SAND2022-12922

SANDIA REPORT
SAND20XX-XXXX
Printed Click to enter a date

Sandia National Laboratories

Improving and testing machine learning methods for benchmarking soil carbon dynamics representation of land surface models

Umakant Mishra, Sagar Gautam

Prepared by
Sandia National Laboratories
Albuquerque, New Mexico
87185 and Livermore,
California 94550

Issued by Sandia National Laboratories, operated for the United States Department of Energy by National Technology & Engineering Solutions of Sandia, LLC.

Available to DOE and DOE contractors from
U.S. Department of Energy
Office of Scientific and Technical Information
P.O. Box 62
Oak Ridge, TN 37831

Telephone: (865) 576-8401
Facsimile: (865) 576-5728
E-Mail: reports@osti.gov
Online ordering: http://www.osti.gov/scitech

Available to the public from
U.S. Department of Commerce
National Technical Information Service
5301 Shawnee Rd
Alexandria, VA 22312

Telephone: (800) 553-6847
Facsimile: (703) 605-6900
E-Mail: orders@ntis.gov
Online order: https://classic.ntis.gov/help/order-methods/

ABSTRACT

Representation of soil organic carbon (SOC) dynamics in Earth system models (ESMs) is a key source of uncertainty in predicting carbon climate feedbacks. The magnitude of this uncertainty can be reduced by accurate representation of environmental controllers of SOC stocks in ESMs. In this study, we used data of environmental factors, field SOC observations, ESM projections and machine learning approaches to identify dominant environmental controllers of SOC stocks and derive functional relationships between environmental factors and SOC stocks. Our derived functional relationships predicted SOC stocks with similar accuracy as the machine learning approach. We used the derived relationships to benchmark the coupled model intercomparison project phase six ESM representation of SOC stocks. We found divergent environmental control representation in ESMs in comparison to field observations. Representation of SOC in ESMs can be improved by including additional environmental factors and representing their functional relationships with SOC consistent with observations.

ACKNOWLEDGEMENTS

This study was supported by the Laboratory Directed Research and Development program of Sandia National Laboratories. Sandia National Laboratories is a multimission laboratory managed and operated by National Technology and Engineering Solutions of Sandia, LLC, a wholly owned subsidiary of Honeywell International, Inc., for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA-0003525.

CONTENTS

LIST OF FIGURES

This page left blank

EXECUTIVE SUMMARY

Representation of soil organic carbon (SOC) dynamics in Earth system models (ESMs) is a key source of uncertainty in predicting future carbon climate feedbacks. The magnitude of this uncertainty can be reduced by accurate representation of environmental controllers of SOC stocks in ESMs. In this study, we used data of environmental factors, field SOC observations, ESM projections, and machine learning approaches to (1) identify dominant environmental controllers of SOC stocks in observations and ESMs, (2) derive functional relationships between environmental factors and SOC stocks, (3) use the derived functional relationships to predict SOC stocks and compare prediction accuracy of the two approaches, and (4) benchmark the environmental control representation of SOC in coupled model intercomparison project phase six ESMs.

Our results show divergent environmental control representation of SOC stocks in ESMs in comparison to field observations. Out of the 50 environmental factors we investigated, 14 were identified as dominant environmental predictors of global SOC stocks. Our results show diurnal temperature, drought index, cation exchange capacity and precipitation as most important observed environmental controllers of global SOC stocks. Random Forest (RF) model prediction of observed SOC stocks at global scale resulted in R^2 and RMSE of 0.61 and 0.46 kg m^{-2} respectively. However, in ESMs, precipitation, temperature, and net primary productivity explained >96% variability of modeled SOC stocks. Our results show control of temperature on SOC stocks better constrained in ESMs in comparison to the controls of precipitation and net primary productivity. Representation of SOC in ESMs can be improved by including additional environmental factors in model representations and representing the functional relationships of environmental factors with SOC stocks consistent with observations.

Various federal agencies and private foundations have shown interest in supporting studies using ML. As a result, use of ML approaches is increasing in diverse scientific applications including SOC storage and dynamics. DOE-BER has shown interest in building a research program to use ML to enhance Earth system predictability. Our study documents a unique use of ML to advance the representation of SOC stocks in ESMs.

ACRONYMS AND TERMS

Acronym/Term	Definition
SOC	Soil organic carbon
ESM	Earth system model
ML	Machine learning
RF	Random forest
GAM	Generalized additive modeling
NLCD	National land cover database
NDVI	Normalized difference vegetation index
NPP	Net primary productivity
VBFI	Valley bottom flatness index
PET	Potential evapotranspiration
MSE	Mean squared error

1. INTRODUCTION

Soils store a large and dynamic fraction of global terrestrial carbon (Sulman et al., 2020), and affect many ecosystem services (Lal, 2013). Soils can act as sources or sinks of atmospheric CO_2, depending on land use, management interventions, and environmental conditions. Observation-based SOC stock estimates show large spatial heterogeneity (Batjes, 2016; Hengl et al., 2014). This observed spatial heterogeneity in SOC stocks is primarily controlled by the soil forming factors: climate, organisms, topography, parent material, and time (Jenny, 1941; McBratney et al., 2003). As a result, different combinations of these environmental factors have widely been used for spatial prediction of SOC stocks at different scales (Adhikari et al., 2020; Mishra et al., 2021; Vitharana et al., 2017). Despite their key role in determining the spatial heterogeneity of SOC stocks and regulating land-atmosphere exchanges of carbon, the control of these environmental factors on SOC stocks are not correctly characterized and represented in current land surface model process representations. As a result, land models poorly represent current SOC spatial heterogeneity (Carvalhais et al., 2014; Todd-Brown et al., 2013), which contributes to large uncertainty in predicting future carbon-climate feedbacks (Arora et al., 2020; Friedlingstein et al., 2014). Therefore, to reduce uncertainty in future carbon-climate feedback projections, it is critical to accurately (i.e., consistent with observations) represent environmental controllers of SOC stocks in ESMs.

A variety of approaches have been applied to predict the spatial heterogeneity and infer environmental controllers of SOC stocks (Lamichhane et al., 2019; Minasny et al., 2013). Among different approaches applied for spatial predictions of SOC stocks, linear regression and ordinary kriging have been most widely used approaches (Minasny et al., 2013; Zhang et al., 2017). Linear regressions quantify the strength and direction of relationships between environmental factors and SOC stocks and have been applied primarily due to their simplicity and ease of interpretation of the results obtained. Ordinary kriging uses the spatial autocorrelation among existing samples to predict the value of SOC stocks at an unsampled location.

However, several recent studies demonstrated use of nonlinear approaches to predict the spatial heterogeneity of SOC stocks. Among nonlinear methods, machine learning (ML) approaches are increasingly being applied to predict soil properties, including SOC stocks (Lamichhane et al., 2019; Padarian et al., 2020; Siewert, 2018). Heuvelink et al. (2020) used a quantile regression forest machine learning approach to predict the annual SOC stock of surface soils of Argentina between 1982 and 2017 and reported a larger temporal variation in comparison to the Intergovernmental Panel on Climate Change Tier 1 approach of predicting SOC change. Ottoy et al. (2017) compared four digital soil mapping approaches to predict SOC stocks at a regional scale and reported boosted regression trees achieved highest prediction accuracy. Authors identified drainage condition, soil type, and vegetation type as important environmental predictors of SOC stocks. Vos et al. (2017) used various data mining approaches to identify and interpret main factors that controlled the cropland SOC stocks. Authors reported land use, land-use history, clay content and electrical conductivity as main predictors of the topsoil SOC stocks, whereas bedrock material, relief and electrical conductivity were main predictors of the subsoil carbon stocks. Bui et al. (2009) reported that SOC in Australian agricultural soils were related to vegetation, biomass, soil moisture and temperature patterns. Authors reported that the structure in the multivariate relationships between environmental factors and soil properties were consistent with principles of pedogenesis and landscape ecology. In a review of digital soil mapping literature, Ma et al. (2019) documented that the pedological knowledge can be used in digital soil mapping and digital soil

mapping can also lead to new knowledge discovery regarding the soil formation. However, Wadoux et al. (2020a) noted that the knowledge discovery based on ML needs to be treated with caution. Interestingly, authors demonstrated how pseudo-covariates not related to any soil-forming factors and processes can also accurately predict soil organic carbon. Therefore, careful pre-selection and preprocessing of pedologically relevant environmental covariates and the posterior interpretation and evaluation of the recognized patterns can only provide meaningful insights.

More recently, ensembles of multiple approaches have been applied to improve the spatial prediction of SOC stocks (Riggers et al., 2019; Vašát et al., 2017). A recent study showed that the median prediction obtained from an ensemble of ML approaches better predicts the spatial heterogeneity of SOC stocks in comparison to individual ML or hybrid approaches, such as regression kriging (Mishra et al., 2020). In many previous studies, ML approaches were used to identify important environmental predictors and predict the spatial variation of SOC stocks. In a recent review of ML applications in soil science, Padarian et al. (2020) identified two primary research needs: (1) identification of parsimonious ML models and (2) interpretability of the applied ML models. Similarly, in another review, Wadoux et al. (2020b) identified the need to incorporate pedological knowledge in ML algorithms to make these approaches more relevant to soil science. These authors identified plausibility, interpretability, and explainability as the greatest challenges in using ML approaches in soil science.

