INNOVATIONS IN MEASURING AND MANAGING FOREST CARBON STOCKS IN CALIFORNIA

A Report for:

California's Fourth Climate Change Assessment

Prepared By:
John J. Battles[1], David M. Bell[2], Robert E. Kennedy[3], David S. Saah[4], Brandon M. Collins[1], Robert A. York[1], John E. Sanders[1], Fernanda Lopez-Ornelas[5]

1 UC Berkeley

2 USDA Forest Service

3 Oregon State University

4 University of San Francisco

5 Spatial Informatics Group

Edmund G. Brown, Jr., *Governor*

August2018
CCCA4-CNRA-2018-014

ACKNOWLEDGEMENTS

Field and remote sensing data used in this project was obtained with funding from the following institutions and programs: The Nature Conservancy, the United States Department of Energy, the United States Forest Service Region 5, the United States Forest Service Forest Inventory and Analysis Program, the Sierra Nevada Adaptive Management Project, and the University of California. This work was supported by the California Fourth Climate Change Assessment. Additional support was provided by the USDA National Institute of Food and Agriculture, McIntire Stennis project CA-B-ECO-0144-MS. Mr. Klaus Scott from the California Air Resources Board provided valuable advice and technical assistance throughout this project. We appreciate his help.

PREFACE

California's Climate Change Assessments provide a scientific foundation for understanding climate-related vulnerability at the local scale and informing resilience actions. These Assessments contribute to the advancement of science-based policies, plans, and programs to promote effective climate leadership in California. In 2006, California released its First Climate Change Assessment, which shed light on the impacts of climate change on specific sectors in California and was instrumental in supporting the passage of the landmark legislation Assembly Bill 32 (Núñez, Chapter 488, Statutes of 2006), California's Global Warming Solutions Act. The Second Assessment concluded that adaptation is a crucial complement to reducing greenhouse gas emissions (2009), given that some changes to the climate are ongoing and inevitable, motivating and informing California's first Climate Adaptation Strategy released the same year. In 2012, California's Third Climate Change Assessment made substantial progress in projecting local impacts of climate change, investigating consequences to human and natural systems, and exploring barriers to adaptation.

Under the leadership of Governor Edmund G. Brown, Jr., a trio of state agencies jointly managed and supported California's Fourth Climate Change Assessment: California's Natural Resources Agency (CNRA), the Governor's Office of Planning and Research (OPR), and the California Energy Commission (Energy Commission). The Climate Action Team Research Working Group, through which more than 20 state agencies coordinate climate-related research, served as the steering committee, providing input for a multisector call for proposals, participating in selection of research teams, and offering technical guidance throughout the process.

California's Fourth Climate Change Assessment (Fourth Assessment) advances actionable science that serves the growing needs of state and local-level decision-makers from a variety of sectors. It includes research to develop rigorous, comprehensive climate change scenarios at a scale suitable for illuminating regional vulnerabilities and localized adaptation strategies in California; datasets and tools that improve integration of observed and projected knowledge about climate change into decision-making; and recommendations and information to directly inform vulnerability assessments and adaptation strategies for California's energy sector, water resources and management, oceans and coasts, forests, wildfires, agriculture, biodiversity and habitat, and public health.

The Fourth Assessment includes 44 technical reports to advance the scientific foundation for understanding climate-related risks and resilience options, nine regional reports plus an oceans and coast report to outline climate risks and adaptation options, reports on tribal and indigenous issues as well as climate justice, and a comprehensive statewide summary report. All research contributing to the Fourth Assessment was peer-reviewed to ensure scientific rigor and relevance to practitioners and stakeholders.

For the full suite of Fourth Assessment research products, please visit www.climateassessment.ca.gov. This report explores new approaches to measure forest biomass and develops new methods to quantify the trade-off between biomass storage and stability for fire prone forests.

ABSTRACT

California has a pressing need to measure and manage forest carbon. Fusion of satellite-based data with plot-level information provides a promising means to measure forest biomass at relevant spatial and temporal scales. Key questions remain regarding accuracy and feasibility. Over a century of fire suppression complicates managing forest carbon in California's dry forests. Live tree biomass is at risk of loss due to wildfire. Here, we evaluated the performance of an emerging technology using Landsat imagery, forest inventory data, and gradient nearest neighbor imputation (referred to as LT-GNN) to measure annual aboveground live tree biomass (AGB) across multiple spatial scales. We also developed a means to quantify the trade-off between biomass storage and stability for fire-prone forests.

We relied on two independent estimates of AGB to evaluate LT-GNN results: local assessments calculated from field-data and airborne light detection, and county estimates calculated from Forest Inventory and Assessment plot results. We also used repeated measurements conducted in Forest Inventory and Analysis plots to quantify the ability of LT-GNN to detect trends in AGB. Finally, we extended a field experiment at Blodgett Forest Research Station in Georgetown, California to gain insights into biomass dynamics of fire-prone forests.

LT-GNN is a promising method to monitor live tree biomass. Its success at interpolating county-level tree biomass suggests an application-ready means to track annual biomass at a policy-relevant scale. However, improvements are needed to track change under stable conditions. At finer scales applications must be pursued with more caution. In particular, LT-GNN did not accurately predict AGB in an old-growth redwood forest. At Blodgett, we quantified the trade-off between biomass storage and stability. Fuel treatments did lower the overall biomass stored, but more biomass survived fire compared to the untreated forest. However, trade-off between biomass storage and stability critically depends on the probability of fire occurring in these stands.

Keywords: aboveground live tree biomass, LandTrendr, Gradient Nearest Neighbor, biomass stability, biomass storage, carbon carrying capacity, forest biomass monitoring system

Please use the following citation for this paper:

Battles, John, David Bell, Robert Kennedy, David Saah, Brandon Collins, Robert York, John Sanders. (University of California, Berkeley). 2018. ***Innovations in Measuring and Managing Forest Carbon Stocks in California.*** California's Fourth Climate Change Assessment, California Natural Resources Agency. Publication number: CCCA4-CNRA-2018-014.

HIGHLIGHTS

- The LandTrendr-Gradient Nearest Neighbor system is a promising means to monitor live tree biomass in California's forest.
- The success of this approach at interpolating county-level tree biomass suggests an application-ready means to track annual biomass at a policy-relevant scale.
- However, improvements need to be made in LT-GNN's ability to track change under stable conditions and to capture AGB in high-biomass forests.
- We presented a framework for evaluating the trade-off between biomass storage and stability in fire-prone forest and applied it to a well-studied Sierran mixed conifer forest.
- Fuel treatments lowered the overall biomass stored but more of this biomass survived a fire compared to the untreated forest.
- We proposed the term "stable aboveground biomass" to describe the fraction of live tree biomass on a site that is capable of surviving a problem wildfire.

TABLE OF CONTENTS

1: Innovations in Measuring Forest Carbon

1.1 Introduction

Remote sensing combined with field inventory data can provide spatially and temporally consistent biomass estimates. Woodall et al. 2015

As noted by the US Forest Service's framework for carbon accounting (Woodall et al. 2015), the fusion of remote sensing with field data provides a potential means to improve forest carbon monitoring. As one of the few jurisdictions in the world to enact mandatory greenhouse gas (GHG) emissions reductions, California has a pressing need to measure its forest carbon (Forest Climate Action Team 2018). Yet, measuring forest carbon at the appropriate temporal and spatial resolution to inform GHG reduction strategies has proven challenging.

The forests of California are vast, diverse, and dynamic. There are nearly 130,000 km^2 of forest that account for 31% of the land area in the state (Christensen et al. 2017). Forests in the state represent more than 200 different vegetation classes (Landfire 2010) and include some of the most carbon dense ecosystems on earth (Gonzalez et al. 2010). These forests are also rapidly changing due to a host of drivers like global warming, urban development, severe wildfire, warm droughts, invasive diseases, and insect pests (Forest Climate Action Team 2018). Recent revisions in the Forest Inventory and Analysis (FIA) program have greatly improved the inventory of forest carbon (Bechtold and Patterson 2005). However, given the design of the sampling, methods that rely exclusively on FIA data will be limited to a temporal resolution of every 10 years (sensu Woodall et al. 2015) and a spatial resolution at the state and county levels.

The fusion of satellite-based data with plot-level information provides a means to improve the resolution of forest biomass maps (Gonzalez et al. 2015, Kennedy et al. 2018, Appendix A). However key questions remain about their accuracy and feasibility. The approach developed by Gonzalez et al. (2015) relies on LANDFIRE, an integrated national product to track change in forest carbon (Landfire 2010). Thus, the timing is constrained by the release of LANDFIRE updates. Also, as noted by Gonzalez et al. (2015), their method underestimated growth in intact mature forests. Kennedy et al. (2018) describe a system that integrates time series signals derived from the Landsat Thematic Mapper with FIA inventory data to measure annual changes in forest biomass. However, a major concern is feasibility. Their method is computationally and analytically expensive (Kennedy et al. 2018). Thus, neither method fully satisfies California's need for a reliable, cost-effective, system for monitoring forest carbon (Forest Climate Action Team 2018).

In this chapter, we evaluated the performance of an emerging technology, the Landsat-based Detection of Trends in Disturbance and Recovery (LandTrendr, Kennedy et al. 2010), to measure forest carbon in California at yearly intervals across multiple spatial scales. The assignment of carbon values relies on results from the FIA plot data and a nearest neighbor imputation technique – the gradient nearest neighbor (GNN) method (Ohmann and Gregory 2002). The combination of LandTrendr segmentation and GNN imputation has proven a reliable means to track changes in forest structure (Ohmann et al. 2012), including aboveground live biomass (Kennedy et al. 2018). Specifically, we tested the hypothesis that LandTrendr

coupled with GNN imputation represents a repeatable, cost-effective approach to measure aboveground live biomass in California's forests.

1.2 Methods

1.2.1 Overview

In this chapter, we focused on the measurement of aboveground live tree biomass (AGB). AGB is directly tied to carbon storage in that 1 grams of woody biomass (dry) typically contains 0.47 grams carbon (McGroddy et al. 2004). We recognize that live trees represent only a fraction of the forest carbon (Woodall et al. 2015). However, it is a large pool that sequesters CO_2 as long as the tree remains alive. Given the longevity of live trees, they contribute to mitigation strategies. Moreover, the aboveground portion can be reliably measured in the field and linked to signals gleaned from remote sensing (Appendix A). Finally, aboveground live tree biomass tends to be a good indicator of forest carbon accumulation (Keith et al. 2009).

Our biomass monitoring scheme builds on the LandTrendr-GNN (LT-GNN) framework described in Kennedy et al. (2018). However, the LT-GNN estimates tested here include three innovations: a version of the LandTrendr algorithm implemented on the Google Earth Engine platform; alternatives for GNN imputation that account for uncertainties in gradient space; and the extension of LT-GNN mapping in California from 1990-2016. These innovations are described in more detail below. To evaluate the LT-GNN results, we relied on two independent estimates of AGB: local assessments calculated from field data and airborne light detection and ranging (LiDAR) and county estimates calculated from FIA plot results. We also used the repeated measurements conducted in FIA plots to quantify the ability of LT-GNN to detect trends in AGB.

1.2.2 LandTrendr-GNN framework

LT-GNN provides a valuable approach for translating the multi-decadal Landsat archive into a spatially and temporally explicit representation of forest biomass, and thus carbon stock, patterns. LT-GNN was developed under a grant from the USDA National Institute for Food and Agriculture by the Kennedy lab at Oregon State University and the Landscape Ecology, Modeling, Mapping and Analysis (LEMMA) team, a collaborative Oregon State University and USDA Forest Service research team currently led by David Bell. LT-GNN involves three phases (Figure 1.1): LandTrendr algorithm implementation, integration of LandTrendr and GNN, and plot-based analyses and comparisons. The LandTrendr algorithm (Figure 1.1; LandTrendr) uses pixel-level temporal segmentation procedures to remove noise from the Landsat image time-series (LTS) and retains the major trends in the spectral signal through time (described in detail in Kennedy et al. 2010). The resulting temporally-smoothed LTS data are then used for biomass modeling purposes. Moreover, the introduction of a new Landsat sensor (i.e., Landsat 8) raised a concern about temporal consistency of the images -- an important consideration since LandTrendr relies on limiting noise in the LTS. Thus, as part of the effort to develop LT-GNN as an operational means to monitor forest carbon, we moved the LandTrendr algorithm into Google's Earth Engine platform in a manner that allows incorporation of the evolving Landsat sensor suite. See Appendix B for details on the transition. All of the LT-GNN results presented here were based on imagery processed using the Google Earth Engine implementation of LandTrendr.

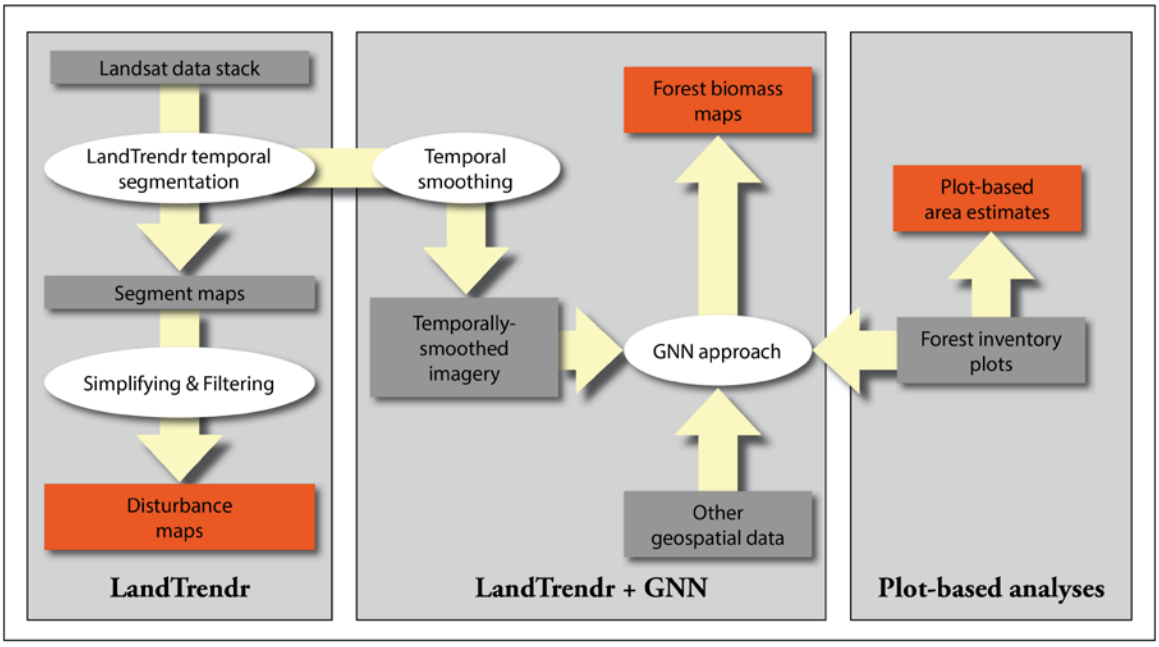


Figure 1.1: Representation of the LandTrendr-GNN framework

Once produced, temporally-smoothed LTS imagery is integrated with other geospatial data and forest inventory data in the GNN modeling method to produce forest biomass maps (Figure1.1; LandTrendr+GNN). GNN is a *k* nearest neighbor (*k*NN) imputation methodology designed explicitly for landscape vegetation mapping (Ohmann and Gregory 2002). Imputation mapping is a method for substituting observed values (i.e., forest inventory data) to replace missing data for all pixels in an area, resulting in a map as output. Specifically, a *k*NN imputation assigns those missing values based on some set of observations that are similar to the pixel of interest in terms of environmental conditions (i.e., minimizes distance in environmental or gradient space). In the case of GNN, the environmental space is defined by a constrained ordination (direct gradient analysis) – canonical correspondence analysis (CCA; ter Braak 1986) that relates forest attributes from forest inventory plots (i.e., species matrix) to geospatial data extracted at those same plot locations (i.e., environment matrix) – for measuring and weighting distances in nearest neighbor calculations (Ohmann and Gregory 2002). Past experience indicates that climate, topography, and ecoregion best define the environmental matrix (Ohmann et al. 2014). Based on the CCA, both forest inventory plots and pixels to which we wish to impute data can be placed in gradient space. In this gradient space, *k*NN imputation then proceeds by allocating the (weighted or unweighted) mean of the *k* nearest neighbors for a given pixel (i.e., minimum distances to pixels in gradient space) to each pixel, producing the vegetation maps of interest.

Central to the performance of LT-GNN is the *k*NN imputation method. Several methods exist for assessing imputation uncertainties, including parametric methods and non-parametric methods (McRoberts et al. 2007, McRoberts 2012, Bell et al. 2015). Specifically, we developed LT-GNN biomass maps using three implementations of GNN: $k = 1$ (GNN_{k1}), bootstrapping approximation (GNN_{ba}), and $k = 10$ (GNN_{k10}). GNN_{k1} is the traditional GNN implementation, where the nearest neighbor to the pixel in gradient space is imputed to that pixel (e.g., Ohmann and Gregory 2002). This method results in maps where unrealistic combinations of attributes cannot be produced (i.e., each prediction was observed at least once in the field). However,

model precision cannot be directly assessed. For larger values of *k*, estimates of uncertainties can be assessed. For GNN_{ba}, an approximation of non-parametric bootstrapping can be used to assess precision at the scale of pixels for a $k = 1$ *k*NN process (Bell et al. 2015). For GNN_{k10}, variation in neighbors within and among pixels forms the basis of model-based inference (McRoberts et al. 2007), with predictions being the unweighted means of the *k* nearest neighbors. We constructed AGB predictions based on GNN_{k1}, GNN_{ba}, and GNN_{k10} in order to access the accuracy for all three predictions and to estimate the uncertainty for results based on GNN_{k10}. For details on the GNN modeling see Appendix C.

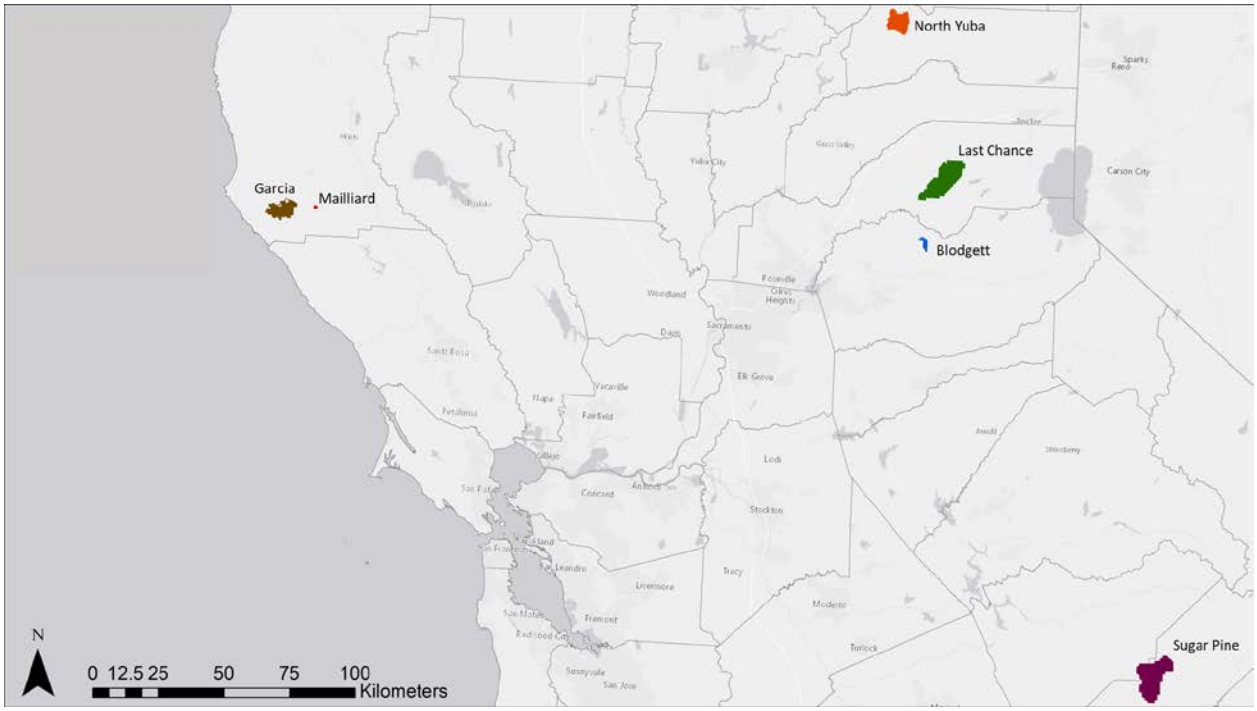


Figure 1.2: Map showing location of the study sites in California where the LiDAR-derived maps of aboveground live tree biomass were developed with field data.

1.2.3 Evaluating LT-GNN predictions

1.2.3.1 LiDAR-derived AGB maps

For the plot-to-landscape scale (900 m^2 to >100 km^2) assessment of LT-GNN, we developed AGB maps for six sites where we had concurrent field inventory and LiDAR data. These study areas (Figure 1.2) represent a range of forest conditions found (Table 1.1) in two common and biomass-dense forests in California: the Sierran conifer forest and the coast redwood forest (Gonzalez et al. 2010, Gonzalez et al. 2015). Such maps based on airborne LiDAR metrics combined with reliable concurrent field measurements produce highly reliable estimates of forest biomass (Dubayah and Drake 2000, Zolkos et al. 2013). Thus, we considered the LiDAR results as high fidelity baselines to evaluate the LT-GNN predictions at the local scale.

Table 1.1: Summary of information on field plot data collection for the study areas. AGB is the aboveground live tree biomass; COV is the plot-level coefficient of variation in AGB. AGB reported as megagrams per hectare (Mg ha^{-1}).

Site	Plot number	Plot size (ha)	Year(s) sampled	AGB median (max-min) (Mg ha^{-1})	COV
Garcia	28	0.1	2005	290 (68-472)	48
Mailliard	12	0.1	2005	606 (336-1392)	47
North Yuba	36	0.1	2005	317 (0-1118)	68
Last Chance	103	0.05	2013	136 (0-1625)	119
Blodgett	525	0.04	2013-15	298 (3- 603)	69
Sugar Pine	121	0.05	2007-08	348 (34-1667)	71

Our goal in the production of the LiDAR-derived AGB maps was to apply proven approaches consistently across all six sites (Figure 1.3). Details are available in Appendix D. Briefly we calculated plot-level AGB estimates using the same regional allometric equations (FIA 2010) included in the LT-GNN models. These plot values were fitted to forest height and cover metrics obtained from the LiDAR with Fusion software (McGaughey 2018). We used a machine learning algorithm, VSURF (Genuer et al. 2015), to select the LiDAR variables included in the regression modeling (sensu Kennedy et al. 2018) that produced the AGB transfer functions. These transfer functions were applied to landscape level forest height and cover metrics to generate a map of AGB for the entire site (Figure 1.3).

1.2.3.2 Comparing LiDAR-derived AGB maps to LT-GNN results

To compare LiDAR-derived AGB map to the LT-GNN map, we used orthogonal regression (Carroll and Ruppert 1996) to correct for the effects of error in the predictors (i.e., LiDAR estimates). Also, as noted by Smith (2009), a symmetric fit is the appropriate approach to measuring the relationship between the two estimates. Legendre and Legendre (2012) recommend reduced major axis regression (also known as standard major axis regression) when the two variables are strongly correlated. We reported the slope of the regression, the correlation coefficient (r), and the bias, calculated as mean error (ME):

$$ME = \sum_{i=1}^{n} \frac{(GNN_i - LiDAR_i)}{n} \qquad \text{Equation 1.1}$$

where *i* represents the estimate for a given pixel, n is the total number of estimates, GNN_i is the GNN estimate of AGB for pixel i, and $LiDAR_i$ is the LiDAR estimate of AGB for pixel i.

For comparisons from the plot-to- stand level, we evaluated three implementations of GNN: k = 1 (GNN_{k1}), bootstrapping approximation (GNN_{ba}), k = 10 (GNN_{k10}). For the site level comparisons, we relied on GNN_{k10} results in order to coincide with the county-level analysis (see Section 1.2.2.3). We limited our comparisons to forested regions in our study areas. Forests

were defined using the most recent vegetation classification system provided by the California Fire and Resource Assessment Program (FRAP Vegetation 2015). Life form classifications as "CONIFER" and "HARDWOOD" were included in our analysis. On average, forests occupied more than 92% of the area within the six study sites.

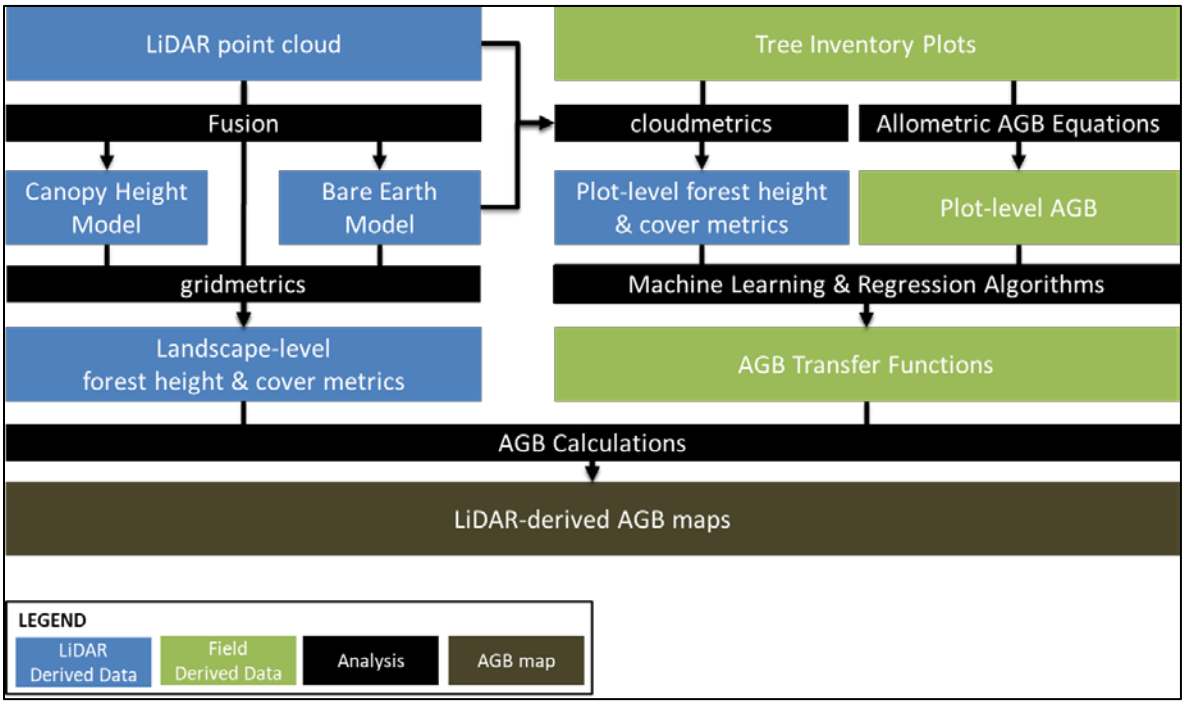


Figure 1.3: Analytical workflow for production of baseline LiDAR maps of aboveground live tree biomass (AGB).

1.2.3.3 County-level comparisons

To gain a statewide perspective, we compared county-level estimates of AGB from LT-GNN to results from FIA plot measurements. Forests can be found in all counties of California, but the total area of forest lands and their dominance varies by county (Figure 1.4). Northwestern California and the Sierra Nevada mountains are the areas most dominated by forest landscapes. In this study, GNN maps of AGB were most appropriate in forest lands, so inference of AGB status were restricted to forest lands based on United States Geological Service (USGS) land cover data (Grossmann et al. 2008).

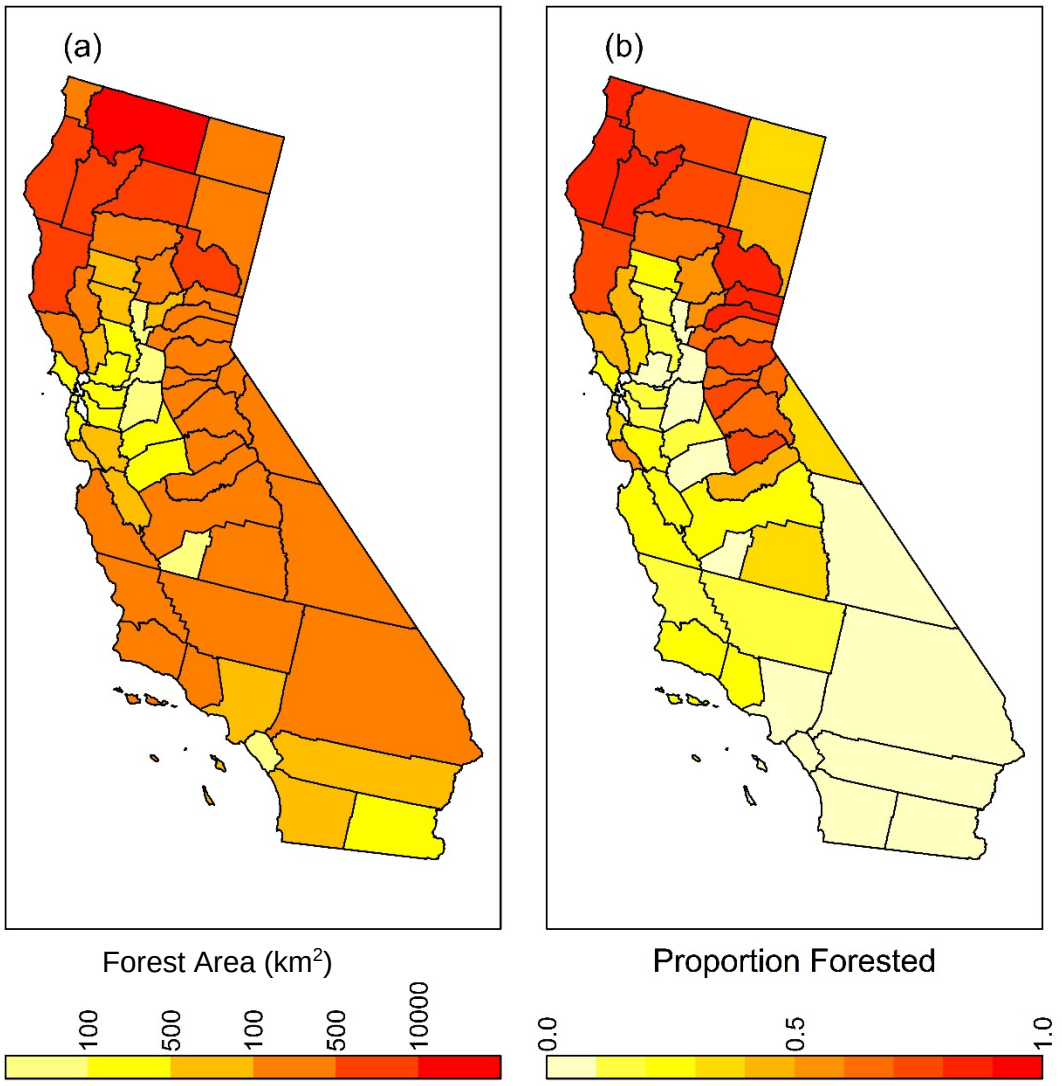


Figure 1.4: The distribution of forests in California by county. Map of (a) total forest area (km^2) and (b) proportion of each county forested. These results were based on forest/non-forest classifications from the USGS Gap Analysis Program (Grossman et al. 2008).

To provide county-level assessments of AGB distributions, we used a model-based estimation technique for *k*-nearest neighbor imputation (McRoberts et al. 2007). In this case, we estimated mean and standard deviation for GNN_{k10} AGB predictions for each county based on results for 2006. Since FIA plots are measured on a ten-year cycle, we used a single measurement of each forested FIA sampling location from 2001-2010 to calculate a mean FIA estimate of AGB for each county assuming the FIA plots in the county represented a random sample. As we did for the LiDAR AGB results, we quantified the relationship between the two county-level estimates using reduced major axis regression. The fit of the regression was quantified with Pearson correlation coefficient and the bias with mean error (Equation 1.1).

Note that this GNN accuracy assessment using FIA plot data is not technically independent since FIA plots are used in the GNN interpolation. However, we leveraged a modified leave-

out-out validation technique to mitigate this dependence (Ohmann and Gregory 2002). Specifically, for the purposes of accurate assessment, predictions for each plot were generated by ignoring the actual plot during imputation (i.e., no plot was self-assigned during validation). Thus, while each plot can influence the gradient space defining nearest neighbors, we avoided comparing imputed plots to themselves during accuracy assessments.

