Science of the Total Environment xxx (xxxx) xxx

Contents lists available at ScienceDirect

Science of the Total Environment

journal homepage: www.elsevier.com/locate/scitotenv

Accounting for two-billion tons of stabilized soil carbon

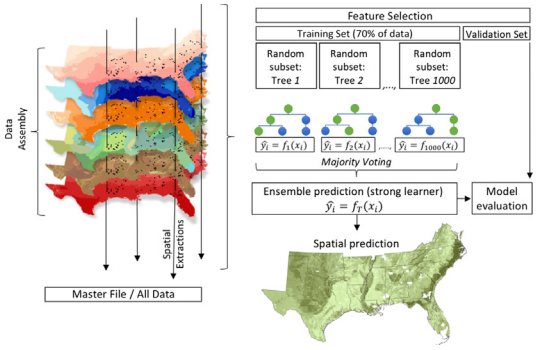
C. Wade Ross [a,b,*], Sabine Grunwald [a], Jason G. Vogel [c], Daniel Markewitz [d], Eric J. Jokela [c], Timothy A. Martin [c], Rosvel Bracho [c], Allan R. Bacon [a], Colby W. Brungard [b], Xiong Xiong [a]

[a] University of Florida, Soil and Water Sciences Department, 2181 McCarty Hall A, PO Box 110290, Gainesville, FL 32611, USA
[b] New Mexico State University, Department of Plant and Environmental Sciences, MSC 3Q, PO Box 30003, Las Cruces, NM 88003, USA
[c] University of Florida, School of Forest Resources and Conservation, 136 Newins-Ziegler Hall, Gainesville, FL 32611, USA
[d] Warnell School of Forestry and Natural Resources, University of Georgia, Athens, GA 30602, USA

HIGHLIGHTS

- Environmental heterogeneity is a large source of soil carbon variability.
- Strategic feature selection is used to identify critical, regional-scale relationships.
- Precipitation, nitrogen, and soil moisture have the strongest association with topsoil C.
- Parent material and elevation have the strongest association with subsoil C.
- Approximately 2.6 Pg C are stored in the upper 1 m of production forestland soil.

GRAPHICAL ABSTRACT



ARTICLE INFO

Article history:
Received 17 July 2019
Received in revised form 17 September 2019
Accepted 21 September 2019
Available online xxxx

Editor: Jay Gan

Keywords:
Soil carbon
Forest soils
Carbon cycle
Machine learning
Feature selection
Data mining

ABSTRACT

The pedosphere is the largest terrestrial reservoir of organic carbon, yet soil-carbon variability and its representation in Earth system models is a large source of uncertainty for carbon-cycle science and climate projections. Much of this uncertainty is attributed to local and regional-scale variability, and predicting this variation can be challenging if variable selection is based solely on *a priori* assumptions due to the scale-dependent nature of environmental determinants. Data mining can optimize predictive modeling by allowing machine-learning algorithms to learn from and discover complex patterns in large datasets that may have otherwise gone unnoticed, thus increasing the potential for knowledge discovery. In this analysis, we identify important, regional-scale determinants for top- and subsoil-carbon stabilization in production forestland across the southeastern US. Specifically, we apply recursive feature elimination to a large suite of socio-environmental data to strategically select a parsimonious, yet highly predictive covariate set. This is achieved by recursively considering smaller and smaller covariate sets—or features—by first training the estimator on the full set to obtain feature importance. The least important features are pruned, and the procedure is recursively repeated until a desired number of covariates is identified. We show that although carbon ranges from 0.3 to 8.2 kg m^{-2} in the topsoil (0 to 20 cm), and from 0.4 to 17.6 kg m^{-2} in the subsoil (20 to 100 cm), this variability is predictably distributed with precipitation, soil moisture, nitrogen and sand content, gamma ray emissions, mean annual minimum temperature, and elevation. From our spatial predictions, we estimate that 2.6 Pg of soil carbon is currently stabilized in the upper 100 cm of production forestland, which covers 34.7 million ha in the southeastern US.

* Corresponding author at: Department of Plant and Environmental Sciences, New Mexico State University, P.O. Box 30003, Las Cruces, NM 88003-8003, USA.
E-mail address: cwross@nmsu.edu (C.W. Ross).

https://doi.org/10.1016/j.scitotenv.2019.134615
0048-9697/

1. Introduction

Soil is a critical component of the global climate system, serving as both a sink and a source of CO_2 by actively exchanging carbon with the atmosphere (Davidson, 2016; Luo et al., 2015). Resource management, carbon cycle projections, and policy all rely on accurate representations of soil carbon, yet global estimates remain highly uncertain despite comprehensive efforts to quantify this large reservoir (Gianelle et al., 2010). Depending on the depth modeled, published estimates range from 863 to over 3800 Pg C (Batjes, 2014; Eglin et al., 2010; Sanderman et al., 2017; Watson et al., 2000). Much of the uncertainty is attributed to the vertical and horizontal variability of soil, the scale of analysis, and the subsequent loss of information that occurs during spatial aggregation (Jobbágy and Jackson, 2000; Ross et al., 2013; Xiong et al., 2015), which is particularly problematic when aggregating non-linear data over large areas (Easterling, 1997).

Because soil carbon stabilization is governed by a multitude of non-linear relationships, the strength of relationships can vary with the scale of analysis (Miller et al., 2015; Xiong et al., 2016). Across larger spatial scales, for example, carbon sequestration varies with plant productivity, which in turn is affected by atmospheric CO_2 (Roy et al., 2016), growing season length (Hilton et al., 2017), and resource availability (Eskelinen and Harrison, 2015). However, soil carbon responds to socio-environmental conditions that can vary dramatically at different temporal scales and across regional and sub-regional scales. Factors affecting soil carbon persistence include temperature, precipitation, and acidity (Chen et al., 2018; Schmidt et al., 2011), as well as management (Noormets et al., 2015), and disturbance from land use change (Ross et al., 2016; Xiong et al., 2014b), fire (Godwin et al., 2017), and erosion (Pimentel, 2006). Characterizing these factors at regional scales may be required to upscale soil carbon to global estimates and to refine our understanding of soil carbon stabilization (Mulder et al., 2016).

A recent US Department of Agriculture funded Coordinated Agricultural Project referred to as PINEMAP (Pine Integrated Network: Education, Mitigation, and Adaptation Project) addressed this issue by establishing a monitoring network across the southeastern US to refine our understanding of carbon storage and dynamics in managed forests at the regional-scale (Will et al., 2015). Forests cover 99 million hectares of land in the southeast and account for almost one third of all forested lands in the conterminous US (Oswalt et al., 2014). Not only are these forests an economically-important resource—providing approximately 60% and 16% of the US and global industrial wood supply by volume (Oswalt et al., 2014)—but are ecologically important as well, and sequestered enough aboveground carbon (176 Tg C yr^{-1}) to mitigate 42% of the regions CO_2 emissions between 2000 and 2005 (Lu et al., 2015). About one third of the regions forests are pinelands, of which 19% are comprised of managed pine plantations (Wear and Greis, 2013). The most dominant species—loblolly pine (*Pinus taeda* L.)—accounts for more than two thirds of all planted tree species in the region (Wear and Greis, 2013).

Intense silvicultural production cycles in this region are a large source of land-cover change, which subsequently affects the region's carbon cycle. An accurate estimate of soil-carbon distribution in southeastern production forestland is therefore a critical step towards further resolving carbon-cycle science in this region, and to identify factors potentially affecting soil carbon at the global scale. By identifying important regional-scale associations, we hypothesize that our models will provide improved estimates of soil carbon stock when compared with those derived from global models. In this analysis, we develop a data-driven approach to model topsoil (0 to 20 cm) and subsoil (20 to 100 cm) carbon, which is based on a regional compilation (N = 2,564) of soil samples collected from PINEMAP research sites. Variable selection is performed by applying recursive feature elimination to a comprehensive set (N = 73) of environmental predictors to identify parsimonious covariate sets (N = 5) for each depth interval, which are used with the random forest (RF) algorithm to produce soil carbon prediction maps for top- and subsoil depth intervals.

2. Material and methods

2.1. Study area

Our random forest models were trained on data collected from the PINEMAP Tier 2 network, which consisted of 106 research sites with 2 to 3 replicates (on average) at each site, for a total of 322 plots (Fig. 1). Tier 2 research sites were chosen to capture the region-wide variation in soil, landscape positions, and climate that characterize the native geographic range of loblolly pine.

