Remote Sensing of Environment 280 (2022) 113147

Contents lists available at ScienceDirect

Remote Sensing of Environment

journal homepage: www.elsevier.com/locate/rse

Direct use of large-footprint lidar waveforms to estimate aboveground biomass

Wenge Ni-Meister [a,*], Alejandro Rojas [a], Shihyan Lee [b]

[a] Department of Geography and Environmental Science, Hunter College of CUNY, New York, NY 10065, United States of America
[b] Science Application International Corp. and NASA Goddard Space Flight Center, Mail Code: 616.2, Greenbelt, MD 20771, United States of America

ARTICLE INFO

Edited by Jing M. Chen

Keywords:
Aboveground biomass
Biomass indices
Lidar waveforms

ABSTRACT

Many studies have established the strong connections between aboveground biomass and lidar height metrics; however, these relationships are site-specific. Field data required to derive these relationships are not readily available in many cases. We developed a model to estimate plot-level aboveground biomass density (AGBD) directly from large-footprint lidar waveform measurements. An individual tree-based aboveground biomass (AGB)-height allometric relationship was scaled up to the plot level using lidar-waveform sensed tree height and crown size distribution characteristics. The AGBD was estimated based on a waveform/foliage profile-weighted height-based allometric equation. The AGBD-height scaling exponent was then built on the allometric relationships of tree height with stem diameter and crown volume with tree height. Global vegetation structure data analysis demonstrated that one general model (scaling exponent ~ 1.6–1.8) works reasonably well across all global forest biomes except boreal forests (scaling exponent ~ 0.9). We applied the model to estimate aboveground biomass in two distinct geographic regions: temperate deciduous/conifer forests in the northeastern USA and a montane conifer forest in Sierra National Forest in California. Local vegetation structural data analysis leads to a consistent height scaling exponent for these two distinct biomes, slightly different from the global data analysis results. This model produced optimal AGBD estimates using the local height scaling exponent value. Adequate AGBD estimates with the general height scaling exponent value were also provided by our model. Our analysis suggests one general allometric relationship between plot-level AGBD and large-footprint lidar waveforms. Integrating local structure allometric relationships improve the predictive accuracy of the model. Our model outperformed the lidar height metrics-based approach for AGBD estimates and overcame the biomass underestimation problem using height metrics for high biomass regions. This model could potentially serve as a general and robust model for monitoring forest carbon stocks using large-footprint lidar waveform measurements such as the Global Ecosystem Dynamics Investigation (GEDI) mission at the continental and global scales. The model could be a framework for integrating a demography-based terrestrial ecosystem model and GEDI global mission measurements to improve global carbon stock and flux estimates.

1. Introduction

Forests account for about 92% of the terrestrial vegetation carbon pool (https://www.globalchange.gov/) (Goldstein et al., 2020) and store up to 80% of all the biomass on Earth (Bar-On et al., 2018; Goldstein et al., 2020; Reichstein and Carvalhais, 2019). Forest aboveground biomass (AGB) is a significant contributor to total terrestrial carbon storage. Accurate measurements of forest aboveground biomass over large spatial scales are crucial for quantifying terrestrial carbon exchanges and are critical for successfully implementing climate change mitigation policies and sustainable forest management. Despite their importance, significant uncertainties remain concerning terrestrial carbon budgets, limiting the ability of current science to quantify their roles and predict their trajectory in the global carbon cycle (Fisher and Koven, 2020; Reichstein and Carvalhais, 2019).

It is neither practical nor cost-effective to collect field aboveground biomass data for large regions. Satellite observations with dense sampling in space and time are essential to characterize the heterogeneity of ecosystem structure and estimate terrestrial ecosystem carbon storage at large scales (Schimel et al., 2015). Lidar directly measures the vertical

* Corresponding author at: Department of Geography and Environmental Science, Hunter College of The City University of New York, 695 Park Ave., New York, NY 10021, United States of America.
E-mail address: Wenge.Ni-Meister@hunter.cuny.edu (W. Ni-Meister).

https://doi.org/10.1016/j.rse.2022.113147
Received 8 August 2021; Received in revised form 10 May 2022; Accepted 24 June 2022
Available online 18 July 2022
0034-4257/

structure of vegetation by providing vertical density profiles (wave forms) of vegetation elements (stems, leaves) that reflect back laser pulses. Such data have excellent potential to provide large-scale estimates of forest carbon storage (Schimel et al., 2015; Wulder et al., 2012). Various extensive spatial vegetation lidar data have become more widely available, including airborne discrete-return lidar, airborne waveform, satellite waveform lidar, and ground-based lidar (Coops et al., 2021). The Global Ecosystem Dynamics Investigation (GEDI) mission collects high-quality measurements of vertical forest structure in temperate and tropical forests between 51.6° N and 51.6° S latitudes at 25 m footprint resolution (Dubayah et al., 2020). GEDI will provide over 10 billion waveforms during its four-year lifetime. A multitude of high-resolution GEDI measurements will provide better quantifications of canopy height, vertical foliage profiles, and aboveground biomass at global and continental scales. Lidar has much fewer saturation problems than conventional optical remote sensing and radar technology and is recognized as a state-of-the-art remote sensing technology for mapping aboveground biomass (Ni-Meister, 2015).

Various waveform lidar sensors at different footprint sizes provide three-dimensional vegetation structure and aboveground biomass measurements at different spatial scales. Many studies have successfully used spaceborne and airborne full-waveform lidar data to estimate aboveground biomass characteristics across various biomes. For example, the airborne Land, Vegetation, and Ice Sensor (LVIS) waveform lidar has been used to estimate aboveground biomass density (AGBD, total AGB per unit area) in conifer forests in the Cascade Mountain Range in Oregon and Washington, USA (Lefsky et al., 2005) and Sierra National Forest, California, USA (Swatantran et al., 2011); temperate deciduous forests in Annapolis, Maryland, USA (Lefsky et al., 1999) and the White Mountains, New Hampshire, USA (Anderson et al., 2006; Anderson et al., 2008; Ni-Meister et al., 2010a), tropical forests in La Selva, Costa Rica (Drake et al., 2002; Dubayah et al., 2010). The spaceborne waveform-based Geosciences Laser Altimeter System (GLAS) on the Ice, Cloud, and Land Elevation Satellite (ICESat) at approximately 70 m footprint scale has been employed to estimate aboveground biomass in a conifer forest in Oregon and a deciduous forest in Tennessee, the USA, and tropical forests in Santarem, in the state of Para, Brazil (Lefsky et al., 2005). GEDI and ICESat-2 data have been used to estimate aboveground biomass in the Western USA (Duncanson et al., 2020; Silva et al., 2021).

The most common approach to estimating aboveground biomass from large-footprint waveform lidar is using lidar height metrics calculated based on waveform measurements. Aboveground biomass density is often estimated as a linear regression of lidar height metrics, i.e., RH100, RH75, RH50, and RH25, where RHX is defined as the height relative to the ground (denoted as "RH") at which there is X% of the accumulated total waveform energy from the last detectable return. Recent LVIS and GEDI height metrics include more RH values: RH10 to RH95 at 5% intervals, then RH96, RH97, RH98, RH99, and finally RH100. Lidar waveform directly links to vegetation structure (Ni-Meister et al., 2001; Ni-Meister et al., 2018). Each lidar height metric calculated from full waveforms represents specific vegetation structure characteristics associated with aboveground biomass. For example, RH100 is a good predictor of aboveground biomass in dense forests in the northeastern USA. However, the model fails to predict aboveground biomass in a newly-harvested forest (Ni-Meister et al., 2010a). Generally, RH50 is better related to aboveground biomass than RH 100 because the latter measures the height of the canopy top; RH50 includes both canopy height and canopy density, and it has been widely used to estimate aboveground biomass in many forests. However, RH75 was the best predictor for aboveground biomass in Sierra National Forest, CA, USA (Swatantran et al., 2011). RH80 works best in south-central Sweden (Saarela et al., 2020), and (Sheridan et al., 2014) used RH90 for Pacific Northwest. Multiple variable regression analysis is often employed to obtain the best results and identify the optimal relationship between aboveground biomass and different lidar height metrics (Choi et al., 2013; Margolis et al., 2015; Nelson, 2010; Nelson et al., 2017). Many lidar biomass models are often derived by empirical regression analysis and calibrated to site-specific fitting parameters. More than 50 different lidar parameters have been used to identify optimal parameters for estimating AGBD (Nelson et al., 2017). However, these parameters and coefficients used in the biomass estimates are site-specific, and field aboveground biomass data are required for each study site to develop the optimal aboveground biomass models.

Applying this height metrics-based regression approach at large or global scales is difficult without a dense ground validation network. Building such a dense global network is either too expensive or impractical. Ultimately, to advance beyond regressions on RH metrics, what is needed is a generalizable understanding of the fundamental relationship between aboveground biomass and full lidar waveform measurements. This study aims to develop a general model to fully describe the vital connection between aboveground biomass and lidar waveform measurements.

This study takes the following approach: The canopy community is composed of multi-cohorts that are ensembles of identical individual trees. For individual trees, the woody stem is the main AGB component, which can be calculated from the stem geometry (diameter, taper factor, and height) and the wood density. At the plot level, the total AGB derives from the sum of the population of trees of different cohorts that form the canopy community. Lidar waveform samples the vertical profile of density of canopy elements, which are an integration of that community; each peak of the waveforms reflects the crown sizes of the trees and their densities of one cohort in the vegetation community. Therefore, it should be possible to derive AGB based on individual-level tree allometry (relations of lidar observable height to stem diameter and crown size) and on population density as reflected in the lidar waveform caused by crown densities for each cohort. We link individual tree geometry allometry with tree biomass and lidar waveform with both height stratification of a canopy and density of tree crowns for each cohort. The challenge then is to 1) determine the appropriate allometric relations at the individual level and globally relevant parameterizations, and 2) to incorporate the lidar waveform as a convolution of tree density and crown sizes.

