373 mm and 22.3°C, respectively. Florida soils were formed mainly in the Quaternary period from continental and marine sediments under the influence of karst terrain and oscillating water tables and include Spodosols (32%), Entisols (22%), Ultisols (19%), Alfisols (13%) and Histosols (11%) [*Natural Resources Conservation Service* (*NRCS*), 2006b].

[10] Land use/land cover (LULC) consists mainly of wetlands (28%), pinelands (18%), and urban or barren lands (15%), whereas agriculture, rangelands and improved pasture occupy 9%, 9% and 8% of FL, respectively [*Stys et al.*, 2003]. The topography is relatively flat, with elevations below 114 m and 0% to 5% slopes across most of FL.

[11] Field sampling in FL took place from 1965 to 1996 as part of the Florida Soil Characterization project (http://flsoils.ifas.ufl.edu), as described in *Vasques et al.* [2010b]. Representative sampling sites were chosen ad hoc for the purposes of soil survey within each county with the help of aerial photographs, map unit delineations and supporting maps. A total of 1,288 soil profiles (Figure 1) were visited and described by horizon to a depth of 2 m or more, among which 1,037 sites with SOC measurements were used in this study.

[12] Soil samples were air-dried and sieved (2 mm). In mineral horizons, SOC was measured using the Walkley-Black modified acid-dichromate method (WB). In organic horizons, SOC was calculated by multiplying soil organic matter (SOM) measured by loss on ignition (LOI) by the van Bemmelen factor (0.58) [*NRCS*, 1996].

2.1.2. Santa Fe River Watershed

[13] The SFRW is located between latitudes 29.63 and 30.21 N, and longitudes 82.88 and 82.01 W (Figure 1). Dominant soil orders include Ultisols (47%), Spodosols (27%) and Entisols (17%) [*NRCS*, 2006a], while LULC include primarily pinelands (30%), wetlands (14%), improved pasture (13%), rangelands (13%) and upland forests (13%) [*Stys et al.*, 2003]. Elevations range from 2 to 92 m, and slopes from 0% to 5% in the watershed.

[14] A stratified random design based on land use and soil order combinations was used to select 130 sampling sites (Figure 1). At each site, composite soil samples were collected within a 2-m radius at fixed depths (0–30, 30–60, 60–120 and 120–180 cm) using an auger. Collected samples were air-dried, sieved (2 mm) and ball-milled. Total C was measured by high temperature combustion (HTC) on a FlashEA 1112 Elemental Analyzer (Thermo Electron Corp., Waltham, Mass.).

2.1.3. University of Florida Beef Cattle Station

[15] The BCS is located within the SFRW between latitudes 29.91 and 29.94 N, and longitudes 82.47 and 82.51 W (Figure 1). Soils are mostly dominated by Ultisols (78%), followed by Entisols (13%) and Inceptisols (5%) [*NRCS*, 2006a], whereas LULC include primarily improved pasture (39%), wetlands (25%) and rangelands (13%) [*Stys et al.*, 2003]. The topography consists of flat to slightly undulating slopes of up to 5% with elevations ranging from 13 to 43 m.

[16] Similar to the SFRW, sampling sites were located according to a random design stratified by land use and soil order, totaling 152 sites. Soils were sampled at fixed depths (0–30, 30–60, 60–120 and 120–180 cm), air-dried and sieved (2 mm). Soil organic matter was measured by LOI and multiplied by the van Bemmelen factor to obtain SOC.

2.2. Calculation of Profile Soil Total Carbon at 0–100 cm

[17] To harmonize soil C analytical methods, WB- and LOI-SOC measurements were converted to equivalent HTC-TC. Representative TC samples from mineral and organic horizons of the FL data set were chosen by randomly selecting about 14 samples from each decile of the frequency distribution of SOC, totaling 144 samples. The representativeness of the selected samples was confirmed by Welch's t tests [*Welch*, 1947] against mineral (t = 0.80; p-value = 0.42) and organic (t = −0.12; p-value = 0.91) samples from the complete FL data set, respectively, calculated from ln-transformed data.

[18] Using these samples, conversion factors were obtained by fitting simple linear regressions crossing the origin to estimate HTC-TC as a function of WB-SOC (R^2 = 0.94) or LOI-SOC (R^2 = 0.97), according to, respectively:

$$TC_{HTC} = 0.98 \times SOC_{WB}, \tag{1}$$

and

$$TC_{HTC} = 0.90 \times SOC_{LOI}, \tag{2}$$

where TC_{HTC} is TC measured by HTC, in %, SOC_{WB} is SOC measured by WB, in %, and SOC_{LOI} is SOC derived by multiplying LOI-SOM by the van Bemmelen factor, in %.

[19] A consistent TC data set across the three study areas was then compiled by calculating TC concentration within 0–100 cm as the depth-weighted average TC across horizons or depth intervals, according to:

$$TC = \sum_{i=1}^{n} TC_i \times D_i \Big/ \sum_{i=1}^{n} D_i, \tag{3}$$

where TC is soil total C at 0–100 cm, in %; TC_i is soil total C at the i-th horizon or depth interval, in %; D_i is the depth of the portion of the i-th horizon or depth interval constrained within 0–100 cm, in cm; and n is the number of horizons or depth intervals containing at least a portion within 0–100 cm.

2.3. Regression Modeling of Soil Total Carbon

[20] The influence of the spatial extent and resolution of input data on TC predictive models was tested using stepwise multiple linear regression (SMLR), with a F probability of 0.05 for including and removing variables. Predictive models for TC were based on the *SCORPAN* [*McBratney et al.*, 2003] conceptual model of soil formation, according to:

$$TC = f(s, c, o, r, p, a, n), \tag{4}$$

where TC is soil total C, s represents soil, c represents climate, o represents organisms including human activity, r represents relief, p represents parent material, a represents age (or time), and n represents spatial position. They were derived using ln-transformed TC (LnTC), which approximated a normal distribution.

[21] Samples from each study area were randomly separated into a training set (∼70%) used to derive the models and an independent validation set (∼30%) to validate the models. The total number of training and validation samples was 696 and 341 in FL, 93 and 37 in the SFRW, and 106 and 46 in the BCS, respectively.

[22] The SMLR models of LnTC were derived as a function of twenty four collocated environmental properties

Table 1. Environmental Properties Used as Explanatory Variables in the Stepwise Multiple Linear Regression Models of ln-Transformed Soil Total Carbon

SCORPAN Factor	Properties	References
Soil	Soil clay, sand and silt contents, in %; pH in 1:1 water; Available water capacity, in cm cm^{-1}; Saturated hydraulic conductivity, in μm s^{-1}	*NRCS* [2006a]
	Soil taxonomic order: Alfisol, Entisol, Spodosol, Ultisol, and Other (Histosol + Inceptisol + Mollisol + Vertisol)	*NRCS* [2006a]
	Soil drainage class: Poorly drained, Somewhat poorly drained, Moderately well drained, Well drained, and Excessively drained	*NRCS* [2006a]
	Soil hydrologic group: A, B, C, and D	*NRCS* [2006a]
Organisms	Land use/land cover: Agriculture, Grassland, Pineland, Urban or barren land, Wetland, and Upland vegetation (Forest + Scrub + Coastal + Exotic vegetation)	*Stys et al.* [2003]
	Landsat ETM+ bands 1, 2, 3, 4, 5 and 7, in digital number; Tasseled cap indices 1, 2 and 3, in digital number; Normalized difference vegetation index	*Stys et al.* [2003]
Relief	Elevation, in m; Slope, in %; Compound topographic index	*USGS* [1999]
	Aspect: East-, West-, North-, and South-facing slope	*USGS* [1999]

represented as individual geographic information system layers, according to:

$$TC[x,y] = \left(\beta_0 + \sum_{i=1}^{p} \beta_i \times F_i[x,y]\right) + e, \qquad (5)$$

where TC is soil total C, x and y are geographic coordinates, β_0 is the model intercept, β_i is the regression coefficient of the i-th selected explanatory environmental property, F_i is the i-th selected explanatory environmental property, p is the number of selected explanatory environmental properties, with $i = 1, 2, \ldots, p$, and e is the model residual.

