Title
Climate-Driven Limits to Future Carbon Storage in California's Wildland Ecosystems

Permalink
https://escholarship.org/uc/item/77z1c10d

Journal
AGU Advances, 2(3)

ISSN
2576-604X

Authors
Coffield, Shane R
Hemes, Kyle S
Koven, Charles D
et al.

Publication Date
2021-09-01

DOI
10.1029/2021av000384

Peer reviewed
Climate-Driven Limits to Future Carbon Storage in California's Wildland Ecosystems

Shane R. Coffield1, Kyle S. Hemes2, Charles D. Koven1, Michael L. Goulden1,4, and James T. Randerson1

1Department of Earth System Science, University of California, Irvine, CA, USA, 2Stanford Woods Institute for the Environment, Stanford University, Stanford, CA, USA, 3Climate and Ecosystem Sciences Division, Lawrence Berkeley National Lab, Berkeley, CA, USA, 4Department of Ecology and Evolutionary Biology, University of California, Irvine, CA, USA

Abstract Enhanced ecosystem carbon storage is a key component of many climate mitigation pathways. The State of California has set an ambitious goal of carbon neutrality by 2045, relying in part on enhanced carbon sequestration in natural and working lands. We used statistical modeling, including random forest and climate analog approaches, to explore the climate-driven challenges and uncertainties associated with the goal of long-term carbon sequestration in forests and shrublands. We found that seasonal patterns of temperature and precipitation are strong controllers of the spatial distribution of aboveground live carbon. RCP8.5 projections of temperature and precipitation are estimated to drive decreases of 16.1% ± 7.5% in aboveground live carbon by the end of the century, with coastal areas of central and northern California and low/mid-elevation mountain areas being most vulnerable. With RCP4.5 projections, declines are less severe, with 8.8% ± 5.3% carbon loss. In either scenario, increases in temperature systematically cause biomass declines, and the spread of projected precipitation across 32 CMIP5 models contributes to substantial uncertainty in the magnitude of that decline. Projected changes in the environmental niche for the 20 most biomass-dominant tree species revealed widespread replacement of conifers by oak species in low elevation regions of central and northern California, with a corresponding decline in carbon storage depending on expected migration rates. The spatial patterns of vulnerability we identify may allow policymakers to assess where carbon sequestration in aboveground biomass is an appropriate part of a climate mitigation portfolio, and where future climate-driven carbon losses may be a liability.

Plain Language Summary Many climate mitigation policies, including those of California, rely in part on increased carbon uptake by forests and shrublands. However, these natural ecosystems are also being impacted by climate change, likely making the goal of increasing biomass carbon more difficult to achieve. In this study, we used a variety of statistical models to estimate the impact of rising temperature and changing precipitation on California ecosystems' carbon storage. We found that in either moderate or severe warming scenarios, aboveground live carbon will decrease substantially. Decreases are driven by the rising temperature, while uncertainty in future precipitation leads to substantial uncertainty in the exact magnitude of those decreases. We also modeled several different tree species separately, finding that climate change will likely favor oak species at the expense of conifers. Lastly, some areas of California appear more vulnerable to carbon loss than others—in particular, the northern and central coasts, low/mid-elevation mountain areas, and places where there are currently forest carbon offset projects. The spatially explicit projections we provide may help with the design of land management and climate policies to anticipate the impacts of climate change, and focus carbon offsets and conservation efforts where they will be most effective.

1. Introduction

Terrestrial ecosystems are currently large carbon sinks, sequestering approximately 30% of anthropogenic emissions globally over 1850–2018 (Friedlingstein et al., 2019). Their past and present ability to sequester carbon, as well as the many other ecosystem services they provide, make “natural climate solutions” an appealing class of climate mitigation strategies (Anderegg et al., 2020; Griscom et al., 2017). In fact, enhanced...
ecosystem carbon storage in forests is a key component of many climate mitigation pathways that keep global temperature rise below 1.5°C or 2°C (Roe et al., 2019).

A prime example of using terrestrial ecosystems toward natural climate solutions can be found in California, home to one of the most ambitious climate change mitigation policies globally. The state’s Natural and Working Lands Climate Change Implementation Plan (California Air Resources Board, 2019) seeks to contribute carbon dioxide removal consistent with the statewide goal of carbon neutrality by 2045. The plan involves reversing the net land carbon flux, which is currently positive (i.e., a source of carbon) (California Air Resources Board, 2019; Sleeter et al., 2019), such that the land will sequester an additional 23 MtC by 2045 and 230 MtC by 2100. For reference, 230 MtC corresponds to approximately 4.2% of the estimated current total ecosystem carbon stock of 5,500 MtC in California (California Air Resources Board, 2019).

Despite this reliance on forests for climate change mitigation, there is considerable uncertainty regarding the future ability of forests to take up and store carbon due to changing temperature and precipitation regimes, disturbance, and other indirect climate change feedbacks (Anderegg et al., 2020; Sperry et al., 2019). Many of these climate-related threats are already apparent in observations. For example, over the twentieth century, as a result of fire suppression management as well as climatic shifts, California forests generally became denser, with smaller trees, less biomass, and an increase in the dominance of oaks relative to pines (McIntyre et al., 2015). More rapid climate change and intense drought stress in this century have caused shifts in plant communities, including widespread mortality (Anderegg et al., 2013; Breshears et al., 2005; Goulden & Bales, 2019), range contractions (Kelly & Goulden, 2008), and shifts in hydraulic trait composition (Trugman et al., 2020). It is estimated that California’s 2012–2015 drought killed 41% (Stovall et al., 2019) to 49% (Fettig et al., 2019) of trees in the central and southern Sierra Nevada, disproportionally ponderosa pines and larger trees at lower elevations. The result has been a shift in forest composition and redistribution of major species. These direct climate-driven changes, along with the effects of land management and increasing severity of wildfires, have caused California’s total terrestrial carbon stocks to decrease (California Air Resources Board, 2019; Fellows & Goulden, 2008) and they pose continued risks to carbon storage into the future (Anderegg et al., 2020; Galik & Jackson, 2009; Lalonde et al., 2018; McDowell et al., 2020). Moreover, future climate change-driven shifts to carbon storage capacity have direct implications for the long-term success of current carbon sequestration projects. Strategies that assume the carbon carrying capacity will remain static across the landscape risk sequestering carbon into vulnerable ecosystems that may undergo a transition to a lower carbon state. Likewise, these strategies may miss opportunities to accelerate storage into locations where the carbon carrying capacity will become more favorable in the future.

