

Article

# Development of a Regional Lidar-Derived Above-Ground Biomass Model with Bayesian Model Averaging for Use in Ponderosa Pine and Mixed Conifer Forests in Arizona and New Mexico, USA

Karis Tenneson <sup>1,\*</sup>, Matthew S. Patterson <sup>1,2</sup> , Thomas Mellin <sup>3</sup>, Mark Nigrelli <sup>4</sup>, Peter Joria <sup>3</sup> and Brent Mitchell <sup>1</sup>

<sup>1</sup> RedCastle Resources, Inc. Contractor to USDA Forest Service Geospatial Technology and Applications Center, Salt Lake City, UT 84119, USA; maspatte@uw.edu (M.S.P.); brentmitchell@fs.fed.us (B.M.)

<sup>2</sup> Department of Urban Design and Planning, College of the Built Environment, University of Washington, Seattle, WA 98195, USA

<sup>3</sup> USDA Forest Service, Southwestern Regional Office, Albuquerque, NM 87102, USA; tmellin@fs.fed.us (T.M.); pjoria@fs.fed.us (P.J.)

<sup>4</sup> USDA Forest Service, Coconino National Forest, Flagstaff, AZ 86001, USA; mnigrelli02@fs.fed.us

\* Correspondence: krtenneson@fs.fed.us; Tel.: +1-801-975-3768

Received: 29 January 2018; Accepted: 6 March 2018; Published: 12 March 2018

**Abstract:** Historical forest management practices in the southwestern US have left forests prone to high-severity, stand-replacement fires. Reducing the cost of forest-fire management and reintroducing fire to the landscape without negative impact depends on detailed knowledge of stand composition, in particular, above-ground biomass (AGB). Lidar-based modeling techniques provide opportunities to increase ability of managers to monitor AGB and other forest metrics at reduced cost. We developed a regional lidar-based statistical model to estimate AGB for Ponderosa pine and mixed conifer forest systems of the southwestern USA, using previously collected field data. Model selection was performed using Bayesian model averaging (BMA) to reduce researcher bias, fully explore the model space, and avoid overfitting. The selected model includes measures of canopy height, canopy density, and height distribution. The model selected with BMA explains 71% of the variability in field-estimates of AGB, and the RMSE of the two independent validation data sets are 23.25 and 32.82 Mg/ha. The regional model is structured in accordance with previously described local models, and performs equivalently to these smaller scale models. We have demonstrated the effectiveness of lidar for developing cost-effective, robust regional AGB models for monitoring and planning adaptively at the landscape scale.

**Keywords:** forest biomass; aboveground biomass; airborne lidar; monitoring; regional forest inventory; variable selection; Bayesian model averaging; multiple linear regression

---

## 1. Introduction

### 1.1. Background

Costs and damages from large, high-severity wildfires have been steadily escalating, particularly in the forests of the western United States [1,2]. In Ponderosa pine (*Pinus ponderosa*) and mixed conifer forests of the southwest USA, a legacy of fire suppression, historical logging practices, and grazing has increased fire risk [3,4]. These activities altered the natural fire cycle due to the increased stand density, accumulation of surface and ladder fuel loads, and regrowth of fire-intolerant trees [5–7]. As a result, forests in the southwestern USA that were once characterized by frequent, low-intensity

fires are now experiencing catastrophic stand-replacement crown fires [3]. Rising temperatures, extended fire seasons, earlier snowmelt, and ongoing drought will continue to increase future wildfire potential [2,8–11]. Interactions between altered fire regimes, land use, and climate change are projected to continue intensifying the occurrence, size and severity of wildfires [12,13].

Knowledge of forest structure is needed in order to effectively manage and restore forests facing these altered fire regimes and environmental conditions. In particular, above-ground biomass (AGB) is useful for making management decisions such as controlled burns and stand thinning when conducting landscape restoration. Repeating extensive ground-based forest inventories to acquire these data is time consuming, labor intensive, and expensive.

Congress appropriated a consistent funding source through the Collaborative Forest Landscape Restoration Program (CFLRP, part of the Omnibus Public Land Management Act of 2009) in recognition of these urgent problems many communities are facing [14]. CFLRP offers competitive awards for implementing large-scale, collaborative, cross jurisdictional restoration plans. CFLRP encourages long-term restoration treatments (10-year period) across national forests, but that also extend across other land ownerships (e.g., federal, state, tribal, and private land) in order to reduce fire risk to vulnerable communities [14,15]. The need to monitor forests with fragmented ownership adds to the expense of accessing ground plots. Estimation techniques that use Earth-observing data provide an alternative toolkit for monitoring landscapes over time [16]. The advantages of lidar (Light Detection and Ranging) based approaches include: (1) the ability to collect and process spatially explicit data representing the horizontal and vertical conditions of the landscape over large spatial extents, (2) coverage of difficult to reach terrain and properties, and (3) accurate estimation of forest structure parameters in a timely and economical fashion (reviews by [17–23]).

Landscape scale restoration efforts aim to create conditions where natural fires can be managed for resource objectives without fear of escalation, thereby reducing fire suppression costs [1,24–26]. Over 280,000 km<sup>2</sup> (70 million acres) of these forests are in need of restoration [1,27]. Awarded projects are required to monitor social, ecological, and economic outcomes for 15 years to adapt future decisions based on new understanding of treatment efficacy and identification of potential unintended negative consequences [28–32].

The need for ongoing monitoring is especially important given the emergence of novel conditions resulting from the interactions of climate change and land use [33–35]. Restoration treatments are often guided by historical reference conditions and the natural range of variability [36]; however, forests are expected to experience new conditions outside this range [2,11–13]. Assessing the efficacy of restoration treatments under new ecological, social, and economic conditions is essential to adapting strategies aimed at increasing resilience of desired forest systems [33].

### 1.2. Project Goals

CFLRP projects have invested in the acquisition of both lidar data and the associated field data used to develop models for estimating forest structure. Since the collection of field data for each new project is expensive, a framework that allows project managers to update forest inventories by applying previously developed models to newly acquired lidar data would result in substantial savings. Regression models have been demonstrated to effectively link forest inventory parameters, including AGB, measured in the field with coincident lidar canopy metrics—aggregated to the plot boundary [17–23]. The resulting models can be applied to the full lidar data to create a continuous GIS raster layer of the forest inventory parameters across the study area. However, few studies have assessed lidar-based inventories in Ponderosa pine and mixed conifer forests of the southwestern USA [37–39]. Hall [39] developed Ponderosa pine and Douglas fir forest inventory models for the Front Range of the Rocky Mountains, CO. Sherrill and colleagues [37] use lidar metrics to explain 76% of the variation in field-estimated AGB values in subalpine forests of the Central Rockies. Finally, Kim and colleagues [38] developed models to estimate live and dead AGB in Ponderosa pine and mixed conifer forests in the North Rim of the Grand Canyon National Park with a RMSE of 46.01 Mg/ha

(23.66% RMSE). Our large sample size, full range of AGB conditions, and expansive spatial footprint enable us to build on their research. However, the range of AGB values in these three studies did not include stands with high biomass; the majority of values were under 300 Mg per hectare. Sample sizes were small, ranging from 36 [37] to 58 [38]. Further, these studies were built on small data sets and represent a limited spatial extent. Lidar-based regional models have been developed to estimate biomass in boreal, temperate deciduous, temperate coniferous, and tropical forests [40–44]. Regional models for southwestern US forests have not yet been explored.

The goal of this work is to develop a regional above-ground biomass (AGB) estimation model using lidar metrics from field data [45] and apply it to new lidar acquisitions. We use AGB estimates from field inventories, discrete-return lidar data, and environmental data to develop a regional lidar model that estimates AGB in Ponderosa pine and mixed conifer forests in the southwest United States. Information from the lidar data is supplemented with data from other sources to explain differences in forest structure due to contrasting environmental conditions, site productivity, and species composition [37,46–50]. AGB was estimated at plots from ground-based data using regional allometric equations.

Lidar-derived height, canopy density, and canopy shape metrics are combined with environmental data that represent forest composition and are regressed against the field-based AGB estimates to derive predictive equations. We integrated data from seven distinct field data collection efforts; five were used for model development and the remaining two were used as an independent validation. We evaluate the reliability of the model with two data sets: a reserved subset of the model development data and two coincident field data collection efforts. The latter validation data set was used to assess the transferability of the model on a new lidar acquisition.

While certain methods (stepwise regression) are popular, no standard approach has emerged in the literature as a robust, bias-free method to select which lidar metrics are best suited for estimating biomass from the large (overwhelming) pool of candidates. To fully explore the model space, and to avoid researcher bias and overfitting, we use Bayesian model averaging (BMA) methods to specify our model structure from the ensemble of candidates. For a good introduction to BMA, see [51–53].

The lidar data are from acquisitions collected with similar flight and sensor specifications. We include AGB estimates from seven field data campaigns with similar, but inconsistent plot size protocols and sample designs. While we recognize this is not ideal for model development, it does allow us to examine the influence of cost-saving plot size protocols on our model errors. We investigate the impact of allowing plot size to be determined by average stand stem density, discuss the implications of these inconsistencies on lidar-based AGB estimations, and make recommendations that attempt to balance the need for immediate field inventory savings vs. long-term costs of monitoring these landscapes.

We use AGB as our case study to test this approach for two reasons. First, studies have demonstrated accurate biomass estimates can be predicted with lidar data [20,21,23]. We expect that if the method works well for AGB it should then be applicable for prediction of other forest structure metrics that are correlated with lidar derivatives. González-Ferreiro and colleagues [54–57] demonstrated the effectiveness of using lidar to estimate canopy fuel characteristics. The second relates to the common need of fire restoration projects to monitor effectiveness at reducing fuel densities and prioritizing areas with high fuel loads. Accurate estimates of biomass are also important for forest management, habitat conservation, and global carbon accounting [58]. Biomass provides information about growth, health, and productivity. It is a key parameter in estimating carbon stock, timber production, wildlife habitat, fire behavior, fire impact, and for ecological modeling. Crucially, Sherrill and colleagues [38] have demonstrated success in separating biomass estimates between live and dead vegetation; this is an invaluable metric for assessing fire risk within a landscape.

## 2. Materials and Methods

### 2.1. Study Area

This study takes place in Ponderosa pine and frequent-fire (dry) mixed conifer forests in Arizona and New Mexico, USA. Sites were selected based on where lidar acquisitions and coincident field data were available. We include three regions in our analysis: the Four Forest Restoration Initiative (4FRI), the Southwest Jemez Mountain Landscape Restoration Project—both part of the network of Collaborative Forest Landscape Restoration Programs— and the 2009 Kaibab Forest Health Focus (Figure 1).

The 4FRI includes 10,000 km<sup>2</sup> of forest in northern Arizona and spans across the Coconino, Tonto, and Apache-Sitgreaves National Forests. It is located on the Mogollon Plateau, a northwest-southeast plateau running along the southern edge of the Colorado Plateau into the White Mountains. It is characterized by a gently sloping topography cut by small, steep drainages and infrequent rock outcroppings. The Southwest Jemez Mountain Landscape Restoration Project includes 850 km<sup>2</sup> of forestlands in the upper and middle Jemez River watersheds of which 512 km<sup>2</sup> was covered by a lidar acquisition. It is located just to the southwest of the Valles Caldera National Preserve in mountains that were formed primarily by volcanic activity. The Jemez River originates within the preserve and flows southwest through the center of the study area, eventually joining the Rio Grande. The Kaibab Forest Health Focus manages land on the Kaibab Plateau. This includes the northern section of the Kaibab National Forest and the Grand Canyon National Park-North Rim, Arizona. It is part of the larger Colorado Plateau and is bounded to the south by the Grand Canyon with tributary canyons on the east and west.

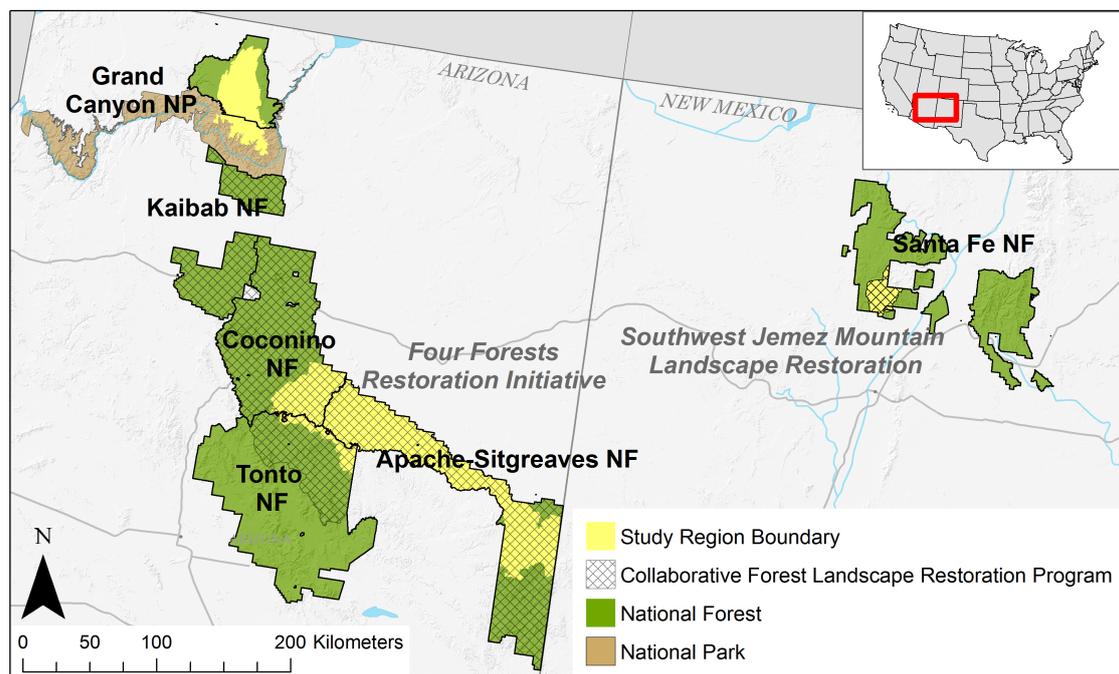


Figure 1. Location of study regions.

