Peer-reviewed version available at Remote Sens. 2018, 10, 442; doi:10.3390/rs10030442

Article

Development of a Regional Lidar-derived Forest Inventory Model with Bayesian Model Averaging for use in Ponderosa Pine and Mixed Conifer Forests in Arizona and New Mexico, USA

Karis Tenneson [1*], Matthew S. Patterson [1,2], Thomas Mellin [3], Mark Nigrelli [4], Peter Joria [3], Brent Mitchell [1]

1 RedCastle Resources, Inc. Contractor to USDA Forest Service Geospatial Technology and Applications Center, Salt Lake City, UT 84119, USA; krtenneson@fs.fed.us (K.T.); brentmitchell@fs.fed.us (B.M.)
2 Interdisciplinary PhD Program in Urban Design and Planning, University of Washington, Seattle, WA 98195, USA; maspatte@uw.edu
3 USDA Forest Service, Region 3 Regional Office, Albuquerque, NM 87102, USA; tmellin@fs.fed.us (T.M.); pjoria@fs.fed.us (P.J.)
4 USDA Forest Service, Coconino National Forest, Flagstaff, AZ 86001, USA; mnigrelli02@fs.fed.us
* Correspondence: krtenneson@fs.fed.us; Tel.: +1-801-975-3768

Abstract: Historical forest management practices in the southwestern US have left forests prone to high intensity, stand-replacement fires. Effective management to reduce the cost and impact of forest-fire management and allow fires to burn freely without negative impact depends on detailed knowledge of stand composition, in particular, above-ground biomass (AGB). Lidar-based modeling techniques provide opportunities to reduce costs and increase ability of managers to monitor AGB and other forest metrics. Using Bayesian Model Averaging (BMA), we develop a regionally applicable lidar-based statistical model for Ponderosa pine and mixed conifer forest systems of the southwestern USA, using previously collected field data. The selected regional model includes a mid and low canopy height metric, a canopy cover, and height distribution term. It explains 72% 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 developed is structured in accordance with previously described models fit to local data, and performs equivalently to models designed for smaller scale application. Developing regional models for broad scale application provides a cost-effective, robust approach for managers to monitor and plan 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

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 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].

Landscape scale restoration efforts are being implemented to create conditions where natural fires can be left to burn without fear of escalation, thereby reducing fire suppression costs [1,14–16]. Over 280,000 km^2 (70 million acres) of these forests are in need of restoration [1,17]. Government agencies, civil society organizations, and regional and local stakeholder groups have been collaborating to develop restoration strategies, identify priority areas for treatment, and implement activities to reduce rising costs and threats of extreme fires to communities and landscapes [18–20]. 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 the urgent problem these communities are facing [21]. CFLRP offers competitive awards to communities that are implementing large-scale, collaborative, cross jurisdictional restoration plans. Three selected projects cover Ponderosa pine and mixed conifer southwest US forests. Awarded projects are required to monitor social, ecological, and economic outcomes for at least 15 years after implementation begins [22]. The monitoring process aids in understanding treatment performance and the identification of negative unintended consequences of treatments; which then informs future decisions in an adaptive planning cycle [23–26].

The need for ongoing monitoring is especially important given the emergence of novel conditions resulting from the interactions of climate change and land use [27–29]. Restoration treatments are often guided by historical reference conditions and the natural range of variability [30]; 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 [27].

Repeating extensive ground-based forest inventories is time consuming, labor intensive, and expensive. 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 [21,31]. 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 [32]. The advantages of lidar-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 [33–39]).

CFLRP projects have invested in lidar acquisitions and collection of field data necessary to train the models that estimate forest structure. On-going field-based inventories are expensive, and project managers have expressed the need for a framework that allows them to use statistical models developed from previous lidar acquisitions and coincident field data collection efforts to update forest inventories on newly acquired lidar data. When new lidar acquisitions become available, the ability to use relationships developed from existing data to update lidar-based forest inventories can result in substantial savings from reduced field data collection efforts.

The goal of this work is to develop a regional biomass estimation model using lidar metrics from field data [40] 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 above-ground biomass 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 [41–46]. Above ground biomass (AGB) was estimated at plots from ground-based data using regional allometric equations.

Lidar-derived height, canopy density and volume metrics are combined with environmental data and regressed against the field-based estimates. The relationship between stand characteristics and lidar metrics vary between tree species, especially those with different crown shapes [47–49]. Therefore we test a combination of information from optical remote sensing and lidar to represent potential differences in plots with deciduous trees. We estimate the magnitude of seasonal variation of greenness (NDVI) from a Landsat time-series analysis. Environmental data includes topographic information and ecological response units [50]. We integrated data from seven field data collection efforts; five were used for model development and the remaining two were for an independent validation. We evaluate the reliability of the model using independent validation data from all sites and test the transferability of the model on a new lidar acquisition and two coincident field data collection efforts. No standard approach has emerged in the literature to select which lidar metrics from the large (overwhelming) pool of candidates are best suited for estimating biomass. We use Bayesian model averaging methods to specify our model structure from the ensemble of candidates.

The lidar data is 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 a plot size determined by average stand stem density, discuss the implications of these inconsistencies on lidar-based AGB estimations, and make recommendations that attempts 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. Studies have demonstrated accurate biomass estimates can be predicted with lidar data [36,37,39]. 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 [51–54] 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 [55].

Biomass provides information about the 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. Few studies have assessed lidar-based inventories in Ponderosa pine and mixed conifer forests of the southwestern USA [41,56,57]. Those that have were built on small data sets covering a limited lidar footprint. Lidar-based regional models have been developed to estimate biomass in boreal, temperate deciduous, temperate coniferous, and tropical forests [49,58–61]. Regional models for southwestern US forests have not been explored. Finally, Sherrill and colleagues [56] have demonstrated success in separating biomass estimates between live and dead vegetation in these regions; an invaluable metric for assessing fire risk of these landscapes.

1.1. Model selection with Bayesian model averaging methods

Lidar data can be aggregated in a plethora of ways to represent forest attributes such as canopy height. Many of the metrics have utility in producing estimates of aboveground biomass, total wood volume, and other landscape measures that are important for management decisions [33,34,36–39]. These metrics can be grouped into three categories that have clear biological interpretation and have direct analogs to ecologically significant variables. 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 [42,62].

Various methods are available to select a parsimonious set of metrics to use as predictors [62]. One possible approach is to engage in dimensional reduction of the data such as a principle component analysis and canonical correlation analysis [41,49]. However, the resultant factors can be difficult or impossible to meaningfully interpret. A commonly used approach is stepwise regression on a large pool of lidar metrics, however this approach can result in over fit models and is sensitive to multicolinearity issues [63,64]. A different approach is to avoid choosing a single model and generate a distribution of possible models that represent the inherent uncertainty that arises when many possible predictors (with possibly conflicting interpretations) exist [65]. This process is referred to as Bayesian

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 January 2018 doi:10.20944/preprints201801.0275.v1

Peer-reviewed version available at Remote Sens. 2018, 10, 442; doi:10.3390/rs10030442

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 [66]. This approach reduces the possibility of researcher bias in variable selection than typical step-wise regression approaches [65].

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 [67]. 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 [66]. For a good introduction to BMA, see [65,67,68].

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 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.

2. Materials and Methods

2.1. Study Area

This study takes place in Ponderosa pine and frequent-fire (dry) mixed conifer forests in U.S. National Forests in Arizona and New Mexico, USA. Sites were selected based on lidar acquisitions covering fire-adapted, Ponderosa pine and mixed conifer forests within active landscape-scale restoration projects. These regions include the Four Forest Restoration Initiative (4FRI), the Southwest Jemez Mountains Collaborative Landscape Restoration Project (both part of the network of Collaborative Forest Landscape Restoration Programs), and the 2009 Kaibab Forest Health Focus (Figure 1). The Four Forest Restoration Initiative (4FRI) covers 10,000 km^2 of ponderosa pine forest in northern Arizona, including lands in the Coconino, Tonto, and Apache-Sitgreaves National Forests. The Southwest Jemez Mountain Landscape Restoration Project expands across 850 km^2 of ponderosa pine, mixed conifer, and pinyon-juniper woodlands in the upper and middle Jemez River watersheds. A portion of the Santa Fe National Forest lies within it. The Kaibab Plateau is in the northern section of the Kaibab National Forest and the Grand Canyon National Park-North Rim, Arizona.

Elevation of the study sites ranges from 1,700 to 3,100 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 [69]. 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*) [70]. 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 [70–72]. The forest matrix includes patches of aspen stands, meadows, and forest openings generated by disturbances such as fire, wind throw, and timber harvesting [70]. 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 [69,72,73].

