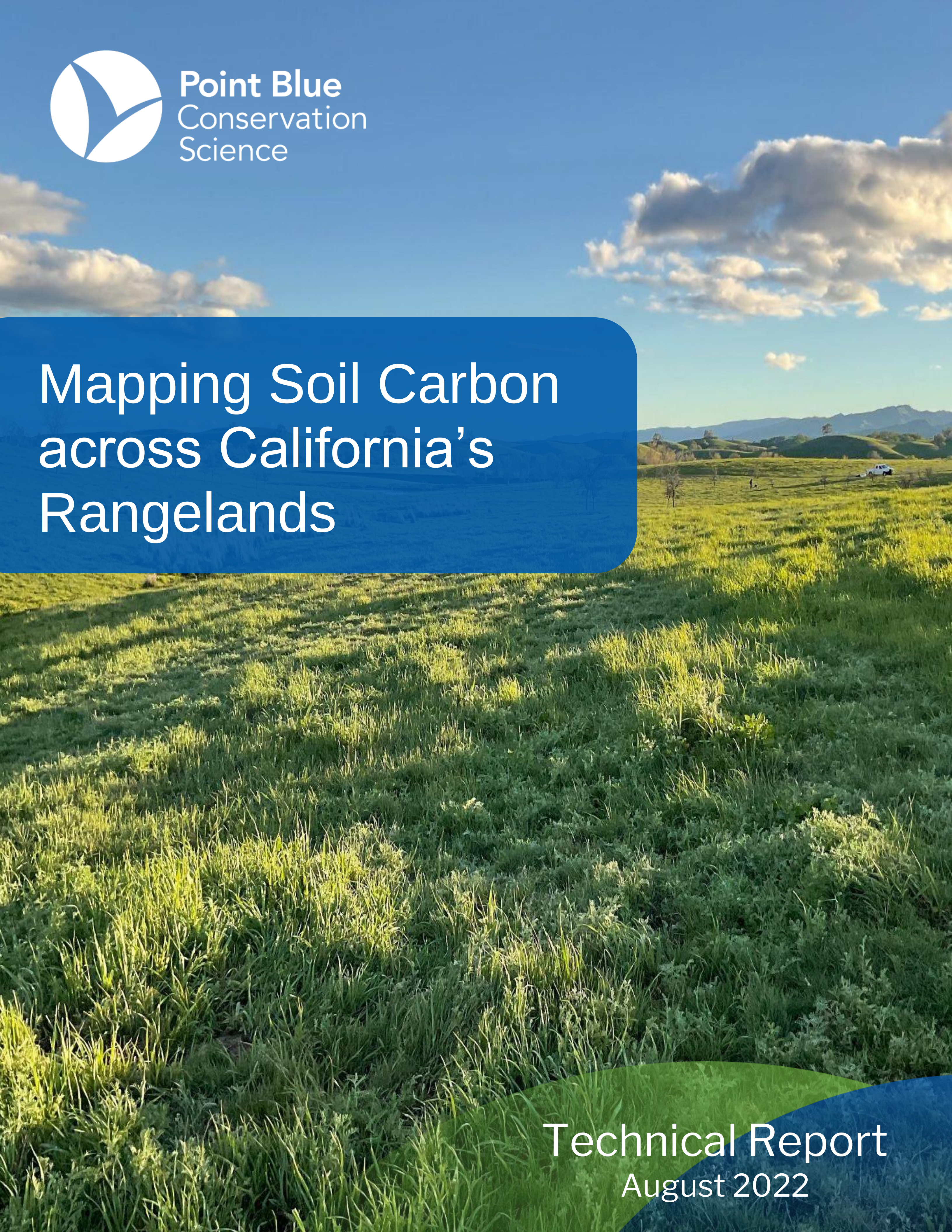


Point Blue Conservation Science

Mapping Soil Carbon across California’s Rangelands

Technical Report

August 2022

Mapping Soil Carbon across California Rangelands

Final Technical Report – August 2022

Prepared by

Point Blue Conservation Science

Sam Veloz, PhD
Nathan Elliott
Libby Porzig, PhD
Chelsea Carey, PhD

Suggested citation:
Veloz, S.*, Elliott, N., Porzig, L., C.J. Carey. 2022. Mapping California Rangeland Soil Carbon: A Technical Report. Point Blue Conservation Science (Contribution 2424), Petaluma, CA.

*Corresponding author: sveloz@pointblue.org

Point Blue Conservation Science – Point Blue's 160 scientists work to reduce the impacts of climate change, habitat loss, and other environmental threats while developing nature-based solutions to benefit both wildlife and people.

Conservation science for a healthy planet

3820 Cypress Drive, #11 Petaluma, CA 94954
T 707.781.2555 | **F** 707.765.1685

TABLE OF CONTENTS

Contents

TABLE OF CONTENTS 2
INTRODUCTION 3
OBJECTIVES 4
METHODS 4
Study Area 4
Soil Field Data 5
Variable Selection 5
Modeling Carbon Dynamics 7
OUTCOMES AND CONCLUSIONS 9
Mapping Average Carbon Stocks 9
Mapping Change Over Time 17
Investigating patterns of changes in carbon stocks 18
Conclusions 19
Acknowledgements 19
Appendix 1. Partial Dependence Plots 20
Appendix 2. Estimated Carbon Stocks on Rangelands by County 26
References 27

INTRODUCTION

Soil organic carbon (hereafter 'soil carbon') is an important natural resource that plays a critical role in helping natural and managed ecosystems mitigate and adapt to climate change (Bossio et al. 2020). As a primary part of soil organic matter, soil carbon also provides many other adjacent services, including supplying nutrients to plants, providing food and habitat for soil biodiversity, storing and purifying water, and contributing to soil structure (Bradford et al. 2019). Unfortunately, a significant portion of soil carbon has been lost globally due to past land use and mismanagement (Sanderman et al. 2017), a trend that is now exacerbated by climatic stressors such as drought (Deng et al. 2021). Thus, there is both a strong need and opportunity to protect and rebuild soil carbon through stewardship, which would result in a myriad of benefits to humans and the ecosystems on which we depend.

Accurate estimates of regional soil carbon levels are needed to guide both policymakers and practitioners in their endeavor to protect and rebuild this critical natural resource (Carey et al. 2020). From a policy perspective, access to estimates of soil carbon can inform carbon inventories and climate scoping plans (e.g., CARB 2017), and can help to spatially prioritize protection of high soil carbon areas. It can also ensure funds and resources are allocated to restoration projects in areas with the greatest need or potential. From a land management perspective, the same information can help producers better understand how their land compares to the broader landscape, identify areas where low levels of soil carbon may be a resource concern, set reasonable targets, and execute management strategies through Carbon Farm Plans. Combined, these actions can help to de-risking public and private investment in soil carbon stewardship.

Collecting empirical data via in-field soil sampling is a reliable way to produce estimates of soil carbon levels. However, collecting soil carbon data at high densities across broad spatial scales is challenging. Digital soil mapping offers a way to leverage empirical field measurements to generate continuous estimates of soil carbon across a property, a region, and even a continent—filling in the gaps and thereby supporting both policy and practice.

For the past seven years, Point Blue has been collecting soil carbon data across California's rangelands using standardized monitoring methods as part of the Rangeland Monitoring Network (RMN; Porzig et al. 2018). This effort has produced a dataset that offers a unique opportunity to map soil carbon levels across California's rangelands using digital soil mapping approaches. With increasing interest in managing soil carbon from policymakers and practitioners, we aimed to use these data to create predictive maps of soil carbon stocks across California's rangelands.

OBJECTIVES

The objectives of this project were to combine RMN soil data with publicly available geographic information service (GIS) data to predict soil carbon across California's rangelands; to determine the accuracy of those predictions; and to produce maps that can be used to inform carbon stewardship of these rangeland soils. In addition to predicting static ("baseline") carbon stocks, we aimed to predict changes in carbon stocks between 2015-2021 as well. As an additional way to identify sites that are ripe for management intervention, we also compared surface carbon stocks with deeper carbon stocks, with the assumption that sites with less surface carbon than expected based on deep carbon values have the potential for improvement. Our final objective was to assess how soil carbon changed over time at sites that had lower than expected surface carbon, which would provide additional insights to guide strategic management interventions.

METHODS

To achieve our objectives, we combined existing soil data collected through the RMN with explanatory variables that we expected to be good predictors of soil carbon stocks at sites that we sampled, and then evaluated model fit. Then, using the same GIS layers for predictor variables, we applied the model to predict carbon across unsampled spatial locations and generated predictive maps for the region. We also analyzed the unexplained variation between soil carbon stocks in shallow and deeper soil horizons to estimate legacy effects of land and inform future management. This process is described in more detail below.

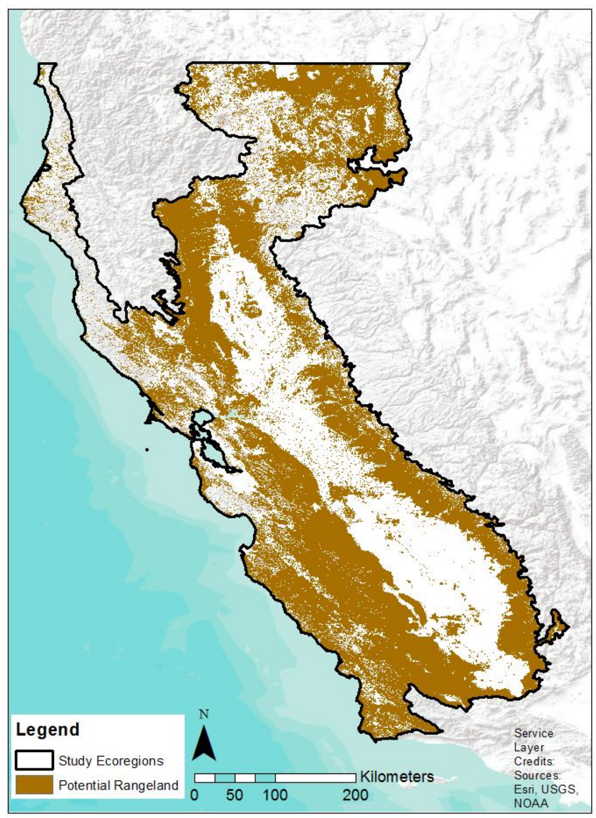


Figure 1. Project study area of rangelands in Northern and Central California, set by ecoregion boundaries and potential rangeland habitat as described in text.