Future ecosystem projections based on climate broadly fall into two categories: statistical models and dynamical (or processed-based) models. Statistical models often leverage the tight spatial relationships between climate and vegetation, which are typically described by bioclimate schemes (Holdridge, 1947; Whittaker, 1975). These spatial relationships can be extrapolated temporally to model past or future vegetation. An early example of this is by Prentice and Fung, who applied a bioclimate scheme to estimate vegetation biomass globally during the last glacial maximum, assuming steady state, i.e., that vegetation is in equilibrium with climate (Prentice & Fung, 1990). More recent examples often use machine learning methods such as random forests (RF) (Gómez-Pineda et al., 2020; Iverson et al., 2004; Prasad et al., 2006; Rehfeldt et al., 2012; Rogers et al., 2017) or the Maxent model (Phillips et al., 2006) to capture climate niches and make projections based on future climate (e.g., Loarie et al., 2008). Decision tree-based statistical methods including random forests can be useful in ecological modeling by uncovering hidden structures in the data and outperforming simpler regression techniques, especially at larger geographic scales (Prasad et al., 2006). Decision trees assume no underlying relationship between response and predictor variables (linear, quadratic, etc.) but instead construct decision rules, which optimally parse and partition the data based on predictor variables. Additionally, techniques like cross-validation and pruning can be used to find optimal tree size and avoid overfitting. Ensemble methods based on randomized collections of trees, that is, random forests, further protect against overfitting and bias by randomly subsetting the out-of-sample test data and candidate variables across individual trees (Breiman, 2001). In general, these methods allow easier interpretation and visualization than more complex or deep learning methods, allowing insight into the key predictors and underlying relationships.
Another statistical approach relevant for ecological forecasting is the calculation of climate analogs, which involves connecting present and future climates by nearest distance in multidimensional climate space (Koven, 2013; Mahony et al., 2017; Williams et al., 2007). This approach allows for identification of particularly novel future climates, and the geographic distance between analogs can inform whether species' migration may be able to keep pace with expected climate change. Areas of highly novel future climates require particularly large levels of extrapolation for statistical niche models, and may indicate locations where process-based model approaches need to be prioritized.

Process-based models offer some advantages over statistical ecological niche models because they are able to represent dynamic processes such as establishment and mortality, competitive interactions, wildfire, effects of carbon dioxide on water use, and climate change impacts on net primary productivity and decomposition (Fisher et al., 2018). Statistical models have received criticism for not representing these processes explicitly (Hampe, 2004; Jackson et al., 2009), but in many cases statistical models have been shown to perform similarly to (or better than) process-based models (Hijmans & Graham, 2006; Kearney et al., 2010; Keenan et al., 2011; Morin & Thuiller, 2009). Even dynamical models have been criticized for the credibility of their representation of complex ecological processes, especially due to limited quantitative understanding of the factors that control species range limits, competition, dispersal, migration, and the long-term physiological impacts of rising CO₂ (Bachelet et al., 2008; Neilson et al., 2005; Rehfeldt et al., 2012). Statistical approaches informed by species abundance observations have the potential to capture the combined interactions of drought and fire that might be contained in the structure of vegetation. They can also constrain some aspects of other processes, for example, by considering migration rates, which to our knowledge have not yet been rigorously integrated into process-based models.

Regardless of approach, previous research has disagreed about the direction and magnitude of terrestrial carbon change in response to future climate in the Western United States (Foster et al., 2019; Lenihan et al., 2003; Rogers et al., 2011), though there appears to be a growing consensus of high vulnerability especially in California (Lenihan et al., 2008a; Loarie et al., 2008; Sleeter et al., 2019; Thorne et al., 2017). What makes this area of research especially challenging is the many possible trajectories of, and interactions between, land management (Cameron et al., 2017; Thorne et al., 2017), wildfires (Westerling & Bryant, 2007), climate scenarios and precipitation change (Thorne et al., 2017), impacts of biotic agents on tree mortality (Stephenson et al., 2019; Trugman et al., 2021), and migration potential (Higgins et al., 2003; Rogers et al., 2017).

In this study, we use a variety of eco-statistical approaches to project end-of-century aboveground live (AGL) carbon storage in California’s wildland ecosystems, isolating the impact of climate change from other global change drivers. Our work builds upon multiple previous studies quantifying climate-driven vulnerability, and through our statistical approaches we offer a more comprehensive analysis of uncertainty arising from different dimensions—namely (a) eco-statistical approach, (b) climate scenarios, (c) climate models, and (d) tree migration rates. The fourth component, migration, is often not accounted for in the current generation of process-based models, and is typically only accounted for in statistical models in the simplest of terms (assuming either no migration or unlimited migration). The uncertainties associated with future wildland carbon distributions and the spatial patterns of vulnerability that we quantify may allow policymakers to identify where multi-decadal carbon sequestration in aboveground biomass is an appropriate part of a climate mitigation portfolio, and where it may be a liability in a future climate.

2. Methods

2.1. Data

2.1.1. Climate Data

We obtained the downscaled modeled climate data for 2006–2099 from the Bias-Corrected Spatially Downscaled (BCSD) CMIP5 Climate Projections data set (Brekke et al., 2013; Maurer et al., 2007). This 8° resolution data set was chosen for its monthly temporal resolution and inclusion of all 32 CMIP5 models for RCP4.5 and RCP8.5. We used the mean daily precipitation and temperature variables and averaged them each to four seasons, giving us eight input variables to our models. With all models, we used 2006–2015 average as “present” and 2090–2099 average as “future.” Because there is substantial variability in precipitation
change across the models (Figure S1), we also grouped the 32 models into three moisture availability scenarios: "dry" (average of the eight models showing the greatest precipitation decrease for California), "mean" (of all 32 models), and "wet" (average of the eight models showing the greatest precipitation increase for California). On average for California, these moisture response scenarios correspond to a precipitation decrease of 95 mm/y (−16%) for the dry scenario, an increase of 50 mm/y (+9%) for the mean of all models, and an increase of 227 mm/y (+39%) for the wet scenario with RCP8.5 (Figure S2).