1.2.3.4 Change detection

In addition to comparing biomass stocks, we assessed the performance of LT-GNN in predicting plot-level change over time. We examined change at 3,412 FIA plots with at least two measurements in time and selected the first and last measurement at each location. We then compared predicted versus observed at each location to examine how LT-GNN performed in assessing AGB change. For analysis, we took advantage of LandTrendr's ability to categorize the underlying dynamics captured in the repeated plot measurements. Specifically, we divided the observed changes into three scenarios: disturbed, recovering, and stable. Disturbed plots were those experiencing a LandTrendr disturbance (lasting one to three years) during the measurement interval. Recovering plots were those with a LandTrendr disturbance prior to, but not including, the measurement interval. Stable plots had no LandTrendr disturbances (Kennedy et al. 2010). Again, we relied on regression and bias analysis to quantify the fit, precision, and accuracy between observed and predicted AGB.

1.3 Results

1.3.1 Local-level Assessment of LT-GNN

The predictive models of AGB based on LiDAR metrics provided reasonable to excellent projections across the six sites (Table 1.2) compared to LiDAR-derived maps. In all but one case, cubic transformations were necessary to meet assumptions of linear modeling. Interestingly the two sites in the coast redwood forest (Garcia and Mailliard) straddled the range of performance in terms of fit (i.e., R^2) even though the field sampling and remote sensing data were identical (Appendix D).

Table 1.2: Evaluation of predictive AGB models derived from LiDAR cloudmetric parameters for the six study areas in California's conifer forests. RMSE = root mean squared error; rRMSE= relative root mean squared error (Equation D-1).

Site/Date	Equation form	R^2 (adjusted)	RMSE (Mg ha^{-1})	rRMSE (%)
Garcia 2005	natural log	0.52	9	4
Mailliard 2005	cubic	0.81	26	4
North Yuba 2005	cubic	0.87	53	15
Last Chance 2012-13	cubic	0.87	36	14
Blodgett 2014	cubic	0.66	54	16
Sugar Pine 2007	cubic	0.72	45	11

At finer spatial scales (30 meters, 90 meters, and 150 meters), LT-GNN underestimated AGB compared to LiDAR results for Mailliard (Figure 1.5A), the most biomass-dense study area (Table 1.1). While the fit varied by scale and GNN method (Figure 1.5A), the prediction was consistently low (range of mean error, i.e., ME: -316 to -328 Mg ha^{-1}). The match was much better at the other five study areas (Figure 1.5B). There was a clear pattern of improving correlation with the more complex imputation methods and increasing pixel size (Figure 1.5). In terms of GNN method, GNN_{k10} results had the highest correlation (r_{mean} = 0.61) with LiDAR values at all sites while GNN_{k1} had the least (r_{mean} = 0.49). Interestingly the magnitude of bias, as measured by the absolute value of the ME, did not vary by imputation method or scale.

The results from the reduced major axis regression analysis (RMA) confirmed the tendency for LT-GNN to under-predict AGB at fine spatial scales (Figure 1.6). The median slope across all sites, scales, and methods (n = 54) with LT- GNN AGB on the y-axis and LiDAR AGB on the x-axis was 0.81. For three sites (Blodgett, Garcia, and Sugar Pine), there were combinations of scale and method that produced slopes where 1 was in included in the 95% confidence intervals. For the remaining sites (Mailliard, North Yuba, and Last Chance), the largest slope was less than 1 for all combinations. Despite better fits to the RMA regression, GNN_{k10} results were not among the best performing models, defined as the combination of scale and method that produced the slope closest to 1. The best models were based on GNN_{k1} (four sites) and GNN_{ba} (two sites).

The discrepancy in performance was driven by the presence of high biomass areas (Figure1. 7). The best LT-GNN model for Blodgett (Figure 1.6A) and Sugar Pine (Figure1.6C) matched the LiDAR results. These two comparisons had few LiDAR AGB estimates > 1,000 Mg ha^{-1}. In contrast, the best LT- models (i.e., RMA slopes closest to 1) for Mailliard (Figure 1.6B) and Last Chance (Figure 1.6D) under-predicted LiDAR AGB. The main driver was lower estimates of AGB for the more biomass-dense pixels (i.e., LiDAR AGB > 1,000 Mg ha^{-1}).

This plateau in the LT-GNN AGB estimate occurred across all sites. For example, while the best LT-GNN model for Blodgett has a RMA slope close to 1 (Figure 1.6A), there were clearly locations at Blodgett where LT-GNN under-predicted AGB relative to LiDAR (Figure1. 7). The mismatch at Mailliard (Figure 1.8) was simply magnified by the abundance of high biomass areas.

Three of the study areas were organized into stands defined by consistencies in forest composition, topography, and management. These stands provided a meaningful spatial scale for evaluation. At the stand scale, the results paralleled the pixel-level findings although the differences in fit and slope among the three GNN methods were smaller while ME was larger (Figure1.9).

The stand level analysis highlighted an apparent contradiction in the results for Last Chance where the slope of the regression was < 1 but the ME was positive. This difference in the metrics was due to a counter balancing in the match between LT-GNN and LiDAR AGB that was present in the fine-scale results (e.g., Figure 1.5, Figure1.6D) but can be more clearly seen at the stand level (Figure 1.10). The slope of the fit was reduced by the tendency of LT-GNN to underestimate AGB in high biomass areas. The relatively fewer points near the high end of the range dragged down the overall slope. At the same time, LT-GNN overestimated AGB relative to LiDAR in the low biomass areas. Although the LT-GNN underestimate was of smaller

magnitude, there were many more stands at lower AGB (Figure 1.10) that, when aggregated, resulted in a positive estimation bias at the stand level.

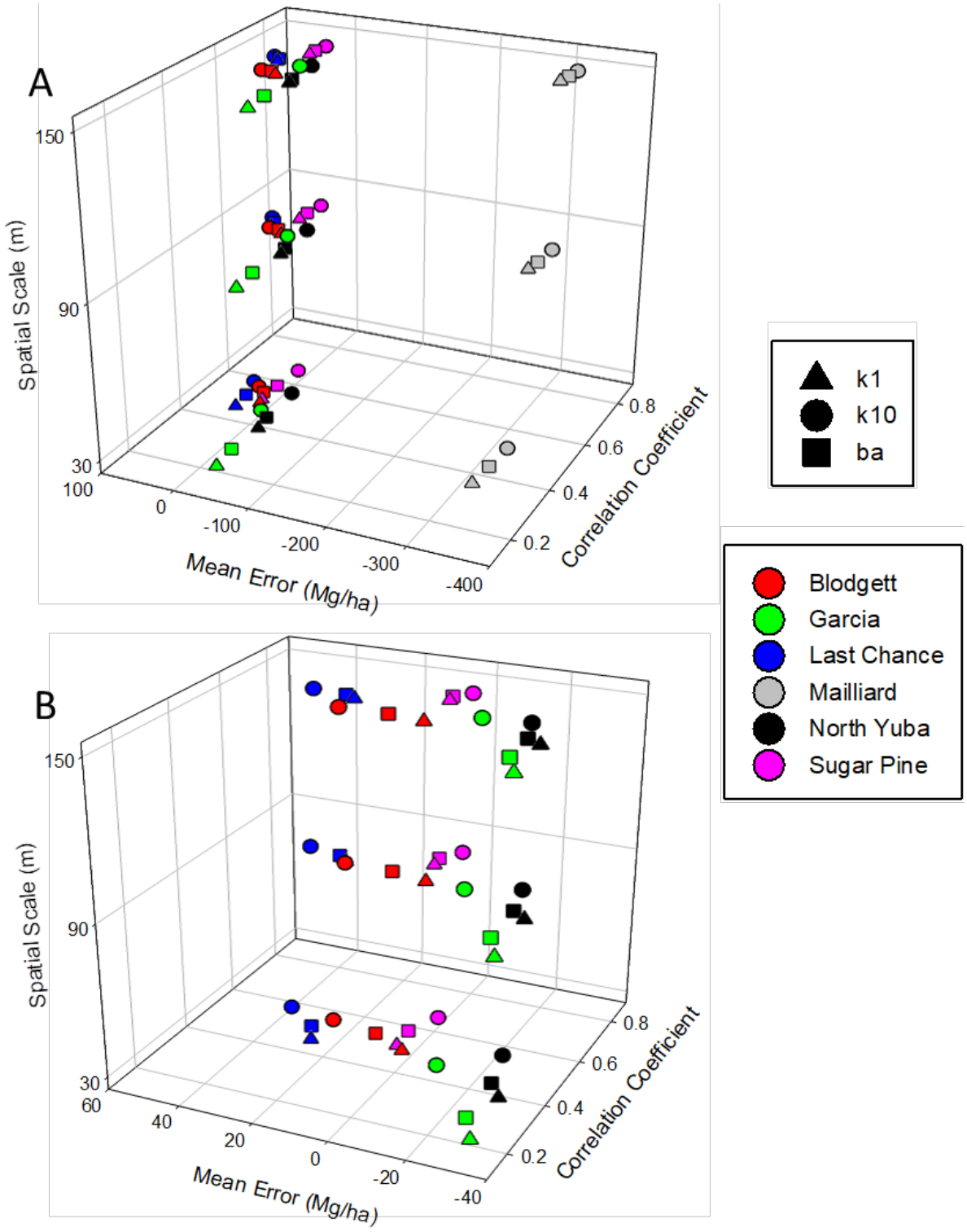


Figure 1.5: Summary of the LT-GNN vs LiDAR estimates of AGB as a function of correlation coefficient (r), mean error, spatial scale, and method of GNN interpolation. A) Performance for all six study areas. B) Results with Mailliard excluded.

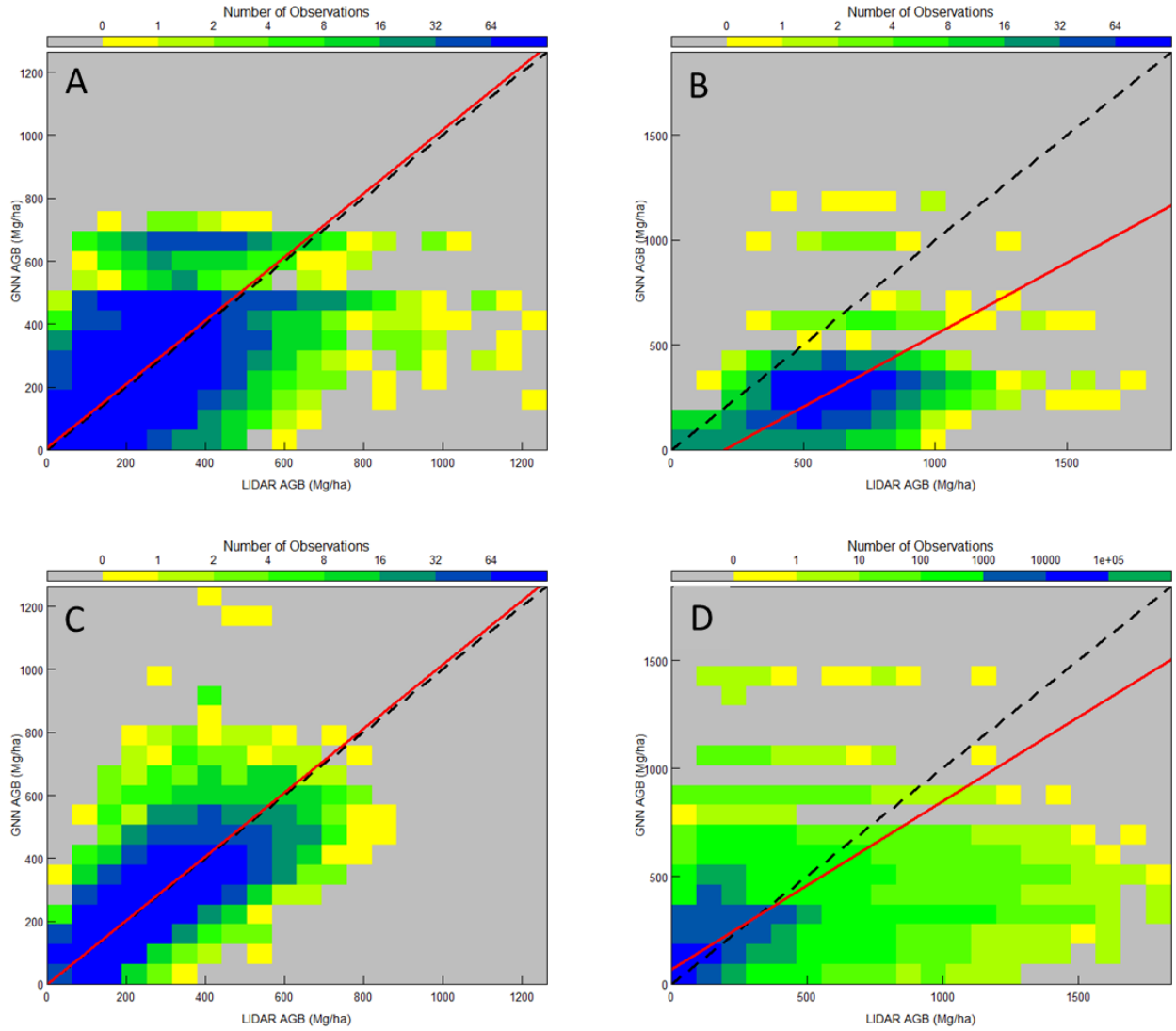


Figure 1.6: Relationship between LT-GNN and LiDAR estimates for the best performing GNN models as defined by the slope of the reduced major axis regression (RMA). The solid red line is the RMA line and the black dotted line is the 1:1 fit. A) For Blodgett, scale = 30 m and method = GNN_{ba}. B) For Mailliard, scale = 30 m and method = GNN_{k1}. C) For Sugar Pine, scale = 150 m and method = GNN_{ba}. D) For Last Chance, scale = 30 m and method = GNN_{k1}.

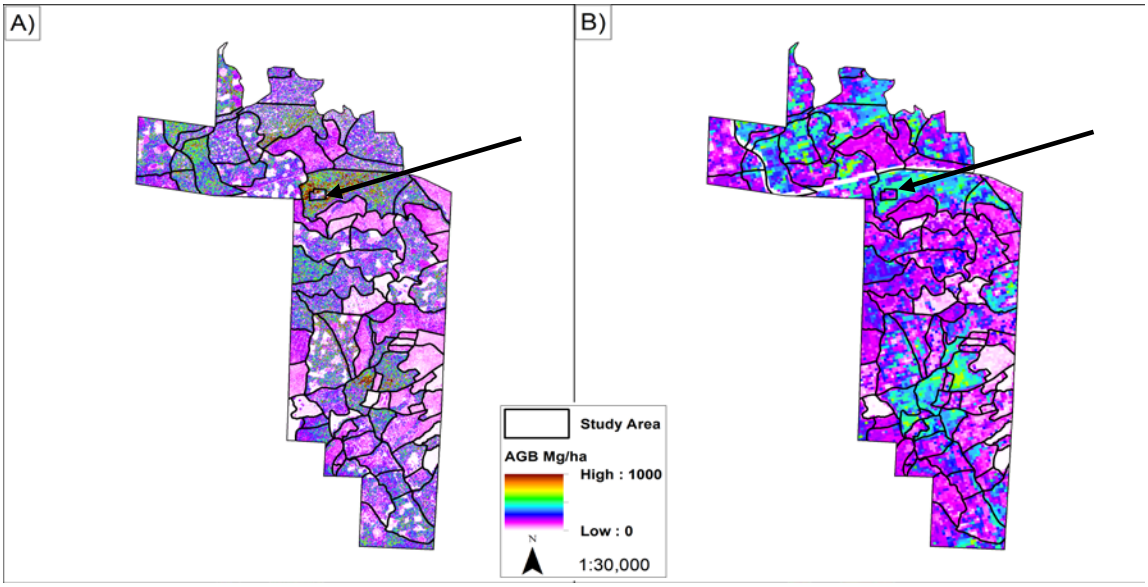


Figure 1.7: Comparative maps of LT-GNN and LiDAR AGB estimates for Blodgett. A) LiDAR based AGB projections. B) Best performing LT-GNN projection of AGB (scale = 30 m; method = GNN_{ba}). Note irregular northern boundary due to gaps in the LiDAR datasets. The AGB color scale is consistent between the two maps. The arrows depict an area of high biomass.

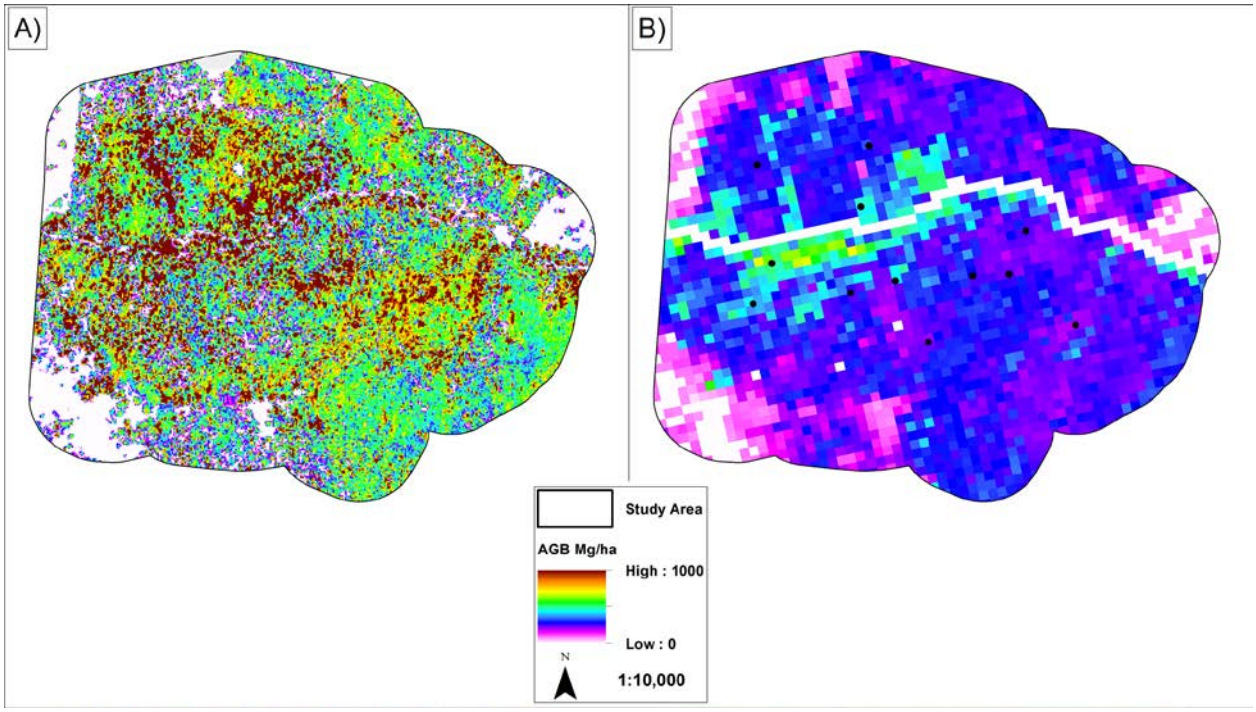


Figure 1.8: Comparative maps of LT-GNN and LiDAR AGB estimates for Mailliard. A) LiDAR based AGB projections. B) Best performing LT-GNN projection of AGB (scale = 30 m and method = GNN_{k1}). The AGB color scale is consistent between the two maps.

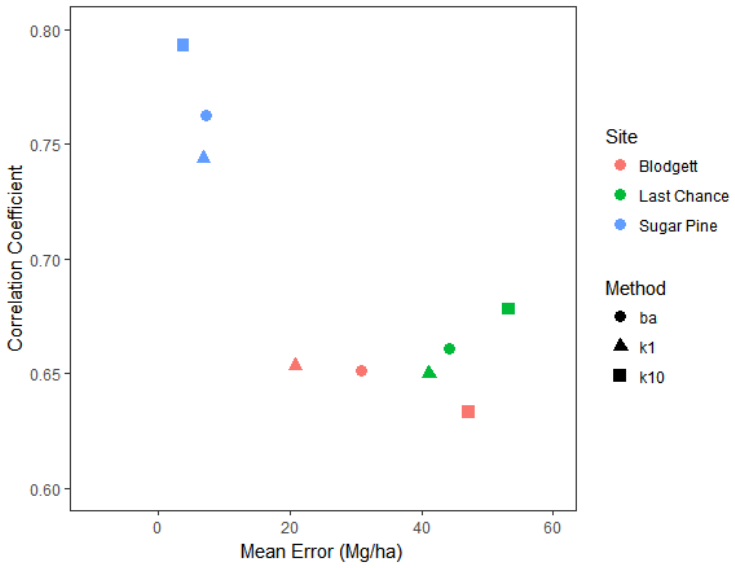


Figure 1.9: Summary of the correlation coefficient (r) and mean error between LT-GNN and LiDAR estimates of AGB for three study areas with well-defined stands. The scale of the analysis defined the area of the stands. At Blodgett, stands ranged from 1 to 45 ha in area; at Last Chance from 1 to 73 ha; and at Sugar Pine from 2 to 302 ha. To calculate the LT-GNN AGB, the base pixel size (i.e., 30 m) was used.

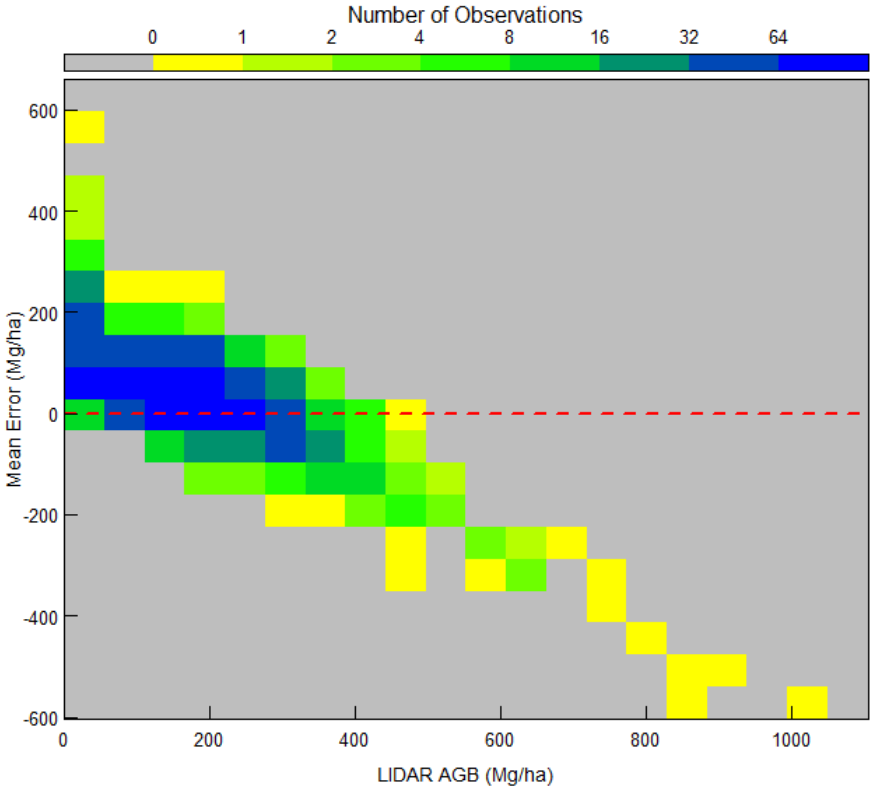


Figure 1.10: Trends in mean error with LiDAR AGB for Last Chance. LT-GNN$_{k1}$ at a spatial scale = 30 m was used to calculate stand-level AGB (same as Figure 1.6D). Dotted red line marks the zero line.

Table 1.7: Site-level comparisons of AGB between LiDAR and LT-GNN estimates. To calculate GNN AGB, GNN_{k10} method was used at a spatial scale = 30 m. Values in parentheses indicate RMSE for LiDAR and standard deviation for LT-GNN.

Site/Date	LiDAR (Mg ha^{-1})	LT-GNN (Mg ha^{-1})	ME (Mg ha^{-1})
Garcia 2005	225 (9)	219 (7.9)	-6
Mailliard 2005	596 (24)	273 (20.5)	-323
North Yuba 2005	272 (41)	234 (6.4)	-38
Last Chance 2012-13	194 (27)	173 (9.8)	-21
Blodgett 2014	230 (37)	257 (14.7)	27
Sugar Pine 2007	269 (30)	267 (7.4)	2

We also compared AGB estimates for the entire study sites (Table 1.7). Overall, LT-GNN produced results with greater precision. The coefficient of variation (CV = standard deviation/mean) for LT-GNN ranged from 2.8% (Sugar Pine 2007) to 5.7% (Blodgett 2014); the relative root mean square error (rRMSE = RMSE/mean) for LiDAR ranged from 4% (Garcia 2005) to 16% (Blodgett 2014). In terms of accuracy, the mean error between LiDAR and LT-GNN was comparable in magnitude to the precision in the LiDAR estimate with the one exception being Mailliard 2005.

1.3.2 County-level Assessment of LT-GNN

County-level estimates of mean and total AGB reflected forest dominance patterns across the study area. Generally speaking, both mean and total AGB were greatest in northwestern CA and the Sierra Nevada (Figure 1.11) where the proportion of land classified as forest was also greatest (Figure 1.4). This pattern likely reflects the increasing dominance of closed canopy coniferous forests composed of large conifers in these regions as opposed to more sparsely distributed and shorter stature woodland tree species associated with discontinuous forest landscapes.

The standard deviations in mean AGB estimates were greatest along coastal California, especially in the Bay Area (Figure 1.12A). In contrast, the coefficients of variation were not as consistently localized (Figure 1.12B) and were negatively related to the forest area within a county (Figure 1.13). This correlation between CV and forest area likely reflects small FIA sample sizes in counties with little forest area. Note that the CV exceeded 5% in only a handful of counties, indicating generally high precision for county-level estimation.

At the county-level, LT-GNN AGB paralleled FIA values (Table 1.8, Figure 1.13). In fact the RMA regression produced a close 1:1 fit (Figure 1.13) with an overall mean error equal to 0.7 Mg ha^{-1}. However, the two estimates did diverge in some instances, even for counties with extensive forests. For example, LT-GNN underestimated AGB for Madera (2,590 km^2 of

forestland, ME = - 29.3 Mg ha^{-1}) and overestimated ABG for Shasta (7,549 km^2, ME = 24.1 Mg ha^{-1}).

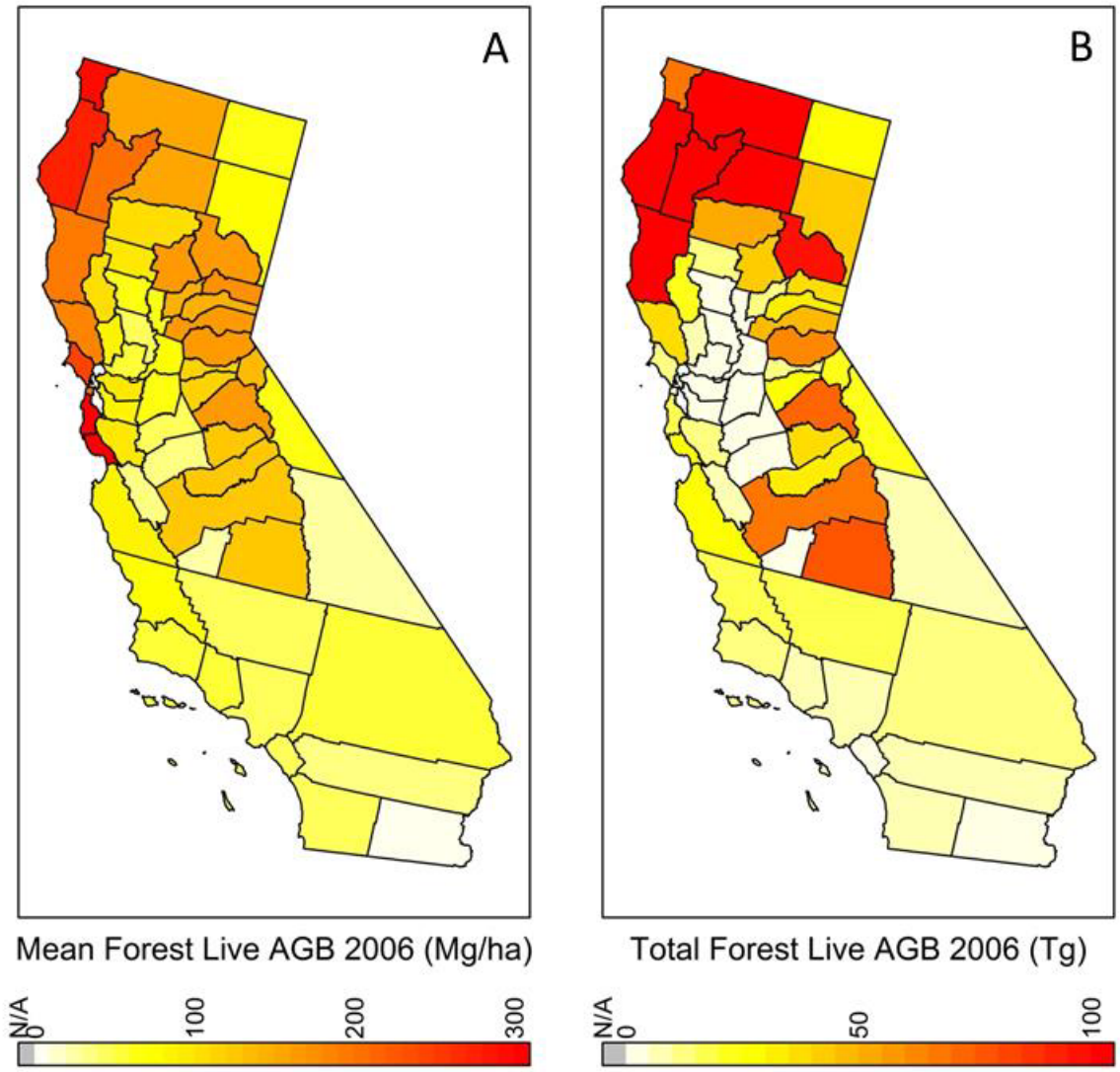


Figure 1.11: Distribution of AGB in California by county from LT-GNN. Maps of (A) mean live AGB in forests in megagrams per hectare (Mg/ha) and (B) total live AGB in forests in 2016 in teragrams (Tg).

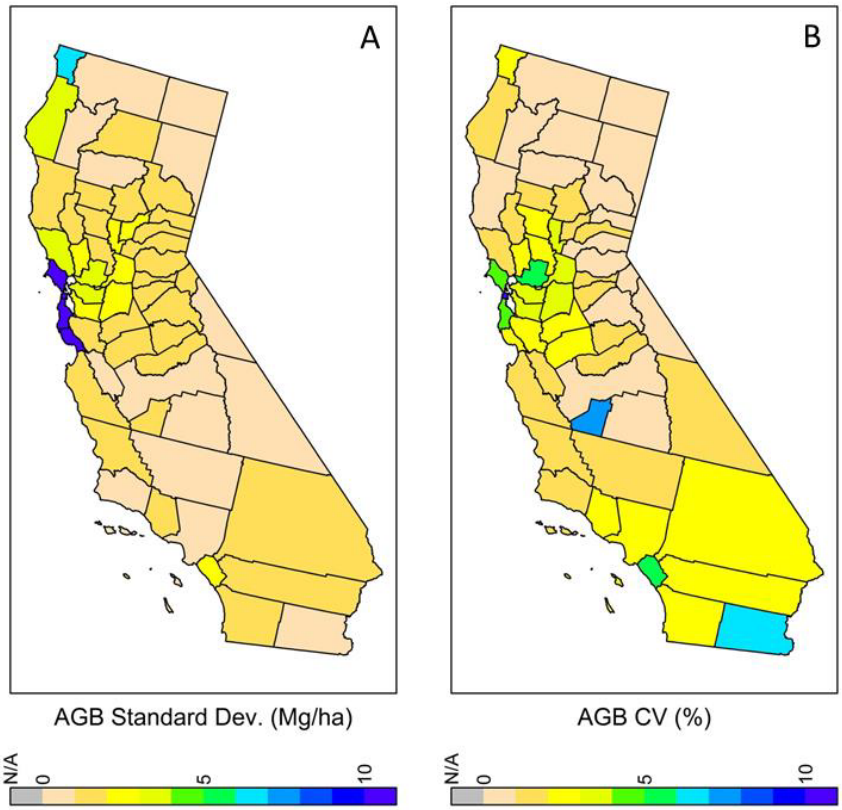


Figure 1.12: Performance of LT-GNN-based estimates of AGB by county in California. Map of (A) standard deviation of AGB in forests in 2006 (Mg ha^{-1}) and (B) coefficient of variation of AGB in forests in 2006.

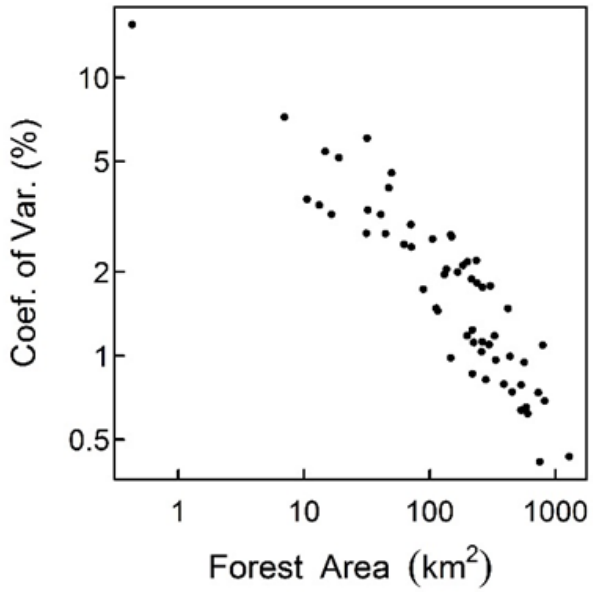


Figure 1.13: Log-log representation of the relationship between county-level coefficient of variation for live forest AGB and the area of the county that is forested.