Climate of the study area is classified as a warm and humid temperate region with hot summers (Kottek et al., 2006); however, temperature, humidity, and precipitation vary considerably across this region (Fig. 2). Mean annual precipitation ranges from 1100 mm yr^{-1} in the Northern Inner Piedmont of Virginia to 1590 mm yr^{-1} in the Southern Pine Plains and Hills of Louisiana and mean annual temperature ranges from 13.5 °C in the Northern Inner Piedmont of Virginia to 20.3 °C in the Eastern Florida Flatwoods (Abatzoglou, 2013). While the soils of this region are very diverse, research sites are primarily positioned on Ultisols (61%), Alfisols (23%), Spodosols (12%), Entisols (2%), and Inceptisols (2%) (Soil Survey Staff, 2013).

2.2. Soil sampling

Soil samples were collected between 2012 and 2015 at eight random locations with four locations randomly assigned to one of two composite samples for analysis. Four depth intervals were sampled for each location from most of the 322 plots, resulting in 2,564 soil samples. Sampled depth intervals at each plot were 0 to 10, 10 to 20, 20 to 50, and 50 to 100 cm and plot size averaged 0.2 ± 0.14 ha (mean ± one standard deviation). Each composite sample was sieved (2 mm) to remove coarse fractions (roots and stones); stone mass and total mass of the air-dried samples were recorded. A subsample (approximately 20% of air-dried mass) was weighed and then dried for 48 h at 65 °C before recording the oven-dried weight and analyzing soil carbon concentration via dry combustion. We assume that the reported soil carbon concentration is predominantly organic in nature, as pH in surface soils averaged 4.8, exceeding 5.5 only 12 times. Soil bulk density (g cm^{-3}) was measured via Eq. (1), and soil carbon content (g cm^{-2}) was derived according to Eq. (2).

$$D_b = \frac{M_s}{V_t} \quad (1)$$

where D_b is the bulk density (g cm^{-3}), M_s = dry soil weight (g), V_t is the soil core volume (cm-3).

$$SC = c * D_b * D_Z \quad (2)$$

where SC is soil carbon (kg m-2), c is carbon concentration (%), Db is bulk density (g cm^{-3}), and D_Z is the depth interval (cm).

2.3. Environmental covariates

Environmental covariates were assimilated from 16 datasets acquired from national and international agencies (Table 1). These data were chosen to represent the environmental factors from the

Please cite this article as: C. W. Ross, S. Grunwald, J. G. Vogel et al., Accounting for two-billion tons of stabilized soil carbon, Science of the Total Environment, https://doi.org/10.1016/j.scitotenv.2019.134615

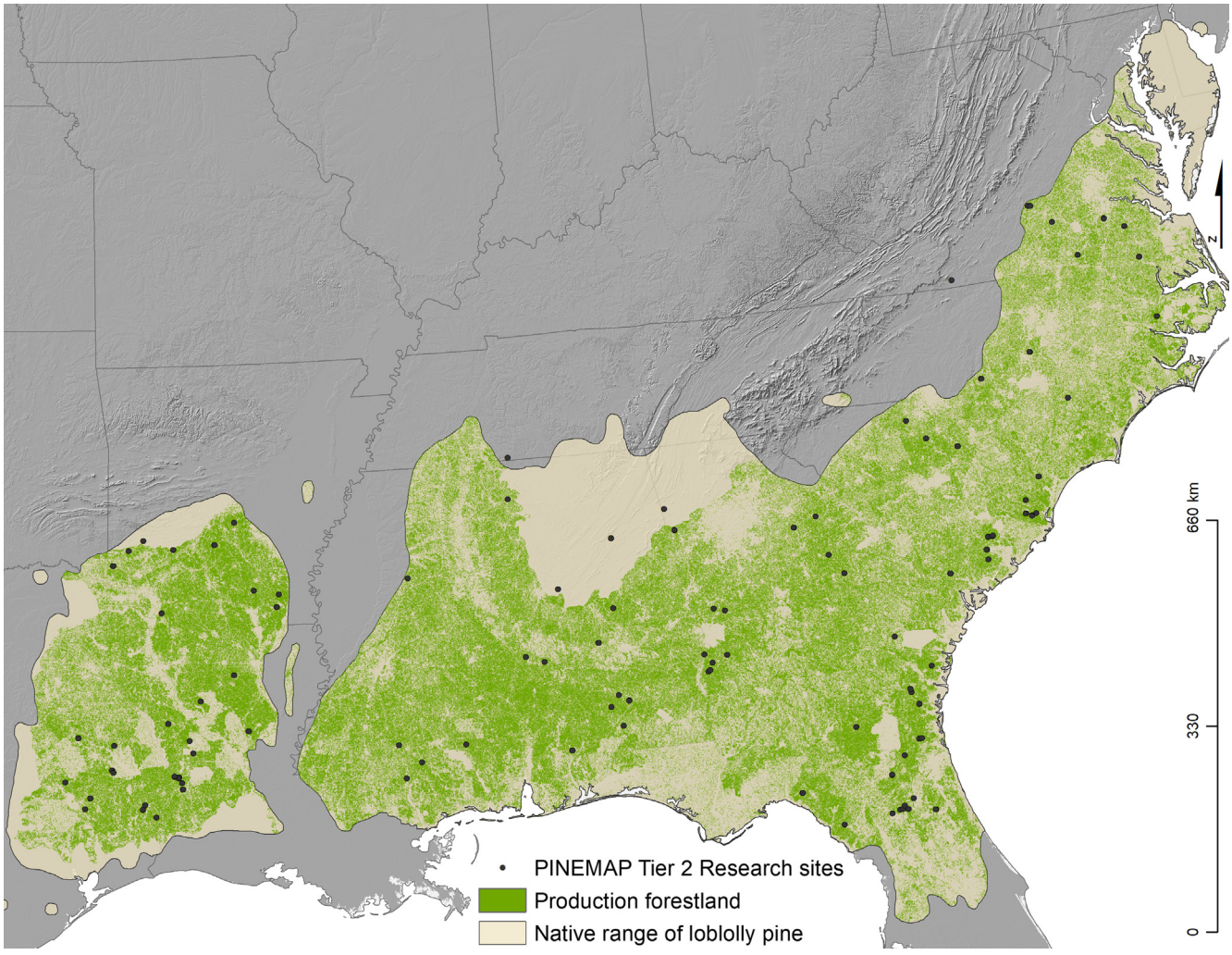


Fig. 1. PINEMAP (Pine Integrated Network: Education, Mitigation, and Adaption Project) Tier 2 research sites and areas identified as production forestland by Marsik et al. (2018) within the native range of loblolly pine (*Pinus taeda* L.).

STEP-AWBH modeling framework (Grunwald et al., 2011; Xiong et al., 2014a), and contained information relating to the lithosphere, atmosphere, anthroposphere, hydrosphere, and biosphere. All data were projected to Albers Equal Area Conic projection at 250 m^2 resolution. Bilinear interpolation was used to resample continuous data, and nearest neighbor was used to resample categorical data. The grid cell attribute information from each covariate was extracted at the point location for each research site and merged into a single data frame that consisted of 106 study sites, corresponding to 322 observations and 73 covariates.

2.4. Variable selection and predictive modeling

A fundamental goal of predictive modeling is to generalize beyond the relationships learned during model training; however, generalizing correctly becomes increasingly difficult as the number of covariates increase—a phenomenon referred to as the "curse of dimensionality" (Bellman, 1961; Domingos, 2012). Variable selection was therefore performed using recursive feature elimination (Kuhn, 2008) to identify a parsimonious, yet highly predictive covariate set for top- and subsoil models. This was achieved by first identifying pairwise-correlations and removing the variable with the largest mean absolute correlation. A random forest regressor was then fit to the remaining covariates to obtain feature importance for each covariate, and the least important feature was pruned from the modeling matrix. This procedure was repeated recursively until the most important covariates (N = 5) remained.

Predictive relationships were modeled with the *randomForest* package (Breiman, 2001; Liaw and Wiener, 2002) available in R 3.4.4 (R Core Team, 2019). Top- and subsoil models were trained on their respective parsimonious covariate sets. A full random forest (RF) model with 1,000 individual regression trees and one covariate (mtry = 1) was used at each split as determined by the tuneRF algorithm included in the *randomForest* package. Model performance was evaluated with an independent validation set by randomly selecting 30% of the observations, leaving 70% of the observations for model training. The two sample Kolmogorov-Smirnov test was conducted to verify that the distribution of soil carbon was similar between training and validation sets (Massey, 1951). Associations between explanatory and response variables was evaluated with accumulated local effects (ALE) plots from the *iml* package (Molnar et al., 2018).