2.1.2. Vegetation and Carbon Data

We incorporated data for vegetation or carbon from several different data sets, depending on what was most appropriate for each eco-statistical approach (Table 1). Our main quantity of interest throughout the study was aboveground wildland carbon density, which we obtained upon request from the California Air Resources Board. This data set was available for California at 30 m for 2014 (California Air Resources Board, 2018; Gonzalez et al., 2015) and is a direct extension of the data set described in Gonzalez et al., 2015 for years 2001 and 2010. For one of our approaches, we also first partitioned California's wildlands into two groups: forest versus shrubland or grassland, based on 30 m land cover data from the US Geological Survey National Land Cover Database (Homer et al., 2020); CA Air Resources Board aboveground wildland carbon density (California Air Resources Board, 2019). For another approach where we searched for present-day analogs to future climates, we needed carbon data that extended beyond California, including the Western United States and Mexico. For this larger domain we used a global product for 100 m aboveground biomass from the European Space Agency's Climate Change Initiative, available for the year 2017 (Santoro & Cartus, 2019). We scaled these biomass data by 47% to represent carbon, following common practice (Gonzalez et al., 2015). Finally, in our last approach, we modeled carbon densities separately for different tree species. Those species-level data were obtained from Oregon State University Landscape Ecology, Modeling, Mapping & Analysis (LEMMA), available at 30 m for 2012 (Kennedy et al., 2018). The LEMMA data set is based on a nearest neighbors approach, matching all pixels to their most similar inventory plot in terms of spectral and environmental characteristics. We considered the top 39 species in California by biomass, which account for 99% of aboveground forest biomass.

2.1.3. Data Processing

In all cases, our study area was the wildland areas of California at an eighth-degree resolution, matching the resolution of the climate data set. We excluded any 1/8" pixels that were less than 50% wildland land cover for the purpose of our analysis. Here we considered "wildland" as forest, shrub, grass, or barren cover and excluded urban, agriculture, or water cover as classified by the NLCD. The remaining data set contained 2258 pixels (approximately 345,000 km² or 81% of California). For the pixels that were kept in the analysis,

<table>
<thead>
<tr>
<th>Eco-statistical approach</th>
<th>Description</th>
<th>Data sources</th>
</tr>
</thead>
<tbody>
<tr>
<td>1. RF regression of carbon density</td>
<td>Random forest regression to project future carbon density based on seasonal climate predictors</td>
<td>CA Air Resources Board aboveground wildland carbon density (California Air Resources Board, 2018; Gonzalez et al., 2015)</td>
</tr>
<tr>
<td>2. RF classification of dominant vegetation type</td>
<td>Random forest classification to project future vegetation type (forest or shrubland/grassland) based on seasonal climate predictors; translated to carbon based on ecoregion averages</td>
<td>USGS National Land Cover Database (Homer et al., 2020); CA Air Resources Board aboveground wildland carbon density (California Air Resources Board, 2019)</td>
</tr>
<tr>
<td>3. Climate analogs</td>
<td>Assigned future carbon density equal to the carbon density from the location of the most similar climate in the present</td>
<td>ESA Climate Change Initiative biomass (Santoro &amp; Cartus, 2019)</td>
</tr>
<tr>
<td>4. Tree species niche models</td>
<td>Random forest regression to project future carbon density for each of 39 species based on seasonal climate predictors</td>
<td>LEMMA species-level biomass (Kennedy et al., 2018)</td>
</tr>
</tbody>
</table>

Notes. Each approach is fundamentally distinct, requiring different data sources based on the spatial domain and quantities of interest. In the third and fourth approaches, we also explored dispersal and establishment as factors limiting future carbon densities. In every approach, we compared results for RCP4.5 versus RCP8.5 and for dry versus wet climate models.

we also kept track of the valid subpixel land cover fraction, which we used to scale back our model estimates for total carbon. For those calculations, we also excluded barren areas, effectively preventing our models from adding biomass carbon to rock-covered areas at high elevations in the Sierra Nevada mountains (or to other areas, such as deserts, which we assumed will not support biomass in the future).

As a supplementary analysis, we also explored the sensitivity of our eighth-degree carbon and land cover data sets to disturbance history (including fire and harvest) in California. We used fire and harvest polygons from the California Department of Forestry and Fire Protection (CALFIRE, 2021; FRAP, 2019) for 1995–2014 to filter our data at a 30-m resolution before averaging to the final eighth-degree resolution. The purpose of the analysis was to quantify whether our training data, and the resulting model projections described below, would be substantially different if we excluded young, recently disturbed forests stands, which could have low carbon density or be classified as grass or shrublands in the NLCD map.

2.2. Eco-Statistical Approaches

2.2.1. RF Regression of Carbon Density

In our first and simplest eco-statistical modeling approach, we fit random forest regression models to estimate the present spatial distribution of aboveground live carbon density as a function of eight predictor climate variables: four seasons of temperature and precipitation averaged for 2006–2015. Random forest models were developed using the scikit-learn machine learning package in Python (Pedregosa et al., 2011) and validated with tenfold cross validation. We used the default forest size of 100 decision trees and chose a maximum number of 25 leaf nodes, which optimized outgroup performance as measured by root mean square error (RMSE). We then fit a single random forest model to all 2258 data points, explored error structure and variable importance, and applied the model to the 2090s climate data. We report a total percent change based on the difference between the sum of modeled present and future AGL carbon density. We repeated this analysis (and all others below) to compare RCP4.5 versus RCP8.5 and dry versus wet models.

For this RF approach to modeling aboveground live carbon density, we also added an analysis of the contributions of temperature versus precipitation to carbon change under RCP8.5. We compared the spread of total projected biomass change across the 32 CMIP5 models when (a) temperature changes but precipitation is held constant, (b) precipitation changes but temperature is held constant, and (c) both temperature and precipitation change.