Many studies have investigated individual tree-level allometric relationships of aboveground biomass with tree height and crown size structure characteristics. A strong relationship between aboveground biomass and tree height exists due to the power-law scaling relationship between tree height and stem diameter (Chave et al., 2014; Feldpausch et al., 2011; Feldpausch et al., 2012; Henry et al., 2010; Ploton et al., 2015). However, power-law relationships vary with species, climate, and stand age, failing to account for the asymptotic nature of faster tree height growth in young trees than in mature trees (Chave et al., 2014; Jucker et al., 2017). Some studies found that stem diameter and aboveground biomass correlate well with crown width even in large trees (Goodman et al., 2014; Henry et al., 2010; Ploton et al., 2015). Jucker et al. (2017) developed general allometric models to estimate stem diameter and aboveground biomass from tree height and crown diameter using a global vegetation structure database. These findings support the idea that both stem diameter and crown size affect scaling of AGB with height beyond a simple power law, where height distribution could potentially be derived from small-footprint lidar measurements. However, all these studies focused on the individual tree level.

Plot-level tree height size distribution estimates have been made from small footprint lidar measurements with great success (Fischer et al., 2020; Spriggs et al., 2017; Taubert et al., 2021). In this study, we developed a model to estimate plot-level AGBD directly using large-footprint full-waveform lidar measurements. Our aboveground biomass model is based on tree height stratification measured by lidar waveforms and the allometric relationships among height, stem diameter, and crown size, where crown size convolved with the density of trees at each height level result in the waveform bulge. This model is general enough to provide reasonable aboveground biomass estimates

for different forest biomes and geographic regions with limited field data. This type of model reduces the site-specific calibration effort for aboveground biomass estimates.

The text structure below is arranged as follows: the general AGBD model and the model generality assessment using a global vegetation structure dataset are demonstrated in Section 2. The test sites and datasets for model evaluation are described in Section 3. The model performance in our test sites is explained in Section 4. Finally, the results and conclusions are discussed and summarized in Sections 5 and 6.

2. The lidar waveform-based aboveground biomass model

We started at the individual tree scale and first developed a height-based aboveground biomass allometric equation for individual trees. Then, we expanded the height-based aboveground biomass allometric equation to plot-level by summing the biomass of all the trees within each plot. Lastly, we linked the plot-level biomass estimates with the waveforms.

2.1. Height-based biomass allometric equation at the individual tree level

A general model of the total aboveground biomass of a tree (kg) is expressed as the product of wood volume and wood specific gravity and a tapering factor (Chave et al., 2005; Chave et al., 2014):

$$AGB_{\text{tree}} = \frac{1}{10} F \left(\rho \frac{\pi}{4} D^2 H\right)^k \quad (1)$$

where H is tree height in m, D is the stem diameter at breast height in cm, ρ is the wood specific gravity (oven-dry wood over green volume in g/cm^3), $k < 1$, $k = 0.976$ was the best fit for measurement for tropical forest (Chave et al., 2014). F is the tree taper factor, *with* $F = 0.6$ for broadleaf species and $F = 0.333$ for a perfect conical shape (Chave et al., 2005).

Stem diameter, D, in Eq. (1) is related to height, H, through their allometric relationship:

$$H = b\, D^{\beta} \quad (2)$$

where H is in m and D is in cm, b is the coefficient, and β is the allometric scaling exponent. The relationship between tree height and stem diameter varies with species and climate conditions (Chave et al., 2014). Niklas (1994) found $\beta = 0.535$; Ketterings et al. (2001a) estimated an average $\beta = 0.62$ from their field data. Zianis and Mencuccini (2004) used a fractal geometry model, in conjunction with field data analysis, and found the scaling exponent β is between 0 and 1, but varies with tree height/size (or stand age), concurring with the findings by Ketterings et al. (2001b). Pilli et al. (2006) concluded β appeared to be related to tree stage but independent of species and site, using forty-nine datasets of different species, sites, and stand ages. A recent analysis of the Forest Inventory and Analysis (FIA) data revealed that β varies with species, tree height, and geographic conditions (Duncanson et al., 2015). Jucker et al. (2017) found a single tree height and stem diameter allometric function to estimate stem diameter without introducing systematic bias for different forest biomes. Incorporating different scaling relationships among forest types, biogeographic regions, and functional groups helped improve the predictive accuracy of the model (Shenkin et al., 2020).

By converting from the diameter to height through their relationship in Eq. (1) and approximating $k = 1$ (Chave et al., 2014), we solved for the individual tree aboveground biomass (AGB_{tree}) allometric equation as a function solely of height:

$$AGB_{\text{tree}} = \frac{\pi}{40} b^{-\frac{2}{\beta}} F \rho H^{\left(1+\frac{2}{\beta}\right)} \quad (3)$$

Height-based AGB equations are not widely used because tree height is not as frequently measured as tree stem diameter. However, (Agee, 1981) found that the allometric scaling exponent, $1 + \frac{2}{\beta}$, ranged from 2.2 to 2.7 for conifer forests in the Pacific Northwest, USA, and that tree height was usually better related to aboveground biomass than stem diameter. Lidar can measure tree height, but the relation to AGB needs further refinement, which we present here.

2.2. The lidar waveform based biomass allometric equation at the plot level

Vegetation canopy layers are comprised of cohorts of plants that are ensembles of identical individuals. Assume the canopy layer has n cohorts, for the *ith* cohort, height ranges $z_{i-1} < z < z_i$, z_{i-1} and z_i are lower bound and upper bound heights of *ith* cohort, and $z_0 = 0$ at the ground for the near ground cohort number. The *ith* cohort has n_i trees in it, $n_i = \lambda(z_{i-1}) - \lambda(z_i)$, $\lambda(z_i)$ is the accumulated tree count density from the canopy top to tree height z_i. If the height of the *kth* tree in *ith* cohort is z_{ik}, then aboveground biomass density at the plot level is the summation of biomass for all trees from all cohorts within a plot, described as follows,

$$AGBD = \frac{\pi}{40} b^{-\frac{2}{\beta}} F \rho \sum_{i=1}^{n} \sum_{k=1}^{n_i} z_{ik}^{\left(1+\frac{2}{\beta}\right)} \quad (4)$$

Assuming the similar plant function type, the same allometric relationships, and constant wood specific gravity ρ, and tapering factor F, with a plot, the continuous form of Eq. (4) is,

$$AGBD = \frac{\pi}{40} b^{-\frac{2}{\beta}} F \rho \sum_{i=1}^{n} \int_{z_{i-1}}^{z_i} -\frac{d\lambda(z)}{dz} z^{\left(1+\frac{2}{\beta}\right)} dz \quad (5)$$

where $\lambda(z)$ is the accumulated tree count density from the canopy top to tree height z. $\frac{d\lambda(z)}{dz}$ is the tree count density per unit height interval (1/m^3), i.e., tree height distribution at different height intervals. The integration goes from the lower bound, z_{i-1}, to the upper bound of the canopy, z_i, for *ith* cohort resulting in a negative sign $\frac{d\lambda(z)}{dz}$. Eq. (5) implies that tree height and its distribution are the most significant variables for aboveground biomass density.

Large footprint lidar waveforms record backscattering energy intensity at each height interval and are directly affected by the vertical and horizontal vegetation structure (Ni-Meister et al., 2001; Ni-Meister et al., 2018). As demonstrated in previous empirical studies, the strong relationships of aboveground biomass with height metrics directly derived from lidar waveforms suggest a strong link between lidar waveforms and aboveground biomass. Horizontally, at each height interval, strong photon returns imply dense tree counts or large tree crowns if not dense, and large foliage volume density at that height level. Vertically, taller trees measured by lidar are associated with larger aboveground biomass and vice versa. Thus the vertical distributions of tree height and crown size, the key variables controlling aboveground biomass are directly measured by lidar waveforms. We developed an approach to link lidar waveforms with tree height and size distributions and plot-level aboveground biomass density.

The geometric optical and canopy radiative (GORT) model provides a fundamental basis for linking lidar waveforms with plot-level aboveground biomass. GORT successfully simulates canopy vegetation lidar waveforms as a function of laser pointing angle and vegetation structure parameters, including canopy crown size, vegetation height, height variation, vegetation density, foliage volume density, and leaf and background reflectivity coefficient ratio (Ni-Meister et al., 2001; Ni-Meister et al., 2018). The GORT modeled waveforms have been fully evaluated in single-layered and multilayered (multi-cohorts) conifer and deciduous forests using LVIS measurements (Ni-Meister et al., 2018). Lidar waveforms and vegetation structure characteristics are directly linked by canopy gap probability, which is defined as the probability of light penetrating vegetation without hitting any canopy element in the

