

Multi-decadal time series of remotely sensed vegetation improves prediction of soil carbon in a subtropical grassland

CHRIS H. WILSON,^{1,6} T. TREVOR CAUGHLIN,² SAMI W. RIFAI,² ELIZABETH H. BOUGHTON,³
MICHELLE C. MACK,⁴ AND S. LUKE FLORY⁵

¹*School of Natural Resources and Environment, University of Florida, 103 Black Hall, PO Box 116455, Gainesville, Florida 32611 USA*

²*School of Forest Resources and Conservation, University of Florida, 136 Newins-Ziegler Hall, Gainesville, Florida 32611 USA*

³*Macarthur Agroecology Research Center, 300 Buck Island Ranch Road, Lake Placid, Florida 33852 USA*

⁴*Center for Ecosystem Science and Society, Northern Arizona University, PO Box 5620, Flagstaff, Arizona 86011 USA*

⁵*Agronomy Department, University of Florida, 3127 McCarty Hall B, Gainesville, Florida 32611 USA*

Abstract. Soil carbon sequestration in agroecosystems could play a key role in climate change mitigation but will require accurate predictions of soil organic carbon (SOC) stocks over spatial scales relevant to land management. Spatial variation in underlying drivers of SOC, such as plant productivity and soil mineralogy, complicates these predictions. Recent advances in the availability of remotely sensed data make it practical to generate multidecadal time series of vegetation indices with high spatial resolution and coverage. However, the utility of such data largely is unknown, only having been tested with shorter (e.g., 1–2 yr) data summaries. Across a 2,000 ha subtropical grassland, we found that a long time series (28 yr) of a vegetation index (Enhanced Vegetation Index; EVI) derived from the Landsat 5 satellite significantly enhanced prediction of spatially varying SOC pools, while a short summary (2 yr) was an ineffective predictor. EVI was the best predictor for surface SOC (0–5 cm depth) and total measured SOC stocks (0–15 cm). The optimum models for SOC in the upper soil layer combined EVI records with elevation and calcium concentration, while deeper SOC was more strongly associated with calcium availability. We demonstrate how data from the open access Landsat archive can predict SOC stocks, a key ecosystem metric, and illustrate the rich variety of analytical approaches that can be applied to long time series of remotely sensed greenness. Overall, our results showed that SOC pools were closely coupled to EVI in this ecosystem, demonstrating that maintenance of higher average green leaf area is correlated with higher SOC. The strong associations of vegetation greenness and calcium concentration with SOC suggest that the ability to sequester additional SOC likely will rely on strategic management of pasture vegetation and soil fertility.

Key words: *enhanced vegetation index; Google Earth Engine; Landsat time series; remote sensing; soil carbon sequestration; soil organic carbon; subtropical grasslands.*

INTRODUCTION

Soil carbon sequestration (SCS) is a promising strategy for mitigating global climate change (Lal 2004, 2010, Stockmann et al. 2013), particularly in managed ecosystems that have suffered soil organic carbon (SOC) losses due to agricultural conversion, land degradation, or overgrazing (Conant et al. 2001, Conant and Paustian 2002a, DeGryze et al. 2004, Rees et al. 2005). However, it is difficult to manage for SCS at meaningful scales because predicting and quantifying SOC outside of experimental plots remains a considerable challenge (Eigenbrod et al. 2010, Vasques et al. 2010, O'Rourke et al. 2015). This difficulty is due in part to the natural heterogeneity in underlying drivers of SOC, especially soil and vegetation properties. Moreover, we have limited knowledge of how

these drivers interact at the scale of individual land management units, which is necessary to quantify and incentivize ecosystem services (hereafter “management-relevant spatial scale”; Saby et al. 2008, Power 2010, Swain et al. 2013, O'Rourke et al. 2015). Although recent work has emphasized the potential utility of simple measures of plant traits for improving SOC predictions across diverse ecosystems at national scales (Yang et al. 2008, Manning et al. 2015), such work has not explored whether continuous variations in vegetation properties (e.g., green leaf area) improve SOC predictions across more homogeneous management units (e.g., pastures or forest stands). This gap in our current knowledge is critical because efforts by ecosystem scientists and land managers to promote SCS ultimately rely on development of cost effective and efficient methods for predicting and quantifying SOC at management-relevant scales.

Across the largest spatial scales (i.e., bioregions or continents), climatic factors, especially mean annual temperature (MAT) and mean annual precipitation (MAP), are assumed to dominate SOC dynamics through their

Manuscript received 3 November 2016; revised 10 March 2017; accepted 27 March 2017. Corresponding Editor: Willem J. D. van Leeuwen.

⁶E-mail: chwilson@ufl.edu

combined effects on vegetation productivity, microbial metabolism, and rates of soil weathering (Jenny 1961, Chapin et al. 2012). Landscape-scale variations in soil properties, such as clay content, mineralogy, and pH, also exert control over plant composition, microbial activity, and the potential for stabilization of carbon (Burke et al. 1989, Conant and Paustian 2002b, Cotrufo et al. 2013). These mechanisms are reflected in basic ecosystem process models (e.g., CENTURY; Parton et al. 1993, Bolker et al. 1998, Evans et al. 2011) that are used to predict SOC across regional or continental gradients (Burke et al. 1989, Schimel et al. 1994). For agricultural land managers seeking to enhance SCS as part of a portfolio of ecosystem services (Chapin F.S. 2009), understanding factors leading to variations in SOC stocks at the scale of their working landscapes is vital for designing proper monitoring protocols and to spatially target interventions. At management-relevant scales, ranging from tens to thousands of hectares, climatic variables such as MAT and MAP are essentially constant, so variations in SOC are largely due to the interaction of vegetative (e.g., productivity and composition; Conant et al. 2001), topographic and edaphic (e.g., soil texture, soil moisture, and pH; Kemmitt et al. 2006) factors. In general, predicting SOC within homogenous landscapes should be more difficult than predicting SOC across heterogeneous areas with significant variation in edaphic and climatic factors. Thus, the best predictive models for SOC at management-relevant scales should emerge by combining two types of information: (1) soil properties that predict variations in soil capacity to stabilize carbon and constrain microbial decomposition (Davidson and Janssens 2006) and (2) vegetation data on plant productivity and composition within and across management units (i.e., pastures, forest stands).

Remotely sensed vegetation indices may be an efficient method for providing data to improve SOC predictions at management-relevant scales. Recent advances in the accessibility of pre-processed remote sensing data (e.g., Google Earth Engine; Moore and Hansen 2011) have greatly increased the practicality of extracting long time series of vegetation data for specific sites. These multi-decadal time series of remotely sensed vegetation data include well-validated greenness indices, such as the Normalized Difference Vegetation Index (NDVI) and the Enhanced Vegetation Index (EVI), which reflect the biomass and productivity of vegetation (Hill et al. 2004, Cook and Pau 2013, Gu et al. 2013). In an Australian pastureland, Hill et al. (2004) utilized satellite-derived NDVI to parameterize simple light-use-efficiency models that accurately predicted pasture growth rate ($R^2 \sim 0.7$). However, thus far the use of EVI and NDVI to predict SOC has been limited to simple data summaries over brief time periods of only one or two years (e.g., Yang et al. 2008, Vasques et al. 2010), which is far too short for any plant-driven biophysical process to substantially alter total SOC stocks (Smith 2004). Thus, any predictive power from snapshot summaries must derive from one of two sources: (1) discriminating coarse

variations in vegetation composition such as grassland from forest or (2) autocorrelation originating from a process whereby higher SOC consistently leads to higher greenness. By contrast, scientists and land managers seeking to understand and predict SOC within management units characterized by relatively homogeneous vegetation cover need to know (1) whether variations in vegetative greenness correlate with variations in underlying SOC and (2) whether any correlation between greenness and SOC is stable and consistent across time or if there is significant interannual variability. Analyzing multi-decadal time series of vegetation indices is an ideal way to disentangle these possibilities.