2.2.2. RF Classification of Dominant Vegetation Type

In our second approach, we chose a categorial variable of dominant vegetation type (namely, forest or shrub) as our target variable and repeated the methodology of the previous approach, with random forest classification models in place of regression models. The dominant vegetation type came from the NLCD, where deciduous, evergreen, and mixed forests were grouped together as “forest,” and shrub/scrub and grassland/herbaceous were grouped together as “shrub.” Instead of RMSE, we considered classification accuracy (number of correct classifications as a fraction of total number of classifications) as our performance metric. To estimate total carbon change from this approach, we applied the mean carbon density across the forest or shrub pixels in the corresponding ecoregion from the present. For this averaging step, we used the Level III Ecoregions as defined by the EPA, helping to account for the different carbon densities of different forest regions in California.

2.2.3. Climate Analogs

Our third approach leveraged the concept of climate analogs, first introduced by Williams et al., 2007 and then revised by Mahony et al., 2017, to project changes in carbon. The main idea is to find, for every pixel under 2090s climate, the most similar (“analog”) pixel from the present climate. In the original Williams et al., 2007 approach, the distances between future and present climates were expressed as a standardized Euclidean distance (SED) in climate space (in our case, an eight-dimensional space of our eight variables). A more statistically robust metric presented by Mahony et al., 2017 is the Mahalanobis distance, which also accounts for the number of dimensions and correlation between variables. The Mahalanobis distance $D_{ij}$ between the future climate of a focal point, $j$, and another point, $i$, in the present, is described as
\[ D_{ji} = \sqrt{D_{ji}^2} = \sqrt{\left( b'_j - a'_i \right)^T \left[ R_j \right]^{-1} \left( b'_j - a'_i \right)} \] (1)

where \( a' \) is a row vector of present climate data averages (in our case, of length 8), and \( b' \) is a row matrix of future climate data averages. Both \( a' \) and \( b'_j \) are normalized by the interannual climatic variability of the present climate at location \( j \). \( \left[ R_j \right] \) is the correlation matrix of the eight climate variables at \( j \) in the present, calculated across 10 years of data. We then assigned a carbon density to each pixel in the future equal to that of its best present analog indicated by the minimum \( D_{ji} \).

In using the future climate as the reference, the minimum Mahalanobis distance represents the novelty of the future climate at a given point. This novelty can then be interpreted in a more meaningful way as a “sigma dissimilarity,” that is, a multivariate z-score from a chi-square distribution with eight degrees of freedom. The sigma dissimilarity represents the departure of the future climate from historical variability (Mahony et al., 2017).

The climate analog analyses also allow determination of a climate analog velocity as the distance between the reference point and the geographical point that minimizes Mahalanobis distance divided by the time interval between the means of the two periods considered. By enlarging or decreasing the search area for the two points, a climate velocity limit can be applied to the metric, to capture dispersal limitations to ecosystem change. We thus explored the sensitivity of this approach to a maximum climate velocity by varying the search area over which a potential analog could be found for a given pixel. We calculated and compared carbon change for three different search areas: within 100 km of a given pixel, within 500 km of a given pixel, or within the entire domain of the climate data set (United States and Mexico, north of 25°). These different search areas roughly represent different magnitudes of dispersal limitation, the first being the most restrictive. The third and broadest domain is the most permissive, allowing for California ecosystems to reassemble and resemble ecosystems anywhere in the United States or northern Mexico if climatically favorable. While perhaps unrealistic, we include this third, unrestricted scenario in our final results as an end-member for comparison against the more restricted scenarios.

2.2.4. Tree Species Niche Models

In our fourth and final approach, we developed random forest regression models separately for the AGL carbon of each of 39 tree species, accounting for 99% of aboveground live forest carbon. Our RF regression models followed the training and testing methodology outlined in the first approach, RF regression of carbon density. As with the climate analogs, we explored a few “sub-approaches” to test the model sensitivity and the equilibrium assumption. In the first “equilibrium” sub-approach, we added together 20 different models for 20 species which account for 94% of AGL forest carbon (species 21–39 were ultimately excluded due to poor model performance). We also verified whether adding together 20 separate models led to any projections of carbon that were higher than anywhere observed in the present, in which case we might need to consider competition. In the second sub-approach, we grouped the 39 various species together into functional types, modeling conifer versus hardwood species (see Table S1 for full details on the species and their groupings). These first two sub-approaches assume equilibrium with future climate, that is, that the tree species are given infinite time to migrate and fully establish. In the final two sub-approaches, we added a consideration of migration—a fast (500 m/y) and slow (50 m/y) scenario. These migration rates were chosen based on previous studies which estimate rates of tree dispersal and establishment on similar orders of magnitude (Davis, 1983; Higgins et al., 2003; Huntley, 1991; Settele et al., 2015; Solomon, 1997).

For simplicity, each migration rate provided a threshold where, for each future pixel, we forced a given species’ biomass to zero if there was no present-day presence within a distance of (migration rate) \( \times 85 \) years. For reference, those distances are 43 km for the fast scenario, and 4.3 km (effectively one 1/8° pixel in any direction) for the slow scenario. These simple calculations are intended to provide a first-order estimation of the magnitude of variation arising from a tenfold increase in migration capacity in comparison to infinite migration capacity. As a whole, the comparison across these sub-approaches allows us to highlight specific vulnerable species/groups and to quantify the impact that management such as assisted migration could have in increasing California’s total carbon storage.
3. Results

We developed, tested, and applied a variety of statistical models to project future aboveground carbon stocks in response to climate change (Table 1). In all cases, models were driven by eight climate predictors: four seasons of temperature and precipitation. We report performance metrics for the different approaches (Table 2) and spatial patterns of error (Figure S3). Each eco-statistical approach revealed important insights about future carbon stocks, and on average projected losses of 8.8% ± 5.3% due to RCP4.5 climate change and 16.1% ± 7.5% due to RCP8.5 climate change (Table 3). We found high agreement in both magnitude and spatial patterns across the various approaches. The largest sources of variation, in order, were between (a) the dry and wet climate models, (b) the slow migration and equilibrium runs in the tree species niche models, and (c) RCP4.5 and RCP8.5.