Elevation of the study sites ranges from 1700 to 3100 m above sea level. Annual precipitation is bimodal: most falls as snow between November and March, with a smaller amount of precipitation from monsoonal rains and thunderstorms during July and August [59]. Forest composition is strongly influenced by elevational gradients. The colder and wetter conditions of higher elevation areas support dense stands of spruce trees (Engelmann's spruce, *Picea engelmannii* and blue spruce, *Picea pungens*) and mixed fir stands (corkbark fir, *Abies lasiocarpa* var. *arizonica*; white fir, *Abies concolor*;

and subalpine fir, *Abies lasiocarpa*) [60]. A narrow Douglas-fir (*Pseudotsuga menziesii*) belt occurs below the spruce and mixed fir stands. These are followed by Ponderosa pine (*Pinus ponderosa*) stands, located at moderate elevations [60–62]. The forest matrix includes patches of aspen stands, meadows, and forest openings generated by disturbances such as fire, wind throw, and timber harvesting [60]. Twoneedle pinyon (*Pinus edulis*), Utah juniper (*Juniperus osteosperma*), Gambel oak (*Quercus gambelii*), New Mexico locust (*Robinia neomexicana*), and rabbitbrush (*Chrysothamnus viscidiflorus*) are commonly interspersed throughout [59,62,63]. Pinyon-juniper woodlands occur below the Ponderosa pine belt. These woodlands regions are dominated by tree species with multi-stem growth form and smaller maximum heights than the higher elevation species [64]. Due to their different growth form, we masked large areas dominated by these woodland growth forms from the study. However, there are still some small pinyon-juniper patches interspersed throughout the study region.

## 2.2. Data

We use information from seven different field data collection efforts and four lidar acquisitions covering the three study sites. Field data were collected between 2013 and 2015 (Table 1) and lidar was flown between 2012 to 2014 (Table 2). All lidar and field acquisitions were during leaf-on periods. The coincident field and lidar data collection efforts were completed within 1–2 years of each other. In all study regions, with the exception of the Coconino N.F. and the Tonto N.F., field plots were located using a stratified random design based on lidar derivatives. The lidar data sets were typically acquired in the summer and delivered in the winter. The sample design was completed in the spring and the field sampling commenced after snowmelt. Areas of obvious change due to fires were not included in the sampling. Due to contracting issues, the Southwest Jemez Mountain field collection efforts was delayed one year. Given the relatively slow growth rate of Southwest forest communities, conditions were unlikely to change much in the time period.

Lidar data can be aggregated in a plethora of ways to represent forest attributes. Researchers have identified three categories that perform well and have clear biological interpretation with direct analogs to ecologically significant variables [45,65,66]. Variables representing the canopy height distribution, the variability or shape of canopy height distribution, and canopy cover or density are analogous to variables used in aerial stand volume tables that are used in forest inventory [46]. We represent the canopy height distribution with quantile height metrics. These provide information on the density of the canopy biomass at different heights by measuring the different levels of laser penetration into the forest canopy [66,67]. We included quadratic heights to maintain consistency with the study by Sherrill and colleagues [18,37]. L-moments and central moments provide information on the heterogeneity of forest composition within the plot by representing the location, variability, and shape (skewness and kurtosis) of the canopy height distribution. L-moments provide a more robust set of statistics and are comparable to central moment variance, skewness, and kurtosis but less subject to bias in small samples and are more robust to outliers [68]. Similar to Kim and colleagues [38] we estimate plot-level volume by assessing the interaction of canopy density and height quartiles.

The relationship between stand characteristics and lidar metrics vary between tree species, especially those with different crown shapes [44,69,70]. Therefore, we tested a combination of information from optical remote sensing and lidar to represent potential differences in plots with deciduous trees. We explored the predictive power of three ancillary data sets that represent present and potential forest composition, including topographic information, an indicator of phenology, and ecological response units. Topography is an important determinant of forest compositional structure. Phenologic differences can differentiate deciduous from evergreen communities. We estimated the magnitude of seasonal variation of greenness (NDVI) from a Landsat time-series analysis. Ecological response units indicate a combination of potential and current community composition [64].

**Table 1.** Data collection summary for each project.

Study Region	Area (km <sup>2</sup> )	Year	No. & Size (ha) of Plots	Min DBH (cm)	Sample Design	Strata	Use
Kaibab Plateau, AZ	1382	2013–2014	112 (0.04)	20.3	stratified random	95th percentile height & percent canopy returns (>3 m)	model dev.
Coconino N.F., 4FRI, AZ	75 (sampled area); 1136 (total)	2013–2014	508 (0.04), 329 (0.03), 669 (0.02), 160 (0.01)	12.7	systematic	288 stands without current inventory	model dev.
Tonto N.F., 4FRI, AZ	48 (sampled area); 499 (total)	2013–2014	491 (0.04), 453 (0.02), 162 (0.01), 13 (0.008)	12.7	systematic	215 stands without current inventory	model dev.
Apache-Sitgreaves N.F., 4FRI, AZ, Stage 1	1028	2014	15 (0.08), 85 (0.04)	12.7	stratified random	95th percentile height & percent canopy returns (>3 m)	model dev.
Southwest Jemez Mountain, NM	353	2014	6 (0.08), 61 (0.04)	12.7	stratified random	99th percentile height & all returns above the mode divided by 1st returns	model dev.
Apache-Sitgreaves N.F., 4FRI, AZ, Stage 2	294	2015	66 (0.08), 84 (0.04)	12.7	stratified random	95th percentile height & percent canopy returns (>3 m)	model valid.
Apache-Sitgreaves N.F., 4FRI, AZ, Stage 3	1700	2015–2016	25 (0.08), 71 (0.04)	12.7	stratified random	95th percentile height & percent canopy returns (>3 m)	model valid.

**Table 2.** Summary of the lidar specifications for each site. The average and standard deviation pulse density is reported.

Study Region	Date	Area (km <sup>2</sup> )	Instrument	Ave. and SD Pulse Density (Pulses/m <sup>2</sup> )	Field of View (Degrees)	Altitude (m)
Kaibab Plateau, AZ	2012	1853	Leica ALS50 & ALS60	12.9; 5.12	20–28	900–2000
Four Forest Restoration Initiative, AZ, Stage 1	2013	3546	Leica ALS50 & ALS60	9.4; 2.8	28	900
Four Forest Restoration Initiative, AZ, Stage 2	2014	4365	Leica ALS70	15.4; 5.5	28	1200–1400
Southwest Jemez Mountain, NM	2012	526	Leica ALS60	13.3; 4.8	26	900

### 2.2.1. Field Survey

Field information was collected at over 3000 plots in seven different data collection efforts across the three study regions between 2013 and 2015 (Figure 2 and Table 1). We did not have access to a complete probabilistic sample that covers the full extent of Ponderosa pine and mixed conifer forests in southwest US, but the combination of these seven collection efforts provide information across a broad range of forest conditions. Five of the seven efforts were completed in the 4FRI. The first effort was implemented on the western half of the Apache-Sitgreaves N.F.; the second stage in the eastern portion. Data from the second stage of the 4FRI data collection effort were completely withheld from the model development process and used for validation of the regional model. In this section we outline similarities and differences in sample design, protocols determining plot size, and field data collection efforts.

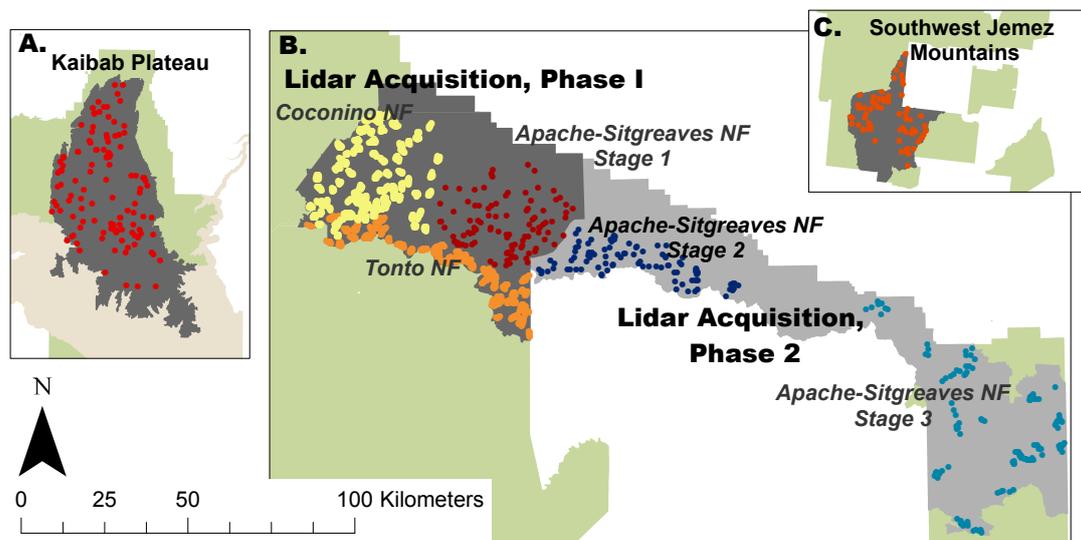
Sample design varied by project (Table 1). Generally, the strategy was to capture the full range of variation in forest structure recorded by the lidar sensor. Plots were located using a stratified random sampling scheme based on lidar-derived canopy structure information for all but two of the field inventory projects (Tonto and Coconino N.F., Table 1) [71,72]. Plots were located in areas where the (maximum) canopy cover reached at least 3 meters in height. Plots in the Tonto and Coconino N.F. were systematically located in stands that lack a current inventory. Minimum sample size per stand was 3 plots, but sample size and spacing between plots varied depending on stand area [73]. Plots for two projects were located based on accessibility. Plots within the Kaibab Plateau were within 250 m of level 2 forest service roads, the Santa Fe plots were within 300 m.

Field protocols that determined plot size varied between projects (Table 1). The plots within the seven field data collection efforts were distributed as follows: 0.4% 0.008 ha plots, 9.7% 0.01 ha plots, 33.9% 0.02 ha plots, 9.9% 0.03 ha plots, 42.7% 0.04 ha plots, and 3.4% 0.08 ha plots. Plot radius was determined by tree density in all data collection efforts, except on the Kaibab Plateau. The plot size in the Kaibab project was 0.04 ha. The default plot size in the other projects was also 0.04 ha,

but was increased or decreased depending on tree density. In the Tonto and Coconino N.F. plot size was decreased in dense stands; in the Southwest Jemez Mountain and three Apache-Sitgreaves N.F. the plot size was increased in low density stands. Overall, 43% of all the plots were 0.04 ha in size; 3.4% were larger; and 54% were smaller.

In stands with high tree density in the Coconino and Tonto N.F. the contractor was allowed to select a plot size such that at least 8 trees (DBH greater than 12.7 cm) were present per plot, on average throughout the stand. Sixty-four percent of the plots in these two data collection efforts met these dense stand conditions and were reduced in size to between 0.03 to 0.008 ha. Twenty percent of the plots in the Coconino N.F. were reduced to 0.03 ha; 40% of the Tonto and 40% of the Coconino N.F. plots were reduced to 0.02 ha; and 10% of the Coconino N.F. and 16% of the Tonto N.F. plots were reduced to 0.01 ha or smaller. One stand in the Tonto N.F. had a tree density that resulted in 13 plots with a size of 0.008 ha. In the Southwest Jemez Mountain and three Apache-Sitgreaves N.F. data collection efforts, plot size was doubled from 0.04 ha to 0.08 ha if there were fewer than 8 trees (DBH greater than 12.7 cm) in the plot. The majority of the 0.08 ha plots are in the Apache-Sitgreaves stage 2 and 3 data collection efforts and were located within the perimeters of the Rodeo-Chediski and Wallow fires respectively.

Information was recorded on all trees with a dbh greater than 12.7 cm in six of the projects and a dbh greater than 20.3 cm for the Kaibab study region. Species and dbh were recorded for each live and standing dead tree that met the minimum dbh threshold for each project. Plot location was measured with a Trimble GeoXH6000 with GPS + GLONASS or a Trimble GeoXH with GPS using accuracy based logging settings. Plot center coordinates were recorded with a minimum of 200 positions in the Coconino and Tonto N.F.; and for a minimum of 10 minutes at the other sites. Differential correction was applied using Pathfinder Office.



**Figure 2.** Map of the seven field data collection efforts and the footprint of the four lidar acquisitions, in grey, subset by region. Plots used for model development are displayed in warm tones (red, orange, yellow); those used to assess transferability of model are blue. The three regions where data collection took place include: (A) Kaibab Plateau, AZ; (B) Four Forest Restoration Initiative, AZ, and (C) Southwest Jemez Mountain Landscape Restoration, NM.

### 2.2.2. Lidar

Four lidar data sets were acquired in our study area between 2012 and 2014 during leaf-on canopy conditions (June to September) (Figure 2). Each acquisition was surveyed with a Leica ALS series sensor with an opposing flight line side-lap greater than or equal to 50% (greater than or equal to 100% overlap) and an average native pulse density greater than or equal to 8 pulses per square meter over

terrain. While some variability in pulse density does occur across the acquisitions, all densities are within roughly one standard deviation of each other, with variability of pulse density between study regions equivalent to variation within each region, and thus should not produce significantly different return profiles. The targeted vertical accuracy (RMSE) for each acquisition was less than or equal to 15 cm. The field of view for each survey was generally between 26° and 28°, except for the 2000 m altitude survey in the North Kaibab which was only 20°. The complete lidar acquisition specifications for each site in this analysis are summarized in Table 2.