Table 1.8: County-level GNN_{k10} estimated means (and standard deviations) for AGB, FIA plot-based estimates, and the mean error. GNN_{k10} estimates were produced for the year 2006. FIA plot-based estimates are the mean AGB of forested plots for each county from plots measured 2001-2010. "NA" refers to counties with no FIA plot measurements.

County	GNN_{k10} (Mg ha^{-1})		FIA (Mg ha^{-1})	Mean Error (Mg ha^{-1})
Alameda	86	(2.8)	79	6.6
Alpine	139	(1.4)	119	19.9
Amador	117	(1.7)	121	-3.9
Butte	161	(1.7)	143	17.7
Calaveras	117	(1.5)	112	5.3
Colusa	60	(1.5)	48	12.0
Contra Costa	99	(3.3)	95	3.9
Del Norte	291	(6.4)	297	-5.8
El Dorado	163	(1.3)	166	-3.0
Fresno	129	(0.8)	140	-11.0
Glenn	95	(1.4)	125	-29.6
Humboldt	277	(3.0)	275	1.6
Imperial	5	(0.3)	11	-5.5
Inyo	28	(0.3)	26	2.5
Kern	47	(0.7)	45	2.3
Kings	24	(1.8)	NA	NA
Lake	109	(1.3)	105	4.3
Lassen	76	(0.7)	75	1.2
Los Angeles	48	(1.0)	42	6.2
Madera	130	(1.5)	159	-29.3
Marin	249	(11.3)	256	-7.0

Table 1.8 (continued)

County	GNN_{k10} (Mg ha^{-1})		FIA (Mg ha^{-1})	Mean Error
Mariposa	120	(1.3)	126	-5.8
Mendocino	200	(1.5)	201	-1.1
Merced	39	(1.1)	41	-2.5
Modoc	61	(0.6)	65	-3.8
Mono	61	(0.6)	50	11.1
Monterey	88	(1.6)	71	16.6
Napa	87	(2.3)	130	-43.2
Nevada	149	(1.7)	138	11.4
Orange	49	(2.6)	15	33.6
Placer	165	(1.4)	164	1.0
Plumas	165	(1.0)	165	0.5
Riverside	40	(1.1)	52	-12.3
Sacramento	72	(2.5)	47	24.7
San Benito	39	(0.8)	46	-6.7
San Bernardino	56	(1.2)	68	-11.5
San Diego	42	(1.1)	37	5.0
San Francisco	204	(31.7)	NA	NA
San Joaquin	66	(2.1)	57	9.5
San Luis Obispo	77	(1.4)	66	10.9
San Mateo	323	(13.0)	268	55.0
Santa Barbara	55	(1.0)	43	11.5
Santa Clara	94	(1.9)	77	17.3
Santa Cruz	361	(10.7)	403	-42.0
Shasta	152	(1.0)	128	24.1
Sierra	172	(1.5)	171	1.4

Table 1.8 (continued)

County	GNN_{k10} (Mg ha^{-1})		FIA (Mg ha^{-1})	Mean Error
Siskiyou	159	(0.7)	151	8.2
Solano	59	(3.0)	81	-22.0
Sonoma	185	(3.5)	209	-24.3
Stanislaus	40	(1.0)	47	-6.9
Sutter	61	(2.2)	62	-1.5
Tehama	102	(0.8)	99	2.9
Trinity	210	(0.9)	205	4.5
Tulare	130	(0.9)	143	-13.0
Tuolumne	161	(1.2)	160	0.5
Ventura	52	(1.1)	39	12.6
Yolo	44	(1.2)	47	-2.9
Yuba	145	(2.5)	158	-12.9

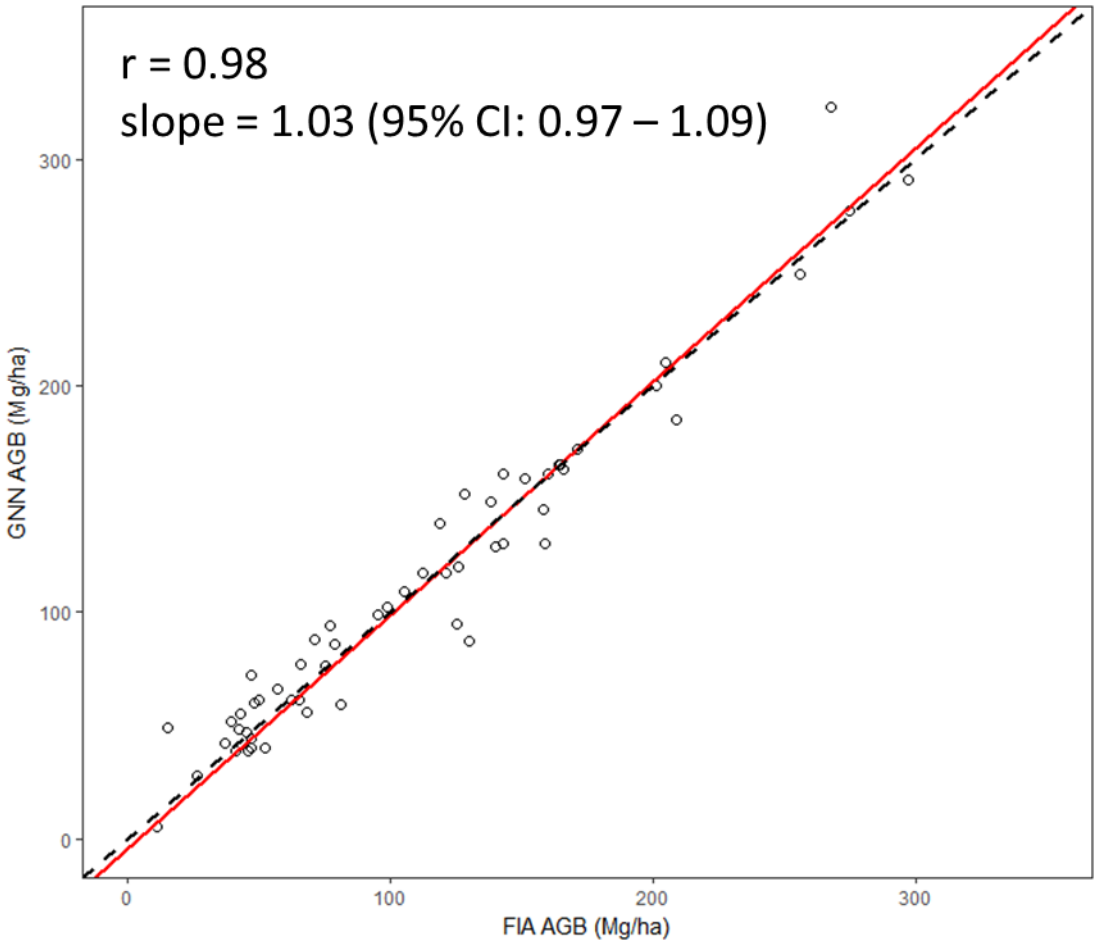


Figure 1.13: Performance of LT-GNN estimates of AGB by county in California compared to FIA values. The black dotted line represents the 1:1 line; the solid red line is the slope of the reduced major axis regression.

1.3.3 Change-detection Assessment of LT-GNN

Overall, LT-GNN tended to underestimate plot-level increases in AGB (Figure 1.14). Most changes were small (near 0, Figure 1.14) and in this region, LT-GNN slightly over-predicted losses and slightly under-predicted gains. These trends were exaggerated at the extremes. LT-GNN detected losses associated with disturbance the best. Indeed, it measured AGB change with no bias (mean error = 0 Mg ha-1) for the disturbed plots (Figure 1.15a). In contrast, LT-GNN underestimated AGB (mean error = -16 Mg ha-1) in stable plots (Figure 1.15c).

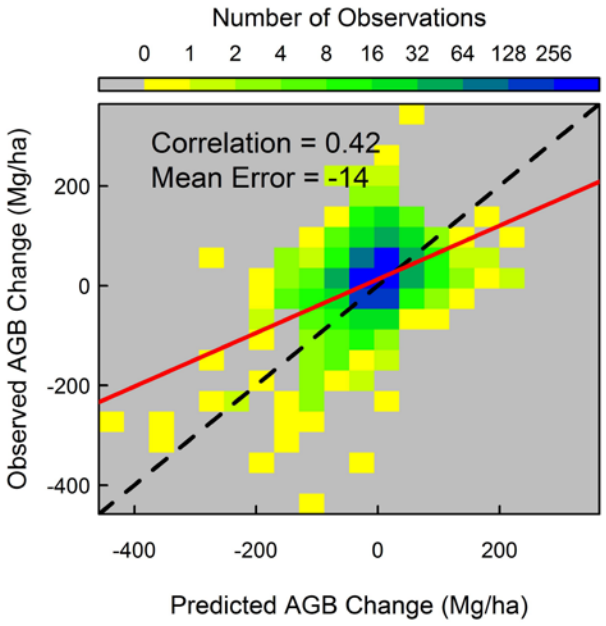


Figure 1.14: Predicted (LT-GNN) vs. observed (FIA) changes in AGB change for 3,412 repeat-measured forest plots. LT-GNN results based on GNN_{k10} imputation at 30-m scale. The dashed line is the 1:1 line. The solid red line is an ordinary least squares regression line between predicted and observed.

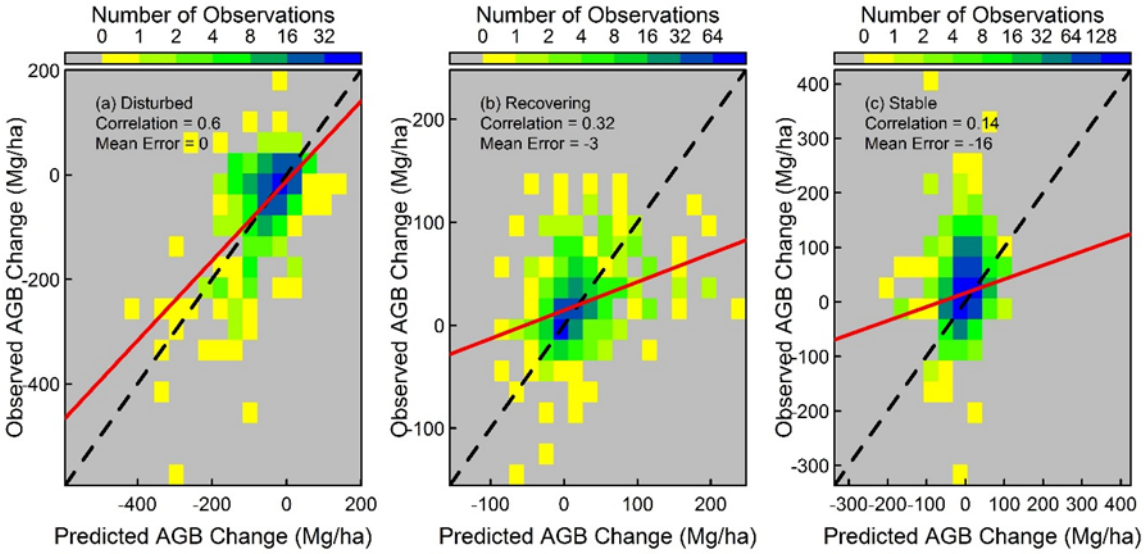


Figure 1.15: Predicted (LT-GNN) vs. observed (FIA) changes in AGB change for three different scenarios: (a) disturbed plots (disturbance noted between measurements), (b) recovering plots (disturbance before first measurement and after 1985), and (c) stable plots (no disturbance observed). For this analysis, only short-term (1-3 year) disturbance events observed with LandTrendr were considered. The dashed line is the 1:1 line. The solid red line is an ordinary least squares regression line between predicted and observed. LT-GNN results based on GNN_{k10} imputation at 30-m scale.

1.4 Conclusions and Future Directions

Woodall et al. (2015) noted the need for annual monitoring of forest carbon at the national level. California shares this policy goal of robust monitoring of forest carbon (Forest Climate Action Team 2018). Recently Kennedy et al. (2018) developed a method that relies on a mix of field data, statistical modeling, and remotely-sensed time-series imagery to track aboveground live tree biomass (AGB) at policy-relevant spatial and temporal scales. As part of California's Fourth Climate Change Assessment, we continued the development of Kennedy et al.'s (2018) approach and tested its performance in California's diverse and extensive forests.

At finer spatial scales, LT-GNN's ability to match LiDAR estimates of AGB varied by site, GNN method, and scale of analysis. Overall, LT-GNN tended to underestimate LiDAR AGB and most combinations of GNN interpolation and spatial scale (i.e., 30 meters, 90 meters, and 150 meters) did not result in a 1:1 relationship. Of the 54 possible combinations, only three returned a 1:1 relationship. Moreover, there were no strong indicators of superior imputation method or spatial scale. While all LT-GNN results at the 150-meter scale were more highly correlated, the high correlation did not necessarily translate into smaller mean errors (ME) or slopes closer to 1. GNN_{ba} tended to have the largest absolute ME relative to the other two methods. Kennedy et al. (2018) found similar "noisy" results at the pixel scale.

We extended the analysis to stands. These stands represented an intermediate scale between pixels and landscapes. Moreover, the structural consistency inherent in the definition of stands minimized the arbitrary variation that occurs when pixels are aggregated systematically. However, the larger size and structural consistency provided no obvious improvement in LT-GNN performance.

The stand-level analysis did detect an anomaly that suggests a potential area for improvement in the LT-GNN framework. At Last Chance, LT-GNN overestimated AGB in stands with low biomass density according to LiDAR (Figure 1.10). In plots with low AGB (< 150 Mg ha^{-1}) at Last Chance, the median shrub cover was 44% (interquartile range: 21% - 64%). Landsat has trouble differentiating tall, dense shrubs from small stands of trees. Shrub fields have 0 AGB by definition in the LiDAR map since only trees with DBH ≥ 5cm were included in the biomass calculations (Appendix D). Thus, any shrubs that were misclassified as trees by Landsat would skew the LT-GNN interpolation toward higher AGB. The LT-GNN overestimate of AGB in low biomass areas was most prevalent at Last Chance (Figure 1.6D), a site with an abundance of shrubs. In contrast, the problem was much less prevalent at Blodgett (Figure 1.6A) – a managed forest where shrub cover is actively suppressed.

At the landscape scale, LT-GNN results matched the LiDAR projections of AGB except for the highest biomass site, the old-growth coast redwood forest at Mailliard Redwood State Natural Reserve. In general, LT-GNN struggles to capture extremes. For example, Kennedy et al. (2018) noted that LT-GNN estimates tended to saturate at 650 Mg ha^{-1} when applied to forests in the Pacific Northwest. Contributing to the problem is the well-established saturation relationships between Landsat multi-spectral data and AGB (e.g., Steininger 2000). Since LT-GNN includes Landsat time-series as part of its gradient analysis, it shares this limitation although the inclusion of additional predictors (Ohmann et al. 2012) mitigates this problem. In addition, the range of possible values in a GNN interpolation routine is constrained by the input data. LT-GNN AGB relied on FIA plot results. Within California, Mailliard represents an extreme in terms of AGB. The mean biomass density at the site was 596 Mg ha^{-1}. At the 30-m scale, 3.6% of

the pixels (76 pixels) exceeded 1,000 Mg ha^{-1}. In contrast, only 0.2% (17 plots) of the FIA plots inventoried between 2000 and 2016 had more than 1,000 Mg ha^{-1} (FIA 2017). Faced by the dual challenges of a saturated Landsat signal and a sparse sample of extreme values, GNN failed to accurately estimate AGB at Mailliard. However, Mailliard is also a spatial isolated in that it is a small reserve of old-growth redwoods, one of the most biomass-dense forest types in the world (Busing and Fujimori 2005), in a location dominated by less massive second-growth forests.

We treated LiDAR AGB as the baseline but acknowledge that LiDAR maps of AGB are also an interpolation of field values. Any faults in the LiDAR maps would be expressed as error in the LT-GNN evaluation. Mitigating these concerns was the fact that despite differences among sites (Table 1.2), all six models ranked among the better performing models for LiDAR biomass estimates in temperate conifer forests (Zolkos et al. 2013). We were also careful to screen for mistakes and outliers in our biomass and LiDAR modeling (Appendix D) improving confidence in the LiDAR baseline (sensu Huang et al. 2017).

At the county-level, LT-GNN estimates of AGB were both precise and accurate compared to FIA results. There were no significant differences in the slope of the regression line comparing AGB from LT-GNN to AGB from FIA plots. Moreover, these county-level estimates were precise. The greatest uncertainty (CV > 5 %) was observed in sparsely forested locales (Figure 1.13). However, in counties with more than 100 km^2 of forest, CV was ≤ 2%. At this scale, GNN is acting to weight FIA plot data from in and outside of a county that best represents the variation in the Landsat data. Thus, it was not surprising that forest-rich counties were well-represented. Our current analysis focused on plots with at least 50% forest area (sensu Ohmann et al. 2014). If we were to relax this constraint and include all plots with some forest land, as FIA does for the purposes of estimation, variability in estimates would likely decline.

On a county-by-county basis, there were some appreciable differences between LT-GNN and FIA results (Table 1.8). One limitation in the current analysis was the handling of the temporal differences in the comparison. To obtain a robust AGB for a county required the inclusion of FIA samples from multiple years. Thus the FIA value for a county was averaged over 10 years while the LT-GNN value was taken at the temporal midpoint. Although impractical for an annual monitoring system, more explicit year-to-year matching of predicted to observed data would likely improve the individual county performances.

A major feature of the LT-GNN forest biomass monitoring system is its potential to track annual changes in AGB (Kennedy et al. 2018). While LT-GNN results were positively correlated to FIA results (Figure 1.14), LT-GNN underestimated AGB increments. Given the average interval between FIA re-measurements (8 years), the annualized underestimate was approximately 1.75 Mg ha^{-1}. Plot-level tests do present the greatest challenge to change detection using LT-GNN. For example, the LT-GNN pixel (30-m) does not account for structural heterogeneity in the FIA plots, a heterogeneity that decreases accuracy (Ohmann et al 2014). Also, the accuracy of LT-GNN estimates improved at coarser-scales like landscapes (Table 1.7) and counties (Figure 1.13). Such spatial averaging would likely limit extreme values where the deviance between predicted and observed AGB was greatest and thereby improve the comparison.

The change-analysis by forest status (Fig 1.15) clearly illustrated the strengths and weaknesses of the LT-GNN approach. It detects AGB losses with remarkable fidelity even at the plot-level (Figure 1.15a). Given that LT-GNN was designed to track disturbances (Kennedy et al. 2010), its success at detecting losses is not surprising. It also does a reasonable job measuring increases in

AGB in plots recovering from disturbance (Figure 1.15b). Both scenarios typically provide strong spectral signatures of change below the saturation threshold. In contrast, LT-GNN did a poor job measuring change in stable forests (Figure 1.15c). Clearly stable scenarios include intact and mature forests that saturate the Landsat spectral indices as discussed above. This underestimate of growth in mature forests was shared by the Landsat-based system used by Gonzalez et al. (2015). Also, changes are small by definition in the stable scenario. In general, small changes are more difficult to detect. Yet capturing them is crucial to AGB monitoring since incremental changes are the most common occurrence.

Other projects have created forest biomass maps for the continental United States. Although these provide only a snapshot of biomass in a single year, they serve as useful benchmarks for performance. Four national-scale biomass maps are in common use; a thorough comparison of the four maps showed that they disagree at the fine scale but agree in aggregate (Neeti and Kennedy 2016). Of the four national-scale maps, two are of a resolution commensurate with our maps: 1) a 100-m resolution map from the NASA Carbon Monitoring System (Hagan et al. 2016). and 2) a 30-m resolution map from the North American Carbon Program's National Biomass and Carbon Dataset for the year 2000 (NBCD 2000). Thus, we compared the CMS and NBCD maps to the LT-GNN maps. For each, we used the LT-GNN map from the nominal year of the national map: 2005 for CMS and 2000 for NBCD. All maps were clipped to the same footprint, and the 100-m CMS product sampled to 30 meters using a nearest-neighbor resampling. For each paired comparison, no-data values in either map were excluded. We compared the national maps to our maps using reduced major axis regression. We then aggregated the maps to 90x90 meter and 150x150 meter pixels and compared each (Figure 1.16).

LT-GNN maps were well correlated with national results ($r \geq 0.74$) in all cases, but national maps saturated at values well below the high AGB observed in California. The relationships between LT-GNN and national maps were relatively unbiased up to approximately 800 Mg ha^{-1}. Higher values present in California's forests were not captured by the national maps.

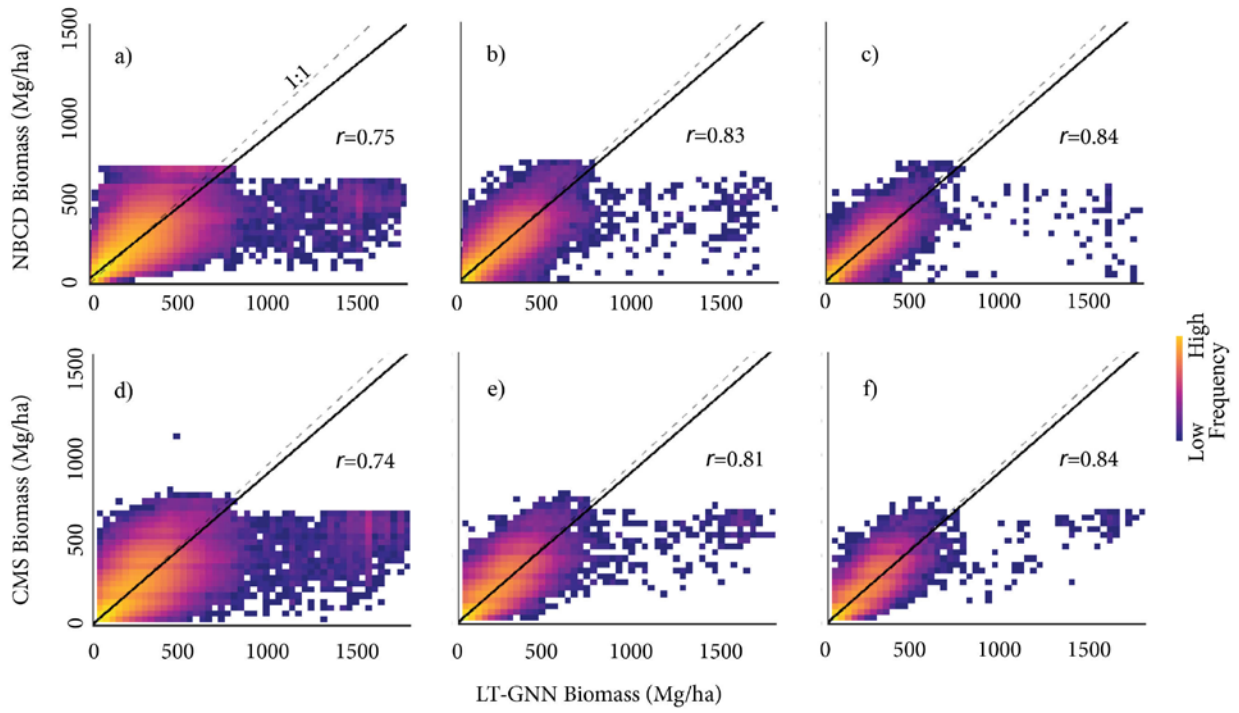


Figure 1.16: Comparing national scale biomass maps to maps derived from LT-GNN for California. a) Density plot of National Biomass and Carbon Dataset (NBCD) vs AGB derived from LT-GNN (scale = 30 m, GNN_{k1}). Dashed line represents a 1:1 relationship; solid line the reduced major axis regression and associated Pearson's correlation (r). Only pixels with non-zero biomass were compared. (b) and (c) as for (a), but at 90 m and 150 m resolution respectively. (d-f) As for (a-c) but for the NASA Carbon Monitoring System (CMS) national biomass product.

LT-GNN is a promising approach to monitor a large pool of carbon in California, namely aboveground live tree biomass. The successful transfer of the LandTrendr algorithm to Google Earth Engine (Appendix B) greatly reduces the costs of processing Landsat images. The success of LT-GNN in interpolating county-level AGB suggests an application-ready means to track annual AGB at a policy-relevant scale. However, key improvements are necessary.

LT-GNN struggled to accurately represent AGB in California's most biomass-dense forests. It also underestimated gains in stable forest. These two weaknesses are related by limitations in the Landsat sensors. In terms of detecting biomass-dense forests, a greater concentration of field measurements in high biomass sites and/or refinement of the gradient analysis could improve the LT-GNN performance in this regard. More plots in extreme cases would provide a larger basis for imputation. In terms of gradient analysis, a temporal component could be added to the imputation model. Thus, in cases of stable forest pixels where there is no or little change in the Landsat indices, the most recently measured FIA plot would be selected. Alternatively, LT-GNN estimates for stable pixels could be calibrated directly via statistical correction of AGB increment or forest growth under stable conditions could be modeled.

Applications of biomass monitoring at finer scale applications must be pursued with more caution. LT-GNN's success in projecting AGB from the plot to the stand level was hit or miss. One near-term application suggested by our results was the ability of LT-GNN to detect AGB

loss (Figure 1.15a). Given the annual resolution, LT-GNN would provide the means to check for large losses in AGB at the project scale.

Despite progress, barriers to the implementation of a LT-GNN monitoring system include the time and expense associated with gradient nearest neighbor analysis (Appendix C). While the LandTrendr processing is automated, the gradient analysis, where new FIA inventory results are interpolated, still must be done "by hand." Other limitations include its current restriction to forest biomes and live vegetation. There is no theoretical reason why the LT-GNN protocol could not be applied to shrublands, but the necessary field data do not exist. The value of the FIA inventory to forest carbon monitoring in California and the United States cannot be overestimated. It provides the essential empirical foundation for all remotely sensed applications. The lack of systematic inventory of vegetation biomass in shrublands precludes development of precise monitoring in these biomes. The extent of drought-related tree mortality in California (USDA Forest Service 2017) prioritizes the need to monitor the fate of the estimated 129 million standing dead trees. The LT-GNN framework is well-suited to this task but would require a rethinking of the relevant indicators and a revision of the field sampling regime. Currently in California, FIA plots are re-measured every decade (PNW-FIADB 2015). Given the longevity of standing dead trees, plots would need to be assessed more frequently.

2: Innovations in Managing Forest Carbon

2.1 Introduction

> *The best way to manage forests to store carbon and to mitigate climate change is hotly debated.* Bellassen and Luyssaert 2014

Nowhere is the international debate noted by Bellassen and Luyssaert (2014) more intensely contested than in California. Forests are currently substantial contributors to California's efforts to reduce its greenhouse gas emissions (Christensen et al. 2017). However, the interaction between a warming climate and increased wildfire severity threatens the carbon carrying capacity of California's forests (Liang et al. 2017). The California Forest Carbon Plan, in recognition of this threat, outlines strategies to sustain the State's forests as reliable carbon sinks (Forest Climate Action Team 2018).

Managing forest carbon is constrained by gaps in our ecological understanding. Even the basic tenet that carbon accumulation in mature forests long-free from disturbance should reach a carrying capacity (Turner 2010) does not always hold (Luyssaert et al. 2008, Harmon and Pabst 2015). The global increase in CO_2 concentrations combined with the regional deposition of atmospheric nitrogen has contributed to increased productivity in some but not all forests. Other factors including warming and moisture stress influence productivity responses (reviewed in Camerero et al. 2015). The evidence from three old-growth conifer forests in the Sierra Nevada reflects this variability: Levine et al. (2016) reported steady increases in aboveground live biomass (AGB); van Mantgem and Stephenson (2007) reported stability in AGB; Wiechmann et al. (2015) noted recent declines in AGB. The relevance of this unpredictability to management is that the baseline, “no-management” trajectory of AGB is uncertain.

The dynamics of the carbon cycle are further complicated by more than a century of fire suppression in California's dry forests (Stephens and Ruth 2005). Among other changes, fire exclusion has led to the loss of large trees, a higher density of shade-tolerant trees, and increased fuel loads (Taylor et al. 2014, McIntyre et al. 2015). The magnitude and continuity of fuel exacerbate fire hazard (Agee and Skinner 2005) and have contributed to an altered fire regime characterized by low frequency but high intensity fires (Miller et al. 2009). To reduce this risk, improving resilience by lowering fuel loads, reducing tree density, and increasing heterogeneity has been recommended as a management response (e.g., North et al. 2009). While there is broad agreement that fuel treatments can reduce fire hazard, the consequences for carbon storage continue to be argued (Hurteau and North 2009, Campbell et al. 2012, Wiechmann et al. 2015). There clearly is a trade-off between the size of the carbon stock and its stability (sensu Hurteau and Brooks 2011) with more carbon literally adding "more fuel to the fire." But as Restaino and Peterson (2013) emphasized, the balance depends critically on the likelihood that the landscape experiences a wildfire, the fate of carbon removed during a treatment, the amount of carbon emitted during a fire, and the rate of carbon accumulation following a disturbance, be it a fire or a fuel treatment. Thus we contend that to manage forest carbon, we need to understand the joint trajectory of carbon accumulation and fire hazard under different treatment regimes. Recently these trade-offs have been productively explored for the Sierra Nevada using landscape modeling to evaluate the carbon consequences of simulated management (e.g., Krofcheck et al. 2017). However long-term field experiments that measure ecosystem responses to actual fire and fuel treatments are exceedingly rare given the complexity and cost of maintaining a measurement and management regime over many years. And yet, results from these field experiments are essential. They inform our understanding of carbon dynamics in fire-prone forests and provide vital reality checks on simulated projections.

In this chapter, we developed a means to quantify the carbon storage-carbon stability trade-off for fire-prone forests. We extended a long-term experiment at the Blodgett Forest Experiment Station in Georgetown, CA designed to evaluate the efficacy of fuel-reduction strategies (McIver et al. 2013). We combined carbon trajectories with periodic estimates of fire hazard (P-torch, Rebain 2010). With these joint trends, we tested the hypothesis that there is a trade-off between carbon storage and carbon stability for a productive, mixed conifer forest.

2.2 Methods

2.2.1 Overview

We extended a field experiment to gain insights into biomass dynamics in fire-prone forests. In 2016, we re-measured sites that were part of the national Fire and Fire Surrogate study in the Sierran mixed conifer forests (Stephens and Moghaddas 2005). Treatments, installed in 2001-2002, include both prescribed fire and mechanical thinning. These repeated field assessments were used to build joint trajectories of AGB and fire hazard in order to assess the trade-off between carbon stored and risk of loss to wildfire.

2.2.2 Study Site

This study was performed at the University of California Blodgett Forest Research Station (Blodgett Forest), approximately 20 km east of Georgetown, California. Blodgett Forest is located in the mixed conifer zone of the north-central Sierra Nevada at latitude 38° 54′ 45″ N,

longitude 120° 39′ 27″ W, between 1100 and 1410 m above sea level and encompasses an area of 1,780 ha (Figure 1.2). Tree species in this area include sugar pine (*Pinus lambertiana*), ponderosa pine (*Pinus ponderosa)*, white fir (*Abies concolor*), incense-cedar (*Calocedrus decurrens*), Douglas-fir (*Pseudotsuga menziesii*), California black oak (*Quercus kelloggii*), tanoak (*Lithocarpus densiflorus*), bush chinkapin (*Chrysolepis sempervirens*), and Pacific madrone (*Arbutus menziezii*).

Fire was a common ecosystem process in the mixed conifer forests of Blodgett Forest before the policy of fire suppression began early in the 20th century. Between 1750 and 1900, median composite fire intervals at the 9-15 ha spatial scale were 4.7 years with a fire interval range of 4-28 years (Stephens and Collins 2004). Forested areas at Blodgett Forest have been repeatedly harvested and subjected to fire suppression for the last 100 years reflecting a management history common to many forests in California and elsewhere in the Western US.

2.2.3 Fuel-reduction Treatments

The primary objective of the treatments was to modify stand structure such that 80% of the dominant and co-dominant trees in the post-treatment stand would survive a wildfire modeled under 80th percentile weather conditions. To meet this objective, three different treatments: mechanical only, mechanical plus fire, and prescribed fire only, as well as untreated control were each randomly applied (complete randomized design) to 3 of 12 experimental units that varied in size from 14 to 29 ha. Total area for the 12 experimental units was 225 hectares. To reduce edge effects from adjoining areas, data collection was restricted to a 10 ha core area in the center of each experimental unit (Stephens and Moghaddas, 2005).