Current ESMs, however, only use the effects of a limited number of environmental factors in representing SOC storage. A recent study that compared SOC stocks from multiple ESMs against observation has indicated that there is still large knowledge gap in both ESMs and observations (Georgiou et al., 2021). It is imperative to compare ESM simulations against global SOC datasets to evaluate model performance and identify key environmental controllers in representing global SOC storage. Benchmarking ESM simulations against observed data is a most common approach for model evaluation (Luo et al., 2012; Todd-Brown et al., 2013; Collier et al., 2018). Through comparing model simulations with observations, we diagnose model's strengths and deficiencies in simulating SOC storage and identify key factors for future improvement. The emerged understanding of SOC storage in benchmarking could further develop to new model structures (by identifying key processes) and new parameterizations (by quantifying key relationships between SOC and environmental variables) of ESMs. Thus, benchmarking analysis of ESMs is an effective tool to reduce uncertainties in predicting SOC dynamics and provide more trustworthy information for environmental management and policymakers (Lauer et al., 2017).

The specific objectives were to (1) use machine learning to select important environmental predictors of SOC stocks, (2) derive empirical relationships between environmental factors and SOC stocks, (3) use the derived functional relationships to predict SOC stocks and compare prediction accuracy of the two approaches, and (4) use the derived relationships to benchmark the environmental control representation in coupled model intercomparison project phase six ESMs.

2. METHODS

2.1 Soil organic carbon observations

We used field SOC measurements from the rapid carbon assessment project of the Natural Resources Conservation Service's Soil Science Division of the USDA (Soil Survey Staff and Loecke, 2016) to achieve our study objectives 1-3. That assessment project was designed to produce a robust estimate of SOC stocks in different kinds of soils and land uses across the conterminous United States based on consistent and dedicated soil sampling. Over 6200 sampling sites across the conterminous United States (Fig. S1) were established following a hierarchical sampling design consisting of major land resource areas as first-level strata, which were further stratified based on land use and land cover and soil types in a nested fashion. Soil samples at observation locations were collected from genetic horizons and were analyzed for SOC concentration and bulk density following the Soil Survey Laboratory Methods Manual (Burt, 2004; Grossman & Reinsch, 2002). However, this study considered SOC stock for only the top 30 cm of soil, calculated after correcting it for coarse fragments (Eq. 1). For soil samples with missing bulk density measurements, a pedotransfer function based on a RF approach was developed (Sequeira et al., 2014) and SOC stock was calculated as

$$SOC_{stk} = \left[(SOC \times BD \times D) \times \left(1 - \frac{CF}{100}\right)\right], \quad (1)$$

where SOCstk is the SOC stock (t C ha−1), SOC is the SOC concentration (g C 100g-soil-1), BD is the soil bulk density (g cm−3), D is the soil layer thickness (cm), and CF is the volumetric fraction of the coarse fragments.

To benchmark the ESM representation of SOC stocks (objective 4), we used the World Soil Information Service (WoSIS) datasets. The World Soil Information Service (WoSIS) compiled SOC profiles across the globe after quality assessment. The 2019 snapshot of WoSIS dataset conserved 111,380 soil profiles with SOC content information (unit: g C g^{-1}) at different soil depths (Batjes et al., 2020). Another dataset we used in this study was compiled from Mishra et al. (2021). This dataset contained 2,546 soil profiles with SOC stock (g C m^{-3}) information in permafrost regions in North America, northern Eurasia, and Qinghai-Tibet Plateau. In total, we used 113,926 soil profiles from these two data sources. Because not all the soil profiles in our database preserve SOC information that covers the whole 0 – 100 cm interval, eventually we used 54000 soil profiles reporting SOC stocks of 0 – 100 cm. Because values of SOC stock across profiles were highly skewed, we used their natural log values in this study.

2.2 Environmental predictors of soil organic carbon stocks

The storage and cycling of SOC stocks are controlled by multiple environmental variables including, climatic variables, soil type, topographic variables, and land cover and land use. For objective 1-3, we compiled 31 environmental variables from different sources and evaluated their usefulness as predictors of SOC in the study area (Table S1). These variables were representative

of major soil forming factors: climate, vegetation, topography, and parent material (Jenny, 1941; McBratney et al., 2003). Seven of the 31 variables were climatic variables, obtained from Parameter-elevation Regressions on Independent Slopes Model and global climate and weather data: the 30-yr (1981 to 2010) annual average of minimum, mean, maximum, and dewpoint temperatures; precipitation as rainfall; rainfall during the wettest and driest quarter in a year; and potential evapotranspiration. Six of the 31 variables described vegetation characteristics: land use, land cover, potential vegetation cover, remote sensing data (median value of surface reflectance during the growing season), net primary production, and ecological regions. Ten variables related to topography were derived from the national digital elevation model at 30-m spatial resolution that was resampled to 100-m grid scale for this study: elevation, slope aspect, slope length factor, multi-resolution valley bottom flatness index, melton ruggedness index, mid-slope position, wetness index, slope height, slope gradient, and valley depth. Five variables described soil environment (parent material and soil climate): soil types, surface geology, natural drainage condition, hydrological unit, and soil temperature regime. For these 31 environmental variables, vector layers were rasterized when necessary, and all the raster layers and point SOC observations were projected to a common Universal Transverse Mercator projection system (NAD 1983). The values of the environmental variables at sampling locations were then extracted and a matrix of SOC stock and 31 predictors (6123 rows, 34 columns) was created for modeling. All the categorical variables were converted to integer variables before using in this analysis.

For the ESM benchmarking (objective 4), we compiled 50 environmental variables from different sources and evaluated their usefulness as predictors of global SOC stocks. The climatic variables include annual average temperature, precipitation, evapotranspiration, drought, and its statistics for different temporal scales. The soil related variables include clay content, sand content, silt content, texture, pH, cation exchange capacity. Land cover variables include IGBP types, vegetation cover, and ESA land cover types. Topographical variables include Elevation and depth to bedrock.

2.3 Dimensionality reduction using Random Forest

We used a Random Forest (RF) regression approach to identify important environmental predictors of SOC stocks. RF is based on a decision tree model and consists of an ensemble of randomized classification and regression trees with a bootstrap aggregation (Breiman, 2001). In RF, a training data set is first randomly drawn with replacement from the original data set. Then, a decision tree is fitted to the training data set by randomly selecting a subset of the input variables at each branch split. Typically, only $p/3$ variables are used to decide a branch split for a regression tree, where p is the number of predictor variables. The process is repeated to build many uncorrelated trees, hence the name "forest", and the prediction is computed by averaging the predictions of each tree. RF is one of the most popular predictive models in ML due to its outstanding performance even with little parameter tuning (Hastie et al., 2001). The RF model was trained by using the "randomForest" package in R (Liaw & Wiener, 2002). The total number of regression trees (ntree) was set to 500, and mtry=10 (31/3) variables were randomly selected to compute a split at each branch. The number of minimum data points to stop growing a tree was set to nodesize=10. Because each tree of the RF is trained with a subset of the original data set, the

model accuracy can be evaluated using a K-fold validation approach. The SOC stock datasets values were positively skewed, and therefore were transformed using natural log function.

We used a 'greedy' approach (Edmonds, 1971) to identify uncorrelated sets of environmental predictors of SOC stocks. In the 'greedy' approach, the environmental predictors were first arranged according to the variable importance rank from the RF model. The Pearson's correlation coefficients between the environmental predictors were calculated, and environmental predictors with absolute value of the correlation coefficients larger than a threshold (taken as 0.6), were removed from the data set.

The variable importance was computed by the random permutation method, where one of the environmental variables is randomly permutated between the out-of-bag samples and the change in the prediction accuracy [R2 (1-Residual sum of square/Total sum of square) and root mean square error] due to the random permutation provides a measure for the importance of the environmental variable (Hastie et al., 2001). The permutation-based importance is one of the most common approaches to assess the relative importance between input variables in the RF approach.