GORT model. For a natural forest stand, assuming that tree crowns are randomly distributed in space, nadir canopy gap probability at height z, ($P(0,z)$), can be expressed as follows (Ni-Meister et al., 2010b; Yang et al., 2010):

$$P(0,z) = e^{-GL_e(z)} \tag{6}$$

where G is a parameter describing leaf orientation distribution, $L_e(z)$ is the cumulative effective leaf area index from the canopy top to height z, which is a product of the clumping factor and accumulative leaf area index,

$$L_e(z) = \gamma\, F_a\, V_{c,i}\, \lambda(z) = -\frac{ln(P(0,z)}{G} \tag{7}$$

where γ, the clumping factor varies with tree structure characteristics (e.g., shape, size, tree count density, and within-crown foliage density) and light incident angle but is constant vertically (Ni-Meister et al., 2010b), F_a is the foliage area volume density ($1/m^3$), $V_{c,i}$ is the crown volume for trees in the *ith* cohort, λ(z) ($1/m^2$) is the accumulative tree count density from the top of the canopy to height, z. Assuming lidar waveform samples a multi-cohort (multi-layer) canopy, each peak in the waveform corresponds to one cohort of the canopy. Each cohort is defined as a canopy layer with a similar tree crown size. For one cohort (one layer) canopy for our test sites, waveforms and derived foliage profiles correspond to tree count density profiles as shown in Figs. 4 and 5 (Section 4.2). For a multi-cohort (multi-layer) canopy, each peak in the waveform corresponds to total crown volume for one cohort of the canopy. Assume a constant crown volume $V_{c,i}$ for each cohort, i and a constant foliage area volume density for the whole canopy, then $\frac{dL_e(z)}{dz}$ is,

$$\frac{dL_e(z)}{dz} = \gamma\, F_a\, V_{c,i}\, \frac{d\lambda(z)}{dz} = -\frac{dln(P(0,z))}{G \bullet dz} \tag{8}$$

Ni-Meister et al. (2001) defined, $\frac{dln(P(0,z))}{dz} = F_{app}(z)$ – apparent foliage profile. Eq. (8) can be written as follows:

$$-\frac{d\lambda(z)}{dz} = \frac{1}{G\,\gamma\,F_a\,V_{c,i}}\,\frac{dln(P(0,z))}{dz} = \frac{F_{app}(z)}{G\,\gamma\,F_a\,V_{c,i}} \tag{9}$$

To correct the impact of multi-cohort canopies on crown size difference, crown volume, $V_{c,i}$ varies with cohorts. Cohorts with taller trees correspond to larger tree volumes. Assume the same allometric relationships within the canopy, we constrain crown volume with cohort tree height, H_i only (Shenkin et al. 2020):

$$V_{c,i} = a\, H_i^{\propto} \tag{10}$$

α is the crown volume and tree height scaling exponent, and a is the coefficient. We approximate cohort tree height, H_i by height z in each cohort. At the plot level, parameters such as leaf orientation factor G, clumping factor, γ, and foliage area volume density F_a are independent of height and assumed to be constant within a plot. Replacing $-\frac{d\lambda(z)}{dz}$ in Eq. (5) by Eq. (9) results in the relationship of the aboveground biomass density at plot level with the lidar waveform as follows,

$$AGBD = \frac{\pi}{40}\, b^{-\frac{2}{\beta}}\frac{F\,\rho}{G\,\gamma\,F_a}\int_{z_0}^{z_n}\frac{dln(P(0,z))}{dz}\,\frac{z^{\left(1+\frac{2}{\beta}\right)}}{a\,z^{\alpha}}dz \tag{11}$$

z_0 is the lower bound of the lowest cohort and z_n is the upper bound of highest cohort (top of canopy). Eq. (12) can be simplified as follows,

$$\begin{cases} AGBD = k_{FP}\, BI_{FP} \\ k_{FP} = \dfrac{\pi}{40}\,\dfrac{b^{-\frac{2}{\beta}}\,F\,\rho}{a\,G\,\gamma\,F_a} \\ BI_{FP} = \displaystyle\int_0^{z_n}\frac{dln(P(0,z))}{dz}\,z^{\left(1+\frac{2}{\beta}\right)-\propto}\,dz \end{cases} \tag{12}$$

BI_{FP} in Eq. (12) refers to the foliage profile-based biomass index. It is an integral from the bottom to the top of the canopy layer of the product of an apparent foliage profile and a power function of tree height. Canopy gap probability, $P(0,z)$, thus $\frac{dln(P(0,z))}{dz}$ can be directly derived from lidar waveforms (Ni-Meister et al., 2018). k_{FP} is the coefficient of the AGBD and foliage profile-based biomass index relationship, which depends on tree taper coefficient, F, wood specific gravity ρ, leaf orientation G, clumping factor γ, and foliage area volume density F_a.

We sought an index directly related to the waveforms to simplify the calculation above. $\frac{d(P(0,z))}{dz}$ is the waveform and $\frac{dln(P(0,z))}{dz} = \frac{1}{P(0,z)}\frac{d(P(0,z))}{dz}$. The difference between $\frac{d(P(0,z))}{dz}$ and $\frac{dln(P(0,z))}{dz}$ is a factor, $\frac{1}{P(0,z)}$, at each height level. We approximated the integrated effect of $\frac{1}{P(0,z)}$ by the averaged canopy gap probability at the top and beneath the canopy, P_{gap}, the biomass index can be directly calculated from the waveforms. The aboveground biomass can be estimated as follows,

$$\begin{cases} AGBD = k_{WF}\, BI_{WF} \\ k_{WF} = \dfrac{\pi}{40}\,\dfrac{b^{-\frac{2}{\beta}}\,F\,\rho}{a\,P_{gap}\,G\,\gamma\,F_a} \\ BI_{WF} = \displaystyle\int_{z_0}^{z_n}\frac{d(P(0,z))}{dz}\,z^{\left(1+\frac{2}{\beta}\right)-\propto}\,dz \end{cases} \tag{13}$$

Eq. (13) indicates that the aboveground biomass can be directly calculated from the waveforms. BI_{WF} is called waveform-based biomass index. The coefficient, k_{WF}, in Eq. (13) is different from the foliage profile-based, k_{FP}, in Eq. (12) by a factor of averaged canopy gap probability, P_{gap}. Eqs. (12) and (13) demonstrate that aboveground biomass at plot level is directly related to biomass index, k_{WF} and k_{FP} are coefficients of the AGBD and waveform-based and foliage-based biomass index relationships respectively, they are functions of tree taper, wood density, leaf orientation, foliage area per crown, and averaged canopy gap probability. Parameters, such as tree taper, wood density, and leaf orientation, are functions of plant functional types. Foliage volume density and averaged canopy gap probability vary with stand age and geographic region. All of these parameters are assumed to be constant vertically within each plot.

The aboveground biomass density model laid out in Eqs. (12) and (13) illustrate that plot-level aboveground biomass density is linearly related to a biomass index, a waveform/foliage profile-weighted height power functions. Waveforms are directly linked to tree height and size distributions. A waveform is a suitable indicator of tree height and tree size distribution within each footprint. In the model, the scaling exponent, $1+\frac{2}{\beta}-\propto$, depends on tree height and diameter allometric scaling exponent(β) and crown volume and tree height allometric scaling exponent (∝). Biomass indices have the dominant structural parameters directly associated with above-ground biomass.

Biomass index is a significant variable influencing AGBD estimates. However, AGBD is also affected by other structural parameters in the coefficient term, and they may vary with stand age, tree density, and species. The main focus of this study is on assessing how biomass indices are related to aboveground biomass density and how the relationships vary across large geographic regions with different ecosystem structural characteristics.

2.3. Assessment of the model generality

In Eqs. (12) and (13), the scaling exponent, $1+\frac{2}{\beta}-\propto$ depends on tree height - stem diameter scaling exponent (β) and the crown volume - tree height scaling exponent (α). The variability of $1+\frac{2}{\beta}-\propto$ determines the generality of this model. We adopted a global database of 108,753 tree measurements of stem diameter, height, and crown diameter to assess the generalization of the model through examining the variations of all the scaling exponents (β, α, and $1+\frac{2}{\beta}-\propto$) for different forest biomes. With available tree height, stem diameter, and horizontal crown

diameter measurements, we analyzed the allometric relationships of a) tree height with stem diameter, b) horizontal crown radius with tree height, and c) crown volume with tree height (Fig. 1). Tree height and stem diameter relationship allow us to assess the variability of β. To investigate the allometric relationship of crown volume with tree height α, we estimated crown volume using horizontal crown size and tree height measurements Assuming an ellipsoid tree crown shape, the crown volumes were computed based on the mathematical relationship of ellipsoid crown volume with the horizontal crown radius (half of the crown diameter) and the vertical crown radius (half of the crown depth which was approximated as half the tree height (Ishii et al., 2003; Sillett et al., 2020)). We calculated the coefficient of determination (R^2), Root Mean Square Error (RMSE), and normalized RMSE by the mean of the predicted variables for all the analyses.