Control units (Control) received no treatment during the study period (2001-2016). Mechanical only treatment units (Thin) had a two-stage prescription; in 2001 stands were crown thinned followed by thinning from below to maximize crown spacing while retaining 28 to 34 m^2ha^{-1} of basal area with the goal to produce an even species mix of residual conifers (Stephens and Moghaddas, 2005). Individual trees were cut using a chainsaw and removed with either a rubber-tired or track laying skidder. During harvests, some hardwoods, primarily California black oak, were coppiced to facilitate their regeneration. All residual trees were well spaced with little overlap of live crowns in dominant and co-dominant trees. Following the harvest, approximately 90% of understory conifers and hardwoods up to 25 cm DBH were masticated in place using an excavator mounted rotary masticator. Mastication shreds and chips standing small diameter live and dead trees in place and this material was not removed from the experimental units. The remaining un-masticated understory trees were left in scattered clumps of 0.04 to 0.20 ha in size.

Mechanical plus fire experimental units (ThinBurn) underwent the same treatment as mechanical only units, but in addition, they were prescribed burned using a backing fire. Fire only units (Burn) were burned with no pre-treatment using strip head-fires. All initial prescribed burning was conducted during a short period (10/23/2002 to 11/6/2002) with the majority of burning being done at night because relative humidity, temperature, wind speed, and fuel moistures were within pre-determined levels to produce the desired fire effects (Knapp et al. 2004). Prescribed fire prescription parameters for temperature, relative humidity, and wind speed were 0-10^0 C, > 35%, and 0.0-5 km hr^{-1}, respectively. Desired ten-hour fuel stick moisture content was 7-10%.

The initial treatments were part of the national Fire and Fire Surrogate (FFS) experiment designed to evaluate the impact of alternative in dry forest across the United States (McIver et al. 2013). We have continued and expanded the effort at Blodgett Forest beyond the original design. Not only have we repeatedly measured the plots, but we also extended the application of prescribed fire. In October 2009, the Burn units were burned a second time using the same prescription parameters as the initial-entry burns. The decision to re-burn these units was based on the objective to re-introduce fire at a frequency consistent with the historical range of variability (Stephens and Collins 2004). In the ThinBurn units the strong shrub response following initial-entry burns limited re-burning opportunities.

2.2.4 Vegetation Measurements

Overstory and understory vegetation was measured in twenty 0.04 ha circular plots, installed in each of the 12 experimental units (240 plots total) in 2001 (PRE), 2003 (POST-1YR), 2009 (POST-7YR), and 2016 (POST-14YR). Individual plots were placed on a systematic 60m grid with a random starting point. Plot centers were permanently marked with a pipe and by tagging witness trees to facilitate plot relocation after treatments. Tree species, DBH (diameter at breast height, breast height = 1.37m), total height, height to live crown base, and crown position (dominant, co-dominant, intermediate, and suppressed) were recorded for all trees ≥ 11.4 cm DBH. Similar information was recorded for all understory trees (trees >1.37 m tall and < 11.4 cm DBH) on 0.004 ha subplots nested in each plot. Canopy cover was measured using a 25 point grid in each 0.04 ha plot with a site tube.

2.2.5 Fuel Measurements

Surface and ground fuels were sampled with two random azimuth transects at each of the 240 plots using the line-intercept method on the same schedule as the vegetation measurements. A total of 480 fuel transects were installed and the same azimuths were used here as done in the original measurements in 2001 (Stephens and Moghaddas 2005). One-hour (0–0.64 cm) and 10-h (0.64– 2.54 cm) fuels were sampled from 0 to 2 m, 100-h (2.54–7.62 cm) fuels from 0–3 m, and 1000-h (>7.62 cm) and larger fuels from 0 to 11.3 m on each transect. Duff and litter depth in cm were measured at 0.3 and 0.9 m on each transect (same points as done previously).

2.2.6 Biomass and Fuel Load Calculations

Aboveground live tree biomass (AGB) was calculated from tree inventory data (species, DBH, and height) using regional biomass equations (FIA 2010) as described in Appendix D. Only live trees ≥ 5 cm DBH were included in the analysis. Surface and ground fuel loads were calculated using appropriate equations developed for California forests. Coefficients required to calculate all surface and ground fuel loads were arithmetically weighted by plot basal area fraction to produce accurate and precise estimates of ground and surface fuel loads (Stephens and Moghaddas, 2005).

2.2.7 Fire Modeling

We modeled potential fire behavior for each inventory plot, at each time step, with the Fire and Fuels Extension (FFE) to the Forest Vegetation Simulator (Reinhardt and Crookston 2003). FFE uses established equations to predict fire behavior and crown fire potential based on user-input tree lists and fire weather (Rebain 2010). We used the conditions during large spread events in two nearby wildfires (2001 Star Fire, 2008 American River Complex) for the weather inputs. By

using actual conditions from nearby wildfires that posed substantial fire control problems we believe predicted fire behavior may better characterize wildfire potential as opposed to using conditions based on fire-weather percentile thresholds.

We varied surface fuel models by treatment type and time period, assigning a "low" and "high" fuel model or each combination. These fuel model assignments were based on measured plot fuel loads and woody shrub cover for each treatment/time period, as well as observed fuelbed characteristics in the field. We modeled fire behavior for each plot, at each time period combination using for both the "low" and "high" surface fuel models.

Our analysis focused the output torching probability (Ptorch) from FFE. The calculation of Ptorch first involves randomly populating 0.01 ha sub-plots from the stand tree list using a Monte Carlo simulation. For each sub-plot FFE computes the surface fire flame length that would be required to cause torching. Next the program computes the height above the ground that the predicted surface fire can ignite crowns based on discussion in Scott and Reinhardt (2001, p13). The torching probability is based on whether this predicted height exceeds the flame length needed to ignite tree crowns. Rebain (2010) explains that torching probability is the proportion of stand area where crowns of larger trees can be ignited by surface fire or flames from burning crowns of small trees. Rebain (2010) argues that this index may better characterize hazard due to torching compared to the conventionally used torching index (see Scott and Reinhardt 2001), due to the lack of dependence on the problematic calculation of canopy base height. We reported Ptorch as the mean of the results from the low and high surface fuel models.

2.2.8 Stable Aboveground Live Biomass

To account for potential losses in AGB due to wildfire, we defined stable aboveground live biomass (SAGB) as:

$$SAGB = AGB * (1 - Ptorch) \qquad \text{Equation 2.1}$$

In effect, we discounted AGB by the probability of it surviving a problematic fire. For example, in a plot with 100 Mg ha^{-1} of AGB and the probability of torching is 0.3, SAGB equals 70 Mg ha^{-1}.

To assess the impact of treatment on AGB and SAGB, we applied a longitudinal analysis. This approach is justified (Stewart-Oaten and Bence 2011) given the rigorous experimental design of the FFS study, namely a randomized and replicated assignment of treatments (including control) with both a pre-treatment measurement and repeated post-treatment measurements. We used mixed-effects linear modeling with the response being a function of the fixed variables of time and treatment. To account for the spatial correlation, plots (i.e., samples) were nested within compartment (i.e., block) and included as random variables in the mixed-effects model (Bolker et al. 2009). Specifically, we defined six a priori models (Table 2.1) with the response as a function of time as the least restrictive case (i.e., the null model). We added a quadratic term (Years^2) to check for changes in the trajectory and included models with the main treatment effects and with treatment-by-time interactions. To stabilize the variance, the response variables were square root transformed. Treatment effects were compared to the baseline trends measured in the Control plots.

We evaluated model performance using an information theoretic approach (Burnham and Anderson 2002). For each model, we calculated Akaike's information criterion (AIC), differences in AIC values relative to the model with the lowest AIC (ΔAIC), and AIC weights

($AICw_i$). $AICw_i$ were calculated to normalize the strength of evidence for a given model and can be interpreted as the probability that a given model is the best Kullback–Leibler model for the data given the candidate set of models; models < 2ΔAIC have substantial empirical support (Burnham and Anderson 2002). Analyses were conducted using the R statistical computing environment version 3.3.0 (R Core Team 2016).

Table 2.1: Linear models of aboveground live biomass (AGB) and stable aboveground live biomass (SAGB). Results from mixed effects linear models include the number of parameters (k) and the AIC metrics (see text for definitions).

A. Aboveground Live Biomass Models	**k**	**AIC**	**ΔAIC**	**AICw**
$AGB^{1/2}$ = f(Years+$Years^2$+Treatment+Treatment:Years)	12	4065.3	0.0	0.71
$AGB^{1/2}$ = f(Years+Treatment+Treatment:Years)	11	4067.1	1.8	0.29
$AGB^{1/2}$ = f(Years+Treatment)	8	4159.6	94.2	0.00
$AGB^{1/2}$ = f(Years+$Years^2$+Treatment)	9	4159.8	94.4	0.00
$AGB^{1/2}$ = f(Years)	5	4161.7	96.4	0.00
$AGB^{1/2}$ = f(Years+ $Years^2$)	6	4161.9	96.6	0.00
B. Stable Aboveground Live Biomass Models	**k**	**AIC**	**ΔAIC**	**AICw**
$SAGB^{1/2}$ = f(Years+ $Years^2$+Treatment+Treatment:Years)	12	4812.5	0.0	0.81
$SAGB^{1/2}$ = f(Years+ $Years^2$+Treatment)	9	4816.1	3.6	0.13
$SAGB^{1/2}$ = f(Years+ $Years^2$)	6	4818.8	6.3	0.03
$SAGB^{1/2}$ = f(Years+Treatment+Treatment:Years)	11	4820.6	8.1	0.01
$SAGB^{1/2}$ = f(Years+ Treatment+Treatment:Years)	8	4823.7	11.2	0.00
$SAGB^{1/2}$ = f(Years)	5	4826.4	13.9	0.00

2.3 Results

Before the treatment, AGB averaged 321 Mg ha^{-1} with the most biomass stored in the ThinBurn sites (mean = 359 Mg ha^{-1}, se = 20 Mg ha^{-1}) and the least in the Burn sites (mean = 285 Mg ha^{-1}, se = 16 Mg ha^{-1}). The imposition of the treatments (PRE to POST-1YR) reduced AGB by 21.5% in the ThinBurn and only 1.3% in the Burn (Figure 2.1). Biomass in the Control sites steadily accumulated throughout the study so that by POST-14yr, Control stored the most AGB (median = 426 Mg ha^{-1}). In contrast, the ThinBurn sites not only lost the most AGB in the treatment but

also failed to recover with only a 13% net gain in AGB between POST-1YR and POST-14YR (Figure 2.1).

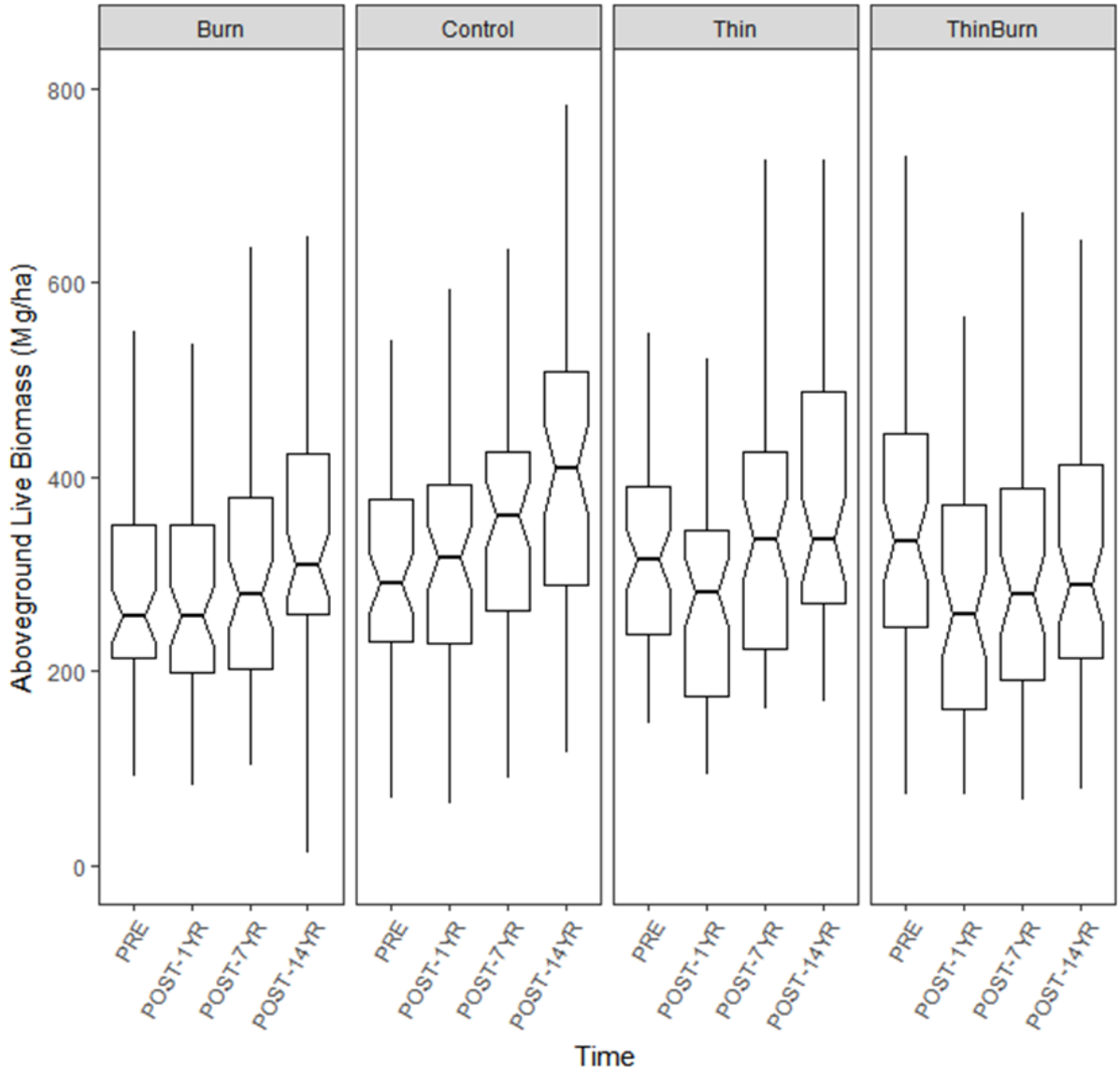


Figure 2.1: Boxplots describing the trends in aboveground live biomass (AGB) in the Fire and Fire Surrogate Study at Blodgett Forest Research Station. The narrow midline indicates the median, the edges of the box represent the 25th and 75th quartiles, and the whiskers define 1.5x the interquartile range.

Prescribed fire greatly reduced the probability of torching (Figure 2.2). In both the Burn and ThinBurn treatments, the probability of a damaging fire was zero in POST-1YR. Ptorch was reduced by about half in the Thin treatment. The reduction in the prescribed fire units was sustained throughout the study. However, note that between the POST-7YR and POST-14YR, the Burn sites were re-burned.

Accounting for the risk of torching dramatically reshaped the trajectory of biomass storage (Figure 2.2). In terms of SAGB, the prescribed fire increased SAGB in POST-1YR by 96% in the Burn sites and 34.8% in the ThinBurn sites. The effects were sustained through the 14 years of

the study. For example, at the ThinBurn sites, SAGB was 97% of AGB in POST-14YR. In other words, 14 years after the initial treatment almost all the live tree biomass present was a low risk of loss due to fire.

These patterns were supported by the statistical analysis. For both AGB and SAGB, the models with the greatest weight of evidence (>0.7 AICw in both cases) included treatment and the interaction between treatment and time (Table 2.1). Based on the model predictions, the Control sites maintained the most AGB but the least SAGB following treatment (Figure 2.4). In contrast, the burned sites (Thin and ThinBurn) supported the least AGB but the most SAGB.

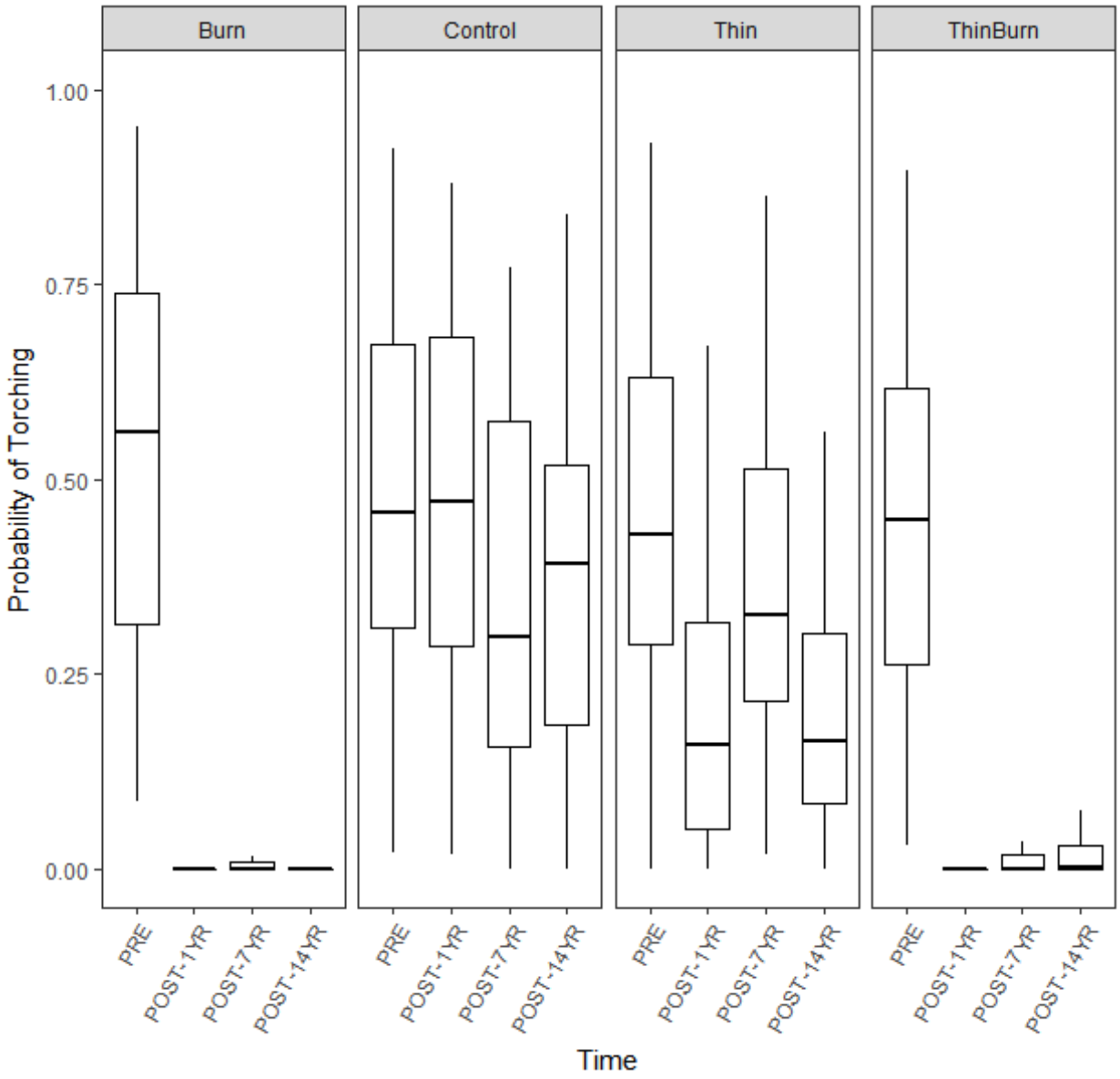


Figure 2.2: Boxplots describing the trends in the probability of torching (Ptorch) in the Fire and Fire Surrogate Study at Blodgett Forest Research Station. The narrow midline indicates the median, the edges of the box represent the 25th and 75th quartiles, and the whiskers define 1.5x the interquartile range.

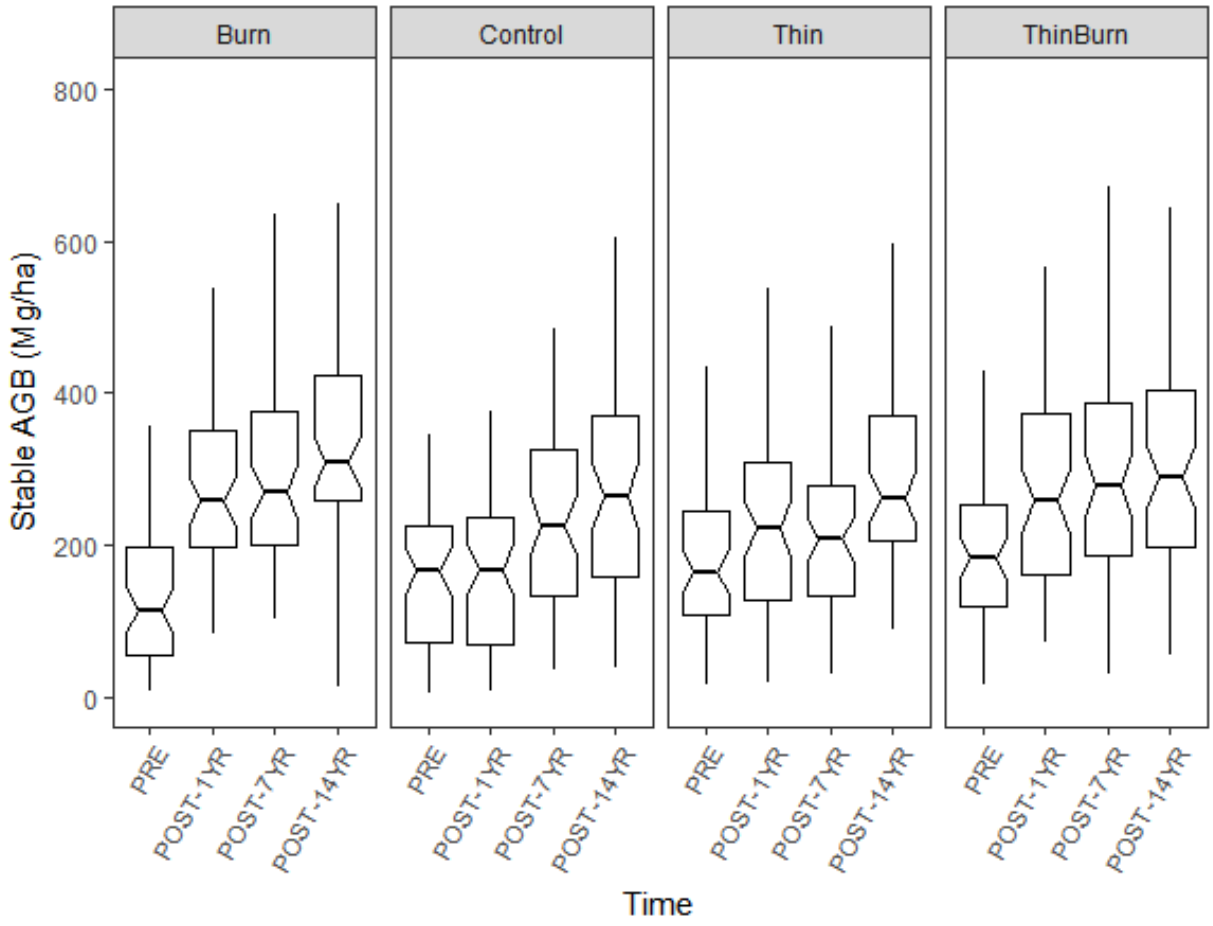


Figure 2.3: Boxplots describing the trends in the stable aboveground live biomass (SAGB) in the Fire and Fire Surrogate Study at Blodgett Forest Research Station. The narrow midline indicates the median, the edges of the box represent the 25th and 75th quartiles, and the whiskers define 1.5x the interquartile range.

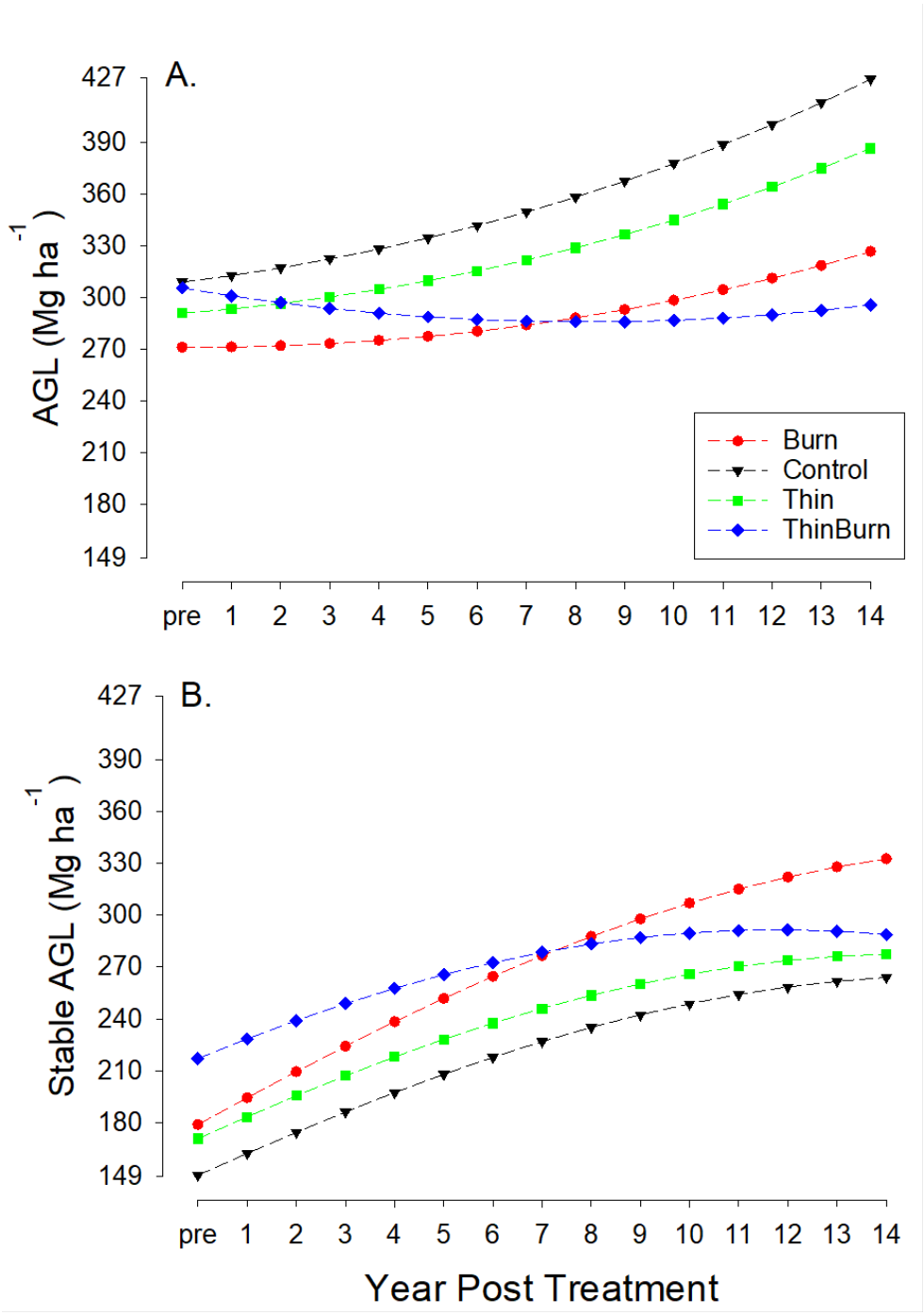
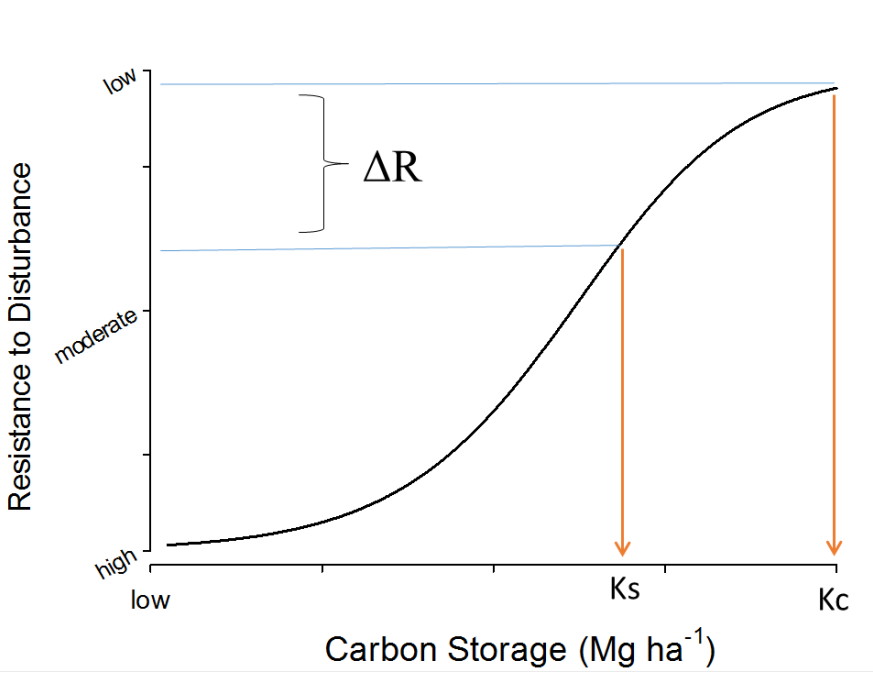


Figure 2.4: Predicted treatment effects on trends in aboveground live biomass (A) and stable aboveground live biomass (B) in the Fire and Fire Surrogate Study at Blodgett Forest Research Station. Results from the best mixed-effects linear model (i.e., greatest AICw, Table 2.1).

2.4 Conclusions and Future Directions

The Sierran mixed conifer forest is a landscape at risk from the combined impact of an altered fire regime, past harvesting, and a warming climate. To reduce this risk, improving resilience by reducing tree density, lowering fuel loads, and increasing heterogeneity is the recommended management response (North et al. 2009, North 2012). A key state condition in the mixed conifer forest is the carbon carrying capacity (Kc). We define Kc as the maximum potential mass of stored carbon in an ecosystem under prevailing environmental conditions and the prevailing disturbance regime (Keith et al. 2010). Collins et al. (2015) reported a 200% increase in live tree carbon stored in contemporary mixed conifer forests where fire was excluded compared to the carbon stored under pre-settlement conditions. More generally, Earles et al. (2014) showed with forest simulation models that Kc is responsive to changes in both fire frequency and drought. They also noted that the stability of the carbon stored tends to decrease at greater magnitudes. We propose that the difference between Kc and the carbon that can be sustained (Ks) is an index of a forest's resistance to disturbance with smaller differences between Kc and Ks indicative of higher resilience (Figure 2.5). We used the results from the Fire and Fire Surrogate Study at Blodgett Forest Research Station to quantify this conceptual trade-off between carbon storage and stability. We specifically focused on the risk posed by a problem wildfire and relied on AGB as an indicator of carbon dynamics.

Figure 2.5: The proposed functional relationship between carbon storage and resistance to disturbance in the mixed conifer forest of the Sierra Nevada. Kc denotes the carbon carrying capacity; Ks denotes a stable level of carbon storage given the risk posed by wildfire, drought, and other disturbances. ΔR is the difference between Kc and Ks.



Over a 14 year period following treatment, the trade-offs between storage and stability were striking (Figure 2.4). Absent treatment (i.e., Control), AGB steadily increased in this productive

second-growth forest thereby maximizing AGB stocks. The burned treatments (i.e., Burn and ThinBurn) stored the least AGB. When loss due to fire was included, the results flipped. The burn treatments had the largest store of stable AGB (SAGB) and the controls the least.

These results were strongly grounded in measured responses over time in a fully replicated and robustly sampled field experiment. Calculations of SAGB did depend on fire mortality prediction from Ptorch, but the Ptorch model was parameterized with forest and fuel data. Moreover, the fire weather scenarios were extracted from problematic wildfires that occurred in the region during the period of study (2000-2016).

The pattern of increasing AGB in absence of fire (Figure 2.1, Control) fits observations for an old-growth mixed conifer forest in the northern Sierra Nevada (Levine et al. 2016). This mesic site has been free of fire and has been adding AGB for the last 50+ years. As of 2013, it stored 512 Mg ha^{-1} of AGB. In 2016, the mean AGB in the Blodgett Control was 441 Mg ha^{-1}, suggesting that under prevailing conditions, it still has not reached it carbon carrying capacity (K_c in Figure2.5). For 2016, our analysis predicted that only 289 Mg ha^{-1} (66%) of the AGB in the untreated forest would survive a wildfire. In contrast, 324 Mg ha^{-1} of SAGB in the burned plots (Burn and ThinBurn, Figure 2.3) would remain after a wildfire (K_s in Figure 2.5).

The trade-off between AGB storage and stability critically depends on the probability of fire occurring in these stands (Restaino and Peterson 2013). Moreover, the efficacy of the fuel treatments decline as the forest grows (e.g., Collins et al. 2011). In addition, our calculation of SAGB did not evaluate the risk posed by megafires (Stephens et al. 2014) or any other uncharacteristically severe disturbance (sensu Millar and Stephenson 2015). Nevertheless, we have demonstrated a quantitative framework for accessing the capacity of Sierran forests to store AGB that incorporates one major threat to its longevity.