2.4 Generalized Additive Models to derive functional relationships between environmental predictors and SOC stocks

RF is a powerful machine learning technique due to its strength in computing nonlinear relations between input and output variables. However, RF is essentially a "black box" model, which does not provide detailed information about the relationships between the input and output variables. The function between predictor variables and response is particularly challenging to tease apart. This makes it difficult to use RF to find a functional relationship between a particular environmental predictor and SOC, particularly when the data points are not uniformly distributed over the high-dimensional feature space. Therefore, we used a Generalized Additive Model (GAM) to derive functional relationships between the RF-identified environmental predictors and SOC stocks. In GAM, the relationship in the data can be modeled as (Hastie & Tibshirani, 1990; Hastie et al., 2001).

$$Y = C + \sum_{i=1}^{p} f_i(X_i). \quad (2)$$

Here, Y is the target variable, e.g., observed SOC; X_i is an environmental variable; f_i is a smooth function; and C is a constant, which is usually a mean of Y. GAMs can generalize multilinear regression, but without the linear assumptions. This is performed by replacing the linear β parameters of the form $Y = C + \sum_{i=1}^{p} \beta_i(X_i)$ with a smoothing function f, usually in the form of additive splines. This allows the influence of individual predictor variable to be decoupled and compared with target variable, without requiring linearity of relationship between the predictor variable and target variable. The thin plate spline is used for the smoother, $f_i(X_i)$ (Wood, 2003). For a one-dimensional problem, the smoothing function is found by minimizing

$$\sum_{i=1}^{N}\left(Y^i - f(X^i)\right)^2 + \lambda \int \left(\frac{d^2 f(x)}{dx^2}\right)^2 dx, \quad (3)$$

in which Y^i and X^i, respectively, denote the target and the input feature, *N* is the total number of data, λ is a penalty parameter. The function that minimizes (3) is given as

$$f(x) = \sum_{i=1}^{N} \delta_i \eta\left(|x - X^i|\right) + \sum_{j=1}^{2} \alpha_j \phi_j(x). \quad (4)$$

Here, δ_i and α_j are unknown parameters, ϕ_j is the *(j-1)*-th order polynomial, and $\eta(r) = r^3$. Furthermore, δ is approximated by a reduced order basis as $\delta = U_k \delta_k$, in which U_k is a rank-k matrix. The rank of U_k denotes the maximum degree of freedom of the thin plate spline. To prevent an overfitting, k is chosen to be four. The unknown parameters, $U_k, \delta_k,$ and, α, are estimated from the data by solving a regularized optimization problem as shown in Wood (2003). The GAM model (2) is then computed by iteratively computing the one-dimensional thin plate splines for each environmental variable, using the backfitting algorithm (Hastie et. al., 2001). For our analysis here we used the "mgcv" package in R to train a GAM model (Wood, 2017) using a Restricted Maximum Likelihood method, and a thin plate spline for the smooth functions (Wood, 2003).

2.5 Earth system model outputs

We downloaded and aggregated the SOC and environmental controller data from three ESMs, Community Earth System Model (CESM) (Hurrell et al., 2013), U.K. Earth System Model (UKESM) (Sellar et al., 2019), and Beijing climate center (BCC) (Wu et al., 2019), to evaluate the environmental controllers of baseline global SOC stocks. ESMs did not report depth-wise soil carbon projection, making direct comparison with depth-specific SOC observation difficult. Many models used in ESMs were designed to simulate the soil carbon for topsoil depth, we assumed that the simulated soil carbon is contained within 1 m of soil profile to simplify comparison with observations.

3. RESULTS AND DISCUSSIONS

3.1 Dominant environmental predictors of SOC stocks

The importance of all the environmental predictors of SOC stocks in descending order, as estimated by RF is provided in Figure 1a. The resulting variable importance shows that soil drainage has the dominant effect on continental US surface SOC stocks, followed by normalized difference vegetation index and dry-season precipitation. We also found that many of the environmental predictors used in this study were correlated with each other (Fig. 1b). While RF offers a good predictive model, it lacks the capability to identify multicollinearity in the environmental predictors (Mishra et al., 2020). Hence, as explained in section 2.4, we removed the correlated variables (resulting in 19 variables) and re-applied the RF approach with the reduced number of environmental predictors.

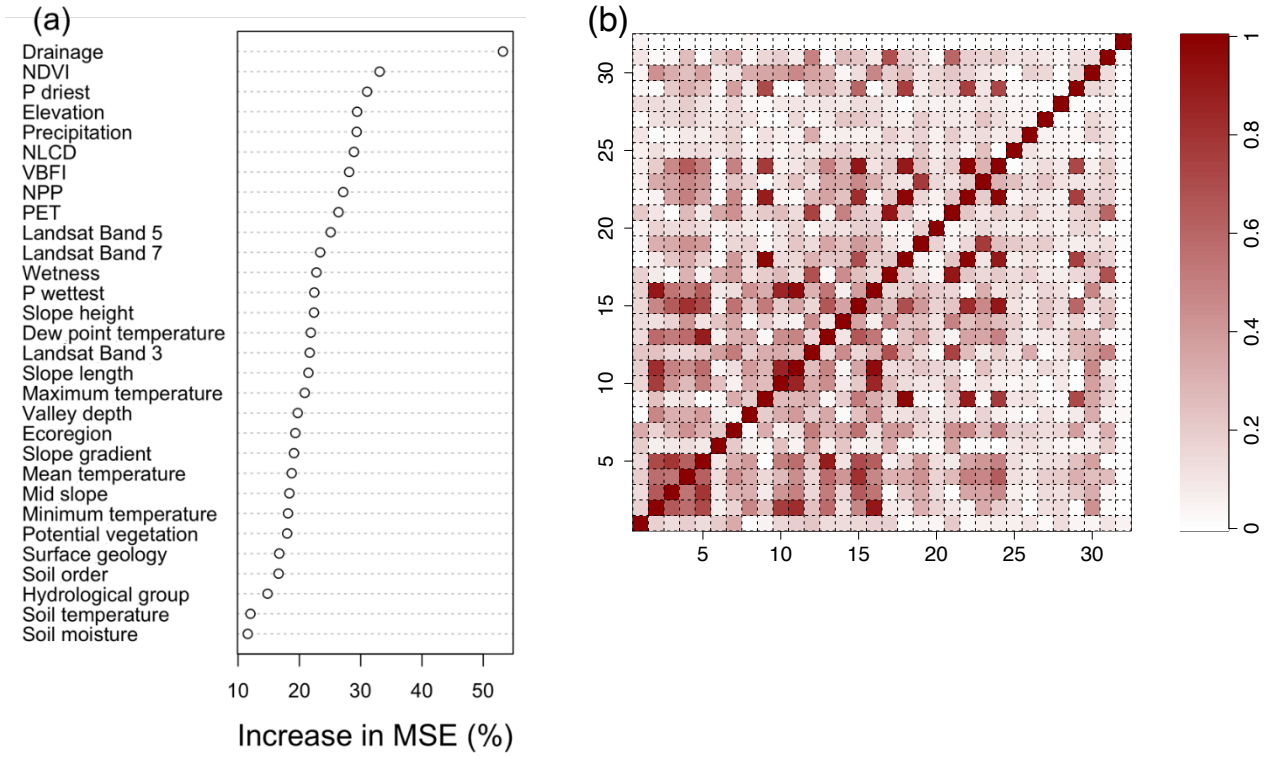

Figure 1. (a) Variable importance for the top 30 variables and (b) absolute values of the correlation coefficients between the variables. The index corresponds to the variable importance rank. MSE is mean squared error, NDVI is normalized difference vegetation index, P driest is precipitation in driest season, NLCD is national land cover database, VBFI is valley bottom flatness index, NPP is net primary productivity, PET is potential evapotranspiration, P wettest is precipitation of the wettest season.

Variable importance ranking changed after correlated environmental predictors were removed (Figure 2a), as did the incremental changes in R^2 with respect to the number of the environmental predictors (Figure 2b). We found significant improvement in the RF performance (R^2 and RMSE)

as the number of environmental predictors increased from 1 to 8 (with the predictors ordered by the RF-inferred importance; Figure 2b). However, after 12 environmental predictors, the improvement in model prediction accuracy was minimal. These results suggested that among all the environmental predictors we used, only 12 environmental predictors were the strongest predictors of SOC stocks.