2.5. Uncertainty assessment

Model predictions were evaluated with the coefficient of determination (R^2, Eq. (3)), root mean squared error (RMSE, Eq. (4)), and the ratio of performance to interquartile distance (RPIQ, Eq. (5)).

Please cite this article as: C. W. Ross, S. Grunwald, J. G. Vogel et al., Accounting for two-billion tons of stabilized soil carbon, Science of the Total Environment, https://doi.org/10.1016/j.scitotenv.2019.134615

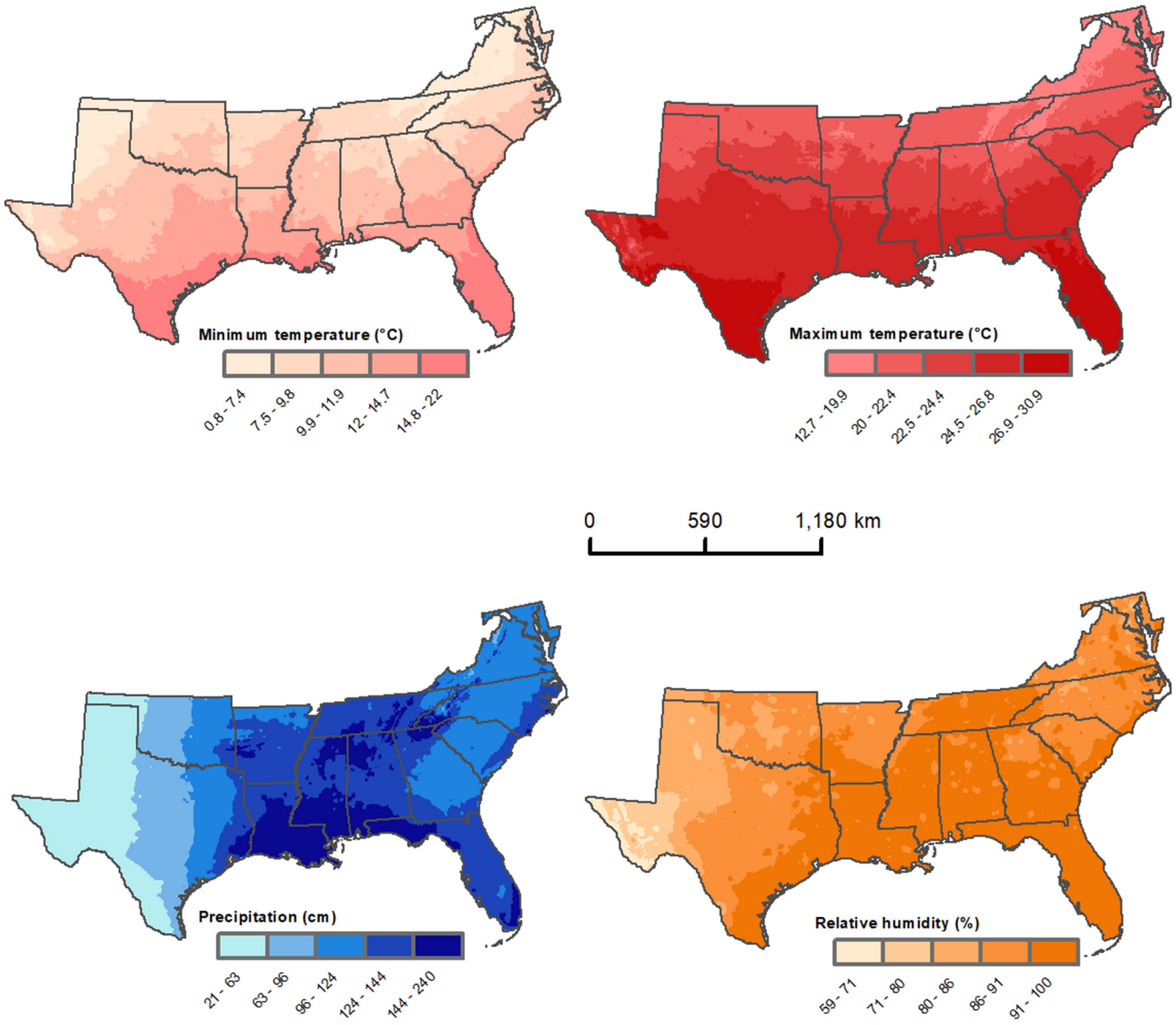


Fig. 2. Long-term mean annual climate (1979 – 2011). Adapted from Abatzoglou (2013).

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \quad (3)$$

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i)^2} \quad (4)$$

$$RPIQ = \frac{IQ}{RMSE} \quad (5)$$

where, $\hat{y}_i$ are the model predicted values, y_i are the observed values, $\bar{y}$ is the mean of observed values, n is the number of predicted or observed values in the held-out dataset (testing) with $i = 1, 2, \ldots, n$, SD is the standard deviation of the testing set, $RMSE$ is the root mean square error, and IQ is the interquartile range. The coefficient of determination (R^2) is a unitless index that measures the size of the model residuals to the mean of the observed values. As such, R^2 indicates how close model predictions are to observations. RMSE is defined as the standard deviation of the model residuals, which indicates the spread of residuals around the line of best fit. RPIQ considers both prediction error and the variation of observed values, and is defined as the range of observed values divided by RMSE. Higher R^2 and RPIQ values translate to better model performance, while lower RMSE indicates better model performance.

3. Results

The concentration of soil carbon (%) generally declines with depth, and soil bulk density increases with depth (Fig. 3). However, the vertical and horizontal distribution of measured soil carbon is highly variable, both within and between PINEMAP research sites. A considerable amount of the observed variation is attributed to extreme, but infrequent values (Table 2).

Carbon contents across USDA soil taxonomy at the suborder level also exhibit a considerable amount of variability, with median carbon stock (kg m^{-2}) greatest in the topsoil of Humods, Udepts, and Aquults, while subsoil carbon is greatest in Orthods, Aquods, and Aquults (Fig. 4).

Precautionary steps taken to ensure that the distribution of carbon was similar between the RF model training and validation sets

Please cite this article as: C. W. Ross, S. Grunwald, J. G. Vogel et al., Accounting for two-billion tons of stabilized soil carbon, Science of the Total Environment, https://doi.org/10.1016/j.scitotenv.2019.134615

Table 1
Covariates used for variable selection with recursive feature elimination.

Variable[a]	Abbreviation	N[a]	Source/Dataset[a]	Scale/Resolution (m)	Date
Gamma ray emissions	gammaAbsDose	1	USGS Gamma-ray	2000	1975 – 1980
Thorium-232	gammaThorium	1	USGS Gamma-ray	2000	1975 – 1980
Potassium-40	gammaPotassium	1	USGS Gamma-ray	2000	1975 – 1980
Uranium-238	gammaThorium	1	USGS Gamma-ray	2000	1975–1980
Parent material age	UNIT_AGE	1	USGS NGM	1:100000	1998
Parent material	ROCKTYPE1	1	USGS NGM	1:100001	1998
Parent material	ROCKTYPE2	1	USGS NGM	1:100001	1998
Elevation	srtmElevation	1	CGIAR SRTM	90	1999
Slope percent	srtmSlopePct	1	CGIAR SRTM	90	1999
Aspect	srtmAspect	1	CGIAR SRTM	90	1999
Curvature	srtmCurv	1	CGIAR SRTM	90	1999
Flow direction	srtmFlowDir	1	CGIAR SRTM	90	1999
Flow accumulation	srtmFlowAcc	1	CGIAR SRTM	90	1999
TWI	srtmTWI	1	CGIAR SRTM	90	1999
Clay content	clay_0to100	2		100	2018
Potassium	k_0to100	2		100	2018
Magnesium	mg_0to100	2		100	2018
pH	ph_0to100	2		100	2018
Sand content	sand_0to100	2		100	2018
Precipitation	Idaho_pr	2	METDATA	4000	1979 – 2010
Maximum RH	Idaho_rmax	1	METDATA	4000	1979 – 2010
Minimum RH	Idaho_rmin	1	METDATA	4000	1979 – 2010
Specific humidity	Idaho_sph	1	METDATA	4000	1979 – 2010
Minimum temperature	Idaho_tmmn	1	METDATA	4000	1979 – 2010
Maximum temperature	Idaho_tmmx	1	METDATA	4000	1979 – 2010
Historical land use	usgsLUhist	1	USGS LULC	30	1975 – 1980
Soil moisture	SMOS	1	NASA SMOS	15,000	2010
Ecoregion	US_L4NAME	4	Ecoregions	1:250000	2012
Biophysical setting	us_110bps	1	LANDFIRE	30	2009
Forest canopy density	us_110cbd	1	LANDFIRE	30	2009
Forest canopy height	us_110cbh	1	LANDFIRE	30	2009
Canopy cover	us_110cc	1	LANDFIRE	30	2009
Canopy height	us_110ch	1	LANDFIRE	30	2009
Vegetation cover	us_110evc	1	LANDFIRE	30	2009
Vegetation height	us_110evh	1	LANDFIRE	30	2009
Vegetation type	us_110evt	1	LANDFIRE	30	2009
Fire regime groups	us_110frg	1	LANDFIRE	30	2009
Mean fire return interval	us_110mfri	1	LANDFIRE	30	2009
Percent low severity	us_110pls	1	LANDFIRE	30	2009
Percent mixed severity	us_110pms	1	LANDFIRE	30	2009
Replacement severity	us_110prs	1	LANDFIRE	30	2009
Succession class	us_110sclass	1	LANDFIRE	30	2009
Disturbance	us_dist[year]	1	LANDFIRE	30	2009
Fuel disturbance	us_fdist[year]	1	LANDFIRE	30	2009
Vegetation disturbance	us_vdist[year]	1	LANDFIRE	30	2009
Land cover	nlcdLU[year]	1	NLCD	30	2001
EVI	modEVI_Annual	1	MODIS4NACP	500	2005
FPAR	modFPAR_Annual	1	MODIS4NACP	500	2005
GPP	modGPP_Annual	1	MODIS4NACP	500	2005
Leaf area index	modLAI_Annual	1	MODIS4NACP	500	2005
NDVI	modNDVI_Annual	1	MODIS4NACP	500	2005
Hydrologic soil groups	HSG	1	HYSOGs250m	250	2018