Study Area

For this project, we focused on the rangeland areas of central and northern California which represented the environmental conditions within which we have field samples. We defined potential rangelands using a union of the rangeland extent in the National Land Cover Database (NLCD) Rangeland Components

Dataset (Rigge et al. 2020) and the classes of 'shrub' or 'herbaceous' in the California Department of Forestry and Fire Protection's vegetation map (CALFIRE-FRAP 2014). We then masked this to the California Ecoregions (Griffiths et al. 2016) of the Central Valley, Central California Foothills, Coast Range, Cascades, and Eastern Cascades (Figure 1).

Soil Field Data

Soil carbon concentrations were measured from 0-10 cm and 10-40 cm depth at 282 sites across California. Samples were collected in a standardized way and analyzed at University of Idaho Analytical Lab via dry combustion. Each point is in one of three cohorts, and each cohort was sampled on a rotating basis every three years from 2015 – 2021. Bulk density measurements were also taken at each site and used to convert carbon concentrations to stocks on a fixed mass basis (Mg C / hectare). This information was then used to calculate average carbon stocks for each point across the time period as well as an average rate of yearly change. See Carey et al. (2020) for a full description of collection methodology.

Variable Selection

We selected initial explanatory variables based on a priori knowledge, which was informed in part by two reviews of soil carbon modeling: Gomes et al (2019) and Keskin et al. (2019). We used prior work that focused on modeling soil carbon in the Pescadero area of California to further focus this list (Veloz et al., 2021) and then eliminated any variables for which there were not statewide layers available. This ruled out using land management practices (e.g., mowing, burning, grazing), some soil information (e.g., water infiltration time and content of iron and aluminum oxides; we were limited to data available in Soil Survey Geographic Database [SSURGO]), and a detailed land use history, among other things. These are all factors that are known to relate to soil carbon across space and time (Delgado-Baquerizo et al. 2018), but for which we were unable to find the necessary datasets for inclusion.

With that in mind, soil carbon data points were attributed with the following explanatory data:

- Landcover type from the National Land Cover Database (NLCD) in 2016, 30 m resolution (Dewitz 2019). Most of the classes for this dataset are non-rangeland classes and are used to describe the land cover across the United States.

- Elevation from the Shuttle Radar Topography Mission Digital Elevation Dataset, 30 m resolution (Farr et al 2007).

- Climate data, including monthly average winter minimum temperature (Dec – Feb) and average summer maximum temperature (Jun – Aug), as well as annual precipitation, runoff, recharge, storage, and climactic water deficit, averaged across 2016 – 2021 from California's Basin Characterization Model v8, 270m resolution (Flint and Flint, 2017).

- Nine measures of annual vegetative productivity derived from MODIS Normalized Difference Vegetation Index (NDVI) data: amplitude (AMP), duration (DUR), end of season NDVI (EOSN), end of season time (EOST), maximum NDVI (MAXN), maximum growing season time (MAXT), start of season NDVI (SOSN), start of season time (SOST), and time-integrated NDVI (TIN), averaged 2016 – 2021, 250m resolution (Meier and Brown 2014).

- Fractional landcover of bare ground, litter, annual, annual herbaceous, shrub, and sagebrush as well as sagebrush and shrub height in 2016 from the NLCD Rangeland Components dataset, 30m resolution (Rigge et al. 2020).

- Soil class, suborder, order, and drainage class; and the weighted average by horizon of pH, sand, silt, and clay from SSURGO (Soil Survey Staff, 2022). We used the gridded version of the SSURGO data, which has a nominal resolution of 10m, but this is derived from a shapefile with non-standard sampling. For example, if SSURGO reported two horizons at a point, one with a depth of 0 - 8 cm and one with a depth of 8 - 100 cm, the SSURGO values for the 0 - 10 cm model would be (4 * 0-8cm val + 1 * 8-100 cm val) / 5. After attribution, we calculated the geometric mean particle size from soil texture data using the method described in Carey et al. (2020).

Prior to modeling carbon stocks, we conducted an exploratory analysis of all explanatory variables to determine their suitability for modeling. We eliminated variables that were highly correlated with one another. If two explanatory variables were highly correlated ($r > 0.8$), we eliminated the one that had higher average correlation with other variables. This resulted in removing the following variables from the model: sagebrush cover, big sagebrush cover, annual herbaceous cover, shrub height, sagebrush height, AMP, DUR, EOSN, and EOST.

A second criteria removed variables that were not well sampled with our data points. For categorical data, we eliminated all variables that had categories covering more than a quarter of our study area for which we had three or fewer sample points. This removed soil class and suborder. For continuous data, we first checked that no variables had >90% of the samples within <10% of the data range. We additionally ensured that variables had sufficient variation across the study area.

Finally, we removed soil drainage class from the model because it was redundant with the more precise, continuous recharge and climatic water deficit from the California Basin Characterization data.

Modeling Carbon Dynamics

Once input variables were selected, we performed a round of exploratory modeling to determine the best model parameters. We applied boosted regression trees using the gbm.step function in the dismo package in R (Hijmans et al. 2022) as described by Elith et al (2008). Boosted regression trees is a machine learning algorithm that combines a set of very simple classification trees. The algorithm iteratively adds new trees to the set and at each step focuses on explaining the remaining unexplained variation from the set of previous trees. The analyst selects the final settings of parameters for the algorithm by balancing the ability of the model to explain the variation in the input data while also being able to predict to data withheld from the creation of the mode.

We tested the following combinations of input parameters: learning rates of 0.0001, 0.0005, 0.001, 0.005, 0.01, 0.05, and 0.1; tree complexities of 1, 2, 3, and 4; and bag fractions of 0.6, 0.7, and 0.8. We ran all combinations once with categorical variables included and once without. We were concerned with including categorical variables since we rarely had adequate field samples that included a sufficient sample of all classes in each variable. Models were evaluated against the training (data used to build the model) and testing data (data withheld from model creation used to validate predictions) using the metrics of tree count, correlation, cross-validated correlation, deviance explained, and cross-validated deviance explained.

We went through this process for a number of different calculated soil carbon response variables and depths. As soil carbon was measured at two depths, we modeled 0 – 10 cm, 10 – 40 cm, and a combined 0 – 40 cm. For each depth, we began by modeling static soil carbon stocks, which were calculated by averaging soil carbon values at each point across years. Given that average soil carbon stock values were approximately 60 times larger than changes over time, we felt justified in modeling the average rather than, for instance, the most recent time point.

However, we were also interested in assessing the change over time. To do this, we started by modeling the yearly change in soil carbon using a Gaussian link. We then created a hurdle model by first using a logistic link boosted regression to predict change / no change (classing change of < 10% as 0) and then using the aforementioned method to predict the amount of change for those points that were predicted to change.

We tested the Gaussian link model applying several different filters to the data to gauge their impact on model performance. We ran the model first on all of the data. Then, to focus in on points that were changing, we ran the model with all points that had relatively constant values of soil carbon across years removed. We tested two removal thresholds: 0.1 Mg C ha/yr, which is on the low end for what is reported in the literature (Conant et al. 2017) and 0.5 Mg C ha/yr, chosen based on the distribution of data. To prevent large changes from skewing the model, we also ran the model removing points that had extremely large changes across the 3 or 6 year sample period (total change > 30% of initial value) or changes that were deemed unreasonable or unlikely (change > 8 Mg C ha/y). We also tried several different transformations of the data, including a binary change/no change threshold, a log transformation, and a multiplicative transformation.

As another way to identify areas in need of management intervention, we compared surface carbon with deeper carbon. In general, we expect the carbon at deeper depths would be correlated with the carbon in shallower depths. However, we also expect that carbon in shallower soil depths would be more sensitive to management actions than carbon within deeper depths (Ward et al. 2016). The relationship between the two depths should therefore tell us something about locations that are underperforming based on past management or other factors. To illustrate areas where this may be the case, we conducted two linear regressions: one of the measured points comparing surface to deep carbon (linear regression 1), and one of the mapped predictions of soil carbon at 10 - 40 cm depths to explain the mapped predictions of soil carbon at 0 - 10 cm (linear regression 2). We then mapped the residuals (mapped prediction - prediction from linear regression) from the second regression model. Negative values on the resulting map can be interpreted as areas where management or other factors may be having an undesirable effect on soil carbon at shallower depths.

To extend our inference, we used the residuals of the first regression model to explain the observed annual change in carbon stocks at each point based on our initial and final field samples. We expect that variation in annual rates of change in carbon stocks will significantly correlate with areas that have higher or lower than expected surface carbon based on the previous analysis, and that this will help to further our understanding of areas to target for future management.

All analysis and most data preparation was performed in R v4.2, using the packages dismo (Hijmans et al. 2022), dplyr (Wickham et al. 2022), gbm (Greenwell et al. 2020), raster (Hijmans 2022), and sp (Pebesma and Bivand 2005). Mapping and some data preparation was also performed in ArcMap v10.8 (ESRI 2021).

OUTCOMES AND CONCLUSIONS

Mapping Average Carbon Stocks

Soil carbon stocks across our network ranged from 3.62 - 60.27 (Mg C/ha) for 0-10 cm and 15.94 - 230 (Mg C/ ha) for 10-40 cm. These values fall within what we would expect for semi-arid rangelands of California (Carey et al. 2020; Silver et al. 2010, Silver et al. 2018).