To test the utility of remote sensing derived vegetation indices for predicting SOC, we posed three specific questions: (1) What is the explanatory power of a two-year summary of EVI data relative to a multi-decadal time series for predicting SOC? (2) What is the importance of EVI data for predicting SOC relative to edaphic factors? (3) How does the predictive value of EVI change at different soil depths? We hypothesized that average EVI would be positively correlated with measured SOC pools, and that this association would be stronger when averaged over longer time periods. Moreover, in line with the mounting evidence that SOC storage is controlled by different factors at different depths (Silveira et al. 2014) and because above and belowground plant inputs are concentrated near the surface, we hypothesized that the relative importance of plant productivity (hence EVI) as a driver of SOC variability would be greater at shallower soil depth.

METHODS

Study site

Our study site was a 4,300 ha-commercial cattle operation, Buck Island Ranch (BIR), located in Lake Placid, Florida, USA just north of Lake Okeechobee. Pastureland and native rangeland are dominant land uses in the northern Everglades watershed, provide numerous ecosystem services, including livestock production, wildlife habitat conservation, and maintaining cultural legacy, and represent relatively benign, low-input agricultural management (Swain et al. 2007). Like most commercial cattle ranches in the watershed, BIR maintains both planted (“improved”) pasture, dominated by *Paspalum notatum* Fluegge (Bahia grass), and “semi-native” pasture, consisting of a mixture of native warm-season wet and dry prairie species and some introduced pasture grasses. Improved pasture is utilized far more intensively for livestock grazing than semi-native pasture, supplying several times more usable forage on a per area basis, and accordingly is fertilized (at a low rate) and limed regularly to improve the quantity and quality of forage (Silveira et al. 2011). The soil series at BIR are mostly from Spodosol and Entisol orders, and are uniformly coarse-textured with the clay and silt fraction accounting for 2–3% of total mineral mass and a range of sand grades accounting

for the remainder (Silveira et al. 2014). There is no significant stone content. Here we focused on SOC in improved pastures given their agronomic relevance (i.e., there is ~1 million ha of planted “improved” pasture in Florida; Bohlen et al. 2009), thus our scope of inference is limited to this particular land-use type.

Soil organic carbon sampling

In July 2014, we sampled 57 plots distributed across 2,000 ha of improved pasture at BIR (Fig. 1). Our focus was on SOC sequestration in improved pasture, thus we focused our sampling on upland pasture and specifically avoided wetlands and tree hammocks. Plots were identified with stratified random sampling in GIS software (ESRI 2009, ArcGIS Desktop 9.3.1, Environmental Systems Research Institute, Redlands, California, USA) with a minimum spacing of 150 m. To quantify SOC concentration and soil bulk density, we collected one 0–15 cm soil sample from each plot using a hammer core (AMS, Inc. American Falls, ID, 5.08 cm × 15.24 cm Signature SCS Complete/354.26) and computed SOC stock for 0–15 cm. To separately analyze SOC concentration from 0–5 cm and 5–15 cm depth fractions, while also averaging over fine-scale variability in soil properties, 12 subsamples were collected from a circle of 5 m radius using a narrow gauge soil auger (AMS, Inc. American Falls, ID, Soil Step Probe 83 cm/401.4), and divided into 0–5 cm and 5–15 cm depth

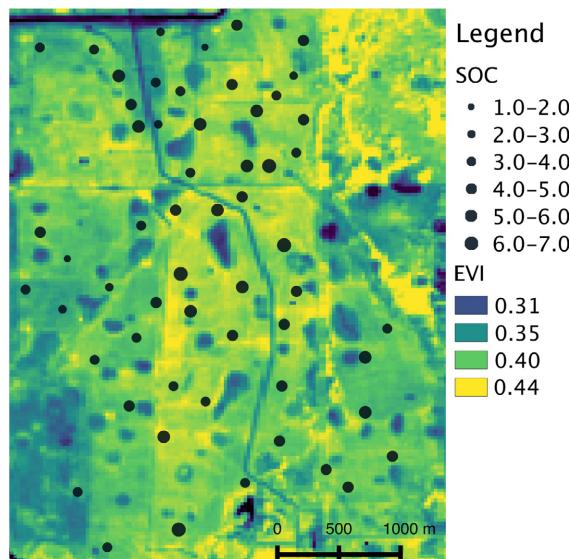


FIG. 1. Map showing spatial layout of soil sampling plots. Dot sizes are proportional to observed size of measured soil organic carbon (SOC) stock (in units of kg C/m² in 0–15 cm depth fraction), ranging from 1.78 kg C/m² to 6.34 kg C/m². Background map contains mean Enhanced Vegetation Index (EVI), with lighter yellow colors indicating higher EVI, and darker blue colors indicating lower EVI (ranging from 0.34 to 0.42). Note that the distinct patches of dark blue represent either wetlands or tree hammocks, both of which we deliberately avoided in our soil sampling where we focused on planted pasture. [Color figure can be viewed at wileyonlinelibrary.com]

fractions. The 12 subsamples for each depth fraction were bulked together in the field.

Soils were oven-dried at 60°C prior to being passed through a 2-mm sieve to remove plant litter. Essentially no stones were found. A subsample from each soil sample was prepared for %C/%N analysis by grinding for 15 min in a Spex SamplePrep Mill, SPEX SamplePrep. Metuchen, NJ, and was then sent to the University of Florida’s (UF) Light Stable Isotopes Lab for analysis on a Carlo Erba NA150.0 CNS Elemental Analyzer (EA), Carlo Erba Strumentazione, Milan, Italy. Separate subsamples were sent to the UF IFAS Analytical Laboratories for measurements of pH, and Mehlich III extractable Ca, K, Mg, and P.

For the 0–15 cm soil sample, we converted the SOC concentration from the EA to a measure of SOC mass by using the bulk density measured from each core. We then standardized the resulting SOC stock estimate to units of kg/m², according to the following formula:

$$\text{SOC Stock} = \text{SOC} \left(\frac{\text{g}}{100 \text{ g}} \right) \times \text{BD} \left(\frac{\text{g}}{\text{cm}^3} \right) \\ \times \left(100^2 \frac{\text{cm}^2}{\text{m}^2} \times 15 \text{ cm} \right) \times \left(\frac{1 \text{ kg}}{1000 \text{ g}} \right)$$

Acquisition of GIS data

To obtain LiDAR surface elevation data, BIR was flown on 13 April 2006, with 36 flight lines covering 72 km². Each flight line was 10.55 km long. Data was collected using an Optech 1233 Airborne Laser Terrain Mapper (Teledyne Optech, Toronto, Canada) mounted on a twin-engine Cessna 337. The LiDAR data were used to produce a DEM (digital elevation map), which we used for topographic measurements (more detail in Appendix S1).