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 [50]. 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.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 January 2018 doi:10.20944/preprints201801.0275.v1

Peer-reviewed version available at *Remote Sens.* **2018**, *10*, 442; doi:10.3390/rs10030442

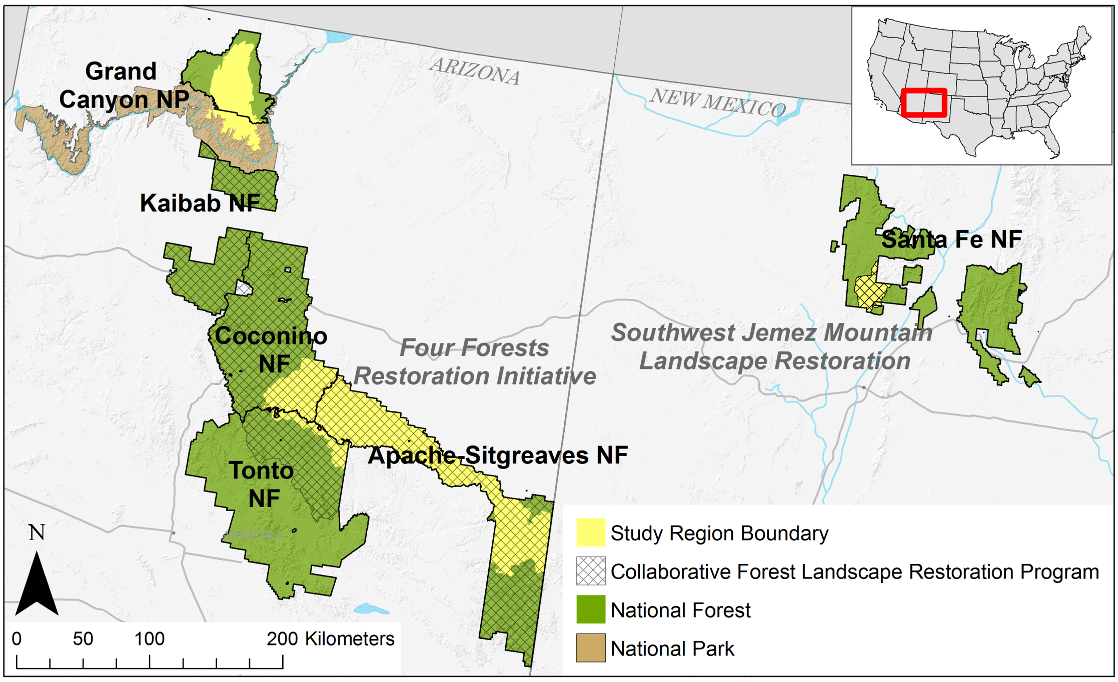

Figure 1. Location of study regions.

2.2. Data

2.2.1. Field Survey

We use information from seven different data collection efforts in Arizona and New Mexico. Data were collected at over 3,000 plots in seven data collection efforts between 2013 and 2015 (Figure 2 and Table 1) in the Kaibab Plateau, Four Forest Restoration Initiative, and the Southwest Jemez Mountains. Five field data collection efforts were completed in multiple stages within the 4FRI: the first effort was implemented on the western half in 2013 to 2014; the second stage of data collection began the following year in the eastern portion of the Apache-Sitgreaves National Forest. Data from the second stage were used for model validation. Currently, there is not complete sample that covers the full extent of Ponderosa pine and mixed conifer forests in southwest US forests to estimate AGB, but the combination of these seven collection efforts provide information across a range of forest conditions.

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) [74,75]. Plots were only placed in areas where the max canopy cover was greater than 3 meters. 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 [76]. 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.

The majority of plots were 0.04 and 0.02 ha in area. The plot size in the Coconino and Tonto NF varied based on average stand density. To decrease time and costs of data collection, in dense forest stands plot size was decreased from 0.04 ha [76]. The plot size in the Apache-Sitgreaves and Southwest Jemez Mountains projects was increased to 0.08 ha if there were fewer than 8 trees in a 0.04 ha plot.

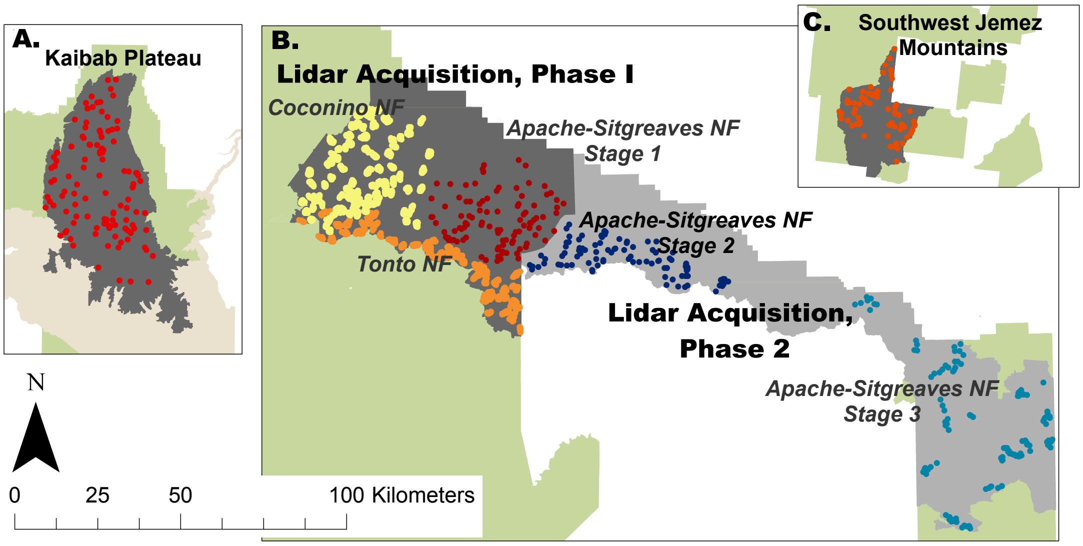

Figure 2. Map of the seven field data collection efforts and the footprint of the four lidar acquisitions, in grey, subset by region. Plots used for model development are displayed in warm tones (red, orange, yellow); those to assess transferability of model are blue. The three regions where data collection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

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.

 doi:10.20944/preprints201801.0275.v1

Peer-reviewed version available at *Remote Sens.* **2018**, *10*, 442; doi:10.3390/rs10030442

Table 1. Data collection summary for each project.

Study region	Area (km^2)	Year	No. & size (ha) of plots	Min DBH (cm)	Sample design	Strata	Use
Kaibab Plateau, AZ	1,382	2013-2014	112 (0.04)	20.3	stratified random	95th percentile height & percent canopy returns (>3 m)	model dev.
Coconino NF, 4FRI, AZ	75 (sampled area); 1,136 (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 NF, 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 NF, 4FRI, AZ, Stage 1	1,028	2014	15 (0.08), 85 (0.04)	12.7	stratified random	95th percentile height & percent canopy returns (>3 m)	model dev.
Southwest Jemez Mountains, 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 NF, 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 NF, 4FRI, AZ, Stage 3	1,700	2015-2016	25 (0.08), 71 (0.04)	12.7	stratified random	95th percentile height & percent canopy returns (>3 m)	model valid.

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. 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.

Table 2. Summary of the lidar specifications for each site.

Study region	Date	Area (km^2)	Instrument	Ave. pulse density ($pulses/m^2$)	Field of view (degrees)	Altitude (m)
Kaibab Plateau, AZ	2012	1,853	Leica ALS50 & ALS60	10.75	20-28	900-2000
Four Forest Restoration Initiative, AZ, Stage 1	2013	3,546	Leica ALS50 & ALS60	9.4	28	900
Four Forest Restoration Initiative, AZ, Stage 2	2014	4,365	Leica ALS70	15.4	28	1200-1400
Southwest Jemez Mountains, NM	2012	526	Leica ALS60	13.3	26	900

Canopy structure metrics were calculated from the raw lidar point cloud using FUSION software [77]. 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 [41,56]. Fractional canopy cover[1], metrics were calculated using this static cover threshold and the dynamic thresholds of mean and mode values [79]. We included canopy volume measures as a product of height quantiles and canopy density. 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 [67,80]. Most highly correlated variables existed as pairs with almost perfect correlation. The lidar metrics are listed in Table 3 (excluded metrics are located in Table 8 in Supplemental Section).

[1] also called canopy point density [42], laser intercept index [78]

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 1^{st}, 10^{th}, 30^{th}, 60^{th}, 90^{th} 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	3^{rd} and 4^{th} L-moments
	L-CV, L-skew., L-kurt.	L-moment coefficient of variation, skewness, and kurtosis
Canopy Cover & Density	CC	fractional canopy cover: number of all returns (> 3m) divided by total number of all returns
	$\text{Cov}_{>\text{mean height}}$:all	mean height cover: number of all returns above the mean divided by total number of all returns
	$\text{Cov}_{>3}$:1^{st}	number of all returns (> 3m) divided by total number of 1^{st} returns
	$\text{Cov}_{\text{all} > \text{mode}}$:all $_{\text{first}}$	number of all returns above the mode divided by total number of 1^{st} returns
Volume	P01*CC, P10*CC, P30*CC, P60*CC, P90*CC	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

2.2.3. Topography

In addition to field level forest structure measures and lidar data, we included two data sets derived from Earth Observing data. Elevation, slope, and aspect at each plot was determined from the Shuttle Radar Topography Mission (SRTM) digital elevation data with a resolution of 1 arc-second [81]. The topographic derivatives were calculated and sampled at each plot center in the Google Earth Engine platform [82].

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 [83]. The β's are the coefficients, f is the frequency, and t is time. Finally, to more easily interpret the cosine and sine coefficients (β_{cos} and β_{sin}) we convert these to amplitude, (β_A), and phase, (β_ϕ), (Equations 2 and 3) [83]. These calculations and sampling were performed in the Google Earth Engine platform [82].