3.1. Four Statistical Approaches to Project Future Carbon Stocks

In our first, simplest RF regression approach, the most important predictors of carbon density were fall, winter, and spring precipitation, with an average $R^2$ of 0.85 between out-of-sample predicted and observed carbon density (Table 2, Figure S4). While less important than fall, winter, and spring precipitation, temperature also enhanced model performance. The importance of summer and winter temperature in particular indicates that climate warming will cause changes in the distribution of carbon stocks. In agreement with other approaches, this RF regression revealed largest losses in the Northern California Coast ecoregion and foothills of the Sierra Nevada, with some potential for gain at high elevations (Figures 2, S5).

The RF classification approach identified specific areas of major plant type transitions, namely, between forest and shrubland ecosystems due to climate change. We found widespread conversion of forest to shrubland, even in wetting scenarios, especially in the lower elevation areas of the Sierra Nevada and Southern
Cascades (Figure 3). With RCP8.5 mean warming, our RF classification model projected a loss of 28.0% of forested area. This loss of forest area corresponded to a smaller decrease of 18.5% for AGL carbon density, given that these are on average less carbon-dense forest areas, and persistent shrublands account for a non-negligible amount of the state’s carbon stocks.

Table 2

<table>
<thead>
<tr>
<th>Eco-statistical approach</th>
<th>Performance</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1) RF regression of carbon density</td>
<td>RMSE = 13.4 $R^2 = 0.85$</td>
</tr>
<tr>
<td>(2) Classification of dominant veg. type</td>
<td>Classification accuracy = 90.1%</td>
</tr>
<tr>
<td>Confusion matrix</td>
<td></td>
</tr>
<tr>
<td>Pred. shrub</td>
<td>Pred. forest</td>
</tr>
<tr>
<td>Obs. shrub</td>
<td>66.0%</td>
</tr>
<tr>
<td>Obs. forest</td>
<td>4.6%</td>
</tr>
<tr>
<td>(4) Tree species niche models</td>
<td></td>
</tr>
<tr>
<td>20 species models’ average</td>
<td>RMSE = 2.0 $R^2 = 0.66$</td>
</tr>
<tr>
<td>Conifer model average</td>
<td>RMSE = 12.4 $R^2 = 0.80$</td>
</tr>
<tr>
<td>Hardwood model average</td>
<td>RMSE = 5.7 $R^2 = 0.80$</td>
</tr>
</tbody>
</table>

Notes. Models are fundamentally different approaches to describing future vegetation and carbon storage, with different relevant metrics for each. For the random forest regression type models (approaches 1 and 4), we report the average root mean square error (RMSE) and the coefficient of determination ($R^2$) between the predictions and observations at present-day. Both represent performance on out-of-sample data during cross-folding validation. For the random forest classification of forest-vs-shrub, we report the classification accuracy (expressed as a percent representing the number of correct classifications relative to the total number of classifications) and confusion matrix. For the third, climate analogs approach, there is not a singular model being fit to the present-day data for which to report a performance metric, but the goodness of fit of the analogs is described by the climatic novelty map shown in Figure S8.

Table 3

<table>
<thead>
<tr>
<th>Projected Change in Aboveground Live Carbon</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
</tr>
<tr>
<td>Eco-statistical approach</td>
</tr>
<tr>
<td>(1) RF regression of carbon density</td>
</tr>
<tr>
<td>(2) RF classification of dominant veg. type</td>
</tr>
<tr>
<td>(3) Climate analogs</td>
</tr>
<tr>
<td>Full domain</td>
</tr>
<tr>
<td>Restricted to 500 km</td>
</tr>
<tr>
<td>Restricted to 100 km</td>
</tr>
<tr>
<td>(4) Tree species niche models</td>
</tr>
<tr>
<td>20 species, equilibrium</td>
</tr>
<tr>
<td>Conifer versus hardwood, equilibrium</td>
</tr>
<tr>
<td>20 species, fast migration (500 m/yr)</td>
</tr>
<tr>
<td>20 species, slow migration (50 m/yr)</td>
</tr>
<tr>
<td>Average</td>
</tr>
<tr>
<td>Standard deviation</td>
</tr>
</tbody>
</table>

Notes. We estimated net carbon losses from climate change for several different scenarios and statistical modeling approaches. For comparison, the total change aligning with the State’s carbon sequestration goals is +4.2%. The largest differences, in order, are (1) between the dry and wet climate models, (2) the slow migration and equilibrium runs in the tree species niche models, (3) between RCP4.5 and RCP8.5.
The climate analogs approach provided similar patterns of change as the previous two approaches, with mean carbon loss of 14.0% for RCP4.5 and 23.2% for RCP8.5. We found no clear evidence that the magnitude of carbon change was sensitive to the restriction in search radius (in some, but not all climate scenarios, further restricting the search radius led to more carbon loss). Also, in quantifying how future climates will compare to present climates (Figures S6 and S7), we found that specific areas of California like the southern deserts, northern coasts, and Central Valley may have little resemblance to any present-day areas of the United States or Mexico (Figure S8).

The climate analogs approach provided similar patterns of change as the previous two approaches, with mean carbon loss of 14.0% for RCP4.5 and 23.2% for RCP8.5. We found no clear evidence that the magnitude of carbon change was sensitive to the restriction in search radius (in some, but not all climate scenarios, further restricting the search radius led to more carbon loss). Also, in quantifying how future climates will compare to present climates (Figures S6 and S7), we found that specific areas of California like the southern deserts, northern coasts, and Central Valley may have little resemblance to any present-day areas of the United States or Mexico (Figure S8).

Figure 2. Present observed aboveground live carbon density (a) and our modeled change with (b) RCP4.5 and (c) RCP8.5 climate change from the random forest regression. Areas of greatest vulnerability to climate-driven carbon loss are the northern coasts and low/mid-elevation Sierras, with some potential for carbon gain at high elevations. See Figure S5 for similar maps showing wet and dry models.