[a] Abbrevations: TWI, topographic wetness index; RH, mean annual maximum relative humidity; EVI, mean annual enhanced vegetation index; FPAR, mean annual fraction of photosynthetically active radiation; GPP, mean annual gross primary production; NDVI, mean annual normalized difference vegetation index; USGS, United States Geological Survey; United States Department of Agriculture; NGM, national geologic map; CGIAR, Consultative Group for International Agricultural Research; SRTM, shuttle radar topography mission; METDATA, meteorological data; N, number of variables; LULC, land use land cover; NASA, National Aeronautics and Space Administration; SMOS, mean annual soil moisture and oceanic salinity; MODIS, mean annual moderate resolution imaging spectroradiometer; HYSOGs, hydrologic soil groups.

indicates that these subsets resemble the whole set (N = 322). The Kolmogorov-Smirnov test confirms that topsoil (p = 0.66) and subsoil (p = 0.51) training sets share a common distribution with their respective validation sets.

3.1. Feature selection and model evaluation

The distribution of top- and subsoil carbon are modeled with parsimonious covariate sets, which were identified with recursive feature elimination. Predictors include mean annual precipitation and minimum temperature (Abatzoglou, 2013), soil moisture (Kerr et al., 2001), soil nitrogen and sand content (Ramcharan et al., 2018), gamma ray emissions of potassium (^{40}K) and thorium (^{232}Th) (Duval et al., 2005), and elevation (Jarvis et al., 2008). Evaluation of the ALE plots (Fig. 5) indicates that precipitation has the largest effect for predicting topsoil carbon variability, followed by soil nitrogen, gamma-ray emissions of ^{40}K, mean annual soil moisture, and sand content. Sand content had the largest effect on subsoil carbon predictions, followed by gamma-ray emissions of ^{232}Th, elevation, mean annual soil moisture, and mean annual minimum temperature.

Model evaluation with the validation set indicates that random forest was able to explain 69% and 67% of top- and subsoil carbon variability, respectively (Table 3). This corresponds to a RMSE of 0.77 kg m^{-2} for topsoil carbon, and 1.3 kg m^{-2} for subsoil carbon. Predicted topsoil carbon—which ranges from 1.7 to 9.8 kg m^{-2} with a mean value of 4.0 ± 0.8 kg m^{-2}—compares well with measured soil carbon, which ranges from 1.1 to 12.6 kg m^{-2} with a

Please cite this article as: C. W. Ross, S. Grunwald, J. G. Vogel et al., Accounting for two-billion tons of stabilized soil carbon, Science of the Total Environment, https://doi.org/10.1016/j.scitotenv.2019.134615

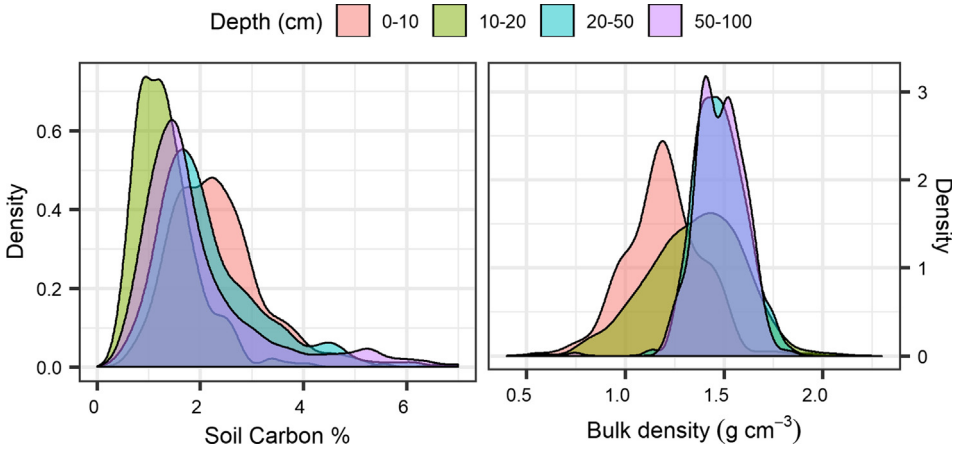


Fig. 3. Kernel density estimates of measured soil organic carbon concentration (left) and soil bulk density (right) for each depth interval.

Table 2
Descriptive statistics for topsoil (0 to 20 cm) and subsoil (20 to 100 cm) carbon.

		Soil carbon (kg m^{-2})							
Dataset		Min	Mean	Med	Max	SD	Kurtosis	Variance	Skewness
Whole	Topsoil	1.1	3.7	3.5	12.6	1.5	5.6	2.2	1.7
	Subsoil	1.3	4.3	3.6	22.0	2.5	9.2	6.1	2.4
Training	Topsoil	1.1	3.6	3.4	12.6	1.4	6.6	2.1	1.7
	Subsoil	1.4	4.3	3.6	22.0	2.4	12.7	5.8	2.7
Validation	Topsoil	1.5	3.8	3.7	10.3	1.6	3.8	2.5	1.6
	Subsoil	1.3	4.3	3.5	14.0	2.6	3.2	6.9	1.8

Min, minimum; Med, median; Max, Maximum; SD, standard deviation.

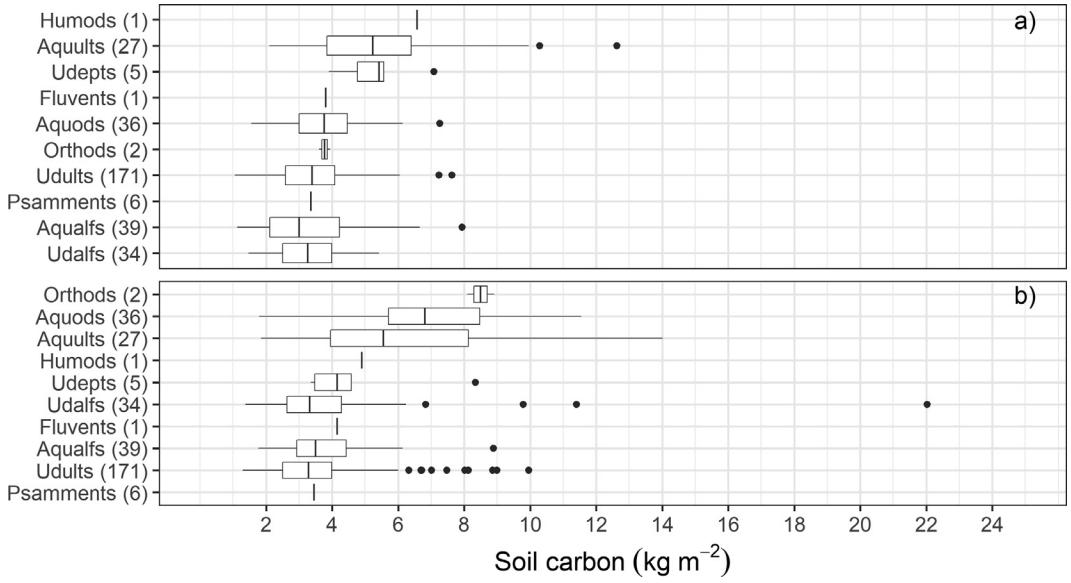


Fig. 4. a) Topsoil (0 – 20 cm) and b) subsoil (20 – 100 cm) carbon content by suborder as sampled from Tier 2 research sites (N = 322). Numbers in parenthesis denote the number of observations for each suborder.

mean of 3.7 ± 1.5 kg m^{-2} (Table 2). Likewise, predicted and measured subsoil carbon were comparable, which ranged from 2.5 to 9.6 kg m^{-2} with a mean value of 4.6 ± 1.0 kg m^{-2}, and from 1.3 to 22.0 kg m^{-2} with a mean value of 4.3 ± 2.5 kg C m^{-2}, respectively.