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

$$\beta_A = \sqrt{\beta_{cos}^2 + \beta_{sin}^2} \tag{2}$$

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

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 January 2018 doi:10.20944/preprints201801.0275.v1

Peer-reviewed version available at *Remote Sens.* **2018**, *10*, 442; doi:10.3390/rs10030442

2.2.5. Ecological Response Unit

Ecological Response Units are a system of mapped ecosystem types [50]. 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. Analysis

2.3.1. Field AGB Estimates

Aboveground biomass (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 [84]. Aboveground biomass and total volume per plot were computed by summing the estimates for all trees within the plot. Analysis was performed using the Region 3 variant of the Forest Vegetation Simulator (FVS) [84,85]. 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 [86,87].

2.3.2. Lidar AGB Estimates

We tested the inclusion of three categories of lidar metrics: 1) variables representing the canopy height distribution, 2) the density of the canopy, and 3) the interaction between height and density of the canopy metrics (Table 3). 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. Upon inspection of linear fits between AGB and the height and canopy density metrics, we tested models built with a log transform of AGB and one with no transform.

Models were created using the Bayesian Adaptive Sampling (BAS) package [88] 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 was 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 is largely left to speak for itself. 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

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 January 2018 doi:10.20944/preprints201801.0275.v1

Peer-reviewed version available at Remote Sens. 2018, 10, 442; doi:10.3390/rs10030442

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 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 [89]. 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 $2 < a \leq 4$ [90]. 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 [91]. We set a to 3.

$$\frac{g}{1+g} \sim Beta(1, \frac{a}{2} - 1) \tag{4}$$

2.4. Accuracy Assessment

2.4.1. Model fit

We used BMA to fit a population of multiple regression linear and log-linear models using ordinary least squares to the data. These were specified with and without quadratic height terms, with and without volume interactions, and with and without a log transform of the response (AGB). The performance and goodness of fit of the highest probability (the single model with the highest probability of occurrence) and median models from the linear and log-linear BMA run was assessed. 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).

The median and highest probability models were fit using ordinary least squares regression. To evaluate the performance of each model we report the coefficient of determination estimates (R^2 and adjusted R^2).The root mean square error (RMSE), percent root mean square error, bias, and mean bias are reported for the median probability and highest probability linear models. Issues with multicolinearity and reliability of predictor estimates were assessed using percent relative standard error (PRSE) and variance inflation factors. Variance inflation factors greater than 5 suggest issues with multicolinearity [92], although Graham [63] 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 [93]. We used these thresholds to trim significant predictors from our models. We maintained the raw terms of significant interaction predictors even if these metrics indicate they are not significant [94]. Terms refer to the covariate variables, predictors encompass terms and combinations of terms.

The BMA object produced in R using BAS was also used to generate predictions, 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. The same error metrics were calculated for these predictions, and they were compared with the performance of the median and highest probability multiple regression models.

Error metrics were calculated on estimates of all three sets of data: the model training data, the 25% withheld from the training data, and the data from the two new lidar acquisitions. Recall that 25% of the data was withheld from the Kaibab Plateau, Coconino N.F., Tonto N.F., Apache-Sitgreaves N.F. Stage 1, and Southwest Jemez Mountains projects for model validation. We also report error and bias metrics for each project. We qualitatively assessed model fit from scatter plots of the observed versus predicted values and marginal plots for each model [95].

Root mean square error and bias provide information on fixed bias, a bias when values are higher (or lower) across the whole range of measurement. Other methods provide additional information on the potential for proportional bias, when estimates diverge progressively along the range of values.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 January 2018 doi:10.20944/preprints201801.0275.v1

Peer-reviewed version available at Remote Sens. 2018, 10, 442; doi:10.3390/rs10030442

We examined proportional bias with ordinary least squares and major axis (MA) regression analysis on the field and lidar based estimates. 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, it is questionable to use OLS to assess proportional bias, as errors exist in both the lidar (predictors) and field based estimates (observed). Field based estimates of AGB include uncertainty due to natural variability, measurement error, allometric model error, and model selection choices [92,96]. Therefore we also present results from major axis regression, which fits errors or natural variability on both variables symmetrically [97–99]. 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 [100].

2.4.2. Model Transferability

We evaluated the regional transferability of our model(s) by applying the final biomass model to independent observations from the Four Forest Restoration Initiative Phase 2 lidar acquisition. We report RMSE, percent RMSE, bias, and percent bias.

3. Results

3.1. Summary Statistics of Field Data Estimates

The average aboveground biomass of the sample of data used for model construction was 122.3 ($\pm$ 1.8) tons per hectare, and 114.6 ($\pm$ 2.9) in the subset of plots used to validate the model development (25% of the data). The composition of the plots from the data used for model construction was 72.8% Ponderosa pine forest, 25.5% mixed conifer, 0.5% spruce-fir forest, 0.5% pinyon-juniper Woodland, 0.4% herbaceousgrassland, and 0.3% deciduous (narrowleaf cottonwood and shrub, alder, and willow). The two additional data sets used to assess model transferability had average aboveground biomass of 71.1 ($\pm$ 5.5) and 89.5 ($\pm$ 5.7). Table 4 includes average biomass values for each data collection effort, including sample and population estimates. For most projects, the average biomass of the field plot samples at each project site is higher than the population average for the entire site (when the sample weights are taken into account). This reflects our sample strategy designed to represent the full range of forest conditions. Only a selection of stands were sampled in the Coconino and Tonto NF. A population estimate for the selected Coconino and Tonto NF 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 regions have been sampled within the southwestern mixed conifer forests. The projects were selected to represent the range of conditions present in the forests, but the sample frame does not cover the full spatial extent of these forests.

Table 4. Sample and population summary statistics from the field data for each project. Estimates are provided for both the model development and validation data subsets. Mean and standard error of AGB and elevation are reported.

	Model Construction Data			Validation Data		
Study region	**$AGB_{population}$ (Mg ha^{-1})**	**AGB_{sample} (Mg ha^{-1})**	**$Elev._{sample}$ (m)**	**$AGB_{population}$ (Mg ha^{-1})**	**AGB_{sample} (Mg ha^{-1})**	**$Elev._{sample}$ (m)**
All Model Dev. Sites	**-**	**122.3 ± 1.8**	**2090 ± 4**	**-**	**114.6 ± 2.9**	**2090 ± 7**
Kaibab Plateau, AZ	121.3 ± 7.2	132.2 ± 9.3	2502 ± 18	126.8 ± 18.8	139.7 ± 19.3	2510 ± 35
Coconino NF, 4FRI, AZ	-	128.5 ± 2.4	2160 ± 2	-	123.6 ± 4	2154 ± 4
Tonto NF, 4FRI, AZ	-	113.9 ± 2.8	1913 ± 5	-	101 ± 4.6	1903 ± 8
Apache-Sitgreaves NF, 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 Mountains, NM	109 ± 6.7	117.7 ± 9.8	2493 ± 23	109.6 ± 7	94.2 ± 13.1	2475 ± 36
Transferability Validation Sites						
Apache-Sitgreaves NF, 4FRI, AZ, Stage 2		-	-	57.2 ± 2.6	71.1 ± 5.5	2076 ± 10
Apache-Sitgreaves NF, 4FRI, AZ, Stage 3		-	-	85.2 ± 3.6	89.5 ± 5.7	2570 ± 13

3.2. AGB Estimation and Model Validation

We analyzed six models, the median probability and highest probability models from the BMA object and the BMA for two versions of the data (log-transformed and not) (Table 5). The error metrics of the estimates derived from the BMA ensemble were nearly identical to those of the median probability model. For each BMA model population, the median probability and the highest probability model were the same. The raw biomass model performed better than the model fit using log-transformed AGB. It explained 72% of the variation in the field based AGB estimates, had lower validation error values, and negligible bias. It is also the more parsimonious model.

Table 5. Model summary statistics of the estimates from the median probability model (MPM), the highest probability model (HPM), and from the Bayesian model average (BMA) object. RMSE, percent RMSE, bias and percent bias were all calculated on the data used to construct the statistical models.

Model	**Height Metrics**	**Canopy Cover and Density**	**Volume**	**Environ.**	**R^2**	**Adj. R^2**	**RMSE (Mg/ha)**	**RMSE%**	**Bias (Mg/ha)**	**Bias%**
MPM, ln(AGB)	P30 P60, QP60 P90 MAD Med.	CC	P90*CC	elevation slope NDVI Ampl	0.69	0.69	45	39	-1e-14	-1e-16
HPM, ln(AGB)				*same as MPM*						
BMA Object, ln(AGB)							47.19	40.92	6.97	0.06
MPM, ABG	P60, QP60 MAD Med. P30	$Cov_{>3}$:1st CC	P30*CC P60*CC		0.72	0.72	45	36.8	-1e-16	-8e-17
HPM, AGB				*same as MPM*						
BMA Object, AGB							44.92	36.74	-1e-13	-1e-17

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 January 2018 doi:10.20944/preprints201801.0275.v1

Peer-reviewed version available at Remote Sens. 2018, 10, 442; doi:10.3390/rs10030442

The median and highest probability raw biomass multiple regression model consisted of five terms and nine predictors (Table 5 and 6). These include the second order polynomial of the 60th percentile height, an indicator of the canopy height distribution (median of absolute deviation from the overall median), relationship between canopy cover returns and all first returns, and a lower and mid height canopy volume metric composed of the canopy cover with the 30th and 60th percentile heights. The lidar derivatives alone are used to estimate AGB; no information on topography, phenology, or ecological response units was included in the model.