We were able to model average soil carbon stocks with a high level of accuracy. Our best model for the 0 - 10 cm depth had a correlation between the observed average soil carbon stocks and predicted soil carbon stocks of 0.723 (SE ± 0.022; residual deviance = 27.1%). Our best model for the 10 - 40 cm depth was slightly better, having a correlation of 0.845 (SE ± 0.023; residual deviance = 17.5%. Our best model for the combined 0 - 40cm depth had a correlation of 0.847 (SE ± 0.015; residual

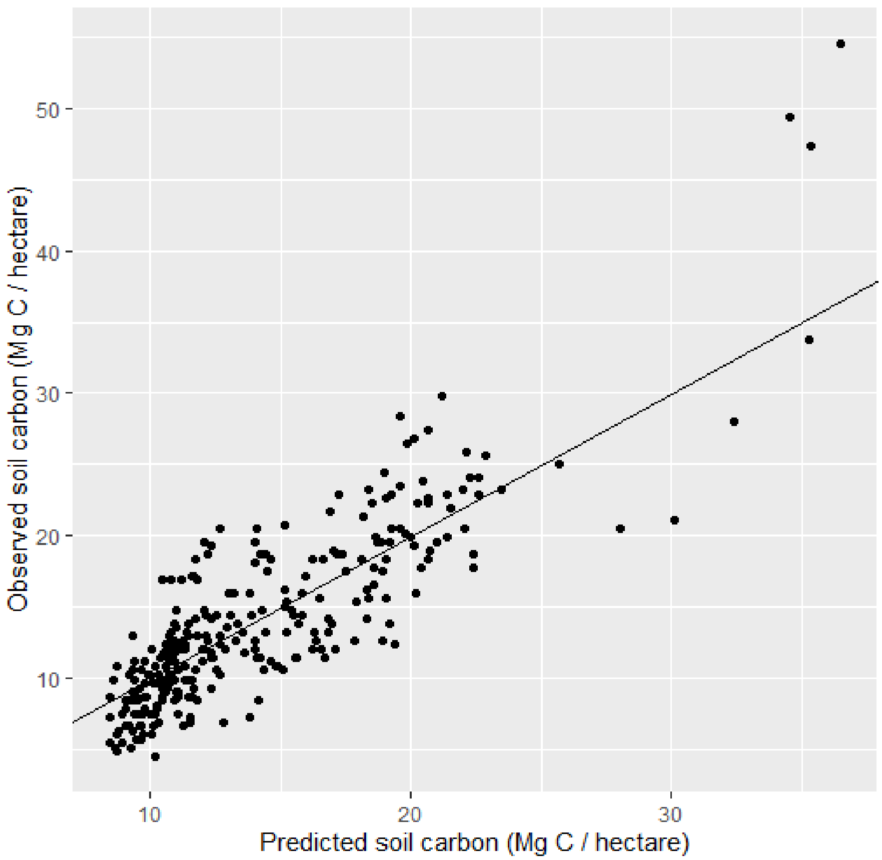


Figure 2. Observed vs. predicted soil carbon at the 0 - 10 cm depth. The black line is the 1:1 line. Points above the line are where predictions are less than observed. Points below the line are where predictions are greater than observed.

deviance = 17.9%). Observed vs predicted values are shown in figures 2 and 3 for the 0 - 10 cm and 10 - 40 cm depth models respectively.

Models without landcover included as a categorical variable (NLCD) performed similarly to or slightly better than models with landcover. However, given the lack of sampling in non-rangeland (i.e., most) landcover categories, we chose to use the models without categorical landcover included as a covariate.

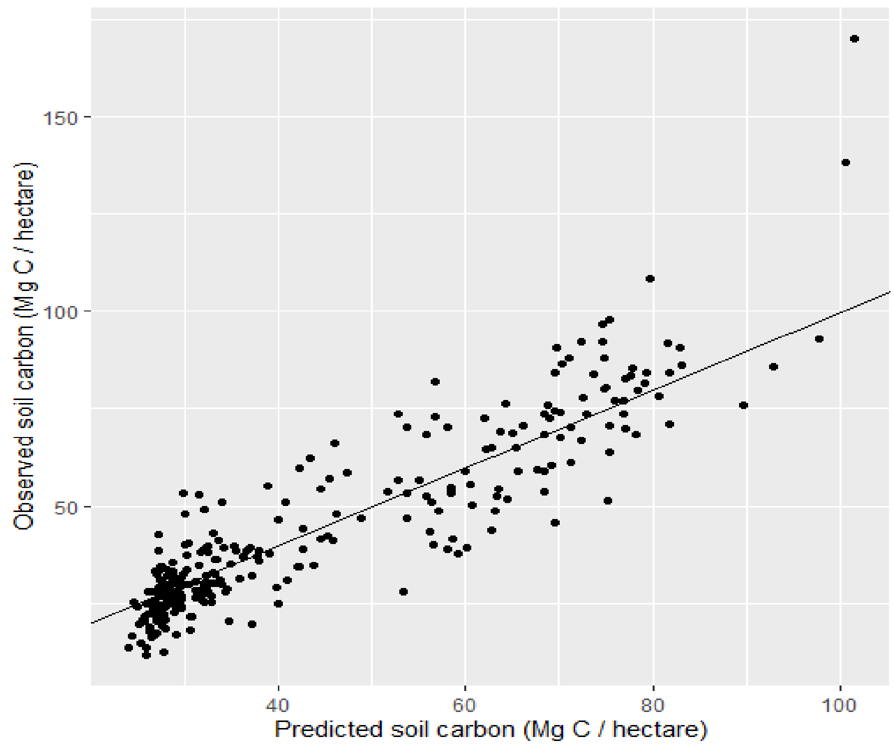


Figure 3. Observed vs. predicted soil carbon at the 10 - 40 cm depth. The black line is the 1:1 line. Points above the line are where predictions are less than observed. Points below the line are where predictions are greater than observed.

In the 0 - 10 cm model, the most important predictor was the start of season NDVI, with a relative influence of 22.1% (Table 1, left column). This was followed by the average minimum temperature in winter (17.6%), the start of season time (8.3%), average max temperature in summer (6.3%), and maximum NDVI (6.3%). Eight predictors had a relative influence of between 2 and 5%; they were, in descending order, shrub cover, growing season duration, climatic water deficit, bare ground, annual recharge, maximum growing season time, annual runoff, and time-integrated NDVI. No other predictors had a relative influence greater than 2%.

In the 10 - 40 cm model, the most important predictor was the average minimum temperature in winter, which had an importance of 45.7% (Table 1, right). Start of season NDVI came next (13.5%), followed by average maximum summer temperature (11.8%). Seven predictors had a relative influence of between 2 and 5%; they were, in descending order, shrub cover, bare ground, growing season duration, climatic water deficit, mean soil particle size, start of season time, and maximum NDVI. The remaining variables had a relative influence of less than 2%.

0 - 10 cm soil depth				10 - 40 cm soil depth			
Variable	Relative Influence %	Variable	Relative Influence %	Variable	Relative Influence %	Variable	Relative Influence %
Start of season NDVI	22.14	Maximum growing season time	3.25	Winter minimum temperature	45.71	Maximum growing	1.72
Winter minimum temperature	17.62	Annual runoff	2.19	Start of season NDVI	13.51	Time-integrated NDVI	1.23
Start of season time	11.94	Time-integrated NDVI	2.09	Summer maximum temperature	11.84	Annual runoff	0.98
Summer maximum temperature	8.28	pH	1.49	Shrub cover	4.06	Soil water storage	0.82
Maximum NDVI	6.33	Soil water storage	1.45	Bare ground	3.68	Annual precipitation	0.75
Shrub cover	4.52	Mean soil particle size	1.09	Growing season duration	3.65	Elevation	0.63
Growing season duration	4.14	Annual precipitation	0.69	Climatic water deficit	3.02	pH	0.55
Climatic water deficit	4.12	Leaf litter	0.63	Mean soil particle size	2.52	Leaf litter	0.38
Bare ground	3.72	Elevation	0.55	Start of season time	2.25	Annual recharge	0.32
Annual recharge	3.28	Annual herbaceous cover	0.48	Maximum NDVI	2.09	Annual herbaceous cover	0.28

Table 1. Relative influence of predictor variables in our top models of average soil carbon on rangeland points in central and northern California. The left two columns show the relative influence of variables in the 0 - 10 cm model; the right two columns show the relative influence of variables in the 10 - 40 cm model.

Our model of carbon at 10 - 40 cm was heavily driven by one variable, winter temperature; its relative influence was twice that of the top variable in the 0 - 10 cm model. Temperature variables were more important in the 10 - 40 cm model than the 0 - 10 cm model. Winter minimum temperatures and summer maximum temperatures had a combined influence of 57% in the deeper model, as compared to 26% in the shallow model.

In contrast, vegetation phenology metrics were more influential in predicting carbon stocks at the shallower depth of 0 - 10 cm than at the 10 - 40 cm depth. The top shallow depth models had SOSN and SOST as the first and third variables in importance (average influences of 22 and 12%, respectively), with MAXN, MAXT, and DUR all having an average influence of 3 - 6%. The deep models placed only one phenology metric in the top three variables, SOSN had the second highest relative influence.

Neither model was heavily influenced by the fractional land cover variables included. Shrub, litter, and bare ground coverage had relative influences of <5% in both models. Included soil characteristics (soil particle size, pH, and soil water storage) had even less impact, with a maximum influence of 2.5% (soil particle size, 10 - 40 cm).

Maps of our predicted soil carbon are provided in figures 4 and 5. We predicted higher soil carbon stocks in areas with greater coastal influence for both depths. Additionally, we predicted relatively high soil carbon stocks in the very north-east of California within the Modoc plateau. In contrast, we predicted relatively low carbon stocks at both depths around the border of the Central Valley and moderate soil carbon in the foothills of the Sierra Nevada Mountains (Figures 4 and 5). These patterns are expected based on known relationships between soil carbon storage, temperature, and precipitation (Delgado-Baquerizo et al. 2018).