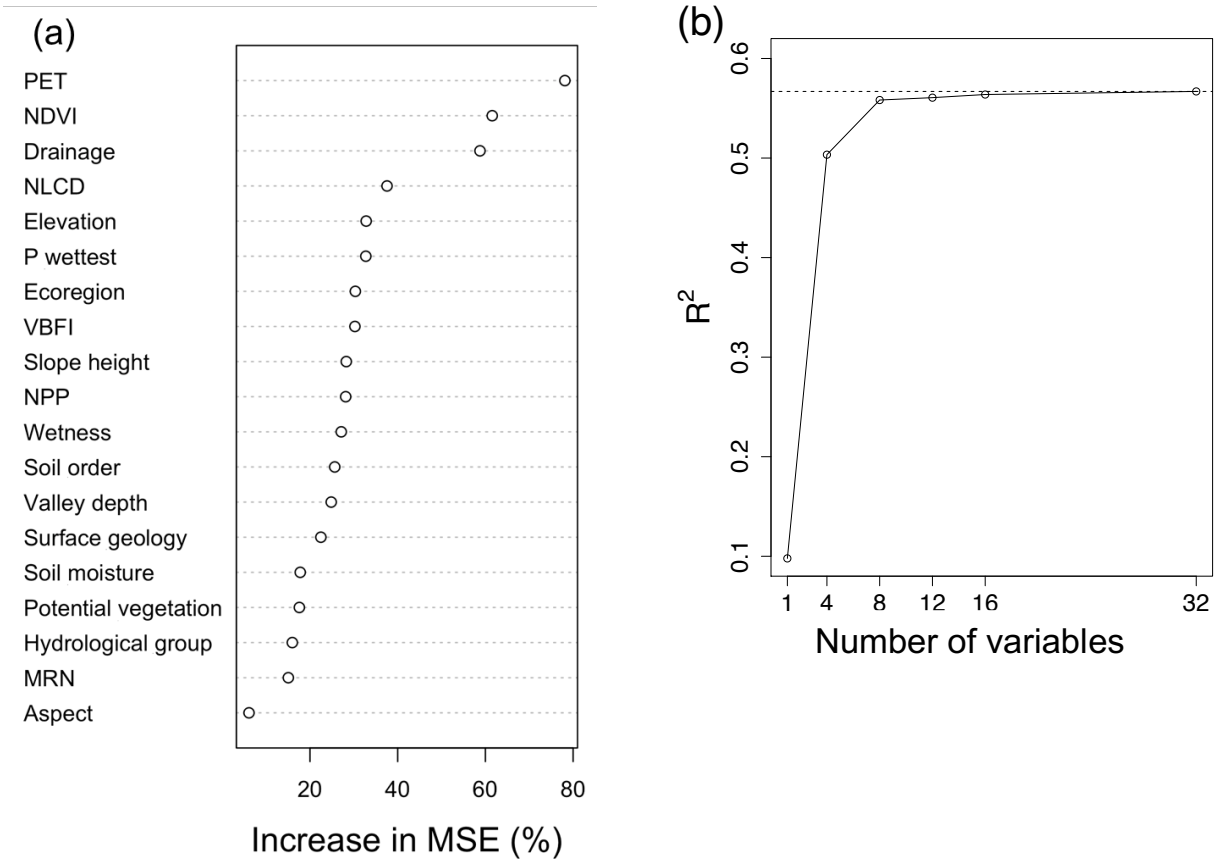

Figure 2. (a) Variable importance after removing correlated variables. (b) Changes in the model accuracy in terms of the number of the input variables of the Random Forest model. MSE is mean squared error, PET is potential evapotranspiration, NDVI is normalized difference vegetation index, NLCD is national land cover database, P wettest is precipitation of the wettest season, VBFI is valley bottom flatness index, NPP is net primary productivity, MRN is melton ruggedness number.

3.2 Nonlinear controls of environmental factors on SOC stocks

Using the 12 most important environmental predictors identified by RF as an input feature set, we trained the GAM approach to fit the log-transformed SOC stocks (Fig. 3). The constant term in the GAM approach was C = 3.98. R^2 and RMSE were 0.52 and 0.69, respectively. The prediction accuracy of GAM was slightly lower than for the RF approach (R^2 = 0.56, RMSE = 0.66). While

RF considers high-order nonlinear interactions between the environmental predictors, in our GAM approach, SOC is modeled by a linear combination of nonlinear functions of each environmental predictor, not considering interactions between them, which may have resulted in a slightly lower prediction accuracy. Figure 3 shows the GAM-inferred relationships between environmental factors and log-transformed SOC stocks with respect to the 12 most important variables. Potential evapotranspiration, normalized difference vegetation index, and soil drainage condition are the three most important variables from RF (Figure 2a).

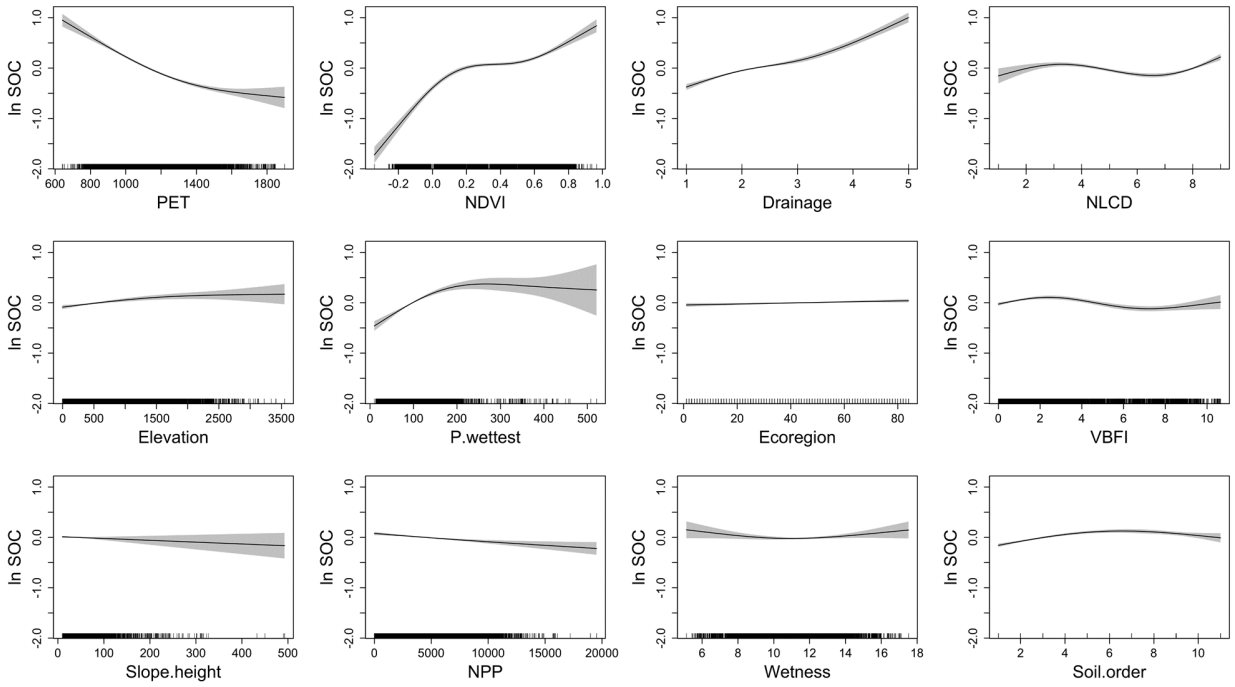

Figure 3. Variable-wise prediction of ln SOC by the Generalized Additive Model. The shade around the solid line indicates 95% confidence interval. The minor ticks on the horizontal axis denote the values of data. ln SOC is observed SOC (Ln Kg m^{-2}), PET is potential evapotranspiration, NDVI is normalized difference vegetation index, NLCD is national land cover database, P wettest is precipitation of the wettest season, VBFI is valley bottom flatness index, and NPP is net primary productivity.

The functional relationships between SOC stocks and environmental predictors were produced as splines by GAM. We next developed explicit analytical expressions by fitting the splines obtained from GAM. Figure 3 shows that the changes of SOC stocks with respect to many of the environmental variables (n = 6) are essentially negligible after considering the uncertainty. Hence, we identified only the following 6 important environmental variables: potential evapotranspiration, normalized difference vegetation index, soil drainage condition, precipitation of the wettest season, elevation, and net primary productivity;

- Potential evapotranspiration (PET):

$$\begin{cases} z = \dfrac{PET - 641}{1000}, \\ Y_{PET} = exp(0.44 - 1.24z - 1.51z^2 + 0.05z^3) - 0.6. \end{cases}$$

- Normalized difference vegetation index (NDVI):

$$Y_{NDVI} = \begin{cases} 0.078 + 1.87(NDVI16 - 0.4)^{1.62} if NDVI16 > 0.4, \\ 0.078 - 4.36|NDVI16 - 0.4|^{2.44} Otherwise. \end{cases}$$

- Soil drainage:

Soil drainage	1	2	3	4	5
$Y_{Soildrainage}$	-0.38	-0.05	0.15	0.50	1.00

- Elevation:

$$\begin{cases} z = \dfrac{Elevation}{1000} \\ Y_{Elevation} = 0.17 - exp\{-1.34 - 0.75z(1 + 0.1z^2)\} \end{cases}$$

- Precipitation:

$$\begin{cases} z = \dfrac{Precipitation}{250} \\ Y_{Precipitation} = 0.38 - exp\{-0.15 - 3.24z^{1.5}\} \end{cases}$$

- Net primary productivity (NPP):

$$Y_{NPP} = 0.077 - 1.68 \times 10^{-5} NPP$$

The fitted curves accurately represented the splines from GAM (Figure 4). The log-transformed SOC stocks from the GAM approach were computed using the following equation,

$$lnSOC = Y_{PET} + Y_{NDVI} + Y_{Soildrainage} + Y_{Elevation} + Y_{Precipitation} + Y_{NPP} + 3.98$$

Here, $lnSOC$ is log transformed SOC stocks, Y_{PET} is the functional relation of PET with SOC stocks, Y_{NDVI} is the functional relation of NDVI with SOC stocks, $Y_{Soildrainage}$ is the functional relation of soil drainage with SOC stocks, $Y_{Elevation}$ is the functional relation of elevation with

SOC stocks, $Y_{Precipitation}$ is the functional relation of precipitation with SOC stocks, and Y_{NPP} is the functional relation of net primary productivity with SOC stocks.