Figure 3. Results from RF classification of dominant vegetation type. Both RCP4.5 and RCP8.5 scenarios result in a net conversion of forest into shrubland, especially in the foothills of the Sierra Nevada and central coast. This type conversion is relevant as it would likely be associated with increased fire risk. Total loss of aboveground live carbon is approximately three times larger with the more extreme warming scenario of RCP8.5.
Finally, our fourth approach with tree species niche models using RF regression quantified how specific tree species could be impacted by varying degrees of climate change, and how migration capacity could substantially limit total carbon stocks (Figures 4, 89). The niche models projected carbon density declines of 30.7% for conifer species such as Douglas fir and Ponderosa pine with RCP8.5 climate change. On the other hand, oak species such as canyon live oak were projected to increase their total AGL carbon density by 43.7%. This replacement occurred especially in the low-to-mid elevation areas of the Sierra Nevada and Southern Cascades, and the general pattern of climate change favoring oaks over conifers agrees with previous research (McIntyre et al., 2015). Coast redwoods in particular showed high vulnerability at the southern ends of their range (south of San Francisco), in agreement with Fernández et al. (2015).

Migration strongly constrained our projections of total future ecosystem carbon. When we assumed that all tree species would be able to geographically adjust to reach equilibrium with future climate, carbon loss was only 2.8% for either RCP4.5 or RCP8.5. Carbon loss increased to 11.3% for RCP8.5 when we imposed a limit to migration consistent with an upper bound on tree migration rate observations (500 m/y), and to 20.0% using a more conservative estimate of possible migration rates (50 m/y).

3.2. Climate Drivers and Uncertainty

We also explored the role of uncertainty in future precipitation, and the entangled effects of temperature and precipitation. Across all approaches, wet models resulted in less carbon loss, and often carbon gain. On average, for RCP8.5, we projected 32.4% carbon loss with dry models and 1.3% carbon gain with wet models (Table 3). The carbon gain in the latter case suggests that the increased moisture availability in these wettest eight climate models is sufficient to compensate for the effects of 3–4°C of warming on water demand. To further understand the sensitivity of carbon density to climate controls, we compared RF regression model projections with RCP8.5 mean climate change but with temperature or precipitation held constant. We found that rising temperatures systematically drive carbon loss, while the variation in future precipitation contributes substantial uncertainty to the magnitude of loss (Figure 5). This finding was true for species-specific approaches as well, where temperature change explained most of the spatial patterns including large losses of coastal redwood in the south and increased favorability of hardwoods over conifers.

3.3. Spatial Patterns of Vulnerability

Lastly, we quantified several aspects of the spatial pattern of vulnerability (Figures 2 and 6), most notably with respect to elevation. Coastal areas and low-to-mid-elevation areas of the Sierra Nevada showed the greatest future carbon declines, whereas high elevation areas may offer the most potential for increased carbon storage. Based on our tree species niche models, these losses are largely explained by the shifts in redwood range on the coast and loss of conifers in favor of oaks in the Sierra Nevada (Figure 4).

The implications of carbon storage changes over the coming century in California will have important economic and policy impacts. We observed particular vulnerability in some of the areas where there are existing forest carbon offset projects as a part of California’s Cap-and-Trade system (California Air Resources Board, 2015) (Figure 7). The offset project protocol legally requires landowners to measure and verify carbon permanence for 100 years after any credits are issued. Credits are calculated by comparing carbon stocking to a 100-year modeled baseline, which could only evolve based on climate-driven risk if the crediting period were renewed every 25 years. We found that offset projects are located in disproportionately vulnerable parts of the state, such as the low-elevation regions of the Southern Cascades and Northern Coast ecoregions. With RCP8.5 mean climate change for our first RF regression approach, the average offset area loses 23.1% AGL carbon, while the state total projected loss is 15.5% across all ecosystems or 10.4% for forests (Figure 7b). Anticipation of these projected changes could inform more realistic baselines in order to minimize losses to such vulnerable areas and constrain expectations around forest management policy.
Figure 4. Selected results from tree species niche models, for the equilibrium scenario and RCP8.5 multi-model mean. Several species shift from low elevation to high elevation in the Sierra Nevada. Coast redwood shows high vulnerability in the southern part of its range, which may be compensated for by large increases in density in the north. In general, conifer species show future carbon losses while hardwood species show carbon gains.
4. Discussion

4.1. Climate Change Effects

Several important insights emerge from some of our key findings, especially where there is consistency across approaches. These insights can help build a more resilient future for California despite great climate uncertainties. The first is that a reduction in emissions from RCP8.5 to RCP4.5 leads to approximately half the end-of-century carbon losses from ecosystems. This result emphasizes how global emissions, the vast majority of which are not in California’s jurisdictional control, will determine the fate of California’s natural and working land carbon stocks. The second insight is that uncertainty in wet season precipitation regimes (Figure S1), under either climate scenario, will drive the majority of the variability in carbon storage. Investments in science that yield more robust, multi-decadal precipitation projections for the state will not only benefit the state’s water-strapped economy, but also our understanding of its natural and working land carbon future. Third, these results allow the state to consider climate mitigation portfolios that do not assume carbon storage stationarity. Doing so could result in proactive strategies that involve minimizing carbon losses in vulnerable areas, adding biomass in areas that will become more favorable for carbon storage (particularly high elevations), and assisting with the redistribution of key species. These projections could also be leveraged to establish more appropriate future baseline scenarios against which carbon sequestration projects, like those featured in California’s Cap & Trade Offset program, could be assessed.

The individual modeling approaches also provided complementary perspectives and insight. For example, one major conclusion from our RF classification of vegetation type is the substantial loss of forest cover (28%) with RCP8.5 mean climate change. Such large-scale conversion of forest to shrub or grasslands may be driven by climate-related disturbances such as wildfire (Abella & Fornwalt, 2015; Coop et al., 2016; Lauvau et al., 2016; Rother & Veblen, 2016; Savage & Mast, 2005; Tepley et al., 2017) and drought- and insect-driven mortality (Anderegg et al., 2013, 2020; Stephenson et al., 2019; Trugman et al., 2021), which have already been documented as important drivers of vegetation change across the Western United States.