Canopy structure metrics were calculated from the raw lidar point cloud using FUSION software [74]. At each plot, we generated canopy height distributions and density metrics from the lidar point cloud using the relative height measure at each return. Relative height is the difference in terrain surface height (from the digital terrain model provided by the vendor) and the Z coordinate of each point. Canopy returns were points with a relative height above 3 m; in these forests anything lower is typically ground, stones, and low-lying vegetation [37,38]. Fractional canopy cover and density metrics were calculated using this static cover threshold and the dynamic thresholds of mean and mode values. We included a product of height quantiles and canopy density, which we refer to as canopy volume [18,37].

Metrics with a correlation in excess of 0.94 with other lidar variables were removed to reduce problems associated with highly collinear predictor variables, particularly ambiguous interpretation issues [52,75]. Most highly correlated variables existed as pairs with almost perfect correlation. The variable in each pair that had the higher correlation with the field AGB estimate was kept for analysis, the other was removed. The lidar metrics are listed in Table 3 (excluded metrics are located in Table S1).

**Table 3.** Table of assessed metrics.

	Variable	Definition
Height Metrics	Mode	height at mode
	Qmode	quadratic mode height
	P01, P10, P30, P60, P90	height at which the 1st, 10th, 30th, 60th, 90th percent of the points are below
	QP01, QP10, QP30, QP60, QP90	quadratic quantile heights
Height Distribution	SD	standard deviation
	Skewness, Kurtosis	skewness and kurtosis
	MAD Med., MAD Mode	median of absolute deviations from the overall median and mode
	L3, L4	3rd and 4th L-moments
	L-CV, L-skew., L-kurt.	L-moment coefficient of variation, skewness, and kurtosis
Canopy Cover & Density	CD	canopy density: number of all returns (>3 m) divided by total number of all returns
	Cov <sub>&gt;mean height</sub> :all	mean height cover: number of all returns above the mean divided by total number of all returns
	Cov <sub>&gt;3:1st</sub> Cov <sub>all &gt;mode</sub> :all <sub>first</sub>	number of all returns (>3 m) divided by total number of 1st returns number of all returns above the mode divided by total number of 1st returns
Volume	P01*CD, P10*CD, P30*CD, P60*CD, P90*CD	product of percentile height measures and canopy density
Environment	Elevation, Aspect, Slope	elevation, aspect, slope
	NDVI Ampl.	NDVI amplitude: a time-series analysis of seasonal greenness to represent phenology
	ERU	ecological response units which were aggregated into five categories (1) Colorado Plateau or Great Basin grassland, montane or subalpine grassland, (2) Pinyon-juniper woodland, (3) narrowleaf cottonwood and shrub, Arizona alder and and willow, willow and thinleaf alder, (4) spruce-fir, Mixed conifer that is freq. fire or mixed with aspen, (5) Ponderosa pine, with or without willow or evergreen oak

### 2.2.3. Topography

Elevation, slope, and aspect at each plot were determined from the Shuttle Radar Topography Mission (SRTM) digital elevation data with a resolution of 1 arc-second [76]. The topographic derivatives were calculated and sampled at each plot center in the Google Earth Engine platform [77].

### 2.2.4. Phenology

Phenology was represented by the amplitude of the seasonal difference in the normalized difference vegetation index (NDVI). We measured amplitude using a harmonic regression time-series analysis on Landsat images recorded between 2012 to 2015. Multiple linear regression is performed on NDVI observations, assuming that there is a sine curve (a harmonic, or Fourier transform) with a frequency of one cycle per year that describes the annual variation in NDVI, Equation (1) [78]. The  $\beta$ 's are the coefficients,  $f$  is the frequency, and  $t$  is time. Finally, to more easily interpret the cosine and sine coefficients ( $\beta_{cos}$  and  $\beta_{sin}$ ) we convert these to amplitude, ( $\beta_A$ ), and phase, ( $\beta_\phi$ ), (Equations (2) and (3)) [78]. These calculations and sampling were performed in the Google Earth Engine platform [77].

$$NDVI = \beta_1 + \beta_t t + \beta_{cos} * \cos(ft) + \beta_{sin} * \sin(ft) + e \quad (1)$$

$$\beta_A = \sqrt{\beta_{cos}^2 + \beta_{sin}^2} \quad (2)$$

$$\beta_\phi = \text{atan}(\beta_{cos}, \beta_{sin}) \quad (3)$$

### 2.2.5. Ecological Response Unit

Ecological Response Units are a system of mapped ecosystem types [64]. They were created using a combination of information on plant associations and structure characteristics that would occur under natural disturbance regimes and biological processes. For the analysis similar ecological response units were grouped into broader categories: herbaceous and grasslands (montane and subalpine grasslands, and Colorado Plateau and Great Basin grasslands), alder and willow (Arizona alder/willow and willow/thinleaf alder units), mixed conifer (frequent fire mixed conifer and mixed conifer with aspen units), and Ponderosa pine (Ponderosa pine, Ponderosa pine with willow, and Ponderosa pine with evergreen oak units).

## 2.3. Field AGB Estimates

AGB (stem, branch, and foliage biomass) and volume were calculated for each tree (live or standing dead) using the recorded DBH values, species-specific allometric equations, and wood densities [79]. AGB per plot was computed by summing the estimates for all trees within the plot. Analysis was performed using the Central Rockies variant of the Forest Vegetation Simulator (FVS) [79,80]. To take into account the sample design of the field data collection efforts, all sample-based summary statistics were calculated using functions within the survey package in R [81,82].

## 2.4. Estimating AGB with Lidar Metrics and Ancillary Data

Parametric linear multiple regression models have been successfully used to estimate AGB using some combination of metrics representing (1) the canopy height distribution (location, scale, and shape), (2) the density of the canopy, and (3) the interaction between height and density of the canopy metrics (Table 3) [42,46,65]. A visual inspection of linear fits between AGB and the height and canopy density metrics and a box cox analysis suggested a log transform of the response variable, AGB, similar to relationships identified in other studies (e.g., [46]). Therefore, we tested models built with a log transform of AGB and one with no transform. We also tested ecologically important environmental variables, including ecological response units, topography (elevation, slope, and aspect), and the amplitude of seasonal differences in the normalized difference vegetation index. We did not assess

non-parametric model functions, such as regression trees, because our response variable and covariates did exhibit strong linear relationships.

#### 2.4.1. Variable Selection with Bayesian Model Averaging and Stepwise Regression

Various methods are available to select a parsimonious set of metrics to use as predictors. One possible approach is to engage in dimensional reduction of the data, such as a principle component analysis and canonical correlation analysis [37,44,65]. However, the resultant factors can be difficult or impossible to meaningfully interpret. Stepwise regression is another commonly used approach to select covariates from a large pool of metrics (e.g., [18,38,44,45]), however this approach can result in overfit models and is sensitive to multicollinearity issues [83,84].

An alternative approach has emerged that avoids choosing a single model and instead generates a distribution of possible models that represents the inherent uncertainty that arises when many possible predictors (with possibly conflicting interpretations) exist [51]. This process is referred to as Bayesian Model Averaging (BMA), and it allows for a large pool of possible models to be enumerated and evaluated for how well they fit the data and for the uncertainty of model fit and parameter values to be clearly represented [85]. This approach reduces the possibility of researcher bias in variable selection than typical step-wise regression approaches [51]. Interestingly, the use of non-specific inference—that is, no single model is used to generate predictions, rather a large suite of possible models are used to produce an average prediction—often outperforms the use of any single model [52].

Enumeration and evaluation of large numbers of models can be problematic; however, for normal multiple linear regression, well-behaved, closed form solutions that allow for direct model comparison exist [85]. For managers, use of Bayesian model averaging for prediction may be problematic, because the Bayesian model averaging object used to generate the model is not intuitive. It also needs to be reproduced by a user for prediction on new data sets, as compared to the utility and ease of use of a single-model specification. Therefore, we present a final, single model from the ensemble generated by BMA. In order to provide information on how BMA methods compare to other established model selection methods, we examined the classic stepwise approach to model specification. We used bidirectional stepwise elimination. Both the BMA and stepwise model selection methods assessed multiple linear regression models fit using ordinary least squares.

Four model selection processes were assessed: (1) the multiple linear model selected from the stepwise regression process, (2) the Bayesian model average object, (3) highest probability multiple linear model from the BMA, and (4) the median probability multiple linear model from the BMA. The highest probability model is the single model with the highest probability of occurrence. The median model includes the set of lidar and biophysical variables that occurred in the population of models more than 50% of the time (posterior probability was greater than or equal to 0.5). We generated predictions using the BMA object using the top 10,000 models ('top' meaning highest posterior probability). While it is possible to use the full population of models to generate predictions, enumerating the full ensemble of models ( $2^{\text{number of covariates}}$ ) is computationally impractical, and most models have very low probability of occurring. These four model selection processes were completed on both the non-transformed and log-linear transformation of AGB. In total we report on performance of eight models.

#### 2.4.2. BMA Specifications

Bayesian model averaging was implemented with the Bayesian Adaptive Sampling (BAS) package [86] for use in R. BAS allows for more rapid exploration of model space than typical Markov Chain Monte Carlo methods, flexible model and prior specification, includes good diagnostic and predictive tools, and is well documented and under active development. We used a version of BAS that combines Markov chain Monte Carlo (MCMC) with the BAS algorithm, as MCMC approaches tend to be more tolerant of strong correlation between predictors (some of the remaining covariates were under the threshold, but are still correlated). Data were randomly subset for the purposes of model

development, with 25% of the data reserved for validation of the model fits. As the focus of this work is on deriving models useful for landscape-level management of forests at risk for destructive fires, we are primarily interested in sites with medium to high AGB. Sites with zero AGB were removed from the data before models were fit to the data.

A key component of Bayesian statistical approaches is specifying the prior distribution. The prior distribution provides a way of using known information to adjust how we view new data. If we have strong beliefs about the world, we are more critical of new data; conversely if we have a lot of uncertainty about the world, new data are largely left to speak for themselves. Two prior distributions need to be assigned before fitting of a model can be conducted using BAS: the prior that describes beliefs about model sizes, and the prior that describes beliefs about how likely it is that the coefficients of the model will be non-zero. While the initial pool of variables is quite large, many are correlated and may in effect be duplicates, so a large model is not necessary or likely to occur. Thus a truncated Poisson model with a mean of 10 covariates and a cut-off of 30 covariates was used; the cutoff sets the probability of larger models to zero. For the coefficients, the majority are likely to be zero, but some are expected to be highly significant. Zellner's  $g$  is often used to specify the prior for model coefficients, as it flexibly allows for a varying degree of belief about the coefficients to be included: a large  $g$  suggests little prior knowledge (and causes the coefficients to closely approximate their ordinary least squares counterparts), and a small  $g$  suggests strong skepticism that the coefficient will be non-zero [87]. We used a hyper- $g$ , a Beta prior on the shrinkage factor of the form in Equation (4), where  $a$  is a parameter in the range  $2 < a \leq 4$  [88]. The benefit of a hyper- $g$  is that we specify a moderately informative prior, splitting the difference between  $g$  approaching infinity and 0; but limit the risk of unintended consequences on the posterior results by allowing Bayesian updating of  $g$  to be used to adjust outcomes [89]. We set  $a$  to 3.

$$\frac{g}{1+g} \sim \text{Beta}\left(1, \frac{a}{2} - 1\right) \quad (4)$$

### 2.5. Model Evaluation and Assessment of Lidar AGB Estimates

Model performance was assessed in multiple stages using three subsets of the data. Five of the seven field data sets were used for model development and the remaining two were used as an independent validation to assess transferability of the developed model to new lidar acquisitions. The model development data set was split into a model calibration and validation set with approximately 75% used for calibration and 25% for validation. The model development data sets (calibration and validation) were used to complete the variable selection and trim multicollinear predictors from the model. The model development validation and transferability validation data were used to assess performance of the final trimmed model overall, by project site, and also by each plot size category. We assess the performance of the selected model when applied to new lidar acquisition using data completely independent of the model development phase—one lidar acquisition with two coincident field data collection efforts, which we refer to as the transfer validation data.

#### 2.5.1. Comparison of Product from Four Model Selection Procedures

Model performance of the eight models was assessed based on three criteria: (1) minimized model error and bias, (2) minimized multicollinearity and maximized reliability of estimates of each predictor, and (3) parsimony, the model with the fewest number of terms and predictors. We used these data subset to assess model performance in two stages. The first was completed within the model selection process by BMA and stepwise regression, this was implemented on the model calibration data set. To evaluate the performance of each of the six selected multiple linear regression models we report the coefficient of determination estimates ( $R^2$  and adjusted  $R^2$ ) calculated from the model calibration data. The second assessment phase was completed with the model validation data set (the 25% withheld from the five model development field data collection efforts). The root mean

square error (RMSE), percent root mean square error, bias, and mean bias are reported using this data subset for all eight models. We visually assess normality, constant variance, and independence of residuals by inspecting marginal plots. Issues with multicollinearity and reliability of predictor estimates were assessed using percent relative standard error (PRSE) and variance inflation factors. We present generalized collinearity factors of coefficients since we have interaction terms [90].

### 2.5.2. Model Refinement to Reduce Variance Inflation and Increase Reliability of Model Coefficients

Variance inflation factors greater than 5 suggest issues with multicollinearity [91], although Graham [83] cautions that values as low as two can have serious impacts on models. PRSE values of greater than 20% are considered unreliable in ecological studies [92]. We used these thresholds to trim our selected model structure. We maintained the raw terms of significant interaction predictors even if these metrics indicate they are not significant [93]. Terms are the covariate variables; model predictors include terms and combinations of terms (e.g., a second degree polynomial includes two predictors and one term—or three predictors when considering the constant).