3.2. Soil carbon stock

From our prediction maps (Fig. 6), we estimate that 2.6 ± 0.8 Pg C of soil carbon is currently stabilized in the top 100 cm of soil in production forestlands, which cover 34.7 million ha of land (Marsik et al., 2018). Although the concentration of soil carbon

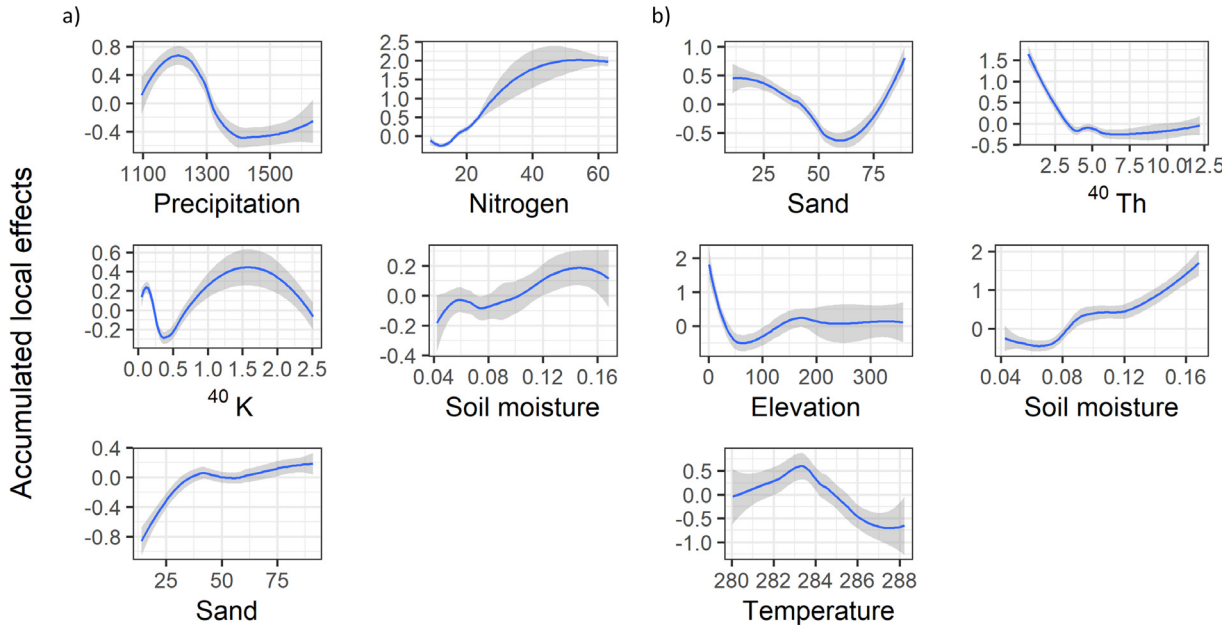


Fig. 5. Accumulated local effects for a) topsoil (0 – 20 cm) and subsoil (20 – 100 cm) covariate sets, which are ranked from highest feature importance (top left) to lowest feature importance (bottom) for each panel. The x-axis represents the units of the independent variables, and the y-axis represents the mean effect each independent variable has on the prediction. Variables include mean annual precipitation (mm), soil nitrogen (mg/g), mean annual soil moisture (m^3 m^{-3}), sand content (%), and gamma ray emissions of potassium (% K). Panel (b) shows the accumulated local effects for the subsoil model, which includes sand content (%), gamma ray emissions of thorium (ppm eTh), elevation (m), soil moisture (m^3 m^{-3}), and mean annual minimum temperature (K).

Table 3
Top- and subsoil model evaluation with training and validation sets for two soil depth intervals (0–20 cm and 20–100 cm).

	0 – 20 cm		20 – 100 cm	
	Training	Validation	Training	Validation
R^2	0.86	0.69	0.79	0.67
RMSE	0.63	0.77	1.28	1.29
RPIQ	2.64	2.10	1.50	1.96

R^2, coefficient of determination; RMSE, Root mean square error; RIPQ, Ratio of performance to inter-quartile distance. The training and validation set consisted of 225 and 97 observations, respectively.

declines with depth, we show that subsoils account for *ca.* 54% (1.4 ± 0.4 Pg) of the total carbon stock, and that this carbon is predictably distributed with a relatively small covariate set. These results are in general agreement with other estimates, such that previous assessments have reported that *ca.* half of Earth's soil carbon is stabilized in the subsoil (Jobbágy and Jackson, 2000; Rumpel and Kögel-Knabner, 2011).

Soil carbon generally increases with proximity to the coast, although there is considerable spatial variability as evidenced in Fig. 6. While our spatial patterns are in general agreement with those produced by Hengl and colleagues (2017), further evaluation indicated that the SoilGrids predictions tended to overestimate carbon (0 to 100 cm) in this system. When tested against our validation set, the SoilGrids predictions achieved an R^2 of 0.31 and a RMSE of 7.6 kg C m^{-2}. Hengl and colleagues (2017) were able to explain 63.5% of the variation in soil carbon globally; however, our comparison reveals that SoilGrids was unable to explain the regional-scale variability in our system of interest to any satisfying degree. Integrating regional-scale soil carbon measurements with a data-mining framework allowed us to optimize our models by identifying the most important predictors for our system and scale of analysis. This comparison not only underscores the need for regional-scale models, but also supports the usefulness of strategic feature selection for model optimization in specific regions or scales of analysis.

4. Discussion

This study integrates data mining with extensive field sampling to express the region-wide variation of soil carbon in pine plantations across the southeastern US. We identify a parsimonious, yet highly predictive covariate set by utilizing strategic feature selection. For our topsoil model, precipitation, nitrogen, and ^{40}K had the largest effect for predicting carbon variability. Examination of the ALE plots indicates that precipitation and ^{40}K have non-linear relationships with topsoil carbon across the region, while soil moisture and nitrogen show a strong positive association (Fig. 5). The positive effect of water availability, and more importantly, nitrogen availability is well established for productivity and carbon sequestration in southern pine forests (Bracho et al., 2018; Jokela et al., 2004; Vogel et al., 2011). Fertilization increases both net primary productivity and biomass, and often the turnover of fine roots, which can aid soil carbon sequestration (Haile et al., 2010) by enhancing soil structure (Sarkhot et al., 2008) (e.g., increased aggregation).

Soil moisture has a strong positive association with subsoil carbon as well; however, sand content, elevation, and ^{232}Th have slightly stronger effects according to the ALE plots (Fig. 5). The gamma-ray data correspond primarily to the mineralogy and geochemistry of the parent material in erosional landscapes—such as the Piedmont region. In depositional landscapes like the Coastal Plain, gamma ray emissions correspond primarily to the geochemistry and mineralogy of the parent material from which the sediments were derived (Wilford and Minty, 2006).