Tree height and stem diameter show a significant allometric relationship with R^2 ranging from 0.68 to 0.84. However, the woodland and savanna biome has R^2 = 0.56, RMSE = 3.2–5 m, and NRMSE = 0.24–0.59, making it the biome with the largest NRMSE and the largest uncertainty of the height and stem diameter relationship. The scaling exponent values are relatively uniform across different biomes, ranging from 0.58 (temperate mixed forests) to 0.75 (woodland and savanna). Crown radius and height show strong allometric relationships with R^2 = 0.41–0.57, RMSE = 0.55–1.46 m, and NRMSE = 0.36–0.53, with tropical forest, woodland, and savanna biomes the most significant NRMSE

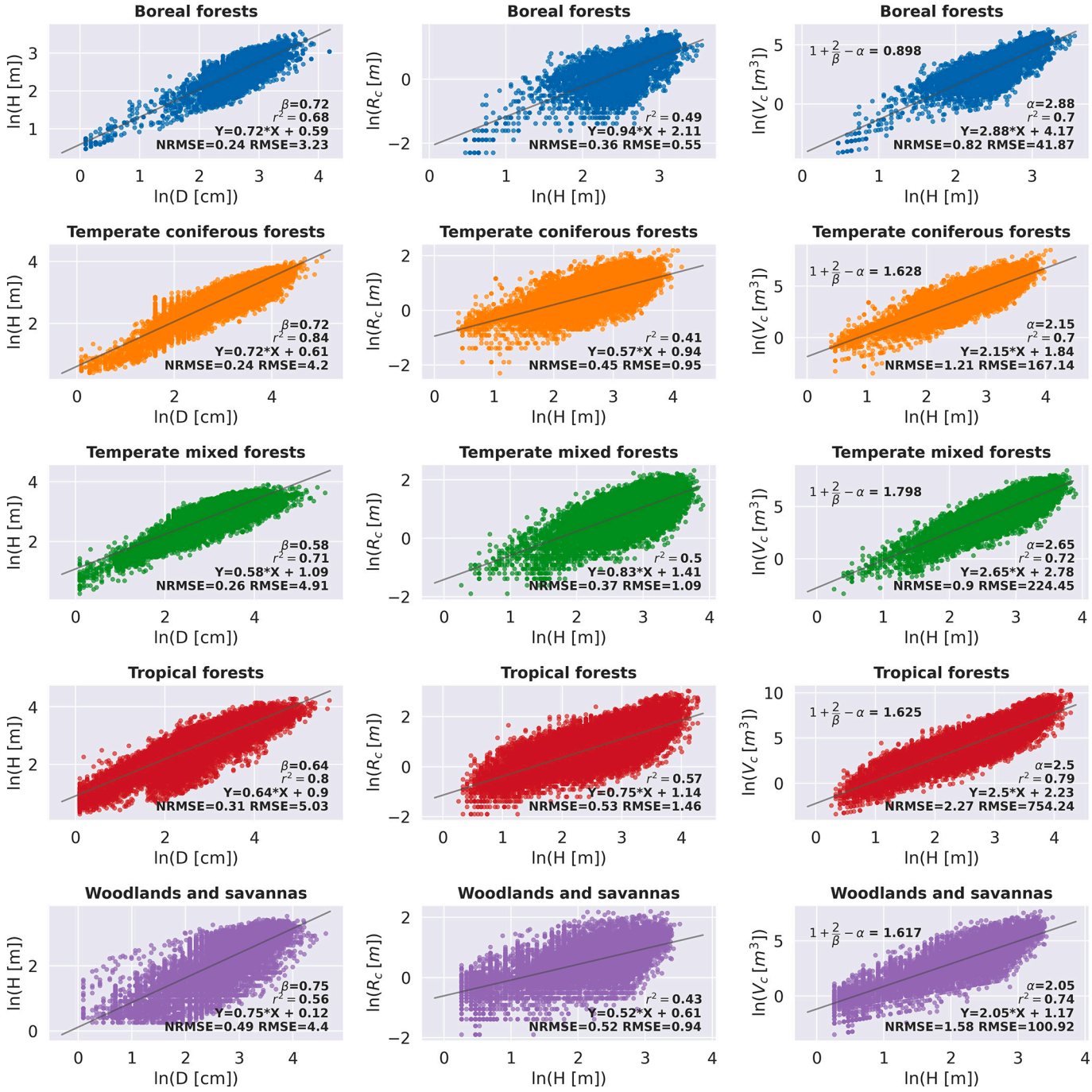


Fig. 1. Allometric scaling relationships of tree height (H) with stem diameter (D) (first column), crown radius (R_c) with tree height (H) (second column), and crown volume (V_c) with tree height (H) (last column) for different forest biomes. The lower right corner of each subplot: The H and D scaling exponent, β and V_c and H scaling exponent, ∝ and associated statistics. Upper left corner of the third column: the resulting AGBD and height scaling exponent, $1+\frac{2}{\beta}-\propto$ for each biome.

values. The allometric relationships between volume with tree height are significant, with $R^2 = 0.7–0.79$, RMSE = 42–754 m^3, and NRMSE = 0.82–2.27. The crown volume - tree height scaling exponent varies among forest biomes (2.05–2.88). These values range from highest to least in boreal forests, temperate mixed forests, tropical forests, and temperate coniferous forests, with the least in woodland and savanna. Based on the β and α values, the resulting AGB and height scaling exponent, $1+\frac{2}{\beta}-\alpha$, in Eqs. (12) and (13) are relatively constant (1.6–1.8) for all forest biomes except boreal forests (0.9). This result suggests that the AGBD and height scaling exponent in Eqs. (12) and (13) are relatively constant across all forest biomes, except boreal forests. This analysis suggests that one general biomass model can be applied to different forest biomes except for boreal forests at large scales. The local variations of β and α values directly lead to local variations in $1+\frac{2}{\beta}-\alpha$. It is also essential to explore the local variability of the AGBD and height scaling exponent.

3. Initial model performance assessment

We selected two forest biomes with distinct forest structures and geographic regions to evaluate the model's performance. We used LVIS full-waveform measurements collected in these regions to calculate aboveground biomass indices and compared these indices with field measured aboveground biomass. We could not model the exact AGBD due to limited knowledge of the coefficients (k_{FP} *and* k_{WF}) in the model. The following sections provide a detailed description of the study sites, data processing, and the assessment results.

3.1. Test sites

Our test sites were temperate deciduous/conifer forests in New England (NE) and montane conifer forests in the Sierra National Forest in California (CA) (Fig. 2). These two biomes have different climates, topography, different forest stand ages, and tree species composition.

The NE sites comprised three intensive ecological study areas: Harvard Forest (HF) in Massachusetts, Bartlett Experimental Forest (BEF) in the White Mountain region, New Hampshire, and Ecosystem Research Forest in Howland, Maine. Harvard Forest is located in the transition zone of hardwoods-white pine-hemlock. The dominant species are red maple (*Acer rubrum*), red oak (*Quercus rubra*), white birch (*B. papyrifera*), yellow birch (*Betula allleghaniensis*), beech (*Fagus grandifolia*), white pine (*Pinus strobus*), and hemlock (*Tsuga canadensis*). BEF in NH was established to study secondary deciduous and coniferous forest dynamics and ecology. The primary tree species in this area were American beech (*Fagus grandifolia*), red maple (*Acer rubrum*), eastern hemlock (*Tsuga canadensis*), sugar maple (*Acer saccharum*), yellow birch (*Betula alleghaniensis*), paper birch (*Betula papyrifera*), red spruce (*Picea rubens*), and balsam fir (*Abies balsamea*), with some localized small stands of eastern white pine (*Pinus strobus*). The Forest Ecosystem Research site in Howland, ME, was located within the Northern Experimental Forest of the International Paper Co. The site was an assemblage of small plantations, multi-generation clearings, and sizable natural forest stands. The stands were mixed hemlock (*Tsuga*), spruce (*Picea*), fir (*Abies*), aspen (*Populus*), and birch (*Betula*) species.

The CA study site in the Sierra National Forest (37°00′N, 119°10′W) were on the western slope of central Sierra Nevada in CA, USA. The site has a Mediterranean climate with hilly slopes. Elevations ranged from 853 to 2743 m. The dominant tree species in this region are red fir (*Abies magnifica*), white fir (*Abies concolor*), ponderosa pine (*Pinus ponderosa*), and California black oak (*Quercus kellogi*).

3.2. Field data and data processing

Datasets used in this study included vegetation structure field data collected in 2003, 2007, and 2008. A total of 73 plots were sampled, 48 in New England and 25 in Sierra Forest.

Two field datasets were collected in New England, one in 2007 and one in 2003. The 2007 dataset includes tree data collected in 28 plots at three forest sites in New England: Harvard Forest in MA, Bartlett Experimental Forest, NH, and Howland Ecosystem Research Forest, ME. Two stands were selected for each site, in a total of six stands, named Hardwood and Hemlock from Harvard Forest, MA; B2 and C2 from Bartlett Experimental Forest, NH and Tower and Shelterwood from Howland Ecosystem Research Forest, ME (Fig. 2). Each stand comprises five circular plots), ranging from 30 to 50 m with a center plot (CT) and four corner plots (NE, NW, SE, and SW) in between. One exception was the Tower stand in Howland, with only three plots (CT, NO, and SO) (Fig. 2). Each plot was circular with a 25 m radius for the Harvard Hardwood stand, MA, and 20 m for the rest of the stands. For all six stands, half (Harvard Hemlock in the Harvard Forest, and Tower and Shelterwood in the Howland Forest) were dominated by conifer trees, and the other half (Harvard Hardwood stand and both B2 and C2 stands in Bartlett Forest) by deciduous trees. Species information and trunk

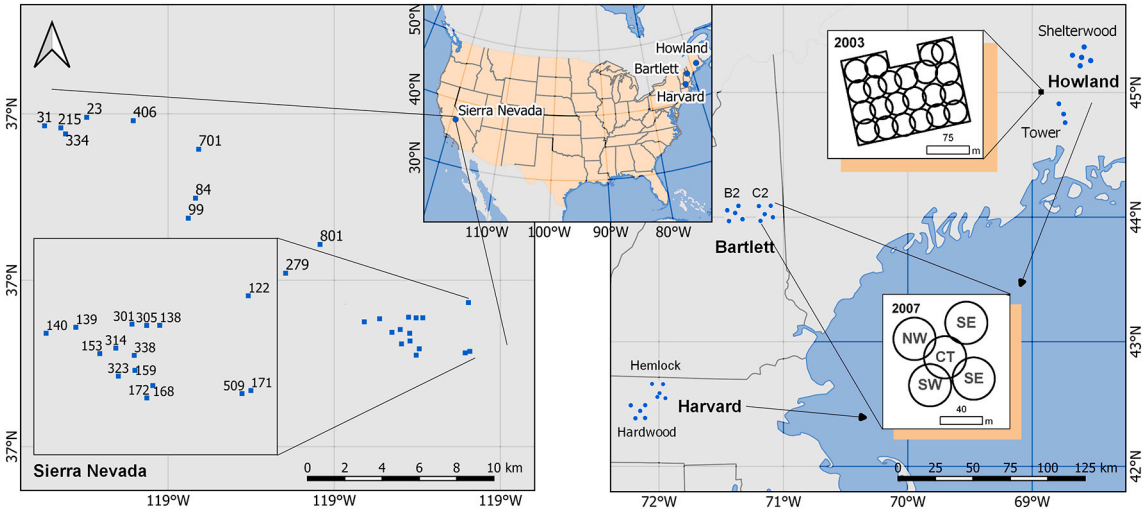


Fig. 2. Locations of test sites - The New England (NE) sites are Harvard Forest, MA, Bartlett Experimental Forest, NH, and the Forest Ecosystem Research site in Howland, ME. Two forest stands were selected for each NE site, with six stands in total; and each forest stand encompasses five (only three for the Howland tower stand) 20 m- or 25 m-circular radius plots with additional 20 plots extracted from the stem map data collected in 2003 in Howland, ME. The CA site in the Sierra Nevada includes 25 33 m × 33 m square plots.