### 2.5.3. Model Performance by Project and Plot Size

We report model bias and error metrics for the trimmed model overall, and we also include a summary of these metrics for each observed field data project and for each plot size category. Model bias can exist as fixed or proportional bias. Root mean square error and bias provide information on fixed bias—when values are higher (or lower) across the whole range of measurement. Proportional bias occurs when estimates diverge progressively along the range of values. We tested for proportional bias of the final trimmed model by fitting a linear model between field and lidar based AGB estimates using ordinary least squares and major axis regression [94]. Ordinary least squares regression of observed vs. predicted values is a popular method used in other studies; therefore we include it so that our results can be compared to these studies. However, since errors exist in both the lidar (predictors) and field based estimates (observed) ordinary least squares assumptions are not met, and therefore, it is ill-suited to assess proportional bias in this situation. Field-based estimates of AGB include uncertainty due to natural variability, measurement error, allometric model error and model selection choices [91,95]. Therefore, we also present results from major axis regression, which fits errors or natural variability on both variables symmetrically [96–98]. It is impossible to know if error is indeed symmetric between the two, but this approach strikes us as a more realistic assessment tool. Major axis regression was implemented using the Model II Regression package, `lmodel2`, from the R CRAN repository [99].

## 3. Results

### 3.1. Summary Statistics of Field Data Estimates

The average AGB of the model calibration sample was 122.3 ( $\pm 1.8$ ) Mg per hectare; the model validation sample average was 114.6 ( $\pm 2.9$ ) Mg per hectare. The composition of the model calibration sample was 72.8% Ponderosa pine forest, 25.5% mixed conifer, 0.5% spruce-fir forest, 0.5% pinyon-juniper woodland, 0.4% herbaceous-grassland, and 0.3% deciduous (narrowleaf cottonwood and shrub, alder, and willow). The transferability validation sample—stage 2 and stage 3 Apache-Sitgreaves N.F. data collection efforts—had average AGB of 71.1 ( $\pm 5.5$ ) and 89.5 ( $\pm 5.7$ ), respectively.

Table 4 includes average AGB values for each data collection effort, including sample and population estimates (statistic is adjusted based on the sample weights). For most projects, the average AGB of the field plot samples at each project site is higher than the population average for the entire site. This reflects our sample strategy aim to represent the full range of forest conditions. A population estimate for the selected Coconino and Tonto N.F. stands and the full model development data set is not appropriate, because the spatial extent of the combined projects is not a meaningful ecological or political unit. Only a selection of stands were sampled in the Coconino and Tonto N.F.; similarly only a selection of Ponderosa pine and mixed conifer forests of the southwest US were sampled.

**Table 4.** Sample and population summary statistics from the field data for each project. Estimates are provided for both the model calibration and validation data subsets. Mean and standard error of above-ground biomass (AGB) and elevation are reported.

Study Region	Model Construction Data			Validation Data		
	AGB <sub>population</sub> (Mg ha <sup>-1</sup> )	AGB <sub>sample</sub> (Mg ha <sup>-1</sup> )	Elev <sub>sample</sub> (m)	AGB <sub>population</sub> (Mg ha <sup>-1</sup> )	AGB <sub>sample</sub> (Mg ha <sup>-1</sup> )	Elev <sub>sample</sub> (m)
<b>All Model Dev. Sites</b>	-	<b>122.3 ± 1.8</b>	<b>2090 ± 4</b>	-	<b>114.6 ± 2.9</b>	<b>2090 ± 7</b>
Kaibab Plateau, AZ	121.3 ± 7.2	132.2 ± 9.3	2502 ± 18	126.8 ± 18.8	139.7 ± 19.3	2510 ± 35
Coconino N.F., 4FRI, AZ	-	128.5 ± 2.4	2160 ± 2	-	123.6 ± 4	2154 ± 4
Tonto N.F., 4FRI, AZ	-	113.9 ± 2.8	1913 ± 5	-	101 ± 4.6	1903 ± 8
Apache-Sitgreaves N.F., 4FRI, AZ, Stage 1	103.4 ± 4.4	107.9 ± 7.3	2238 ± 12	92.8 ± 5.7	93.5 ± 9	2230 ± 18
Southwest Jemez Mountain, NM	109 ± 6.7	117.7 ± 9.8	2493 ± 23	109.6 ± 7	94.2 ± 13.1	2475 ± 36
<b>Transferability Validation Sites</b>						
Apache-Sitgreaves N.F., 4FRI, AZ, Stage 2	-	-	-	57.2 ± 2.6	71.1 ± 5.5	2076 ± 10
Apache-Sitgreaves N.F., 4FRI, AZ, Stage 3	-	-	-	85.2 ± 3.6	89.5 ± 5.7	2570 ± 13

### 3.2. Assessing Results of Alternative Variable Selection Procedures

We analyzed eight models: the model selected through stepwise regression, the median probability and highest probability models from the BMA object and the BMA for two versions of the response variable—raw AGB and log-transformed AGB (Table 5). For each BMA model population, the median probability (the model composed of parameters with at least a 50% chance of being included in any model) and the highest probability model (the single model with the highest probability of occurring in the population of models) were the same. For brevity we refer to both of these models as the median probability model throughout the remainder of the document.

The stepwise variable selection process returned the largest (least parsimonious) models. The log-transformed AGB model specified from the stepwise selection process included 16 terms and 20 predictors. The ERU consists of 5 binary factors, one of which is represented within the intercept value. The stepwise model from the linear AGB included 11 terms and 17 predictors. The increase in terms and predictors did not substantially improve model performance as indicated by the similar adjusted coefficient of determination values. The log-transformed stepwise model had an adjusted coefficient of determination of 0.70; the adjusted R<sup>2</sup> value for the log-transformed median probability model was 0.69. The models without a transformed response variable, AGB, explained more of the variation in field-estimated AGB values. The adjusted R<sup>2</sup> values were 0.71 and 0.72 for the stepwise and BMA selected models respectively.

We selected the median probability (which is also the highest probability) model from the BMA object built with the non-transformed AGB value as the response variable. This model performs similar to the stepwise model but with far fewer terms and predictors. The root mean square error of the non-transformed AGB BMA ensemble was nearly identical to that of the median probability model with a difference less than 1 Mg per hectare. The median probability model explained 72% of the variation in the field based AGB estimates, had an RMSE percent value of 35.8, and negligible percent bias of −0.019. The results and figures reported in the subsequent sections are from this model structure.

**Table 5.** Model summary statistics of the estimates of the multiple linear models resulting from four model selection procedures— stepwise regression, median probability model (MPM) identified from Bayesian model averaging (BMA), the highest probability model (HPM) identified from BMA, and from the BMA ensemble object— each fit to a response variable, AGB, with no transform and with a log transform. R<sup>2</sup> and adjusted R<sup>2</sup> were calculated on the model calibration data. Root mean square error (RMSE), percent RMSE, bias and percent bias were all calculated on the validation data (the records withheld from model construction, not the transfer validation data set).

Model	Height Metrics	Canopy and Density	Cover	Volume	Environ.	R <sup>2</sup>	Adj. R <sup>2</sup>	RMSE (Mg/ha)	RMSE%	Bias (Mg/ha)	Bias%
Stepwise, ln(AGB)	P30, QP30 P60	CD		P60*CD P90*CD	elevation slope NDVI Ampl. aspect ERU	0.70	0.70	41.21	35.97	4.76	0.042
	P90, QP90 MAD Med. SD L-CV										
MPM, ln(AGB)	P30 P60, QP60 P90 MAD Med.	CD		P90*CD	elevation slope NDVI Ampl.	0.69	0.69	41.11	35.9	4.84	0.042
HPM, ln(AGB)	<i>same as MPM</i>										
BMA Object, ln(AGB)								41.19	35.96	4.99	0.044
Stepwise, AGB	P10, QP10 P30, QP30	Cov <sub>&gt;3;1st</sub> CD		P10*CD P30*CD P60*CD	elevation slope NDVI Ampl.	0.72	0.71	40.7	35.53	−2.34	−0.02
	P60, QP60 MAD Med. L-CV Kurtosis										
MPM, AGB	P60, QP60 MAD Med. P30	Cov <sub>&gt;3;1st</sub> CD		P30*CD P60*CD		0.72	0.72	41.01	35.8	−2.13	−0.019
HPM, AGB	<i>same as MPM</i>										
BMA Object, AGB								40.92	35.72	−2.19	−0.019

### 3.3. Median Probability AGB Model Structure

The median probability raw AGB multiple regression model includes five terms and eight predictors—nine including the intercept (Tables 5 and 6). The identified significant predictors include the second order polynomial of the 60th height percentile, an indicator of the canopy height variation (the median of absolute deviation from the overall median), a canopy density measure (all canopy returns divided by total number of first returns), and a lower and mid height canopy volume proxy (the product of the canopy density with the 30th and 60th percentile heights). The raw terms that make up the interactions are also included in the model; these include the 30th and 60th percentile heights, and canopy density. Only the lidar derivatives were selected to estimate AGB; no information on topography, phenology, or ecological response units was included in the model.

The variance inflation factor of the 30th height percentile and the two canopy volume proxies exceeded 5. The high PRSE metric of these predictors also suggests there are issues with these estimates that need to be remedied. Therefore we removed the canopy density normalized by first returns (Cov<sub>>3;1st</sub>) and the volume term with the 60th percentile height. The trimmed model included one less term, two fewer predictors, and explained 71% of the variation in field AGB estimates (both R<sup>2</sup> and adjusted R<sup>2</sup> were 0.71). Table 6 includes the model specifications of the full and trimmed models. For both models, at least one predictor of all terms was significant at  $p < 0.001$ . The 60th percentile was not significant in the trimmed model. However, it was kept in the model because it is a term in the polynomial predictor, which is significant [93]. The raw and interaction terms had PRSE and generalized variance inflation factor values that exceeded the recommended thresholds, but colinearity of predictors created from the same terms is expected. An examination of the marginal model plots shows that the quadratic height term, QP60, improves model fit by pulling the estimates of plots with high and low biomass values closer to those observed in the field; without the quadratic term they are

under- and over-predicted, respectively. All subsequent analysis was conducted using this trimmed model (Table 6).

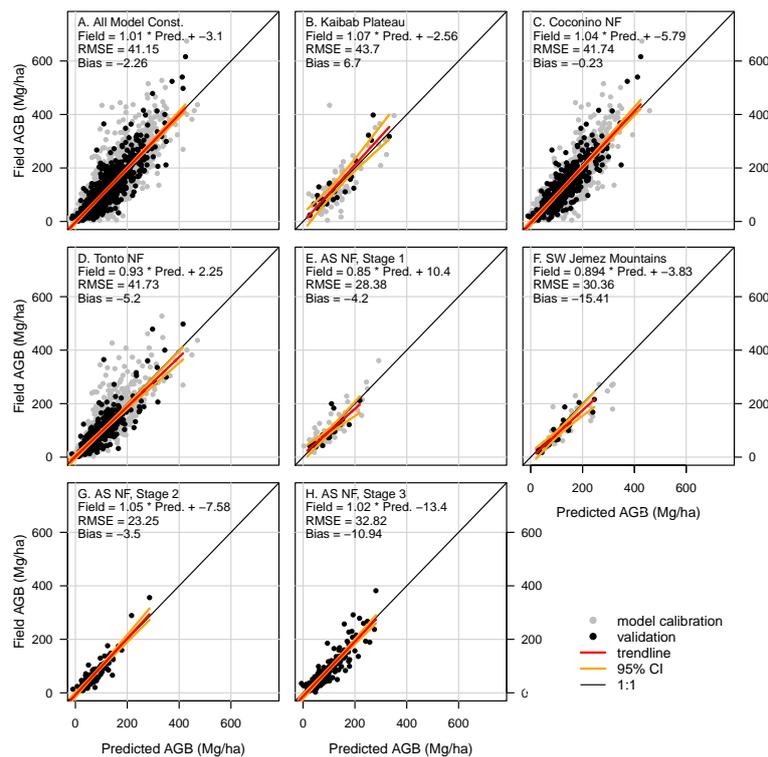
**Table 6.** Full and trimmed (non-transformed) AGB median probability model specifications. GVIF is the generalized variance inflation factor. The significance of the relationship between each predictor and the response is indicated as follows: \* is less than 0.05, \*\* is less than 0.01, and \*\*\* less than 0.001; others are less than 1.

Predictors	Full Model					Trimmed Model				
	Coef.	Std. Error	Signif.	PRSE	GVIF	Coef.	Std. Error	Signif.	PRSE	GVIF
Intercept	−33.62	16.1	*	47.89		−9.78	13.93		142.5	
<i>Canopy Height Metrics</i>										
P30	−1.21	1.59		131.35	7.52	−4.66	1.19	***	25.63	5.62
P60	−837.93	351.15	*	41.9	3.19	−68.73	249.3		362.7	2.55
QP60	396.45	59.73	***	15.07	3.19	457.74	54.91	***	12	2.55
<i>Canopy Height Distribution</i>										
MAD Median	10.02	1.58	***	15.77	2.57	10.94	1.54	***	14.08	2.48
<i>Canopy Cover and Density</i>										
Cov <sub>&gt;3;1st</sub>	0.44	0.098	***	22.16	2.53	<i>removed due to variance inflation issues</i>				
CD	0.11	0.23		209.17	3.72	1.01	0.16	***	15.57	2.55
<i>Canopy Volume</i>										
P30*CD	0.15	0.033	***	33.19	10.08	0.24	0.015	***	6.47	4.69
P60*CD	0.083	0.028	**	22.54	11	<i>removed due to variance inflation issues</i>				

### 3.4. Trimmed Median Probability Model Performance, Overall and by Field Data Collection Site

The overall percent root mean squared error between the field observed AGB and the AGB estimate from the trimmed model was 35.93% for the model validation sample. It was slightly lower for the two transfer validation data sets: 31.18% and 32.83% for the phase 2 and phase 3 projects within Apache-Sitgreaves N.F. Three of the five samples with AGB field-estimates above 400 Mg per hectare had percent RMSE values above 41% (Figure 3, Table 7). These included the Kaibab Plateau calibration sample, and both the calibration and validation sample from the Tonto N.F. projects. Upon visual inspection of the lidar point cloud and field survey photographs, we found the Kaibab Plateau calibration data set had a plot with obvious edge effect error caused by a tree with large biomass at the edge of the plot border. This partially explains the large RMSE value for this sample.