To gain insight into variations in pasture productivity and biomass, for each sampled point ($n = 57$), we extracted time series of EVI, a widely used vegetation index, on a 32-d interval at 30-m pixel resolution from 1 January 1984 to 31 December 2011. Each time series was composed of cloud and shadow masked, atmospherically corrected surface reflectance data processed by the USGS using the Landsat Ecosystem Disturbance Adaptive Processing System (Masek et al. 2006) from Landsat 5’s TM multi-spectral sensor using Google Earth Engine (see Hansen et al. 2013 [discussion in supplementary information]). The open-access Landsat archive has a global extent and images dating back to 1972 (Wulder et al. 2016). Earth Engine is a cloud computing platform designed for rapid processing of large remote sensing data sets through a distributed computing architecture. The EVI data points were calculated as follows (following Huete et al. 2002):

$$\text{EVI} = 2.5 \times \frac{\text{B4} - \text{B3}}{\text{B4} + 6 \times \text{B3} - 7.5 \times \text{B1} + 1}$$

Here, B4 corresponds to the near infrared band, B3 represents the red band, and B1 represents the blue band. EVI

values are therefore dimensionless quantities constrained to be between 0 and 1 in practice. Whereas the better-known NDVI exhibits a non-linear saturating response with increased leaf area index (LAI) above 2–3, the EVI index greatly reduces the saturating response present in NDVI so that more information can be retrieved from high biomass or high LAI vegetation (Huete et al. 2002). For instance, EVI was found to correlate linearly with pasture LAI/biomass up to LAI ~ 5 (the entire tested range) by Houborg and Soegaard (2004).

Statistical analysis

To fully leverage the information contained in the EVI time series, we first decomposed each time series record ($n = 57$) into two components: a linear temporal trend from 1984 to 2012, which represents whether the EVI was trending higher or lower, and a periodic component representing the amplitude of the seasonal phenology, using a simplified form of Fourier analysis based on Bradley et al. (2007; further details in Appendix S2). However, analysis showed that only the linear components of the EVI time series were useful in predicting SOC, and only when the intercept (starting value) and trend were taken in combination, which is a formulation that is mathematically similar to simply taking the mean of the entire time series. In particular, there was no benefit to knowing the particular trend per se; the average value sustained was critical, not whether EVI increased or decreased over time. Thus, for subsequent analyses we only used the simple arithmetic mean of the entire time series record.

To isolate the predictive value of summarizing a long time-series of EVI data compared to using only the most recent two years, we first compared the model fit (using R^2) between these two predictors. Then, since we had previously ruled out temporal trends, seasonal phenology or variance per se as important aspects of our long EVI time series, we investigated whether all dates in the time series contributed equally to predictive power, or if there was significant variation in contribution among dates. To accomplish this task, we performed a simple exploratory analysis where we plotted R^2 from a univariate regression model predicting surface SOC concentrations (0–5 cm; $n = 57$) with mean EVI ($n = 57$) derived by taking the mean for each year in the time series ($n = 28$). To estimate whether high R^2 observations within any given year were likely to arise from chance sampling variation, we performed a simple Monte Carlo analysis to randomly generate 10^5 vectors of $n = 57$ EVI observations within the range of EVI we observed in this study system, fitted univariate least squares regression models between the random vectors and our observed SOC stock, and plotted resulting sampling distribution of R^2 values (see Appendix S3).

Next, we compared the relative importance of EVI to the other environmental covariates in the context of a multivariate regression across all responses. Other environmental predictors included elevation (from

LiDAR-derived DEM), soil pH, and Mehlich III extractable P, K, Ca, and Mg. Note that silt and clay content, as noted above in *Study site*, is uniformly low across our sampled points ($n = 57$), and thus we did not include data on soil mineralogy. To simplify comparison among factors, we combined the information contained in the Mehlich III soil audit by taking the arithmetic mean of the standardized values for each element (Appendix S4), but we also analyzed calcium separately due to its potential role in SOC stabilization processes (i.e., due to cation bridging; Stockmann et al. 2013). Thus, our candidate predictors included mean EVI, elevation, soil pH, Ca, and soil fertility index. Prior to analysis, we standardized all predictor variables by dividing by two standard deviations and centering around the mean so that we could interpret coefficients on the same scale (Gelman 2008). Each coefficient was read as the impact on the response variable of moving from a typical low to a typical high value for that predictor. By contrast, response variables were standardized to have mean zero and unit standard deviation, and therefore each regression coefficient could be interpreted as the effect size of that predictor in units of standard deviations of the response. The purpose of standardizing both predictor and response variables was to enable relative importance comparisons of each predictor both within models (i.e., compared to other predictors) and across models where the original scale of responses could differ. We present the results of these models in the form of coefficient plots, which show the point estimate from least squares regression surrounded by 50% and 95% CIs. Additionally, each coefficient plot also displays the constrained point estimate from LASSO regression, and we provide R^2 model fit summaries from both the unconstrained (least squares) and constrained (LASSO) models, with the latter computed based on out-of-sample predictions that correct for over-fitting.

To perform predictive model selection among these candidate variables, we implemented LASSO (least absolute shrinkage and selection operator) regression (Tibshirani 2011). LASSO constrains the sum of coefficients in a linear regression to be less than or equal to a constant tuning parameter, a form of penalization that reduces estimates of small and/or noisy effects, sometimes all the way to zero (thus eliminating that covariate; Tibshirani 1996). We used cross-validation to set the tuning parameter for our LASSO regression models using the R package “glmnet” (Friedman et al. 2010). We overlaid the LASSO point estimates onto the unpenalized least squares coefficient plots. Next, we validated the predictive contribution of our time series EVI measure by performing LASSO regression on models with and without the EVI predictor and comparing the out-of-sample predictive performance using leave-one-out cross-validation (LOOCV). Specifically, we iteratively refit a LASSO regression model, each time holding out a single data point to predict with the refitted model ($N = 57$), and then computed a LOOCV R^2 metric for

each model. In this way, we validated the predictive value of multi-decadal time series EVI on SOC stock estimation and compare its contribution to quantifying surface SOC vs. SOC contained in lower horizons.

RESULTS

Observed SOC stock (0–15 cm) ranged widely from 1.78 kg C/m² to 6.34 kg C/m², with a mean value of 4.27 kg C/m² (sample CV = 27%; Fig. 1). Surface SOC concentrations (0–5 cm) were even more variable, ranging from 1.04% to 6.34%, while deeper SOC concentrations (5–15 cm) were between 0.63% and 2.55%.

Averaging across the $N = 57$ sampled plots, the EVI data captured fluctuations in seasonal phenology between 0.14 and 0.71 (Fig. 2a). Across the 1984 to 2011 time series, plots varied substantially in both estimated intercepts and slopes, indicating variations in temporal trends (Fig. 2b), and also differed in the amplitude of seasonal EVI (Fig. 2c). However, only the combination of linear terms was statistically significant for predicting SOC, and was equivalent to simply taking the arithmetic mean of the entire time series (Appendix S1). Overall, mean EVI across the 28-yr record varied between 0.35 and 0.44. Measured soil fertility parameters varied widely among the sampled points; for instance, in the 0–15 cm cores pH was between 4.08 and 6.70 and P was between 4.13 and 38.57 ppm.

In univariate regressions with soil carbon as a response variable, R^2 values of models with mean EVI from a short time interval (two most recent years) were uniformly low, with R^2 values of 0.03 (0–15 cm SOC stock), 0.04 (0–5 cm SOC concentration), and 0.04 (5–15 cm SOC concentration). In contrast, for mean EVI derived from the whole 28-yr time series, R^2 varied between 0.24 (0–5 cm SOC concentration) and 0.31 (0–15 cm SOC stock, with $P < 0.0001$ for all models; Fig. 3a, c, e vs. b, d, f).

