Digital Soil Mapping
Across Paradigms,
Scales and Boundaries
Chapter 13
Estimation of the Actual and Attainable Terrestrial Carbon Budget

P. Chaikaew, S. Grunwald and X. Xiong

Abstract Organic carbon is a key component of the terrestrial system that affects the physical, chemical, and biological processes. Changes in both the soil and the terrestrial carbon storage occur due to the interactions of natural ecological processes and anthropogenic activities. Research gaps to quantify soil and terrestrial carbon still exist. To discern between the actual and attainable carbon pools is critical to identify suitable adaptation and management alternatives to optimize natural carbon capital in the context of regional imposed changes, such as land use and climate change. Our objectives were to: (i) assess the spatially explicit relationships between observed soil organic carbon (SOC) and environmental factors and (ii) assess actual (TerrC\text{actual}) and attainable (TerrC\text{attain}) terrestrial carbon capital considering below-ground (soil) and above-ground (biomass) carbon. We collected 234 soil samples in the topsoil (0–20 cm) in 2008 and 2009 across the Suwannee River Basin in Florida, USA, based on a random design stratified by land cover/land use and soil suborders. For above-ground carbon assessment, we derived data from the LANDFIRE project which provided a high-resolution map of year-2000 baseline estimates of biomass carbon. A comprehensive set of 172 soil-environmental and human covariates was assembled from multiple data sources to predict and validate observed SOC stocks and TerrC\text{actual} using Random Forest (RF). The STEP-AWBH conceptual model (with S: Soil, T: Topography, E: Ecology, P: Parent material, A: Atmosphere, W: Water, B: Biota, and H: Human factors) provided the conceptual modeling framework to model TerrC\text{attain} that was implemented using RF and simulated annealing in combination. In the simulation, the STEP factors were kept constant, but the AWBH factors were varied by ±10, ±20, and ±30 %. The combined factors which amount to the highest modeled...
terrestrial carbon stocks were postulated to equal the attainable terrestrial carbon stocks. Results suggest that the TerrC_{attain} stocks showed slightly larger amounts than the TerrC_{actual} stocks across the basin. The TerrC_{actual} was 190 Tg C and the maximum TerrC_{attain} was 195 Tg C. Biotic, soil, parent material, topographic, and water-related factors played important roles in determining SOC storage, while human factors were only weak predictors. Although mean annual precipitation and monthly mean temperature in summer months were significant to explain both SOC and terrestrial carbon stocks, they showed moderate/minor influence on carbon storage. The land use/land cover variables were the strongest factors predicting soil and terrestrial carbon stocks. These findings suggest that land use adaptions have much potential to achieve TerrC_{attain}, specifically conversions from cropland to land use systems with larger net primary productivity. Bare soils, which represent marginal soils, also have potential to elevate carbon storage through management adaptions that do not compete with other land uses.

**Keywords** Actual carbon stocks · Attainable carbon stocks · Terrestrial carbon · Soil organic carbon · Random forest · Simulated annealing · STEP-AWBH

### 13.1 Introduction

Carbon sequestration has become an important policy option to mitigate the increasing atmospheric greenhouse gases (GHG) that pose threats to a warming global climate. The terrestrial biosphere can sequester significant amounts of anthropogenic carbon dioxide (CO₂) by the natural carbon uptake process through plant biomass and soils. However, how ecosystem factors and carbon dynamics in terrestrial systems interact with each other that determine critically important ecosystem services is not well understood yet.

Numerous digital soil mapping studies have modeled soil organic carbon (SOC) across large regions (Bélanger and Pinno 2008; Wang et al. 2011; Wu et al. 2009; Xiong et al. 2014), and terrestrial carbon has been assessed at global and continental scale by Lal (2008) and Dickson et al. (2014). Yet, these carbon assessments focus on actual conditions without providing clues of attainable or potential carbon that could be sequestered in a landscape. To identify those site-specific adaptations that demonstrate most promise to attain the largest amount of carbon storable in soil and biomass is of interest because they can guide land management, adaptation, mitigation, decision making, and policy implementations to achieve a more carbon neutral global state.

The STEP-AWBH model was developed for pixel-specific assessment of soil properties from a suite of soil-environmental factors (Grunwald et al. 2011). This conceptual model embraces soil (S), topography (T), ecology (E), parent material
(P), atmosphere (A), water (W), biota (B), and human (H) factors together to account for the effects on soil genesis. The STEP-AWBH model is flexible enough to be implemented with various geostatistical techniques, ensemble regression, and data mining methods to predict soil and terrestrial properties. For example, Xiong et al. (2014) applied the STEP-AWBH modeling concept to model SOC stocks in Florida using various data mining techniques, such as regression trees, bagged trees, random forest (RF), and support vector machines, from a large set of predictors (210 STEP-AWBH variables).