Overall model bias was negligible for the model calibration and validation data sets, with values of  $-1 \times 10^{-15}$  and  $-0.02$  respectively. (Table 7). For most project samples the percent bias values ranged between  $-5$  and  $5\%$  between project samples. The exceptions were the 1) Southwest Jemez Mountain data collection validation at  $-16.36\%$  and calibration data at  $-19.25\%$ , the phase 3 Apache-Sitgreaves collection at  $-12.22\%$ , and the Tonto N.F. validation data set at  $-5.15\%$ . However, the 95% confidence intervals of the ordinary least squares trend line between the field and lidar based estimates includes the 1:1 line across the range of values at each of the seven sites, indicating the bias is not significantly different than zero (Figure 3). The 95% confidence interval on the trend line for the Southwest Jemez Mountain project does, however, not fully enclose the 1:1 line (Figure 3f). Further, the Southwest Jemez Mountain data collection effort has the largest negative bias. This project has moderate AGB values; none of the plots have a field-estimated AGB over 400 Mg/ha.



**Figure 3.** Scatter plot of the field-measured AGB vs. the predicted values from the trimmed median probability model (Table 6). Plots include data from: (A) all calibration and validation data used during model construction (data in plots b-f); (B) Kaibab Plateau, (C) Coconino N.F. (4FRI); (D) Tonto N.F. (4FRI); (E) Apache-Sitgreaves N.F. (4FRI) Stage 1; (F) Southwest Jemez Mountain; (G) transfer validation data from the Stage 2 Apache-Sitgreaves N.F. (4FRI) project; and (H) transfer validation data from the Stage 3 Apache-Sitgreaves N.F. (4FRI) project. Red and orange lines depict the linear fit and 95% confidence interval band of field-measured AGB versus predicted values, fit to the model validation data subset. Black line is 1:1.

**Table 7.** Trimmed median probability model validation statistics. RMSE, percent RMSE, bias and percent bias were all calculated on the independent validation data sets.

Project Site	Validation or Model	n	RMSE (Mg/ha)	RMSE%	Bias (Mg/ha)	Bias%
Model Construction Data	Validation	793	41.15	35.93	-2.26	-0.02
	Calibration	2271	45.29	37.04	$-2 \times 10^{-13}$	$-1 \times 10^{-15}$
Kaibab Plateau, AZ	Validation	25	43.7	31.28	6.7	4.79
	Calibration	87	55.19	41.75	-1.01	-0.77
Coconino N.F., 4FRI, AZ	Validation	448	41.74	33.76	-0.23	-0.2
	Calibration	1218	43.84	34.13	-0.51	-0.4
Tonto N.F., 4FRI, AZ	Validation	272	41.73	41.34	-5.2	-5.15
	Calibration	847	47.43	41.66	2.061	1.81
Apache-Sitgreaves N.F., 4FRI, AZ, Phase 1	Validation	27	28.38	30.36	-4.2	-4.5
	Calibration	73	28.2	26.13	0.16	0.15
Southwest Jemez Mountain, NM	Validation	21	30.36	32.23	-15.41	-16.36
	Calibration	46	43.82	37.22	-22.66	-19.25
Transferability Validation Data						
Apache-Sitgreaves N.F., 4FRI, AZ, Phase 2	Validation	96	23.25	32.71	-3.5	-4.92
Apache-Sitgreaves N.F., 4FRI, AZ, Phase 3	Validation	150	32.82	36.66	-10.94	-12.22

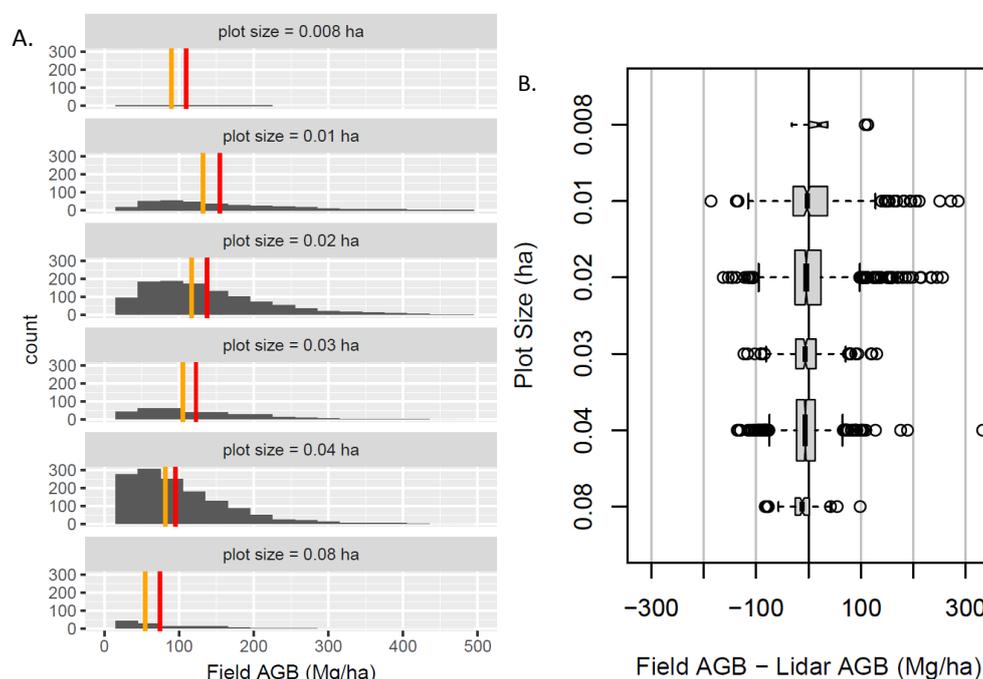
### 3.5. Influence of Inconsistent Plot Size

In this section we assess how AGB estimation performance varies with regard to the different plot size protocols used in the seven data collection efforts. In the Coconino N.F. and Tonto N.F. when the plots were located in stands with high tree density the plot size was reduced, proportional to tree

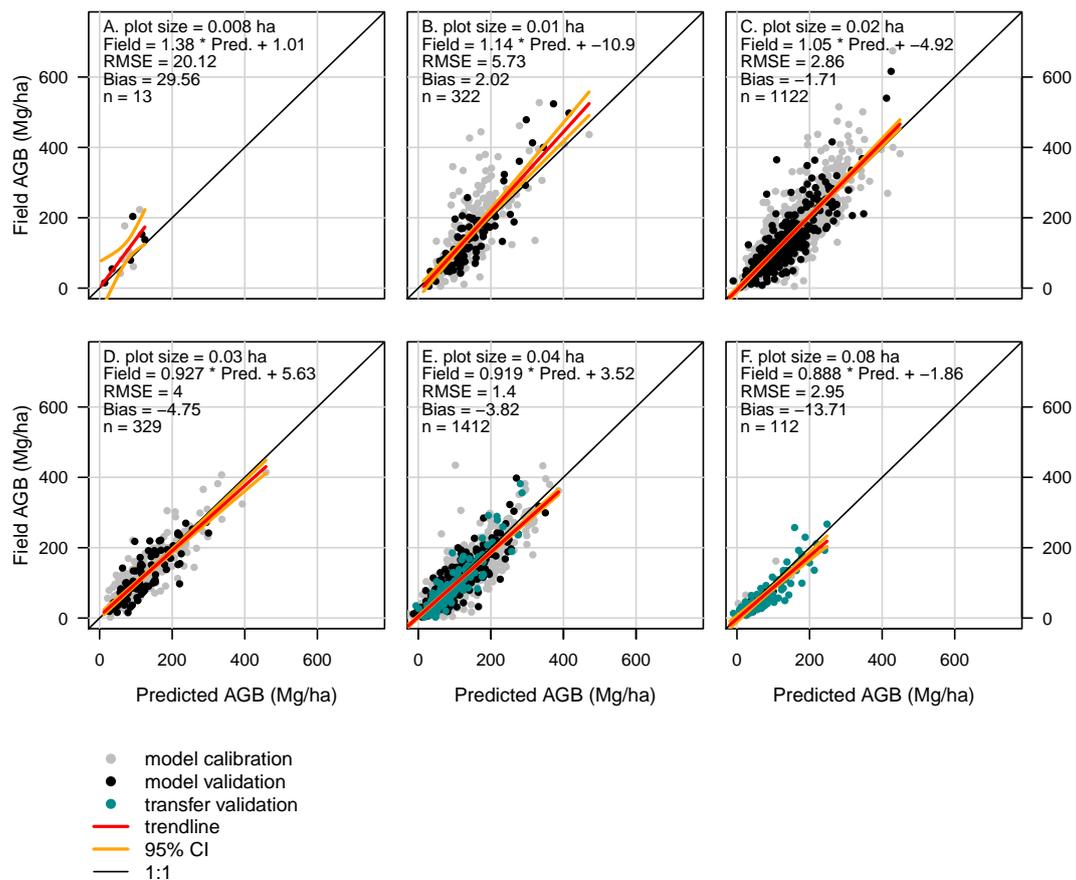
density. Conversely, in the three Apache-Sitgreaves N.F. and Southwest Jemez Mountain data collection efforts plots with low tree density were doubled in size to 0.08 ha. A majority of the smallest plots are composed of a high density of small stature trees, whereas forest structure in the largest plots (0.08 ha) was defined by larger, thinly-dispersed trees. An examination of the AGB distributions by plot size confirms that the smallest and the largest plot sizes—the 0.008 ha and 0.08 ha plots, respectively—both have relatively low AGB values (Figures 4a and 5a,f).

The error rates are not consistent across all plot sizes (Figures 4b and 5). The residuals, the difference between the field-based estimates and the lidar-estimated AGB values, are centered around zero for the four plot sizes ranging from 0.04 to 0.01 ha. However, the smallest and largest plot sizes have residuals that are not equal to zero (paired *t*-test *p*-value = 0.04 and  $<1 \times 10^{-5}$ ) (Figure 4b). The RMSE value of the smallest plots is 20.12 and the bias is 29.56 Mg per hectare (Figure 5a). These 0.008 ha plots are all located in one stand that had a high density of small DBH trees in the Tonto N.F.; these plots have lidar-based AGB estimates higher than those estimated from the field (average difference of 31.19 Mg/ha). The inverse is true for the large, sparsely populated stands (0.08 ha plots) where lidar-based estimates are on average 11.57 Mg/ha less than field-estimates and a bias of  $-13.71$  Mg/ha. The RMSE value of these plots is relatively small at 2.95 Mg/ha.

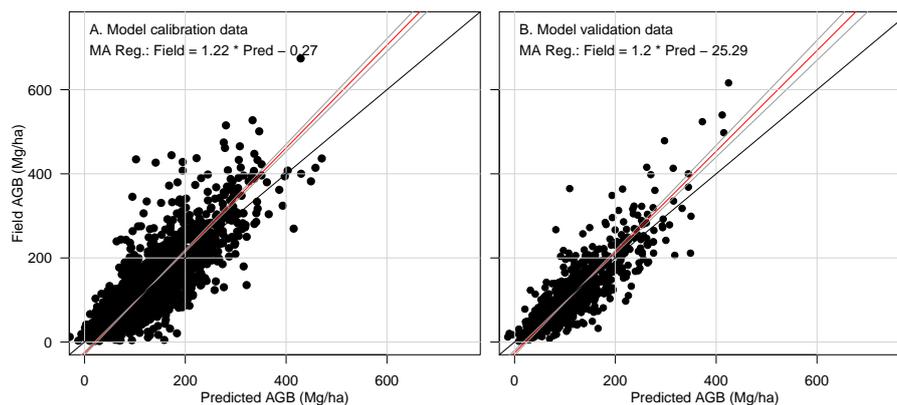
Figure 5 shows that the majority of plots with high field-based AGB estimates (above 400 Mg/ha) are located in plots smaller than 0.04 ha—most are within 0.01 and 0.02 hectare plots. The plots with high field-estimated AGB values ( $>400$  Mg/ha) consistently have lower lidar-based estimates. The average disagreement (residual between field- and lidar-estimates) is 142.33 Mg/ha. A trend line fit between the field- and lidar-based estimates using major axis regression indicates that there is proportional disagreement between the lidar and field estimates (Figure 6). The effects are most evident in these plots with high AGB field values.



**Figure 4.** Distribution of AGB and trimmed median probability model disagreement by plot size: (A) Histogram of field-estimated AGB for each of the plot sizes. The red line is the group mean, orange is group median and (B) Box plot of disagreement between field-estimated and lidar-estimated AGB from the trimmed median probability model for each of the different plot sizes. Plot width is proportional to the square root of the number of observations in each group. Notches indicate a 95% confidence interval of the mean [100].



**Figure 5.** Scatter plot of field-measured AGB versus predicted values from the raw AGB trimmed median probability regression model (Table 6) for each of the plot sizes used in the study. Plot sizes for each window are: (A) 0.008 ha, (B) .01 ha, (C) 0.02 ha, (D) 0.03 ha, (E) 0.04 ha, and (F) 0.08 ha. Red and orange lines depict the linear fit and 95% confidence interval band of field-measured AGB versus predicted values, fit to both the model validation and transfer validation data subsets. Black line is 1:1.



**Figure 6.** Scatter plot and major axis regression trend line of field-measured AGB versus predicted values from the raw AGB trimmed median probability linear regression model (Table 6). Plots include data from: (A) all model calibration data, (B) all validation data from the model construction data set. Red dashed line is linear fit from the major axis regression of field-measured AGB versus predicted values on the independent validation data subset. Grey lines are 95% confidence intervals from 1000 permutations. Black line is 1:1.