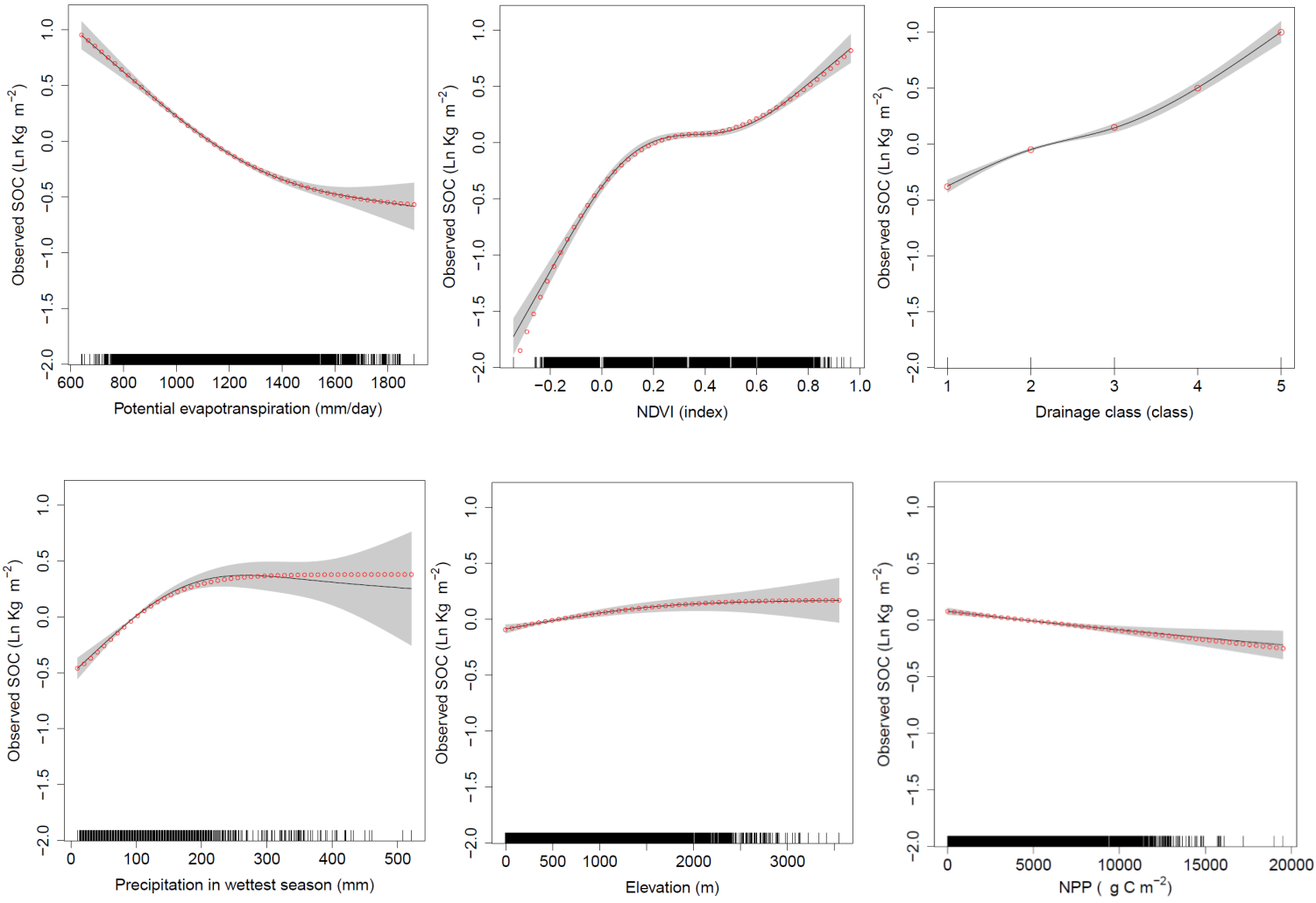

Figure 4. Curve fittings of the splines from the Generalized Additive Model. The solid lines are the expectation values from the Generalized Additive Model, and the circles are computed from the fitting curves. The shade around the solid line indicates 95% confidence interval.

Our results show that the analytical model we developed using only 6 environmental predictors (Fig. 5) showed similar prediction accuracy as that obtained from the GAM with 12 variables. Using only the first three environmental predictors (potential evapotranspiration, normalized difference vegetation index, and soil drainage condition) together with the constant term (3.98), the analytical model achieved an R2 of 0.48, indicating relatively marginal importance of the remaining three environmental factors (elevation, precipitation, and net primary productivity). Figure 5a and 5b show the comparison between the GAM model with all the 12 environmental variables and the analytical model with 6 environmental variables in predicting SOC stocks.

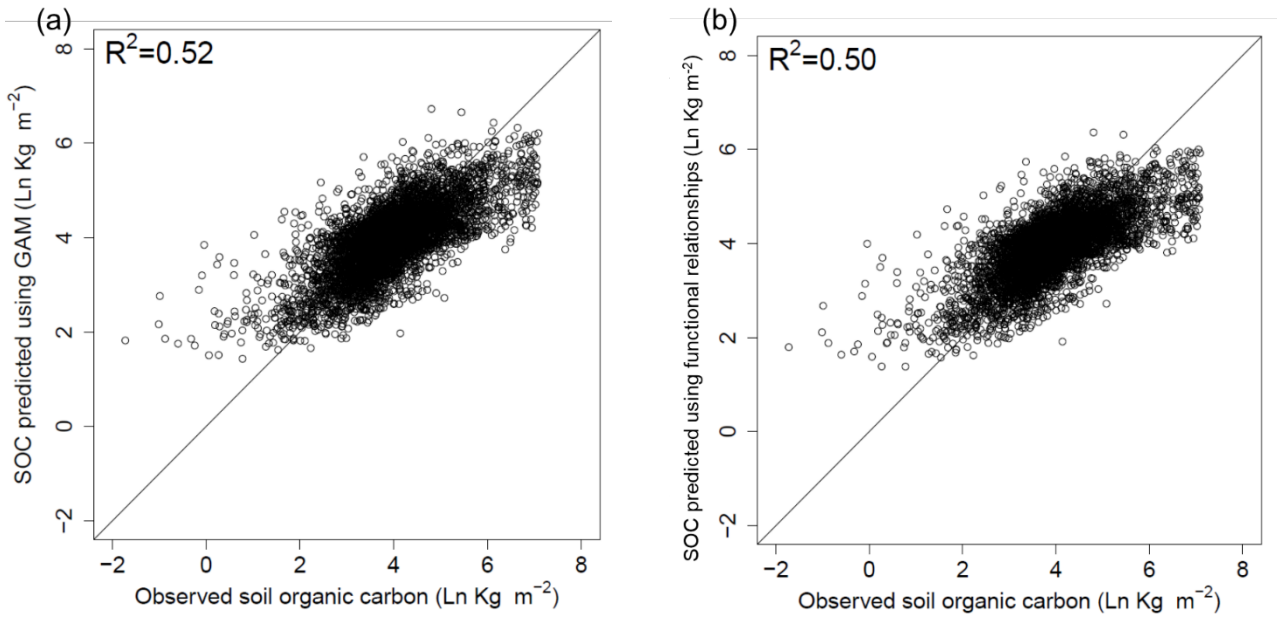

Figure 5. Comparison of the model predictions between (a) GAM (Generalized Additive Model) with 12 variables and (b) analytical model with 6 variables.

We developed an approach to derive analytical expressions for observationally-derived environmental controls on SOC stocks. In this approach, we first identified the dominant environmental predictors of continental U.S. surface SOC stocks using a RF approach. We then derived mathematical equations which captured environmental controls on SOC stocks using a GAM approach. The mathematical relations we derived produced comparable prediction accuracy consistent with the RF approach, using only a subset of environmental predictors used in the RF approach. Our study demonstrates a novel use of ML to improve understanding of nonlinear controls of environmental factors on SOC stocks. Our approach of deriving analytical relationships between environmental factors and SOC stocks can be used to evaluate ESM representations of environmental controls on SOC stocks. However, we note that our study quantified these relationships at a much finer resolution (100 m) than typically used in ESM land models for global simulations (~10-100 km). Therefore, a first step in evaluating ESM land model SOC predictions using these derived analytical relationships would be to run the models at fine resolution using appropriate forcing, initial conditions, and site characteristics. Such an analysis could point to deficiencies in the models' mechanistic representations so that evaluation at ESM resolutions could focus on spatial scaling methods.

Our analysis identified 6 environmental factors (potential evapotranspiration, vegetation index, soil drainage condition, precipitation, elevation, and net primary productivity) as dominant predictors of continental US surface SOC stocks among the 31 environmental predictors we evaluated. Out of these 6 environmental factors, potential evapotranspiration, soil drainage condition, and normalized difference vegetation index were the most important environmental predictors of SOC stocks. Elevation and net primary productivity showed marginal importance in

predicting continental US surface SOC stocks, although these are key environmental controls in current ESM land models.