The variance inflation factor of the two canopy cover and volume (product) predictors exceeded 10. 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 cover normalized by first returns ($Cov_{>3}$:1st) and 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 (R^2 was 0.71; adjusted R^2 was 0.71). Table 6 includes the full model specification. At least one predictor of all terms were significant at $p < 0.001$. The 60th percentile was not significant. It was included in the model because it is a term in the polynomial predictor, which is significant [94]. 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; with out the quadratic term they are under and over predicted, respectively. All subsequent analysis was conducted using this trimmed model (Table 6).

Table 6. Final prediction model for AGB and the correlation coefficient between AGB and the selected covariate. The significance of the relationship between each predictor and the response is indicated as follows: * is less than .05, ** is less than 0.01, and *** less than 0.001; others are less than 1.

	Full Model				Trimmed Model			
Predictors	**Coef.**	**Std. Error**	**Signif.**	**PRSE**	**Coef.**	**Std. Error**	**Signif.**	**PRSE**
Intercept	-33.62	16.1	*	47.89	-9.78	13.93		142.5
Canopy Height Metrics								
P30	-1.21	1.59		131.35	-4.66	1.19	***	25.63
P60	-837.93	351.15	*	41.9	-68.73	249.3		362.7
QP60	396.45	59.73	***	15.07	457.74	54.91	***	12
Canopy Height Distribution								
MAD Median	10.02	1.58	***	15.77	10.94	1.54	***	14.08
Canopy Cover and Density								
$Cov_{>3}$:1st	0.44	0.098	***	22.16	*removed due to variance inflation issues*			
CC	0.11	0.23		209.17	1.01	0.16	***	15.57
Canopy Volume								
P30*CC	0.15	0.033	***	33.19	0.24	0.015	***	6.47
P60*CC	0.083	0.028	**	22.54	*removed due to variance inflation issues*			

3.3. Model Performance by Site

The overall percent root mean squared error between the field observed AGB and the predicted AGB using the trimmed model was 35.23% for the validation data set withheld during model development. It was 31.18% and 32.83% for the two new validation data sets used to assess the efficacy of transferring the model to new lidar and field data acquisitions. The disagreement, expressed as percent RMSE, between predictions and field observed estimates from the Kaibab Plateau, Coconino NF and Tonto NF were slightly larger than at the other data collection sites (Table 7). These three projects were the only data collection efforts with AGB field estimates above 400 Mg per hectare (Fig. 3).

Peer-reviewed version available at *Remote Sens.* **2018**, *10*, 442; doi:10.3390/rs10030442

Bias was negligible overall for the model development sites; the model development construction and validation data had a percent bias below a tenth of a percent; the fully independent data from the Phase 2 lidar aquisition and stage 2 and 3 field data collection efforts on the eastern half of the Apache-Sitgreaves NF has a slightly negative bias of 4.69% and 10.89% (Table 7). These slight negative biases are occurring on sites that have moderate AGB estimates; none of the plots have field estimated AGB above 400 Mg/ha. However, at most sites the 95% confidence intervals of the OLS trend line between the field and predicted estimates includes the 1 to 1 line, indicating the bias estimates may not be significantly different (Fig. 3). The 95% confidence interval on the trendline for the Southwest Jemez Mountain project does, however, not fully enclose the 1:1 line (Fig. 3, f).

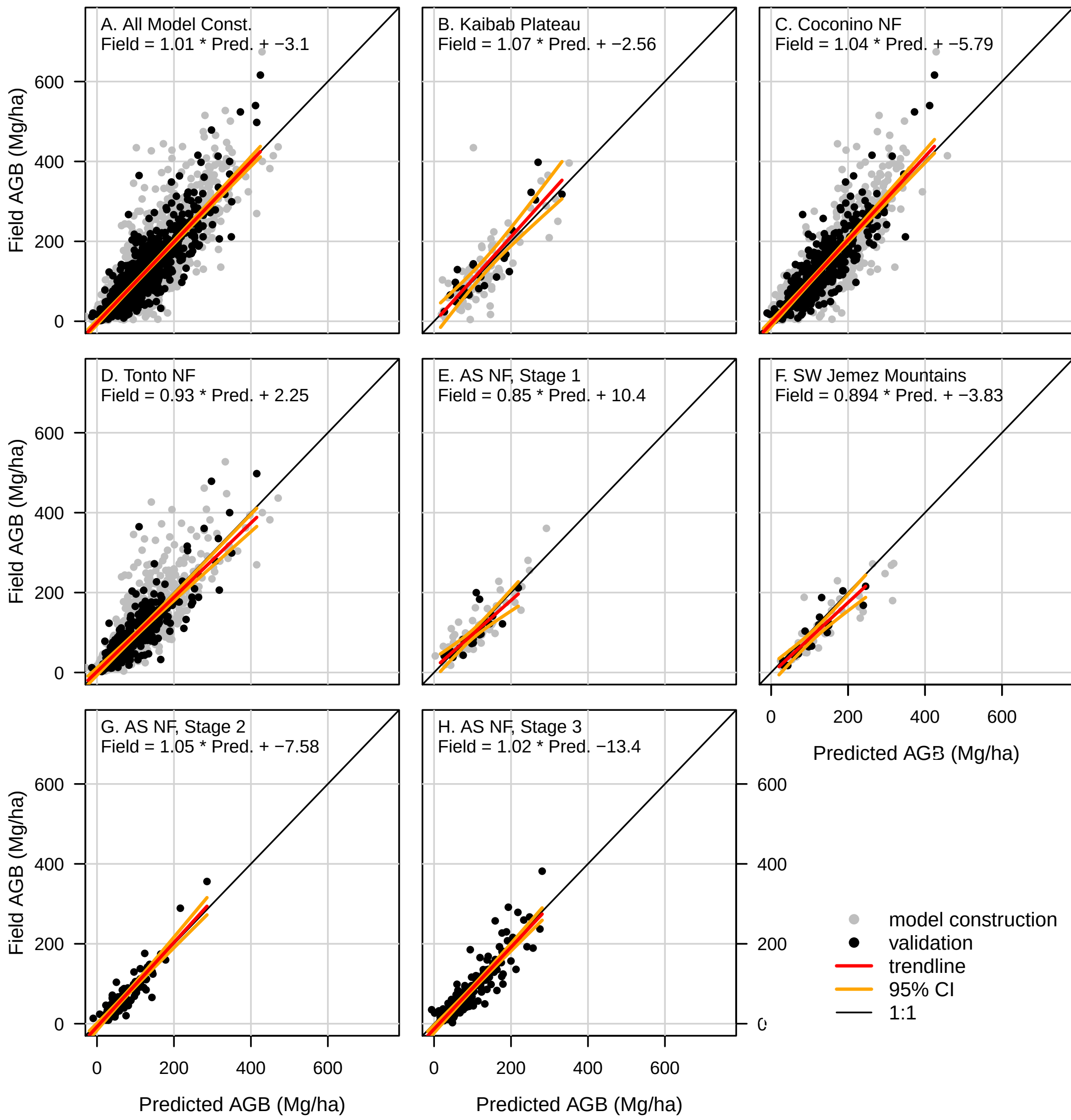

Figure 3. Scatter plot and ordinary least squares regression trend line of field measured aboveground biomass versus predicted values from the raw biomass regression model (Table 6). Plots include data from: (**a**) all data used for model construction (data in plots b-f), (**b**) Kaibab Plateau, (**c**) Coconino NF (4FRI), (**d**) Tonto NF (4FRI), (**e**) Apache-Sitgreaves NF (4FRI) Stage 1, (**f**) Southwest Jemez Mountains, (**g**) Apache-Sitgreaves NF (4FRI) Stage 2, and (**h**) Apache-Sitgreaves NF (4FRI) Stage 3. Red and orange lines are the linear fit and 95% confidence interval band of field measured aboveground biomass versus predicted values on the independent validation data subset. Black line is 1:1.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 January 2018 doi:10.20944/preprints201801.0275.v1

Peer-reviewed version available at Remote Sens. 2018, 10, 442; doi:10.3390/rs10030442

Table 7. 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	35.23	-2.26	-0.019
	Model	2271	45.29	37.04	-2e-13	-1e-15
Kaibab Plateau, AZ	Validation	25	43.7	32.85	6.7	5.03
	Model	87	55.19	41.43	-1.01	-0.76
Coconino NF, 4FRI, AZ	Validation	448	41.74	33.7	-0.23	-0.19
	Model	1218	43.84	33.99	-0.51	-0.4
Tonto NF, 4FRI, AZ	Validation	272	41.73	39.31	-5.2	-4.9
	Model	847	47.43	42.42	2.06	1.84
Apache-Sitgreaves NF, 4FRI, AZ, Phase 1	Validation	27	28.38	29.05	-4.2	-4.3
	Model	73	28.2	26.17	0.16	0.15
Southwest Jemez Mountains, NM	Validation	21	30.36	27.7	-15.41	-14.06
	Model	46	43.82	31.21	-22.66	-16.14
Transferability Validation Data						
Apache-Sitgreaves NF, 4FRI, AZ, Phase 2	Validation	96	23.25	31.18	-3.5	-4.69
Apache-Sitgreaves NF, 4FRI, AZ, Phase 3	Validation	150	32.82	32.66	-10.94	-10.89

3.4. Influence of Inconsistent Plot Size

This analysis is conducted with data from seven data collection efforts and the field protocols that determined plot size varied between projects (Table 1). The plots within the sample 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 NF plot size was decreased in dense stands; in the Southwest Jemez Mountains and three Apache-Sitgreaves NF 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 NF 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 through out the stand. 64% 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. 20% of the plots in the Coconino were reduced to 0.03 ha; 40% of the Tonto and 40% of the Coconino NF plots were reduced to 0.02 ha; and 10% of the Coconino and 16% of the Tonto plots were reduced to 0.01 ha or smaller. One stand in the Tonto had a tree density that resulted in 13 plots with a size of 0.008 ha. In the Southwest Jemez Mountains and three Apache-Sitgreaves NF data collection efforts, plot size was doubled from 0.04 ha to 0.08 if there were fewer than 8 trees (DBH greater than 12.7cm) 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.