Indeed, the transition from crystalline-derived soils in the Piedmont to sedimentary-derived soils of the Coastal Plain was effectively captured by the gamma ray data, as evidenced in Fig. 6. Carbon stabilization in the Piedmont is heavily influenced by the regions moderately deep and well-drained soils that formed in

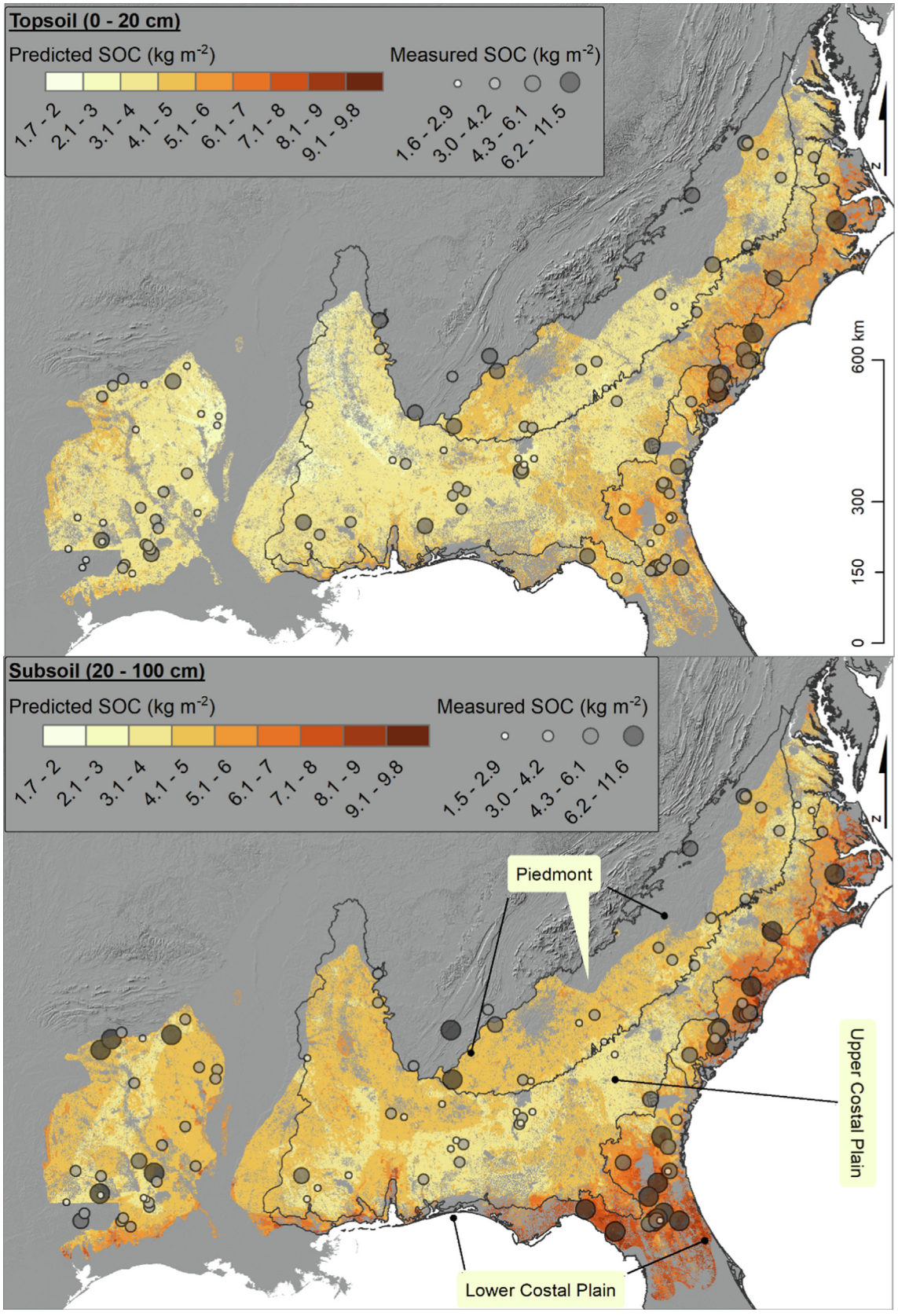


Fig. 6. Top- and subsoil organic carbon predictions produced from the parsimonious covariate sets at 250 × 250 m resolution. Predictions were masked to production forestland as shown in Fig. 1.

Please cite this article as: C. W. Ross, S. Grunwald, J. G. Vogel et al., Accounting for two-billion tons of stabilized soil carbon, Science of the Total Environment, https://doi.org/10.1016/j.scitotenv.2019.134615

material weathered from metamorphosed granite (i.e., granitic gneiss). Soil texture in the Piedmont is generally silty clay loam or clay loam, which effectively aids soil carbon stabilization via adsorption of organic matter to the clay surface. Conversely, Upper and Lower Coastal Plain soils formed primarily from parent materials of marine deposits, and therefore tend to be sandy and nutrient poor. However, the density of soil carbon is generally greater in the Lower Coastal Plain (Fig. 6), where carbon stabilization is aided by low landscape positions (i.e., elevation), high mean-annual precipitation, and shallow water tables. These factors have a profound effect on soil development, which is reflected by the extent of 'wet soils' (e.g., Aqualfs, Aquults, and Aquods) and carbon-rich spodic (Bh) horizons—a key diagnostic feature that is ubiquitous to soils of the lower Coastal Plain (Gonzalez et al., 2018).

4.1. Implications

While the coordinated implementation of silvicultural treatments has increased productivity and thus carbon sequestration in southeastern US pine plantations, global warming and the associated changes in climate may alter forest productivity, structure, and species geographic range. Minimum temperature currently constrains productivity at the northern extent of the region, leading to the assumption that an expanded growing season in response to warmer temperatures could enhance primary production (Nedlo et al., 2009). However, the response of soil carbon to increased mean annual temperature is not yet fully resolved as clearly demonstrated by the lack of temperature as a topsoil carbon predictor, as well as the non-linear response to temperature in the subsoil (Fig. 5). This is an important consideration in the context of forest management and carbon cycle science in the southeastern US, as the region is expected to be warmer and potentially drier in the coming decades (Ingram et al., 2013).

Although precipitation is expected to increase across the southeastern US in response to warmer temperatures, the region will likely experience more intense, but less-frequent rainfall events during the growing season (Li et al., 2010). As such, soil moisture and water availability will likely decrease as drought conditions increase in response to global warming, as well as increased runoff during high-intensity rainfall events in areas that have moderately-high and high-runoff potential (Ross et al., 2018). Loblolly pine is a moderately drought tolerant species; however, the strong positive association of soil moisture with top- and subsoil carbon suggests that these soils could function as a net source of atmospheric CO_2 in response to long-term drought conditions.

There is also concern that the negative effects of drought on net primary productivity could be exacerbated due to increased leaf area in response to nitrogen and phosphorus fertilization, thus increasing evapotranspiration and water stress in these systems. While the interactive effects of drought and fertilization have rarely been investigated experimentally, Bracho and colleagues (2018) recently reported that fertilization did not exacerbate the effects of experimental drought conditions (30% throughfall reduction) in managed loblolly pine plantations at four locations across the southeastern US. Total and heterotrophic soil respiration reportedly decreased in response to fertilization treatments; however, it remains unclear if this suppression had any quantifiable effect on soil carbon, and if soil carbon stock is stabilized against future changes in temperature, precipitation, and management. Results from our analysis indicate that nitrogen has a strong positive association with topsoil carbon (Fig. 5), suggesting that land management (i.e., fertilization) has had a positive effect on soil carbon stabilization in production forestland across the region.

Identifying associations that are important for carbon stabilization in specific regions and/or ecosystems is critical to further resolving carbon dynamics in response to global change. While feature selection was able to reduce model uncertainty by identifying important predictors for our system and scale and of analysis, further research is required to fully resolve soil-carbon dynamics in response to global change. For example, is soil carbon in the Lower Coastal Plain stabilized against potentially lower water tables due to population growth and increased demand for food and water? Or will warmer and drier conditions increase carbon mineralization and atmospheric CO_2? How will the interactive effects of increased nitrogen deposition (or fertilization) and drought conditions affect soil carbon in production forestlands?

5. Conclusions

We demonstrate the application of strategic feature selection to identify covariates that are important for soil-carbon stabilization across a large and highly-variable region. We opted for a parsimonious covariate-set (N = 5) to increase model interpretation while avoiding the "curse of dimensionality". Mean annual precipitation and gamma-ray emissions of ^{40}K have non-linear associations with topsoil carbon, while sand content, nitrogen, and soil moisture show strong, positive associations. Although temperature is often used to explain soil-carbon variation, our results indicate that other covariates were more important for explaining topsoil-carbon stabilization. In the subsoil, topo-edaphic conditions and soil moisture were more important for explaining carbon stabilization, which was primarily attributed to landscape position and water-table depth. From our spatial predictions, we estimated that 2.6 ± 0.8 Pg of soil carbon is currently stabilized in production forestlands, which corresponds to 34.7 million ha. The data-mining strategy presented here has significance for other modeling applications, as it is often not known which predictors will achieve the best data-model fit in a given region due to scale-dependent relationships.

Declaration of Competing Interest

The authors declare that there is no conflict of interest regarding the publication of this article.

Acknowledgements

The Pine Integrated Network: Education, Mitigation, and Adaptation project (PINEMAP) was a Coordinated Agricultural Project funded by the USDA National Institute of Food and Agriculture [Award #2011-68002-30185]. We would like to thank all PINEMAP team members for their contributions, with a special thanks to Marshall A. Laviner, Madison K. Akers, Joshua Cucinella, Tom Fox, Risa Patarasuk, and Beijing Cao for their contributions to this research.