3: References

Agee, J. K., and C. N. Skinner. 2005. Basic principles of forest fuel reduction treatments. Forest Ecology and Management 211:83-96.

Bell, D. M., M. J. Gregory, and J. L. Ohmann. 2015. Imputed forest structure uncertainty varies across elevational and longitudinal gradients in the western Cascade Mountains, Oregon, USA. Forest Ecology and Management 358:154-164.

Bechtold, W.A. and P.L. Patterson. 2005. The enhanced forest inventory and analysis program: national sampling design and estimation procedures. USFS Gen. Tech. Rep. SRS-80.

Bellassen, V., and S. Luyssaert. 2014. Carbon sequestration: Managing forests in uncertain times. Nature 506:153-155.

Bolker, B.M., M.E. Brooks, C.J. Clark, S.W. Geange, J.R. Poulsen, M.H.H. Stevens, and J.S.S. White, J.S.S. 2009. Generalized linear mixed models: a practical guide for ecology and evolution. Trends in Ecology and Evolution 24:127-135.

Burnham, K. P., and D. R. Anderson. 2002. Model Selection and Multimodel Inference: A Practical Information -Theoretic Approach. Springer, New York.

Busing, R. T., and T. Fujimori. 2005. Biomass, production and woody detritus in an old coast redwood (*Sequoia sempervirens*) forest. Plant Ecology 177: 177–188.

Camarero, J. J., A. Gazol, J. D. Galván, G. Sangüesa-Barreda, and E. Gutiérrez. 2015. Disparate effects of global- change drivers on mountain conifer forests: warming-induced growth enhancement in young trees vs. CO2 fertilization in old trees from wet sites. Global Change Biology 21:738-749.

Campbell, J. L., M. E. Harmon, and S. R. Mitchell. 2012. Can fuel-reduction treatments really increase forest carbon storage in the western US by reducing future fire emissions? Frontiers in Ecology and the Environment 10:83-90.

Carroll, R. J., and D. Ruppert. 1996. The use and misuse of orthogonal regression in linear errors-in-variables models. The American Statistician 50:1-6.

Christensen, G.A., A.N. Gray, O. Kuegler, N.A. Tase, and M. Rosenberg. 2017. AB 1504 California Forest Ecosystem and Harvested Wood Product Carbon Inventory: 2006 - 2015. Final Report. California Department of Forestry and Fire Protection agreement no. 7CA02025. Sacramento, CA: California Department of Forestry and Fire Protection and California Board of Forestry and Fire Protection.

Collins, B. M., S. L. Stephens, G. B. Roller, and J. J. Battles. 2011. Simulating fire and forest dynamics for a landscape fuel treatment project in the Sierra Nevada. Forest Science 57:77-88.

Collins, B. M., J. M. Lydersen, R. G. Everett, D. L. Fry, and S. L. Stephens. 2015. Novel characterization of landscape-level variability in historical vegetation structure. Ecological Applications 25:1167-1174.

Dubayah, R. O. and J.B. Drake. 2000. LiDAR remote sensing for forestry. Journal of Forestry 98:44-46.

Earles, J. M., M. P. North, and M. D. Hurteau. 2014. Wildfire and drought dynamics destabilize carbon stores of fire-suppressed forests. Ecological Applications 24:732-74

FIA. 2010. Regional biomass equations used by the Forest Inventory and Analysis program to estimate bole, bark, and branches (updated 13-Jan-2010), U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, Portland, Oregon.

FIA. 2017. Forest inventory data for California from the Forest Inventory and Analysis Program. FIADB_1.6.0.02.

Forest Climate Action Team. 2018. California Forest Carbon Plan: Managing Our Forest Landscapes in a Changing Climate. Sacramento, CA. 178p.

FRAP Vegetation (2015). FVEG15_1 accessed November 2017 (http://frap.fire.ca.gov/data/frapgisdata-sw-fveg_download).

Genuer, R., J Poggi, and C. Tuleau-Malot. 2015. VSURF: An R Package for Variable Selection Using Random Forests. The R Journal 7:19-33.

Gonzalez, P., G. P. Asner, J. J. Battles, M. A. Lefsky, K. M. Waring, and M. Palace. 2010. Forest carbon densities and uncertainties from LiDAR, QuickBird, and field measurements in California. Remote Sensing of Environment 114:1561-1575.

Gonzalez, P., J.J. Battles, B.M. Collins, T. Robards, and D.S. Saah. 2015. Aboveground live carbon stock changes of California wildland ecosystems, 2001-2010. Forest Ecology and Management 348:68-77.

Grossmann, E.B., J.L. Kagan, H.K. May, M.J. Gregory, and C. Tobalske. 2008. Final report on land cover mapping methods, map zones 2 and 7, Pacific Northwest ReGAP. Institute for Natural Resources, Oregon State University, Corvallis, OR.

Hagan, N., N. Harris, S. S. Saatchi, T. Pearson, C.W. Woodall, S. Ganguly, G.M. Domke, B.H. Braswell, B.F. Walters, J.C. Jenkins, S. Brown, W.A. Salas, A. Fore, Y. Yu, R.R. Nemani, C. Ipsan, and K.R. Brown. 2016. CMS: Forest Carbon Stocks, Emissions, and Net Flux for the Conterminous US: 2005-2010. ORNL DAAC, Oak Ridge, Tennessee, USA. http://dx.doi.org/10.3334/ORNLDAAC/1313)

Harmon, M. E., and R. J. Pabst. 2015. Testing predictions of forest succession using long-term measurements: 100 yrs of observations in the Oregon Cascades. Journal of Vegetation Science 26:722-732.

Huang, W., A. Swatantran, L. Duncanson, K. Johnson, D. Watkinson, K. Dolan, J. O'Neil-Dunne, G. Hurtt, and R. Dubayah. 2017. County-scale biomass map comparison: a case study for Sonoma, California. Carbon Management 8:417-434.

Hurteau, M., and M. North. 2009. Fuel treatment effects on tree-based forest carbon storage and emissions under modeled wildfire scenarios. Frontiers in Ecology and the Environment 7:409-414.

Hurteau, M. D., and M. L. Brooks. 2011. Short- and Long-term Effects of Fire on Carbon in US Dry Temperate Forest Systems. Bioscience 61:139-146.

Keith, H., Mackey, B.G. and Lindenmayer, D.B., 2009. Re-evaluation of forest biomass carbon stocks and lessons from the world's most carbon-dense forests. Proceedings of the National Academy of Sciences, 106:11635-11640.

Keith, H., B. Mackey, S. Berry, D. Lindenmayer, and P. Gibbons. 2010. Estimating carbon carrying capacity in natural forest ecosystems across heterogeneous landscapes: addressing sources of error. Global Change Biology 16:2971-2989.

Kennedy, R. E., Z. Yang, and W. B. Cohen. 2010. Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr - Temporal segmentation algorithms. Remote Sensing of Environment 114:2897-2910.

Kennedy, R.E., J. Ohmann, M. Gregory, H. Roberts. D.M. Bell, V. Kane, M. J. Hughes, W. Cohen, S. Powell, N. Neeti, S. Hooper, J. Kane, D.L. Miller, J. Perkins, J. Braaten, and R. Seidl. 2018. An empirical, integrated forest biomass monitoring system. Environmental Research Letters 13:2.

Krofcheck, D. J., M. D. Hurteau, R. M. Scheller, and E. L. Loudermilk. 2017. Restoring surface fire stabilizes forest carbon under extreme fire weather in the Sierra Nevada. Ecosphere 8(1).

Liang, S., M. D. Hurteau, and A. L. Westerling. 2017. Potential decline in carbon carrying capacity under projected climate-wildfire interactions in the Sierra Nevada. Scientific Reports 7:2420.

Landfire. 2010. http://www.landfire.gov/lf_120.php.

Legendre, P. and Legendre, L.F., 2012. Numerical Ecology 24. Elsevier.

Levine, C. R., F. Krivak-Tetley, N. S. van Doorn, J. S. Ansley, and J.J. Battles. 2016. Long-term demographic trends in a fire-suppressed mixed-conifer forest. Canadian Journal of Forest Research 46:745-52.

Luyssaert, S., E. D. Schulze, A. Boerner, A. Knohl, D. Hessenmoeller, B. E. Law, P. Ciais, and J. Grace. 2008. Old-growth forests as global carbon sinks. Nature 455:213-215.

McGaughey, R.J., 2018. FUSION/LDV: Software for LiDAR Data Analysis and Visualization Version 3.7. United States Department of Agriculture Forest Service Pacific Northwest Research Station (http://forsys.sefs.uw.edu/fusion/FUSION_manual.pdf).

McGroddy, M.E., T. Daufresne, and L.O. Hedin. 2004. Scaling of C:N:P stoichiometry in forests worldwide: implications of terrestrial Redfield-type ratios. Ecology 85:2390–2401.

McIntyre, P. J., J. H. Thorne, C. R. Dolanc, A. L. Flint, L. E. Flint, M. Kelly, and D. D. Ackerly. 2015. Twentieth-century shifts in forest structure in California: Denser forests, smaller trees, and increased dominance of oaks. Proceedings of the National Academy of Sciences 112:1458-1463.

McIver, J. D., S. L. Stephens, J. K. Agee, J. Barbour, R. E. Boerner, C. B. Edminster, K. L. Erickson, K. L. Farris, C. J. Fettig, and C. E. Fiedler. 2013. Ecological effects of alternative

fuel-reduction treatments: highlights of the National Fire and Fire Surrogate study (FFS). International Journal of Wildland Fire 22:63-82.

McRoberts, R. E., E. O. Tomppo, A. O. Finley, and J. Heikkinen. 2007. Estimating areal means and variances of forest attributes using the *k*-Nearest Neighbors technique and satellite imagery. Remote Sensing of Environment 111: 466-480.

McRoberts, R. E. 2012. Estimating forest attribute parameters for small areas using nearest neighbors techniques. Forest Ecology and Management 272:3-12.

Millar, C. I., and N. L. Stephenson. 2015. Temperate forest health in an era of emerging megadisturbance. Science 349:823-826.

Miller, J. D., H. D. Safford, M. Crimmins, and A. E. Thode. 2009. Quantitative Evidence for Increasing Forest Fire Severity in the Sierra Nevada and Southern Cascade Mountains, California and Nevada, USA. Ecosystems 12:16-32.

Neeti, N., and R. Kennedy, R. 2016. Comparison of national level biomass maps for conterminous US: Understanding pattern and causes of differences. Carbon Balance and Management 11:19

North, M. P. 2012. Managing Sierra Nevada Forests. General Technical Report PSW-GTR-237 US Forest Service: 165-175.

North, M. P. Stine, K. O'Hara, W. Zielinski, and S. Stephens. 2009. An ecosystem management strategy for Sierran mixed-conifer forests. 2nd printing, with addendum. Gen. Tech. Rep. PSW-GTR-220. Albany, CA: U.S. Department of Agriculture, Forest Service, Pacific Southwest Research Station.

NBCD 2000. National Biomass and Carbon Dataset. https://doi.org/10.3334/ORNLDAAC1161)

Ohmann, J. L. and M. J. Gregory. 2002. Predictive mapping of forest composition and structure with direct gradient analysis and nearest-neighbor imputation in coastal Oregon, USA. Canadian Journal of Forest Research 32:725-741.

Ohmann, J. L., M. J. Gregory, H. M. Roberts, W. B. Cohen, R. E. Kennedy, and Z. Yang. 2012. Mapping change of older forest with nearest-neighbor imputation and Landsat time-series. Forest Ecology and Management 272:13-25.

Ohmann, J. L., M. J. Gregory, H. M. Roberts. 2014. Scale considerations for integrating forest inventory plot data and satellite image data for regional forest mapping. Remote Sensing of the Environment 151:3-15.

PNW-FIADB. 2015. A Data Dictionary and User Guide for the PNW-FIADB database. U.S. Department of Agriculture, Forest Inventory and Analysis, Pacific Northwest Research Station, Portland, Oregon.

R Core Team. 2016. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria. https://www.R-project.org

Rebain, S. A. (compiler). 2010 (revised March 23, 2015). The Fire and Fuels Extension to the Forest Vegetation Simulator: Updated Model Documentation. Internal Report, U.S.

Department of Agriculture, Forest Service, Forest Management Service Center, Fort Collins, USA.

Reinhardt, E., and N. L. Crookston. 2003. The Fire and Fuels Extension to the Forest Vegetation Simulator. General Technical Report RMRS-GTR-116, U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station, Ogden, UT, USA.

Restaino, J. C., and D. L. Peterson. 2013. Wildfire and fuel treatment effects on forest carbon dynamics in the western United States. Forest Ecology and Management 303:46-60.

Scott, J. H., and E. D. Reinhardt. 2001. Assessing crown fire potential by linking models of surface and crown fire behavior. Research Paper RMRS-RP-29, U. S. Department of Agriculture, Forest Service, Rocky Mountain Research Station.

Smith, R.J. 2009. Use and misuse of the reduced major axis for line-fitting. American Journal of Physical Anthropology 140: 476-486.

Steininger, M.K. 2000. Satellite estimation of tropical secondary forest above-ground biomass: Data from Brazil and Bolivia. International Journal of Remote Sensing 21: 1139–1157.

Stephens, S. L., and B. M. Collins. 2004. Fire regimes of mixed conifer forests in the north-central Sierra Nevada at multiple spatial scales. Northwest Science 78:12-23

Stephens, S.L. and J.J. Moghaddas. 2005. Experimental fuel treatment impacts on forest structure, potential fire behavior, and predicted tree mortality in a California mixed conifer forest. Forest Ecology and Management 215:21-36.

Stephens, S. L., and L. W. Ruth. 2005. Federal forest-fire policy in the United States. Ecological Applications 15:532-542.

Stephens, S.L., N. Burrows, A. Buyantuyev, R.W. Gray, R.E. Keane, R. Kubian, S. Liu, F. Seijo, L. Shu, K.G. Tolhurst, and J.W. Van Wagtendonk. 2014. Temperate and boreal forest mega-fires: characteristics and challenges. Frontiers in Ecology and the Environment 12:115-122.

Stewart-Oaten, A. and J.R. Bence. 2001. Temporal and spatial variation in environmental impact assessment. Ecological Monographs 71:305-339.

Taylor, A. H., A. M. Vandervlugt, R. S. Maxwell, R. M. Beaty, C. Airey, and C. N. Skinner. 2014. Changes in forest structure, fuels and potential fire behavior since 1873 in the Lake Tahoe Basin, USA. Applied Vegetation Science 17:17-31.

ter Braak, C. J. F. 1986. Canonical correspondence analysis: A new eigenvector technique for multivariate direct gradient analysis. Ecology 67: 1167–1179.

Turner, M. G. 2010. Disturbance and landscape dynamics in a changing world. Ecology 91:2833-2849.

USDA Forest Service. 2017. News Release: Record 129 Million Dead Trees in California. https://www.fs.usda.gov/Internet/FSE_DOCUMENTS/fseprd566303.pdf

van Mantgem, P. J., and N. L. Stephenson. 2007. Apparent climatically induced increase of tree mortality rates in a temperate forest. Ecology Letters 10:909-916.

Wiechmann, M. L., M. D. Hurteau, M. P. North, G. W. Koch, and L. Jerabkova. 2015. The carbon balance of reducing wildfire risk and restoring process: an analysis of 10-year post-treatment carbon dynamics in a mixed-conifer forest. Climatic Change 132:709-719.

Woodall, C.W., J.W. Coulston, G.M. Domke, B.F. Walters, D.N. Wear, J.E., Smith, H.E. Andersen, B. J. Clough, W.B. Cohen, D.M., Griffith, and S.C. Hagen. 2015. The US Forest Carbon Accounting Framework: Stocks and Stock Change, 1990-2016. US Forest Service, General Technical Report NRS-154.

Zolkos, S. G., S. J. Goetz, and R. Dubayah. 2013. A meta-analysis of terrestrial aboveground biomass estimation using LiDAR remote sensing. Remote Sensing of Environment 128: 289-298.

APPENDIX A: Comparison of Carbon Storage Estimates

A.1 Introduction

California has a pressing need to measure and manage its forest carbon as part of its effort to reduce greenhouse gas emissions (ARB 2015). However, measuring forest carbon at the required spatial (entire state) and temporal resolution (annual) is a challenge. The measurement includes two main sources of uncertainty: 1) the allometric models used to estimate tree-level biomass; and 2) the method used to scale tree-level carbon estimates to the landscape (Zhao et al. 2012, Duncanson et al. 2017). In order to guide our development of a robust means to track annual changes in forest carbon stocks in California, we first review existing approaches and then compare the performance in three forest landscapes.

A.2 Review: Allometric Modeling of Tree Biomass

At their core, all estimates of forest carbon rely on the allometric property of vascular plants:

$M \alpha D^{8/3}$ (Equation A-1)

where M is tree mass and D is stem diameter (Enquist and Niklas 2001). Note that approximately 50% of the (bio)mass of tree is composed of carbon. However, the measurement of M requires that trees be harvested and all parts carefully measured, typically using dimensional analysis (Whittaker and Woodwell 1968). With direct measures of M and D, allometric models are statistically fit using the equation (Whittaker et al. 1974):

$M = aD^b$ (Equation A-2)

where M represents the biomass of the tree, D the diameter of the stem (usually measured at breast height, 1.37 m) and a, b are estimated parameters. Often to simplify the analysis, equation 2 is transformed to its linear form:

$M = a + b * ln(D)$ (Equation A-3)

While the theory based on biomechanical principles predicts no variance in the allometric relationship, in practice the scaling coefficient (b) can deviate from the expected $8/3^{rd}$ power. For example, in Jenkins et al.'s (2003) national synthesis, the scaling coefficient never reached $8/3$ (2.67) and ranged from a low of 2.26 to a high of 2.48 for tree species groups (e.g., pine trees, mixed hardwoods). The variance is due to site and species differences in tree height and wood density. Thus, the predictive ability of the allometric models can be improved by including tree

height and developing species-specific equations. Alternatively, specific wood gravity can be used instead of species to account for differences in wood properties (Chave et al. 2005, Chojnacky et al. 2014, Temesgen et al. 2015).

As a result of efforts to accommodate variation in tree allometry, existing equations take different forms. For common species with large ranges, there are multiple allometric equations developed from different data sets and fit to different model forms. The decision on which equation to apply can lead to large differences in tree mass estimates for the same diameter (Melson et al. 2011). This "model uncertainty" translates to important differences in carbon stock estimates for forested landscapes (Zhao et al. 2012, Ducanson et al. 2017). At the same time, the uncertainty can be controlled if a consistent set of models are developed and applied over time. Indeed, developing such a consistent set of models is a priority task of the US Forest Service Forest Inventory and Assessment (FIA) Program (Weiskittel et al. 2015a).

Several problems confront the FIA effort to develop a nationally consistent approach to predict tree biomass (Weiskittel et al. 2015b). One is that the performance of models varies depending on the scope of the application. Locally derived models tend to perform better in terms of accuracy and precision when they are applied locally in comparison to regional or national models. At the same time, national models are more efficient at summarizing national trends in tree biomass but this efficiency comes at the cost of resolution at the local and regional level (Zhou and Hemstrom 2009, Westfall 2012, Weiskettel et al. 2015). Another challenge for FIA is that data on tree mass are limited at the national level. Often individual studies only sampled part of the size range for well-sampled species. Other species were not well-sampled or even measured at all. Thus, pseudo data were generated to fit equations for the entire size range of well-sampled species, and sparsely-sampled species were lumped into larger taxonomic groups (Jenkins et al. 2003). A third challenge is that in the western forests (including California), there exists an effective regional approach for estimating wood biomass from volume equations and wood density (FIA 2010). This regional approach was built on an extensive legacy dataset collected to inform timber management. Given the commercial imperative to accurately estimate wood volume, the performance of these equations was well-vetted. However, the regional equations focus on the merchantable part of the stem. They do not include the tops and limbs of the tree. In an effort to resolve differences between the national "Jenkins" equations and the regional equations, the FIA developed the Component Ratio Method (CRM, Woodall et al. 2011) where aspects of both approaches are combined to refine the estimate of whole-tree biomass. Currently CRM is the national standard for estimating tree biomass, but FIA reports alternative estimates of tree biomass using regional equations and many forestry/ecosystem studies use Jenkins equations. Not surprisingly, the plethora of choices permeates forest carbon estimates in California (Huang et al. 2017). Moreover, the efforts to refine tree biomass predictions that are both nationally consistent and regionally accurate continue (e.g., Weiskittel et al. 2015a).

A.3 Review: Scaling of Tree-Level Biomass Estimates

Forest carbon estimates over large landscapes are often calculated as the product of surface areas of vegetation types, classified by satellite remote sensing (Table A-1), and the carbon densities, derived from field measurements of trees and allometric equations, summed over all vegetation types (e.g. Achard et al. 2004, DeFries et al. 2007, Harris et al. 2012). The number of vegetation types that satellites with moderate spectral or spatial resolutions can accurately

discriminate, generally five to twenty classes (e.g. Bartholomé and Belward 2005, Loveland et al. 2000), can limit the possible carbon density assignments to a few discrete values.

In contrast, other methods use Light Detection and Ranging (LiDAR) or high-resolution satellites such as QuickBird, Ikonos, or WorldView (Table A-1) to sense the physical dimensions of trees to which aboveground biomass directly correlates (e.g., Gonzalez et al. 2010, Saatchi et al. 2011, Huang et al. 2017). With these systems, forest carbon content equals the product of the area and the carbon density of each pixel, where carbon density is calculated by applying allometric equations to field measurements of individual trees and correlated to canopy height metrics estimated by LiDAR or tree crown diameter estimated by high-resolution satellite data. This method generates raster coverage of the spatial distribution of forest carbon density with continuous values.

A recent meta-analysis of the performance of remote sensors in estimating aboveground forest biomass/carbon density (Zolkos et al. 2013) concluded that the active sensors were more accurate than the passive sensors (Table A-1). In addition, various implementations of LiDAR were better than any satellite-based active sensing platform. In general, the greater detail with which the tree form was resolved (e.g., height, crown width), the better the performance. However, as Zolkos et al. (2013) note, the accuracy of the estimate also depended on the field measurements of tree biomass/carbon. Despite the promise of airborne LiDAR, the current costs of its acquisition make it impractical for state-level monitoring of forest carbon.

Table A-1: Summary of major remote sensing technologies used for assessing carbon stocks.

Sensor	Scale	Platform	Mode	Detection	Access
MODIS	250 m	Satellite	Passive	36 spectral bands	Public
Landsat	30 m	Satellite	Passive	9 spectral bands	Public
QuickBird	1m	Satellite	Passive	5 spectral band	Commercial
GLAS	1 km	Satellite	Active	Light returns	Public
PALSAR	10 m	Satellite	Active	Microwave returns	Public
LiDAR	<1 m	Aircraft	Active	Light returns	Commercial
Aerial photos	50 m	Aircraft	Passive	Visible spectrum	Public (NAIP)

MODIS = Moderate Resolution Imaging Spectroradiometer; GLAS = Geoscience Laser Altimeter System; PALSAR = Phased Array type L-band Synthetic Aperture Radar; LiDAR = Light Detection and Ranging; NAIP = National Agriculture Imagery Program.

The Landsat sensor provides a workable alternative to scale tree-level estimates of biomass (Gonzalez et al. 2014, Pflugmacher et al. 2014). Two applications of Landsat to forest carbon mapping have been realized to date: LandfireC (Gonzalez et al. 2014) and LandTrendr (Pflugmacher et al. 2014). LandfireC relies on vegetation type, cover, and height assignments from LANDFIRE. LANDFIRE combines data from several sources but primarily Landsat

images to produce fine-grained spatial units (Rollins 2009) over which field data can be applied. LandTrendr uses the same sensor (i.e., Landsat) but directly links tree biomass estimates to changes in the spectral signature via gradient nearest neighbor analysis (GNN; Ohmann and Gregory 2002, Ohmann et al. 2012).

A.4 Comparison of Approaches

To inform our current mapping efforts, we developed multiple estimates of carbon density for three forest landscapes in California (Table A-2). We choose three sites where the specific goal was to use LiDAR and robust field sampling in order to accurately access carbon stores. All three sites span a diverse and relatively large forest landscape. LiDAR for all sites was acquired using an Optech ALTM 2050 system on an airplane flying at an altitude of ~800 m and average velocity of 260 km per hour. The ALTM 2050 acquired up to three returns per pulse at a pulse frequency of 50 kHz, scan frequency of 38 Hz, and a maximum scan angle of 15°, creating a swath width of ~580 m. The point density is about 2–4 returns per square meter. Optech, Inc. rates the RMSE precision of individual point locations surveyed by the ALTM 2050 as ±15 cm vertical and ±50 cm horizontal. Flights were flown during the growing season of 2005 (Gonzalez et al. 2010, Chen et al. 2012).

Two sites are in the Sierra Nevada, North Yuba and Sagehen. The North Yuba carbon area in located north of Downieville, California in the Tahoe National Forest. The area encompasses 57.5 km^2 within the North Yuba Old Forest Emphasis Area, an area managed to maintain and restore pre-European settlement forest conditions. The area was harvested during the mid-1800's and has an existing network of roads and trails in addition to disturbance from mining activities. Sagehen refers to the Sagehen Creek Experimental Forest near Truckee, CA in the Tahoe National Forest. Sagehen contains a more structurally diverse landscape (35 km^2) that spans from mature mixed conifer forest to young pine plantations. Garcia/Mailliard refers to the Garcia River Forest (GRF) located south of Boonville, California. The carbon area encompasses 58 km^2 on the eastern side of the GRF. Before The Conservation Fund acquired the land, private industry timber companies had extensively harvested the area. In addition to the GRF carbon area, a nearby state reserve was also sampled to represent larger and older trees than are currently present in the GRF. The Mailliard Redwood State Natural Reserve is a 200-hectare area of old forest dominated by coast redwood, located due east of the Garcia River Forest.

Table A-2: Comparison of forest carbon density estimates for three forest landscapes in California. Superscripts refer to the source of the estimate: [1]Gonzalez et al. 2010; [2]Battles et al. 2014; [3]This study, see Chapter 1; [4]Chen et al. 2012.

Landscape (forest type, area)	Carbon density (Mg ha^{-1})	95% CI (Mg ha^{-1})	Year	RME	Allometric equation source
North Yuba (mixed conifer, 58 km^2)					
LiDAR[1]	140	1	2005		Researcher (local)

LandfireC[2]	120	70	2008	-0.14	FIA regional
LT-GNN$_{CRM}$[3]	112	--	2005	-0.20	FIA CRM
LT-GNN$_{Jenkins}$[3]	152	--	2005	0.09	Jenkins
Garcia/Mailliard					
(coast redwood, 59 km^2)					
LiDAR[1]	82	1	2005		Researcher (local)
LandfireC[2]	120	80	2008	0.46	FIA regional
LT-GNN$_{CRM}$[3]	95	--	2005	0.16	FIA CRM
LT-GNN$_{Jenkins}$[3]	111	--	2005	0.35	Jenkins
Sagehen					
(mixed conifer, 35 km^2)					
LiDAR[4]	74	--	2005		Researcher (local)
LandfireC[2]	73	32	2008	-0.01	FIA regional
LT-GNN$_{CRM}$[3]	53	--	2005	-0.28	FIA CRM
LT-GNN$_{Jenkins}$[3]	84	--	2005	0.14	Jenkins

$RME_i = \frac{X_i - Lidar_i}{Lidar_i}$ where RME_i is the relative mean error of carbon estimate for landscape i; X_i is the carbon density estimate i; $LiDAR_i$ is the LiDAR based estimate of carbon density for landscape i.

Allometric equations for each site were chosen from the best available published equations. Criteria for selection included species, the size range of tree sampled for biomass, and the forest habitat where tree biomass data were collected. The expectation was that these local equations would produce the most accurate estimate of forest carbon density for the site. Given the consistency in the remote sensing input, field sampling methodology, and data analysis, we set the LiDAR estimates of carbon density as the benchmark for comparison. We then obtained LandfireC (Battles et al. 2014) and LandTrendr-GNN (LT-GNN; Kennedy et al. unpublished) estimates for the same footprint at each site. The FIA regional biomass equations were used to parameterize LandFireC while two realizations of LT-GNN were developed using either the FIA CRM or the Jenkins equations. Both LandfireC and LT-GNN used forest inventory data from the FIA plots (Bechtold and Patterson 2005) to build the models.

LandfireC and LT-GNN$_{CRM}$ matched the LiDAR estimates with similar overall accuracy but LandfireC overestimated carbon density (mean RME = 0.10) while LT-GNN$_{CRM}$ tended to underestimate carbon (mean RME = -0.11). Of the two, the LT-GNN$_{CRM}$ estimates were more consistently close to the LiDAR value while for LandFireC, the three estimates varied from very close to very far from the LiDAR benchmark (Table A-2). LT-GNN$_{Jenkins}$ was the least accurate and most biased. It overestimated carbon density in all three cases with an average relative mean error of 0.19.

A.5 Conclusions and Future Directions

The performance of LandfireC and LT-GNN builds confidence in the prospects for a robust method for monitoring forest carbon in California using resources (i.e., Landsat, FIA) in the public domain. Both efforts were built independently, yet they converged toward the LiDAR-based benchmark with a mean relative error of approximately 10%. On the other hand, the choice of allometric equations clearly matters. The consistent bias in the LT-$GNN_{Jenkins}$ matches the difference between Jenkins and CRM results (Domke et al. 2012). In their comparison for trees in the Pacific Northwest, Zhou and Hemstrom (2009) showed that the Jenkins estimate was 17.3% larger than the estimate from the CRM method and 14.4% larger than the FIA regional (Table A-3). The difference between the CRM and FIA regional equations was much smaller (Table A-3).

Table A-3: Differences among allometric model estimates of aboveground live tree carbon density for a western conifer forest. Percent difference (%Diff) calculated from results in Zhou and Hemstrom 2009.

%Diff = (COL-ROW)/COL	FIA Regional	FIA CRM	Jenkins
FIA Regional	–	-3.5%	+14.4%
FIA CRM	+3.4%	–	+17.3%
Jenkins	-16.8%	-21.0%	–

An obvious conclusion is that the use of a consistent set of allometric equations would improve the precision of forest carbon monitoring in California as well as for the United States. And yet, a common practice is to maintain multiple approaches. Jenkins is by far the easiest to calculate and requires the least amount of field data (diameter, species group). Both the FIA regional and CRM methods require tree diameter as well as tree height measurements. Also, the volume equations in the FIA regional are complex. For example, the volume equation for quaking aspen has 18 fitted coefficients applied in a series of complicated linear equations (Woodall et al. 2011, p. 10). The CRM method inherits this complexity and extends it with several extra steps.

These problems are not limited to California. Weiskittel et al. (2015b) outline a national effort to improve tree biomass estimates. Preliminary results from a new method, referred to as "CRM2," shows promise (Weiskittel et al. 2015a). However, based on recent discussions with our federal colleagues, the decision to revise the FIA methodology is 18 to 24 months away.

We shared the insights from this review with our agency partners. Clearly a consistent set of tree biomass equations is essential to evaluation of forest carbon monitoring methods. Among the available approaches, the FIA regional equations were preferred. The regional FIA office considers these equations more appropriate and accurate for forests in California than the national standard (PNW-FIADB 2015). Moreover, there is a legacy of their use in California forest assessments. Therefore, all of the analyses and evaluations in this project use consistent applications of the regional tree biomass equations maintained by the Pacific Northwest FIA program.

A.6 References

Achard, F., H.D. Eva, P. Mayaux, H.J. Stibig, and A. Belward. 2004. Improved estimates of net carbon emissions from land cover change in the tropics for the 1990s. Global Biogeochemical Cycles 18: GB2008. 10.1029/2003GB002142.

ARB (2016). California Air Resources Board. AB 32 Scoping Plan. http://www.arb.ca.gov/cc/scopingplan/scopingplan.htm

Bartholomé, E. and A.S. Belward. 2005. GLC2000: A new approach to global land cover mapping from Earth observation data. International Journal of Remote Sensing 26: 1959-1977.

Battles, J.J., P. Gonzalez, T. Robards, B.M. Collins, and D.S. Saah. 2014. Final Report: California Forest and Rangeland Greenhouse Gas Inventory Development. California Air Resources Board Agreement 10-778.

Bechtold, W.A. and P.L. Patterson. 2005. The enhanced forest inventory and analysis program: national sampling design and estimation procedures. Gen. Tech. Rep. SRS-80. Asheville, NC: U.S. Department of Agriculture, Forest Service, Southern Research Station.

Chave, J., C. Andalo, S. Brown, M. Cairns, J. Chambers, D. Eamus, H. Fölster, F. Fromard, N. Higuchi, and T. Kira. 2005. Tree allometry and improved estimation of carbon stocks and balance in tropical forests. Oecologia 145:87-99.

Chen, Q., G. V. Laurin, J. J. Battles, and D. Saah. 2012. Integration of airborne LiDAR and vegetation types derived from aerial photography for mapping aboveground live biomass. Remote Sensing of Environment 121:108-117.