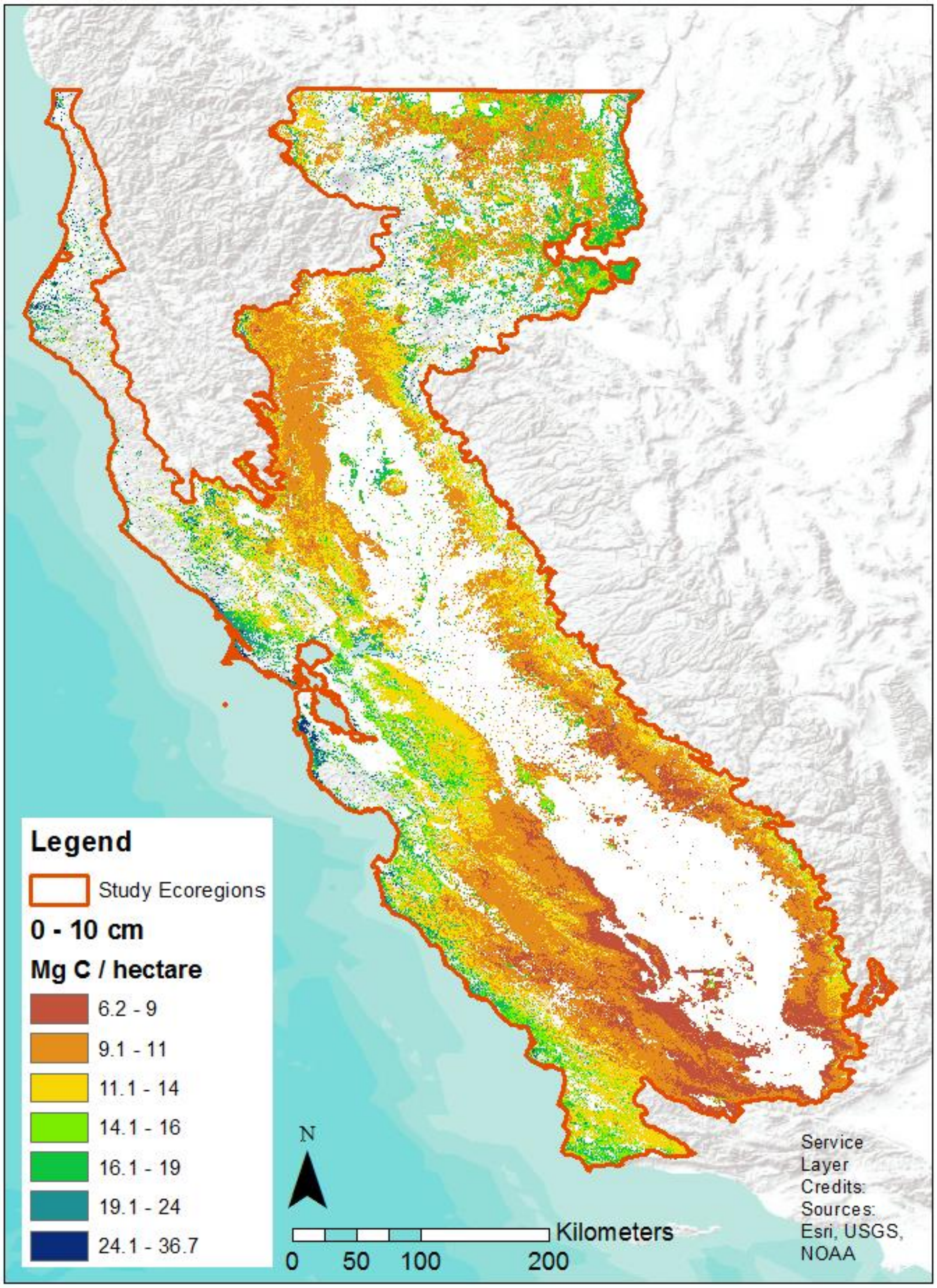


Figure 4. Predicted soil carbon stocks in California rangelands at 0 - 10 cm depth. An online version of the map is available here https://www.arcgis.com/home/item.html?id=cd4d39fa5cb44216a1b064e3120b8ba0

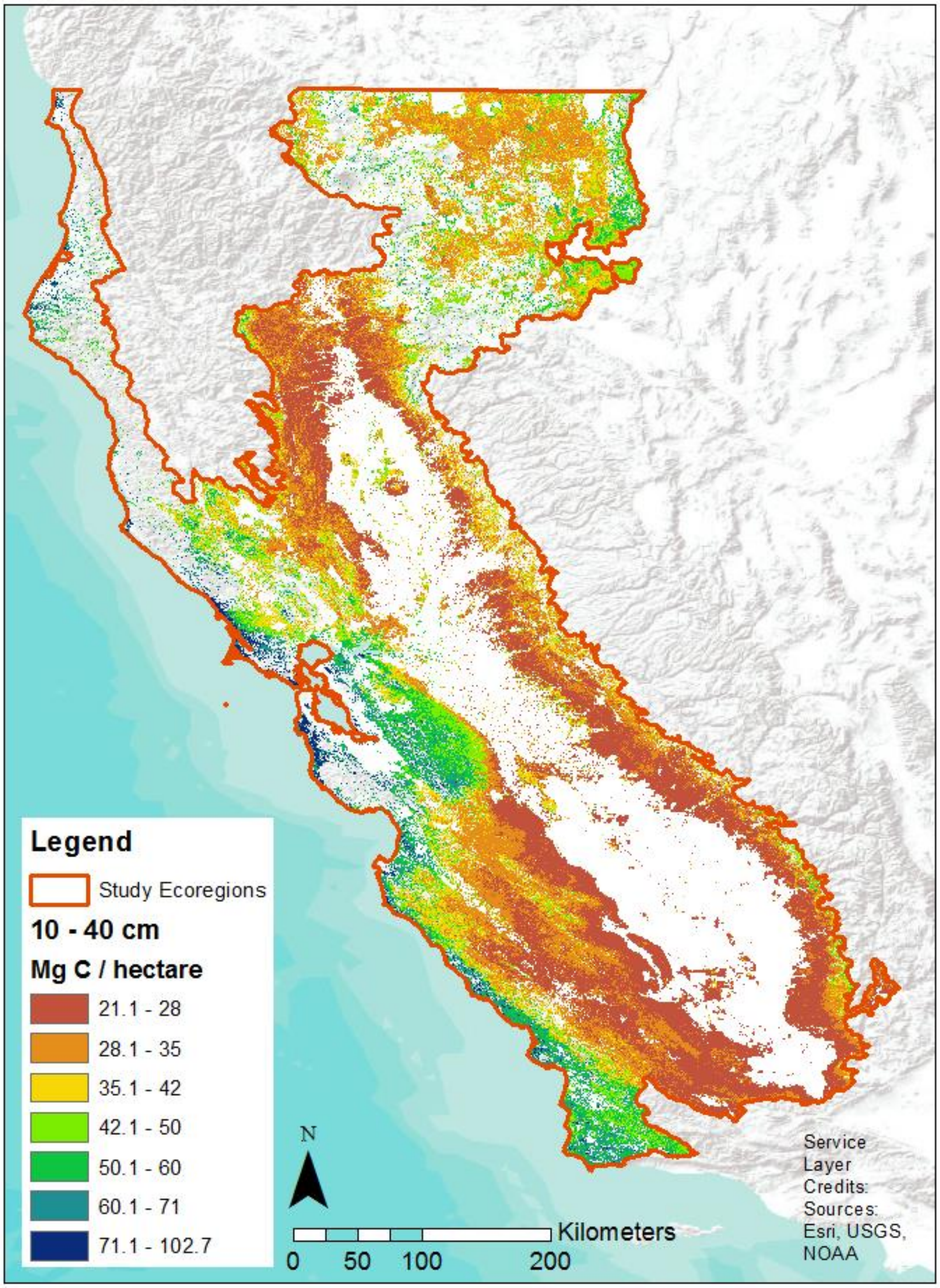


Figure 5. Predicted soil carbon stocks in California rangelands at 10 - 40 cm depth. An online version of the map is available from here https://www.arcgis.com/home/item.html?id=4ce144c080144dbbbbbc10bac9933e6b

We expected that carbon stocks at shallower depths should be highly correlated with carbon stocks at deeper depths. In agreement with our expectation, we found that observed average carbon stocks at 10 - 40 cm depth is a very good predictor of observed carbon average carbon stocks at 0 - 10 cm (R^2 = .895, $p < 0.001$, slope = 0.2020 ± 0.004, intercept = 2.300 Figure 6). Deviations from this pattern may signal locations where management or other factors are having an influence on surface carbon stocks.

Given that we found such a tight fit between the observed carbon stocks at shallower and deeper depths, we then used the mapped predictions of carbon stocks to visualize where on the landscape carbon stocks in shallower depths are deviating from expectations based on soil carbon stocks at deeper depths. We found a significant linear regression explaining predicted soil carbon at 0 - 10 cm using predicted soil carbon at 10 - 40 cm with similar parameters that we found when analyzing the observed soil carbon stocks ($p<0.001$, slope 0.252, +- 0.0001, intercept=2.867). The R^2 of the regression was 0.78 meaning that just over 20% of the variation in predicted soil carbon at 0 - 10 cm is not well explained by predicted soil carbon at 10 - 40 cm. We found that much of our study area has higher predicted soil carbon at 0 - 10 cm than we would expect based on the predicted soil carbon at 10 - 40 cm (Figure 7). However, many of the locations that have the highest predicted soil carbon at the 0 - 10 cm depth (Figure 4) also have lower predicted soil carbon than we would expect based on the predicted soil carbon at 10 - 40 cm depth (red areas in Figure 8). This aligns with research that has concluded carbon-rich soils are more sensitive to losses with drought and management than carbon-poor soils (Canarini et al. 2017, Bellamy et al. 2005). In contrast, the high predicted soil carbon at 0 - 10 cm depth in the Modoc Plateau (Figure 4) is largely

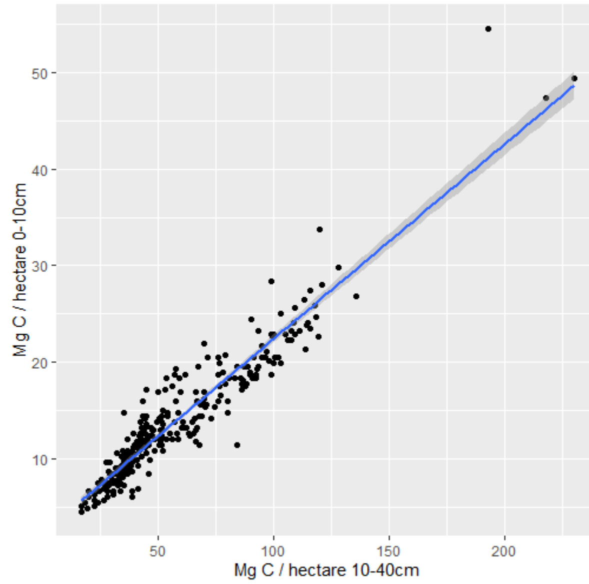


Figure 7. Linear regression of observed average soil carbon stocks at 10 - 40 cm depth vs observed average soil carbon stocks at 0 - 10 cm depth. Each point represents a field sample. The blue line is the predicted slope of the regression and the shaded area is the standard error.

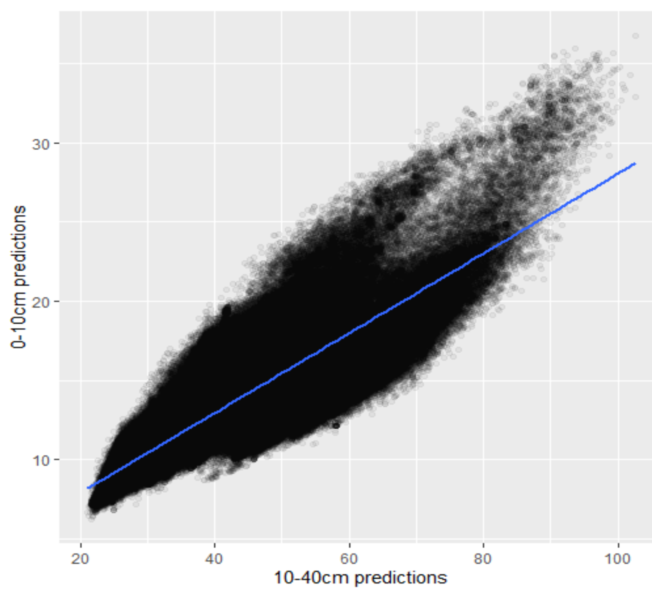


Figure 6. Scatter plot and linear regression of predicted soil carbon at 10 - 40 cm depth vs. predicted soil carbon at 0 - 10 cm depth. The blue line represents the predicted slope from the linear regression. Points above the line have positive residuals and points below the line have negative residuals. Each point is a pixel from the predicted maps in figures 4 and 5.

higher than we would expect based on predicted soil carbon at the 10 - 40 cm depth, providing nuance to our results (blue pixels Figure 8).