[23] The environmental properties included soil survey data (nine polygon layers, scale of 1:24,000), digital elevation model (DEM) and topographic derivatives (four raster layers, 30-m resolution), LULC data (one raster layer, 30-m resolution), and reflectance data and derivatives (ten raster layers, 30-m resolution) obtained from a histogram-matched mosaic of Landsat ETM+ (Enhanced Thematic Mapper Plus) images covering FL (Table 1). Categorical variables (e.g., soil order and LULC) were converted to indicator (i.e., binary) variables to be included in the models. Model accuracy was evaluated using the coefficient of determination (R^2) calculated using the training (R_t^2) or validation set (R_v^2), respectively.

2.3.1. Influence of the Spatial Extent on Soil Total Carbon Models

[24] To prepare the data matrix for modeling, all environmental properties were converted to 30-m-resolution raster layers and extracted to the TC point observations. At each extent (FL, SFRW and BCS), a SMLR model of LnTC was derived using the training set (70% of data) and validated using the validation set (30% of data), respectively. The transferability of the LnTC models among the three extents was evaluated by applying the model derived at one specific extent to the validation data of the other two extents. A general framework of the methodology is presented in Figure 2.

2.3.2. Influence of the Resolution of Input Data on Soil Total Carbon Models

[25] The influence of the resolution of input data on LnTC models was tested in FL and in the SFRW. The twenty four environmental properties (Table 1) were resampled (i.e., aggregated) to seven resolutions (30, 60, 120, 240, 480, 960 and 1920 m) using either bilinear convolution (i.e., averaging within a 2 × 2 pixel window) for continuous properties or nearest neighbor assignment for categorical properties. Doubling the resolution at every resample level caused the new value of the property to match the average of the four cells included in the resample window, assuring data consistency across resolutions.

[26] For each resolution, all environmental properties were extracted to the TC point observations. Based on these data, for each resolution a SMLR model of LnTC was derived using the training set (70% of data) and validated using the validation set (30% of data), respectively. The transferability of the LnTC models among the seven resolutions of input data was evaluated by applying the model derived for one specific resolution to the validation data of the other six resolutions.

3. Results and Discussion

3.1. Descriptive Statistics

[27] The FL data set had the largest TC range (0.03%–54.59%) and the most variable TC among the three extents (Table 2), which was expected since FL covers a larger area that encompasses the other two extents. For the same reason, the SFRW had a larger TC range (0.12%–17.03%) and more variable TC than the BCS (0.48%–8.65%). The frequency distributions of the training and validation sets had similar properties in all cases, with the range of the training set encompassing the range of the validation set.

3.2. Influence of the Spatial Extent on Soil Total Carbon Models

[28] The spatial extent influenced the selection of explanatory properties and quality of SMLR models of LnTC (Table 3). All LnTC models were significant (p-value $\sim$ 0) and the most accurate models were derived in FL ($R_t^2 = 0.61$) and in the SFRW ($R_t^2 = 0.50$). However, independent validation of the models in the respective extent was considerably better in FL ($R_v^2 = 0.60$) than in the other extents. In contrast, LnTC models derived in the BCS had the lowest quality ($R_t^2 = 0.10$; $R_v^2 \sim 0.00$).

[29] Florida was the largest extent and encompassed a higher variation of LnTC and explanatory environmental properties than the SFRW and BCS. In addition, the number of observations in FL was eight times larger than in the other areas (Table 2), which provided a more representative sample of LnTC and its correlations with environmental properties. Nevertheless, sampling density in FL (0.007 samples km^{-2}) was smaller than in the other extents (SFRW, 0.04 samples km^{-2}; BCS, 27.2 samples km^{-2}). Conversely, the BCS was

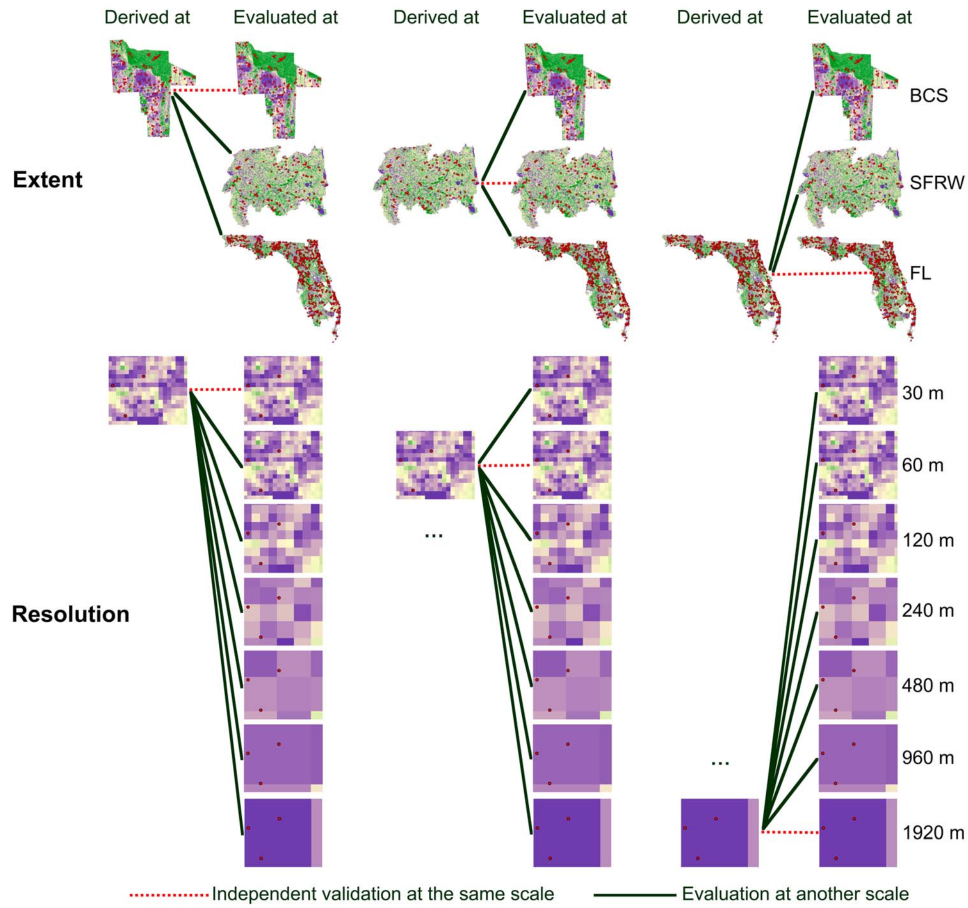

Figure 2. Overview of the framework used to test the influence of the spatial extent and resolution of input data on the stepwise multiple linear regression models of ln-transformed soil total carbon. BCS is the University of Florida Beef Cattle Station, FL is Florida, and SFRW is the Santa Fe River watershed.

the smallest extent, constraining the variation of LnTC and co-located environmental properties, which reflected in a poorer model relative to the larger extents.