Various earlier studies also used different combination of these environmental factors to predict SOC stocks at different scales and in different environmental conditions (Gonçalves et al., 2021; Lamichhane et al., 2019; Minasny et al., 2013; Mishra et al., 2020; Mishra et al., 2021). Garten et al. (2009) reported that the control of soil moisture on bulk SOC and its fractions were greater than the controls of elevated CO2 and temperature individually at a field scale. Consistent with this finding, the dominant controls of potential evapotranspiration, soil drainage condition, and precipitation demonstrate the control of soil moisture on SOC stocks across the continental U.S. Our results show that SOC stocks decreased exponentially with increases in potential evapotranspiration. In our dataset, higher potential evapotranspiration values are in the Southern U.S. (Fig. S1), which has higher air temperatures and solar radiation in comparison to other parts of the U.S. Higher air temperatures and longer duration of solar radiation cause drier soil conditions, promoting SOC mineralization and lower total SOC stocks (Das et al., 2019; Hungate et al., 2002; Sherrod et al., 2005).

Our results show lower SOC stocks in excessively drained soils (number 1) and higher SOC stocks in poorly drained soils (number 5) across the continental US. Excessively drained soils are generally coarse-textured soils with high saturated hydraulic conductivity. Similarly, poorly drained soils are often fine-textured soils with more of their pore space filled with water for longer periods of time. Our results are consistent with findings of earlier studies, which showed mean soil carbon concentration significantly differed across different soil drainage classes (Raymond et al., 2013; Wickland et al., 2010). Poorly and very poorly drained soils have lower soil respiration rates (Davis et al., 2010; Webster et al., 2008) compared to well-drained soils (Davidson et al., 1998; Savage & Davidson, 2001), resulting in higher SOC preservation. Some studies suggest precipitation has a strong positive correlation with SOC (Alvarez & Lavado, 1998; Burke et al., 1989; Evans et al., 2011), while other studies show precipitation has little to no influence on SOC (Doetterl et al., 2015; Percival et al., 2000). Our results show increased SOC stocks with increases in precipitation up to 200 mm y-1. Beyond 200 mm y-1, the impact of precipitation on SOC stocks was small. Considering precipitation as a proxy for soil moisture content, control of precipitation on SOC stocks is higher in drier areas of the continental US than in areas with higher precipitation.

Our results indicate that with increased vegetation index, continental U.S. surface SOC stocks increased nonlinearly. We found large increases in SOC stocks as annual average NDVI values increased from -0.2 to 0.2, but the relationship between SOC and NDVI flattened at higher NDVI values (>0.2 to 1). This relationship could be due to nonlinear relationships between cholorphyll concentration of green biomass and the calculated NDVI values (Yoder & Waring, 1994). Vegetation properties have been documented as strong predictors of SOC stocks (Guo et al., 2016; Jobbágy & Jackson, 2000; Li et al., 2010) and widely used in statistical and process-based models to predict SOC stocks (Gautam et al., 2020; Mishra et al., 2021).

3.3 Dominant environmental controllers and their relationships with SOC stocks in Earth system models

Out of the 50 global environmental factors we evaluated, only 14 were dominant predictors of global SOC stocks. These 14 environmental factors explained 60% of variability in SOC stocks in observations. In contrast, CMIP6 ESMs used only 8 environmental factors and explained >95% variability in ESM SOC representations. The variable importance of these environmental factors is presented in Fig. 6.

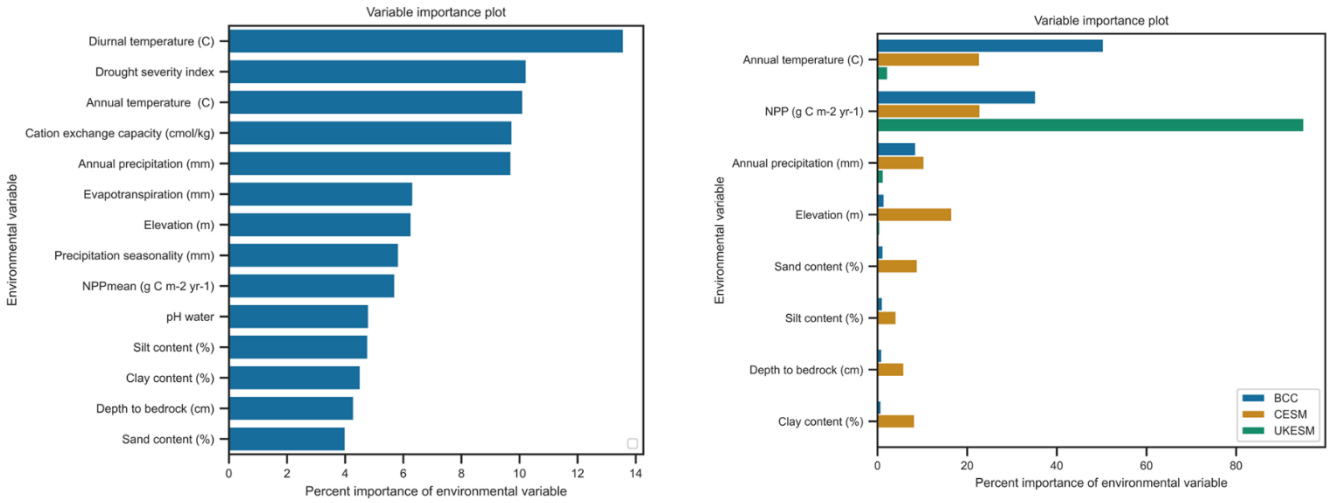

Figure 6: Divergent environmental controls of global SOC stocks in observations (left) and in Earth system models (right).

The RF models in CMIP6 ESM datasets produced near prefect predictions of ESM SOC stocks (average $R^2 = 0.95$) using only three environmental factors (precipitation, net primary productivity and temperature). The R^2 values obtained by RF for UKESM, CESM, and BCC model were 0.99,0.89, and 0.98 respectively.

We evaluated the control of temperature, precipitation and net primary productivity on SOC stocks which were found dominant controllers both in observations and ESMs. The nonlinear relationships between environmental factors and SOC stocks are not consistent between field observations and ESM representations (Fig. 7). In ESMs, the control of temperature on SOC stocks were consistent with observations within the range of 0 to 20°C, but the control of temperature on SOC stocks was not consistent with observations either at higher or lower temperatures (< 0°C and >20°C). Our results showed different relationships describing the control of precipitation on SOC stocks in observations and ESMs. In observations, SOC stocks are found increasing with the increase in precipitation. But, in ESMs, the control of precipitation on SOC stocks has been represented differently among ESMs, and with observations. Among ESMs, CESM shows relatively close functional relationship with observations, but other two ESMs shows different control of precipitation on global SOC stocks. We found that the ESM SOC stocks showed significantly higher sensitivity to NPP than what was observed in field observations. Field

observations of SOC stocks showed the saturating relationship with the NPP, but ESMs represent exponential increase of SOC stocks with increase in NPP.

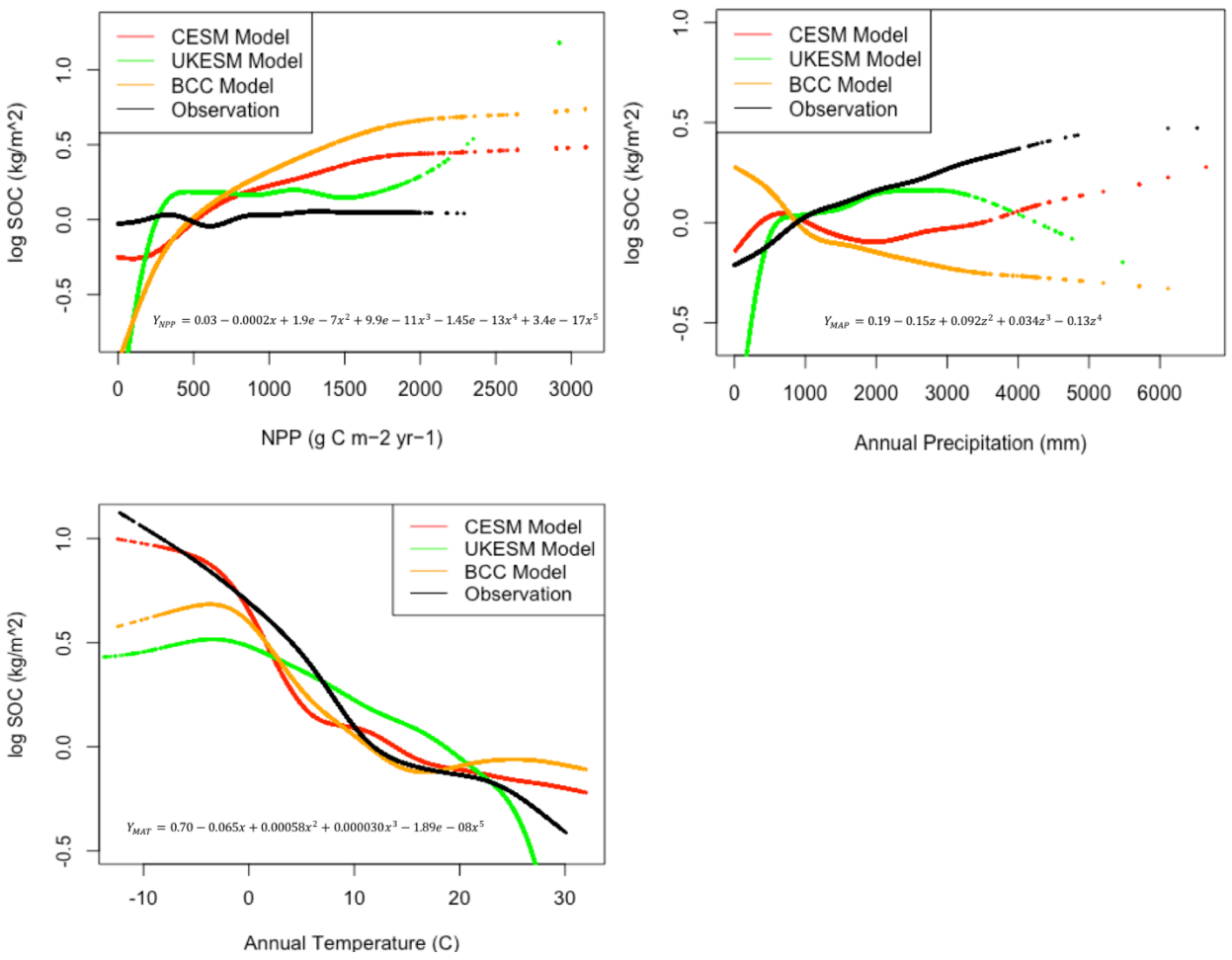

Figure 7: Relationships between net primary productivity (NPP), annual precipitation and annual temperature with observed SOC stocks (black line) and ESM representations (colored lines).