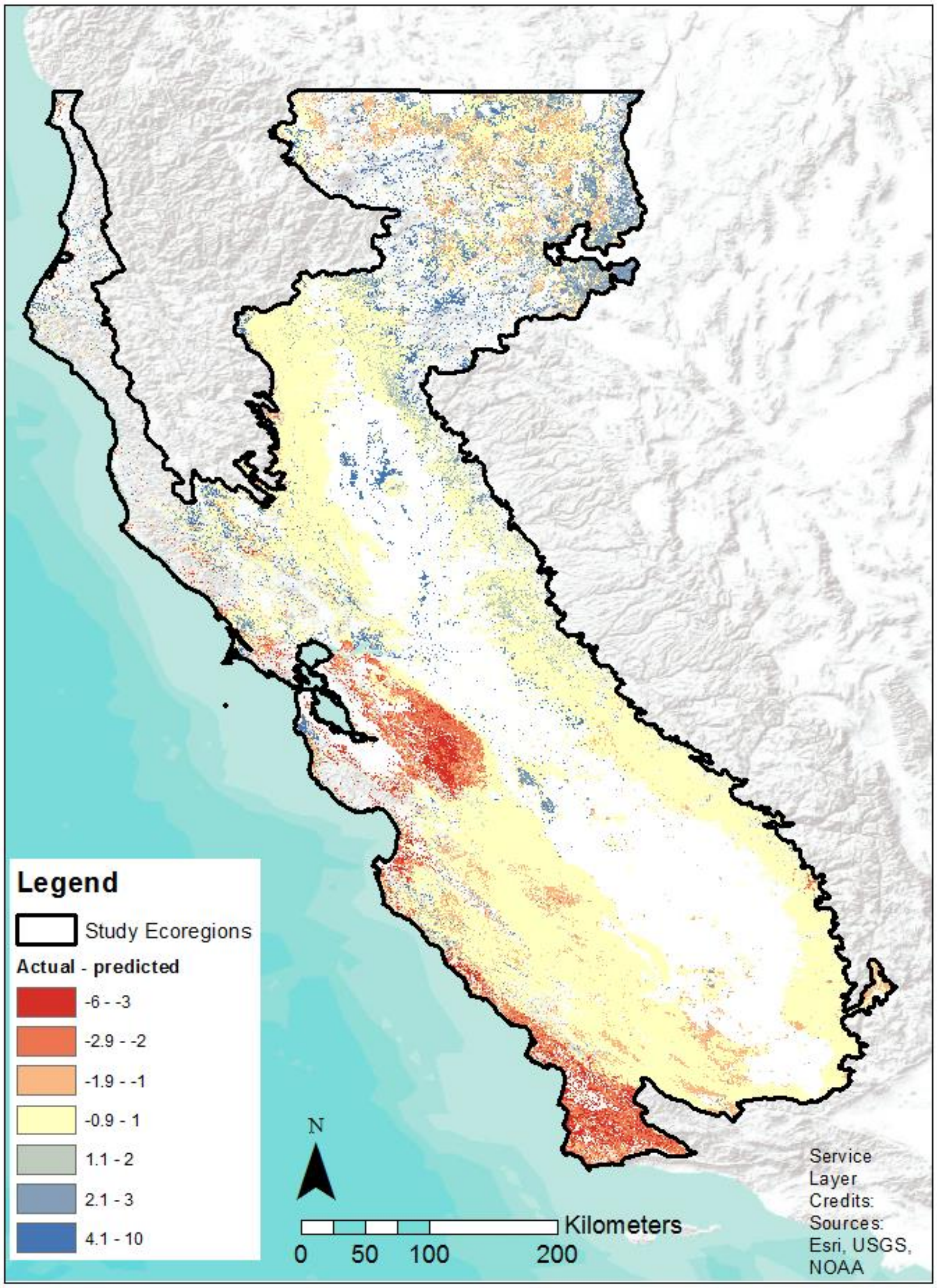


Figure 8. Map of the residuals from the linear regression in Figure 6. Red areas indicate locations where predicted soil carbon at 0 - 10 cm is lower than we would expect based on the soil carbon predicted at 10 - 40 cm. Values close to 0 are where predicted soil carbon at 0 - 10 cm is about what we would expect based on predictions at 10 - 40 cm. Blue areas are where soil carbon at 0 - 10 cm is greater than what we expect based on predicted soil carbon at 10 - 40 cm depth. Online version of the data is available here https://www.arcgis.com/home/item.html?id=bded3890324e4629b8a32c128c12af39

Mapping Change Over Time

We found relatively large rates of carbon change in our observed data. The average annual change was -0.59 Mg C/ha/y with 50% of the values between -1.38 and 0.40 Mg C/ha/y, which falls within the higher end of what is typically documented in the literature for rangeland systems (Conant et al. 2017).

Unfortunately, overall performance in modeling the change in soil carbon was extremely poor. No model explained more than 33% of the observed variation in yearly soil carbon change. Our best model had a cross validated correlation of 0.236 and a residual deviance of 67%. Moreover, models that did converge adequately were likely overfit, as cross-validated correlations averaged about 70% lower than the training correlation (0.15 vs 0.54).

In all tested methods, models of change converged too quickly (number of trees < 50) except at the lowest learning rates and tree complexities. This suggests either that there was insufficient variation in the response variable compared to our explanatory variables or that none of the chosen explanatory variables were very useful in predicting change. As outlined in the methods above, we tested several ways of filtering and transforming our response variable (change in carbon), none of which significantly changed model performance. None of our data filtering or transformations made a significant impact on modeling, showing broadly similar explanatory power compared to the unfiltered and untransformed data and also showing evidence of overfitting. Modeling change as either positive (>0) or negative (<0) was unsuccessful, with the best model having a cross-validated correlation of only 0.094. The initial binary hurdle model of change / no change fared slightly better, with our best model having a cross-validated correlation of 0.21. These also showed strong evidence of overfitting, with cross-validated correlations averaged about 70% lower than the training correlation across all models (0.58 vs 0.16).

We believe that there are two major causes for our inability to accurately predict change in soil carbon. First, the distribution of yearly carbon change was challenging. While the change was normally distributed, most points experienced little change, making it hard for the model to pick out a signal. The fact that models usually converged too quickly indicates that there was insufficient variation in the response variable, at least with regards to our predictor variables.

Second, none of our variables appear to have a large causal relationship with the change in soil carbon over a period of three to six years. Our initial exploratory modeling showed hints of this, as no covariate had a correlation with soil carbon change of > 0.15. This finding—which is borne out by looking at the correlations of individual covariates with carbon change–is surprising given we included covariates that are known to moderate carbon dynamics (Singh et al. 2018). Still, predictors such as calcium content, mineralogy (iron and aluminum oxides), and historical land use—which were not included here—have also been shown as important (Rasmussen et al. 2018). It's possible that their inclusion would have improved model performance.

Investigating patterns of changes in carbon stocks

To identify whether areas that have unexpectedly low surface carbon are also losing carbon over time, we used the residuals from the linear regression in Figure 7 to test for significant correlations with observed changes in soil carbon stocks at the shallower depth. When we used the residuals from this linear regression as a predictor of the annual change in soil carbon stocks at the 0 - 10 cm depth, we found a small but significant negative correlation ($R^2 = 0.05$, $p < 0.001$, Figure 9). While this model did not explain much of the variation in the data, it does indicate that sites with less average soil carbon at 0 - 10 cm depth than predicted by average soil carbon at 10 -40 cm depth are experiencing the highest positive annual rates of carbon change. The inverse was true at sites with more soil carbon at 0 - 10 cm than expected based on average soil carbon at 10 - 40 cm depth (Figure 9). These results extend the conclusions above, which suggest that carbon-rich soils are more susceptible to loss than relatively carbon-poor soils.

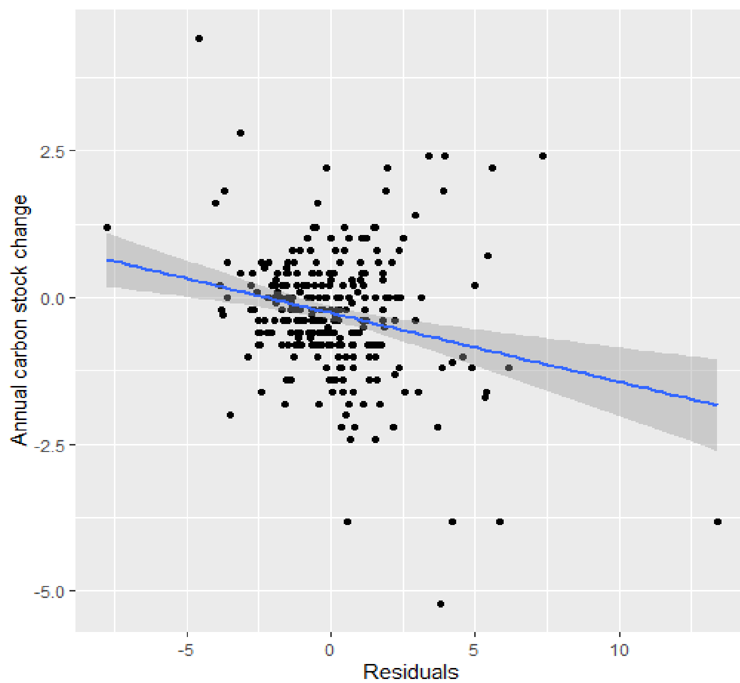


Figure 9. Linear regression of the residuals from the regression in Figure 8 vs the observed annual rate of change in carbon stocks. Each point represents a field sample. The blue line is the predicted slope of the regression and the shaded area indicates the standard error.