diameters were recorded for each live stem, excluding stems with a diameter at breast height (DBH) < 10 cm. We measured tree heights and horizontal crown radii for the selected trees. Additionally, a complete stand map was assembled at the Forest Ecosystem Research site in Howland, ME, in 2003 (see Fig. 2 for the locations of the stem map area). We extracted 20 circular plots with a 20-m radius to be comparable with the other field data. In combination with the 2007 and 2003 datasets, 48 plots from NE sites were used in this study.

The CA data were collected in 2008, similarly to the NE sites. Twenty-five one-hectare stands were selected. Each stand was divided into nine (three × three) 33 m × 33 m square plots. Species information, trunk diameter, and tree height were collected for all live stems, excluding stem diameter D < 10 cm. No single crown radius was measured for the CA site. We used the data collected in the center plots for this study due to the lack of height measurements in other plots.

We estimated the aboveground biomass for each tree using species-specific allometric equations and then aggregated them into the plot level. We chose allometric equations to best approximate local conditions. The aboveground biomass of the NE sites was calculated based on DBH-biomass allometric equations (Tritton and Hornbeck, 1982). For the CA site, we used the allometric equations of AGB based on both DBH and tree height, developed by the United States Department of Agriculture (USDA) Forest Service (Waddell and Hiserote, 2005). Tree heights were calculated for the NE sites.

3.3. LVIS data and LVIS data processing

The LVIS flew over the NE sites in 2009 and the CA site in 2008. LVIS, a GEDI airborne sensor, is an airborne laser altimeter system emitting laser energy at the 1064 nm wavelength at a footprint size of 20 or 25 m. Both datasets have an above 0–5° field of view and a footprint size of 20 m. The LVIS and field data for the CA site were collected in the same year (2008). However, LVIS and the field data share two-year (2009 vs. 2007) and six-year (2009 vs. 2003) gaps for NE sites. Considerable variation in the relationship between field data and LVIS data in NE sites could be caused by a significant time lap.

To extract LVIS data for the corresponding plots, we used LVIS data with footprint centers located within a 25 m radius circular area from each plot center so that the selected LVIS observation could occupy an area slightly larger than the plot range to reduce geolocation uncertainty (Ni-Meister et al., 2010a). We normalized each waveform in each plot first, and then extracted canopy gap probability and apparent foliage density profiles using the approach presented in (Ni-Meister et al., 2001), which provided direct inputs for biomass index calculation.

3.4. Calculating biomass indices and comparing them with field AGBD

We first calculated both waveform-based and foliage-based biomass indices based on Eqs. (12) and (13). We then aggregated footprint-level biomass indices into the plot level through arithmetic averaging. We adopted different height scaling exponent values, ranging between 1.0 and 3.0. Plot-level biomass indices were compared with field-measured AGBD in NE and CA sites. Lidar height metrics (RH 75 and RH50) were compared with field-measured AGBD. A linear regression of RH100, RH75, and RH50 was developed to estimate AGBD.

4. Evaluation results

This section presents a detailed analysis to better understand the model and its model performance. We first compared the traits of aboveground biomass, tree cover, tree height, and topographic conditions in these two biomes to understand the different vegetation structure characteristics in our sites. Next, as we discussed before, a waveform is a good indicator of tree height and size distribution. We examined the relationships between waveforms and tree height distribution patterns for both the NE and CA sites. Due to the lack of data, we could not add crown size to this analysis. We expect that both sites are one cohort (one layer) canopy, the tree size does not vary so much within each plot. Next, we attempted to extract the height scaling exponent in the model based on available field-measured vegetation structural data from NE sites. We used height and AGB measurements to derive and evaluate the height scaling exponent estimation for both the NE and CA sites. Lastly, we discussed our biomass indices evaluation results.

4.1. Comparison of different forest stand conditions

The whisker plots of the plot-level AGBD, vegetation cover, maximum vegetation height (RH100), and surface topography (slope) (Fig. 3) illustrate the significant differences in vegetation structure characteristics and environmental conditions between the NE and CA sites. Vegetation cover *(1 - P_{gap})*, where P_{gap} calculated using LVIS data indicates forest density. RH100, directly extracted from the LVIS data product, may be linked to the maximum vegetation height within each plot. Vegetation in the NE sites was relatively uniform. AGBD (first and third quantiles) for the NE sites ranged from roughly 200–300 Mg/ha with minimal variation. Most NE sites had full canopy cover (> 90%), but maximum tree heights for each plot varied 20–30 m. All the NE sites were flat, with a slope < 3^0. The NE sites were dense and uniform forests, with high canopy cover and minimal variations in AGBD.

In contrast to the NE sites, the CA site was composed of a complex forest structure with significant canopy cover and tree height variations. AGBD varied from 100 to 1500 Mg/ha. The Giant Sequoia plot had an AGBD >1500 Mg/ha, but some plots were very sparse, with AGBD <100 Mg/ha. Some plots were multi-layer canopies with 70–80 m in height and some small trees underneath. Vegetation was sparse for most CA plots; vegetation cover ranged from 20 to 80%, with a median of 45%. The maximum height for each plot in the CA site varied from 20 to 80 m. The slope of the CA site varied significantly, from 3.4^0 to $24.4^{0,}$ with an average slope of 9.6^0.

4.2. Relationship between lidar waveforms and tree height distribution

To demonstrate the relationships of waveforms with height distribution, Fig. 4 illustrates the similarities between tree height distribution and the corresponding waveforms and vertical foliage profile for the 28 NE plots sampled in 2007. The 20 plots of the 2003 stem map had a vegetation structure similar to the Tower stand and were not included here. As discussed before, not all trees had tree height measurements. The tree heights displayed here were calculated based on a plant functional type (PFT) and DBH and height allometric equations (Albani et al., 2006). These heights match the field measured tree heights adequately (Fig. 6). In Fig. 4, the tree height distribution refers to the number of trees per 1 m vertical grid at the corresponding height for each 20 m or 25 m radius circular plot. Total tree count density (# tree counts/ha) and AGBD are included for each plot.

The vertical tree height distribution showed distinct features for the three NE sites. We discussed their differences as follows. The vegetation structure was quite similar in the Bartlett B2 and C2 stands. Both stands had trees with different stand ages from mature trees (dominant and codominant trees) to intermediate and young stands (tree height < 10 m) (Fig. 4, rows 1–2), except for the C2-NW plot with many small trees. The features suggest that the B2 and C2 stands were very dense forests with different tree stand ages and most tree densities >800/ha.

Waveforms and foliage profiles for Bartlett B2 and C2 stands have similar features to tree height distribution. The waveforms in each plot have a relatively symmetric vertical pattern along the middle of the canopy. For the plots with vertically evenly distributed tree heights (C2-CT, C2-NE, C2-NW, and C2-NE), the vertical foliage profiles were featured with evenly distributed foliage vertically from the lower to the upper canopy, corresponding to vertically evenly distributed tree height distributions. Some waveforms in C2-CT, C2-NE, and C2-SE suggest that