[30] Soil and hydrologic properties were the most important variables in the FL and SFRW models (Table 3). Vegetation properties were included in all extents and normalized difference vegetation index (NDVI) was the only variable selected to explain LnTC patterns in the BCS. The selection of soil properties in the FL model suggests that the regional spatial distribution of LnTC is mainly controlled by soil patterns and underlying processes that regulate C dynamics (e.g., aggregation, humification, decomposition, among others). In contrast, in the BCS (i.e., locally), absence of soil properties in the model suggests that another factor, perhaps human-induced vegetation (i.e., land use) control, is dictating TC dynamics. In effect, the BCS has been managed for pasture and forage production in the last decade involving fertilization and plowing.

[31] The literature suggests from site-specific studies that vegetation and land use show a close relationship with soil C [*Ross et al.*, 1999; *Guo and Gifford*, 2002; *Pärtel et al.*, 2008]. At the field scale (BCS), our study confirms this land use/vegetation-soil C association; however, vegetation properties were less significant to infer on soil C at larger escalating regional extents (SFRW and FL). These findings have important implications for assessing soil C dynamics regionally as they challenge the established expectation that land use is the primary source of variation affecting TC.

Table 2. Descriptive Statistics of Soil Total Carbon (TC) and ln-Transformed TC (LnTC) at the Three Nested Study Areas (Spatial Extents)[a]

Statistics	FL	SFRW	BCS
	TC (%)		
N	1,037	130	152
Mean	1.75	0.85	1.90
Standard deviation	6.08	1.78	1.13
Median	0.42	0.51	1.74
Range	0.03–54.59	0.12–17.03	0.48–8.65
Skewness	5.97	7.26	2.62
	LnTC (ln%)		
N	1,037	130	152
Mean	−0.65	−0.58	0.51
Standard deviation	1.11	0.68	0.51
Median	−0.87	−0.67	0.56
Range	−3.56–4.00	−2.09–2.83	−0.73–2.16
Skewness	1.79	2.05	0.20

[a]FL is Florida, SFRW is the Santa Fe River watershed, and BCS is the University of Florida Beef Cattle Station. *N* is the number of observations.

Table 3. Stepwise Multiple Linear Regression Models of ln-Transformed Soil Total Carbon Derived at Different Spatial Extents With 30-m Resolution

		Extents[b,c]					
		FL		SFRW		BCS	
		Regression Coefficients					
SCORPAN Factor	Properties[a]	Unstd	Std	Unstd	Std	Unstd	Std
	Intercept	2.89		−1.69		0.13	
Soil	ClayCnt	−0.04	−0.33				
	SandCnt	−0.03	**−0.57**				
	AWC			14.86	**0.50**		
	Ksat	−0.003	−0.12				
	PoorDrn	0.51	0.23				
	SwPoorDrn	0.18	0.06				
	HydroGrpD	0.22	0.07				
	Alfisols	−1.04	−0.36				
	Entisols	−0.77	−0.27				
	Spodosols	−0.58	−0.23				
	Ultisols	−0.91	−0.37				
Organisms	Wetland			1.18	0.37		
	TC2	0.01	0.08				
	NDVI					1.14	**0.31**
		R^{2d}					
		Train	Val	Train	Val	Train	Val
		0.61	0.60	0.50	0.11	0.10	0.00

[a]Abbreviations in the “Properties” column are as follows: ClayCnt is soil clay content; SandCnt is soil sand content; AWC is soil available water capacity; Ksat is soil saturated hydraulic conductivity; PoorDrn is poorly drained soil; SwPoorDrn is somewhat poorly drained soil; HydroGrpD is soil hydrologic group D; TC2 is tasseled cap index 2; and NDVI is normalized difference vegetation index.
[b]FL is Florida; SFRW is the Santa Fe River watershed; BCS is the University of Florida Beef Cattle Station; Unstd is unstandardized; and Std is standardized.
[c]Numbers in bold indicate the most important explanatory variable in the model according to the standardized coefficient.
[d]R^2 is the coefficient of determination; Train is training; and Val is validation.

[32] Studies exploring the influence of the spatial extent on soil-landscape or soil C variation are rare [*McBratney*, 1998]. Among the few examples, the influence of the extent on landscape pattern metrics was investigated by *Wu et al.* [2002] and *Wu* [2004], who found a less predictable effect of the extent than of the resolution (further discussed in a later section in this paper) on landscape metrics.

3.3. Transferability of Soil Total Carbon Models Among Spatial Extents

[33] Predictive LnTC models were reasonably transferable between FL and the SFRW (R_v^2 = 0.40 in both cases), but not transferable to/from the BCS (Figure 3). A similar process relationship and an adequate overlap of the range of LnTC and environmental properties in the SFRW could explain the good model transferability between the SFRW and FL. In contrast, the BCS was not large and variable enough to adequately capture the regional variation of LnTC present in the larger extents, suggesting that a domain-specific subset of processes may be controlling LnTC.

[34] These results suggest that the SFRW is, to a certain extent, representative of the soil-landscape relationships found in FL. The LnTC models for FL and the SFRW selected very different predictors, but models cross-applied between these extents explained 40% of the LnTC variance in the other extent, respectively (Figure 3). Thus, in one direction, the correlation between soil properties and LnTC observed in FL had a counterpart in the SFRW; in the other direction, available water capacity (AWC) and an indicator variable for wetland (the only variables selected in the SFRW model) were sufficient to explain 40% of LnTC variation when applied in FL. These differences between models support the hypothesis that the scale of observation (i.e., the extent) is important for model making. Interestingly, the more complete model derived in FL adjusted better to the validation data in the SFRW than the model derived in the SFRW itself (Figure 3). This can be an artifact of the regression method used, which constrained the variables in the SFRW to only those that were significant at the 95% confidence level. Moreover, a more global model (FL) applied to a smaller nested extent lost predictive power, i.e., the R_v^2 declined from 0.60 (FL) to 0.40 (SFRW) to 0.04 (BCS), whereas the LnTC model developed in the SFRW (R_v^2 of 0.11) improved predictive capabilities to an R_v^2 of 0.40 at a larger extent (FL).

[35] Model transferability and the selection of explanatory variables by the models at escalating extents were related to differences in variances and correlations of soil and environmental properties. For example, properties which impart control on LnTC at the SFRW (AWC and wetland indicator) are replaced by more complex interactions of soil and hydrologic variables in FL. This demonstrates the sensitivity of LnTC models to increasing environmental variances, where the relationships between LnTC and *SCORPAN* properties, notably soil properties, depend on the spatial extent.