Relative importance of different dates in the EVI time series varied dramatically (Fig. 4a, b). One year showed an R^2 of 0.35, comparable to the mean of the entire record, while several other years were very low (practically zero), and many were in between (Fig. 4). However, exclusion of the high R^2 year did not significantly impact predictive value of the rest of the time series (i.e., mean EVI was still a strong predictor). Moreover, our multiple observations of years with $R^2 > 0.1$ are extremely unlikely to have resulted by chance variation (i.e., $P \ll 0.00001$, Appendix S3). Comparing EVI (mean of 28-yr period) to our suite of other environmental covariates for predicting soil C revealed differences across responses (i.e., between 0 and 15 cm SOC stock, 0–5 cm SOC concentration and 5–15 cm SOC concentration). For explaining variations in the SOC stock (0–15 cm, Fig. 5a), EVI and calcium availability were both significant positive predictors and appeared to be of

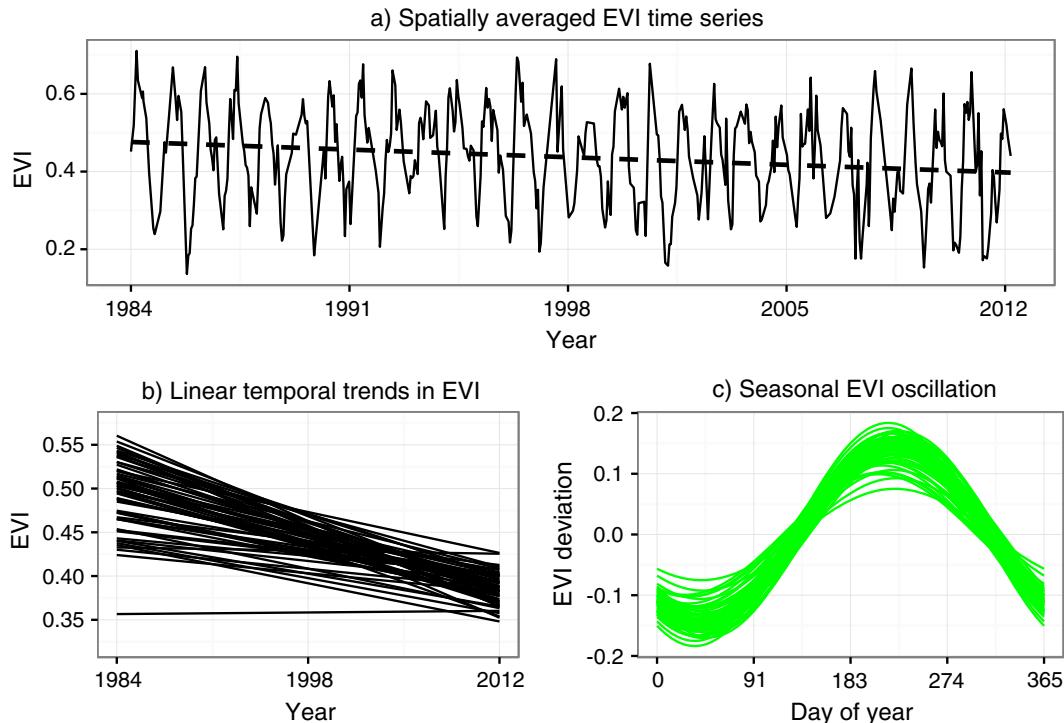


FIG. 2. (a) Spatial average of EVI for all 57 plots across the time series. (b) Linear temporal trends for each plot and (c) estimated seasonal phenology for each point from first order Fourier Series expansion. [Color figure can be viewed at wileyonlinelibrary.com]

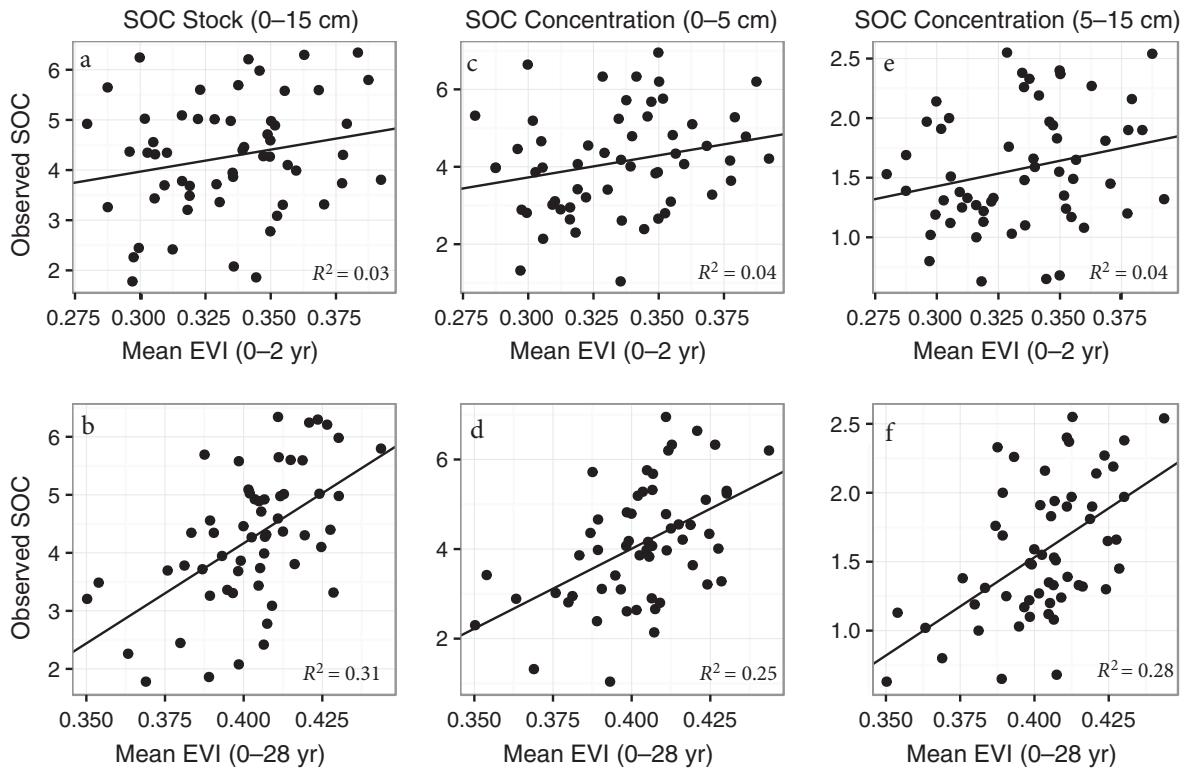


FIG. 3. Assessing univariate regression models in terms of model fit (R^2), for (a, b) SOC stock 0–15 cm, (c, d) SOC concentration 0–5 cm, and (e, f) SOC concentration 5–15 cm, comparing a snapshot approach to EVI (simple mean of the most recent 2 yr), vs. taking mean of the full (28 yr) time series.

similar importance (i.e., coefficient estimates were similar at around 0.9 as were 95% CI coverage), while the other predictors were ambiguous (95% CI crossed zero), and LASSO shrunk their coefficients to zero. Thus, moving from a low to a high value for both EVI and Ca corresponded to a positive effect of about 0.9 standard deviations in SOC stock. For surface SOC concentration (Fig. 5b), EVI, elevation, and calcium, and soil fertility were all retained by LASSO, and EVI was clearly most important. Higher values of EVI, Ca, and fertility positively correlated to greater SOC concentration, while higher elevations corresponded to lower SOC

concentration. By contrast, in the 5–15 cm depth fraction (Fig. 5c), EVI was still statistically significant but had a smaller estimated effect size (0.5) than Ca availability (1.35) and was similar to pH in importance. In that model, LASSO shrunk pH to zero but retained both EVI and Ca.