Conclusions

Despite being limited by the somewhat coarse resolution of spatial data, we were able to produce accurate spatial predictions of soil carbon stocks in rangelands across the state of California. As expected, we found that soil carbon levels were generally highest in cooler, wetter sites like those of the Central Coast. The five counties with the largest soil stocks were, in descending order Monterey, San Luis Obispo, Modoc, Lassen and Kern (Appendix 2). Some of our results also confirm that climate change is likely to negatively affect rates of carbon storage in rangeland soils. Indeed, our models consistently predicted lower carbon stocks in areas with more extreme seasonal temperatures and with higher rates of climatic water deficit. Although there is still high uncertainty among future climate projections related to changes in precipitation, increasing summer temperatures are ubiquitously predicted (Bedsworth et al 2018). This will act to increase climatic water deficit and dry out soils, which based on our results and others, will result in negative effects on soil carbon stocks.

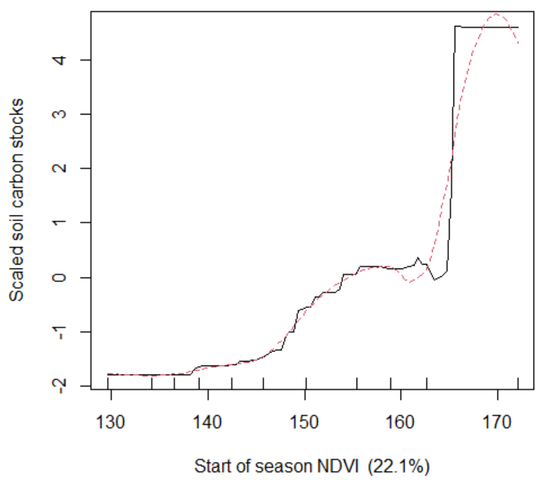
Our results also join a growing body of literature that suggest carbon-rich soils are more susceptible to losses with drought or mismanagement than their relatively carbon-poor counterparts. It stands to reason then, that actions to protect and restore soil carbon through management may be most important in these higher carbon areas. These actions should focus on managing vegetation productivity–including growing season length–through, for instance, prescribed grazing, compost applications, or range seeding; our models found that variables related to vegetation productivity were the best predictors of soil carbon stocks. Future monitoring efforts that are designed to evaluate the predictions we produce in this report will serve an important role in continuing to improve our understanding of carbon dynamics of California's rangelands across space and through time.

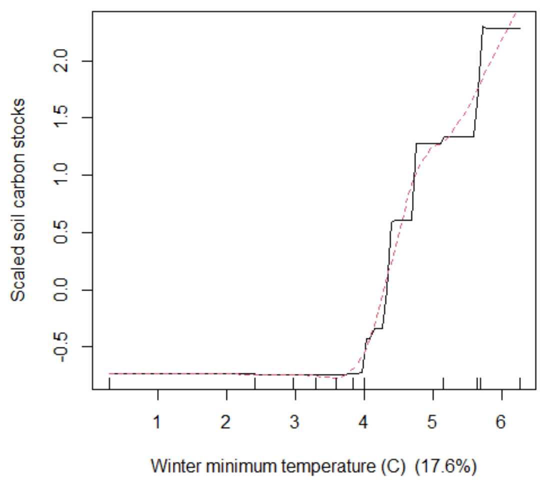
Acknowledgements

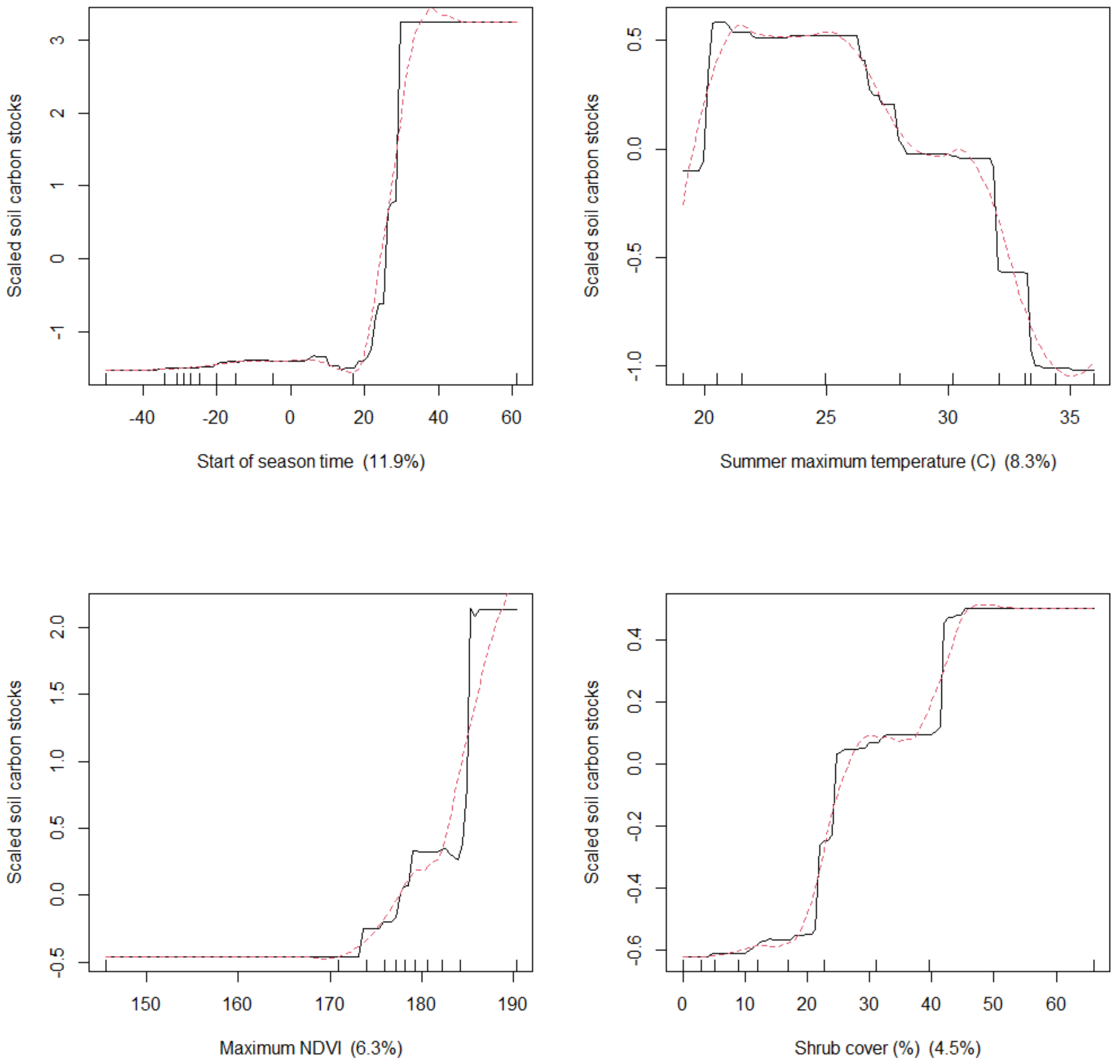
We would like to acknowledge the team of Point Blue biologists and field technicians for conducting field work and collecting much of the data that made this report possible. We would also like to thank the participating landowners and managers for allowing access to the study areas, the Natural Resources Conservation Service for their support and partnership, and the TomKat Ranch Educational Foundation for access, support, and vision while building the monitoring network. Financial support for the monitoring network has been provided by the TomKat Ranch Educational Foundation, the S.D. Bechtel Jr. Foundation, the Patterson Foundation, the Hellman Foundation, the 11th Hour Project, and three anonymous donors. Funding for the generation of this report was provided by Terraffirm and TomKat Ranch Educational Foundation.

Appendix 1. Partial Dependence Plots

Partial dependence plots of the top nine influential variables in the 0 - 10 cm model of average soil carbon. Each plot shows predicted soil carbon stocks (scaled to mean of 0) across the range of values of each explanatory variable. All other variables are held constant at their mean value to illustrate the general shape of the predicted response. The red line is a smoothed version of the prediction. Vertical ticks on the x axis indicate where there are higher densities of data points used to build the model.






Scaled soil carbon stocks
3
2
1
0
-1
-40
-20
0
20
40
60
Start of season time (11.9%)
Scaled soil carbon stocks
0.5
0.0
-0.5
-1.0
20
25
30
35
Summer maximum temperature (C) (8.3%)
Scaled soil carbon stocks
2.0
1.5
1.0
0.5
0.0
-0.5
150
160
170
180
190
Maximum NDVI (6.3%)
Scaled soil carbon stocks
0.4
0.2
0.0
-0.2
-0.4
-0.6
0
10
20
30
40
50
60
Shrub cover (%) (4.5%)