This study adopts the STEP-AWBH model and applies it to assess the attainable capacity of a terrestrial ecosystem to store carbon. Similar to the SOC sequestration, the attainable terrestrial carbon (TerrC\textsubscript{attain}) is set by factors that limit the inputs of carbon to the system (e.g., residue from vegetation, and fertilization that stimulates biomass production/net primary production), which can be modified by humans (e.g., implementation of best management practices, reduction of burning of fossil fuels for energy consumption, and wetland restoration) (Ingram and Fernandes 2001; Stockmann et al. 2013). The actual terrestrial carbon (TerrC\textsubscript{actual}) is controlled by factors that modulate carbon storage (e.g., drainage, tillage, land use management, soil respiration, or photosynthesis) and depends on a combination of environmental landscape factors, past and current anthropogenic forcings, and socioeconomic drivers. We postulate that human-induced management strategies combined with environmental fluctuations of climate, land use, and hydrology contribute to TerrC\textsubscript{attain} that is constraint by site-specific soil-landscape conditions. Importantly, which of these coupled human-environmental combinations achieve the maximum attainable carbon storage in terrestrial systems is usually not known.

The objectives of this study were to: (i) assess the spatially explicit relationships between measured SOC stocks and environmental factors and (ii) assess the environmental value of actual (TerrC\textsubscript{actual}) and attainable (TerrC\textsubscript{attain}) terrestrial carbon capital across the Florida portion of the Suwannee River Basin, (FL-SRB) consisting of below-ground and above-ground carbon.

### 13.2 Study Area

The FL-SRB is located in north-central Florida with an approximate area of 19,665 km\textsuperscript{2} (Fig. 13.1). Dominant soils are sand-rich which inherently does not promote SOC accretion. On the other hand, soils formed in aquic conditions are carbon-rich which are prominent in the FL-SRB. The soil temperature regimes are mixed with 86 % of the area’s soil classified as thermic and 14 % as hyperthermic (Natural Resources Conservation Service (NRCS), 2006). The ecological landscape conditions such as land use/land cover (LULC) and hydrology are complex.
13.3 Materials and Methods

13.3.1 Above-Ground and Below-Ground Carbon Data

A total of 234 soil samples in the topsoil (0–20 cm) were collected between 2008 and 2009 across the FL-SRB and its buffer area (20 km around the FL-SRB) based on the random design stratified by LULC and soil suborders. Each soil sampling location was georeferenced with a differential global positioning system. Total carbon (TC) was measured by a combustion gas analyzer (Shimadzu TOC-V/SSM-5000) at 900°C, while inorganic carbon (IC) was analyzed by treating soil samples with 42.5% phosphoric acid (H₃PO₄) and then combusting them at 200°C. The SOC concentration was calculated by subtracting the IC concentration from the obtained TC concentration (mg kg⁻¹). The SOC concentration for each site was converted to stock units (kg C m⁻²) using measured bulk density and soil depth. For above-ground carbon assessment, we derived data from the LANDFIRE project which provided a high-resolution map of year-2000 baseline estimates of above-ground biomass carbon (National Biomass Carbon Data, NBCD 2000 Version 2) (Kellnorfer et al. 2013). The above-ground live and dry biomass in kg C m⁻² was extracted at each of the soil sampling locations. The site-specific TerrC_actual stocks were derived by the summation of the measured SOC stocks in
the top 20 cm of the soil and above-ground biomass from the NBCD database (kg C m\(^{-2}\)) based on 234 soil sampling locations.

### 13.3.2 Environmental and Anthropogenic Covariates

A comprehensive set of 172 environmental and human covariates representing the STEP-AWBH factors was compiled from multiple data sources using ArcGIS software. Predictors included 31 categorical and 141 continuous data types and are described in detail in Chaikaew (2014).

The S factor was described by 15 soil properties (e.g., soil taxonomic order and soil texture) and ten soil–water variables (e.g., surface soil moisture and drainage class); the T factor was represented by 9 topographic attributes (e.g., slope and compound topographic index); the E factor was described by 2 ecological variables (e.g., ecoregion and physiographic province); the P factor was represented by 2 parent material variables (e.g., environmental geology and surficial geology); the A factor was described by 3 atmospheric factors (e.g., precipitation, temperature, and solar radiation); the B factor was represented by 16 vegetation factors (e.g., LULC, canopy coverage, and cropland); and the H factor was described by 5 human covariates (e.g., population growth and household income).