References

Abatzoglou, J.T., 2013. Development of gridded surface meteorological data for ecological applications and modelling. Int. J. Climatol. 33, 121–131. https://doi.org/10.1002/joc.3413.

Batjes, N.H., 2014. Total carbon and nitrogen in the soils of the world. Eur. J. Soil Sci. 65, 10–21. https://doi.org/10.1111/ejss.12114_2.

Bellman, R., 1961. Curse of dimensionality. Adapt. Control Process. Guid. Tour Princet, NJ.

Bracho, R., Vogel, J.G., Will, R.E., Noormets, A., Samuelson, L.J., Jokela, E.J., Gonzalez-Benecke, C.A., Gezan, S.A., Markewitz, D., Seiler, J.R., Strahm, B.D., Teskey, R.O., Fox, T.R., Kane, M.B., Laviner, M.A., McElligot, K.M., Yang, J., Lin, W., Meek, C.R., Cucinella, J., Akers, M.K., Martin, T.A., 2018. Carbon accumulation in loblolly pine plantations is increased by fertilization across a soil moisture availability gradient. For. Ecol. Manag. 424, 39–52. https://doi.org/10.1016/j.foreco.2018.04.029.

Breiman, L., 2001. Random forests. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324.

Chen, S., Wang, W., Xu, W., Wang, Yang, Wan, H., Chen, D., Tang, Z., Tang, X., Zhou, G., Xie, Z., Zhou, D., Shangguan, Z., Huang, J., He, J.-S., Wang, Yanfen, Sheng, J.,

Tang, Li, X., Dong, M., Wu, Y., Wang, Q., Wang, Z., Wu, J., Chapin, F.S., Bai, Y., 2018. Plant diversity enhances productivity and soil carbon storage. Proc. Natl. Acad. Sci. 115, 4027–4032. https://doi.org/10.1073/pnas.1700298114.

Davidson, E.A., 2016. Projections of the soil-carbon deficit. Nature 540, 47–48. https://doi.org/10.1038/540047a.

Domingos, P., 2012. A few useful things to know about machine learning. Commun ACM 55, 78–87. https://doi.org/10.1145/2347736.2347755.

Duval, J.S., Carson, J.M., Holman, P.B., Darnley, A.G., 2005. Terrestrial radioactivity and gamma-ray exposure in the United States and Canada (Open File Report). US Geological Survey.

Easterling, W.E., 1997. Why regional studies are needed in the development of full-scale integrated assessment modelling of global change processes. Glob. Environ. Change 7, 337–356. https://doi.org/10.1016/S0959-3780(97)00016-2.

Eglin, T., Ciais, P., Piao, S.L., Barre, P., Bellassen, V., Cadule, P., Chenu, C., Gasser, T., Koven, C., Reichstein, M., Smith, P., 2010. Historical and future perspectives of global soil carbon response to climate and land-use changes. Tellus B 62, 700–718. https://doi.org/10.1111/j.1600-0889.2010.00499.x.

Eskelinen, A., Harrison, S.P., 2015. Resource colimitation governs plant community responses to altered precipitation. Proc. Natl. Acad. Sci. 112, 13009–13014. https://doi.org/10.1073/pnas.1508170112.

Gianelle, D., Oechel, W., Miglietta, F., Rodeghiero, M., Sottocornola, M., 2010. Cataloguing soil carbon stocks 1476 1476 Science 330. https://doi.org/10.1126/science.330.6010.1476-c.

Godwin, D.R., Kobziar, L.N., Robertson, K.M., 2017. Effects of fire frequency and soil temperature on soil CO2 efflux rates in old-field pine-grassland forests. Forests 8, 274. https://doi.org/10.3390/f8080274.

Gonzalez, Y.N., Bacon, A.R., Harris, W.G., 2018. A billion tons of unaccounted for carbon in the southeastern United States. Geophys. Res. Lett. 45, 7580–7587. https://doi.org/10.1029/2018GL077540.

Grunwald, S., Thompson, J., Boettinger, J., 2011. Digital soil mapping and modeling at continental scales: finding solutions for global issues. Soil Sci. Soc. Am. J. 75, 1201–1213.

Haile, S.G., Nair, V.D., Nair, P.K.R., 2010. Contribution of trees to carbon storage in soils of silvopastoral systems in Florida, USA. Glob. Change Biol. 16, 427–438. https://doi.org/10.1111/j.1365-2486.2009.01981.x.

Hengl, T., de Jesus, J.M., Heuvelink, G.B.M., Gonzalez, M.R., Kilibarda, M., Blagotić, A., Shangguan, W., Wright, M.N., Geng, X., Bauer-Marschallinger, B., Guevara, M.A., Vargas, R., MacMillan, R.A., Batjes, N.H., Leenaars, J.G.B., Ribeiro, E., Wheeler, I., Mantel, S., Kempen, B., 2017. SoilGrids250m: Global gridded soil information based on machine learning. PLOS One 12,. https://doi.org/10.1371/journal.pone.0169748 e0169748.

Hilton, T.W., Whelan, M.E., Zumkehr, A., Kulkarni, S., Berry, J.A., Baker, I.T., Montzka, S.A., Sweeney, C., Miller, B.R., Elliott Campbell, J., 2017. Peak growing season gross uptake of carbon in North America is largest in the Midwest USA. Nat. Clim. Change 7, 450–454. https://doi.org/10.1038/nclimate3272.

Ingram, K.T., Dow, K., Carter, L., Anderson, J., Sommer, E.K., 2013. Climate of the Southeast United States: variability, change, impacts, and vulnerability. Island Press, Washington, D.C..

Jarvis, A., Reuter, H.I., Nelson, A., Guevara, E., 2008. Hole-filled SRTM for the globe Version 4. CGIAR Consort. Spat. Inf, CGIAR-CSI, p. 2008.

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

Jokela, E.J., Dougherty, P.M., Martin, T.A., 2004. Production dynamics of intensively managed loblolly pine stands in the southern United States: a synthesis of seven long-term experiments. For. Ecol. Manag. 192, 117–130.

Kerr, Y.H., Waldteufel, P., Wigneron, J.P., Martinuzzi, J., Font, J., Berger, M., 2001. Soil moisture retrieval from space: the Soil Moisture and Ocean Salinity (SMOS) mission. IEEE Trans. Geosci. Remote Sens. 39, 1729–1735. https://doi.org/10.1109/36.942551.

Kottek, M., Grieser, J., Beck, C., Rudolf, B., Rubel, F., 2006. World map of the Köppen-Geiger climate classification updated. Meteorol. Z. 15, 259–263. https://doi.org/10.1127/0941-2948/2006/0130.

Kuhn, M., 2008. Building predictive models in R using the caret package. J. Stat. Softw. 28, 1–26.

Li, W., Li, L., Fu, R., Deng, Y., Wang, H., 2010. Changes to the North Atlantic subtropical high and its role in the intensification of summer rainfall variability in the southeastern United States. J. Clim. 24, 1499–1506. https://doi.org/10.1175/2010JCLI3829.1.

Liaw, A., Wiener, M., 2002. Classification and regression by randomForest. R News 2, 18–22.

Lu, X., Kicklighter, D.W., Melillo, J.M., Reilly, J.M., Xu, L., 2015. Land carbon sequestration within the conterminous United States: regional- and state-level analyses. J. Geophys. Res. Biogeosciences 120, 379–398. https://doi.org/10.1002/2014JG002818.

Luo, Y., Ahlström, A., Allison, S.D., Batjes, N.H., Brovkin, V., Carvalhais, N., Chappell, A., Ciais, P., Davidson, E.A., Finzi, A., Georgiou, K., Guenet, B., Hararuk, O., Harden, J.W., He, Y., Hopkins, F., Jiang, L., Koven, C., Jackson, R.B., Jones, C.D., Lara, M.J., Liang, J., McGuire, A.D., Parton, W., Peng, C., Randerson, J.T., Salazar, A., Sierra, C. A., Smith, M.J., Tian, H., Todd-Brown, K.E.O., Torn, M., van Groenigen, K.J., Wang, Y.P., West, T.O., Wei, Y., Wieder, W.R., Xia, J., Xu, Xia, Xu, Xiaofeng, Zhou, T., 2015. Towards more realistic projections of soil carbon dynamics by Earth system models. Glob. Biogeochem. Cycles 2015GB005239. https://doi.org/10.1002/2015GB005239

Marsik, M., Staub, C.G., Kleindl, W.J., Hall, J.M., Fu, C.-S., Yang, D., Stevens, F.R., Binford, M.W., 2018. Regional-scale management maps for forested areas of the southeastern United States and the US Pacific Northwest. Sci. Data 5,. https://doi.org/10.1038/sdata.2018.165 180165.