#### 4. Discussion

As expected, using the ensemble of models from BMA performed well for estimation of AGB. However, broadly adopting the BMA approach for prediction may not be appropriate, given the added complexity and computational cost, particularly when scaling the estimation up to the raster level. A single multiple regression equation is more practical for technicians to implement, only requires access to a GIS with raster processing capacity, and produces results similar to the more complicated method. Further, models with clear biological interpretations and that can be related to ecological theory are typically preferred and are easier check for obvious disconnects with ground conditions. The BMA approach can be used to select a single multiple regression equation, by identifying the median or highest probability model. Even when selecting a single model, the BMA approach considers far more of the possible model space and reduces the possibility of researcher bias in variable selection than typical step-wise regression approaches [51].

We have demonstrated the BMA median and highest probability models are robust and perform well in this application. The highest and median probability models identified by the BMA process produced a parsimonious, interpretable model that explained 72 percent of the variation in the field-based AGB estimates of the sample of plots. Close agreement in magnitude between RMSE from the data used to build the model and the cross validation data, as well as consistent performance across the regions involved in the study suggest that the model is not over fit and is suitable for generalization in Ponderosa pine and mixed conifer forests in the southwestern US. The terms and predictors arrived at via our approach can be reasonably interpreted to have direct analogs to ecologically significant variables. Lidar height, density, and distribution metrics correspond to variables used in forest inventory aerial stand volume tables.

Our final, trimmed model contains variables that relate strongly to the vertical and horizontal extent and central tendency of the canopy, as well as its density and volume. Both the relationship of stem diameter to biomass and stem height to biomass have been well-studied in the biomass and allometric estimation literature, and scaling relationships are well supported by empirical work [91]. Lidar cannot directly measure stem diameter in most forest types, but canopy extent does relate to stem diameter; crown radius and crown area can both be related to stem diameter via simple power laws of the form  $Y = Y_0 M^b$ , where  $Y$  is a biological variable of interest (crown area),  $Y_0$  is a normalization constant,  $M$  is a measured biological variable (stem diameter), and  $b$  is the scaling exponent [101]. The set of lidar metrics selected in our final model describes the location of the thickest part of the canopy, and biomass estimation theory indicates this should be strongly correlated to biomass of the site. Similarly, the median of absolute deviations from the overall median (MAD\_median) describes the vertical variability of the canopy and may help the model to account for the complex of intermediate tree crown in the over-story and suppressed trees in the understory [46]. This combined with the height metrics represents the vertical distribution and extent of canopy.

##### 4.1. Model Bias

Our model exhibited a pattern where plots with large AGB field-estimates were under-predicted by the lidar model. We observe this pattern in scatter plots of observed field vs lidar predictions of other studies (e.g., [38,46,102–104]). This disagreement can be partially explained by known errors associated with plot sizes, discrepancies between a minimum DBH requirement in the field and lidar sensors that return pulses from vegetation regardless of DBH thresholds, error structure of the field-based estimates, and model structure. Sheridan and colleagues [103] remedied the issue by using a square root transform of the response; this transformation exacerbated our model bias. Estimates of AGB from field data include measurement error, allometric model error, and error from the choice of allometric model [91]. The magnitude of these measurement errors increases with biomass [105]. Our log-transformed model performed similar to our natural AGB model, however the performance of these models might shift if we had more information about the error structure of the high biomass sites. An examination of the marginal plots of the high biomass sites that were

well-predicted (300–400 Mg/ha) and those that were under-predicted (>400 Mg/ha) indicate negligible differences in lidar metric characteristics between the two groups.

Model errors decrease with increasing plot size [23,65,106]. The relationship is non-linear and asymptotic, and the influence levels off at a plot size of around 0.2 ha (well above our maximum plot) [23,65]. This is partially explained by the discrepancy between the amount of AGB estimated from field measurement vs. lidar returns due to edge effects. Lidar sensors record information from trees with stems outside the plot boundary but with crowns that extend into the plot; conversely, AGB from a tree near the inside edge of a plot may be less than the amount represented by the portion of the canopy recorded by the lidar sensor. A larger plot radius has a smaller perimeter-to-area ratio, mitigating discrepancies between field and laser measurement protocols at plot edges [65,106]. Co-registration errors are reduced in larger plots due to the higher degree of spatial overlap [65]. Gobakken and Naesset [106] reported that plots larger than 0.03 ha were generally unaffected by positional errors of 5 m or less; however 0.02 ha plots exhibited substantial biases in the estimation of height, basal area, and volume due to slight positional mis-registrations. Small plots have substantial variation around canopy height quantiles which increases disagreement between lidar predictions and field-based estimates.

Forty-four percent of the plots in our study are 0.02 ha or smaller, increasing the positional errors as well as the possibility that the lidar and plot data do not represent the same conditions. The increased positional errors and edge effects for locations with large trees (with larger canopies that extend into the study area) that are captured in one dataset but not the other likely contribute to poor model performance in the upper ranges of AGB. Plot size was linked to stem density in the sample design, so a large proportion of the sites with high biomass values were recorded on small plots. All but three of our plots with AGB values in excess of 400 Mg/ha were recorded on plots 0.02 ha or smaller; these same plots tend to have field AGB estimates far in excess—on average 142.33 Mg/ha—of the lidar based estimates. Additionally, there are errors associated with extrapolation of the field-based AGB estimates for large stature trees. The allometric equations used to convert field dbh measurements to biomass estimations were developed on a sample of trees that had a max dbh of 25; just under 2% of the trees in our sample had a larger dbh.

We have identified biases in our model that have implications for determining when estimates will be accurate enough for different management applications. The model under predicts AGB in areas with high field AGB estimates (>400 Mg/ha). This has real consequences to management in terms of carbon accounting and perhaps in the identification of fuel loads. For example, the model will likely yield a lower, conservative estimate of total carbon at the landscape scale. However, as areas with very high AGB make up a small proportion of these forested landscapes, we consider these estimates to still be relevant and the model useful for application at broad scales. We also have some reason to question the sensitivity of the model to discern differences in structure of low AGB plots with a high density of small trees vs. a low density of mature trees. This warrants further investigation to determine the suitability of the model in prioritizing where to apply some restoration treatments, such as stand thinning. To refine the model, we suggest an intensified collection of data in areas with AGB in excess of 400 Mg/ha and across a range of low AGB conditions. Finally, data collection efforts that cover the full extent of Ponderosa pine and mixed conifer forests are required to get more precise model error estimates; Johnson and colleagues [102] describe limitations to the application of models developed with data sampled from a narrow definition of forests to regions with tree cover that are not within that definition. Understanding these implications is especially important to determining whether lidar-based models perform well at the interface of public forest and communities, where the costs of fire and fire suppression are the greatest.

By combining data from projects with different plot size protocols, we are in an interesting position to examine the potential unintended consequences of cost savings efforts—determining plot size based on stem density—on lidar-based monitoring products. While allowing contractors to collect information on smaller plots in high density stands reduces time and costs of field data

collection, our findings suggest that these savings have practical implications on the ability to monitor the landscape and may cost more in the long term. Field data protocols that will assist in remedying disagreement between field and model predictions include consistent plot sizes with a minimum size of at least 0.04 ha.

#### 4.2. Relationship to Other Modeling Efforts

Our model is consistent with other studies conducted in the same region [37–39]. Hall [39] proposed a model using the proportion of ground returns that were not intercepted by the canopy fit using a sample of Ponderosa pine and Douglas fir plots in the Front Range of the Rocky Mountains, CO. Their model had a coefficient of determination similar to ours, 0.74. Sherrill and colleagues [37] used a canonical correlation analysis to predict AGB with a coefficient of determination of 0.76 and a RMSE of 36.5 Mg/ha on a sample from subalpine forests of the Central Rockies. Kim and colleagues [38] proposed a lidar-based model fit to estimate live and dead AGB in Ponderosa pine and mixed conifer forests in the North Rim of the Grand Canyon National Park, a small subset of the forests we have examined. Their best model for (non-transformed) live AGB had an RMSE of 46.01 Mg/ha (23.66% RMSE) and a coefficient of determination of 0.76. Our large sample size, full range of AGB conditions, and expansive spatial footprint enable us to build on their research. The data from these three studies had a limited range of AGB values; a max less than 300 Mg/ha [37], a max below 400 but with only 4 plots above 150 Mg/ha [39], and a max less than 400 Mg/ha [38]. Sample sizes were small, ranging from 36 [37] to 58 [38]. However, plot sizes were larger; 0.1 ha [38] and 0.32 ha [39].

The lidar covariates Kim [38] selected for their live AGB model are nearly identical to our total AGB model when you take into account the strongly co-linear nature of many lidar derivatives (Table S1). Their model included a volume product (mean height and canopy cover), 20th percentile height, mean height, and variation of the height metrics. Our proposed model structure includes the addition of theoretically sound predictors that improve on their model limitations. The Kim and colleagues [38] model did not include volume metrics on multiple height quartiles nor did they assess quadratic height quartiles. We found these to be valuable; the 30th percentile metric appeared in our model as a volume metric ( $P30 \times CD$ ) as did the quadratic term of the mean height equivalent. The inclusion of the quadratic height term, QP60, also improved our model fit, reducing the tendency of the model to under-predict plots with high AGB values and over-predict those with low values. A comparison of their scatter plot of predicted to observed values indicates that their model under-predicts high AGB plots (starting at about 250 Mg/ha) and over-predicts low AGB plots, especially those with close to zero AGB [38].

#### 4.3. Management Implications

If management to prevent fires that generate ecological regime changes is a goal, it is important to identify and characterize high AGB stands. AGB is a key variable to track for management in terms of fire behavior and restoration to a fire-adapted system. Excess fuels are often densely growing, small-diameter trees with low bases to their crowns, a condition that increases risk of high-severity, stand-replacing fires. Our model builds on existing data sets and allows for regional assessment and identification of high AGB stands based on lidar data and minimal investment in manual collection of stand characteristics. Once such stands have been identified, further work can go forward in designing effective management strategies.

It is vital to consider that management must take into account more than ecological concerns in National Forests and CFLRP projects; such forests are managed for multifactorial purposes, and economic and community considerations need to be balanced with ecological concerns. Once identified, thinning of high AGB stands can create biomass as a key product. Some of the biomass will be in the form of trees large enough to be utilized for traditional timber uses, while the rest of the biomass will be in the form of small diameter timber that is most often used for poles, chipped, or burned in piles onsite. Considering possible revenue streams that can be generated in CFLRP and

USFS forests is important, as it can be a challenge to get treatments to pay for themselves, particularly in cases of removal of trees too small for traditional timber uses.

## 5. Conclusions

We present a cost effective approach to use previous data collection efforts to assist in developing models and estimates of AGB—and potentially other forest structure attributes—for new lidar acquisitions where suitable field plot data is not available. While this approach still requires a limited field work campaign to validate the performance of the model on new lidar data, use of this predictive model reduces the size of the field data collection efforts, offering significant time and cost savings. Lidar-based regional AGB models have been developed for boreal, temperate deciduous, temperate coniferous, and tropical forests [40–44]. This study presents a novel contribution by being among the first to develop a regional AGB lidar-based model for Ponderosa pine and mixed conifer forests of the southwest USA. The BMA model selection produced a parsimonious, interpretable model that explained 71% and 72% of the variation in the field based AGB estimates of the sample of plots, for the trimmed and full models respectively. The terms and predictors arrived at via our approach can be reasonably interpreted to have direct analogs to ecologically significant variables: lidar height, density, and distribution metrics correspond to variables used in forest inventory aerial stand volume tables.

Model root mean square error was 45.29 Mg/ha; comparable to other published regional lidar-based AGB models [23]. The final biomass models performed well when they were used to predict observed values in the 4FRI stage 2 and 3 lidar datasets (independent dataset acquired later in the analysis). The RMSE of the model cross validation and the two transferability validation data sets were 41.15, 23.25, and 32.82 Mg/ha respectively. Close agreement in magnitude between RMSE from the data used to build the model and the validation data, as well as consistent performance across the regions and field collection designs involved in the study suggest that the model is not overfit and is suitable for generalization. The lidar data used in this analysis were collected using a Leica ALS series lidar sensor with identical range of flight specifications. As the lidar industry evolves, instrument development advances, and new sensors become operational (e.g., multi-wavelength lidar, Geiger-mode, or single photon systems) the transferability of this regional model will need to be reevaluated and parameterized to match new technologies. The cover and height percentile metrics in our model are relatively more robust than others across a variety of lidar sensor platforms. However, point cloud metrics, such as the (MAD\_median), are known to be sensitive to variations in the technical properties of sensors [107–109]. The model presented here is trained on data from Ponderosa pine and mixed conifer forests in the southwest US and lidar with similar data acquisition specifications (see specs in Table 2). It should only be applied when the domain of a new lidar acquisition with similar specifications covers these forest types.

### *Directions for Future Research*

As new validation data become available they can be used to refine the model with Bayesian model updating techniques. This approach can be used to improve the known model shortcomings due to the influence of high disagreement between field and model AGB estimates at the upper range of AGB due to small plot size. Hierarchical Bayesian models have proven to be robust in individual tree biomass estimation models [110,111].

The focus of this research was on AGB, but we expect this approach can be duplicated to develop regional lidar-based models to monitor other forest structure attributes that are well suited to estimation by lidar (e.g., see [22]). Examples of forest characteristics of particular importance in these fire-prone forests include timber volume, canopy fuels [54,55,57], monitoring management intensity [112], and standing dead biomass [38]. Recognizing the broad applicability of lidar acquisitions (hazards, terrain mapping, etc.) and the decreased unit cost as scanned surface increases, agencies are partnering to form lidar consortiums to fund the continued acquisition of lidar covering a large spatial extent.

Therefore, the application of this methodology has the possibility to provide estimates of important biological characteristics of large areas at relatively low cost, using large volumes of already extant data.

**Supplementary Materials:** The following are available online at [www.mdpi.com/2072-4292/10/3/442/s1](http://www.mdpi.com/2072-4292/10/3/442/s1), Table S1: Correlation statistics of excluded lidar metrics; metrics that had a correlation of 0.94 or greater were not considered in the BMA.