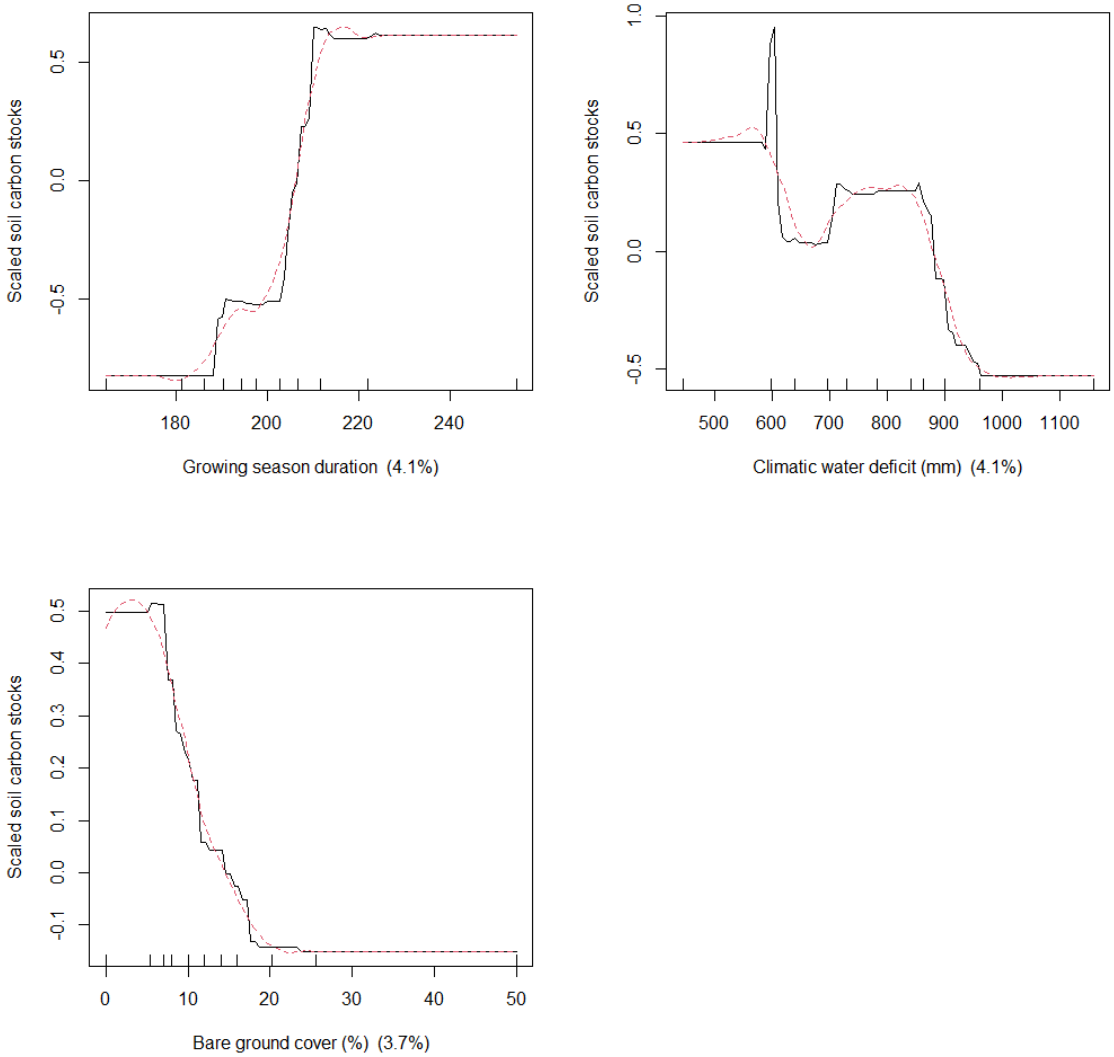
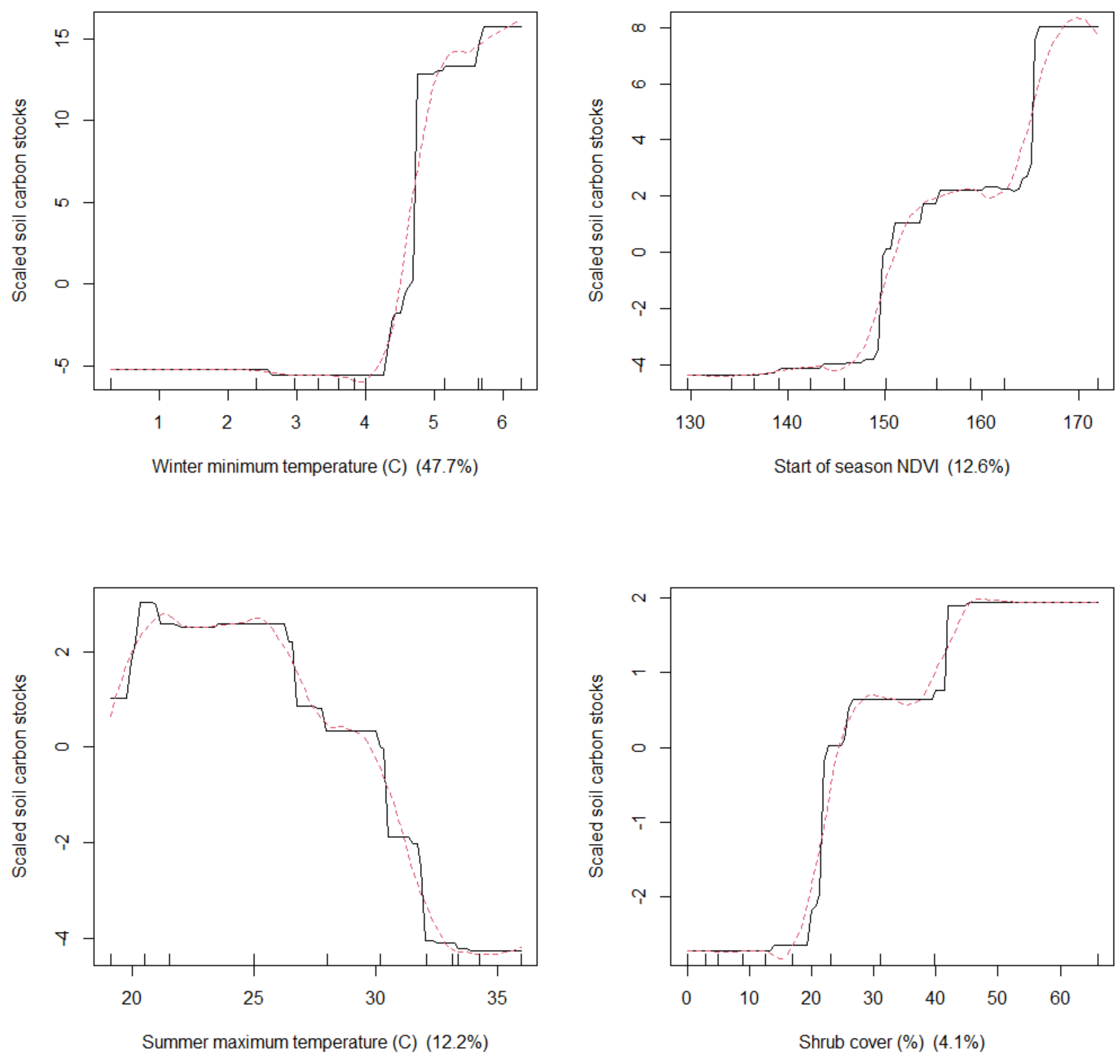
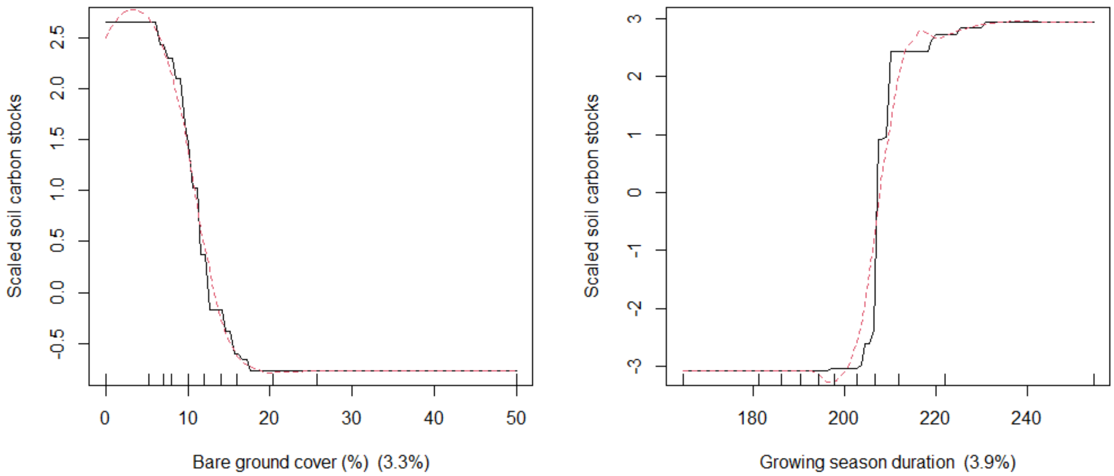
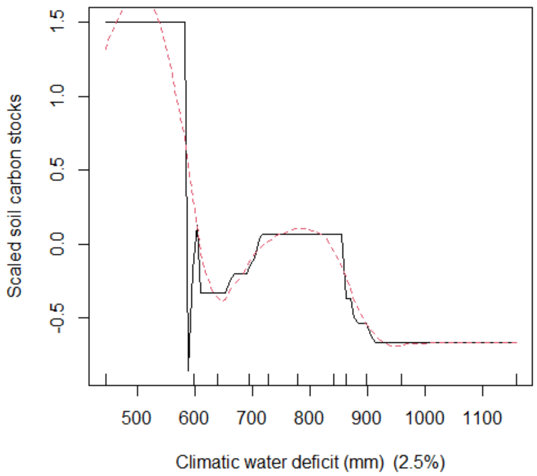
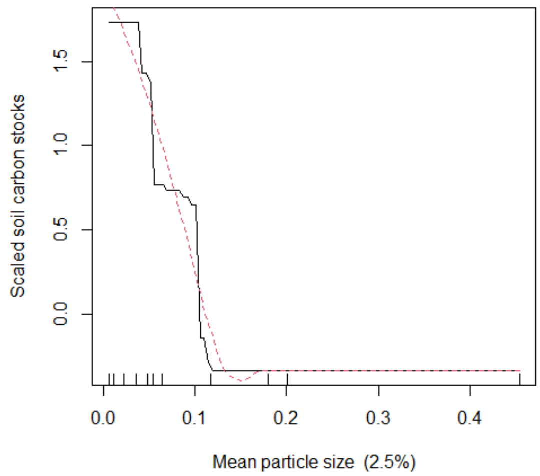

Scaled soil carbon stocks
0.5
0.0
-0.5
180
200
220
240
Growing season duration (4.1%)
Scaled soil carbon stocks
1.0
0.5
0.0
-0.5
500
600
700
800
900
1000
1100
Climatic water deficit (mm) (4.1%)
Scaled soil carbon stocks
0.5
0.4
0.3
0.2
0.1
0.0
-0.1
0
10
20
30
40
50
Bare ground cover (%) (3.7%)

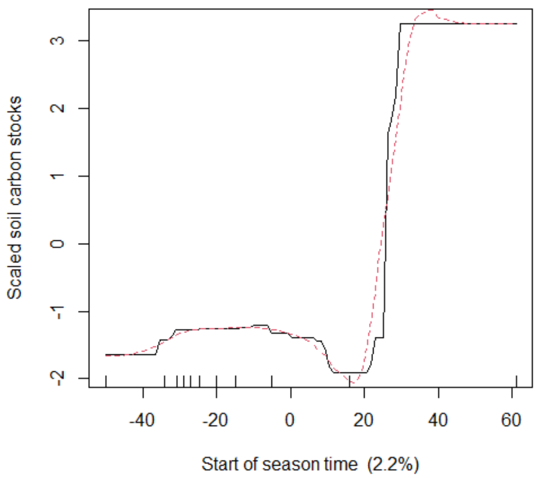
Figure A2. Partial dependence plots of the top nine influential variables in the 10 - 40 cm model of average soil carbon. Each plot shows predicted soil carbon stocks (scaled to mean of 0) across the range of values of each explanatory variable. All other variables are held constant at their mean value to illustrate the general shape of the predicted response. The red line is a smoothed version of the prediction. Vertical ticks on the x axis indicate where there are higher densities of data points used to build the model.