Comparing cross-validated R^2 among our selected models with and without EVI, we found that inclusion of EVI enhanced the (out-of-sample) model fit for both the SOC stock (0–15 cm), and the surface SOC concentration by 17% and 7% of variance explained, respectively (Fig. 6a–d). These improvements corresponded to relative improvements of 74% and 21% in R^2 , respectively. By comparison, inclusion of EVI alongside pH and Ca in the 5–15 cm model only marginally increased out-of-sample model performance (Fig. 6e, f). Overall, our LASSO models explained between 40% and 46% of variance under leave-one-out cross-validation.

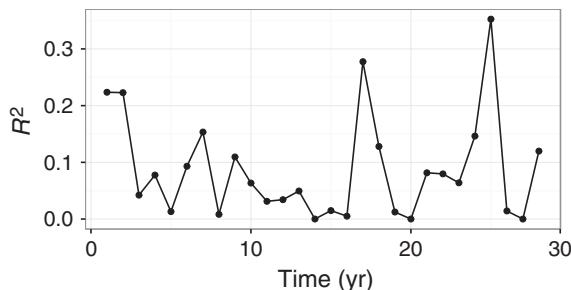


FIG. 4. Exploratory analysis of interannual stochasticity in relationship between mean EVI and SOC measured in 2014. Plot of model fit (R^2) using each year's EVI summary to predict surface SOC concentration.

DISCUSSION

Across grazed subtropical pastures, we found that a long time series of vegetation greenness data significantly enhanced predictions of SOC stocks. EVI was a crucial predictor for total SOC stocks because its addition to our model improved the out-of-sample predictive fit by 74% relative to a model containing only elevation

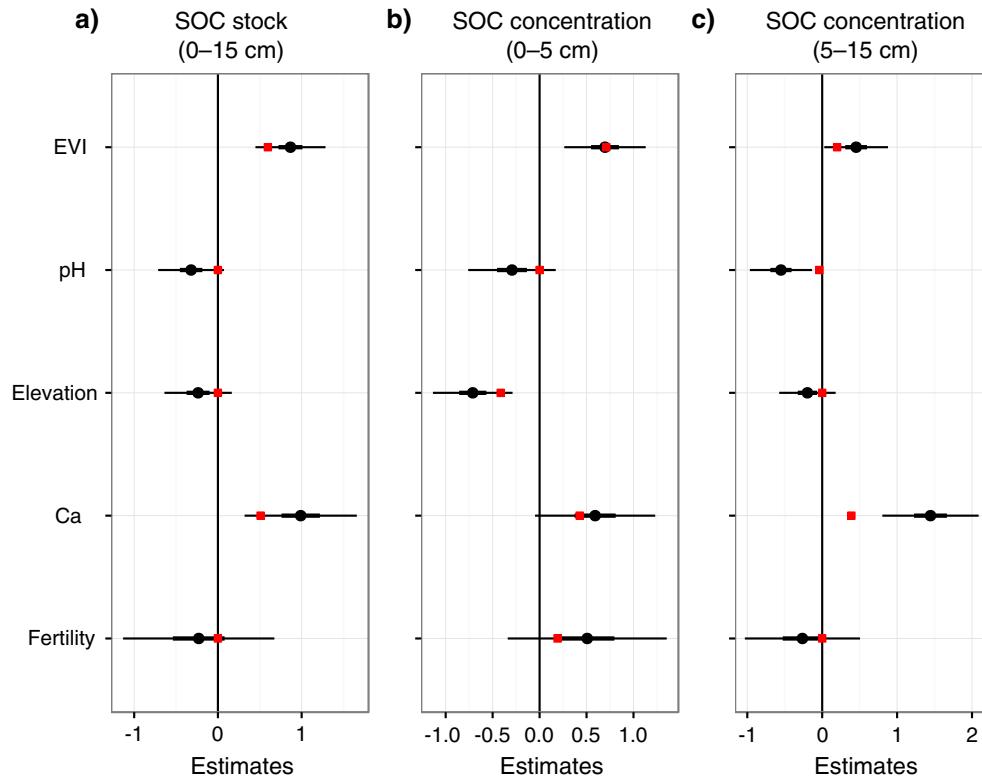


FIG. 5. Coefficient plot from multiple regression models comparing the relative importance of weighted mean EVI (0–28 yr) to our other landscape (elevation), and edaphic (soil pH, P, K, Ca, and Mg) variables for SOC stock 0–15 cm (a), SOC concentration 0–5 cm (b), and SOC concentration 5–15 cm (c). Predictors were all standardized to same scale prior to analysis, and coefficients represent estimates of effect size in units of standard deviation of the response variable. We report point estimates with a black dot surrounded by 50% (solid lines) and 95% (dashed lines) confidence intervals resulting from ordinary least squares regression. Coefficients with 95% CI that do not overlap zero can be considered statistically significant. In addition, we overlay red squares to indicate the point estimates obtained by using the LASSO to shrink the estimates for covariates with limited predictive power. Finally, we report both adjusted R^2 from our multiple regression models estimated via least squares (upper left corner) and out-of-sample R^2 from the LASSO regression, which corrects for over-fitting rampant with unconstrained least squares regression. [Color figure can be viewed at wileyonlinelibrary.com]

and edaphic factors. Moreover, EVI was more important than elevation and edaphic factors in explaining surface SOC concentrations (0–5 cm), and played a greater role in improving the out-of-sample predictive fit for surface SOC compared to deeper SOC (5–15 cm). In contrast, a short time interval of EVI (2-yr mean) failed to deliver any discernable predictive power (contrary to Yang et al. [2008] but consistent with Vasques et al. [2010]). The significant improvements in predictive power arising from the use of a long time-series of remote sensing data demonstrates that across a relatively homogeneous landscape significant spatial variation in SOC stocks can be related to variations in long term vegetation greenness. Finally, when we quantified the relative predictive value of different dates across the entire time series, we found that the coupling of EVI and SOC stocks was highly stochastic across time. In total, our results demonstrate the value of leveraging recent computational advances to acquire and process long time series remote sensing data for the purposes of improving ecosystem SOC predictions.

As we hypothesized, EVI was a stronger predictor of surface SOC than deeper SOC, and a different suite of factors best predicted surface compared to deeper SOC. A stronger coupling between pasture EVI and surface SOC concentrations could arise for at least two, non-exclusive reasons: (1) higher EVI implying a greater rate of plant litter inputs, which concentrate in the upper 5 cm of soil or (2) upper 5 cm SOC/SON stocks being more significant for predicting nutrient mineralization and hence grass production than deeper stocks. Moreover, both processes could be reinforced in a positive feedback cycle. Targeted experimental fieldwork clearly would be necessary to disentangle these causal pathways. For instance, ion-exchange resin membranes (Durán et al. 2013) could be deployed across the landscape to assess whether net mineralization of N and P related to total SOC stocks or pasture production and consequent spectral properties. Despite less predictive power with greater depth, mean EVI was equivalent to calcium as the most important predictor of total measured SOC stocks in the 0–15 cm fraction, suggesting great potential utility in