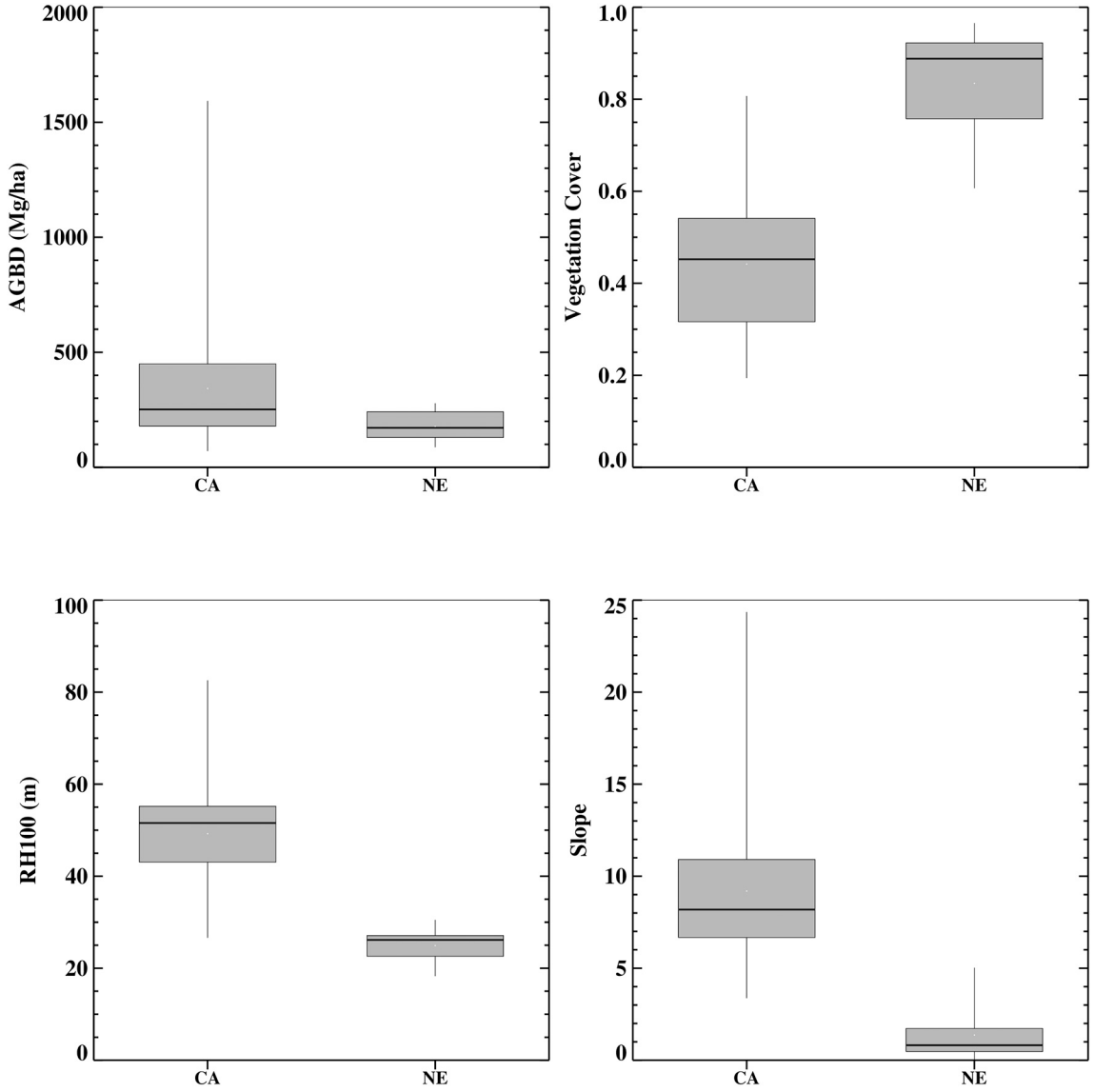


Fig. 3. Whisker plots of aboveground biomass density, vegetation cover, tree height, and the topographic slope for both the CA and NE sites. Both plot-level tree height and vegetation cover were estimated using LVIS data.

maximum tree heights should be taller than modeled maximum tree heights, indicating that tree height variability was suppressed by modeled tree heights. One waveform in each C2-CT, C2-NE, C2-SE, and C2-SW plot has only ground returns and no vegetation returns. Overall, waveforms and foliage profiles corresponded reasonably well with tree height distribution for all B2 and C2 plots.

Compared to the B2 and C2 stands, the two Harvard stands, Hardwood and Hemlock had relatively different vegetation structure patterns. Both stands were less dense (around 500–600/ha) (Fig. 4, rows 3–4) than the B2 and C2 stands. However, both were dominant/codominant forest stands, with a few small trees in each plot. The waveforms and foliage profiles of both stands echo similar features. Peak laser returns in waveforms at the top of the canopy indicate that both stands were featured by dominant/codominant and intermediate trees, agreeing with what was observed in the field data. Foliage profiles with most foliage were at the upper canopy layer and decreased downward. Both peaks in waveforms and foliage profiles in these two stands are closer to the canopy top, different from what was observed in the B2 and C2 stands with uniform foliage distribution from the lower to the upper part of the canopy layer. This feature reinforces that most trees in these two Harvard stands were mature dominant and codominant trees, corresponding to tree height distributions.

Two conifer stands in Howland had distinct features in tree height distribution and corresponding waveforms and foliage profiles (Fig. 4, rows 5–6). The sparest stand, Shelterwood, was a recently-harvested conifer forest, with a tree count density of <716/ha of various small and intermediate trees. Tree heights were almost evenly distributed vertically. Most were <20 m for all five plots. In contrast, the Tower stand, the densest forest stand for all the NE sites, had a tree count density above 1600/ha with many intermediate and young trees for all three plots. Most tree heights were between 10 and 15 m.

The corresponding waveforms and foliage profiles had similar patterns to the height distribution. The Tower stand had uniform waveforms with the peak canopy returns in the middle of the canopy. The foliage profile in the Tower stand indicated a peak foliage density in the middle of the canopy. In contrast, waveforms in Shelterwood had weaker canopy returns, but relatively more robust ground returns than the Tower or other NE stands. The foliage density profiles in Shelterwood had much less foliage than in the other stands. Also, at least half of the waveforms in each Shelter plot exhibited significant ground returns with minimal returns from vegetation, mainly due to cutting.

Fig. 5 illustrates the similarity of tree height distribution and

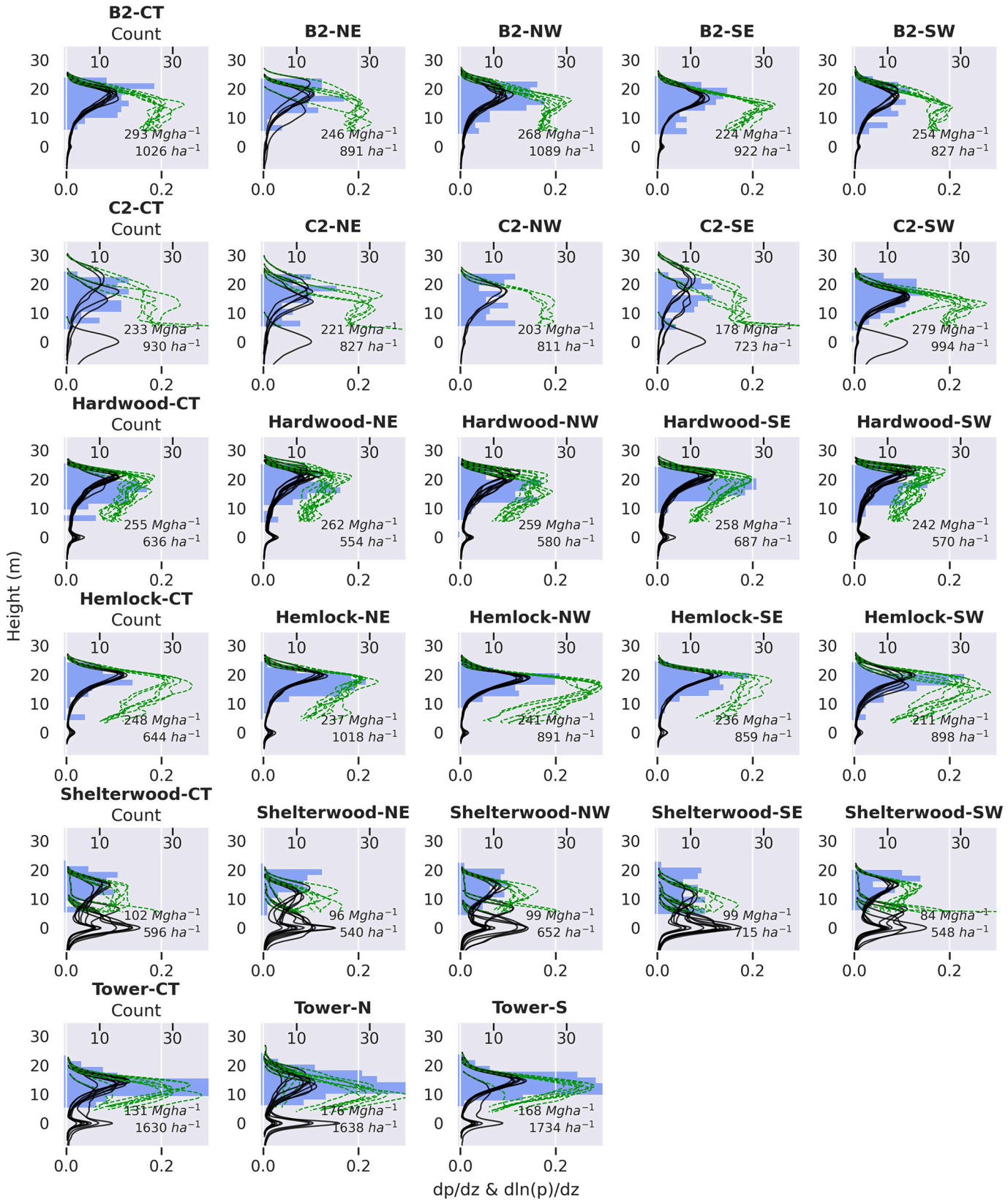


Fig. 4. Overlay of normalized LVIS waveforms (black), apparent foliage profiles (green), and tree height distribution (tree counts per 1 m vertical grid) in each of the 20 m/25 m diameter circular plots in NE sites. The title regards the stand name followed by the plot name. Aboveground biomass density (Mg/ha) and total tree count density (tree counts/ha) for each plot are marked in the lower right corner of each panel. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