Chojnacky, D. C., L. S. Heath, and J. C. Jenkins. 2014. Updated generalized biomass equations for North American tree species. Forestry 87:129-151.

DeFries, R., F. Achard, S. Brown, M. Herold, D. Murdiyarso, B. Schlamadinger, and C. de Souza. 2007. Earth observations for estimating greenhouse gas emissions from deforestation in developing countries. Environmental Science and Policy 10: 385-394.

Domke, G. M., C. W. Woodall, J. E. Smith, J. A. Westfall, and R. E. McRoberts. 2012. Consequences of alternative tree-level biomass estimation procedures on US forest carbon stock estimates. Forest Ecology and Management 270:108-116.

Duncanson, L., W. Huang, K. Johnson, A. Swatantran, R.E. McRoberts, and R.Dubayah. 2017. Implications of allometric model selection for county-level biomass mapping. Carbon Balance and Management 12:18.

Enquist, B. J., and K. J. Niklas. 2001. Invariant scaling relations across tree-dominated communities. Nature (London) 410:655-660.

FIA. 2010. Regional biomass equations used by Forest Inventory and Analysis Program to estimate bole, bark, and branches (updated 13-Jan-2010), USDA Forest Service, Pacific Northwest Research Station.

FIA. 2017. Forest inventory data for California from the Forest Inventory and Analysis Program. FIADB_1.6.0.02.

Gonzalez, P., G. P. Asner, J. J. Battles, M. A. Lefsky, K. M. Waring, and M. Palace. 2010. Forest carbon densities and uncertainties from LiDAR, QuickBird, and field measurements in California. Remote Sensing of Environment 114:1561-1575.

Harris, N.L., S. Brown, S.C. Hagen, S.S. Saatchi, S. Petrova, W. Salas, M.C. Hansen, P.V. Potapov, and A. Lotsch. 2012. Baseline map of carbon emissions from deforestation in tropical regions. Science 336: 1573-1576.

Huang, W., A. Swatantran, L. Duncanson, K. Johnson, D. Watkinson, K. Dolan, J. O'Neil-Dunne, G. Hurtt, and R. Dubayah. 2017. County-scale biomass map comparison: a case study for Sonoma, California. Carbon Management 8:417-434.

Jenkins, J. C., D. C. Chojnacky, L. S. Heath, and R. A. Birdsey. 2003. National-scale biomass estimators for United States tree species. Forest Science 49:12-35.

Loveland, T.R., B.C. Reed, J.F. Brown, D.O. Ohlen, Z. Zhu, L. Yang, and J.W. Merchant. 2000. Development of a global land cover characteristics database and IGBP DISCover from 1-km AVHRR data. International Journal of Remote Sensing 21: 1303-1330.

Melson, S. L., M. E. Harmon, J. S. Fried, and J. B. Domingo. 2011. Estimates of live-tree carbon stores in the Pacific Northwest are sensitive to model selection. Carbon Balance and Management 6:1.

Ohmann, J. L., and M. J. Gregory. 2002. Predictive mapping of forest composition and structure with direct gradient analysis and nearest neighbor imputation in coastal Oregon, USA. Canadian Journal of Forest Research 32: 725–741.

Ohmann, J. L., M. J. Gregory, H. M. Roberts, W. B. Cohen, R. E. Kennedy, and Z. Yang. 2012. Mapping change of older forest with nearest-neighbor imputation and Landsat time-series. Forest Ecology and Management 272: 13–25.

Pflugmacher, D., W. B. Cohen, R. E. Kennedy, and Z. Yang. 2014. Using Landsat-derived disturbance and recovery history and LiDAR to map forest biomass dynamics. Remote Sensing of Environment 151:124-137.

PNW-FIADB. 2015. A data dictionary and user guide for the PNW-FIADB database (national version FIADB 5.1.6). Compiler: J. Thompson. Forest Inventory and Analysis Program, Pacific Northwest Research Station, Portland, Oregon.

Rollins, M.G. 2009. LANDFIRE: a nationally consistent vegetation, wildland fire, and fuel assessment. International Journal of Wildland Fire 18: 235–249.

Saatchi, S.S., N.L. Harris, S. Brown, M. Lefsky, E.T.A. Mitchard, W. Salas, B.R. Zutta, W. Buermann, S.L. Lewis, S. Hagen, S. Petrova, L. White, M. Silman, and A. Morel. 2011. Benchmark map of forest carbon stocks in tropical regions across three continents. Proceedings of the National Academy of Sciences 108: 9899-9904.

Temesgen, H., D. Affleck, K. Poudel, A. Gray and J. Sessions. 2015. A review of the challenges and opportunities in estimating above ground forest biomass using tree-level models. Scandinavian Journal of Forest Research 30: 326-33.

Weiskittel, A., J. Frank, D. Walker, P. Radtke, D. Macfarlane, and J. Westfall. 2015a. Advancing individual tree biomass prediction: assessment and alternatives to the component ratio method. PNW-GTR-931. USDA Forest Service, Pacific Northwest Research Station.

Weiskittel, A. R., D. W. MacFarlane, P. J. Radtke, D. L. Affleck, H. Temesgen, C. W. Woodall, J. A. Westfall, and J. W. Coulston. 2015b. A call to improve methods for estimating tree biomass for regional and national assessments. Journal of Forestry 113:414-424.

Westfall, J. A. 2012. A comparison of above-ground dry-biomass estimators for trees in the northeastern United States. Northern Journal of Applied Forestry 29:26-34.

Whittaker, R. H., and G. M. Woodwell. 1968. Dimension and production relations of trees and shrubs in the Brookhaven Forest, New York. The Journal of Ecology56:1-25.

Whittaker, R. H., F. H. Bormann, G. E. Likens, and T. G. Siccama. 1974. The Hubbard Brook Ecosystem Study: Forest biomass and production. Ecological Monographs 44:233-254.

Woodall, C.W., L. S. Heath, G. M. Domke and M. C. Nichols. 2011. Methods and equations for estimating aboveground volume, biomass, and carbon for trees in the U.S forest inventory, 2010. Gen Tech Rep NRS-88 USDA Forest Service, Northern Research Station.

Zhao, F., Q. Guo, and M. Kelly. 2012. Allometric equation choice impacts LiDAR-based forest biomass estimates: A case study from the Sierra National Forest, CA. Agricultural and Forest Meteorology 165:64-72.

Zhou, X. and M.A. Hemstrom. 2009. Estimating aboveground tree biomass on forest land in the Pacific Northwest: a comparison of approaches. Res. Pap. PNW-RP-584. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station.

Zolkos, S. G., S. J. Goetz, and R. Dubayah. 2013. A meta-analysis of terrestrial aboveground biomass estimation using LiDAR remote sensing. Remote Sensing of Environment 128: 289-298.

APPENDIX B: Description of Development of GEE Capable Version of the LandTrendr Algorithm

B.1 Introduction

A central goal of our project is testing the hypothesis that a Landsat-based approach to mapping carbon represents a repeatable, cost-effective approach to measure forest carbon stocks in California. In our review (Appendix A), we found that the carbon density map developed using an existing version of our Landsat-based approach compared favorably to finer-resolution estimates from LiDAR. Those existing Landsat-based carbon density maps were produced for a different project, and only extended to the year 2012 because of reliance on a particular set of sensors. To move forward, we must move our entire Landsat-based approach into an operational mode. The Landsat-based approach involves two broad efforts: LandTrendr (Landsat-based detections of trends in disturbance and recovery, (Kennedy et al. 2010)) algorithms to extract a signal from the time-series of Landsat data, and the GNN (Gradient Nearest Neighbor; (Ohmann and Gregory 2002)) algorithms to estimate carbon density by linking the LandTrendr signal with forest plot data. This report focuses on the first effort, detailing how we have moved the computationally expensive and resource-intensive LandTrendr algorithms into Google's Earth Engine platform in a manner that allows incorporation of the evolving Landsat sensor suite.

B.2 Background

The LandTrendr algorithms distill a simplified, coherent signal from large Landsat time-series datasets. The algorithms are computationally intensive, running many different statistical tests and fitting routines on each 30 by 30m pixel in large Landsat image time-series. They also require large archives of Landsat imagery to be accessible to the algorithms. In the past, Co-Investigator Kennedy's lab has utilized expert remote sensing and data management professionals to shepherd vast Landsat archives through the algorithm to create maps. This is not tenable in an operational context for a state agency.

Google's Earth Engine (GEE) platform has substantial promise to overcome this challenge, but comes with challenges of its own. GEE already stores the entire Landsat archive on its cloud-based data servers, eliminating much of the need to store large datasets locally. Additionally, its image processing system distributes large computing loads across Google's cloud-based architecture, eliminating the need for powerful distributed local computing resources. Implementing new algorithms on this platform is not trivial, however. For algorithms that do not already exist in the core GEE library, developers must work directly with Google to translate existing code into the code of the GEE library, and then must pass a vetting process and testing process to ensure compatibility with the GEE platform.

Even if the algorithms can be ported to GEE, we must extend the functionality to seamlessly incorporate new satellite data. Our existing biomass maps relied heavily on two satellites, one of which no longer functions. A replacement sensor now exists, but integration must ensure that the new sensor faithfully links to the temporal record established by its predecessors.

Thus, at least two steps are needed to ensure that the LandTrendr core algorithms can be reliably run in an operational mode from here forward. First, the software that runs the

LandTrendr algorithms must be transferred faithfully to the programming languages underlying GEE. At the time of project onset, Kennedy and collaborator Yang had been working extensively with the GEE developers to convert the core LandTrendr algorithms into GEE environment, but further development was required, as well as testing to ensure that the translation was successful. Second, we must bring recent Landsat data into the same time-series environment to allow extension of our prior maps, which heretofore had only existed up to the year 2012. This required a new set of tests and development.

B.3 Translation of LandTrendr Algorithms to GEE

The core of LandTrendr is temporal segmentation, where each pixel's spectral time series is simplified into straightline segments separated by vertices or breakpoints (Figure B-1). Such simplification serves the dual purpose of distilling a large dataset into a few key parameters as well as reducing year-to-year noise. When the spectral index being treated relates to vegetation density, segments can be labeled as capturing vegetative growth, loss, or stability. Segments that show abrupt (short-duration) loss of vegetation are typically related to disturbance, such as clearing of forest for urban development or harvest.

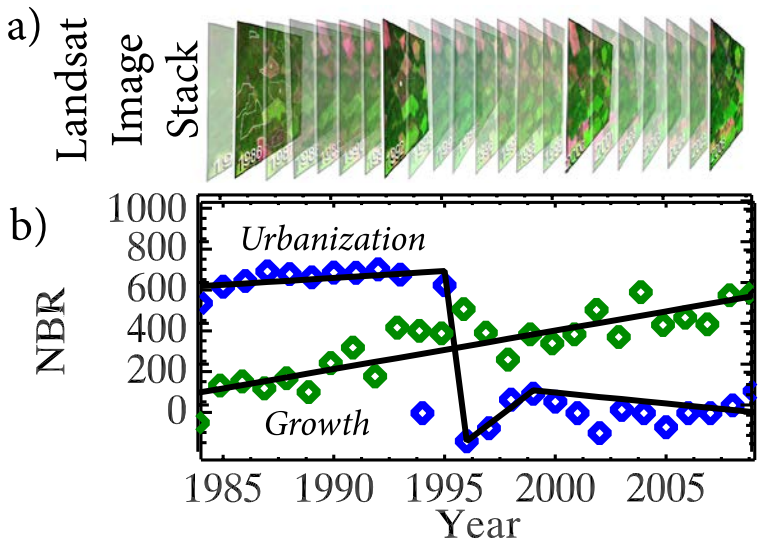


Figure B-1: Temporal segmentation of the Landsat archive. From a stack of cleaned yearly Landsat Thematic Mapper ™ images (a), algorithms fits straight line segments (black lines) to each pixel's spectral trajectory (colored symbols). The segments capture the salient differences among land surface processes.

The algorithms to conduct temporal segmentation were originally written in the programming language IDL (Interactive Data Language). Noel Gorelick, one of the original developers of GEE, worked with Zhiqiang Yang, a collaborator of Kennedy, to translate the code into core library functions of GEE. This process had already begun when the current project was proposed, but required further debugging and subsequent testing. Through iterative testing and improvement, the algorithm was successfully run over large test areas in early 2016, leading to the phase of testing conducted under this project.

To test the implementation of LandTrendr in GEE, we compared the LT-GEE algorithm's temporal segmentation to expert interpretation of time-series done at more than 5,000 sample sites nationally. The expert dataset was developed previously for a separate project by Dr. Warren Cohen of the USDA Forest Service. To collect validation points, Cohen utilized the TimeSync interpretation interface that he and our collaborator Yang developed specifically to validate temporal segmentation outputs from LandTrendr (Cohen et al. 2010). TimeSync interpretation follows a standard protocol. Interpreting time series of spectral data for individual Landsat pixels (30m by 30m in size), the TimeSync analyst manually defines the breakpoints that bound straight-line segments. Because these segments conceptually match the straight-line segments of the LandTrendr algorithm, the comparison is one of testing whether breakpoints found by the algorithm match those found by the analyst.

For these 5000+ nationally-distributed points, we evaluated agreement in several ways. First, we assigned the label disturbance, stable, or growth to every single year of every interpreted TimeSync plot and similarly to each plot's corresponding LT-GEE temporal segmentation score (resulting in 96,606 comparisons, Table B-1). Because of the potential for slight geographic mis-registration between the TimeSync plot and the LT-GEE data, we used the mean value of a three-by-three pixel window centered on TimeSync-interpreted plot location for the LT-GEE data.

Table B-1: Accuracy assessment of TimeSync plots and Landtrendr-GEE temporal segmentation score.

		LANDTRENDR-GEE				
		DISTURBANCE	RECOVERY	STABLE	No. Truth	Producers Acc
TIMESYNC	DISTURBANCE	**1527**	133	1556	3216	0.47
	RECOVERY	1154	**9158**	28898	39210	0.23
	STABLE	2632	2482	**49066**	54180	0.91
	No. Predictions	5313	11773	79520	96606	
	Users Accuracy	0.29	0.78	0.62	Overall	0.62

Some caution is needed in interpreting these results. The TimeSync interpretation process brings substantially more information to bear on labeling, including high resolution aerial photographs, as well as spatial context and the natural human ability to infer pattern from subtle spatial cues. Thus, the analyst can label temporal segments for very subtle phenomena, and can even use expert knowledge to label a segment in a somewhat counterintuitive manner: a temporal drop in a vegetation index may still be labeled recovery or stable, for example, because of a larger signal inferred by the analyst. The time-series algorithm, on the other hand, only works with the numeric values of a single spectral time series and can take none of the context into account. Thus, these comparisons with TimeSync data represent the highest bar of accuracy standard possible. In a similar comparison of 10 different change detection algorithms (including but not limited to LandTrendr) conducted for the project under which Cohen originally interpreted the TimeSync plots, agreement scores such as these were the norm, if not at the higher end of accuracy.

Because of these issues, it is important to further parse the comparison results. First, we split the TimeSync disturbance and recovery categories into their constituent process types, and compared those with the magnitude of spectral change recorded by LandTrendr-GEE (Table B-2). Several key patterns emerge: 1) Where LandTrendr-GEE overestimates disturbance, it is primarily for low-magnitude processes; 2) Where LandTrendr-GEE underestimates disturbance, it is similarly for low-magnitude processes; and 3) Stability and recovery dominate the dataset, and disturbance processes are an extreme minority of observations. This latter fact obscures the ability of the algorithm to detect disturbance. Moreover, the strict year-by-year comparison adds a potential double penalty to the algorithm in cases where noise or haze mask the algorithm's ability to match disturbance event timing perfectly. Off by one year, the algorithm would receive double penalties of false negative in one year and false positive in the next.

Thus, we conducted a final parsing of the dataset. Here, we focused only on traditional disturbance events labeled by the TimeSync analyst, such as fire, harvest, mechanical clearing, and wind. We then evaluated whether the LandTrendr-GEE algorithm captured the event within one year of the analyst's call. Of the 1895 such individual TimeSync derived events, the algorithm correctly captured 1364 (or 72%) of them.

In sum, our evaluation showed that the GEE implementation of the LandTrendr algorithm appeared to be functioning in a manner consistent with the existing published version of the code.

Table B-2: Parsing of LandTrendr-GEE and Timesync agreement by category and magnitude.

		LandTrendr-GEE Disturbance				LandTrendr-GEE Stable			LandTrendr-GEE Recovery			
		Greater change			Subtle change				Subtle change			Greater char
		-1500_-250	-250_-50	-50_-10	-10_-5	-5_-0.001	0	0.001_5	5_10	10_50	50_250	250
TimeSync Disturbance	DEBRIS	0	0	1	0	0	0	0	0	0	0	0
	DELAY	0	13	21	3	7	2	3	2	11	3	0
	FIRE	51	44	27	5	11	1	11	1	6	3	0
	HARVEST	241	460	266	67	109	31	101	53	47	6	11
	MECHANICAL	4	32	40	15	16	3	12	6	10	5	2
	OTHER DISTURBANCE	1	3	2	0	3	1	3	2	2	1	0
	SITE-PREPARATION FIRE	10	41	21	2	7	1	5	2	20	9	1
	STRESS	2	27	225	238	421	14	335	35	5	1	0
	WATER	0	0	6	5	3	2	0	0	1	0	0
	WIND	0	12	11	4	7	14	10	0	2	0	1
	RECOVERY	24	173	957	881	3596	1219	15385	7817	7728	1100	330
Stable	STABLE	30	294	2251	2380	13536	2254	26153	4529	1790	102	564
	OTHER NON-DISTURBANCE	0	1	22	18	96	3	71	9	8	2	2

B.4 Integration of Landsat 8 Reflectance Data

Landsat 8 is the most recent version of the Landsat sensor family, and will be critical for ongoing monitoring of forest condition. Launched in 2013, the Landsat 8 OLI (Operational Land Imager) spectral sensor was designed to carry on the legacy of the sensors that had been carried on Landsats 5 and 7. Those latter sensors were the data on which the prior LandTrendr-derived biomass maps had been produced. While Landsat 7 still records image data, Landsat 5 ceased functioning in November 2011. To ensure continuity of measurement, we need to ensure that Landsat 8 OLI data could be seamlessly included in the time series of the prior image archive.

Because the design characteristics of the Landsat 8 OLI sensor are slightly different from those of Landsat 7, a data transformation is needed. We explored two possible approaches to transform Landsat 8 spectral data to match with Landsat 7 spectral data. The first approach was to utilize transformation coefficients cited in Roy et al. (2016), developed as regressions using more than 59 million individual pixels in near-coincident Landsat 7 and Landsat 8 images over the continental United States. Recognizing that the differences in spectral bands between the sensors may cause vegetation-specific changes, we chose to test whether a more localized correction approach may be more effective. -This second method was developed in the Kennedy lab by Dr. Joseph Hughes, a postdoc working on this project. Approximately 4000 small regions (0.1 x 0.1 degrees) were selected across the globe in areas where Landsats 7 and 8 overlapped (on scene margins), and biome type noted. Surface reflectance values were obtained for each sensor for near-coincident observations (defined as being within 5 days of each other). We omitted cloud and shadows, as well pixels that were water or snow in one image but not the other, both as determined by the cloud/shadow algorithm CFMask (the standard used by the USGS). Large changes in spectral value (greater than 0.3 reflectance units) were also omitted, assuming they corresponded to either real change on the ground or to inadequate cloud screening. Regressions between L8 and L7 bands were calculated for each global biome type, and coefficients stored for correction at the biome level.

We then tested the impact of these two transformation approaches on the temporal segmentation process. Two 1° x 1° locations were examined between 1985 and 2016: 1. A region in the western Willamette Valley of Oregon, with strong disturbance signals from forest harvest in the coastal range; and 2. An area around the Great Smoky Mountains, a protected area with low disturbance. We chose these areas to test a range of forest types and to ensure that our analysts could interpret the potential change patterns in each study area. Both locations are a mix of forest, agriculture, and urban; all analyses are limited to forested areas (pixels with median NBR [normalized burn ratio spectral index] over the timeseries > 0.6). For all pixels (30 by 30m cell size) in each study area, we ran the LT-GEE temporal segmentation on time-series of surface reflectance data from the appropriate sensor, and noted the vertex years found by the algorithm. We documented vertex values for fitting with no correction, with the Roy et al. transformations, and with our biome-specific transformation (noted below as the "mjh" transformation). For each of those conditions, we tested two conditions within the Landsat 8 era: where all possible images are allowed from either Landsat 7 or Landsat 8 sensors, and where only Landsat 8 images are allowed after Landsat 8 data became available. If the correction of the L8 data were incorrect, we would expect to see a step function (and hence, more vertices found by the GEE-LandTrendr algorithm) in the 2012 and 2013 time period.

In both study areas, we found evidence to suggest that the generic coefficients reported by Roy et al. (2016) are sufficient to bring Landsat 8 data into agreement with Landsat 7 data. In the Willamette Valley study area (Figure B-2), the Roy et al. coefficients resulted in the fewest vertices in 2012, suggesting a better correction of Landsat 8 data.

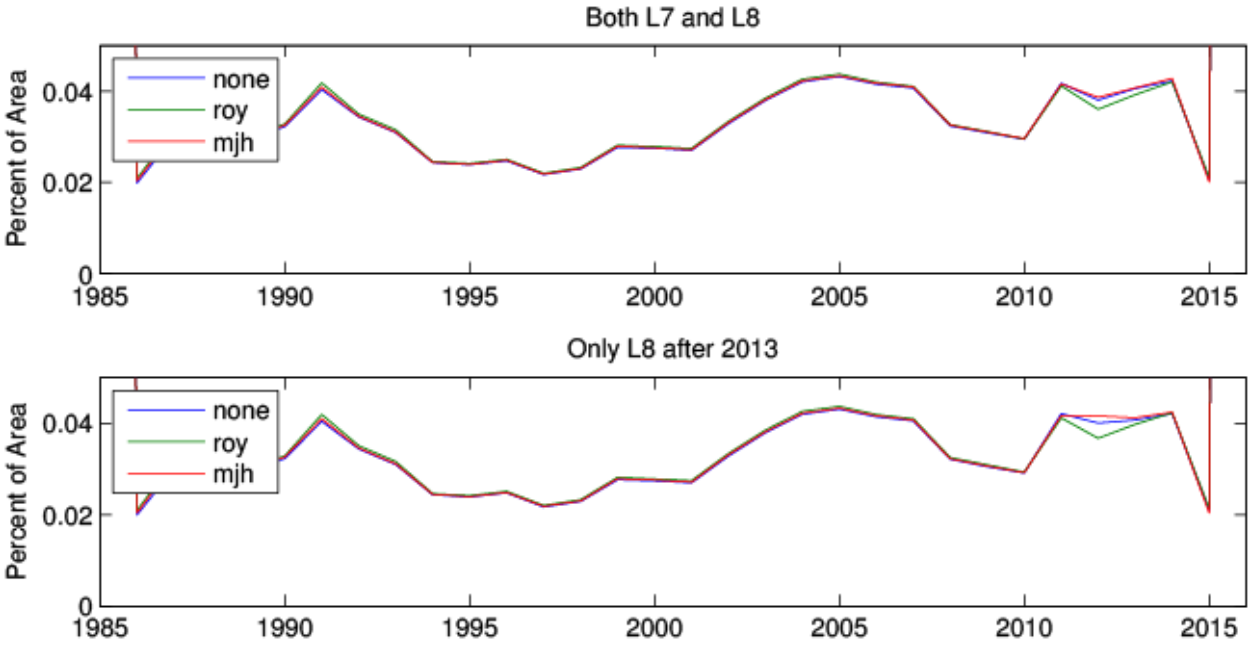


Figure B-2: Vertex counts from GEE-LandTrendr fitting in the Willamette Valley study area. Vertex counts in the year 2012 and 2013 are lowest for the Roy et al. (2016) transformation coefficients, suggesting they introduce the least artifacts into the time-series fitting.

In the Smokey Mountains study area, neither transformation was appreciably different from applying no transformation at all (Figure B-3).

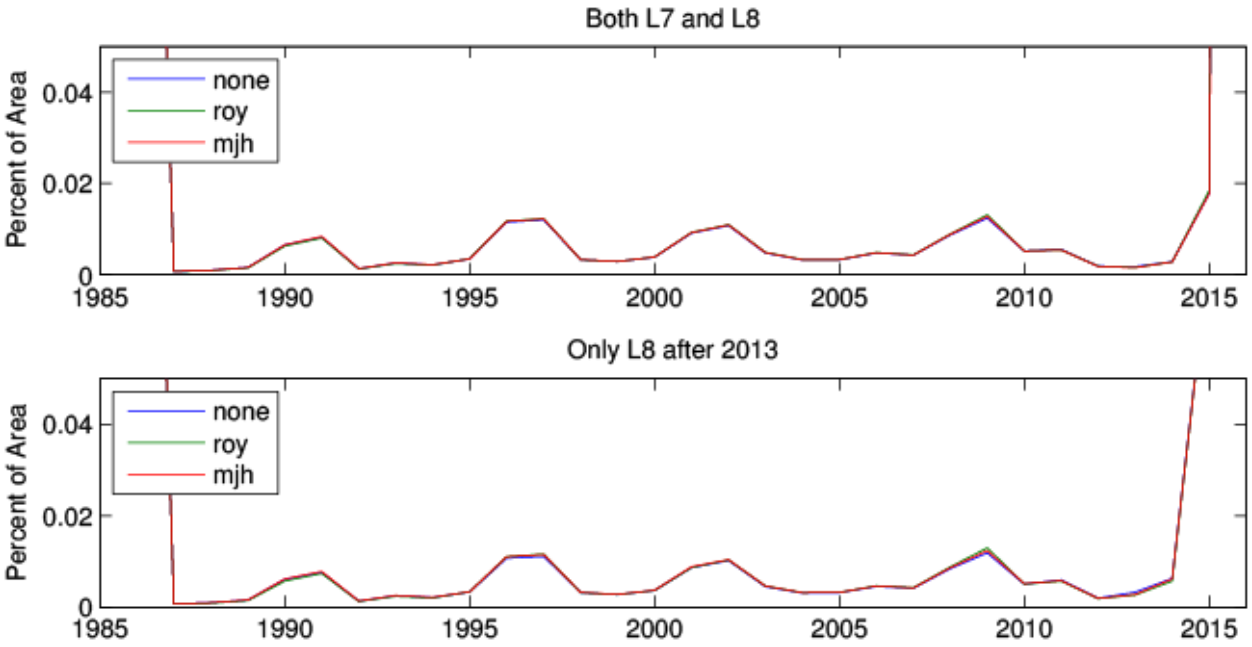


Figure B-3: Vertex counts from GEE-LandTrendr fitting in the Smokey Mountains study area. No appreciable effect of the Landsat 8 transition is seen in the 2012-2013 era for any of the approaches, suggesting that any of the transformation (including no transformation) would be appropriate.

From these tests, we concluded that best practices argued for the use of the Roy et al. (2016) coefficients to integrate Landsat 8 data into our GEE-LandTrendr temporal segmentation. In the case where little correction was needed (the Smokey Mountains example), the Roy coefficients caused no harm. In the case where disturbance was common (the Willamette Valley example), the Roy coefficients appeared to be the most reliable. Moreover, the methodology of the Roy transformations is public and well-documented.

Thus, from this point forward, we applied the Roy et al. (2016) coefficients Landsat 8 OLI data for our time-series analysis.

B.5 Extension of Segmentation Results to Other Indices

Temporal segmentation is the core of the LandTrendr process, but only operates on a single spectral index at a time. When chosen carefully, that single index may reasonably capture the timing of a diverse range of events. However, no single spectral index can adequately characterize the depth of information needed to model biomass. Therefore, an essential next step is to extend the segmentation information from the original index to additional spectral indices. Described in (Kennedy et al. 2015), the process involves forcing the target indices into a temporal segmentation (temporal fitting) that matches the timing of the segmentation determined by the original index (temporal congruence). Temporal fitting and temporal congruence each play an important role. The former removes the noise from the target indices, while the latter ensures that no unrealistic combinations of indices occur. We call the process "fitting to vertices", or FTV, and like the original segmentation, this project required that we move the code into the core GEE code libraries.

Through the late summer and early fall of 2016, collaborator Yang worked with the GEE team to test and debug these FTV routines. For our test case, we used an area of the Sierra Nevada Mountains. Our goal was to create images of the indices in the tasseled-cap transformation (Kauth and Thomas 1976): the tasseled-cap Brightness (TCB), Greenness (TCG), and Wetness (TCW) indices. The three indices mathematically capture most of the variability in the original Landsat spectral signal, but far more efficiently. They form the foundation of the biomass-estimation process we developed for our prior forest-biomass mapping project with the USDA.

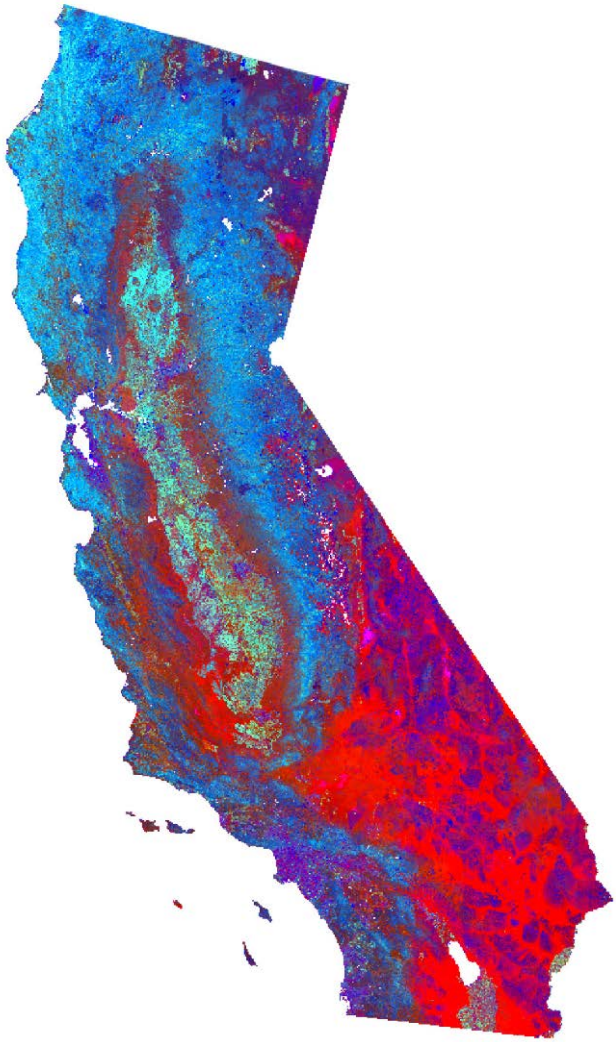


Figure B-4: GEE-LandTrendr-derived FTV-fitted tasseled-cap imagery for the year 2010 or the state of California. Blue tones indicate forest, red tones indicate barren, and purples indicate urban. Similar images were generated for every year from 1986 to 2016.

The resultant images met the goals of the FTV process. Images for the entire state of California (Figure B-4) were generated from 1986 to 2016, using source imagery from Landsats 5, 7, and 8. The images showed minimal to no residual impacts of cloud or cloud shadow, and no systematic bias at the Landsat 7/8 transition.

More importantly, the FTV images appropriately capture temporal evolution of spectral signal. As an example, we examined the spectral properties of the forests near Sequoia and Kings Canyon National Parks in the southern Sierras from 2010 to 2015. These forests are among those that have experienced stress and mortality from the California drought. As expected, the spectral properties of the FTV images show spatial patterns of change in indices related to vegetation and canopy structure (Figure B-5).

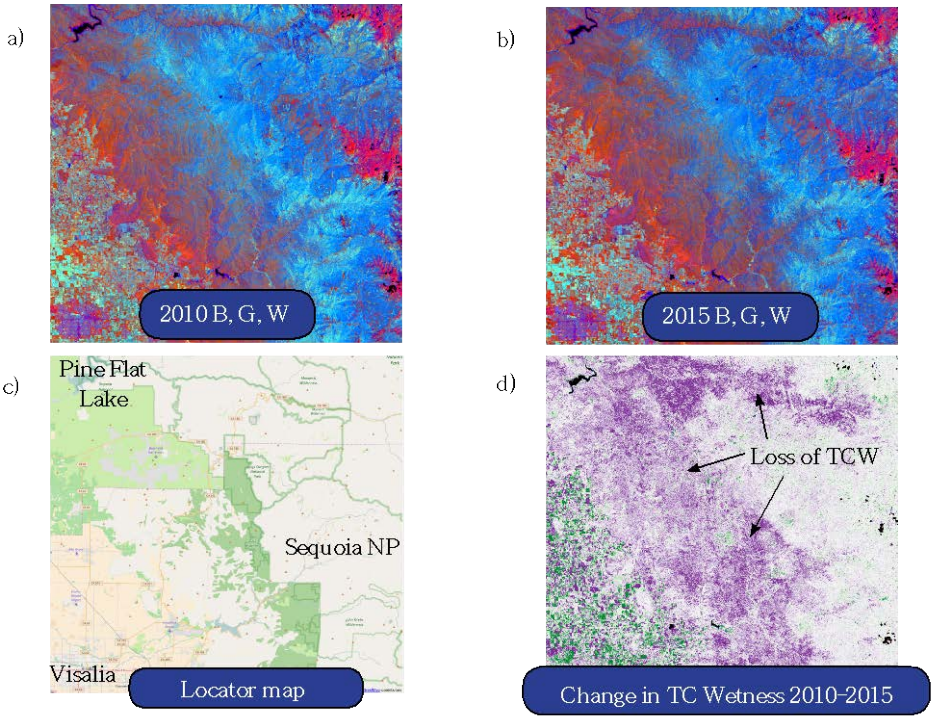


Figure B-5: Utilizing fitted imagery to track forest change concurrent with drought in the southern Sierras. Tasseled-cap Brightness (B), Greenness (G), and Wetness (W) displayed in the red, green, and blue colors, respectively, for the year (a) 2010 and (b) 2015, for the area (c) including parts of Sequoia National Park and Kings Canyon National Park. Broad swaths of loss in TCW (shown in purple tones in (d)) are typically associated with loss in vigor or in mortality amount conifer dominated forests.