Another set of insights comes from the climate analogs approach, which quantifies the degree of novelty of the emerging climate regimes across the state. The spatial pattern of novelty highlights the areas where the emerging climate regime is most novel as compared to the historical climate regime (Figures S8b and S8c). The less novel areas are where historically informed statistical approaches are interpolating within the existing climate variability and thus most likely to have some predictive power. In contrast, the more novel areas (in particular, along the Northern California Coast and central part of the Southern Cascades) are where statistical approaches are fundamentally extrapolating and may require approaches based on process representation to understand their future trajectories.

From the separate species niche models, we projected that certain tree types like oaks will be favored over conifers. Certain key tree species like redwood were projected to see substantial range shifts, with total carbon density being limited by migration capacity as a result. The pattern of increasing total carbon loss with lower migration potential supports the idea that the velocity of climate change is a source of ecosystem vulnerability (Ackerly et al., 2010; Loarie et al., 2009), and that management activities such as assisted migration could have a large impact on carbon storage and conservation of key species. For coast redwood in particular, large increases in density at the northern end of its coastal range and even expansion toward the Sierra Nevada could theoretically offset losses at the southern end of its range, but only if migration rates do not become limiting. Realistically, redwoods take hundreds of years to grow, so range shifts resulting...
AGU Advances

4.2. Assumptions and Limitations

This study was based on a few major assumptions. First, we chose a fairly simple representation of environmental drivers, that is, four seasons of temperature and precipitation, and focus on exploring the influence of different statistical approaches and climate scenarios. We did not explicitly represent other drivers as input variables—such as other hydrological variables, soil characteristics, or vulnerability to insects—although many of these processes are implicitly represented in the climate-carbon relationships derived from our statistical models. Regarding hydrology, our approach captured some of the seasonal structure and aspects of water balance (e.g., length and intensity of the dry season, high temperature limits) but did not explicitly account for other hydrological variables like drought indices (Madakumbura et al., 2020) and vapor pressure deficit which should be explored in future studies. We also did not consider the potential increasing interannual variability of precipitation with climate change (Swain et al., 2018), which may be another source of ecosystem vulnerability and additional decreases in carbon stocks.

Second, our methodology involving fairly coarse spatial resolution and fitting of climate drivers to empirical data is an imperfect representation of finer scale ecological dynamics. Due to limitations in both data and modeling, we cannot capture the exact fundamental niches of different species and vegetation types as a function of landscape and watershed position, but rather estimate the realized climate niches at an eighth-degree within California. In the results presented in the main text, we also have not explicitly accounted for land use legacy, disturbance history, or forest age. We provide a supplemental analysis (Table S2 and Figure S10) showing that an initial attempt to account for post-fire and post-harvest impacts on carbon stocks in our random forest regression or classification had only a minimal impact on our carbon density projections. While the supplemental analysis does not change any of our major conclusions, which are more focused on comparison of different climate scenarios and statistical approaches, it highlights the importance of regarding with caution the interpretation of individual grid cell changes given the heterogeneity of fire and harvest effects. Relatedly, embedded in our projections is an assumption of a set of climate-fire-management interactions which do not change markedly over the next century. If the State undertakes a fundamentally different approach to fire and land management, like widespread forest fuel reduction treatments (Agee & Skinner, 2005) or fire regimes intensify beyond the current range of observations, these could dampen or amplify some of our projected changes.

Third, we assumed that other potential factors to mitigate carbon loss (i.e., CO₂ fertilization and acclimation), would be negligible compared to the scale of spatial reorganization of vegetation represented in our models. Regarding CO₂, there is a lack of agreement in the literature on the extent to which carbon storage will be enhanced by rising CO₂, especially considering the concurrent changes to drought frequency (Birami et al., 2020; Jiang et al., 2020; Lenihan et al., 2008b; Needham et al., 2020; Sperry et al., 2019; Swann et al., 2016) and declining nitrogen availability in some cases (Luo et al., 2004; Wamelink et al., 2009). For example, one study found that 55–71% of climate projections have enough CO₂ increase to offset the temperature-driven mortality, depending on the extent of acclimation (Sperry et al., 2019), whereas another found that heat and drought erased any benefits of increased CO₂ (Birami et al., 2020). Acclimation may play a substantial role, though the extent of which is uncertain (Sperry et al., 2019). Due to this uncertainty, in net carbon gains are implausible on the timescale of this century. In addition, our methodology does not explicitly capture coastal fog as a moisture source, and there is evidence of declining fog frequency in the last century (Johnstone & Dawson, 2010), highlighting the importance of considering a broader set of climate drivers in future work.
and the lack of elevated CO$_2$ experiments in semi-arid forest ecosystems, our analysis targeted only the climate-driven ecosystem response. Our scenarios may be representative of the full ecosystem response, including CO$_2$ fertilization, if changes in water use efficiency due to CO$_2$ are small in comparison to the effects of effects of 3°–4° of warming. During the historical era, the accelerating effects of large-scale drought and fire mortality and across California and the western United States seem to suggest that, so far, the magnitude of climate impacts on forests is substantially larger than the benefits from rising CO$_2$.

Finally, our analysis considered aboveground live carbon and did not attempt to model dead or belowground carbon pools, which account for a majority (83%) of ecosystem carbon in California (California Air Resources Board, 2019). In order to compensate for a projected aboveground live loss of 11.3% and meet state goals of 4.2% total ecosystem carbon increase, these other pools would have to increase by 6.1%. There is considerable uncertainty in the carbon dynamics of the dead pools and whether changes in these pools could compensate for losses in aboveground biomass. Drier conditions may slow decomposition and minimize carbon losses of litter and coarse woody debris; however, the buildup of these pools would also

Figure 7. Vulnerability of California forest carbon offset projects to RCP8.5 mean climate change. Thirty-two forest carbon offset projects are based in California, in the Northern California Coast and Southern Cascades ecoregions (black polygons in panel (a)). For RCP8.5 mean climate change, grid cells where projects are located are projected to lose 23.1% of AGL carbon (orange, b), while the average forested grid cell loses 10.4% of aboveground live (AGL) carbon (blue, b), and the total statewide expected loss is 15.5% (Table 3). For RCP4.5 mean climate change, project grid cells are projected to lose 6.5% of AGL carbon compared to average forest loss of 3.1%. The disproportionate vulnerability of these critical areas is likely to impede their capacity to store carbon for 100+ years as required per the forest carbon offset protocol.
increase fire risk. The direct effects of warming, in contrast, may accelerate decomposition, increasing losses from litter and soil carbon pools (Davidson et al., 1998; Davidson & Janssens, 2006).