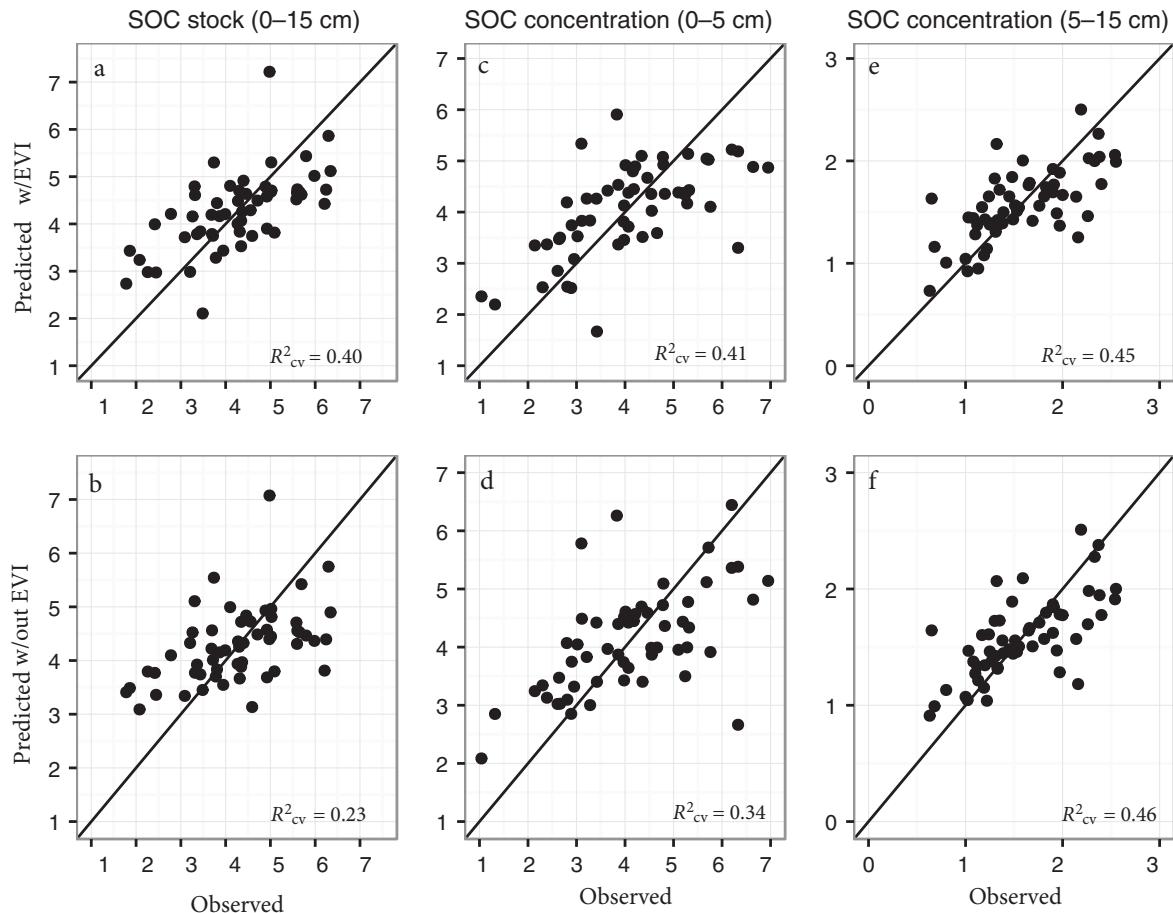


FIG. 6. Out-of-sample predictive model assessment for LASSO models. Specifically, we compared predictive power for models with (w/) and without (w/out) our weighted EVI predictor, for (a, b) 0–15 cm SOC stock, (c, d) 0–5 cm SOC concentration, and (e, f) 5–15 cm SOC concentration. Predictions were generated via leave-one-out cross-validation. We assess the contribution of EVI as a predictor based on improvements in the cross-validated model fit metric (R^2). Conversely, over-fitting would be indicated by a decline in cross-validated R^2 .

estimating variations in SOC across management relevant scales (tens to thousands of hectares).

Aside from EVI, the best predictors for surface SOC pools were elevation and calcium, while for deeper SOC pools, calcium availability appeared to be of over-riding importance. The importance of elevation at shallow but not deeper depths is likely related to soil water distribution given that we observed development of a muck horizon in low-lying wet pastures. By contrast, at depth, higher pH values predicted lower SOC (although LASSO shrunk the pH coefficient to zero), whereas higher calcium was positively associated with SOC at both depths, a contrast that is somewhat counter-intuitive. Soil pH has many well-known effects on microbial activity and function, and lower pH can inhibit microbial decomposition of SOC (Kemmitt et al. 2006). The consistent and strong importance of calcium for all responses, but particularly for 5–15 cm SOC, is striking. Since this relationship only existed for calcium and not for any other of the cations or for phosphorous, autocorrelation between SOC stocks

and cation/anion-exchange capacity alone seems unlikely to drive the observed relationship. Rather, it may be that calcium plays an underappreciated role in stabilizing SOC in these soils through mineral-organic complexation (Fornara et al. 2010, Clarholm et al. 2015, Keiluweit et al. 2015), a possibility that deserves to be tested experimentally.

In this ecosystem, we found evidence of significant interannual variability in the coupling of vegetation phenology and SOC pools. EVI from some years in isolation delivered almost the same predictive power as the time series as a whole, whereas EVI from others years had almost no correlation. A primary utility of long time series is to smooth out any sampling effects that could arise from using only a single or few years of data as a composite measure. For instance, if we had averaged over an arbitrary 4-yr time window with, for example, years 11–14 of the time series, our predictive power would be extremely low, whereas years 21–24 would deliver excellent predictions. The source of interannual variability in

this system is unknown, but perhaps relates to climatic factors (e.g., duration and intensity of drought) or complex interactions of climate with landscape processes such as grazing and fire. Future work should investigate the generality of patterns within and across land uses in grazing landscapes. In the meantime, given the stochasticity in EVI-SOC relationships, we encourage further research to develop and analyze multi-decadal time series as a best practice. Fortunately, the advent of Google Earth Engine and cloud computing make the difference in difficulty between acquiring, for example, 5 yr and 30 yr of data, essentially trivial.

The correlation between EVI and SOC stocks could be due to higher primary productivity (indexed by EVI) driving higher SOC stocks, or higher SOC stocks driving greater productivity, or both. The high predictive power of certain individual years in our EVI time series (e.g., years 11–14) favors the latter possibility. Although the survey design of the present study precludes definitive disentanglement of these causal pathways, note that for the purpose of predicting SOC, both causal pathways should strengthen a positive association between EVI and SOC. Another major challenge in using satellite-derived vegetation indices to link vegetation and SOC pools is that recent empirical and theoretical work suggests that SOC may primarily originate from root system production and turnover (Rasse et al. 2005). However, the relationship between green leaf area and root production in grazed grasslands is complex, especially over short timescales where defoliation due to grazing or haying temporarily removes green leaf area and potentially alters root:shoot allocation patterns (Briske and Richards 1995, Dawson et al. 2000). Thus, future work will require more extensive experimentation to understand how vegetation indices may relate to variations in above- and belowground allocation patterns.

Leave-one-out cross-validation verified that a long time series of EVI is a powerful predictor of total SOC stock and surface SOC pools in this ecosystem. Likewise, penalized regression via LASSO (Tibshirani 1996, Friedman et al. 2010) retained EVI as a predictor across all depths. We suggest that these findings have at least two practical implications for predicting SOC at management relevant scales. Most obviously, future work with geostatistical models (e.g., Vasques et al. 2010) should acquire the longest time-series possible for EVI and/or NDVI for use as a spatial covariate, rather than continue to utilize only short time interval summaries of 1 or 2 yr. However, in assessing the overall (out-of-sample) fit of our best model it was clear that some significant sources of variation were unexplained. While there are no meaningful variations in climate or clay and silt content across our study site, it is worth noting that our best proxy of soil moisture (LiDAR elevation) is not perfect, in large part because ranch water management (i.e., drainage and sub-surface irrigation; Swain et al. 2013) can occasionally over-ride elevation in determining soil moisture at this site. More importantly, just as previous

work has incorporated remote sensing derived vegetation indices into models of plant and pasture production (e.g., Hill et al. 2004), we suggest that future work could improve estimates of SOC by utilizing these unprecedented data sets to parameterize ecosystem carbon models (e.g., CENTURY Parton et al. 1993, or a more modern microbially driven model such as MIMICS; Wieder et al. 2014) to estimate variability in SOC across space and evaluate differential spatiotemporal trends. For example, given a model and field data linking EVI/NDVI to plant productivity, a spatially explicit process-based SOC model could be run using 40 yr of variable input rates (i.e., net primary production) across the study site(s) of interest as predicted by the EVI/productivity model in each individual Landsat pixel. A spatially resolved hydrological model also could be included to better constrain variable rates of decomposition across the landscape. We argue that integration of spectral data into process models is a promising path toward generating more accurate predictive surfaces of SOC at management-relevant spatial scales.