### 13.3.3 Modeling the Relationships Between SOC and STEP-AWBH Factors

This study is embedded in the STEP-AWBH modeling concept which explicitly combines spatially and temporally explicit environmental and human variables that model the evolution of the soil ecosystem (Grunwald et al. 2011) (Eq. 13.1).

\[
SA(z, p_x, t_c) = f\left\{ \sum_{j} S_j(z, p_x, t_c), T_j(p_x, t_c), E_j(p_x, t_c), P_j(p_x, t_c) \right\} ;
\]

\[
\int_{i=0}^{m} \left\{ \sum_{j} A_j(p_x, t_i), W_j(p_x, t_i), B_j(p_x, t_i), H_j(p_x, t_i) \right\} (13.1)
\]

where \(SA\) is the target soil (or terrestrial) realization, \(S\) represents the ancillary soil properties, \(T\) represents the topographic properties, \(E\) represents the ecological properties, \(P\) represents the parent material and geologic properties, \(A\) represents the atmospheric properties, \(W\) represents the water properties, \(B\) represents the biotic
properties, $H$ represents the human-induced forcings, $j$ is the number of predictors, $j = 1, 2, \ldots, n$, $p_x$ is a pixel with size $x$ (width = length = $x$) at a site specific on land, $t_c$ is the current time, $t_i$ is the time to $t_c$ with time steps $i = 0, 1, 2, \ldots, m$, and $z$ is the soil depth. The spatially explicit STEP factors capture the relative stable soil-forming factors within a human time frame, while the AWBH factors account for time-dependent variation (Thompson et al. 2012).

The RF regression method was used to identify the most powerful environmental predictive factors to model SOC. The model was randomly split into a calibration set (70 %, $n = 164$) and a validation set (30 %, $n = 70$). Model performance was assessed using the coefficient of determination ($R^2$), root-mean-square error (RMSE), and residual prediction deviation (RPD), and ratio of prediction error to interquartile range (RPIQ) was reported for error assessment of the RF model.

### 13.3.4 Assessing Terrestrial Carbon Stocks

The most powerful predictors ($n = 43$) in the first quantile of all STEP-AWBH variables of the SOC model were selected and used to predict the observed $\text{TerrC}_{\text{actual}}$ stocks using the RF model. To assess the $\text{TerrC}_{\text{attain}}$ stocks, we posit that the STEP factors are not expected to substantially change within a human time frame (e.g., past decades), whereas AWBH factors are likely to be variable in time and may increase or decrease. Thus, the latter were used to vary within a range of upper and lower bounds to simulate $\text{TerrC}_{\text{attain}}$.

The same 43 STEP-AWBH predictors were used to predict $\text{TerrC}_{\text{attain}}$ stocks at the 234 sites using a simulated annealing (SA) approach (Kirkpatrick et al. 1983). In the $\text{TerrC}_{\text{attain}}$ model, the STEP factors were kept constant. The AWBH factors were varied by ±10, ±20, and ±30 %, respectively, by keeping AWBH factors constant, except for one of them that was increased/decreased one-by-one with the respective percentage value, until all factor combinations were assessed within reasonable upper and lower bounds. The factor combination which amounted to the highest terrestrial carbon stocks simulated at the 234 sites for the FL-SRB was postulated to equal the attainable terrestrial carbon stocks that could be obtained based on dynamic AWBH variables and relatively stable STEP conditions.

To characterize the spatial distribution of actual terrestrial carbon stocks across the basin, the regression kriging (RK) technique (Odeh et al. 1995) was used. First, the residuals of $\text{TerrC}_{\text{actual}}$ at the 234 sites were kriged and then added back to the estimates of $\text{TerrC}_{\text{actual}}$ to create interpolated surfaces showing terrestrial carbon stocks.
13.4 Results

13.4.1 Variable Importance and Spatial Variation in Soil Organic Carbon

The variables that emerged in the first quantile of the RF-SOC model showing predictive power were as follow: biota > soil > topography. Minor, yet significant, predicting variables represented water, atmospheric properties, and parent material. Soil taxonomic variables, such as soil great group, suborder, and subgroup, were highly interrelated with SOC stocks as shown in the top ten explanatory variables. Areas with poorly to very poorly drained soils tended to accumulate SOC content, whereas areas with well-drained to excessively drained soils had a tendency to have net losses of SOC. These results suggest that topographic and soil/water-related variables play a dominant role to infer on SOC storage in the basin.