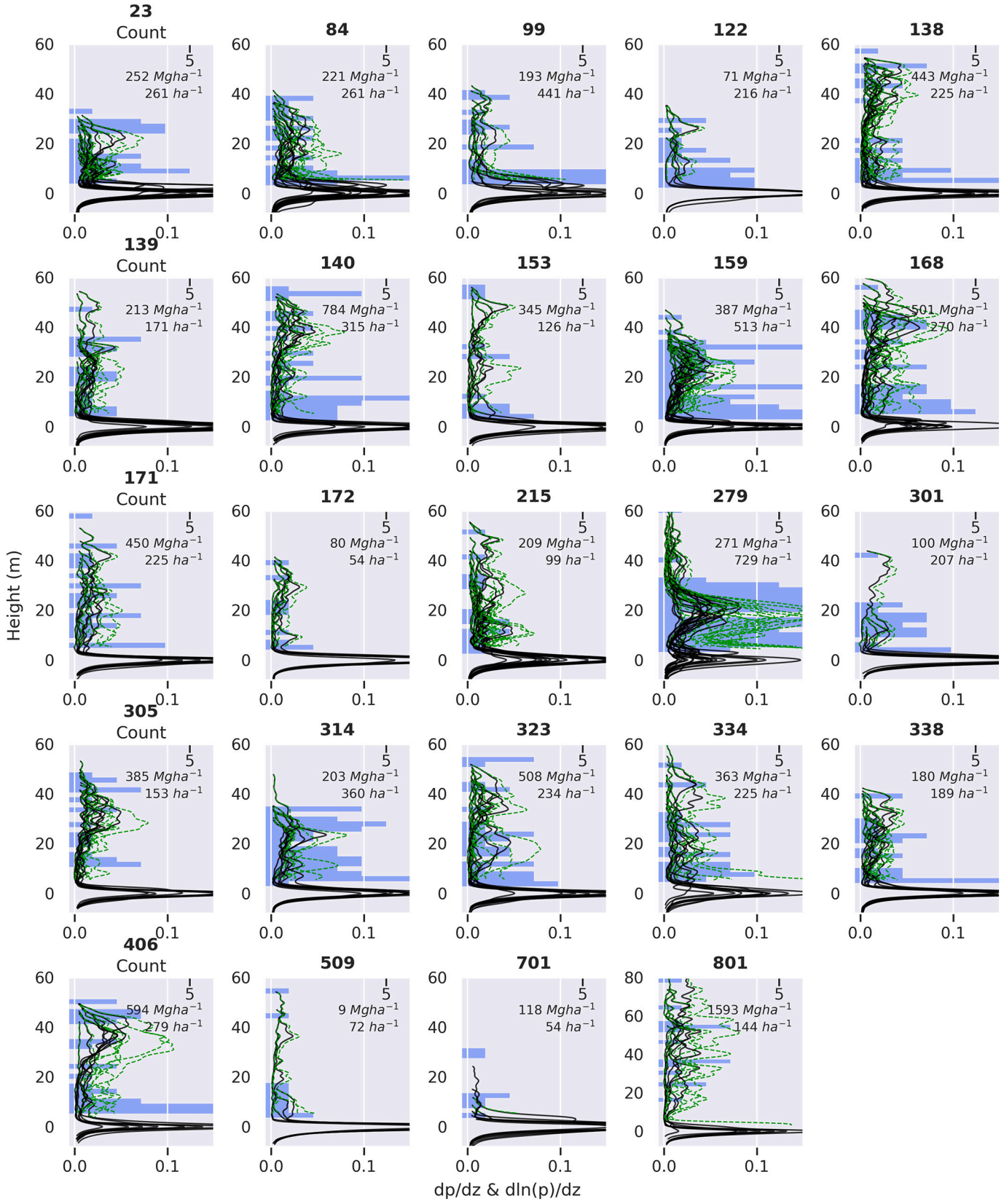


Fig. 5. Overlay of normalized LVIS waveforms (black), apparent foliage profiles (green), and tree height distribution (tree counts per 1 m vertical grid) for the 25 plots of the CA site. The title regards the plot number. Aboveground biomass density (Mg/ha) and total tree count density (tree counts/ha) for each plot are marked in the upper right corner of each panel. (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

waveforms and foliage profiles for the 25 plots of the CA site. Note that the tree height distribution refers to the vertical distribution of tree count per 1 m vertical interval for each 30 m × 30 m horizontal grid area. Total tree count density (# tree counts/ha) and AGBD for each plot were included.

The CA site was much sparser, and had taller trees comparing to the NE sites. Most plots had a tree density between 100 and 300 trees/ha, except for a few dense plots (plot #279, #159, #99, and #314). Some very sparse plots (plots # 701, #509, #172, and # 215) had tree count density <100 trees/ha. Many plots were mixed with very tall dominant/codominant trees (tree height above 40 m and intermediate and young trees (plot #138, #140, #153, #168, #171, #305, #334, #406, and #801). Some plots (plot # 23, #31, #84, #99, #122, #159, #172, #279, #301, #314, #338, #509, and #701) were featured with intermediate and young short trees (tree height < 40 m). The densest plots had mostly young trees: plot #279 with 729 trees/ha with most tree height <30 m, #159 (tree density 513 trees/ha) with many small and intermediate trees, #99 (tree density 441 trees/ha), and #314 (tree density 360/ha). Tree height for most plots had much larger variations than the NE sites.

The corresponding waveforms and foliage density profiles of the plots reflected similar vegetation structure characteristics. There were three distinct features: First, waveforms had weaker peak canopy returns and more robust ground returns for most CA plots, indicating sparse forests. One exception is plot # 279, with similar vegetation and ground peak returns. Compared to the NE stands, the differences between the waveform and foliage profiles were relatively small due to sparseness. Second, the vegetation structure had significant variations among the plots. Tree height ranged from 5 to 10 m (plot # 509 and #701) to 80 m (plot # 801). Many waveforms indicated taller trees (> 40 m) in the CA plots, while the NE plots had a maximum height lower than 30 m. However, waveforms had significantly short trees or almost no vegetation returns for some CA plots (i.e., plots # 701 and # 509) with a total canopy gap probability of 0.93 and 0.81, respectively. Some had mixed tall and short trees within plots. These features were consistent with what was observed in the tree distribution profiles. Finally, some plots were featured with various waveforms and foliage profiles.

Canopy peak returns in waveforms within a plot occurred at different heights, and were less uniform than the NE plots. For example, plot #301 had one waveform with a canopy peak at around 40 m and one waveform with a canopy peak at 10 m corresponding to one tall tree (40 m) and many young trees as observed in tree height profiles. Overall, the features shown in the waveforms and foliage profiles were consistent with the tree distribution profiles. This similarity between waveforms/foliage profiles and tree height distribution supported our hypothesis that waveforms and foliage profiles are good indicators of tree height distribution. The similarity of vertical tree distributions with waveforms and foliage profiles was in agreement with (Stark et al., 2015; Stark et al., 2012). They found that vertical foliage profiles and waveforms were associated with tree size distribution and forest demography in tropical Amazon forests.

4.3. AGB - height allometric relationships

The AGB-height scaling exponent present in our model could be estimated using two approaches: one is using the formula we derived in our model, $1+\frac{2}{\beta}-\propto$ which is based on the allometric relationships of height with stem diameter (β) and crown volumes with tree height ($\propto$). It could also be estimated based on the AGB and height measurements if available. To verify the validity of height and aboveground biomass allometric relationships presented in our model, we first attempted to analyze the relationships of tree height with stem diameter and crown volume with tree height to derive the height scaling exponent, $1+\frac{2}{\beta}-\propto$ with available sampled tree height, stem diameter, and crown diameter measurements in the NE sites. Then we analyzed the AGB - tree height allometric relationships in both NE and CA sites based on the calculated AGB and the tree height measurements.

The natural logarithms of tree height and stem diameter show a near-linear relationship, with $R^2 = 0.66$ and scaling exponent $\beta = 0.61$ (Fig. 6, top left panel). However, it seems that the linear relationship varies at young and mature trees. The natural logarithms of the horizontal crown radius and stem diameter imply a relatively weak linear relationship, with $R^2 = 0.14$ (Fig. 6, top right panel). The NE sites were mixed forests, half deciduous forests, and half coniferous forests. Crown shape differences between these two forest types might cause this weak relationship. The natural logarithms of the crown volume and tree height relationship were more robust, with $R^2 = 0.39$ and the scaling exponent $\alpha = 1.89$ (Fig. 6, middle left). The AGB - height scaling exponent was estimated to be $1+\frac{2}{\beta}-\propto = 2.39$. The natural logarithms of the AGB and tree height also indicates a near-linear relationship, with slightly different slopes for young and mature trees. The suggested $1+\frac{2}{\beta}-\propto = 2.56$ for the NE sites (Fig. 6, middle right panel). These two values were similar. Verifying the formula for the height scaling exponent was valid.

For the CA site, at the individual tree level, logarithms of tree height and stem diameters showed a linear relationship with $R^2 = 0.64$ and the scaling exponent $\beta = 0.84$ (Fig. 6, bottom left panel). The natural logarithms of AGB and H implied a near-linear relationship, with slightly different slopes for young and mature trees. The AGB - tree height scaling exponent $1+\frac{2}{\beta}-\propto = 2.32$ (Fig. 6, bottom right). This value was within the range of the field measurements (2.2–2.7) in cascade Oregon (Agee, 1981).

This analysis demonstrated that the height scaling exponent calculated based on stem diameter and crown volume with allometric relationships agrees well with the field-measured AGB and height scaling exponent for the NE site. Our analysis also indicated that the height scaling exponent values were close for these two test regions/biomes. In the following section, we evaluated the performance of the model using a scaling exponent of 2.4.