4. OUTOMES AND IMPACTS

Our research activities enhanced the scientific understanding about the environmental controllers of SOC stocks and their representation in existing CMIP6 ESMs. The key outcome of our efforts is an approach to improve existing modeling representation of environmental controls of SOC stocks in the ESMs. Specifically, our findings documented the mathematical algorithms that described the environmental controls of SOC stocks both in models and observations. During this project, we produced data products, publications, and scientific insights of interest to experimentalists, modelers, land managers, and policymaking and implementing agencies.

The capabilities gained through this effort will better position Sandia to integrate measurement and remote sensing data with mechanistic, process-based, and geospatial modeling in order to evaluate and forecast belowground responses to a wide range of environmental perturbations – including climatic change; land use/land cover changes; soil contamination by metals, radionuclides, or organics; or other long-term conditions – that may disrupt the soil properties or processes that are essential to sustaining life on Earth. Thus, there is potential for developing long-term support for this kind of studies from a variety of sponsors. Successful integration of Dr. Mishra's capabilities with multi-disciplinary team will place Sandia, over the long term, in a strategic position to respond to new opportunities in support of DOE mission areas of environmental quality and energy security, e.g., through research as diverse as (1) predicting the consequences of warming high-latitude systems (permafrost regions) on future atmospheric concentrations of greenhouse gases and subsequent feedbacks to climatic change, (2) evaluating the impacts of environments contaminated with metals, radionuclides, and organics and identifying methods for their remediation, and (3) devising systems for sustained production of biofuel feedstocks.

5. CONCLUSIONS

Appropriate representation of environmental controllers on SOC stocks in Earth system land models is required to project realistic rates of change in SOC in response to land use and climate changes, and to understand feedbacks between the land and atmosphere. The non-linear expressions we derived quantify controls of individual environmental factors on SOC stocks in the presence of other environmental factors. Therefore, these observationally derived analytical expressions can be used to benchmark land model representations of environmental factors on SOC stocks. Our analysis showed potential evapotranspiration, normalized difference vegetation index, soil drainage condition, precipitation, elevation, and net primary productivity as important environmental controllers of continental US surface SOC stocks. Out of these six environmental factors, potential evapotranspiration, normalized difference vegetation index, and soil drainage condition explained about 50% of the variability in observed SOC stocks (while the other three environmental variables explained another 6% of the variability). Our derived analytical expressions produced comparable prediction accuracy as the Generalized Additive Modeling and Random Forest approach using only a subset of environmental factors.

We observed different environmental factors as dominant controllers of SOC stocks in observations and ESM representations. The environmental factors which were present both in observations and ESMs, showed different functional relationships between environmental factors and SOC stocks in observations and ESM representations. Future studies should investigate the magnitude of uncertainty in SOC stocks that can be reduced by including additional environmental factors in ESMs consistent with field observations.

REFERENCES

Adhikari, K., & Hartemink, A. E. (2016). Linking soils to ecosystem services—A global review. *Geoderma, 262*, 101-111.

Adhikari, K., Mishra, U., Owens, P., Libohova, Z., Wills, S., Riley, W. J., . . . Smith, D. (2020). Importance and strength of environmental controllers of soil organic carbon changes with scale. *Geoderma, 375*, 114472.

Alvarez, R., & Lavado, R. S. (1998). Climate, organic matter and clay content relationships in the Pampa and Chaco soils, Argentina. *Geoderma, 83*(1-2), 127-141.

Arora, V. K., Katavouta, A., Williams, R. G., Jones, C. D., Brovkin, V., Friedlingstein, P., . . . Ziehn, T. (2020). Carbon–concentration and carbon–climate feedbacks in CMIP6 models and their comparison to CMIP5 models. *Biogeosciences, 17*(16), 4173-4222. doi:10.5194/bg-17-4173-2020

Batjes, N. H. (2016). Harmonized soil property values for broad-scale modelling (WISE30sec) with estimates of global soil carbon stocks. *Geoderma, 269*, 61-68.

Breiman, L. (2001). Random forests. *Machine learning, 45*(1), 5-32.

Burke, I. C., Yonker, C. M., Parton, W. J., Cole, C. V., Flach, K., & Schimel, D. S. (1989). Texture, climate, and cultivation effects on soil organic matter content in US grassland soils. *Soil Science Society of America Journal, 53*(3), 800-805.

Burt, R. (2004). Soil survey laboratory methods manual.

Carvalhais, N., Forkel, M., Khomik, M., Bellarby, J., Jung, M., Migliavacca, M., . . . Weber, U. (2014). Global covariation of carbon turnover times with climate in terrestrial ecosystems. *Nature, 514*(7521), 213-217.

Collier, Nathan, et al. (2018). The international land model benchmarking (ILAMB) system: design, theory, and implementation. *Journal of Advances in Modeling Earth Systems 10(11)*, 2731-2754.

Das, S., Richards, B. K., Hanley, K. L., Krounbi, L., Walter, M., Walter, M. T., . . . Lehmann, J. (2019). Lower mineralizability of soil carbon with higher legacy soil moisture. *Soil Biology and Biochemistry, 130*, 94-104.

Davidson, E. A., Belk, E., & Boone, R. D. (1998). Soil water content and temperature as independent or confounded factors controlling soil respiration in a temperate mixed hardwood forest. *Global change biology, 4*(2), 217-227.

Davis, A. A., Compton, J. E., & Stolt, M. H. (2010). Soil respiration and ecosystem carbon stocks in New England forests with varying soil drainage. *Northeastern Naturalist, 17*(3), 437-454.

Doetterl, S., Stevens, A., Six, J., Merckx, R., Van Oost, K., Pinto, M. C., . . . Venegas, E. Z. (2015). Soil carbon storage controlled by interactions between geochemistry and climate. *Nature Geoscience, 8*(10), 780-783.

Edmonds, J. (1971). Matroids and the greedy algorithm. *Mathematical programming, 1*(1), 127-136.

Evans, S. E., Burke, I. C., & Lauenroth, W. K. (2011). Controls on soil organic carbon and nitrogen in Inner Mongolia, China: A cross-continental comparison of temperate grasslands. *Global Biogeochemical Cycles, 25*(3).

Friedlingstein, P., Meinshausen, M., Arora, V. K., Jones, C. D., Anav, A., Liddicoat, S. K., & Knutti, R. (2014). Uncertainties in CMIP5 climate projections due to carbon cycle feedbacks. *Journal of Climate, 27*(2), 511-526.

Garten, C. T., Classen, A. T., & Norby, R. J. (2009). Soil moisture surpasses elevated CO2 and temperature as a control on soil carbon dynamics in a multi-factor climate change experiment. *Plant and Soil, 319*(1), 85-94. doi:10.1007/s11104-008-9851-6

Gautam, S., Mishra, U., Scown, C. D., & Zhang, Y. (2020). Sorghum biomass production in the continental United States and its potential impacts on soil organic carbon and nitrous oxide emissions. *GCB Bioenergy, 12*(10), 878-890.

Georgiou, K., Malhotra, A., Wieder, W. R., Ennis, J. H., Hartman, M. D., Sulman, B. N., ... & Jackson, R. B. (2021). Divergent controls of soil organic carbon between observations and process-based models. *Biogeochemistry, 156(1),* 5-17.

Gonçalves, D. R. P., Mishra, U., Wills, S., & Gautam, S. (2021). Regional environmental controllers influence continental scale soil carbon stocks and future carbon dynamics. *Scientific reports, 11*(1), 1-10.

Grossman, R., & Reinsch, T. (2002). 2.1 Bulk density and linear extensibility. *Methods of soil analysis: Part 4 physical methods, 5*, 201-228.

Guo, X., Meng, M., Zhang, J., & Chen, H. Y. (2016). Vegetation change impacts on soil organic carbon chemical composition in subtropical forests. *Scientific reports, 6*(1), 1-9.