**Acknowledgments:** This research was performed under contract to the US Forest Service, Geospatial Technology and Applications Center (GTAC) and in support of Southwestern Region management goals. We are grateful for Haans Fisk and the support from GTAC to provide meaningful solutions using Earth observing data to assist in monitoring landscapes and informing forest restoration efforts. Kim McCallum conducted a pilot of this study. The authors would also like to acknowledge the assistance Richard Reynolds, Sandy Boyce, and other researchers working with the Rocky Mountain Research Station in Fort Collins, CO, for providing field data covering the Kaibab Plateau. We are grateful for comments provided by Robert J. McGaughey, Karen Dyson, and four anonymous reviewers, they offered valuable insight and suggestions to help improve the quality of the manuscript.

**Author Contributions:** K.T., M.S.P., T.M., M.N., P.J., and B.M. conceived and designed the experiments; T.M., M.N., P.J., and B.M. processed field and lidar data; K.T. and M.P. analyzed the data and wrote the paper.

**Conflicts of Interest:** The authors declare no conflict of interest.

## Abbreviations

The following abbreviations are used in this manuscript:

BMA	Bayesian model average
CFLRP	Collaborative Forest Landscape Restoration Program
dbh	diameter at breast height
4FRI	Four Forest Restoration Initiative
N.F.	National Forest
PRSE	percent relative standard error
RMSE	root mean square error
RMSPE	root mean square predicted error

## References

1. Brown, R.T.; Agee, J.K.; Franklin, J.F. Forest restoration and fire: Principles in the context of place. *Conserv. Biol.* **2004**, *18*, 903–912.
2. Jolly, W.M.; Cochrane, M.A.; Freeborn, P.H.; Holden, Z.A.; Brown, T.J.; Williamson, G.J.; Bowman, D.M. Climate-induced variations in global wildfire danger from 1979 to 2013. *Nat. Commun.* **2015**, *6*, 7537.
3. Covington, W.W.; Moore, M.M. Post settlement changes in natural fire regimes and forest structure: Ecological restoration of old-growth ponderosa pine forests. *J. Sustain. For.* **1994**, *2*, 153–181.
4. Huffman, D.W.; Ziegler, T.J.; Fule, P.Z. Fire history of a mixed conifer forest on the Mogollon Rim, northern Arizona, USA. *Int. J. Wildland Fire* **2015**, *24*, 680–689.
5. Cooper, E. Changes in vegetation, structure, and growth of southwestern pine forest since white settlement. *Ecol. Mon.* **1960**, *30*, 129–164.
6. Moore, M.M.; Huffman, D.W.; Fule, P.Z.; Covington, W.W.; Crouse, J.E. Comparison of historical and contemporary forest structure and composition on permanent plots in southwestern ponderosa pine forest. *For. Sci.* **2004**, *50*, 162–176.
7. Strahan, R.T.; Sánchez Meador, A.J.; Huffman, D.W.; Laughlin, D.C. Shifts in community-level traits and functional diversity in a mixed conifer forest: A legacy of land-use change. *J. Appl. Ecol.* **2016**, *53*, 1755–1765.
8. Flannigan, M.D.; Stocks, B.J.; Wotton, B.M. Climate change and forest fires. *Sci. Total Environ.* **2000**, *262*, 221–229.
9. McKenzie, D.; Gedalof, Z.E.; Peterson, D.L.; Mote, P. Climatic change, wildfire, and conservation. *Conserv. Biol.* **2004**, *18*, 890–902.
10. Westerling, A.L.; Hidalgo, H.G.; Cayan, D.R.; Swetnam, T.W. Warming and earlier spring increase western US forest wildfire activity. *Science* **2006**, *313*, 940–943.

11. Liu, Y.; Stanturf, J.; Goodrick, S. Trends in global wildfire potential in a changing climate. *For. Ecol. Manag.* **2010**, *259*, 685–697.
12. Seidl, R.; Thom, D.; Kautz, M.; Martin-Benito, D.; Peltoniemi, M.; Vacchiano, G.; Wild, J.; Ascoli, D.; Petr, M.; Honkaniemi, J.; et al. Forest disturbances under climate change. *Nat. Clim. Chang.* **2017**, *7*, 395–402.
13. Abatzoglou, J.T.; Williams, A.P. Impact of anthropogenic climate change on wildfire across western US forests. *Proc. Natl. Acad. Sci. USA* **2016**, *113*, 11770–11775.
14. Schultz, C.A.; Jedd, T.; Beam, R.D. The Collaborative Forest Landscape Restoration Program: A history and overview of the first projects. *J. For.* **2012**, *110*, 381–391.
15. Schoennagel, T.; Nelson, C.R.; Theobald, D.M.; Carnwath, G.C.; Chapman, T.B. Implementation of National Fire Plan treatments near the wildland–urban interface in the western United States. *Proc. Natl. Acad. Sci. USA* **2009**, *106*, 10706–10711.
16. Nagendra, H.; Lucas, R.; Honrado, J.P.; Jongman, R.H.; Tarantino, C.; Adamo, M.; Mairota, P. Remote sensing for conservation monitoring: Assessing protected areas, habitat extent, habitat condition, species diversity, and threats. *Ecol. Indic.* **2013**, *33*, 45–59.
17. Dubayah, R.; Drake, J.B. Lidar remote sensing for forestry. *J. For.* **2000**, *98*, 44–46.
18. Lefsky, M.A.; Cohen, W.B.; Parker, G.G.; Harding, D.J. Lidar remote sensing for ecosystem studies: Lidar, an emerging remote sensing technology that directly measures the three-dimensional distribution of plant canopies, can accurately estimate vegetation structural attributes and should be of particular interest to forest, landscape, and global ecologists. *AIIBS Bull.* **2002**, *52*, 19–30.
19. Hyyppä, J.; Hyyppä, H.; Leckie, D.; Gougeon, F.; Yu, X.; Maltamo, M. Review of methods of small-footprint airborne laser scanning for extracting forest inventory data in boreal forests. *Int. J. Remote Sens.* **2008**, *29*, 1339–1366.
20. Goetz, S.J.; Dubayah, R.O. Advances in remote sensing technology and implications for measuring and monitoring forest carbon stocks and change. *Carbon Manag.* **2011**, *2*, 231–244.
21. Petrokofsky, G.; Kanamaru, H.; Achard, F.; Goetz, S.J.; Joosten, H.; Holmgren, P.; Lehtonen, A.; Menton, M.C.S.; Pullin, A.S.; Wattenbach, M. Comparison of methods for measuring and assessing carbon stocks and carbon stock changes in terrestrial carbon pools. How do the accuracy and precision of current methods compare? A systematic review protocol. *Environ. Evid.* **2012**, *1*, 6.
22. Wulder, M.A.; White, J.C.; Nelson, R.F.; Næsset, E.; Ørka, H.O.; Coops, N.C.; Hilker, T.; Bater, C.W.; Gobakken, T. Lidar sampling for large-area forest characterization: A review. *Remote Sens. Environ.* **2012**, *121*, 196–209.
23. Zolkos, S.G.; Goetz, S.J.; Dubayah, R. A meta-analysis of terrestrial aboveground biomass estimation using lidar remote sensing. *Remote Sens. Environ.* **2013**, *128*, 289–298.
24. Reinhardt, E.D.; Keane, R.E.; Calkin, D.E.; Cohen, J.D. Objectives and considerations for wildland fuel treatment in forested ecosystems of the interior western United States. *For. Ecol. Manag.* **2008**, *256*, 1997–2006.
25. Gaines, W.L.; Harrod, R.J.; Dickinson, J.; Lyons, A.L.; Halupka, K. Integration of Northern spotted owl habitat and fuels treatments in the eastern Cascades, Washington, USA. *For. Ecol. Manag.* **2010**, *260*, 2045–2052.
26. Roccaforte, J.P.; Fule, P.Z.; Covington, W.W. Monitoring Landscape-Scale Ponderosa Pine Restoration Treatment Implementation and Effectiveness. *Restor. Ecol.* **2010**, *18*, 820–833.
27. Schoennagel, T.; Nelson, C.R. Restoration relevance of recent National Fire Plan treatments in forests of the western United States. *Front. Ecol. Environ.* **2010**, *9*, 271–277.
28. Schultz, C.A.; Coelho, D.L.; Beam, R.D. Design and governance of multiparty monitoring under the USDA Forest Service’s Collaborative Forest Landscape Restoration Program. *J. For.* **2014**, *112*, 198–206.
29. Ringold, P.L.; Alegria, J.; Czaplowski, R.L.; Mulder, B.S.; Tolle, T.; Burnett, K. Adaptive monitoring design for ecosystem management. *Ecol. Appl.* **1996**, *6*, 745–747.
30. Stankey, G.H.; Bormann, B.T.; Ryan, C.; Shindler, B.; Sturtevant, V.; Clark, R.N.; Philpot, C. Adaptive management and the Northwest Forest Plan: Rhetoric and reality. *J. For.* **2003**, *101*, 40–46.
31. Stem, C.; Margoluis, R.; Salafsky, N.; Brown, M. Monitoring and evaluation in conservation: A review of trends and approaches. *Conserv. Biol.* **2005**, *19*, 295–309.
32. Larson, A.J.; Belote, R.T.; Williamson, M.A.; Aplet, G.H. Making monitoring count: Project design for active adaptive management. *J. For.* **2013**, *111*, 348–356.

33. Folke, C.; Carpenter, S.; Elmqvist, T.; Gunderson, L.; Holling, C.S.; Walker, B. Resilience and sustainable development: Building adaptive capacity in a world of transformations. *AMBIO J. Hum. Environ.* **2002**, *31*, 437–440.
34. Hobbs, R.J.; Arico, S.; Aronson, J.; Baron, J.S.; Bridgewater, P.; Cramer, V.A.; Epstein, P.R.; Ewel, J.J.; Klink, C.S.; Lugo, A.E.; et al. Novel ecosystems: Theoretical and management aspects of the new ecological world order. *Glob. Ecol. Biogeogr.* **2006**, *15*, 1–7.
35. Hobbs, R.J.; Higgs, E.; Harris, J.A. Novel ecosystems: Implications for conservation and restoration. *Trends Ecol. Evol.* **2009**, *24*, 599–605.
36. Roccaforte, J.P.; Huffman, D.W.; Fulé, P.Z.; Covington, W.W.; Chancellor, W.W.; Stoddard, M.T.; Crouse, J.E. Forest structure and fuels dynamics following ponderosa pine restoration treatments, White Mountains, Arizona, USA. *For. Ecol. Manag.* **2015**, *337*, 174–185.
37. Sherrill, K.R.; Lefsky, M.A.; Bradford, J.B.; Ryan, M.G. Forest structure estimation and pattern exploration from discrete-return lidar in subalpine forests of the central Rockies. *Can. J. For. Res.* **2008**, *38*, 2081–2096.
38. Kim, Y.; Yang, Z.; Cohen, W.; Pflugmacher, D.; Lauver, C.; Vankat, J. Distinguishing between live and dead standing tree biomass on the North Rim of Grand Canyon National Park, USA using small-footprint lidar data. *Remote Sens. Environ.* **2009**, *113*, 2499–2510.
39. Hall, S.A.; Burke, I.C.; Box, D.O.; Kaufmann, M.R.; Stoker, J.M. Estimating stand structure using discrete-return lidar: An example from low density, fire prone ponderosa pine forests. *For. Ecol. Manag.* **2005**, *208*, 189–209.
40. Næsset, E. Practical large-scale forest stand inventory using small-footprint airborne scanning laser. *Scand. J. For. Res.* **2004**, *19*, 164–179.
41. Nelson, R.; Short, A.; Valenti, M. Measuring biomass and carbon in Delaware using an airborne profiling LIDAR. *Scand. J. For. Res.* **2004**, *19*, 500–511.
42. Næsset, E.; Gobakken, T. Estimation of above-and below-ground biomass across regions of the boreal forest zone using airborne laser. *Remote Sens. Environ.* **2008**, *112*, 3079–3090.
43. Asner, G.P.; Mascaró, J.; Muller-Landau, H.C.; Vielledent, G.; Vaudry, R.; Rasamoelina, M.; Hall, J.S.; van Breugel, M. A universal airborne LiDAR approach for tropical forest carbon mapping. *Oecologia* **2012**, *168*, 1147–1160.
44. Lefsky, M.A.; Hudak, A.T.; Cohen, W.B.; Acker, S.A. Geographic variability in lidar predictions of forest stand structure in the Pacific Northwest. *Remote Sens. Environ.* **2005**, *95*, 532–548.
45. Næsset, E. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sens. Environ.* **2002**, *80*, 88–99.
46. Li, Y.; Andersen, H.; McGaughey, R. A comparison of statistical methods for estimating forest biomass from light detection and ranging data. *West. J. Appl. For.* **2008**, *23*, 223–231.
47. Sarrazin, M.J.D.; Van Aardt, J.A.N.; Asner, G.P.; McGlinchy, J.; Messinger, D.W.; Wu, J. Fusing small-footprint waveform LiDAR and hyperspectral data for canopy-level species classification and herbaceous biomass modeling in savanna ecosystems. *Can. J. Remote Sens.* **2012**, *37*, 653–665.
48. Ediriweera, S.; Pathirana, S.; Danaher, T.; Nichols, D. Estimating above-ground biomass by fusion of LiDAR and multispectral data in subtropical woody plant communities in topographically complex terrain in North-eastern Australia. *J. For. Res.* **2014**, *25*, 761–771.
49. Laurin, G.; Chen, Q.; Lindsell, J.; Coomes, D.; Del Frate, F.; Guerriero, L.; Pirotti, F.; Valentini, R. Above ground biomass estimation in an African tropical forest with lidar and hyperspectral data. *ISPRS J. Photogramm. Remote Sens.* **2014**, *89*, 49–58.
50. Strunk, J.; Temesgen, H.; Andersen, H.E.; Packalen, P. Prediction of forest attributes with field plots, Landsat, and a sample of lidar strips. *Photogramm. Eng. Remote Sens.* **2014**, *80*, 143–150.
51. Raftery, A.E. Bayesian model selection in social research. *Sociol. Methodol.* **1995**, *25*, 111–163.
52. Hoeting, J.A.; Madigan, D.; Raftery, A.E.; Volinsky, C.T. Bayesian model averaging: A tutorial. *Stat. Sci.* **1999**, *4*, 382–417.
53. Hoeting, J.A. Methodology for Bayesian model averaging: An update. In Proceedings of the Manuscripts of Invited Paper Presentations, International Biometric Conference, Freiburg, Germany, 21–26 July 2002.
54. Andersen, H.E.; McGaughey, R.J.; Reutebuch, S.E. Estimating forest canopy fuel parameters using LIDAR data. *Remote Sens. Environ.* **2005**, *94*, 441–449.