Massey, F.J., 1951. The Kolmogorov-Smirnov test for goodness of fit. J. Am. Stat. Assoc. 46, 68–78. https://doi.org/10.2307/2280095.

Miller, B.A., Koszinski, S., Wehrhan, M., Sommer, M., 2015. Impact of multi-scale predictor selection for modeling soil properties. Geoderma 239–240, 97–106. https://doi.org/10.1016/j.geoderma.2014.09.018.

Molnar, C., Bischl, B., Casalicchio, G., 2018. iml: An R package for interpretable machine learning. JOSS 3, 786 https://doi.org/10.21105/joss.00786.

Mulder, V.L., Lacoste, M., Richer-de-Forges, A.C., Martin, M.P., Arrouays, D., 2016. National versus global modelling the 3D distribution of soil organic carbon in mainland France. Geoderma 263, 16–34. https://doi.org/10.1016/j.geoderma.2015.08.035.

Nedlo, J.E., Martin, T.A., Vose, J.M., Teskey, R.O., 2009. Growing season temperatures limit growth of loblolly pine (Pinus taeda L.) seedlings across a wide geographic transect. Trees 23, 751–759. https://doi.org/10.1007/s00468-009-0317-0.

Noormets, A., Epron, D., Domec, J.C., McNulty, S.G., Fox, T., Sun, G., King, J.S., 2015. Effects of forest management on productivity and carbon sequestration: a review and hypothesis. For. Ecol. Manag., Carbon, water and nutrient cycling in managed forests 355, 124–140. https://doi.org/10.1016/j.foreco.2015.05.019.

Oswalt, S.N., Smith, Brad, W., Miles, D.P., Pugh, A.S., 2014. Forest resources of the United States, 2012: a technical document supporting the Forest Service 2010 update of the RPA Assessment. Gen Tech Rep WO-91 Wash. DC US Dep. Agric. For. Serv. Wash. Off. 218, P 91.

Pimentel, D., 2006. Soil erosion: A food and environmental threat. Environ. Dev. Sustain. 8, 119–137. https://doi.org/10.1007/s10668-005-1262-8.

Core Team, R., 2019. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria.

Ramcharan, A., Hengl, T., Nauman, T., Brungard, C., Waltman, S., Wills, S., Thompson, J., 2018. Soil property and class maps of the conterminous United States at 100-meter spatial resolution. Soil Sci. Soc. Am. J. 82, 186–201.

Ross, C.W., Grunwald, S., Myers, D.B., 2013. Spatiotemporal modeling of soil organic carbon stocks across a subtropical region. Sci. Total Environ. 461–462, 149–157. https://doi.org/10.1016/j.scitotenv.2013.04.070.

Ross, C.W., Grunwald, S., Myers, D.B., Xiong, X., 2016. Land use, land use change and soil carbon sequestration in the St. Johns River Basin, Florida, USA. Geoderma Reg. 7, 19–28. https://doi.org/10.1016/j.geodrs.2015.12.001.

Ross, C.W., Prihodko, L., Anchang, J., Kumar, S., Ji, W., Hanan, N.P., 2018. HYSOGs250m, global gridded hydrologic soil groups for curve-number-based runoff modeling. Sci. Data 5,. https://doi.org/10.1038/sdata.2018.91 180091.

Roy, J., Picon-Cochard, C., Augusti, A., Benot, M.-L., Thiery, L., Darsonville, O., Landais, D., Piel, C., Defossez, M., Devidal, S., Escape, C., Ravel, O., Fromin, N., Volaire, F., Milcu, A., Bahn, M., Soussana, J.-F., 2016. Elevated CO2 maintains grassland net carbon uptake under a future heat and drought extreme. Proc. Natl. Acad. Sci. 113, 6224–6229. https://doi.org/10.1073/pnas.1524527113.

Rumpel, C., Kögel-Knabner, I., 2011. Deep soil organic matter—a key but poorly understood component of terrestrial C cycle. Plant Soil 338, 143–158. https://doi.org/10.1007/s11104-010-0391-5.

Sanderman, J., Hengl, T., Fiske, G.J., 2017. Soil carbon debt of 12,000 years of human land use. Proc. Natl. Acad. Sci. 201706103. https://doi.org/10.1073/pnas.1706103114.

Sarkhot, D.V., Jokela, E.J., Comerford, N.B., 2008. Surface soil carbon size–density fractions altered by loblolly pine families and forest management intensity for a Spodosol in the southeastern US. Plant Soil 307, 99–111. https://doi.org/10.1007/s11104-008-9587-3.

Schmidt, M.W.I., Torn, M.S., Abiven, S., Dittmar, T., Guggenberger, G., Janssens, I.A., Kleber, M., Kögel-Knabner, I., Lehmann, J., Manning, D.A.C., Nannipieri, P., Rasse, D.P., Weiner, S., Trumbore, S.E., 2011. Persistence of soil organic matter as an ecosystem property. Nature 478, 49–56. https://doi.org/10.1038/nature10386.

Soil Survey Staff, 2013. Gridded soil survey geographic (gSSURGO) database for the conterminous United States. United States Department of Agriculture, Natural Resources Conservation Service.

Vogel, J.G., Suau, L.J., Martin, T.A., Jokela, E.J., 2011. Long-term effects of weed control and fertilization on the carbon and nitrogen pools of a slash and loblolly pine forest in north-central Florida. Can. J. For. Res. 41, 552–567. https://doi.org/10.1139/X10-234.

Watson, R.T., Noble, I.R., Bolin, B., Ravindranath, N., Verardo, D.J., Dokken, D.J., et al., 2000. Land use, land-use change and forestry: a special report of the Intergovernmental Panel on Climate. Change. Cambridge University Press.

Wear, D.N., Greis, J.G., 2013. The southern forest futures project: Technical report. USDA-Forest Service, Southern Research Station, Ashville, NC.

Wilford, J., Minty, B., 2006. In: Developments in Soil Science, Digital Soil Mapping. Elsevier, pp. 207–610. https://doi.org/10.1016/S0166-2481(06)31016-1.

Will, R.E., Fox, T., Akers, M., Domec, J.-C., González-Benecke, C., Jokela, E.J., Kane, M., Laviner, M.A., Lokuta, G., Markewitz, D., McGuire, M.A., Meek, C., Noormets, A., Samuelson, L., Seiler, J., Strahm, B., Teskey, R., Vogel, J., Ward, E., West, J., Wilson, D., Martin, T.A., 2015. A range-wide experiment to investigate nutrient and soil moisture interactions in loblolly pine plantations. Forests 6, 2014–2028. https://doi.org/10.3390/f6062014.

Xiong, X., Grunwald, S., Corstanje, R., Yu, C., Bliznyuk, N., 2016. Scale-dependent variability of soil organic carbon coupled to land use and land cover. Soil Tillage Res. 160, 101–109. https://doi.org/10.1016/j.still.2016.03.001.

Please cite this article as: C. W. Ross, S. Grunwald, J. G. Vogel et al., Accounting for two-billion tons of stabilized soil carbon, Science of the Total Environment, https://doi.org/10.1016/j.scitotenv.2019.134615

Xiong, X., Grunwald, S., Myers, D.B., Kim, J., Harris, W.G., Bliznyuk, N., 2015. Assessing uncertainty in soil organic carbon modeling across a highly heterogeneous landscape. Geoderma 251–252, 105–116. https://doi.org/10.1016/j.geoderma.2015.03.028.

Xiong, X., Grunwald, S., Myers, D.B., Kim, J., Harris, W.G., Comerford, N.B., 2014a. Holistic environmental soil-landscape modeling of soil organic carbon. Environ. Model. Softw. 57, 202–215. https://doi.org/10.1016/j.envsoft.2014.03.004.

Xiong, X., Grunwald, S., Myers, D.B., Ross, C.W., Harris, W.G., Comerford, N.B., 2014b. Interaction effects of climate and land use/land cover change on soil organic carbon sequestration. Sci. Total Environ. 493, 974–982. https://doi.org/10.1016/j.scitotenv.2014.06.088.

Please cite this article as: C. W. Ross, S. Grunwald, J. G. Vogel et al., Accounting for two-billion tons of stabilized soil carbon, Science of the Total Environment, https://doi.org/10.1016/j.scitotenv.2019.134615