Slope was negatively correlated with SOC stocks. This indicates that the relatively flat downslope positions were correlated with high SOC, and vice versa, upslope positions showed the opposite behavior. Even though topographic terrain in Florida is relatively flat (0–5 % slopes), the steeper slopes are found in the northern region of Florida where the FL-SRB is located. Distance from streams or open water, available water capacity (0–25 cm), hydrologic group, ponding frequency class, and soil runoff potential were water variables (W) in the model that demonstrated strong connectivity with SOC stocks. The effect of wetness in soil is considered to be a major factor that controls vegetation growth and the decomposition process that are closely interlinked with SOC gain (Vasques et al. 2012).

The parent material was found to have high influence on SOC, while the ecology variables (E) were not strongly associated with SOC. The influence of parent material on SOC stocks occurs through different sources, such as soil weathering, mineralogy, water permeability, nutrient supply, mineral complexation, structure, that control the pH, and microbe’s habitat affecting plant production and decomposition (Post et al. 2004).

Three climatic variables of long-term (2000–2008) monthly maximum temperature in summer (July, August, and September) were negatively correlated with SOC stocks. The opposite was found for annual average precipitation that was positively correlated with SOC stocks. Ample research has been conducted to study the interactions between climate and SOC which are still debated fiercely because interactions vary geographically around the globe (Bardgett 2011; Ontl and Schulte 2012; Poeplau et al. 2011; Xiong et al. 2014).

Human variables were not considered powerful predictors in the model as they ranked in the middle and bottom of all predictor variables. Fertilizer consumption (74th) had the most influence among the H factor, followed by best management practice (BMP) implementation (100th) of all variables, while population growth ranked near the bottom of all predictors. This indicates that human factors may have an indirect effect (e.g., through land use management and tillage operations) or little impact on SOC.
13.4.2 Model Performance

The prediction model of SOC stocks using 172 STEP-AWBH variables was able to account for 93% of the variation in calibration mode and 44% of the variation in validation mode across the basin. The RMSE value of 2.65 kg C m\(^{-2}\) in the validation set was higher than 1.33 kg C m\(^{-2}\) in the calibration set, and RPD values of 1.28 kg C m\(^{-2}\) in the validation set were lower than 2.58 kg C m\(^{-2}\) in the calibration set. Considering that the minimum of SOC stock was 1.1 kg C m\(^{-2}\) and the maximum of SOC stock in the basin was 28.8 kg C m\(^{-2}\), the errors suggest that the model performed moderately well.

13.4.3 Estimates of the Terrestrial Carbon Stocks

The amount of carbon was almost twofold in the terrestrial ecosystem (mean of 9.2 kg C m\(^{-2}\)) in comparison with topsoil (mean of 5.2 kg C m\(^{-2}\)) at a site-specific basis. The results imply that about 44% of carbon storage is found above ground and 56% of carbon is stored in topsoils. This is a conservative estimate because additional carbon is present in subsoils. The relatively high terrestrial carbon stocks were observed in swamps and forests. The highest mean TerrC\(_{\text{actual}}\) stock values were found in mixed wetland forests 16.6 kg C m\(^{-2}\), followed by swamps (14.6 kg C m\(^{-2}\)) and pineland (11.0 kg C m\(^{-2}\)). The lowest TerrC\(_{\text{actual}}\) stocks were present in row/field crops with a mean value of 1.2 Tg C (2.7 kg C m\(^{-2}\)).

The modeled TerrC\(_{\text{attain}}\) estimates under different environmental forcings (i.e., A WBH forcings) clearly showed that factor combinations (i.e., climatic properties, water, and biota) concomitantly had strong effects on carbon storage. Assuming that predictors are changed by \(\pm 10\%\), our model simulated a minimal increase in terrestrial carbon stocks with mean values of 9.3 kg C m\(^{-2}\) and median values of 8.8 kg C m\(^{-2}\). The mean TerrC\(_{\text{attain}}\) was 9.4 kg C m\(^{-2}\) and the median was 8.8 kg C m\(^{-2}\) when the environmental factors fluctuated by \(\pm 20\). And the mean TerrC\(_{\text{attain}}\) was 9.4 kg C m\(^{-2}\) and the median was 9.0 kg C m\(^{-2}\) when the environmental factors fluctuated by \(\pm 30\).