The smallest and largest plot sizes tend to have low AGB values (Fig 4, a and Fig 5, a and f). The smallest plots (0.008 ha) were typically composed of tightly packed small trees. Conversely, the large plots included trees of varied size that were dispersed through out the plot. These plots suggest the model may under and over predict the AGB in low biomass forests at the edges of the range of high density with small trees and low density with mature trees. However, our moderate plot sizes presumably contain similar stand characteristics, so further work would need to be done to accurately assess the influence of plot size and performance of the model in stands with these characteristics.

The lidar-estimated AGB values for the smallest and largest plot sizes are not equal to zero (paired t-test p-value = 0.04 and <1e-5) (fig. 4, b). The stands with a high density of small DBH trees in the

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 January 2018 doi:10.20944/preprints201801.0275.v1

Peer-reviewed version available at Remote Sens. 2018, 10, 442; doi:10.3390/rs10030442

Tonto NF have plot estimates slightly higher than those reported from the field (average difference of 31.19 Mg/ha). While the reverse is true for the sparsely populated stands (0.08 ha plots), where estimates are on average 11.57 Mg/ha less than field estimates.

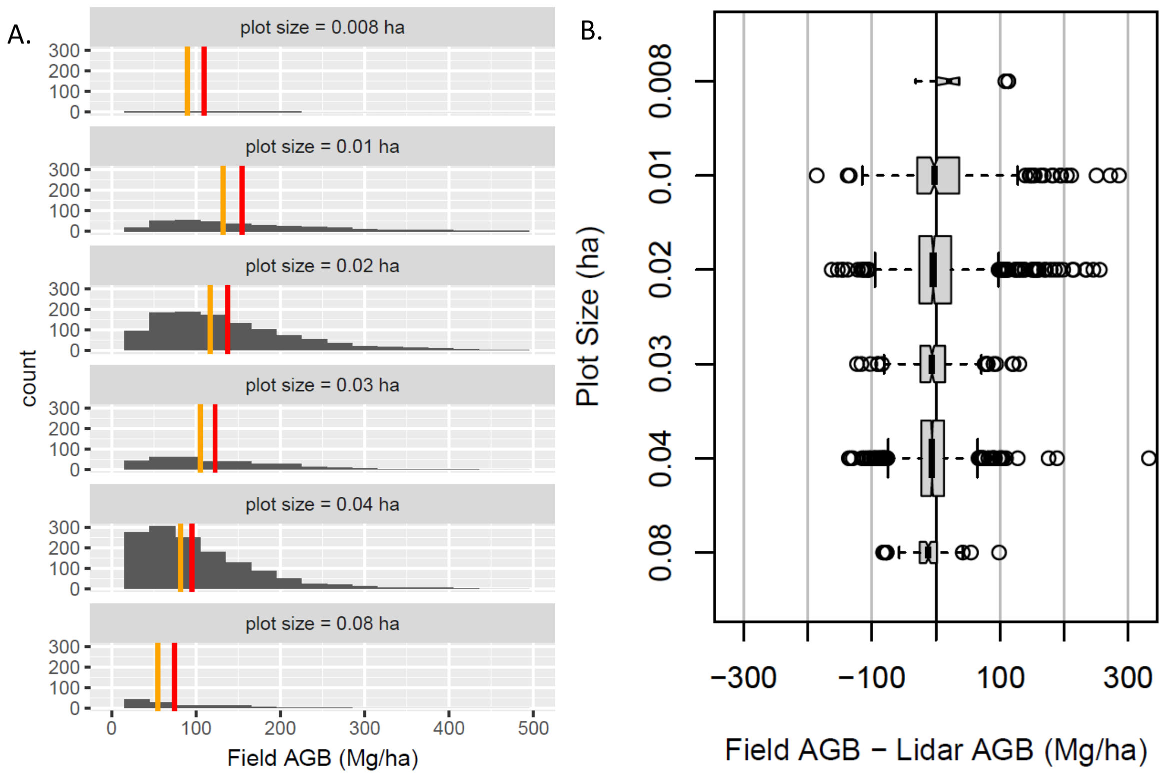

Figure 4. Distribution of AGB and mode disagreement by plot size: (**a.**) Histogram of field estimated AGB for each of the plot sizes. The red line is the group mean, orange is group median and (**b.**) Box plot of disagreement between field estimated and lidar estimated aboveground biomass for each 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 [101]

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 January 2018 doi:10.20944/preprints201801.0275.v1
Peer-reviewed version available at Remote Sens. 2018, 10, 442; doi:10.3390/rs10030442

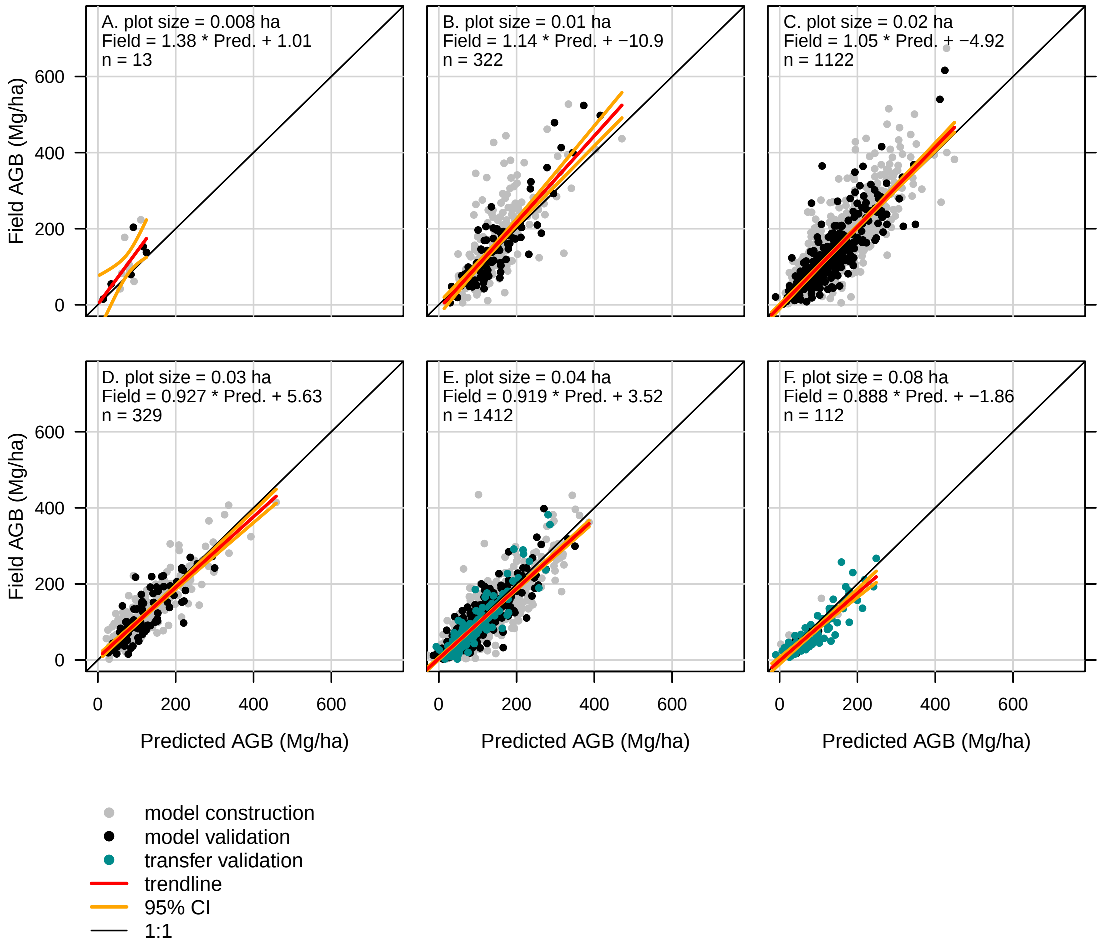

Figure 5. Scatter plot of field measured aboveground biomass versus predicted values from the raw biomass regression model (Table 6) for each of the plot sizes used in the study. Plot sizes for each window are: (**a**) 0.008 ha, (**b**) .01 ha, (**c**) 0.02 ha, (**d**) 0.03 ha, (**e**) 0.04 ha, and (**f**) 0.08 ha. Red and orange lines are the linear fit and 95% confidence interval band of field measured aboveground biomass versus predicted values on the independent validation data subset. Black line is 1:1.