Scaled soil carbon stocks
-0.5
0.0
0.5
1.0
1.5
2.0
2.5
0
10
20
30
40
50
Bare ground cover (%) (3.3%)
Scaled soil carbon stocks
-3
-2
-1
0
1
2
3
180
200
220
240
Growing season duration (3.9%)


Scaled soil carbon stocks
Climatic water deficit (mm) (2.5%)


Scaled soil carbon stocks
Mean particle size (2.5%)


Scaled soil carbon stocks
Start of season time (2.2%)

Appendix 2. Estimated Carbon Stocks on Rangelands by County

County	0 - 10 cm Tg C	10 - 40 cm Tg C	County	0 - 10 cm Tg C	10 - 40 cm Tg C
Monterey	8.46	26.65	Madera	1.49	4.29
San Luis Obispo	8.13	26.31	Napa	1.44	4.21
Modoc	8.51	25.15	Butte	1.52	3.94
Lassen	6.39	18.59	Solano	1.29	3.78
Kern	5.35	15.92	Calaveras	1.33	3.58
Santa Barbara	4.05	15.29	San Mateo	0.94	3.23
Tehama	5.38	14.64	Yolo	1.17	3.22
Siskiyou	3.96	11.82	San Joaquin	1.03	3.19
Shasta	4.26	11.76	Sacramento	0.89	2.29
Fresno	3.68	10.61	Kings	0.74	2.18
San Benito	3.50	10.49	Tuolumne	0.75	2.05
Santa Clara	2.71	9.96	Amador	0.75	1.99
Sonoma	2.65	8.42	El Dorado	0.69	1.91
Stanislaus	2.36	7.66	Placer	0.68	1.88
Tulare	2.61	7.56	Yuba	0.68	1.79
Merced	2.66	7.49	Santa Cruz	0.38	1.29
Lake	2.20	6.45	Sutter	0.32	0.82
Alameda	1.46	5.53	Nevada	0.30	0.80
Humboldt	1.69	5.20	Del Norte	0.22	0.69
Mendocino	1.71	5.16	Plumas	0.25	0.68
Glenn	1.71	4.85	Ventura	0.03	0.10
Marin	1.40	4.84	San Francisco	0.00	0.01
Contra Costa	1.39	4.77	Trinity	0.00	0.00
Colusa	1.63	4.43	**Total**	**106.30**	**321.76**
Mariposa	1.52	4.31			

Table A1. Estimated total carbon stocks (Tg C) on rangelands by depth and county. Counties are in descending order based on total carbon in the 10 – 40 cm depth. Note that some counties are only partially within our study area (see Fig 1).

References

Bedsworth, Louise, Dan Cayan, Guido Franco, Leah Fisher, Sonya Ziaja. (California Governor's Office of Planning and Research, Scripps Institution of Oceanography, California Energy Commission, California Public Utilities Commission). 2018. Statewide Summary Report. California's Fourth Climate Change Assessment. Publication number: SUMCCCA4-2018-013.

Bellamy, P. H., Loveland, P. J., Bradley, R. I., Lark, R. M., & Kirk, G. J. (2005). Carbon losses from all soils across England and Wales 1978–2003. Nature, 437(7056), 245-248.

Bossio, D. A., Cook-Patton, S. C., Ellis, P. W., Fargione, J., Sanderman, J., Smith, P., ... & Griscom, B. W. (2020). The role of soil carbon in natural climate solutions. Nature Sustainability, 3(5), 391-398.

Bradford, M. A., Carey, C. J., Atwood, L., Bossio, D., Fenichel, E. P., Gennet, S., ... & Wood, S. A. (2019). Soil carbon science for policy and practice. Nature Sustainability, 2(12), 1070-1072.

California Air Resources Board (CARB) (2017).
https://ww2.arb.ca.gov/sites/default/files/classic/cc/scopingplan/scoping_plan_2017.pdf

California Department of Forestry and Fire Protections Fire and Resource Assessment Program (CALFIRE-FRAP). 2014. Compilation of "best available" land cover data for the state of California. https://map.dfg.ca.gov/metadata/ds1327.html

Carey, Chelsea J., Jacob Weverka, Ryan DiGaudio, Thomas Gardali, and Elizabeth L. Porzig. 2020. "Exploring Variability in Rangeland Soil Organic Carbon Stocks across California (USA) Using a Voluntary Monitoring Network." *Geoderma Regional* 22: e00304. https://doi.org/10.1016/j.geodrs.2020.e00304.

Canarini, A., Kiær, L. P., & Dijkstra, F. A. (2017). Soil carbon loss regulated by drought intensity and available substrate: A meta-analysis. Soil Biology and Biochemistry, 112, 90-99.

Conant, R. T., Cerri, C. E., Osborne, B. B., & Paustian, K. (2017). Grassland management impacts on soil carbon stocks: a new synthesis. Ecological Applications, 27(2), 662-668.

Delgado-Baquerizo, M., Karunaratne, S.B., Trivedi, P., Singh, B. K. 2018. Chapter 5 - Climate, Geography, and Soil Abiotic Properties as Modulators of Soil Carbon Storage. Singh, B.K. (Editor). Soil Carbon Storage, Academic Press, Pages 137-165, ISBN 9780128127667, https://doi.org/10.1016/B978-0-12-812766-7.00005-6.

Deng, L., Peng, C., Kim, D. G., Li, J., Liu, Y., Hai, X., ... & Kuzyakov, Y. (2021). Drought effects on soil carbon and nitrogen dynamics in global natural ecosystems. Earth-Science Reviews, 214, 103501.

Dewitz, J., 2019, National Land Cover Database (NLCD) 2016 Products: U.S. Geological Survey data release, https://doi.org/10.5066/P96HHBIE.

Elith, J., J. R. Leathwick, and T. Hastie. 2008. "A Working Guide to Boosted Regression Trees." *Journal of Animal Ecology* 77 (4): 802–13. https://doi.org/10.1111/j.1365-2656.2008.01390.x.

Farr, T. G. et al., 2007, The Shuttle Radar Topography Mission, Rev. Geophys., 45, RG2004, doi:10.1029/2005RG000183

Flint, L.E. and Flint, A.L. 2017. California Basin Characterization Model: A Dataset of Historical and Future Hydrologic Response to Climate Change. Version 1.1. U.S. Geological Survey Data Release, https://doi.org/10.5066/F76T0JPB

Flint, L.E., Flint, A.L., and Stern, M.A., 2021, The Basin Characterization Model - A regional water balance software package (BCMv8) data release and model archive for hydrologic California, water years 1896-2020, U.S. Geological Survey data release, https://doi.org/10.5066/P9PT36UI.

Gomes, Lucas Carvalho, Raiza Moniz Faria, Eliana de Souza, Gustavo Vieira Veloso, Carlos Ernesto G.R. Schaefer, and Elpídio Inácio Fernandes Filho. 2019. Modeling and Mapping Soil Organic Carbon Stocks in Brazil. *Geoderma* 340 (November 2018): 337–50. https://doi.org/10.1016/j.geoderma.2019.01.007.

Greenwell, B., Boehmke, B.,Cunningham, J. and GBM Developers (2020). gbm: Generalized Boosted Regression Models. R package version 2.1.8. https://CRAN.R-project.org/package=gbm

Griffith, G.E., Omernik, J.M., Smith, D.W., Cook, T.D., Tallyn, E., Moseley, K., and Johnson, C.B., 2016, Ecoregions of California (poster): U.S. Geological Survey Open-File Report 2016–1021, with map, scale 1:1,100,000, http://dx.doi.org/10.3133/ofr20161021.

Hijmans, R. (2022). raster: Geographic Data Analysis and Modeling. R package version 3.5-15. https://CRAN.R-project.org/package=raster

Hijmans, R. J., Phillips, S., Leathwick, J. and Elith, J. (2022). dismo: Species Distribution Modeling. R package version 1.3-5. https://CRAN.R-project.org/package=dismo

Keskin, Hamza, Sabine Grunwald, and Willie G. Harris. 2019. "Digital Mapping of Soil Carbon Fractions with Machine Learning." *Geoderma* 339 (November): 40–58. https://doi.org/10.1016/j.geoderma.2018.12.037.

Meier, G.A. & Jesslyn F. Brown. 2014. Remote sensing of land surface phenology. USGS Numbered Series 2014-3052. https://doi.org/10.3133/fs20143052

Pebesma, E.J., R.S. Bivand, 2005. Classes and methods for spatial data in R. R News 5 (2), https://cran.r-project.org/doc/Rnews/.

Porzig, E., N.E. Seavy, R. T. DiGaudio, C. Henneman, and T. Gardali. 2018. The Rangeland Monitoring Network Handbook V2.1. Point Blue Conservation Science, Petaluma, California.

Rasmussen, C., Heckman, K., Wieder, W. R., Keiluweit, M., Lawrence, C. R., Berhe, A. A., ... & Wagai, R. (2018). Beyond clay: towards an improved set of variables for predicting soil organic matter content. Biogeochemistry, 137(3), 297-306.

Rigge, Matthew B., Homer, Collin G., Cleeves, L., Meyer, Debbie K., Bunde, Brett, Shi, Hua, Xian, George, Schell, S., Bobo, M., 2020. Quantifying western U.S. rangelands as fractional components with multi-resolution remote sensing and in situ data: Remote Sensing, v. 12, no. 3.

Sanderman, J., Hengl, T., & Fiske, G. J. (2017). Soil carbon debt of 12,000 years of human land use. Proceedings of the National Academy of Sciences, 114(36),

Silver , W L, Vergara, S E, Mayer, A, 2018. Carbon sequestration and greenhouse gas mitigation potential of composting and soil amendments on California's rangelands. California Natural Resources Agency, p. 62.

Silver, W L, Ryals, R, Eviner, V, 2010. Soil carbon pools in California's annual grassland ecosystems. Rangel. Ecol. Manag. 63, 128–136. doi:10.2111/REM-D-09-00106.1.

Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture. Soil Survey Geographic (SSURGO) Database for California. Available online at https://sdmdataaccess.sc.egov.usda.gov. Accessed 2022-04-01.

Veloz, S., Eliott, N. Carey, C. Porzig. L. 2021. Pescadero Soil Carbon Mapping: Technical Report. Point Blue Conservation Science (Contribution No. 2358). Petaluma, CA.

Ward, S. E., Smart, S. M., Quirk, H., Tallowin, J. R., Mortimer, S. R., Shiel, R. S., ... & Bardgett, R. D. (2016). Legacy effects of grassland management on soil carbon to depth. Global change biology, 22(8), 2929-2938.

Wickham, H., François, R., Henry, L. and Müller, K. (2022). dplyr: A Grammar of Data Manipulation. R package version 1.0.9. https://CRAN.R-project.org/package=dplyr