3.4. Influence of the Resolution of Input Data on Soil Total Carbon Models

[36] For both FL and the SFRW the resolution of input data influenced the selection of explanatory environmental properties and quality of SMLR LnTC models. In both areas the LnTC models were significant (p-value ~ 0) for all resolutions and the R_t^2 consistently decreased with an increase in resolution, as expected (Tables 4 and 5). By

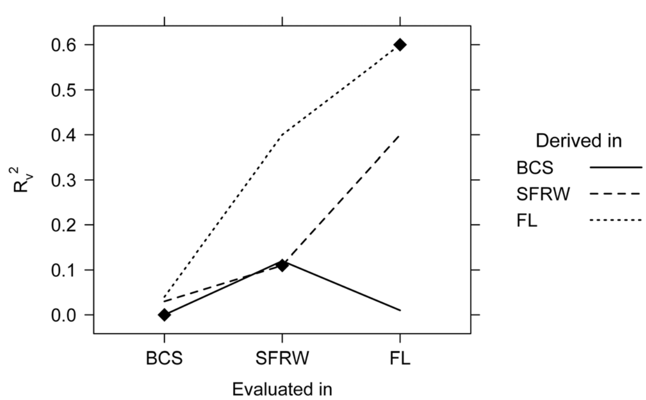

Figure 3. Evaluation of the stepwise multiple linear regression models of ln-transformed soil total carbon among spatial extents. BCS is the University of Florida Beef Cattle Station, FL is Florida, and SFRW is the Santa Fe River watershed. Diamonds indicate validation in the same extent that the model was derived.

Table 4. Stepwise Multiple Linear Regression Models of ln-Transformed Soil Total Carbon Derived for Different Resolutions of Input Data in Florida

		Resolutions[b,c]													
		Regression Coefficients													
SCORPAN		30 m		60 m		120 m		240 m		480 m		960 m		1920 m	
Factor	Properties[a]	Unstd	Std	Unstd	Std	Unstd	Std	Unstd	Std	Unstd	Std	Unstd	Std	Unstd	Std
	Intercept	2.89		3.25		3.81		3.06		3.08		1.30		1.03	
Soil	ClayCnt	−0.04	−0.33	−0.04	−0.32	−0.04	−0.32	−0.04	−0.31	−0.04	−0.31	−0.03	−0.22	−0.02	−0.16
	SiltCnt			−0.02	−0.10	−0.03	−0.15	−0.02	−0.11	−0.02	−0.14	−0.02	−0.09	−0.02	−0.10
	SandCnt	−0.03	**−0.57**	−0.04	**−0.69**	−0.04	**−0.75**	−0.04	**−0.63**	−0.04	**−0.68**	−0.04	**−0.54**	−0.04	**−0.58**
	AWC							3.15	0.12						
	Ksat	−0.003	−0.12	−0.003	−0.15	−0.004	−0.19	−0.003	−0.13	−0.005	−0.18	−0.004	−0.13		
	PoorDrn	0.51	0.23	0.53	0.24	0.39	0.17	0.23	0.10						
	SwPoorDrn	0.18	0.06												
	ModWellDrn													0.25	0.07
	WellDrn											−0.33	−0.10		
	HydroGrpB			−0.18	−0.08										
	HydroGrpD	0.22	0.07												
	Alfisols	−1.04	−0.36	−0.59	−0.20	−0.48	−0.17	−0.32	−0.11			−0.23	−0.08		
	Entisols	−0.77	−0.27	−0.25	−0.09										
	Spodosols	−0.58	−0.23											0.24	0.09
	Ultisols	−0.91	−0.37	−0.43	−0.17	−0.35	−0.14	−0.31	−0.12	−0.24	−0.09				
Organisms	Pineland														
	Urban/Barren							0.20	0.07						
	Wetland			0.32	0.09										
	B3					−0.01	−0.08	−0.03	−0.27						
	B7							0.02	0.19						
	TC2	0.01	0.08	0.01	0.08										
Relief	Slope											0.12	0.15	0.13	0.16
	CTI					0.02	0.07	0.03	0.10	0.04	0.16	0.10	0.35	0.10	0.36
	AspectE													−0.23	−0.07
	AspectN											0.23	0.07		
		R^2[d]													
		Train	Val	Train	Val	Train	Val	Train	Val	Train	Val	Train	Val	Train	Val
		0.61	0.60	0.61	0.63	0.60	0.53	0.53	0.44	0.42	0.40	0.36	0.33	0.33	0.25

[a]Abbreviations in the "Properties" column are as follows: ClayCnt is soil clay content; SiltCnt is soil silt content; SandCnt is soil sand content; AWC is soil available water capacity; Ksat is soil saturated hydraulic conductivity; PoorDrn is poorly drained soil; SwPoorDrn is somewhat poorly drained soil; ModWellDrn is moderately well drained soil; WellDrn is well drained soil; HydroGrpB is soil hydrologic group B; HydroGrpD is soil hydrologic group D; B3 and B7 are Landsat ETM+ bands 3 and 7, respectively; TC2 is tasseled cap index 2; CTI is compound topographic index; AspectE is east-facing slope; and AspectN is north-facing slope.
[b]Unstd is unstandardized, and Std is standardized.
[c]Numbers in bold indicate the most important explanatory variable in the model according to the standardized coefficient.
[d]R^2 is the coefficient of determination; Train is training; and Val is validation.

resampling the input raster layers into progressively larger pixel sizes, in essence we tested the aggregation behavior of environmental properties on modeling of LnTC. This aggregation has the effect of smoothing the variation of environmental predictors, which causes their correlation with LnTC (and thus, R^2) to decrease. In addition, soil and environmental properties that are more homogenous across the landscape show less reduction in variance than heterogeneously distributed properties which may impact model behavior (i.e., the selection of significant variables and the strength of the relationship modeled in the regression equation).

[37] At the FL extent, soil properties were prominently important in the LnTC models across spatial resolutions, sand content being the most important variable for all resolutions (Table 4). Hydrologic properties (e.g., compound topographic index and soil saturated hydraulic conductivity, drainage class and hydrologic group) were also selected as predictors at different resolutions, with regression coefficients indicating a positive association between the amount of C and amount of water in the soil. Properties related to vegetation (LULC and those derived from Landsat ETM+) were selected by the models up to a resolution of 240 m. This suggests that the spatial patterns of LnTC and vegetation were correlated up to a resolution of 240 m, beyond which they lost their connection due to over-aggregation. As previously discussed, this is different from our understanding of site-specific studies that advocate strong linkages between soil C and land use.

[38] At the regional scale, we expected vegetation properties to be highly correlated to LnTC. In this sense, the dominance of soil-related properties in the models could have masked the influence of vegetation on LnTC patterns. In effect, based on our experience [*Vasques et al.*, 2010a], regional patterns of soils in FL are related to vegetation patterns, but also originate from other patterns and processes that span across FL, including the distribution of parent materials and hydrologic gradients. The prevalence of soil properties in the models across spatial resolutions in FL underpins the finding that soil-soil correlations are stronger than soil-landscape correlations. In addition, soil properties derived from the Soil Survey Geographic (SSURGO) database, a polygon layer with a map scale of 1:24,000 [*NRCS*,

Table 5. Stepwise Multiple Linear Regression Models of ln-Transformed Soil Total Carbon Derived for Different Resolutions of Input Data in the Santa Fe River Watershed

		Resolutions[b,c]													
		30 m		60 m		120 m		240 m		480 m		960 m		1920 m	
		Regression Coefficients													
SCORPAN Factor	Properties[a]	Unstd	Std	Unstd	Std	Unstd	Std	Unstd	Std	Unstd	Std	Unstd	Std	Unstd	Std
	Intercept	−1.69		−1.74		−2.01		−2.10		−0.80		0.86		−0.58	
Soil	ClayCnt									0.06	0.25				
	AWC	14.86	**0.50**	16.44	**0.49**	19.62	**0.58**	20.27	**0.59**						
	WellDrn											1.27	0.47		
	ExcessDrn							0.42	0.19					−0.57	**−0.23**
Organisms	Pineland													0.33	0.21
	Urban/Barren											0.36	0.18		
	Wetland	1.18	0.37	1.19	0.37	1.62	0.40								
	B5											−0.07	**−1.42**		
	B7											0.07	1.09		
	TC3									0.02	**0.35**				
Relief	AspectN			−0.42	−0.20										
		R^{2d}													
		Train	Val	Train	Val	Train	Val	Train	Val	Train	Val	Train	Val	Train	Val
		0.50	0.11	0.56	0.22	0.57	0.09	0.32	0.01	0.19	0.00	0.30	0.01	0.11	0.06

[a]Abbreviations in the “Properties” column are as follows: ClayCnt is soil clay content; AWC is soil available water capacity; WellDrn is well drained soil; ExcessDrn is excessively drained soil; B5 and B7 are Landsat ETM+ bands 5 and 7, respectively; TC3 is tasseled cap index 3; and AspectN is north-facing slope.
[b]Unstd is unstandardized, and Std is standardized.
[c]Numbers in bold indicate the most important explanatory variable in the model according to the standardized coefficient.
[d]R^2 is the coefficient of determination; Train is training; and Val is validation.