Finally, the majority of plots with field based AGB estimates above 400 Mg/ha are in plots that are smaller than 0.04 ha, most are 0.01 and 0.02 hectare plots (Fig. 5, b and c). The high AGB plots (>400 Mg/ha) also exhibit under predicted lidar based estimates compared to the field estimates; with an average disagreement of 142.33 Mg/ha. A trend line fit using major axis regression indicates that there is proportional disagreement between the lidar and field estimates. The effects are most evident in these plots with high AGB (Fig. 6). This may be the result of edge effects, plot mis-registration, and other errors associated with smaller plot sizes [39,102–105]. These issues are discussed in more detail below.

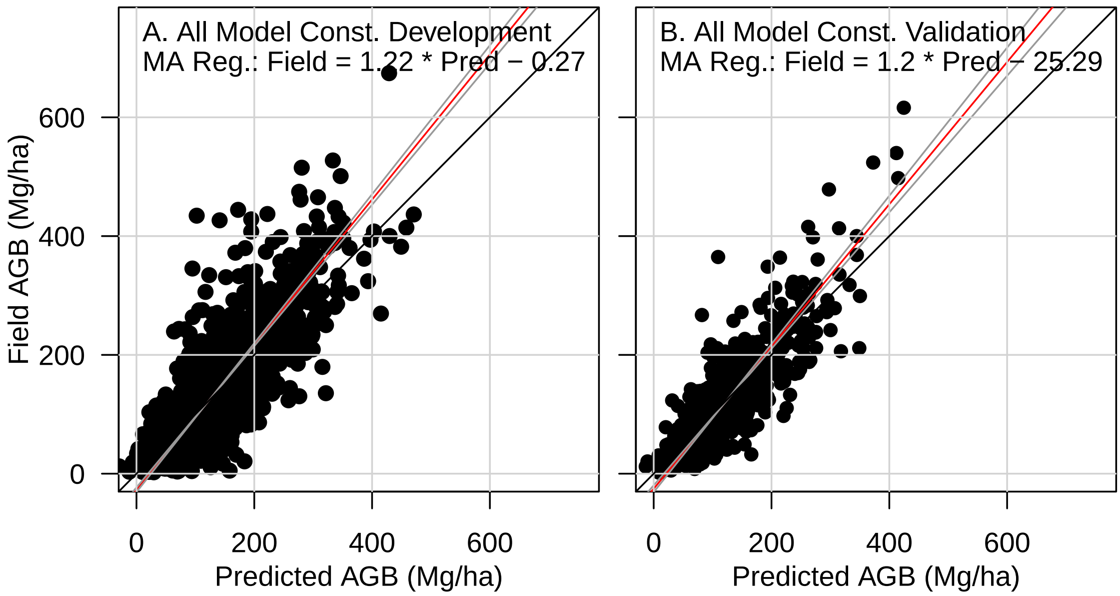

Figure 6. Scatter plot and major axis regression trend line of field measured aboveground biomass versus predicted values from the raw biomass regression model (Table 6). Plots include data from: (**a**) all model development data used during model construction, (**b**) all validation data used during model construction. Red dashed line is linear fit from the major axis regression of field measured aboveground biomass versus predicted values on the independent validation data subset. Grey lines are 95% confidence intervals from 1,000 permutations. Black line is 1:1.

4. Discussion

As expected, the BMA performed well for estimation of AGB, with marginally lower error values for predictions. 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 which can be related to ecological theory are typically preferred and easier to interpret or check for obvious disconnects with ground conditions. The BMA approach can be used to come up with a single multiple regression equation, by identifying the median or highest probability model. 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 [65].

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 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 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

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 January 2018 doi:10.20944/preprints201801.0275.v1

Peer-reviewed version available at Remote Sens. 2018, 10, 442; doi:10.3390/rs10030442

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 [92]. 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 [106]. The set of lidar metrics selected in our final model describe 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 [42]. This combined with the height metrics represents the vertical distribution and extent of canopy.

4.1. Relationship to other Modeling Efforts

Our model is consistent with other studies conducted in the same region [41,56,57]. Hall [57] 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 et al [41] 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' [56] proposed a lidar-based model fit to estimate live and dead aboveground biomass 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 above ground biomass 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 [41], a max below 400 but with only 4 plots above 150 Mg/ha [57], and a max less than 400 Mg/ha [56]. Sample sizes were small, ranging from 36 [41] to 58 to [56]. However, plot sizes were larger; 0.1 ha [56] and 0.32 ha [57].

The lidar covariates Kim [56] 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 et al [56] model did not include volume metrics on multiple height quartiles nor did they asses quadratic height quartiles. We found these to be valuable; the 30th percentile metric appeared in our model as a volume metric (P30*CC) 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 biomass 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 biomass plots (starting at about 250 Mg/ha) and over predicts low biomass plots, especially those with close to zero AGB [56].

4.2. 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., [42,56,107–109]). This disagreement can be partially explained by knowns 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 [108] remedied the issue by using a

Peer-reviewed version available at *Remote Sens.* **2018**, *10*, 442; doi:10.3390/rs10030442

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 choice of allometric model [92]. 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 field estimates. 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 values between the two groups.

Model errors decrease with increasing plot size [39,103,104]. 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)[39,104]. 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 [103,104]. Co-registration errors are reduced in larger plots due to the higher degree of spatial overlap. Gobakken and Naesset [103] 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.

44% 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.

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 biomass 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 biomass 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 biomass in excess of 400 Mg/ha, and across a range of low biomass 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 [107] 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 if lidar based models perform well at the interface of public forest and settlements, 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 collection information on smaller plots in high density stands reduces time and costs on 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

5. Conclusion

The task of identifying the best performing combination of lidar metrics for AGB estimation is a key challenge in the development of regional lidar-based AGB predictive models. No standard approach has been agreed on; approaches range from theory driven hypothesis test of a single lidar derivative to information criteria-based data mining. Studies using a priori candidate models built from allometric theory are not well suited to evaluate which of the suite of lidar metrics that represent a functional trait are the most appropriate (e.g., the forest height profile can be represented with a plethora of related, but distinct lidar metrics). Information theory approaches face issues with spurious relationships, confounding variables, and confirmation bias [64,110,111]. Stephens [112] makes the case for a combination of these methods. Model selection with BMA allows these issues to largely be circumvented through the full exploration of the model space, and assesses probability of both the inclusion of individual parameters in any model, and the probability of any given model[65,66]. Thus, we attempt to blend these two approaches using Bayesian model averaging, verifying our final model is supported by empirical findings and biomass estimation theory, and finally assessing the performance of the model on independent validation data sets. Our final model takes a functional form that aligns with theory and empirical observations on relating biomass to forest height and cover profiles.

Lidar based regional AGB models have been developed for boreal, temperate deciduous, temperate coniferous, and tropical forests [49,58–61]. 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 72 percent of the variation in the field based AGB estimates of the sample of plots. 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 [39]. 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 involved in the study suggest that the model is not overfit and suitable for generalization. The lidar data used in this analysis was 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 [113–115]. 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

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 January 2018 doi:10.20944/preprints201801.0275.v1

Peer-reviewed version available at *Remote Sens.* **2018**, *10*, 442; doi:10.3390/rs10030442

only be applied when the domain of a new lidar acquisition with similar specifications covers these forest types.

We present a cost effective approach to use previous data collection efforts to assist in updating lidar-derived forest inventories. While this approach still requires a 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. Further, as new validation data becomes available it 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 [116,117].

The focus of this research was on aboveground biomass, 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 [38]). Examples of forest characteristics of particular importance in these fire-prone forests include timber volume, canopy fuels [51,52,54], monitoring management intensity Valbuena et al 2016, and standing dead biomass [56]. 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/link, Table S1: Correlation statistics of excluded lidar metrics.

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 January 2018 doi:10.20944/preprints201801.0275.v1

Peer-reviewed version available at Remote Sens. 2018, 10, 442; doi:10.3390/rs10030442

Table 8. Correlation statistics of excluded lidar metrics; metrics that had a correlation of 0.94 or greater were not considered in the BMA.

Variable	Excluded Pair	Corr. Coef.
Height Metrics		
P10	P05	0.94
P30	P20	0.97
	P25	0.99
	P40	0.98
P60	mean height	0.99
	P50	0.99
	P70	0.99
	P75	0.98
P90	P80	0.98
	P95	0.99
	P99	0.97
Height Distribution Metrics		
SD	L2	0.98
	Average absolute deviation from the mean height	0.97
MAD Med.	Interquartile distance	0.94
LCV	Coefficient of variation	0.94
Canopy Cover		
CC	number of 1^{st} returns (> 3m) divided by total number of all returns	0.98
$Cov_{>mean\ height}$:all	number of all returns above the mean divided by total number of 1^{st} returns	0.99
	number of 1^{st} returns above the mean divided by total number of all returns	0.99
$Cov_{all\ >mode}$:all_{first}	number of all returns above the mode divided by total number of all returns	0.98
	number of 1^{st} returns above the mode divided by total number of all returns	0.99

Acknowledgments: This research was performed under contract to the US Forest Service, Geospatial Technology and Applications Center (GTAC) and in support of Region 3 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 three anonymous reviewers, they offered valuable insight and suggestions to help improve the quality of the manuscript.