Based on our results, we encourage greater integration of vegetation spectral data into ranch management. First, given the robust link between greenness metrics and pasture productivity (e.g., Hill et al. 2004), EVI maps such as we present in Fig. 1 could provide insight into spatial patterns of productivity across large management operations. Understanding spatial variations in EVI could enable managers of pastoral systems to target grazing in more productive areas and apply appropriate management to areas with less production, an application of “precision agriculture” concepts that are more often deployed in higher-value commodity production (e.g., Lee and Ehsani 2015). Second, as ecosystem scientists improve our mechanistic understanding of the linkages between productivity and SOC stocks in these systems, similar spatially explicit renderings of EVI data could prove useful for monitoring and predicting future SOC. If confirmed by subsequent experimental investigation, our finding that calcium availability is an independent driver of SOC suggests that increased fertilization with this element (e.g., via lime or gypsum applications) may be a cost-effective approach to enhancing SCS in addition to providing other agronomic benefits (e.g., pH amelioration). Although the observational design of our present work precludes strict estimation of causal effects, we note that the soils in the upper ranges of Ca availability were associated with 0.75 standard deviation higher SOC stocks, representing an additional 10 Mg C/ha.

Overall, we found strong evidence that remote sensing of long-term vegetation EVI can be critical for predicting SOC. In contrast to previous work, these data did not enhance predictive power by discriminating coarse vegetation types (as in large landscape transects such as Yang et al. 2008), but instead reflect deeper relationships between pasture green leaf area and SOC. The relationship between EVI and SOC appears to have a high

degree of interannual variability, a finding that calls for investigation. Nevertheless, remote sensing of vegetation can help scale predictions of SOC from plots to management-relevant units. Accordingly, we call for deeper and more extensive testing of the utility of long time series of remotely sensed vegetation indices for predicting SOC stocks across various ecosystems. These improvements in efficient SOC estimation may help promote adaptive management of grazing lands to optimize SCS in parallel with other ecosystem services such as biodiversity maintenance and livestock production.

ACKNOWLEDGMENTS

Patrick Norby provided significant assistance with collecting field data. Samantha Miller provided helpful advice for processing soils in the lab. Chris H. Wilson was partially supported by NSF DEB DDIG grant 1501686. T. Trevor Caughlin was supported by NSF grant no. 1415297 Science, Engineering and Education for Sustainability Fellow program.

LITERATURE CITED

- Bohlen, P. J., S. Lynch, L. Shabman, M. Clark, S. Shukla, and H. Swain. 2009. Paying for environmental services from agricultural lands: an example from the northern Everglades. *Frontiers in Ecology and the Environment* 7:46–55.
- Bolker, B. M., S. W. Pacala, and W. J. Parton. 1998. Linear analysis of soil decomposition: insights from the century model. *Ecological Applications* 8:425–439.
- Bradley, B. A., R. W. Jacob, J. F. Hermance, and J. F. Mustard. 2007. A curve fitting procedure to derive inter-annual phenologies from time series of noisy satellite NDVI data. *Remote Sensing of Environment* 106:137–145.
- Briske, D., and J. Richards. 1995. Plant responses to defoliation: a physiological, morphological and demographic evaluation. Pages 635–710 in D. J. Bedunah and R. E. Sosebee, editor. *Wildland plants: physiological ecology and developmental morphology*. Society for Range Management, Denver, CO.
- Burke, I. C., C. M. Yonker, W. J. Parton, C. V. Cole, D. S. Schimel, and K. Flach. 1989. Texture, climate, and cultivation effects on soil organic matter content in U.S. Grassland Soils. *Soil Science Society of America Journal* 53:800.
- Chapin, F. S. 2009. Managing ecosystems sustainably: the key role of resilience. Pages 29–53 in C. Folke, G. P. Kofinas, and F. S. Chapin, editors. *Principles of ecosystem stewardship*. Springer, New York, New York, USA.
- Chapin III, F. S., P. A. Matson, and P. M. Vitousek. 2012. *Principles of terrestrial ecosystem ecology*. Second edition. Springer, New York, New York, USA.
- Clarholm, M., U. Skjellberg, and A. Rosling. 2015. Organic acid induced release of nutrients from metal-stabilized soil organic matter—the unbutton model. *Soil Biology and Biochemistry* 84:168–176.
- Conant, R. T., and K. Paustian. 2002a. Potential soil carbon sequestration in overgrazed grassland ecosystems. *Global Biogeochemical Cycles* 16:90–109.
- Conant, R. T., and K. Paustian. 2002b. Spatial variability of soil organic carbon in grasslands: implications for detecting change at different scales. *Environmental Pollution* 116 (Supplement 1):S127–S135.
- Conant, R. T., K. Paustian, and E. T. Elliott. 2001. Grassland management and conversion into grassland: effects on soil carbon. *Ecological Applications* 11:343–355.
- Cook, B. I., and S. Pau. 2013. A global assessment of long-term greening and browning trends in pasture lands using the GIMMS LAI3 g dataset. *Remote Sensing* 5:2492–2512.
- Cotrufo, M. F., M. D. Wallenstein, C. M. Boot, K. Deneff, and E. Paul. 2013. The microbial efficiency-matrix stabilization (MEMS) framework integrates plant litter decomposition with soil organic matter stabilization: do labile plant inputs form stable soil organic matter? *Global Change Biology* 19: 988–995.
- Davidson, E. A., and I. A. Janssens. 2006. Temperature sensitivity of soil carbon decomposition and feedbacks to climate change. *Nature* 440:165–173.
- Dawson, L. A., S. J. Grayston, and E. Paterson. 2000. Effects of grazing on the roots and rhizosphere of grasses. Pages 61–84 in G. Lemaire, J. Hodgson, A. de Moraes, C. Nabinger, and P. C. de F. Carvalho, editors. *Grassland ecophysiology and grazing ecology*. CABI, Wallingford, UK.
- DeGryze, S., J. Six, K. Paustian, S. J. Morris, E. A. Paul, and R. Merckx. 2004. Soil organic carbon pool changes following land-use conversions. *Global Change Biology* 10:1120–1132.
- Durán, J., M. Delgado-Baquerizo, A. Rodríguez, F. Covelo, and A. Gallardo. 2013. Ionic exchange membranes (IEMs): A good indicator of soil inorganic N production. *Soil Biology and Biochemistry* 57:964–968.
- Eigenbrod, F., P. R. Armsworth, B. J. Anderson, A. Heinemeyer, S. Gillings, D. B. Roy, C. D. Thomas, and K. J. Gaston. 2010. The impact of proxy-based methods on mapping the distribution of ecosystem services. *Journal of Applied Ecology* 47:377–385.
- Evans, S. E., I. C. Burke, and W. K. Lauenroth. 2011. Controls on soil organic carbon and nitrogen in Inner Mongolia, China: A cross-continental comparison of temperate grasslands. *Global Biogeochemical Cycles* 25:GB3006.
- Fornara, D. A., S. Steinbeiss, N. P. McNAMARA, G. Gleixner, S. Oakley, P. R. Poulton, A. J. Macdonald, and R. D. Bardgett. 2010. Increases in soil organic carbon sequestration can reduce the global warming potential of long-term liming to permanent grassland. *Global Change Biology* 17:1925–1934.
- Friedman, J., T. Hastie, and R. Tibshirani. 2010. Regularization paths for generalized linear models via coordinate descent. *Journal of Statistical Software* 33:1–22.
- Gelman, A. 2008. Scaling regression inputs by dividing by two standard deviations. *Statistics in Medicine* 27:2865–2873.
- Gu, Y., B. K. Wylie, and N. B. Bliss. 2013. Mapping grassland productivity with 250-m eMODIS NDVI and SSURGO database over the Greater Platte River Basin, USA. *Ecological Indicators* 24:31–36.
- Hansen, M. C., et al. 2013. High-resolution global maps of 21st-century forest cover change. *Science* 342:850–853.
- Hill, M. J., G. E. Donald, M. W. Hyder, and R. C. G. Smith. 2004. Estimation of pasture growth rate in the south west of Western Australia from AVHRR NDVI and climate data. *Remote Sensing of Environment* 93:528–545.
- Houborg, R. M., and H. Soegaard. 2004. Regional simulation of ecosystem CO₂ and water vapor exchange for agricultural land using NOAA AVHRR and Terra MODIS satellite data. Application to Zealand, Denmark. *Remote Sensing of Environment* 93:150–167.
- Huete, A., K. Didan, T. Miura, E. P. Rodriguez, X. Gao, and L. G. Ferreira. 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment* 83:195–213.
- Jenny, H. 1961. Derivation of state factor equations of soils and ecosystems I. *Soil Science Society of America Journal* 25:385.
- Keiluweit, M., J. J. Bougoure, P. S. Nico, J. Pett-Ridge, P. K. Weber, and M. Kleber. 2015. Mineral protection of soil