55. Erdody, T.L.; Moskal, L.M. Fusion of LiDAR and imagery for estimating forest canopy fuels. *Remote Sens. Environ.* **2010**, *114*, 725–737.
56. Hermosilla, T.; Ruiz, L.A.; Kazakova, A.N.; Coops, N.C.; Moskal, L.M. Estimation of forest structure and canopy fuel parameters from small-footprint full-waveform LiDAR data. *Int. J. Wildland Fire* **2014**, *23*, 224–233.
57. González-Ferreiro, E.; Diéguez-Aranda, U.; Crecente-Campo, F.; Barreiro-Fernández, L.; Miranda, D.; Castedo-Dorado, F. Modelling canopy fuel variables for *Pinus radiata* D. Don in NW Spain with low-density LiDAR data. *Int. J. Wildland Fire* **2014**, *23*, 350–362.
58. Fahey, T.J.; Woodbury, P.B.; Battles, J.J.; Goodale, C.L.; Hamburg, S.P.; Ollinger, S.V.; Woodall, C.W. Forest carbon storage: Ecology, management, and policy. *Front. Ecol. Environ.* **2010**, *8*, 245–252.
59. White, M.A.; Vankat, J.L. Middle and high elevation coniferous forest communities of the north rim region of Grand Canyon National Park, Arizona, USA. *Vegetatio* **1993**, *109*, 161–174.
60. Nichol, A.A. *The Natural Vegetation of Arizona*; College of Agriculture, University of Arizona: Tucson, AZ, USA, 1937.
61. Romme, W.H.; Floyd, M.L.; Hanna, D. *Historical Range of Variability and Current Landscape Condition Analysis: South Central Highlands Section, Southwestern Colorado and Northwestern New Mexico*; Colorado Forest Restoration Institute at Colorado State University, and Region 2 of the U.S. Forest Service: Fort Collins, CO, USA, 2009.
62. Rasmussen, D.I. Biotic communities of Kaibab Plateau, Arizona. *Ecol. Monogr.* **1941**, *11*, 229–275.
63. Weng, C.; Jackson, S.T. Late glacial and Holocene vegetation history and paleoclimate of the Kaibab Plateau, Arizona. *Palaeogeogr. Palaeoclimatol. Palaeoecol.* **1999**, *153*, 179–201.
64. Wahlberg, M.M.; Triepke, F.J.; Robbie, W.A.; Strenger, S.H.; Vandendriesche, D.; Muldavin, E.H.; Malusa, J.R. *Ecological Response Units of the Southwestern United States*; USDA, Forest Service Forestry Report FR-R3-XX-XX; Southwestern Region, Regional Office: Albuquerque, NM, USA, 2013; p. 201.
65. Frazer, G.W.; Magnussen, S.; Wulder, M.A.; Niemann, K.O. Simulated impact of sample plot size and co-registration error on the accuracy and uncertainty of LiDAR-derived estimates of forest stand biomass. *Remote Sens. Environ.* **2011**, *115*, 636–649.
66. Lim, K.S.; Treitz, P.M. Estimation of Above ground Forest Biomass from Airborne Discrete Return Laser Scanner Data Using Canopy-based Quantile Estimators. *Scand. J. For. Res.* **2004**, *19*, 558–570.
67. Nilsson, M. Estimation of tree heights and stand volume using an airborne lidar system. *Remote Sens. Environ.* **1996**, *56*, 1–7.
68. Hosking, J.R.M. L-Moments: Analysis and Estimation of Distributions Using Linear Combinations of Order Statistics. *J. R. Stat. Soc. Ser. B* **1990**, *52*, 105–124.
69. Næsset, E. Determination of mean tree height of forest stands using airborne laser scanner data. *ISPRS J. Photogramm. Remote Sens.* **1997**, *52*, 49–56.
70. Nelson, R.; Oderwald, R.; Gregoire, G. Separating the ground and airborne laser sampling phases to estimate tropical forest basal area, volume, and biomass. *Rem. Sens. Environ.* **1997**, *60*, 311–326
71. Hawbaker, T.; Keuler, N.; Lesak, A.; Gobakken, T.; Contrucci, K.; Radeloff, V. Improved estimates of forest vegetation structure and biomass with a LiDAR-optimized sampling design. *J. Geophys. Res. Biogeosci.* **2009**, *114*, doi:10.1029/2008JG000870.
72. Junttila, V.; Finley, A.O.; Bradford, J.B.; Kauranne, T. Strategies for minimizing sample size for use in airborne LiDAR-based forest inventory. *For. Ecol. Manag.* **2013**, *292*, 75–85.
73. Dyess, J.; Youtz, J.; Nicolet, T. *Vegetation Monitoring and Sampling Protocols in Brief*; US Department of Agriculture, Forest Service, Region 3: Albuquerque, NM, USA, 2011.
74. McGaughey, R.J. *FUSION/LDV: Software for LIDAR Data Analysis and Visualization, Version 3.60+*; US Department of Agriculture, Forest Service, Pacific Northwest Research Station: Seattle, WA, USA, 2016; p. 123.
75. Clyde, M.A.; Ghosh, J.; Littman, M.L. Bayesian adaptive sampling for variable selection and model averaging. *J. Comput. Graph. Stat.* **2011**, *20*, 80–101.
76. Farr, T.G.; Rosen, P.A.; Caro, E.; Crippen, R.; Duren, R.; Hensley, S.; Kobrick, M.; Paller, M.; Rodriguez, E.; Roth, L.; et al. The Shuttle Radar Topography Mission. *Rev. Geophys.* **2007**, *45*, 2.
77. Gorelick, N.; Hancher, M.; Dixon, M.; Ilyushchenko, S.; Thau, D.; Moore, R. Google Earth Engine: Planetary-scale geospatial analysis for everyone. *Remote Sens. Environ.* **2017**, *202*, 18–27.

78. Shumway, R.H.; Stoffer, D.S. *Time Series Analysis and Its Applications: With R Examples*; Springer Science and Business Media: New York, NY, USA, 2006; p. 183.
79. Keyser, C.E. *Southern Variant Overview—Forest Vegetation Simulator*; Internal Rep. US Department of Agriculture, Forest Service, Forest Management Service Center: Fort Collins, CO, USA, 2008; p. 80.
80. Dixon, G.E. *Essential FVS: A User's Guide to the Forest Vegetation Simulator*; Internal Rep. US Department of Agriculture, Forest Service, Forest Management Service Center: Fort Collins, CO, USA, 2002; p. 226.
81. Lumley, T. Analysis of complex survey samples. *J. Stat. Softw.* **2004**, *9*, 1–19.
82. Lumley, T. *Survey: Analysis of Complex Survey Samples; R Package Version 3.32*; John Wiley & Sons, Inc: Hoboken, NJ, USA, 2017.
83. Graham, M.H. Confronting multicollinearity in ecological multiple regression. *Ecology* **2003**, *84*, 2809–2815.
84. Hawkins, D.M. The problem of overfitting. *J. Chem. Inf. Comput. Sci.* **2004**, *44*, 1–12.
85. Steel, M.F.J. Bayesian Model Averaging. Wiley StatsRef: Statistics Reference Online, 1–7. 2016. Available online: <http://doi.wiley.com/10.1002/9781118445112.stat07874> (accessed on 10 December 2017).
86. Clyde, M. *BAS: Bayesian Adaptive Sampling for Bayesian Model Averaging; R Package Version 1.4.7*; R Foundation for Statistical Computing: Vienna, Austria, 2017.
87. Zeugner, S. Bayesian Model Averaging with BMS for BMS Version 0.3.0. 2011. Available online: [www.bms.zeugnereu](http://www.bms.zeugnereu) (accessed on 15 July 2017).
88. Liang, F.; Paulo, R.; Molina, G.; Clyde, M.A.; Berger, J.O. Mixtures of g priors for Bayesian Variable Selection. *J. Am. Stat. Assoc.* **2008**, *103*, 410–423.
89. Feldkircher, M.; Zeugner, S. *Benchmark Priors Revisited: On Adaptive Shrinkage and the Supermodel Effect in Bayesian Model Averaging*; IMF Working Paper, WP/09/202; International Monetary Fund: Washington, DC, USA, 2009.
90. Fox, J.; Monette, G. Generalized collinearity diagnostics. *J. Am. Stat. Assoc.* **1992**, *87*, 178–183.
91. Sileshi, G.W. A critical review of forest biomass estimation models, common mistakes and corrective measures. *For. Ecol. Manag.* **2014**, *329*, 237–254.
92. McCune, B.; Grace, J.B. *Analysis of Ecological Communities*; Mjrn Software Design: Gleneden Beach, OR, USA, 2002; p. 304.
93. Weisberg, S. *Applied Linear Regression*, 3rd ed.; John Wiley & Sons: Hoboken, NJ, USA, 2005; p. 299.
94. Piñeiro, G.; Perelman, S.; Guerschman, J.P.; Paruelo, J.M. How to evaluate models: Observed vs. predicted or predicted vs. observed? *Ecol. Model.* **2008**, *216*, 316–322.
95. Van Breugel, M.; Ransijn, J.; Craven, D.; Bongers, F.; Hall, J.S. Estimating carbon stock in secondary forests: Decisions and uncertainties associated with allometric biomass models. *For. Ecol. Manag.* **2011**, *262*, 1648–1657.
96. Jolicoeur, P. Bivariate allometry: Interval estimation of the slopes of the ordinary and standardized normal major axes and structural relationship. *J. Theor. Biol.* **1990**, *144*, 275–285.
97. Legendre, P.; Legendre, L. *Numerical Ecology*, 2nd English ed.; Elsevier Science BV: Amsterdam, The Netherlands, 1998.
98. Ludbrook, J. Linear regression analysis for comparing two measurers or methods of measurement: But which regression? *Clin. Exp. Pharmacol. Physiol.* **2010**, *37*, 692–699.
99. Legendre, P. *Model II Regression User'S Guide, R Edition*; R Foundation for Statistical Computing: Vienna, Austria, 1998.
100. Chambers, J.M.; Cleveland, W.S.; Kleiner, B.; Tukey, P.A. *Graphical Methods for Data Analysis*; Wadsworth & Brooks/Cole: London, UK, 1983.
101. West, G.B.; Brown, J.H.; Enquist, B.J. The fourth dimension of life: Fractal geometry and allometric scaling of organisms. *Science* **1999**, *284*, 1677–1679.
102. Johnson, K.D.; Birdsey, R.; Finley, A.O.; Swantaran, A.; Dubayah, R.; Wayson, C.; Riemann, R. Integrating forest inventory and analysis data into a LIDAR-based carbon monitoring system. *Carbon Balance Manag.* **2014**, *9*, 3.
103. Sheridan, R.D.; Popescu, S.C.; Gatzliolis, D.; Morgan, C.L.; Ku, N.W. Modeling forest aboveground biomass and volume using airborne LiDAR metrics and forest inventory and analysis data in the Pacific Northwest. *Remote Sens.* **2014**, *7*, 229–255.

104. Hayashi, R.; Kershaw, J.A.; Weiskittel, A. Evaluation of alternative methods for using LiDAR to predict aboveground biomass in mixed species and structurally complex forests in northeastern North America. *Math. Comput. For. Nat. Resour. Sci.* **2015**, *7*, 49.
105. Ahmed, R.; Siqueira, P.; Hensley, S.; Bergen, K. Uncertainty of forest biomass estimates in north temperate forests due to allometry: Implications for remote sensing. *Remote Sens.* **2013**, *5*, 3007–3036.
106. Gobakken, T.; Næsset, E. Assessing effects of positioning errors and sample plot size on biophysical stand properties derived from airborne laser scanner data. *Can. J. For. Res.* **2009**, *39*, 1036–1052.
107. Thomas, V.; Treitz, P.; McCaughey, J.H.; Morrison, I. Mapping stand-level forest biophysical variables for a mixedwood boreal forest using lidar: An examination of scanning density. *Can. J. For. Res.* **2006**, *36*, 34–47.
108. Næsset, E. Effects of different sensors, flying altitudes, and pulse repetition frequencies on forest canopy metrics and biophysical stand properties derived from small-footprint airborne laser data. *Remote Sens. Environ.* **2009**, *113*, 148–159.
109. Asner, G.P.; Mascaro, J. Mapping tropical forest carbon: Calibrating plot estimates to a simple LiDAR metric. *Remote Sens. Environ.* **2014**, *140*, 614–624.
110. Price, C.A.; Ogle, K.; White, E.P.; Weitz, J.S. Evaluating scaling models in biology using hierarchical Bayesian approaches. *Ecol. Lett.* **2009**, *12*, 641–651.
111. Tredennick, A.T.; Bentley, L.P.; Hanan, N.P. Allometric convergence in savanna trees and implications for the use of plant scaling models in variable ecosystems. *PLoS ONE* **2013**, *8*, e58241.
112. Valbuena, R.; Eerikäinen, K.; Packalen, P.; Maltamo, M. Gini coefficient predictions from airborne lidar remote sensing display the effect of management intensity on forest structure. *Ecol. Indic.* **2016**, *60*, 574–585.



© 2018 by the authors. Licensee MDPI, Basel, Switzerland. This article is an open access article distributed under the terms and conditions of the Creative Commons Attribution (CC BY) license (<http://creativecommons.org/licenses/by/4.0/>).