Author Contributions: KT, MP, TM, MN, PJ, and BM conceived and designed the experiments; TM, MN, PJ, and BM processed field and lidar data; KT and MP 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
NF: National Forest
OLS: ordinary least squares regression
PRSE: percent relative standard error
RMSE: root mean square error
RMSPE: root mean square predicted error

Preprints (www.preprints.org) | NOT PEER-REVIEWED | Posted: 30 January 2018 doi:10.20944/preprints201801.0275.v1
Peer-reviewed version available at *Remote Sens.* **2018**, *10*, 442; doi:10.3390/rs10030442

MBE: mean bias error

1. Brown, R.T.; J.K. Agee; J.F. Franklin. 2004. 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. *Nature Communications* **2015**, *6*.
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. *Journal of Sustainable Forestry* **1994**, *2(1-2)*, 153-181.
4. Huffman, D.W.; Zegler, T.J.; Fule, P.Z. Fire history of a mixed conifer forest on the Mogollon Rim, northern Arizona, USA. *Int. J. of 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(2)*, 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. *Journal of Applied Ecology* **2016**, *53(6)*, 1755-1765.
8. Flannigan, M. D.; Stocks, B. J.; Wotton, B. M. Climate change and forest fires. *Science of the Total Environment* **2000**, *262(3)*, 221-229.
9. McKenzie, D.; Gedalof, Z. E.; Peterson, D. L.; Mote, P. Climatic change, wildfire, and conservation. *Conservation Biology* **2004**, *18(4)*, 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(5789)*, 940-943.
11. Liu, Y.; Stanturf, J.; Goodrick, S. Trends in global wildfire potential in a changing climate. *Forest Ecology and Management* **2010**, *259(4)*, 685-697.
12. Seidl, R.; Thom, D.; Kautz, M.; Martin-Benito, D.; Peltoniemi, M.; Vacchiano, G.; ... Lexer, M. J. Forest disturbances under climate change. *Nature Climate Change* **2017**, *7(6)*, 395-402.
13. Abatzoglou, J. T.; Williams, A. P. Impact of anthropogenic climate change on wildfire across western US forests. *Proceedings of the National Academy of Sciences* **2016**, *113(42)*, 11770-11775.
14. 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. Manage.* **2008**, *256*, 1997–2006.
15. 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. *Forest Ecology and Management* **2010**, *260(11)*, 2045-2052.
16. Roccaforte, J. P.; Fule, P. Z.; Covington, W. W. Monitoring Landscape-Scale Ponderosa Pine Restoration Treatment Implementation and Effectiveness. *Restoration Ecology* **2010**, *18(6)*, 820-833.
17. Schoennagel, T.; Nelson, C. R. Restoration relevance of recent National Fire Plan treatments in forests of the western United States. *Frontiers in Ecology and the Environment* **2010**, *9(5)*, 271-277.
18. Governor's Forest Health Councils State of Arizona. *The statewide Strategy for Restoring Arizona's Forests*; Aumack, E., T. Sisk, J. Palumbo Eds.; Arizona Public Service: Phoenix, AZ, **2007**; pp. 151.
19. Goldstein, B.E.; Butler, W.H. The U.S. Fire Learning Network: Providing a narrative framework for restoring ecosystems, professions, and institutions. *Soc. Nat. Resour.* **2010**, *23*, 935–951.
20. Western Governors' Association Forest Health Advisory Committee (WGA FHAC). *Forest health landscape-scale restoration recommendations*; WGA, Denver, CO, **2010**.
21. Schultz, C. A.; Jedd, T.; Beam, R. D. The Collaborative Forest Landscape Restoration Program: a history and overview of the first projects. *Journal of Forestry* **2012**, *110(7)*, 381-391.
22. 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. *Journal of Forestry* **2014**, *112(2)*, 198-206.

23. Ringold, P. L.; Alegria, J.; Czaplewski, R. L.; Mulder, B. S.; Tolle, T.; Burnett, K. Adaptive monitoring design for ecosystem management. *Ecological Applications* **1996**, *6(3)*, 745-747.
24. 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. *Journal of forestry* **2003**, *101(1)*, 40-46.
25. Stem, C.; Margoluis, R.; Salafsky, N.; Brown, M. Monitoring and evaluation in conservation: a review of trends and approaches. *Conservation Biology* **2005**, *19(2)*, 295-309.
26. Larson, A. J.; Belote, R. T.; Williamson, M. A.; Aplet, G. H. Making monitoring count: project design for active adaptive management. *Journal of Forestry* **2013**, *111(5)*, 348-356.
27. 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: A journal of the human environment* **2002**, *31(5)*, 437-440.
28. Hobbs, R. J.; Arico, S.; Aronson, J.; Baron, J. S.; Bridgewater, P.; Cramer, V. A.; ... Norton, D. Novel ecosystems: theoretical and management aspects of the new ecological world order. *Global Ecology and Biogeography* **2006**, *15(1)*, 1-7.
29. Hobbs, R. J.; Higgs, E.; Harris, J. A. Novel ecosystems: implications for conservation and restoration. *Trends in Ecology & Evolution* **2009**, *24(11)*, 599-605.
30. 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. *Forest Ecology and Management* **2015**, *337*, 174-185.
31. 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. *Proceedings of the National Academy of Sciences* **2009**, *106(26)*, 10706-10711.
32. 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. *Ecological Indicators* **2013**, *33*, 45-59.
33. Dubayah, R.; Drake, J.B. Lidar remote sensing for forestry. *Journal of Forestry* **2000**, *98*, 44-46.
34. 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. *AIBS Bulletin* **2002**, *52(1)*, 19-30.
35. 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. *International Journal of Remote Sensing* **2008**, *29(5)*, 1339-1366.
36. Goetz, S.J.; Dubayah. R.O. Advances in remote sensing technology and implications for measuring and monitoring forest carbon stocks and change *Carbon Management* **2011**, *2 (3)*, 231-244.
37. Petrokofsky, G.; Kanamaru, H.; Achard, F.; Goetz, S. J.; Joosten, H.; Holmgren, P.; ... 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. *Environmental Evidence* **2012**, *1(1)*, 6.
38. Wulder, M. A.; White, J. C.; Nelson, R. F.; Næsset, E.; Ørka, H. O.; Coops, N. C.; ... Gobakken, T. Lidar sampling for large-area forest characterization: A review. *Remote Sens. of Environ.* **2012**, *121*, 196-209.
39. Zolkos, S. G.; Goetz, S. J.; Dubayah, R. A meta-analysis of terrestrial aboveground biomass estimation using lidar remote sensing. *Remote Sens. of Environ.* **2013**, *128*, 289-298.
40. Næsset, E. Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sens. of Environ.* **2002**, *80(1)*, 88-99.
41. 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. *Canadian Journal of Forest Research* **2008**, *38(8)*, 2081-2096.
42. Li, Y.; Andersen, H; McGaughey, R. A comparison of statistical methods for estimating forest biomass from light detection and ranging data *Western Journal of Applied Forestry* **2008**, *23(4)*, 223-231.
43. 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. *Canadian Journal of Remote Sensing* **2012**, *37(6)*, 653-665.

44. 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. *Journal of Forestry Research* **2014**, *25(4)*, 761-771.
45. Laurin, G.; Chen, Q.; Lindsell, J.; Coomes, D.; Del Frate, F.; Guerriero, L.; ... Valentini, R. Above ground biomass estimation in an African tropical forest with lidar and hyperspectral data. *ISPRS Journal of Photogrammetry and Remote Sensing* **2014**, *89*, 49-58.
46. Strunk, J.; Temesgen, H.; Andersen, H.E.; Packalen, P. Prediction of forest attributes with field plots, Landsat, and a sample of lidar strips. *Photogrammetric Engineering & Remote Sensing* **2014**, *80(2)*, 143-150.
47. Næsset, E. Determination of mean tree height of forest stands using airborne laser scanner data. *ISPRS Journal of Photogrammetry and Remote Sensing* **1997**, *52*, 49-56.
48. 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. of Environ.* **1997**, *60*, 311-326
49. 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. of Environ.* **2005**, *95(4)*, 532-548.
50. 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, **2013**; pp 201.
51. Andersen, H. E.; McGaughey, R. J.; Reutebuch, S. E. Estimating forest canopy fuel parameters using LIDAR data. *Remote sensing of Environment* **2005**, *94(4)*, 441-449.
52. Erdody, T. L.; Moskal, L. M. Fusion of LiDAR and imagery for estimating forest canopy fuels. *Remote Sensing of Environment* **2010**, *114(4)*, 725-737.
53. 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. *International journal of Wildland Fire* **2014**, *23(2)*, 224-233.
54. 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. *International Journal of Wildland Fire* **2014**, *23(3)*, 350-362.
55. 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. *Frontiers in Ecology and the Environment* **2010**, *8(5)*, 245-252.
56. 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. of Environ.* **2009**, *113*, 2499-2510.
57. 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. *Forest Ecology and Management* **2005**, *208(1)*, 189-209.
58. Næsset, E. Practical large-scale forest stand inventory using small-footprint airborne scanning laser. *Scandinavian Journal of Forest Research* **2004**, *19*, 164–179.
59. Nelson, R.; Short, A.; Valenti, M. Measuring biomass and carbon in Delaware using an airborne profiling LIDAR. *Scandinavian Journal of Forest Research* **2004**, *19(6)*, 500–511.
60. Næsset, E.; Gobakken, T. Estimation of above-and below-ground biomass across regions of the boreal forest zone using airborne laser. *Remote Sensing of Environment* **2008**, *112(6)*, 3079-3090.
61. Asner, G. P.; Mascaro, 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**.
62. White, . *Vegetatio* **1993**, *109*, 161–174.
63. Graham, M. H. Confronting multicollinearity in ecological multiple regression. *Ecology* **2003**, *84(11)*, 2809-2815.
64. Hawkins, D. M. The problem of overfitting. *Journal of Chemical Information and Computer Sciences* **2004**, *44(1)*, 1-12.
65. Raftery, A.E. Bayesian model selection in social research. *Sociological Methodology* **1995**, *25*, 111-163.
66. Steel, M.F.J. *Bayesian model averaging.* **2016**, *Wiley StatsRef: Statistics Reference Online*, 1-7. http://doi.wiley.com/10.1002/9781118445112.stat07874
67. Hoeting, J.A.; Madigan, D; Raftery, A.E.; Volinsky, C.T. Bayesian model averaging: a tutorial. *Stat. Sci.* **1999**, *4*, 382–417.