- carbon counteracted by root exudates. *Nature Climate Change* 5:588–595.
- Kemmitt, S. J., D. Wright, K. W. T. Goulding, and D. L. Jones. 2006. pH regulation of carbon and nitrogen dynamics in two agricultural soils. *Soil Biology and Biochemistry* 38:898–911.
- Lal, R. 2004. Soil carbon sequestration impacts on global climate change and food security. *Science* 304:1623–1627.
- Lal, R. 2010. Managing soils and ecosystems for mitigating anthropogenic carbon emissions and advancing global food security. *BioScience* 60:708–721.
- Lee, W. S., and R. Ehsani. 2015. Sensing systems for precision agriculture in Florida. *Computers and Electronics in Agriculture* 112:2–9.
- Manning, P., et al. 2015. Simple measures of climate, soil properties and plant traits predict national-scale grassland soil carbon stocks. *Journal of Applied Ecology* 52:1188–1196.
- Masek, J. G., E. F. Vermote, N. E. Saleous, R. Wolfe, F. G. Hall, K. F. Huemmrich, F. Gao, J. Kutler, and T.-K. Lim. 2006. A Landsat surface reflectance dataset for North America, 1990–2000. *IEEE Geoscience and Remote Sensing Letters* 3:68–72.
- Moore, R. T., and M. C. Hansen. 2011. Google Earth Engine: a new cloud-computing platform for global-scale earth observation data and analysis. *AGU Fall Meeting Abstracts* 43:1.
- O'Rourke, S. M., D. A. Angers, N. M. Holden, and A. B. McBratney. 2015. Soil organic carbon across scales. *Global Change Biology* 21:3561–3574.
- Parton, W. J., et al. 1993. Observations and modeling of biomass and soil organic matter dynamics for the grassland biome worldwide. *Global Biogeochemical Cycles* 7:785–809.
- Power, A. G. 2010. Ecosystem services and agriculture: trade-offs and synergies. *Philosophical Transactions of the Royal Society B* 365:2959–2971.
- Rasse, D. P., C. Rumpel, and M.-F. Dignac. 2005. Is soil carbon mostly root carbon? Mechanisms for a specific stabilisation. *Plant and Soil* 269:341–356.
- Rees, R. M., I. J. Bingham, J. A. Baddeley, and C. A. Watson. 2005. The role of plants and land management in sequestering soil carbon in temperate arable and grassland ecosystems. *Geoderma* 128:130–154.
- Saby, N. P. A., et al. 2008. Will European soil-monitoring networks be able to detect changes in topsoil organic carbon content? *Global Change Biology* 14:2432–2442.
- Schimel, D. S., B. H. Braswell, E. A. Holland, R. McKeown, D. S. Ojima, T. H. Painter, W. J. Parton, and A. R. Townsend. 1994. Climatic, edaphic, and biotic controls over storage and turnover of carbon in soils. *Global Biogeochemical Cycles* 8:279–293.
- Silveira, M. L., A. K. Obour, J. Arthington, and L. E. Sollenberger. 2011. The cow-calf industry and water quality in South Florida, USA: a review. *Nutrient Cycling in Agroecosystems* 89:439–452.
- Silveira, M. L., S. Xu, J. Adewopo, A. J. Franzluebbers, and G. Buonadio. 2014. Grazing land intensification effects on soil C dynamics in aggregate size fractions of a Spodosol. *Geoderma* 230–231:185–193.
- Smith, P. 2004. How long before a change in soil organic carbon can be detected? *Global Change Biology* 10:1878–1883.
- Stockmann, U., et al. 2013. The knowns, known unknowns and unknowns of sequestration of soil organic carbon. *Agriculture, Ecosystems and Environment* 164:80–99.
- Swain, H. M., P. J. Bohlen, K. L. Campbell, L. O. Lollis, and A. D. Steinman. 2007. Integrated ecological and economic analysis of ranch management systems: an example from South Central Florida. *Rangeland Ecology and Management* 60:1–11.
- Swain, H. M., E. H. Boughton, P. J. Bohlen, and L. O. Lollis. 2013. Trade-offs among ecosystem services and disservices on a Florida ranch. *Rangelands* 35:75–87.
- Tibshirani, R. 1996. Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society: Series B (Methodological)* 58:267–288.
- Tibshirani, R. 2011. Regression shrinkage and selection via the lasso: a retrospective. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* 73:273–282.
- Vasques, G. M., S. Grunwald, J. O. Sickman, and N. B. Comerford. 2010. Upscaling of dynamic soil organic carbon pools in a North-Central Florida watershed. *Soil Science Society of America Journal* 74:870.
- Wieder, W. R., A. S. Grandy, C. M. Kallenbach, and G. B. Bonan. 2014. Integrating microbial physiology and physiochemical principles in soils with the microbial-mineral carbon stabilization (MIMICS) model. *Biogeosciences* 11:3899–3917.
- Wulder, M. A., J. C. White, T. R. Loveland, C. E. Woodcock, A. S. Belward, W. B. Cohen, E. A. Fosnight, J. Shaw, J. G. Masek, and D. P. Roy. 2016. The global Landsat archive: status, consolidation, and direction. *Remote Sensing of Environment* 185:271–283.
- Yang, Y., J. Fang, Y. Tang, C. Ji, C. Zheng, J. He, and B. Zhu. 2008. Storage, patterns and controls of soil organic carbon in the Tibetan grasslands. *Global Change Biology* 14:1592–1599.

SUPPORTING INFORMATION

Additional supporting information may be found online at: <http://onlinelibrary.wiley.com/doi/10.1002/eap.1557/full>

DATA AVAILABILITY

Data associated with this paper have been deposited in a Dryad digital repository <https://doi.org/10.5061/dryad.266m0>