2006a], could be less sensitive to aggregation than Landsat ETM+ or DEM derived properties.

[39] In the SFRW, LnTC was primarily controlled by hydrologic patterns (Table 5). Up to a resolution of 240 m, AWC was the most important variable; for the 480-m resolution, tasseled cap index 3, commonly described as a terrain wetness index, substituted AWC in the model as a surrogate hydrologic property. For coarser resolutions, soil drainage class was selected instead of AWC. Selection of environmental properties to predict LnTC for fine resolutions could relate to short-range spatial patterns of LnTC as suggested by *Vasques et al.* [2012]. Accordingly, variables selected for coarse resolutions capture the long-range variation of LnTC associated with large-scale spatial patterns of the corresponding environmental properties.

[40] To visualize the impact of variable selection at different spatial resolutions on LnTC patterns, we produced seven maps showing predicted LnTC in the SFRW, based on the seven models, respectively (Figure 4). The general spatial trend of LnTC captured by the SMLR models for all resolutions up to 120 m was very similar and reflected mainly the distribution of AWC. However, for resolutions above 120 m some areas of high and low LnTC became less differentiable due to the smoothing caused by aggregating environmental data into larger pixels. For example, hot spot areas of LnTC (in red) disappeared in some parts at 240 m, and almost completely at resolutions greater than 240 m. At 480 and 960 m, the spatial distribution of LnTC still showed some resemblance with smaller resolutions, possibly due to the selected hydrologic variables and patterns thereof. Finally, at the coarsest resolution (1920 m), the distribution of LnTC became so generalized that the major spatial trends associated with environmental properties were no longer distinguishable in most parts. Furthermore, the location of high-C areas was no longer in accordance with the smaller resolutions, which would mis-guide adaptation of management to enhance C sequestration, soil health and conservation.

[41] Another effect caused by increasing the resolution was the loss of variability of LnTC in the output maps. For the finest resolutions up to 240 m, the output range of LnTC values was similar to the observed one, whereas for resolutions of 480 m and larger, the range of LnTC was reduced, especially in high-C areas, where LnTC values were underestimated by the models. Thus, using a resolution of input data larger than 240 m may considerably underestimate total LnTC content in the SFRW.

[42] In the SFRW and FL hydrologic patterns play an important role of controlling soil and other environmental properties. The level topography and high annual precipitation create widespread areas of wetlands across the state, where accumulation of TC is fostered by the relatively slower anaerobic decomposition of organic matter. This was confirmed by the coefficients of the hydrologic properties included in the LnTC models for multiple spatial resolutions. For example, in the SFRW, AWC and the indicator variable for wetland received positive coefficients. Similarly, in FL, regression coefficients for AWC, compound topographic index and indicator variables for poorly drained soil and wetland were positive. It was also confirmatory that soil saturated hydraulic conductivity received a negative coefficient in regression equations (decreasing hydraulic conductivity and reduced water percolation correlated with increasing SOC) (Tables 4 and 5).

[43] Other studies have shown that topographic attributes are influenced by the spatial resolution, with an effect on soil-landscape models. *Thompson et al.* [2001] compared DEMs and terrain derivatives with 10- and 30-m resolutions from two sources (field survey and United States Geological Survey (USGS), respectively) to estimate the depth of the surface

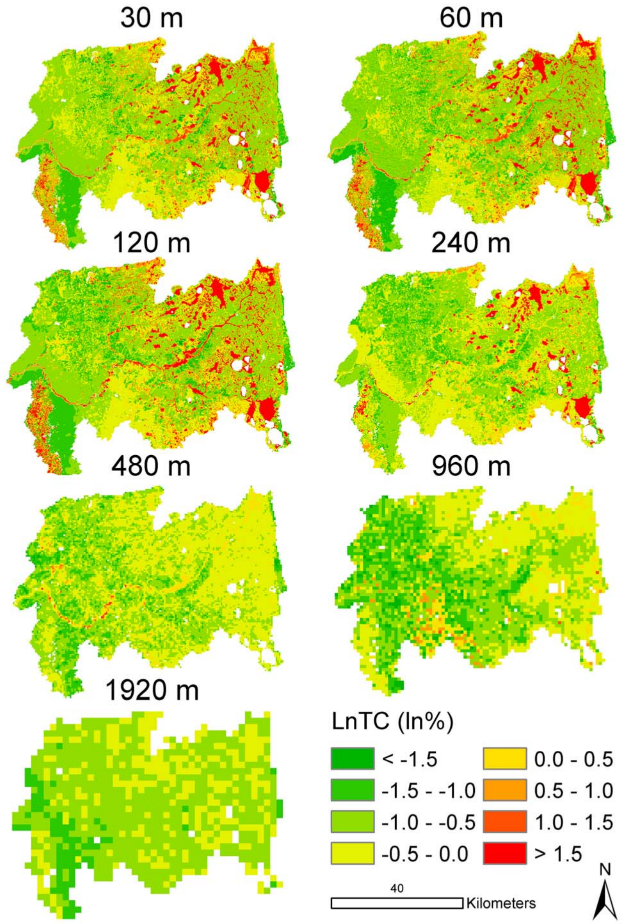

Figure 4. Estimated maps of ln-transformed soil total carbon (LnTC) from the stepwise multiple linear regression models derived for seven resolutions of input data, respectively, in the Santa Fe River watershed.

horizon (A horizon). Results based on the R_v^2 were contradictory, indicating a preference of the model for the 30-m resolution in the case of the field survey DEM and for the 10-m resolution in the case of the USGS DEM. *Wu et al.* [2008] observed changes in the correlations among soil and topographic attributes associated to changes in the resolution of the DEM and derivatives (4–30 m). They found, however, a general consistency in the sign of regression coefficients and significance of soil-topography relationships. *Schulp and Veldkamp* [2008] also observed changes in SOM-landscape correlations from 50- to 200- and 500-m resolutions, obtaining the best SOM predictive model at 500-m resolution, based on soil, groundwater and historical land use variables. *Chaplot et al.* [2000] observed a consistent decrease in the quality of topographic attributes with increases in spatial resolution from 10 to 20, 30 and 50 m, which led to a decreasing trend in the quality of estimated soil hydromorphic index similarly to our observations for LnTC in this study.