68. Hoeting, J.A. Methodology for Bayesian model averaging: an update *Proceedings-Manuscripts of Invited Paper Presentations, International Biometric Conference* **2002**.
69. 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.
70. Nichol, A.A. *The natural vegetation of Arizona*. College of Agriculture, University of Arizona: Tucson, Arizona, USA. **1937**.
71. 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: CO, USA.**2009**.
72. Rasmussen, D.I. Biotic communities of Kaibab Plateau, Arizona. *Ecological Monographs* **1941**, *11*, 229-275.
73. Weng, C.; Jackson, S.T. Late glacial and Holocene vegetation history and paleoclimate of the Kaibab Plateau, Arizona. *Palaeogeography, Palaeoclimatology, Palaeoecology* **1999**, *153*, 179–201.
74. 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. *Journal of Geophysical Research: Biogeosciences* **2009**, *114(G2)*.
75. Junttila, V.; Finley, A. O.; Bradford, J. B.; Kauranne, T. Strategies for minimizing sample size for use in airborne LiDAR-based forest inventory. *Forest Ecology and Management* **2013**, *292*, 75-85.
76. Dyess, J.; Youtz, J.; Nicolet, T. *Vegetation Monitoring and Sampling Protocols in Brief*; US Department of Agriculture, Forest Service, Region 3, USA, **2011**.
77. McGaughey, R.J. /em 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**; pp. 123.
78. Luo, S.; Wang, C.; Xi, X.; Pan, F.; Peng, D.; Zou, J.; ... Qin, H. Fusion of airborne LiDAR data and hyperspectral imagery for aboveground and belowground forest biomass estimation. *Ecological Indicators* **2017**, *73*, 378-387.
79. Véga, C.; Renaud, J. P.; Durrieu, S.; Bouvier, M. On the interest of penetration depth, canopy area and volume metrics to improve Lidar-based models of forest parameters. *Remote Sensing of Environment 2016*, *175*, 32-42.
80. Clyde, M.A, M.A.; Ghosh, J.; Littman, M.L. Bayesian adaptive sampling for variable selection and model averaging *Journal of Computational and Graphical Statistics* **2011**, *20(1)*, 80-101.
81. Farr, T.G.; Rosen, P.A.; Caro, E.; Crippen, R.; Duren, R.; Hensley, S.; Kobrick, M.; Paller, M.; Rodriguez, E.; Roth, L.; Seal, D.; Shaffer, S.; Shimada, J.; Umland, J.; Werner, M.; Oskin, M.; Burbank, D.; Alsdorf, D. The Shuttle Radar Topography Mission. *Reviews of Geophysics* **2007**, *45:2*.
82. 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**, 18-27.
83. Shumway, R.H.; Stoffer, D.S. *Time series analysis and its applications: with R examples*; Springer Science and Business Media, **2006**; pp. 183.
84. 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**; pp. 80.
85. 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**; pp. 226.
86. Lumley, T. Analysis of complex survey samples. Analysis of complex survey samples. *Journal of Statistical Software* **2004**, *9(1)*, 1-19.
87. Lumley, T. *Survey: analysis of complex survey samples; R package version 3.32* **2017**.
88. Clyde, M. *BAS: Bayesian Adaptive Sampling for Bayesian Model Averaging; R package version 1.4.7* **2017**.
89. Zeugner, S. *Bayesian Model Averaging with BMS for BMS Version 0.3.0*; online: *www.bms.zeugner eu*; **2011**.
90. Liang, F.; Paulo, R.; Molina, G.; Clyde, M.A.; Berger, J.O. Mixtures of g priors for Bayesian Variable Selection. *Journal of the American Statistical Association* **2008**, *103*, 410-423.
91. Feldkircher, M; Zeugner, S. Benchmark priors revisited: on adaptive shrinkage and the supermodel effect in Bayesian Model Averaging *IMF Working Paperper*, WP/09/202. **2009**.
92. Sileshi, G. W. A critical review of forest biomass estimation models, common mistakes and corrective measures. *Forest Ecology and Management* **2014**, *329*, 237-254.
93. McCune, B; Grace, J.B. *Analysis of Ecological Communities* Gleneden Beach, OR, USA, 2002; pp 304.
94. Weisberg, S. *Applied linear regression (Vol. 528)*. John Wiley & Sons: Hoboken, NJ, USA, 2005; pp. 299.

95. Piñeiro, G.; Perelman, S.; Guerschman, J. P.; Paruelo, J. M. How to evaluate models: observed vs. predicted or predicted vs. observed?. *Ecological Modelling* **2008**, *216(3)*, 316-322.
96. 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. *Forest Ecology and Management* **2011**, *262(8)*, 1648-1657.
97. Jolicoeur, P. (1990). Bivariate allometry: interval estimation of the slopes of the ordinary and standardized normal major axes and structural relationship. *Journ. of Theoretical Biology* **1990**, *144* 275–285.
98. Legendre, P; L. Legendre. *Numerical Ecology, 2nd English edition*. Elsevier Science BV: Amsterdam, 1998.
99. Ludbrook J. Linear regression analysis for comparing two measurers or methods of measurement: but which regression? *Clinical and Experimental Pharmacology and Physiology* **2010**, *37*, 692–699.
100. Legendre, P. *Model II regression user's guide, R edition*.
101. Chambers, J. M.; Cleveland, W. S.; Kleiner, B.; Tukey, P. A. *Graphical Methods for Data Analysis*. Wadsworth & Brooks/Cole, 1983.
102. Gobakken, T.; Næsset, E. Assessing effects of laser point density, ground sampling intensity, and field sample plot size on biophysical stand properties derived from airborne laser scanner data. *Canadian Journal of Forest Research* **2008**, *38(5)*, 1095-1109.
103. Gobakken, T.; Næsset, E. Assessing effects of positioning errors and sample plot size on biophysical stand properties derived form airborne laser scanner data. *Canadian Journal of Forest Research*, **2009**. *39(5)*, 1036–1052.
104. 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 Sensing of Environment* **2011**, *115(2)*, 636-649.
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 Sensing* **2013**, *5(6)*, 3007-3036.
106. 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.
107. 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 and Management* **2014**, *9(1)*, 3.
108. Sheridan, R. D.; Popescu, S. C.; Gatziolis, 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 Sensing* **2014**, *7(1)*, 229-255.
109. 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. *Mathematical and Computational Forestry & Natural Resource Sciences* **2015**, *7(2)*, 49.
110. Guthery, F. S.; Brennan, L. A.; Peterson, M. J.; Lusk, J. J. Information theory in wildlife science: critique and viewpoint. *Journal of Wildlife Management* **2005**, *69(2)*, 457-465.
111. Mundry, R. (2011). Issues in information theory-based statistical inference—a commentary from a frequentist's perspective. Behavioral Ecology and Sociobiology, 65(1), 57-68.
112. Stephens, P. A.; Buskirk, S. W.; Hayward, G. D.; Martinez Del Rio, C. Information theory and hypothesis testing: a call for pluralism. *Journal of Applied Ecology* **2005**, *42(1)*, 4-12.
113. 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. *Canadian Journal of Forest Research* **2006**, *36(1)*, 34-47.
114. 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 Sensing of Environment* **2009**, *113(1)*, 148-159.
115. Asner, G. P.; Mascaro, J. Mapping tropical forest carbon: Calibrating plot estimates to a simple LiDAR metric. *Remote Sensing of Environment* **2014**, *140*, 614-624.
116. Price, C. A.; Ogle, K.; White, E. P.; Weitz, J. S. Evaluating scaling models in biology using hierarchical Bayesian approaches. *Ecology letters 2009*, *12(7)*, 641-651.
117. 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(3)*, e58241.