As key deliverable from the effort, we handed the FTV images off to our collaborators at the USDA Forest Service (Dr. David Bell) and at Oregon State University (Matt Gregory and Heather Roberts). They will develop statistical approaches to link these data with USDA Forest Service Forest Inventory and Analysis (FIA) plot data to generate yearly maps of aboveground forest biomass.

B.6 References

Cohen, W.B., Zhiqiang, Y., & Kennedy, R.E. 2010. Detecting Trends in Forest Disturbance and Recovery using Yearly Landsat Time Series: 2. TimeSync - Tools for Calibration and Validation. Remote Sensing of Environment, 114: 2911-2924.

Kauth, R.J., & Thomas, G.S. 1976. The tasselled-cap: a graphic description of the spectral-temporal development of agricultural crops as seen by Landsat. In: Second Annual Symposium on the Machine Processing of Remotely Sensed Data. West Lafayette, IN: Purdue University Laboratory of Applied Remote Sensing.

Kennedy, R.E., Yang, Z., Braaten, J., Thompson, C., Antonova, N., Jordan, C., & Nelson, P. 2015. Attribution of disturbance change agent from Landsat time-series in support of habitat monitoring in the Puget Sound region, USA. Remote Sensing of Environment, 166: 271-285.

Kennedy, R.E., Yang, Z., & Cohen, W.B. 2010. Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr - Temporal segmentation algorithms. Remote Sensing of Environment, 114: 2897-2910.

Ohmann, J.L., & Gregory, M.J. 2002. Predictive mapping of forest composition and structure with direct gradient analysis and nearest-neighbor imputation in coastal Oregon, U.S.A. Canadian Journal of Forest Research, 32: 724-741.

Roy, D.P., Kovalskyy, V., Zhang, H.K., Vermote, E.F., Yan, L., Kumar, S.S., & Egorov, A. 2016. Characterization of Landsat-7 to Landsat-8 reflective wavelength and normalized difference vegetation index continuity. Remote Sensing of Environment, 185: 57-70.

APPENDIX C: Detailed Methods: LandTrendr and GNN Imputation to Produce California Map of Aboveground Live Biomass

C.1 Introduction

Assessment and monitoring of biomass and carbon patterns and dynamics across forest landscapes is needed to support management and decision making processes, highlighting the need for spatially complete (e.g., all forest lands in California) and temporally deep (e.g., 1990 baseline to present) vegetation mapping products. This research project aims to continue development of a repeatable, cost-effective, Landsat-based approach to forest carbon stock mapping in California. We have compared existing moderate-resolution Landsat-based and finer-resolution LiDAR based carbon density maps representing annual forest conditions from 1990-2012 (Appendix A). We have also developed a new framework for assimilating the full Landsat 5, 7, and 8 archives (1984-2016) into the Landsat-based detections of trends in disturbance and recovery algorithm (LandTrendr; Kennedy et al. 2010) using the Google Earth Engine (GEE) platform to distribute computational demands within a cloud-computing environment (Appendix B). LandTrendr not only helps to identify disturbance patches, but also extracts the signal (i.e., reduces sensor noise) from the Landsat time series (LTS), improving the spatial and temporal consistency of vegetation, and thus forest carbon stock, and mapping produced using statistical algorithms, such as the gradient nearest neighbor approach (GNN; Ohmann and Gregory 2002, Ohmann et al. 2012). This report describes the integration of LandTrendr from GEE into the GNN framework to produce new disturbance and biomass maps from 1990-2016, along with associated measures of uncertainty.

C.2 Background

The LandTrendr-GNN framework provides a valuable approach for translating the multi-decadal Landsat archive into a spatially and temporally explicit representation of forest biomass, and thus carbon stock, patterns. The LandTrendr-GNN framework was developed by the Landscape Ecology, Modeling, Mapping and Analysis (LEMMA) team, a collaborative Oregon State University and USDA Forest Service research team currently led by Co-Investigator Bell. The LandTrendr-GNN framework involves three phases (Figure C-1): LandTrendr algorithm implementation, integration of LandTrendr and GNN, and plot-based analyses and comparisons. The LandTrendr algorithm (Figure C-1; LandTrendr) uses pixel-level temporal segmentation procedures to remove noise from the LTS and retains the major trends in the spectral signal through time (described in detail in Kennedy et al. 2010). The resulting temporally-smoothed LTS data can then be used for biomass modeling purposes.

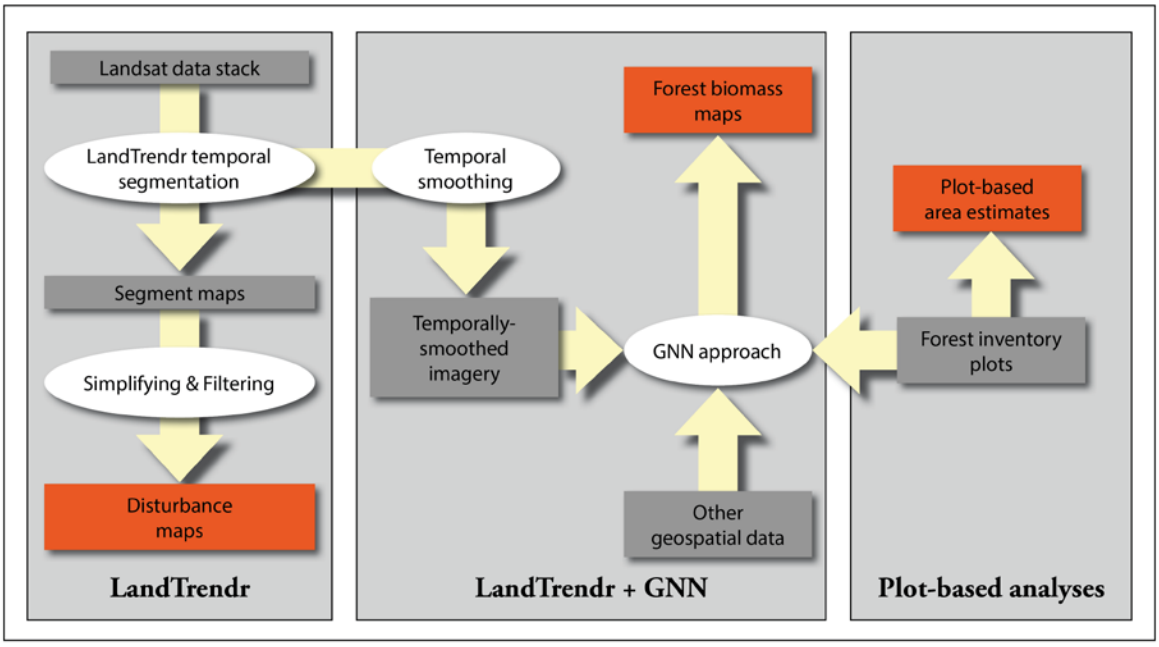


Figure C-1: Representation of the LandTrendr-GNN framework

Temporally-smoothed LTS imagery can be integrated with other geospatial data and forest inventory data in the GNN modeling method to produce forest biomass maps (Figure C-1; LandTrendr+GNN). GNN is a *k* nearest neighbor (*k*NN) imputation methodology designed explicitly for landscape mapping (Ohmann and Gregory 2002). Imputation mapping is a method for substituting observed values (i.e., forest inventory data) to replace missing data for all pixels in an area, resulting in a map as output. Specifically, a *k*NN imputation assigns those missing values based on some set of observations that are similar to the pixel of interest in terms of environmental conditions (i.e., minimizes distance in environmental or gradient space). In the case of GNN, the environmental space is defined by a constrained ordination (direct gradient analysis) – canonical correspondence analysis (CCA; ter Braak 1986) that relates forest attributes from forest inventory plots (i.e., species matrix) to geospatial data extracted at those same plot locations (i.e., environment matrix) – for measuring and weighting distances in nearest neighbor calculations (Ohmann and Gregory 2002). The selection of variables to include in the species and environment matrices represent the modelers' underlying assumptions and will impact the eventual maps. Past experience indicates that, depending on the physiographic region being modeled and mapped (hereafter modeling region, or MR), the relationship between LTS data and forest structure depends on other environmental variables, such as climate, topography, and location (Ohmann et al. 2014).

C.3 Forest Attribute Data

The forest attributes impact the GNN imputation mapping in two ways. First, they enter the CCA as the species matrix and help to define the gradient space with which neighbors are identified and assigned during imputation. The species matrix contains tree basal area grouped by species and tree size class to represent both compositional and structural variation in forest ecosystems (Ohmann et al. 2012). Second, a broader suite of forest attributes can be imputed to individual pixels either from the nearest neighbor directly (i.e., $k = 1$) or the mean of a set of neighbors (i.e., $k > 1$). In this case, we are interested in aboveground live forest biomass which is calculated based on individual tree data for each plot. Three modifications to existing forest attribute data were undertaken for the current project to improve model performance and meet the needs of the end-users: refined plot selection, plot screening to remove outliers, and revised biomass equations.

C.3.1 Plot Selection

In the past, vegetation maps produced using GNN have relied on a mix of forest inventory plots representing an approximate 1-ha footprint, but plot designs still differed (Ohmann and Gregory 2002; Ohmann et al. 2012). For example, U.S. Forest Service Region 6 Current Vegetation Survey (CVS) and U.S. Forest Service Forest Inventory and Analysis (FIA) plots used different numbers of subplots (5 vs. 4), used different diameter thresholds for nested plots, and were measured during different periods of time (1994-2007 vs. 2001-present; Max et al. 1996, Bechtold and Patterson 2005). These differences in design result in differing sampling intensities, and thus differing probability of including certain trees. As a result, including multiple sampling designs can introduce artifacts into vegetation maps.

As part of the current project, we assessed whether changes in sampling design substantially impacted annual time-series of biomass mapping products produced using the LandTrendr-GNN framework. From summer 2016 to fall 2016, Matt Gregory, Heather Roberts, and David Bell examined temporal trends in existing LandTrendr-GNN biomass maps and compared them with maps produced using only FIA annual plots in a subset of modeling regions (MRs 223-227; Figure C-2). These modeling regions were selected because the large number of non-FIA annual plot design plots (i.e., Region 6 CVS plots) present in these modeling regions would result in the most extreme manifestation of mapping artifacts. For the purposes of this exercise, previous modeling workflows and imagery were used, only allowing the forest inventory plots entering the modeling to vary. We found that artifacts caused by changing plot designs were common, especially for threshold-based forest attributes. For example, CVS plots examine a larger footprint for the largest trees, meaning that a shift from CVS to FIA plots occurring in the mid-2000s tended to be associated with erratic inter-annual fluctuations in biomass predictions and an eventual decline in forest biomass predictions for which no clear mechanism was observed (i.e., large disturbance or mortality event). To avoid this erratic behavior, the LEMMA team decided to use FIA annual plots alone for modeling and mapping in the current project. Focusing on a single plot design was not previously possible as a full cycle of FIA annual plot measurements was not completed until 2011.

Figure C-2: Map of modeling regions (MRs) used in this project

C.3.2 Plot Screening

To improve model fitting and eventual forest biomass maps, Bell screened >2,000 FIA plots identified as potential outliers from August 2016 to January 2017. In part, the large number of outliers was due to the addition of 7487 plot measurements that had not previously been incorporated into vegetation mapping for the study area. Of the new plot measurements, 3404 occurred between 2012 and 2015 on the FIA national grid, and 4083 occurred between 2001 and 2015 on R6 CVS locations outside of the FIA national grid. All of the new plots were measured by FIA field crews using FIA's annual sampling protocol. Outliers were identified based on

several criteria, including, but not limited to, inaccurate vegetation predictions and overly variable predictions within the plot footprint. For each outlier, Bell examined tassel-capped brightness, greenness, and wetness from LandTrendr along with color images from the National Agriculture Imagery Program (NAIP) at the plot location. Plots encompassing (1) more than one vegetation condition (e.g., straddling the boundary between a plantation and an old-growth forest), (2) human structures and modifications to the landscape (e.g., roads), and (3) waterways were excluded from modeling. Plots where a disturbance occurred between plot measurement and Landsat imagery dates, or vice versa, were excluded. Finally, plots where the tree data did not appear to be representative of the Landsat or NAIP imagery, potentially indicating inaccurate plot coordinates or missing tree data, were excluded. After all outlier plots were examined, models were run again and new outlier plots were assessed iteratively. When no new outliers were identified, the plot screening was considered complete. This plot screening resulted in an additional 5978 FIA annual plots available for modeling.

C.3.3 Biomass Equations

Biomass equations were updated to reflect regional biomass functions, as used for LandFire (Appendix A). LEMMA already uses both CRM (Heath et al. 2009) and Jenkins (Jenkins et al. 2003) biomass algorithms for aboveground live biomass estimation. From November 2016 to February 2017, Heather Roberts (LEMMA database manager) implemented regional biomass equations (FIA 2010) within our database structure. All biomass equations are applied at the individual tree level and expanded aboveground live tree biomass to the scale of plots (kg ha-1) by multiplying individual tree biomass (kg) by expansion factors (i.e., trees ha-1 represented by each tree). As noted in Appendix A, regional biomass equations are associated with merchantable volumes, while CRM and Jenkins methods are associated with total volumes.

Although the source of LEMMA's data is FIA's National Inventory Management System (NIMS), which contains tree-level regional biomass values, updates made to both databases after LEMMA loaded the data prevented us from importing biomass directly from NIMS for older plots. As part of LEMMA's QA/QC process, we compared our regional biomass values with those in NIMS for plots for which we were confident that no database updates had occurred. When we found some differences in regional biomass between LEMMA and NIMS, we contacted FIA and discovered that they had updated some of their volume equations in 2014. We decided to stick with the FIA volume equations documented in 2010 to be consistent with the methods used to calculate volume on the LiDAR plots. After excluding species whose volume equations had changed, LEMMA's regional biomass values matched FIA's values. In a few cases, bark equations and branch equations return negative biomass values for trees with small diameters and heights. LEMMA followed FIA's lead and zeroed out all negative bark and branch components.

Due to difficulties in differentiating certain fir species (genus *Abies*) in the field, LEMMA has created some species codes that represent multiple species. ABCO (white fir) and ABGR (grand fir) are assigned to ABGRC, and ABMA (California red fir), ABPR (noble fir) and ABSH (Shasta red fir) are assigned to ABPRSH. Since we assign the new codes before we calculate volume, we use the same specific gravity value and volume equation for all species assigned to a combination code. We used the specific gravity and volume equation for the most commonly occurring species in the group. Since we retain the original species codes, we have the option of calculating volume based on the original species codes in future versions of the GNN models.

C.4 Geospatial Data

GNN imputation is, at its heart, a mapping approach that provides a wall-to-wall representation of forest vegetation. As a result, spatially and temporally complete geospatial data are needed as inputs to GNN in order to produce mapped predictions for all pixels and years. These geospatial data (i.e., rasters) enter GNN in two ways. First, both LTS imagery (in this case from LandTrendr) and environmental data (Table 1) for each forest inventory plot included in the species matrix are used to construct the environment matrix as an input into the CCA. Second, the same environmental variables are used to calculate the plot and pixel locations in gradient space, given the results of the CCA. As a consequence, plots can be attributed to pixels as the basis of the *k*NN imputation within GNN. As part of the current project, we examined a new method for selecting candidate Landsat TM images based on a mediod method and explored new combinations of geospatial predictors contained in the environment matrix. In particular, new geospatial data included change metrics for LTS data and soils information (Miller and White 1998), whereas previous GNN mapping has generally included LTS, climate, location, and topographic data (Ohmann et al. 2002, 2012, 2014; Bell et al. 2015).

Table C-1: Geospatial data used in LandTrendr-GNN Framework

Subset	Code	Description
Landsat	TC1	Brightness (i.e., axis 1 of the tassel cap transformation)
	TC2	Greenness (i.e., axis 2 of the tassel cap transformation)
	TC3	Wetness (i.e., axis 3 of the tassel cap transformation)
	NBR	Normalized burn ratio (unitless)
Change	ΔTC1	Mean change in TC1 during previous 6 years
	ΔTC2	Mean change in TC2 during previous 6 years
	ΔTC3	Mean change in TC3 during previous 6 years
	ΔNBR	Mean change in NBR during previous 6 years
Climate	ANNPRE	Mean annual precipitation (ln[mm])
	ANNTMP	Mean annual temperature (°C)
	AUGMAXT	Mean maximum temperature of August (°C)
	DECMINT	Mean minimum temperature of December (°C)
	SMRTP	Ratio of mean temperature (°C) to precipitation (ln[mm]) of May-Sept.
Location	COASTPROX	Distance to the Pacific Ocean (km)
	LAT	Latitude (°)
	LON	Longitude (°)
Topography	ASPTR	Cosine transformation of aspect
	DEM	Elevation from a digital elevation map (m)
	PRR	Potential relative radiation (unitless)
	SLPPCT	Slope (%)
	TPI450	Topographic position index, difference between elevation and mean elevation within a 450-m radius window
Soils	ROCKDEPTH	Rock depth (cm)
	BD_30	Bulk density of soils 0 cm – 30 cm (g cm^{-3})
	PERM_30	Permeability of soils 0 cm – 30 cm (m^2)
	PH_30	Mean pH of soils 0 cm – 30 cm
	RVOL_30	Rock volume of soils 0 cm – 30 cm (cm^3)

C.4.1 Mediod image mosaics

In Fall 2016, new LandTrendr-based spectral data were extracted from GEE using the methods described in Appendix B, including normalized burn ratio (NBR) and three tassel cap indices (brightness, greenness, and wetness) for each year and 30m pixel in the modeling area (Figure C-2). During December 2016/January 2017, Co-Investigators Bell and Kennedy and Collaborators Gregory and Yang discovered LTS imagery quality issues (e.g., clouds) arising from the method for selecting candidate images used in the construction of yearly mosaics that feed into the LandTrendr algorithm. While there are algorithms designed to maintain high quality data of surface reflectance data (e.g., cloud masking using Fmask; Zhu et al. 2015), anomalous observations were still quite common within individual images. Since the Landsat images for LandTrendr were selected based on a target date (i.e., the observation for a given pixel is the clear image with the closest acquisition date to the target date), anomalous observations were often retained and impacted the LandTrendr segmentation, and thus the temporally smoothed imagery used by GNN. In January 2017, Collaborators Gregory and Yang developed a mediod method for image selection which avoids these anomalous observations. The mediod method takes a window of time around the target date, July 1 to August 31 in this case, calculates the median of each pixel for six spectral variables (TM bands 1-5, 7), and then identifies the image with the most similar set of observations for that pixel. As a result, extreme values are avoided and the quality of the imagery entering and exiting LandTrendr are much improved (Figure C-3).

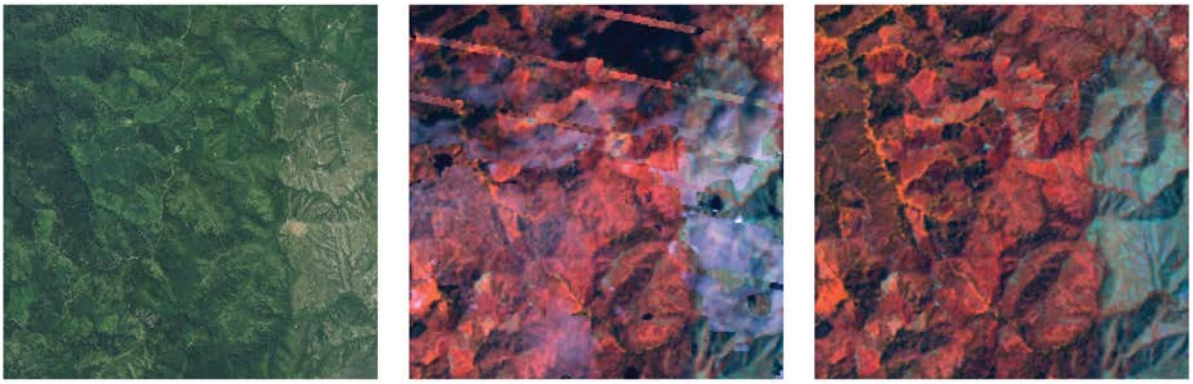


Figure C-3: An example landscape (aerial photo; left) with the associated false color Landsat composite based on images closest to the target date (center) and the mediod method of image selection (right).

C.4.2 Geospatial Predictor Assessment

In November and December 2016, Collaborator Gregory prepared all relevant geospatial data for modeling (Table C-1). These data can be broadly categorized into one of six subsets: Landsat, Change, Climate, Location, Topography, and Soils. Collaborators Gregory and Roberts ran several permutations of GNN with differing combinations of these geospatial data subsets in MR224 (Western Cascade Mountains; Figure C-2). At the time of this analysis, LTS data from LandTrendr acquired using the mediod method (see Section C.4.1) were only available for MR224. Still, MR224 is a good candidate for the current comparison as it includes diverse vegetation landscapes ranging across vegetation types (moist vs. dry forests and warm vs. cool forests) and stand ages (recently harvested to old-growth). Specifically, we compared the

following eight models: a saturated model with all variables, models with a single subset of variables removed (i.e., no soils, no change, no Landsat, no location, no topography, or no climate data), and a model with both soils and change data removed. The final model is similar to the model used to generate previous GNN maps of forest carbon used in Appendix A. Once models had been generated, we performed a modified leave-one-out validation protocol (Ohmann and Gregory 2002) to examine each models performance (Pearson correlation and root mean square error, in this case) for a suite of 37 forest attributes representing live and dead forest density, forest biomass, and species composition. We then calculated the difference between the performance metrics for the saturated models and all other models to assess the impact of excluding certain subsets of geospatial data on GNN predictions.

Our results indicated that Landsat imagery followed by topography were the most important subset of geospatial data while the removal of other subsets exhibited similar impacts on GNN predictive performance (Figure C-4). Removing Landsat variables increased RMSE and decreased correlations by an average of 26% and 39%, respectively. Removing topographic variables increased RMSE and decreased correlations by an average of 5% and 8%, respectively. GNN performance when removing other geospatial data subsets changed by less than 2%. Previous studies have shown that in some regions, location and climate data can greatly improve predictions, such as coastal environments where fog plays a key role in the distribution of some tree species or transitions from moist to dry landscapes where structure and composition are impacted by biogeographic constraints (Ohmann and Gregory 2002). The lack of an effect of removing soil data (one of the geospatial subsets not previously examined) from modeling may reflect the relatively coarse resolution of the source data (1 km; Miller and White 1998) compared to natural variation in soil properties that may manifest at the scale of meters. Additionally, in some areas soils information impacted predictions substantially, creating unrealistic transitions in vegetation characteristics. Surprisingly, Landsat change metrics were not particularly useful, despite previous evidence that disturbance records can greatly improve biomass predictions such that Landsat-based models perform on par with LIDAR based models (Pflugmacher et al. 2012).

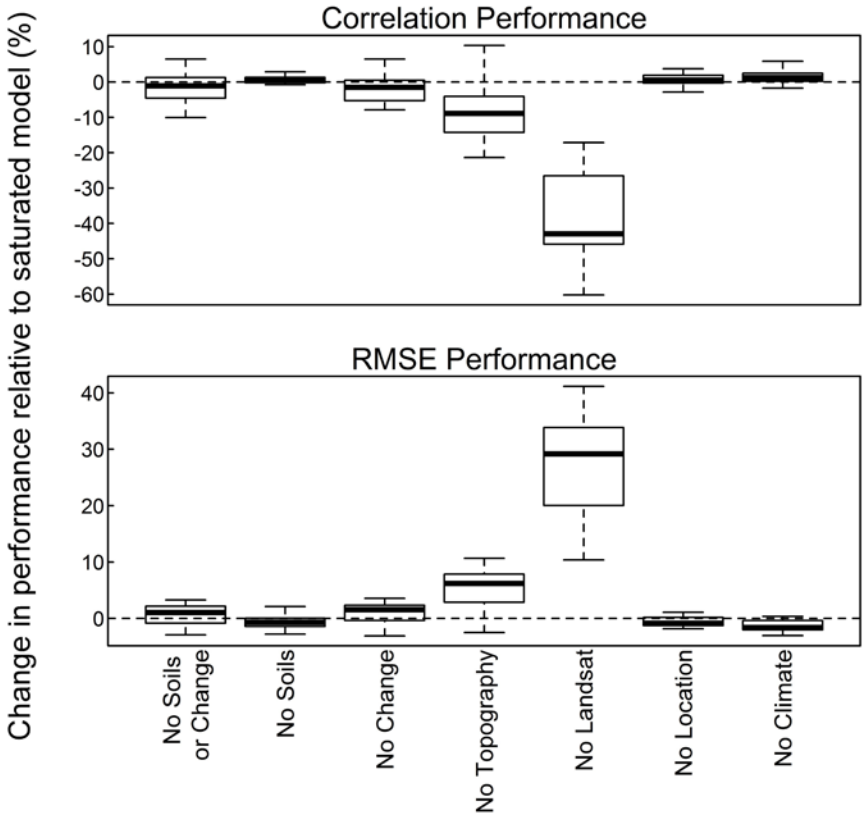


Figure C-4: Percent change in Pearson correlation coefficients and root mean squared error (RMSE) for GNN predictions when removing certain geospatial data subsets, summarized with box-and-whisker diagrams for a set of 37 key forest structural predictions.

To further examine the differences between GNN modeling with and without change metrics, we mapped biomass based on these initial GNN results for MRs 224 and 232 and highlighted areas of divergence. These results indicated that including 6-year change metrics introduced significant artifacts into maps associated with post-disturbance forest recovery. Specifically, seven years after disturbance, biomass increased substantially in a completely unrealistic fashion (Figure C-5). While previous work has indicated that multi-decadal records of forest disturbance and change information can greatly improve forest biomass stock predictions (Pflugmacher et al. 2012), a 6-year average of change in the spectral signals appears to be too short a period of time to effectively incorporate into GNN. Future research should focus on how long a record of disturbance is actually needed to inform modeling. For the purposes of this project, we concluded that the models incorporating climate, location, Landsat data, and topography were most appropriate and the remaining sections focus on this model alone.

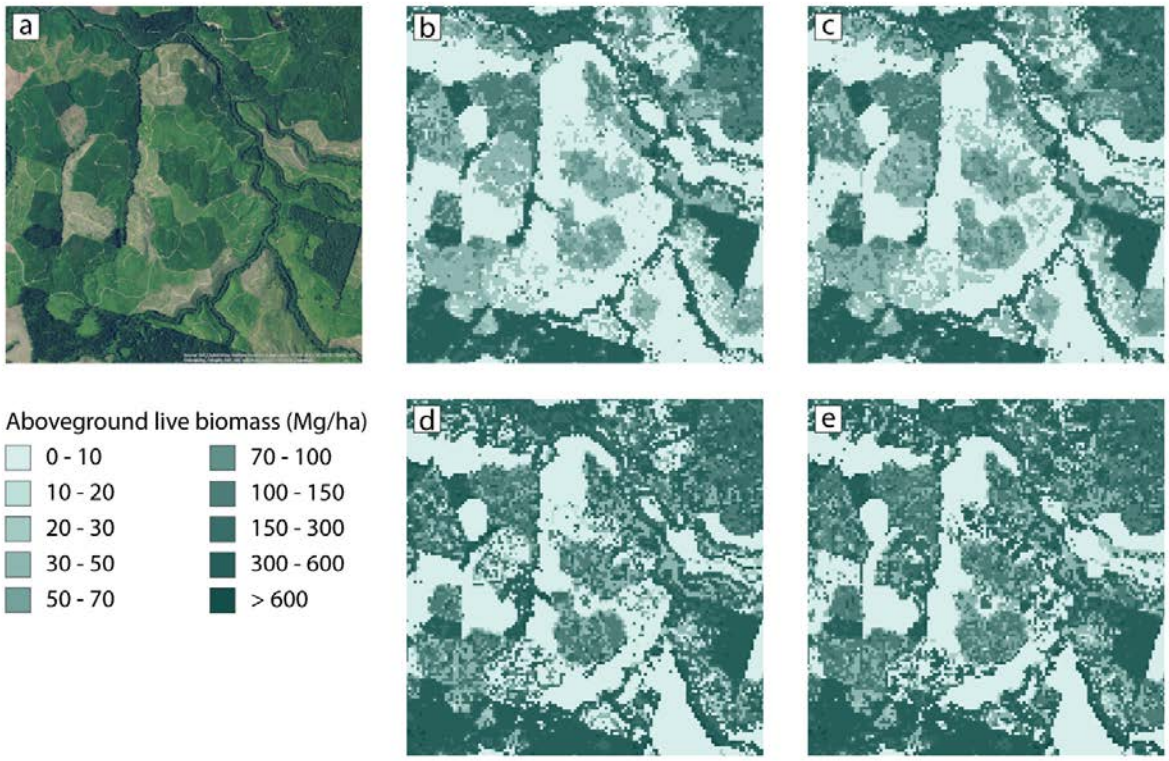


Figure C-5: Within (a) a managed forest landscape (2014 orthophoto), we present a comparison of live aboveground forest biomass maps based on GNN maps (b-c) without Landsat change metrics and (d-e) with Landsat change metrics for (b, d) 2013 and (c, e) 2014. Note that aboveground live biomass predictions within harvests tended to be greater, more spatially heterogeneous, and exhibit more rapid change with GN models incorporated change metrics compared to models without change metrics, all of which are inconsistent with our understanding of early seral forests.

C.5 Forest Biomass and Uncertainty Mapping

Following alterations to forest attribute data and geospatial data, we used the LandTrendr-GNN framework to produce annual aboveground live biomass maps for all modeling regions (Figure C-2) using three implementations of GNN: $k = 1$ (GNN_{k1}), bootstrapping approximation (GNN_{ba}), $k = 10$ (GNN_{k10}). GNN_{k1} is the traditional GNN implementation, where the nearest neighbor to the pixel in gradient space is imputed to that pixel (e.g., Ohmann and Gregory 2002). This method results in maps where unrealistic combinations of attributes cannot be produced (i.e., each prediction was observed at least once in the field). However, model precision cannot be directly assessed. For larger values of k estimates of uncertainties can be assessed. For GNN_{ba}, an approximation of non-parametric bootstrapping can be used to assess precision at the scale of pixels for a $k = 1$ kNN process (Bell et al. 2015). For GNN_{k10}, variation in neighbors within and among pixels forms the basis of model-based inference (McRoberts et al. 2007), with predictions being the unweighted means of the k nearest neighbors. Because GNN_{ba} and GNN_{k10} rely on weighted or unweighted means of the k nearest neighbors, they

underestimate the prevalence of rare features (e.g., uncommon forest structural conditions) or produce unrealistic combinations of structures or species. GNN_{ba} weights nearest neighbors based on their neighborhood rank (i.e., probability of inclusion of neighbor *j* and exclusion of all nearer neighbors in a given bootstrap sample), such that GNN_{ba} represents a compromise between GNN_{k1} and GNN_{k10}. Note that GNN_{k10} can be used as the basis of small area estimation using model-based inference (e.g., McRoberts et al. 2007), either GNN_{ba} or GNN_{k10} can be used in a model-assisted framework (e.g., McRoberts et al. 2016), and all three products might be the basis of an error-adjusted area estimation procedure for areal estimates of forest biomass categories (Olofsson et al. 2013).

For the current task, LEMMA (Bell, Gregory, and Roberts) constructed predictions based on GNN_{k1}, GNN_{ba}, and GNN_{k10} so that methods appropriate for uncertainty assessment (GNN_{ba} and GNN_{k10}) could be compared with the legacy methods for which direct estimation of precision in estimates was not possible (GNN_{k1}). Maps of mean aboveground live biomass (GNN_{k1}, GNN_{ba}, and GNN_{k10}) and associated standard deviations (GNN_{ba} and GNN_{k10}) were generally similar among methods, though standard deviations for GNN_{k10} were greater than standard deviations for GNN_{ba}. GNN maps indicated that live tree biomass was greatest in the higher elevations of the Sierra Nevada Mountains and in the North Coast of California (Figure C-6). Standard deviations in aboveground live biomass predictions tended to increase with mean predictions, consistent with GNN precision in western Oregon (Bell et al. 2015). Coefficients of determination (i.e., standard deviations divided by means) tend to be highest in recently disturbed forests (data not shown).