[44] The effect of the spatial resolution on topographic and landscape indices has also been recognized by a number of studies. For example, *Wu et al.* [2007] observed a deterioration of topographic index distributions at coarse resolutions caused by the smoothing effect on the DEM. *Vaze et al.* [2010] observed a superiority of higher-resolution DEMs to derive topographic indices for hydrological modeling. In terms of landscape indices, *Wu et al.* [2002] and *Wu* [2004] identified power, logarithmic, linear, staircase-like, and in some cases erratic relationships between the spatial resolution and diverse indices, arguing that the characterization of landscape patterns should be performed using multiscale approaches (in their case, scalograms). In a previous study, *Qi and Wu* [1996] showed the influence of the spatial resolution on indices of spatial autocorrelation, finding a decrease in the degree of spatial autocorrelation with an increase in pixel size.

[45] In Florida, a study in the Everglades [*Obeysekera and Rutchey*, 1997] tested multiple resolutions from 20 to 1000 m to describe land cover variability extracted from SPOT imagery using spatial indices with the aim to identify ideal scales to derive spatial models in the region. The authors observed an almost linear decrease in the diversity index when broadening the resolution, indicating loss of information as the pixel size increased. For instance, beyond 700 m, tree islands, very important ecological features, virtually disappeared from the images. Furthermore, they observed self-similarity only for resolutions in the range from 20 to 100 m, suggesting the adoption of a resolution smaller than 100 m to model the Everglades.

[46] Our study indicates that the spatial resolution of 60 m, according to the R_v^2, conveys adequate environmental information from multiple sources of variation to model LnTC. Moreover, going beyond a resolution of 240 m is not recommended due to mischaracterization of the main spatial patterns of LnTC.

3.5. Transferability of Soil Total Carbon Models Among Resolutions of Input Data

[47] As expected, the transferability of the LnTC models among resolutions of input data in both FL and the SFRW showed a general trend of decreasing R_v^2 from finer to coarser resolutions (Figures 5 and 6, respectively). Environmental information was lost due to aggregation from finer to coarser resolutions. Thus, when models derived for finer resolutions were evaluated at coarser resolutions, the variation of explanatory properties was underrepresented, tending to the mean values that resulted from the aggregation process. As a consequence, low and high LnTC predictions were biased toward the mean, thus degrading the Rv^2. In the opposite direction, models transferred better from coarser to finer resolutions because they were exposed to a wider range of variation of the explanatory properties, which was closer to the original variation from which the properties originated (i.e., were resampled from) in the first place.

[48] In both the SFRW and FL the highest R_v^2 were found when evaluating the models at the 30- and 60-m resolutions. As explained before, the 30- and 60-m resolutions carried the most detailed information of the environmental predictors relative to the coarser resolutions. In the case of 60 m, one round of resampling (from 30 to 60 m) still kept most of the information of the predictors to be incorporated in the model. In the cases where validation at the 60-m resolution was better than at the 30-m resolution, it is possible that some of the noise present in the 30-m resolution was smoothed out (i.e., removed) by resampling to 60 m, thus increasing the portion of useful correlation between LnTC and predictors. Currently the GlobalSoilMap.net consortium (http://www.globalsoilmap.net) plans to derive 100-m resolution maps of various soil

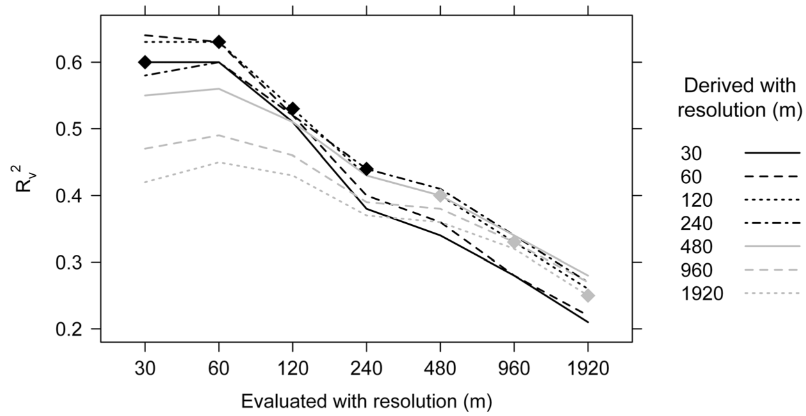

Figure 5. Evaluation of the stepwise multiple linear regression models of ln-transformed soil total carbon among resolutions of input data in Florida. Diamonds indicate validation at the same resolution that the model was derived.

properties across the globe. In Florida, this resolution would convey adequate information on soil and environmental properties for soil C and probably other assessments.

4. Conclusions

[49] Both the spatial extent and resolution of input data influenced the relationships between soil-forming environmental properties and LnTC, and thus the quality of LnTC predictive models. In general, larger extents and smaller resolutions produced more complex models, with more predictor variables than in smaller extents and larger resolutions, respectively. In most cases, soil hydrology was in great part responsible for the spatial distribution of soil C, suggesting the need to better characterize spatial hydrologic patterns in Florida. This would potentially be the case for other similar areas with level and smooth topography forming interspersed upland and wetland areas associated with soil C degradation and accretion.

[50] In the spatial extent analysis, while to some degree models at larger extents represented (i.e., could be applied at) smaller extents, in the opposite direction regional variations were inadequately accounted for by local models. This raises concerns about situations where regional to global inferences related to soil C dynamics are based solely on field or multi-field experiments. In effect, due to budget or labor constraints, soil C assessments are often limited to specific fields, LULC or soil types, imposing major limitations to transfer knowledge from smaller/finer to larger/coarser scales. Thus, knowing how C behaves regionally as a function of environmental properties and moreover how scale parameters affect these relationships is important to address global C issues.

[51] Furthermore, adequately assessing soil C in areas with C-rich soils is essential to support C sequestration, budgeting and monitoring policies. Thus, the spatial extent must not be prohibitively small and the resolution of input data not prohibitively large that important spots of C-rich soils become unapparent or misrepresented. However, it remains unclear whether a universal, transferable, multiscale soil C model that represents local to regional to global soil variation can be achieved, and under what scale boundaries such a model would satisfactorily work.

[52] The influence of scale parameters on spatial models of soil C, as identified in this study for Florida, can provide important information for the design of future soil C assessment campaigns and for establishing priorities for the

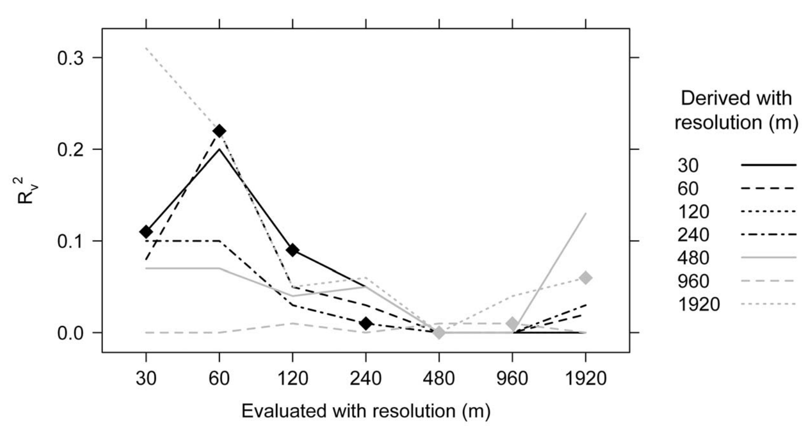

Figure 6. Evaluation of the stepwise multiple linear regression models of ln-transformed soil total carbon among resolutions of input data in the Santa Fe River watershed. Diamonds indicate validation at the same resolution that the model was derived.

collection and preparation of basic ancillary environmental spatial data to support regional-scale assessments of soil C.