The TerrC\(_{\text{actual}}\) stocks in the model indicated that this terrestrial system was close to the saturation condition. The absolute increase from predicted TerrC\(_{\text{actual}}\) to TerrC\(_{\text{attain}}\) (\(\pm 30\%\)) in terrestrial carbon stocks amounted to 4.3 Tg (mean) and 13.2 Tg (median) which are substantial amounts. The predicted TerrC\(_{\text{attain}}\) (\(\pm 30\%\)) for wetlands, pinelands, and hardwood forests were 13.7, 9.4, and 10.5 kg C m\(^{-2}\), respectively, while the predicted TerrC\(_{\text{actual}}\) for wetlands, pinelands, and hardwood forests were 11.9, 9.1, and 10.3 kg C m\(^{-2}\), respectively. The TerrC\(_{\text{actual}}\) and TerrC\(_{\text{attain}}\) stocks in pineland and forests were quite similar in values with differences in TerrC\(_{\text{actual}}\) and TerrC\(_{\text{attain}}\) of 0.3 kg C m\(^{-2}\) (pineland) and 0.3 kg C m\(^{-2}\) (hardwood forests). The actual storage of carbon in the terrestrial ecosystem
amounted to 190 Tg C and the maximum storable attainable carbon was 195 Tg C. The spatial distribution of terrestrial carbon stocks is shown in Fig. 13.2.

13.5 Conclusions

We assessed the relationships between SOC stocks and STEP-AWBH factors in the FL-SRB and found that biotic, soil/water, and topographic factors played crucial roles in determining SOC storage, whereas human factors seemed to be fading from being strong predictors. Among the predictors, maximum temperature in summer time and mean annual precipitation also stood out as controlling factors for SOC storage. The whole basin stores tremendous amounts of carbon, multiple times larger than atmospheric carbon, with about 190 Tg of \( \text{TerrC}_{\text{actual}} \) carbon stocks based on the RK model. There was no single environmental factor that imparted most control on SOC storage, but instead intricate combinations of STP-AWB variables.
Importantly, biotic factors played a larger role to associate with SOC stocks and terrestrial carbon compared to climatic factors. This inherently implies that global climate change will have less of an impact compared to land cover and land use change to achieve the attainable carbon level in this soil landscape. The model predicted that the $\text{TerrC}_{\text{actual}}$ stock values were close to the $\text{TerrC}_{\text{attain}}$ stock values in some instances (e.g., under wetlands) suggesting the presence of saturation in this area, while in other regions (e.g., under row/field crops) ample opportunities exist to enhance carbon sequestration.

The simulated $\text{TerrC}_{\text{attain}}$ values were “point specific (in kg C m$^{-2}$)” assuming no fixed depth (i.e., soil depth/volume or vegetation height/volume) to store carbon. This vantage point liberates us from the constraint to consider a fixed soil volume (kg C m$^{-2}$ and 0–20 cm depth) to accrete carbon up to a saturation limit. As is well known, hydric soils tend to accrete carbon through an increase in soil depth rather than enhancement of carbon density within existing peat layers. Similarly, attainable biomass carbon is not necessarily linearly linked to a specific height of the vegetation and may increase carbon not only through growth but also through changes in the vegetation density.

The major goal was to preserve the carbon stored in the terrestrial system of the FL-SRB and enhance them through optimized carbon management. Land use adaptations have much potential to reach $\text{TerrC}_{\text{attain}}$, specifically land use conversions from cropland to systems with larger net primary productivity (NPP). Bare soils, which represent marginal soils, also bear potential to elevate carbon storage if improved through management. Our study demonstrated a novel approach to assess terrestrial carbon through management, adaptation, and mitigation that is attainable in a subtropical basin consisting of a mosaic of different land uses embedded in a complex soil landscape. This approach is generalizable and transferable to any other landscape setting. Rather than offsetting CO$_2$ emissions, there are other co-benefits of increased levels of carbon sequestration to the ecosystem functions (e.g., improvements in crop productivity, soil security, food security, soil aggregation that enhances nutrient storage, and water holding capacity). The spatially explicit modeling of actual and attainable terrestrial carbon stocks allows identifying and targeting carbon poor areas to implement conservation and carbon management strategies. Hence, our approach has much value to outline pathways into a carbon-rich future.

**Acknowledgements** Soil data collection is supported by the “Rapid Assessment and Trajectory Modeling of Changes in Soil Carbon across a Southeastern Landscape” funded by USDA-NIFA, a core project of the North American Carbon Program. This research was conducted through a PhD fellowship funded by the Royal Thai Government.
References