Hastie, T., & Tibshirani, R. (1990). Exploring the nature of covariate effects in the proportional hazards model. *Biometrics*, 1005-1016.

Hastie, T., Tibshirani, R., & Friedman, J. (2001). The elements of statistical learning. Springer series in statistics. In*:* : Springer.

Hengl, T., de Jesus, J. M., MacMillan, R. A., Batjes, N. H., Heuvelink, G. B., Ribeiro, E., . . . Walsh, M. G. (2014). SoilGrids1km—global soil information based on automated mapping. *PloS one, 9*(8), e105992.

Heuvelink, G. B., Angelini, M. E., Poggio, L., Bai, Z., Batjes, N. H., van den Bosch, R., . . . Olmedo, G. F. (2021). Machine learning in space and time for modelling soil organic carbon change. *European Journal of Soil Science, 72*(4), 1607-1623.

Hungate, B. A., Reichstein, M., Dijkstra, P., Johnson, D., Hymus, G., Tenhunen, J., . . . Drake, B. (2002). Evapotranspiration and soil water content in a scrub-oak woodland under carbon dioxide enrichment. *Global change biology, 8*(3), 289-298.

Jenny, H. (1941). Factors of soil formation. 281 pp. *New York, 801*.

Jobbágy, E. G., & Jackson, R. B. (2000). The vertical distribution of soil organic carbon and its relation to climate and vegetation. *Ecological applications, 10*(2), 423-436.

Lal, R. (2013). Soils and ecosystem services. In *Ecosystem services and carbon sequestration in the biosphere* (pp. 11-38): Springer.

Lamichhane, S., Kumar, L., & Wilson, B. (2019). Digital soil mapping algorithms and covariates for soil organic carbon mapping and their implications: A review. *Geoderma, 352*, 395-413.

Lauer, Axel, et al. (2017). Benchmarking CMIP5 models with a subset of ESA CCI Phase 2 data using the ESMValTool. *Remote Sensing of Environment, 203,* 9-39.

Li, P., Wang, Q., Endo, T., Zhao, X., & Kakubari, Y. (2010). Soil organic carbon stock is closely related to aboveground vegetation properties in cold-temperate mountainous forests. *Geoderma, 154*(3-4), 407-415.

Liaw, A., & Wiener, M. (2002). Classification and regression by randomForest. *R news, 2*(3), 18-22.

Luo, Y. Q., et al. (2012). A framework for benchmarking land models. *Biogeosciences 9(10):* 3857-3874.

McBratney, A. B., Santos, M. M., & Minasny, B. (2003). On digital soil mapping. *Geoderma, 117*(1-2), 3-52.

Minasny, B., McBratney, A. B., Malone, B. P., & Wheeler, I. (2013). Digital mapping of soil carbon. *Advances in agronomy, 118*, 1-47.

Mishra, U., Gautam, S., Riley, W., & Hoffman, F. M. (2020). Ensemble machine learning approach improves predicted spatial variation of surface soil organic carbon stocks in data-limited northern circumpolar region. *Frontiers in big Data, 3*, 40.

Mishra, U., Hugelius, G., Shelef, E., Yang, Y., Strauss, J., Lupachev, A., . . . Riley, W. J. (2021). Spatial heterogeneity and environmental predictors of permafrost region soil organic carbon stocks. *Science advances, 7*(9), eaaz5236.

Mishra, U., & Riley, W. (2015). Scaling impacts on environmental controls and spatial heterogeneity of soil organic carbon stocks. *Biogeosciences, 12*(13), 3993-4004.

Olaya-Abril, A., Parras-Alcántara, L., Lozano-García, B., & Obregón-Romero, R. (2017). Soil organic carbon distribution in Mediterranean areas under a climate change scenario via multiple linear regression analysis. *Science of the Total Environment, 592*, 134-143.

Ottoy, S., De Vos, B., Sindayihebura, A., Hermy, M., & Van Orshoven, J. (2017). Assessing soil organic carbon stocks under current and potential forest cover using digital soil mapping and spatial generalisation. *Ecological indicators, 77*, 139-150.

Padarian, J., Minasny, B., & McBratney, A. B. (2020). Machine learning and soil sciences: A review aided by machine learning tools. *Soil, 6*(1), 35-52.

Percival, H. J., Parfitt, R. L., & Scott, N. A. (2000). Factors controlling soil carbon levels in New Zealand grasslands is clay content important? *Soil Science Society of America Journal, 64*(5), 1623-1630.

Raymond, J. E., Fernandez, I. J., Ohno, T., & Simon, K. (2013). Soil drainage class influences on soil carbon in a New England forested watershed. *Soil Science Society of America Journal, 77*(1), 307-317.

Riggers, C., Poeplau, C., Don, A., Bamminger, C., Höper, H., & Dechow, R. (2019). Multi-model ensemble improved the prediction of trends in soil organic carbon stocks in German croplands. *Geoderma, 345*, 17-30.

Savage, K., & Davidson, E. (2001). Interannual variation of soil respiration in two New England forests. *Global Biogeochemical Cycles, 15*(2), 337-350.

Sequeira, C. H., Wills, S. A., Seybold, C. A., & West, L. T. (2014). Predicting soil bulk density for incomplete databases. *Geoderma, 213*, 64-73.

Sherrod, L. A., Peterson, G. A., Westfall, D. G., & Ahuja, L. R. (2005). Soil organic carbon pools after 12 years in no-till dryland agroecosystems.

Siewert, M. B. (2018). High-resolution digital mapping of soil organic carbon in permafrost terrain using machine learning: a case study in a sub-Arctic peatland environment. *Biogeosciences, 15*(6), 1663-1682.

Soil Survey Staff & Loecke, T. (2016). Rapid Carbon Assessment: Methodology, Sampling, and Summary. United States Department of Agriculture, Natural Resources Conservation Service.

Sulman, B. N., Harden, J., He, Y., Treat, C., Koven, C., Mishra, U., . . . Nave, L. E. (2020). Land use and land cover affect the depth distribution of soil carbon: Insights from a large database of soil profiles. *Frontiers in Environmental Science*, 146.

Todd-Brown, K., Randerson, J., Post, W., Hoffman, F., Tarnocai, C., Schuur, E., & Allison, S. (2013). Causes of variation in soil carbon simulations from CMIP5 Earth system models and comparison with observations. *Biogeosciences, 10*(3), 1717-1736.

Vašát, R., Kodešová, R., & Borůvka, L. (2017). Ensemble predictive model for more accurate soil organic carbon spectroscopic estimation. *Computers & Geosciences, 104*, 75-83.

Vitharana, A., Zhu, X., Du, J., Oberheide, J., & Ward, W. E. (2019). Statistical modeling of tidal weather in the mesosphere and lower thermosphere. *Journal of Geophysical Research: Atmospheres, 124*(16), 9011-9027.

Vos, C., Jaconi, A., Jacobs, A., & Don, A. (2018). Hot regions of labile and stable soil organic carbon in Germany–Spatial variability and driving factors. *Soil, 4*(2), 153-167.

Wadoux, A. M.-C., Minasny, B., & McBratney, A. B. (2020). Machine learning for digital soil mapping: applications, challenges and suggested solutions. *Earth-Science Reviews*, 103359.

Webster, K., Creed, I., Bourbonniere, R., & Beall, F. (2008). Controls on the heterogeneity of soil respiration in a tolerant hardwood forest. *Journal of Geophysical Research: Biogeosciences, 113*(G3).

Wickland, K. P., Neff, J. C., & Harden, J. W. (2010). The role of soil drainage class in carbon dioxide exchange and decomposition in boreal black spruce (Picea mariana) forest stands. *Canadian journal of forest research, 40*(11), 2123-2134.

Wood, S. N. (2003). Thin plate regression splines. *Journal of the Royal Statistical Society: Series B (Statistical Methodology), 65*(1), 95-114.

Wood, S. N. (2017). *Generalized additive models: an introduction with R*: CRC press.

Yoder, B. J., & Waring, R. H. (1994). The normalized difference vegetation index of small Douglas-fir canopies with varying chlorophyll concentrations. *Remote Sensing of Environment, 49*(1), 81-91.

Zhang, G.-l., Feng, L., & Song, X.-d. (2017). Recent progress and future prospect of digital soil mapping: A review. *Journal of Integrative Agriculture, 16*(12), 2871-2885.

DISTRIBUTION

Email—Internal

Name	Org.	Sandia Email Address
Benjamin Cook	8910	bkcook@sandia.gov
Susan Altman	8140	sjaltma@sandia.gov
Amanda Barry	1817	anbarry@sandia.gov
John Gladden	8624	jmgladd@sandia.gov
Anthe George	8620	angerog@sandia.gov
Technical Library	1911	sanddocs@sandia.gov

This page left blank

Sandia National Laboratories

Sandia National Laboratories is a multimission laboratory managed and operated by National Technology & Engineering Solutions of Sandia LLC, a wholly owned subsidiary of Honeywell International Inc. for the U.S. Department of Energy's National Nuclear Security Administration under contract DE-NA0003525.