[53] **Acknowledgments.** Funding for this study was provided by the United States Department of Agriculture (Cooperative Ecosystem Studies Unit, Natural Resources Conservation Service), by the National Research Initiative (Agriculture and Food Research Initiative, National Institute of Food and Agriculture) Competitive Grant 2007-35107-18368 (core project of the North American Carbon Program) and by the University of Florida Alumni Fellowship program. We also thank the Soil and Water Science Department at the University of Florida, the Natural Resources Conservation Service and the Florida Fish and Wildlife Conservation Commission for providing data used in this study.

References

Arrouays, D., J. Daroussin, J. L. Kicin, and P. Hassika (1998), Improving topsoil carbon storage prediction using a digital elevation model in temperate forest soils of France, *Soil Sci.*, *163*, 103–108, doi:10.1097/00010694-199802000-00003.

Arrouays, D., W. Deslais, and V. Badeau (2001), The carbon content of topsoil and its geographical distribution in France, *Soil Use Manage.*, *17*, 7–11, doi:10.1111/j.1475-2743.2001.tb00002.x.

Bian, L., and S. J. Walsh (1993), Scale dependencies of vegetation and topography in a mountainous environment of Montana, *Prof. Geogr.*, *45*, 1–11, doi:10.1111/j.0033-0124.1993.00001.x.

Bourennane, H., S. Salvador-Blanes, S. Cornu, and D. King (2003), Scale of spatial dependence between chemical properties of topsoil and subsoil over a geologically contrasted area (Massif central, France), *Geoderma*, *112*, 235–251, doi:10.1016/S0016-7061(02)00309-9.

Burrough, P. A. (1983), Multiscale sources of spatial variation in soil. I. The application of fractal concepts to nested levels of soil variation, *J. Soil Sci.*, *34*, 577–597, doi:10.1111/j.1365-2389.1983.tb01057.x.

Chaplot, V., C. Walter, and P. Curmi (2000), Improving soil hydromorphy prediction according to DEM resolution and available pedological data, *Geoderma*, *97*, 405–422, doi:10.1016/S0016-7061(00)00048-3.

Chaplot, V., M. Bernoux, C. Walter, P. Curmi, and U. Herpin (2001), Soil carbon storage prediction in temperate hydromorphic soils using a morphologic index and digital elevation model, *Soil Sci.*, *166*, 48–60, doi:10.1097/00010694-200101000-00008.

Chaplot, V., B. Bounthong, and C. Valentin (2010), Soil organic carbon stocks in Laos: Spatial variations and controlling factors, *Global Change Biol.*, *16*, 1380–1393, doi:10.1111/j.1365-2486.2009.02013.x.

Corstanje, R., R. Schulin, and R. M. Lark (2007), Scale-dependent relationships between soil organic carbon and urease activity, *Eur. J. Soil Sci.*, *58*, 1087–1095, doi:10.1111/j.1365-2389.2007.00902.x.

Desai, A. R., B. R. Helliker, P. R. Moorcroft, A. E. Andrews, and J. A. Berry (2010), Climatic controls of interannual variability in regional carbon fluxes from top-down and bottom-up perspectives, *J. Geophys. Res.*, *115*, G02011, doi:10.1029/2009JG001122.

Field, C. B., J. Sarmiento, and B. Hales (2007), The carbon cycle of North America in a global context, in *The First State of the Carbon Cycle Report (SOCCR): The North American Carbon Budget and Implications for the Global Carbon Cycle*, edited by A. W. King et al., pp. 21–28, Natl. Clim. Data Cent., Asheville, N. C.

Florinsky, I. V., R. G. Eilers, G. R. Manning, and L. G. Fuller (2002), Prediction of soil properties by digital terrain modeling, *Environ. Model. Softw.*, *17*, 295–311, doi:10.1016/S1364-8152(01)00067-6.

Gower, S. T., J. G. Vogel, J. M. Norman, C. J. Kucharik, S. J. Steele, and T. K. Stow (1997), Carbon distribution and aboveground net primary production in aspen, jack pine, and black spruce stands in Saskatchewan and Manitoba, Canada, *J. Geophys. Res.*, *102*(D24), 29,029–29,041, doi:10.1029/97JD02317.

Grunwald, S., J. A. Thompson, and J. L. Boettinger (2011), Digital soil mapping and modeling at continental scales: Finding solutions for global issues, *Soil Sci. Soc. Am. J.*, *75*, 1201–1213, doi:10.2136/sssaj2011.0025.

Guo, L. B., and R. M. Gifford (2002), Soil carbon stocks and land use change: A meta analysis, *Global Change Biol.*, *8*, 345–360, doi:10.1046/j.1354-1013.2002.00486.x.

Henderson, B. L., E. N. Bui, C. J. Moran, and D. A. P. Simon (2005), Australia-wide predictions of soil properties using decision trees, *Geoderma*, *124*, 383–398, doi:10.1016/j.geoderma.2004.06.007.

Houghton, R. A. (2000), Interannual variability in the global carbon cycle, *J. Geophys. Res.*, *105*(D15), 20,121–20,130, doi:10.1029/2000JD900041.

Jenny, H. (1941), *Factors of Soil Formation*, McGraw-Hill, New York.

Jobbágy, E. G., and R. B. Jackson (2000), The vertical distribution of soil organic carbon and its relation to climate and vegetation, *Ecol. Appl.*, *10*, 423–436, doi:10.1890/1051-0761(2000)010[0423:TVDOSO]2.0.CO;2.

Kunkel, M. L., A. N. Flores, T. J. Smith, J. P. McNamara, and S. G. Benner (2011), A simplified approach for estimating soil carbon and nitrogen stocks in semi-arid complex terrain, *Geoderma*, *165*, 1–11, doi:10.1016/j.geoderma.2011.06.011.

Liao, C., Y. Luo, C. Fang, J. Chen, and B. Li (2012), The effects of plantation practice on soil properties based on the comparison between natural and planted forests: A meta-analysis, *Global Ecol. Biogeogr.*, *21*, 318–327, doi:10.1111/j.1466-8238.2011.00690.x.

Mackay, D. S., B. E. Ewers, M. M. Loranty, and E. L. Kruger (2010), On the representativeness of plot size and location for scaling transpiration from trees to a stand, *J. Geophys. Res.*, *115*, G02016, doi:10.1029/2009JG001092.

Martin, J. G., and P. V. Bolstad (2009), Variation of soil respiration at three spatial scales: Components within measurements, intra-site variation and patterns on the landscape, *Soil Biol. Biochem.*, *41*, 530–543, doi:10.1016/j.soilbio.2008.12.012.

Martin, M. P., M. Wattenbach, P. Smith, J. Meersmans, C. Jolivet, L. Boulonne, and D. Arrouays (2011), Spatial distribution of soil organic carbon stocks in France, *Biogeosciences*, *8*, 1053–1065, doi:10.5194/bg-8-1053-2011.

Masek, J. G., and G. J. Collatz (2006), Estimating forest carbon fluxes in a disturbed southeastern landscape: Integration of remote sensing, forest inventory, and biogeochemical modeling, *J. Geophys. Res.*, *111*, G01006, doi:10.1029/2005JG000062.

McBratney, A. B. (1998), Some considerations on methods for spatially aggregating and disaggregating soil information, *Nutr. Cycl. Agroecosyst.*, *50*, 51–62, doi:10.1023/A:1009778500412.