4.3. Implications for Land Management

Land management strategies have the potential to mitigate some of the projected carbon losses reported here. A study by Cameron et al., 2017 found that extremely ambitious implementation of conservation, restoration, and forest management could contribute up to 26 MMTCO₂e/y by 2050, or 135 MtC (2.5%) to total ecosystem carbon by 2050, not considering climate change. Another found that a low population growth and land-use scenario could contribute 215 MtC (3.9%) by 2100, even with RCP8.5 climate change (Sleeter et al., 2019). These potential increases to total ecosystem carbon would be enough to offset our projected loss of aboveground live carbon alone, but likely not enough to offset potential losses in the other larger carbon pools.

More broadly, our spatial patterns of climate-stable and climate-unstable carbon stocks (and habitat types) are relevant for management as they could inform where different actions would be most effective over this century. The current one-size-fits-all strategy to maximize carbon stocks across California forests appears poorly suited to the projected shifting mosaic of carbon with climate change. We suggest a more climate-aware approach—for example, in climate-unstable locations such as the low-mid elevation Sierra Nevada and central and northern coastal ranges, management should focus on stabilizing existing carbon stocks against inevitable climate-driven transitions, rather than incentivizing carbon gain. In these areas, the priorities could be to reduce the risk of catastrophic fire, and thinning and restoration to promote large trees and reduce water stress. Management to increase carbon stocks would be most valuable in select climate-stable locations such as above 2,000 m elevation in the Sierra Nevada and in the northwestern coastal Klamath range. In addition, assisted migration and establishment, though difficult and controversial, could allow key species like coast redwood to relocate to or increase density in climate-favorable habitat regions (McLachlan et al., 2007; Millar & Stephenson, 2015). These species take decades or centuries to reach maturity, and so early action in anticipation of climate change is essential for both conservation and achieving California’s long-term carbon goals.

Our results provide actionable insights about the likely magnitude and uncertainties of terrestrial carbon change over this century. We present a comprehensive statistical analysis showing widespread agreement in the direction and relative magnitude of change across different approaches. We also included a consideration of migration which is often not well accounted for in other modeling approaches. Our findings highlight that the uncertainty in this migration component, as well as in future precipitation, are major scientific bottlenecks for long-term ecological forecasting. Overall, we estimate that rapid warming in the coming decades will drive large declines in aboveground biomass, especially in coastal and low-elevation areas. The losses are in stark contrast to California’s goals of markedly increasing the land carbon sink toward carbon neutrality by 2045. The projected losses also suggest that climate-driven vulnerability should be an important point of continued research in the context of natural climate solution strategies, such as California’s forest carbon offsets program, which rely heavily on sustained carbon dioxide removal by ecosystems.

Conflict of Interest

M. L. Goulden reports grants from the California Strategic Growth Council, and K. S. Hemes had an occasional scientific advising relationship with the California Forest Carbon Coalition. Neither of these impacted the scientific analysis presented in this publication.

Data Availability Statement

The data that support the findings of this study were derived from the following resources available in the public domain. We appreciate the insight and data development work of our colleague Klaus Scott at the California Air Resources Board, who provided the California AGL carbon data layer. https://ww2.arb.ca.gov/nwl-inventory; https://lemma.forestry.oregonstate.edu/data; https://www.mrlc.gov/data/nlcd-2016-land-cover-conus; https://catalogue.ceda.ac.uk/uuid/bedc59f37c9545c981a839eb552e4084; ftp://gdo-dcp.
Acknowledgments

This work is based upon support received from the National Science Foundation (NSF) Graduate Research Fellowship Program under grant number DGE-1839285 (for S. R. Coffield), by the UCOP National Laboratory Fees Research Program under grant number LFR-18-542511 (for S. R. Coffield, J. T. Randerson, C. D. Koven, M. L. Goulden), by the California Strategic Growth Council’s Climate Change Research Program with funds from California Climate Investments as part of the Center for Ecosystem Climate Solutions (for M. L. Goulden, J. T. Randerson, S. R. Coffield, and K. S. Hemes), by NSF under grant number 1633631 (for J. T. Randerson) as part of the University of California, Irvine (UCI) NSF Research Traineeship (NRT) Machine Learning and Physical Sciences (MAPS) Program, by the Department of Energy (DOE) Office of Science’s Reducing Uncertainty in Biogeochemical Interactions through Synthesis and Computation (RUBISCO) Science Focus Area (for J. T. Randerson and C.D. Koven), by the DOE Early Career Research Program (for C. D. Koven), by the Regents of the University of California under prime contract number DE-AC02-05CH12313 for management and operation of LBNL (for C. D. Koven), and by the Stanford Woods Institute for the Environment (for K.S. Hemes). We also acknowledge the Native peoples and land of California. UC Irvine, where most of this research has taken place, is located on the shared ancestral territory of the Acjachemen and Tongva Peoples, many of whom maintain strong and active physical, spiritual, and cultural ties to the region and to the local environment.

References


Brekke, L., Thrasher, B. L., Maurer, E., & Fruit, T. (2013). Downscaled CMIP3 and CMIP5 climate projections: Release of downscaled CMIP5 climate projections, comparison with preceding information, and summary of user needs.


ucnl.niov/pub/1cp/archive/cmip5/bc/5. All input data, model projections, and Python and Google Earth Engine scripts are available in a public repository via Dryad: https://doi.org/10.7280/D1568Z. Data include observations and model estimates of AGL carbon density on a 0.125° × 0.125° grid. The AGL carbon density data set at its original resolution of 30 m is available upon request from CARB.


Reference From the Supporting Information