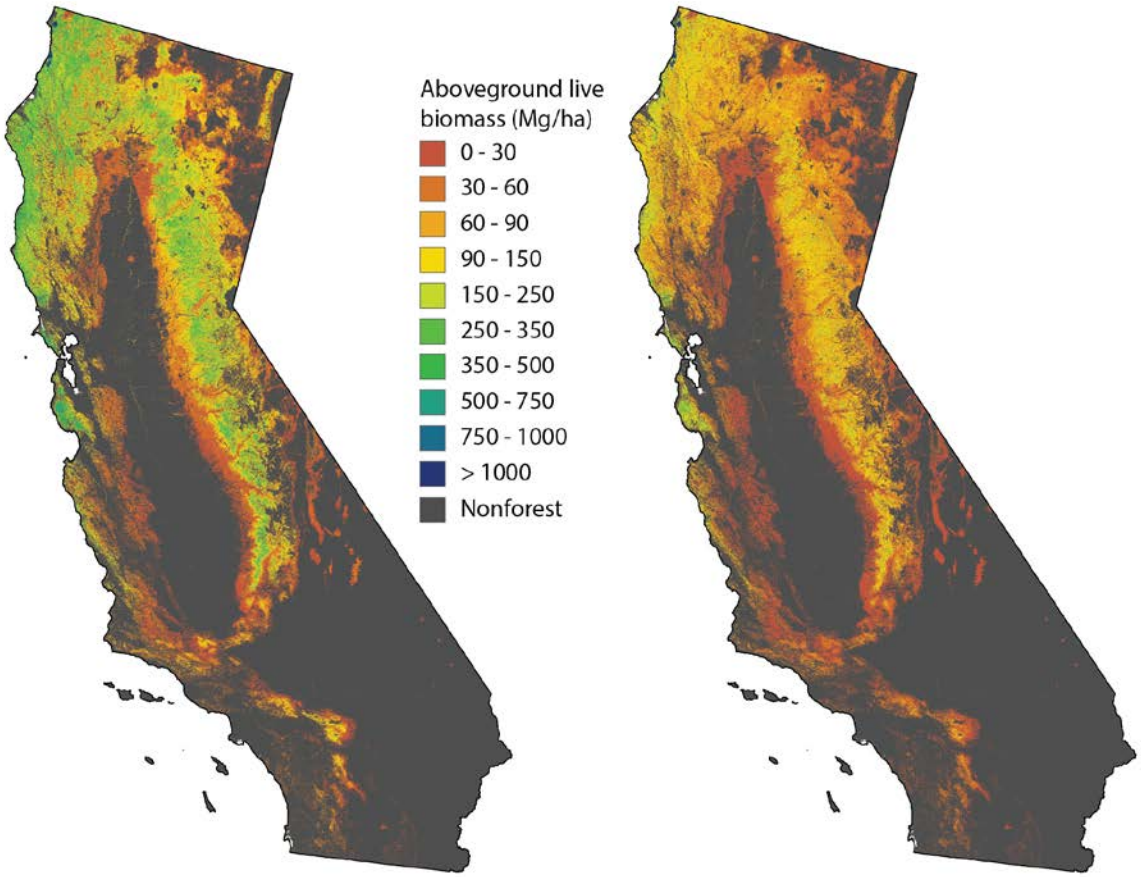


Figure C-6: Maps of (left) mean and (right) standard deviation for aboveground live biomass based on GNN_{k10}.

C.6 Accuracy Assessments

In February 2017, we assessed accuracy of each model implementation (GNN_{k1}, GNN_{ba}, and GNN_{k10}) based on normalized root mean square error (NRMSE) and coefficient of determination (R_2) using a modified leave-one-out procedure. The modified leave-one-out cross-validation compares observations to the nearest neighbor that is not itself, thus avoiding self-imputation that might exaggerate model performance statistics. This method has been shown to perform similarly to traditional leave-one-out cross-validation while greatly reducing the computational intensity of accuracy assessment (Ohmann and Gregory 2002). Accuracy statistics (NRMSE and R^2) were calculated for aboveground live biomass. We found that in all modeling regions but one, the GNN_{k10} implementation exhibited the minimum NRMSE and maximum R^2 (Table C-2). The single exception was MR 231 (California Central Valley; Figure C-2) where GNN_{ba} exhibited the most accurate estimates, though differences from GNN_{k10} in terms of accuracy statistics were miniscule.

Table C-2: Accuracy assessment for each of the three GNN implementations for each MR (see Figure C-2) for aboveground live biomass for 2014.

		NRMSE			R^2		
Modeling Region	**# Plots**	**GNN$_{k1}$**	**GNN$_{ba}$**	**GNN$_{k10}$**	**GNN$_{k1}$**	**GNN$_{ba}$**	**GNN$_{k10}$**
223	1454	0.517	0.492	0.479*	0.534	0.577	0.600*
224	4026	0.565	0.536	0.503*	0.550	0.594	0.642*
225	2704	0.721	0.701	0.682*	0.367	0.401	0.434*
226	850	0.945	0.871	0.796*	0.254	0.366	0.471*
227	3427	0.660	0.626	0.597*	0.474	0.527	0.569*
231	774	0.794	0.769*	0.770	0.530	0.559*	0.558
232	3935	0.719	0.690	0.653*	0.518	0.557	0.603*
233	797	1.048	1.009	0.951*	0.535	0.569	0.617*
234	241	0.915	0.876	0.808*	0.384	0.436	0.520*
MEAN		**0.765**	**0.730**	**0.693***	**0.461**	**0.509**	**0.557***

* indicates the model implementation with the (1) minimum NRMSE or (2) maximum R^2

We also examined scatterplots of predicted vs. observed values to understand how predictive capacity varied as a function of aboveground live forest biomass (Figure C-7). Generally, residual magnitudes increased with aboveground live forest biomass. At low aboveground biomass levels (<500 Mg ha^{-1}; 97% of plots), all three model implementations performed similarly. At high levels (>500 Mg ha^{-1}; 3% of plots), all models, but especially GNN$_{k10}$, exhibited under-predictions consistent with the widely recognized biomass saturation in biomass modeling based on Landsat data (Steininger 2000). For example, aboveground live biomass observations in MR232 (Sierra Nevada Mountains; Figure C-2) ranged up to 1971 Mg ha^{-1}, but maximum predictions at plot locations were 994, 909, and 716 Mg ha^{-1} for GNN$_{k1}$, GNN$_{ba}$, and GNN$_{k10}$, respectively. However, plots with aboveground live biomass exceeding 1000 Mg ha^{-1} were rare, ranging from 0.0% to 1.4% of plots across modeling regions with only one modeling region having more than 0.5% of plots exceeding 1000 Mg ha^{-1}.

The accuracy assessment results (Table C-2 and Figure C-7) indicated that overall accuracy of GNNk_{10} was superior to the other products and that biases for high live aboveground biomass sites represented a small portion of the landscape. Given the more flexible solutions for small area estimation in a model-based approach using GNN$_{k10}$ (see Section C-5), the high performance of GNN$_{k10}$ was a favorable result given the focus of later parts of the project focusing on estimation. Still, given the generally favorable behavior of GNN$_{ba}$, both GNN$_{ba}$ and GNN$_{k10}$ will be shared with collaborators for the remainder of the project. Therefore, we have made annual aboveground live biomass mean and standard deviation maps available to all collaborators via an FTP service.

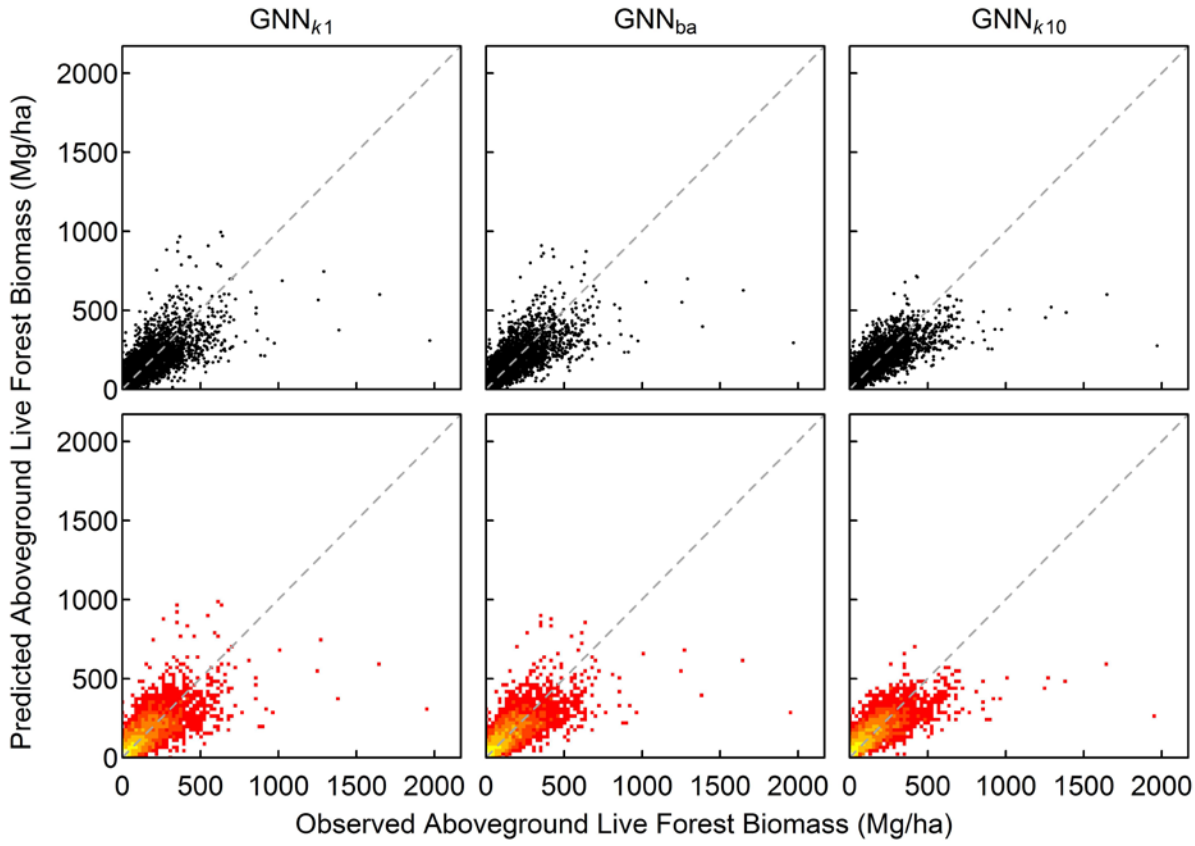


Figure C-7: Scatterplots (upper panels) and heat maps (lower panels) showing observed vs. predicted aboveground live forest biomass for GNN_{k1}, GNN_{ba}, and GNN_{k10}. For heat maps, white areas indicate no data, with red to yellow areas indicating low to high density of plots.

C.7 References

Bechtold, W. A., and P. L. Patterson (editors). 2005. The enhanced Forest Inventory and Analysis program—National sampling design and estimation procedures. Gen. Tech. Rep. SRS-80.

Bell, D. M., M. J. Gregory, and J. L. Ohmann. 2015. Imputed forest structure uncertainty varies across elevational and longitudinal gradients in the western Cascade Mountains, Oregon, USA. Forest Ecology and Management 358:154-164.

FIA. 2010. Regional biomass equations used by FIA to estimate bole, bark, and branches (updated 13-Jan-2010), USDA Forest Service, Pacific Northwest Research Station.

Heath, L.S.; Hansen, M.H.; Smith, J.E.; Smith, W.B.; Miles, P.D. 2009. Investigation into calculating tree biomass and carbon in the FIADB using a biomass expansion factor approach. In: McWilliams, W.; Moisen, G.; Czaplewski, R., comps. 2008 Forest Inventory and Analysis (FIA) Symposium; 2008 October 21-23; Park City, UT. Proc. RMRS-P-56CD. Fort Collins, CO: U.S. Department of Agriculture, Forest Service, Rocky Mountain Research Station. 1 CD.

Jenkins, J.C.; Chojnacky, D.C.; Heath, L.S.; Birdsey, R.A. 2003. National scale biomass estimators for United States tree species. Forest Science. 49: 12-35.

Kennedy, R. E., Z. Yang, and W. B. Cohen. 2010. Detecting trends in forest disturbance and recovery using yearly Landsat time series: 1. LandTrendr - Temporal segmentation algorithms. Remote Sensing of Environment, 114: 2897-2910

Max, T. A., H. T. Schreuder, J. W. Hazard, D. O. Oswald, J. Teply, and J. Alegria. 1996. The Pacific Northwest region vegetation inventory and monitoring system. Research Paper PNW-RP-493.

McRoberts, R. E., G. M. Domke, Q. Chen, E. Næsset, and T. Gobakken. 2016. Using genetic algorithms to optimize *k*-Nearest Neighbors configurations for use with airborne laser scanning data. Remote Sensing of Environment 184: 387-395.

McRoberts, R. E., E. O. Tomppo, A. O. Finley, and J. Heikkinen. 2007. Estimating areal means and variances of forest attributes using the *k*-Nearest Neighbors technique and satellite imagery. Remote Sensing of Environment 111: 466-480.

Miller, D. A., and R. A. White. 1998. A Conterminous United States Multi-Layer Soil Characteristics Data Set for Regional Climate and Hydrology Modeling. Earth Interactions 2-002: 1-26

Ohmann, J. L., and M. J. Gregory. 2002. Predictive mapping of forest composition and structure with direct gradient analysis and nearest neighbor imputation in coastal Oregon, USA. Canadian Journal of Forest Research 32: 725–741.

Ohmann, J. L., M. J. Gregory, H. M. Roberts, W. B. Cohen, R. E. Kennedy, and Z. Yang. 2012. Mapping change of older forest with nearest-neighbor imputation and Landsat time-series. Forest Ecology and Management 272: 13–25.

Ohmann, J. L., M. J. Gregory, H. M. Roberts. 2014. Scale considerations for integrating forest inventory plot data and satellite image data for regional forest mapping. Remote Sensing of Environment *151*:3-15.

Olofsson, P., G. M. Foody, S. V. Stehman, and C. E. Woodcock. 2013. Making better use of accuracy data in land change studies: Estimating accuracy and area and quantifying uncertainty using stratified estimation. Remote Sensing of Environment 139: 121-131.

Pflugmacher, D., W. B. Cohen, R. E. Kennedy, and Z. Yang. 2014. Using Landsat-derived disturbance and recovery history and LiDAR to map forest biomass dynamics. Remote Sensing of Environment 151:124-137

Steininger, M. K. 2000. Satellite estimation of tropical secondary forest above-ground biomass: Data from Brazil and Bolivia. International Journal of Remote Sensing 21: 1139−1157.

ter Braak, C. J. F. 1986. Canonical correspondence analysis: A new eigenvector technique for multivariate direct gradient analysis. Ecology 67: 1167–1179

Zhu, Z., S. Wang, and C. E. Woodcock. 2015. Improvement and expansion of the Fmask algorithm: cloud, cloud shadow, and snow detection for Landsats 4–7, 8, and Sentinel 2 images. Remote Sensing of Environment 159: 269-277.

APPENDIX D: Detailed Methods: LiDAR-based Maps of Aboveground Live Tree Biomass

D.1 Data

D.1.1 Site Descriptions

We compiled results from six study areas (Figure D-1) that represent a range of forest conditions found (Table D-1) in two common and biomass-dense forests in California: the Sierran conifer forest and the coast redwood forest (Gonzalez et al. 2010, Gonzalez et al. 2015). For both forest types, sites included examples of active forest management under state (Garcia and Blodgett) and federal (Last Chance and Sugar Pine) jurisdiction. In addition, two sites are managed as preserves (Mailliard and North Yuba). For detailed descriptions of the study area see Gonzalez et al. (2010) for Garcia, Mailliard, and North Yuba; Hopkinson and Battles (2015) for Last Chance and Sugar Pine; and Olson and Helms (1996) for Blodgett.

D.1.2 Data Descriptions

For these six sites, we have concurrent field inventory and LiDAR data (Table D-2, Table D-3). Circular field plots were stratified by management class at Garcia, slope-position at Mailliard, and forest type at North Yuba (Gonzalez et al. 2010). Grid sampling with a random start was used at the other sites (Hopkinson and Battles 2015, Olson and Helms 1996). Basic field inventory data included species, diameter at breast height (1.37m, DBH) and height of live trees ≥ 5 cm DBH. Height was measured on a subsample of trees at Garcia, Mailliard, and North Yuba. Thus, we developed species and site specific equations using likelihood-based methods (Buckland et al. 1997) to quantify the strength of evidence for alternative models of the relationship between DBH and height. We considered four candidate models in three general functional forms: linear, power, and saturating. We selected the model with the lowest Akaike's information criterion (AIC). For the dominant conifer species, fits (R^2) for the best DBH-height model were greater than 0.81.

Plot centers at all but one site were located with ± 1m. The exception was Blodgett where plot centers have ± 3 m location accuracy. LiDAR was acquired by two well-established contractors: Sanborn Map Company (http://www.sanborn.com/aerial-lidar/) and the National Center for Airborne Laser Mapping (NCALM, http://ncalm.cive.uh.edu/).

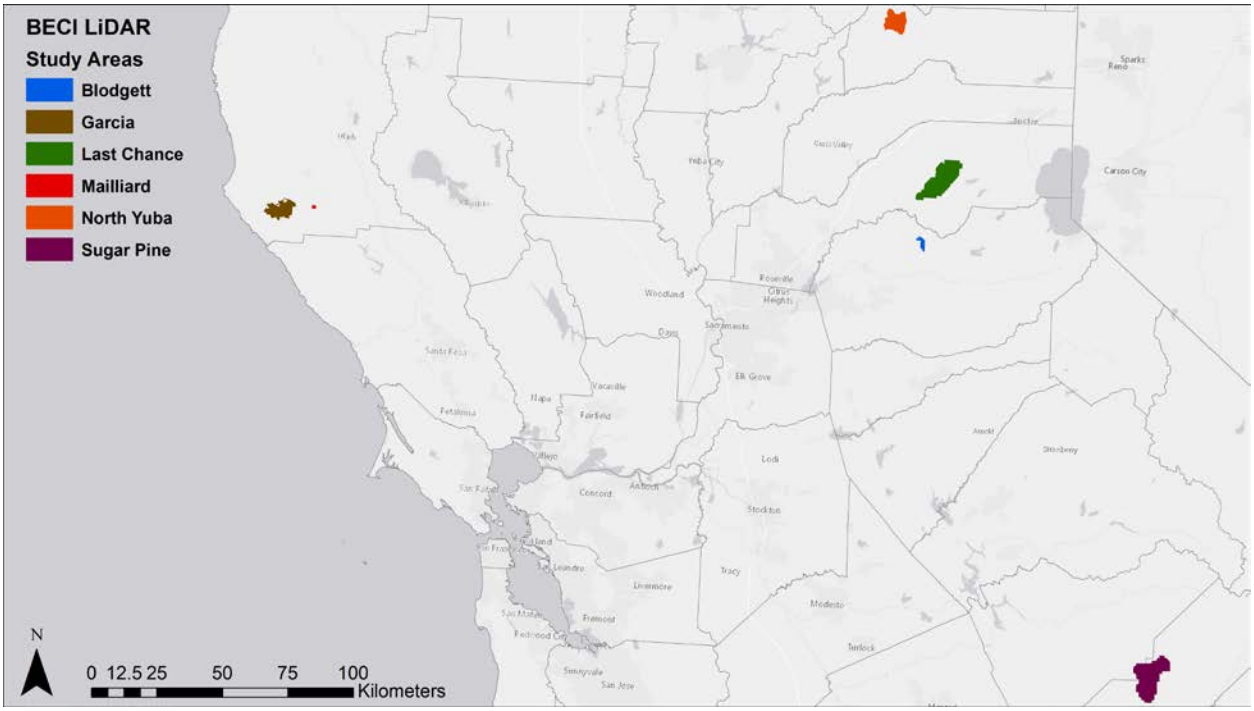


Figure D-1: Map showing location of the study sites in California where we developed LiDAR-derived maps of aboveground live biomass.

Table D-1: Description of the local-scale study areas used in this study. Area refers to the LiDAR footprint in the study area.

Site	Plot number	Plot size (ha)	Year(s) sampled	AGB median (max-min) (Mg ha^{-1})	COV
Garcia	28	0.1	2005	290 (68-472)	48
Mailliard	12	0.1	2005	606 (336-1392)	47
North Yuba	36	0.1	2005	317 (0-1118)	68
Last Chance	103	0.05	2013	136 (0-1625)	119
Blodgett	525	0.04	2013-15	298 (3- 603)	69
Sugar Pine	121	0.05	2007-08	348 (34-1667)	71

Table D-2: Description of the local-scale study areas used in this study. Area refers to the LiDAR footprint in the study area.

Site/owner	Forest type and management regime	Area (km^2)
Garcia River Preserve/The Conservation Fund	Coast redwood forest managed for sustainable timber production.	58.7
Mailliard Redwood State Natural Reserve/California	Old growth coast redwood forest that is maintained as a preserve with no active forest management.	2.1
Last Chance, Tahoe National Forest/US Forest Service	Sierran mixed conifer with management priorities of modifying fire behavior and improving forest health.	135.0
Blodgett Forest Research Station/University of California	Sierran mixed conifer forest that encompasses a range of management regimes from even-aged harvests to old-growth reserves.	9.2
Sugar Pine, Sierra National Forest/US Forest Service	Sierran mixed conifer with management priorities of modifying fire behavior and improving forest health.	142.5

Table D-3: Attributes of LiDAR acquisitions used in this study as reported by the contractors.

Attributes	Garcia Mailliard North Yuba	Last Chance	Sugar Pine	Blodgett
Survey date	Sep 2005	Nov 2012, Aug 2013	Sep 2007	Sep 2014
Contractor	Sanborn	NCALM*	NCALM	NCALM
Instrument	Optech ALTM 2050	Optech GEMINI	Optech GEMINI	Optech GEMINI
Mean pulse density m^2	2-4	10	10	11
Scan angle (°)	±15	±12 to 14	±12 to 14	±18
Survey altitude (m)	800	600-800	600-800	600-800
Horz. accuracy (RMSE, cm)	<50	10	10	<30
Vert. accuracy (RMSE, cm)	<15	5-35	5-35	9-100

*National Center for Airborne Laser Mapping

D.2 Biomass Modeling

D.2.1 Field Inventory

Aboveground live tree biomass (AGB) was calculated from tree inventory data that included species, DBH, and tree height using regional biomass equations (FIA 2010). These equations predict biomass of the entire tree stem from estimates of cubic volume and species-specific wood density. Separate allometric equations were used to calculate the biomass in bark and branches. AGB is the sum of the stem, bark, and branch mass. For western states, the Pacific Northwest Forest Inventory and Analysis program considers estimates of tree biomass using the regional equations to be more accurate than the national analysis (PNW-FIADB 2015). Only live trees ≥ 5 cm DBH were included in the analysis. Plot-level AGB is the sum of the individual tree biomass scaled by plot size and is expressed in megagrams of biomass per hectare (Mg ha^{-1}).

To ensure consistency in the application of the biomass equations (FIA 2010) between the GNN estimates and the LiDAR estimates, we compared results for a test set of trees that included small (DBH = 10.2 cm) and large (DBH = 50.8) examples of all the species in our six study areas. The largest discrepancy was 0.087% with most less than 0.002%. These differences were tracked

to an error in a branch equation and minor differences in the processing of rare cases. These inconsistencies were corrected in the final results.

To ensure quality across multiple datasets, both data (i.e., species, DBH, and height) and results (i.e., tree-level and plot-level AGB) were screened for unusual values. We used DBH-height relationships to identify outliers where outliers were defined as observed heights > ± 100% of predicted values. Original records of the suspect data were inspected, and if necessary, interpolated height values were used in the rare cases (<1 in 1,000) where there was no explanation for the outlier. For example, if there was a note in the records that the top was broken, the shorter than expected tree would be retained, and if not, the predicted value was used. We also evaluated AGB estimates. From the 2016 FIA California database (FIADB v 6.1), we calculated the distribution of tree-level AGB by species and plot-level AGB by forest type (California Mixed Conifer and Redwood). Results less than the 1st percentile or greater than 99th percentile were checked.

D.2.2 LiDAR Processing

We used the USDA Forest Service's FUSION software (version 3.7, McGaughey 2018) to calculate forest height and cover metrics from the LiDAR data (Table D-4). We took the location of the plot center to create a polygon that replicated plot size and then extracted plot-level metrics for the plot-polygons with the FUSION cloudmetrics tool. All structural analyses were normalized to local bare-earth estimates of the height above ground.

Table D-4: Metrics computed by the FUSION LiDAR processing software and tested for inclusion in models in this study. Details on the calculation of these metrics are available in McGaughey (2018).

Height metrics (Returns <2 m height were not used in the calculation of the height metrics.)		**Cover and intensity metrics** (Cover metrics defined as the ratio of returns above a height break.)
Maximum, Minimum, Mode		Total all returns
Variance		Total first returns
Standard deviation		All returns > 2 m
Coefficient of Variation		First returns > 2 m
Interquartile distance		All returns > mean
Kurtosis		First returns > mean
Skewness		All returns > mode
L-moments (L1, L2, L3, L4)		First returns > mode
L-moment skewness		Percentage all returns > 2m
L-moment kurtosis		Percentage first returns > 2 m
AAD (Average Absolute Deviation)		Percentage all returns > mean
MADMedian (Median of the absolute deviations from the overall median)		Percentage first returns > mean
MADMode (Median of the absolute deviations from the overall mode)		Percentage all returns > mode
Cubic mean, quadratic mean		Percentage first returns > mode
Percentile values (1st, 5th, 10th , 20th, 25th, 30th, 40th, 50th, 60th, 70th, 75th, 80th, 90th, 95th, 99th percentiles)		Percentage all returns > 2/ total first returns
Canopy relief ratio		Percentage all returns above mean/ total first returns
		Percentage all returns above mode/ total first returns

We relied on machine learning techniques to select the best LiDAR predictors of AGB. Specifically, we used a variable selection method based on random forests -- VSURF (Genuer et al. 2015). A critical feature of VSURF is the permutation-based approach to variable importance that is designed to reduce redundancy in the selection of predictor variables. For each site, we used VSURF to identify the best predictors for AGB and three transformed variants: square-root, cubic-root, and natural logarithm of AGB. We initially evaluated the effect of the transformations by computing simple random forest analysis (Breiman 2001) for the transformed AGB and the full set of LiDAR variables and then comparing the overall fit of the models. We used the fit, measured by the coefficient of variation (R^2), to prioritize the subsequent linear modeling of predictor variables.

With the VSURF results, we specified a global linear model of AGB that included main effects and all two-way interactions of the predictor variables. We adopted a streamlined information-theoretic approach to compute and then rank all possible subsets of the global model (Barton 2018). Specifically, we used the function "dredge" in the MuMin R package (Barton 2018). Model rank was based on Akaike's information criterion (AIC) with the best models defined as the model with the lowest AIC and all other models where the differences in AIC values relative to the model with the lowest AIC (ΔAIC) was < 2. Burnham and Anderson (2002) regard all candidate models with ΔAIC < 2 to have substantial empirical support.

The top ranked model was evaluated for lack of fit, non-linearity in the relationship between predictor(s) and response, and for heteroscedasticity in the regression residuals (sensu Kennedy et al. 2018). Specifically, fit was assessed with the Rainbow test (Utts 1982); non-linearity with the Ramsey Regression Specification Error Test (Ramsey 1969); and heteroscedasticity with the Breusch-Pagan test (Breusch and Pagan 1979). A candidate was eliminated from consideration if it failed any of the three tests with failure defined as $P < 0.1$. In the infrequent cases where the top model (i.e., ΔAIC = 0) failed a test, we evaluated the remaining models in order based on ΔAIC. We never considered models with $\Delta AIC \geq 2$.

The performance of the best conforming model was quantified with two criteria: adjusted R^2 and root mean squared error (RMSE) based on a 10-fold cross validation analysis (Zhao et al. 2012). For the three sites with < 100 plots (Garcia, Mailliard, and North Yuba), RMSE was calculated with leave-one-out cross validation. We also report relative RMSE (rRMSE):

$$\text{rRMSE} = \frac{\text{RMSE}}{\sum_{i=1}^{n} \frac{AGB_i}{n}} \qquad \text{Equation D-1}$$

where i represents plots, n is the total number of plots, and AGB_i is the LiDAR based estimate for plot i.

D.2.3 LiDAR Biomass Map

For each site, we applied the selected model to produce a map of AGB. Specifically, we used the large area processor in the FUSION gridmetrics tool to create a wall-to-wall estimate of AGB. Given the variation in the LiDAR datasets (e.g., differences in sensor and point density, Table

D-3) and complex mountain terrain, we considered a range of grid cell areas to project the gridmetrics. A 25 m^2 grid cell provided the most consistent responses across the sites. We screened the AGB projection for outliers using upper bounds of 2,000 Mg ha^{-1} for study areas in the Sierran conifer forests (i.e., North Yuba, Blodgett, Last Chance, and Blodgett) and 4,000 Mg ha^{-1} in the coast redwoods (i.e., Garcia and Mailliard). Negative values were set to 0 Mg ha^{-1}. Trimmed values were rare (< 1 in 1,000).

D.3 References

Barton, K. 2018. MuMin package for R. (https://cran.r-project.org/web/packages/MuMIn/index.html).

Breiman, T and A. Cutler. 2015. Breiman and Cutler's Random Forests for Classification and Regression (https://cran.r-project.org/web/packages/randomForest/randomForest.pdf).

Breusch, T.S., and A.R. Pagan., 1979. A simple test for heteroscedasticity and random coefficient variation. Econometrica 47: 1287–1294.

Buckland, S. T., K. P. Burnham, and N. H. Augustin. 1997. Model selection: An integral part of inference. Biometrics **53**:603-618.

Burnham, K. P., and D. R. Anderson. 2002. Model Selection and Multimodel Inference: A Practical Information -Theoretic Approach. Springer, New York.

FIA. 2010. Regional biomass equations used by the Forest Inventory and Analysis program to estimate bole, bark, and branches (updated 13-Jan-2010), U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station, Portland, Oregon.

Genuer, R., J Poggi, and C. Tuleau-Malot. 2015. VSURF: An R Package for Variable Selection Using Random Forests. The R Journal 7:19-33.

Gonzalez, P., G. P. Asner, J. J. Battles, M. A. Lefsky, K. M. Waring, and M. Palace. 2010. Forest carbon densities and uncertainties from LiDAR, QuickBird, and field measurements in California. Remote Sensing of Environment 114:1561-1575.

Gonzalez, P., J.J**.** Battles, B.M. Collins, T. Robards, and D.S. Saah. 2015. Aboveground live carbon stock changes of California wildland ecosystems, 2001-2010. Forest Ecology and Management 348: 68-77.

Hopkinson, P. and J.J. Battles **(**editors). 2015. Learning how to apply adaptive management in Sierra Nevada forests: An integrated assessment. Final Report of the Sierra Nevada Adaptive Management Project. Berkeley, CA: Center for Forestry, UC Berkeley (http://www.snamp.edu/).

Kennedy, R.E., J. Ohmann, M. Gregory, H. Roberts. D.M. Bell, V. Kane, M. J. Hughes, W. Cohen, S. Powell, N. Neeti, S. Hooper, J. Kane, D.L. Miller, J. Perkins, J. Braaten, and R. Seidl. 2018. An empirical, integrated forest biomass monitoring system. Environmental Research Letters 13:2.

McGaughey, R.J., 2018. FUSION/LDV: Software for LiDAR Data Analysis and Visualization Version 3.7. United States Department of Agriculture Forest Service Pacific Northwest Research Station (http://forsys.sefs.uw.edu/fusion/FUSION_manual.pdf).

Olson, C.M. and Helms, J.A. 1996. Forest growth and stand structure at Blodgett Forest Research Station 1933–1995. In Sierra Nevada Ecosystem Project: Final Report to Congress 3: 681-732.

PNW-FIADB. 2015. A Data Dictionary and User Guide for the PNW-FIADB database. U.S. Department of Agriculture, Forest Inventory and Analysis, Pacific Northwest Research Station, Portland, Oregon.

Ramsey, J.B. 1969. Tests for specification error in classical linear least squares regression analysis. Journal of the Royal Statistical Society Series B 31, 350–371.

Utts, J.M., 1982. The rainbow test for lack of fit in regression. Communications in Statistical Theory and Methods 11: 1801–1815.

Zhao, F., Q. Guo, and M. Kelly. 2012. Allometric equation choice impacts LiDAR-based forest biomass estimates: A case study from the Sierra National Forest, CA. Agricultural and Forest Meteorology 165:64-72.