McBratney, A. B., M. L. Mendonça Santos, and B. Minasny (2003), On digital soil mapping, *Geoderma*, *117*, 3–52, doi:10.1016/S0016-7061(03)00223-4.

Meentemeyer, V., and E. O. Box (1987), Scale effects in landscape studies, in *Landscape Heterogeneity and Disturbance*, edited by M. G. Turner et al., pp. 15–34, Springer, New York, doi:10.1007/978-1-4612-4742-5_2.

Meersmans, J., F. De Ridder, F. Canters, S. De Baets, and M. Van Molle (2008), A multiple regression approach to assess the spatial distribution of soil organic carbon (SOC) at the regional scale (Flanders, Belgium), *Geoderma*, *143*, 1–13, doi:10.1016/j.geoderma.2007.08.025.

Mykrä, H., J. Aroviita, J. Kotanen, H. Hämäläinen, and T. Muotka (2008), Predicting the stream macroinvertebrate fauna across regional scales: Influence of geographical extent on model performance, *J. N. Am. Benthol. Soc.*, *27*, 705–716, doi:10.1899/07-074.1.

Natural Resources Conservation Service (NRCS) (1996), *Soil Survey Laboratory Methods Manual*, U.S. Dep. of Agric., Washington, D. C.

Natural Resources Conservation Service (NRCS) (2006a), Soil Survey Geographic (SSURGO) database, scale 1:24,000, U.S. Dep. of Agric., Fort Worth, Tex.

Natural Resources Conservation Service (NRCS) (2006b), State Soil Geographic (STATSGO) database, scale 1:250,000, U.S. Dep. of Agric., Fort Worth, Tex.

Obeysekera, J., and K. Rutchey (1997), Selection of scale for Everglades landscape models, *Landscape Ecol.*, *12*, 7–18, doi:10.1007/BF02698203.

Pärtel, M., L. Laanisto, and S. D. Wilson (2008), Soil nitrogen and carbon heterogeneity in woodlands and grasslands: Contrasts between temperate and tropical regions, *Global Ecol. Biogeogr.*, *17*, 18–24.

Phachomphon, K., P. Dlamini, and V. Chaplot (2010), Estimating carbon stocks at a regional level using soil information and easily accessible auxiliary variables, *Geoderma*, *155*, 372–380, doi:10.1016/j.geoderma.2009.12.020.

Post, W. M., and K. C. Kwon (2000), Soil carbon sequestration and land-use change: Processes and potential, *Global Change Biol.*, *6*, 317–327, doi:10.1046/j.1365-2486.2000.00308.x.

Pringle, M. J., and R. M. Lark (2007), Scale- and location-dependent correlations of soil strength and the yield of wheat, *Soil Tillage Res.*, *95*, 47–60, doi:10.1016/j.still.2006.10.010.

Qi, Y., and J. Wu (1996), Effects of changing spatial resolution on the results of landscape pattern analysis using spatial autocorrelation indices, *Landscape Ecol.*, *11*, 39–49, doi:10.1007/BF02087112.

Ross, D. J., K. R. Tate, N. A. Scott, and C. W. Feltham (1999), Land-use change: Effects on soil carbon, nitrogen and phosphorus pools and fluxes in three adjacent ecosystems, *Soil Biol. Biochem.*, *31*, 803–813, doi:10.1016/S0038-0717(98)00180-1.

Schulp, C. J. E., and A. Veldkamp (2008), Long-term landscape–land use interactions as explaining factor for soil organic matter variability in Dutch agricultural landscapes, *Geoderma*, *146*, 457–465, doi:10.1016/j.geoderma.2008.06.016.

Stys, B., R. Kautz, D. Reed, M. Kertis, R. Kawula, C. Keller, and A. Davis (2003), Digital vegetation and land cover data set for Florida derived from 2003 Landsat ETM+ imagery, resolution 30 m, Fla. Fish and Wildlife Conserv. Comm., Tallahassee, Fla.

Tan, Z. X., R. Lal, N. E. Smeck, and F. G. Calhoun (2004), Relationships between surface soil organic carbon pool and site variables, *Geoderma*, *121*, 187–195, doi:10.1016/j.geoderma.2003.11.003.

Thompson, J. A., and R. K. Kolka (2005), Soil carbon storage estimation in a forested watershed using quantitative soil-landscape modeling, *Soil Sci. Soc. Am. J.*, *69*, 1086–1093, doi:10.2136/sssaj2004.0322.

Thompson, J. A., J. C. Bell, and C. A. Butler (2001), Digital elevation model resolution: Effects on terrain attribute calculation and quantitative soil-landscape modeling, *Geoderma*, *100*, 67–89, doi:10.1016/S0016-7061(00)00081-1.

Turner, M. G., R. V. O'Neill, R. H. Gardner, and B. T. Milne (1989), Effects of changing spatial scale on the analysis of landscape pattern, *Landscape Ecol.*, *3*, 153–162, doi:10.1007/BF00131534.

United States Geological Survey (USGS) (1999), National elevation dataset, scale 1:24,000, U.S. Geol. Surv., Sioux Falls, Iowa.

Vasques, G. M., S. Grunwald, N. B. Comerford, and J. O. Sickman (2010a), Regional modelling of soil carbon at multiple depths within a subtropical watershed, *Geoderma*, *156*, 326–336, doi:10.1016/j.geoderma.2010.03.002.

Vasques, G. M., S. Grunwald, and W. G. Harris (2010b), Spectroscopic models of soil organic carbon in Florida, USA, *J. Environ. Qual.*, *39*, 923–934, doi:10.2134/jeq2009.0314.

Vasques, G. M., S. Grunwald, and D. B. Myers (2012), Associations between soil carbon and ecological landscape variables at escalating spatial scales in Florida, USA, *Landscape Ecol.*, *27*, 355–367, doi:10.1007/s10980-011-9702-3.

Vaze, J., J. Teng, and G. Spencer (2010), Impact of DEM accuracy and resolution on topographic indices, *Environ. Model. Softw.*, *25*, 1086–1098, doi:10.1016/j.envsoft.2010.03.014.

Welch, B. L. (1947), The generalization of 'Student's' problem when several different population variances are involved, *Biometrika*, *34*, 28–35.

Woodcock, C. E., and A. H. Strahler (1987), The factor of scale in remote sensing, *Remote Sens. Environ.*, *21*, 311–332, doi:10.1016/0034-4257(87)90015-0.

Wu, J. (2004), Effects of changing scale on landscape pattern analysis: Scaling relations, *Landscape Ecol.*, *19*, 125–138, doi:10.1023/B:LAND.0000021711.40074.ae.

Wu, J., W. Shen, W. Sun, and P. T. Tueller (2002), Empirical patterns of the effects of changing scale on landscape metrics, *Landscape Ecol.*, *17*, 761–782, doi:10.1023/A:1022995922992.

Wu, S., J. Li, and G. H. Huang (2007), Modeling the effects of elevation data resolution on the performance of topography-based watershed runoff simulation, *Environ. Model. Softw.*, *22*, 1250–1260, doi:10.1016/j.envsoft.2006.08.001.

Wu, W., Y. Fan, Z. Wang, and H. Liu (2008), Assessing effects of digital elevation model resolutions on soil–landscape correlations in a hilly area, *Agric. Ecosyst. Environ.*, *126*, 209–216, doi:10.1016/j.agee.2008.01.026.