4.4. Relationship between aboveground biomass and biomass indices

With height scaling exponent value c = 2.4, we calculated two biomass indices from each waveform, aggregated them at the plot level, and compared them to field measured AGBD for the NE sites, the CA sites, and the combined NE and CA sites, respectively (Fig. 7). Fig. 7 shows that both indices had strong links with AGBD for all the test cases. For the NE sites, the waveform-based and foliage-based biomass indices had coefficients of determination (R^2) of 0.81 and 0.80, root mean squared error (RMSE) of 27.08 Mg/ha and 27.96 Mg/ha, and normalized RMSE based on the mean observed AGBD (NRMSE) of 0.15 and 0.16, respectively. The CA site has higher R^2 values (0.88 for both indices), higher RMSE (108–109 Mg/ha for both indices) and NRMSE (0.31 for both indices). When combined with NE and CA sites, R^2 values reached 0.9 and 0.87 for waveform-based and foliage-based biomass indices, respectively, higher than the NE and CA sites alone. RMSE values (66 Mg/ha and 73 Mg/ha for both indices) and NRMSE (0.28 and 0.31 for both indices) were lower than the CA site alone. Both biomass indices had similar performances for the NE and CA sites alone and the combined CA+NE data.

We also compared the performance of biomass indices with lidar height metrics, including RH 75 and RH50, and linear regression of RH100, RH75, and RH50 for the AGBD estimate. For the NE sites, R^2 values for RH75, RH50, and the regression of RH100, RH75, and RH50 were 0.82, 0.84, and 0.84, respectively. The corresponding RMSE values were 26 Mg/ha, 25 Ma/ha, and 25 Mg/ha, and NRMSE values were 0.15, 0.14, and 0.12. Overall a regression of RH100, RH76, and RH50 had the best performance. Both biomass indices had similar performance with height metrics.

However, biomass indices performed much better than height metrics for CA sites. R^2 values for RH75, RH50, and the regression of RH100,

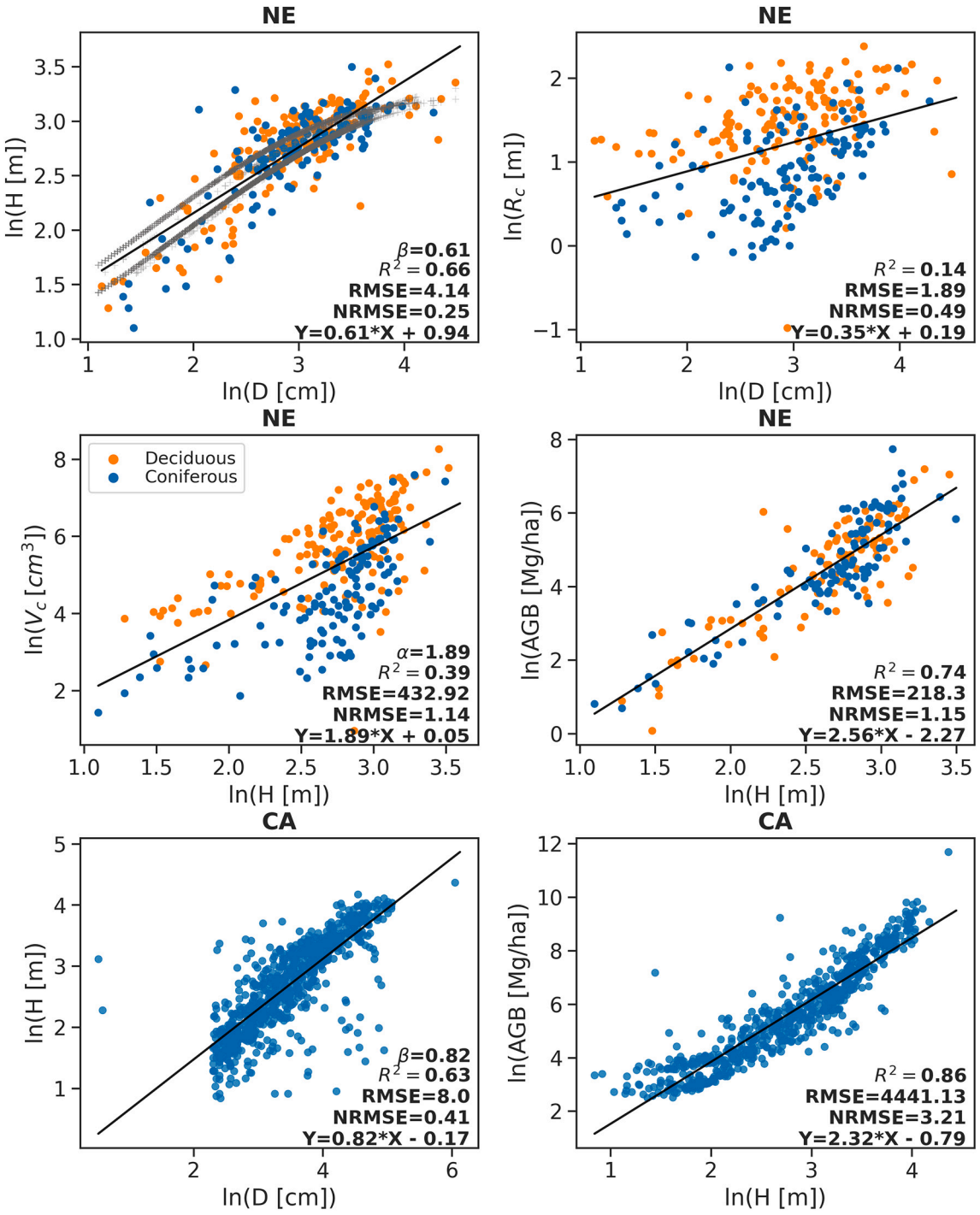
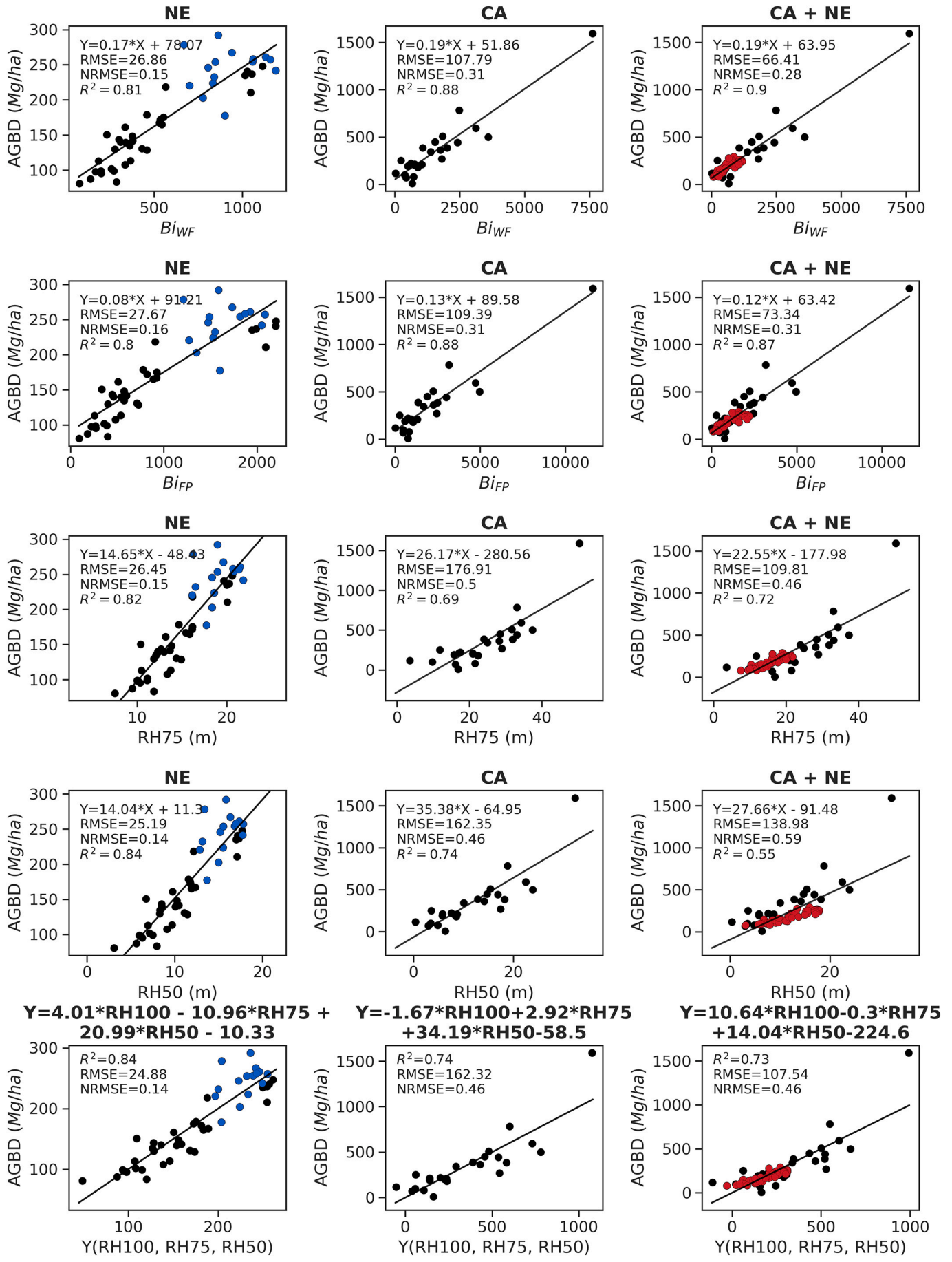


Fig. 6. Allometric relationships of tree height (H) with stem diameter (D), crown radius (R_c) with stem diameter, crown volume with tree height, and AGB with height (H) for the NE sites (top two rows), and tree height (H) with stem diameter (D) and AGB with height in the CA site (bottom panel). Top left panel also includes the modeled tree heights (in grey).

RH75, and RH50, were 0.69, 0.74, and 0.74, respectively, much lower than 0.88 for biomass indices. RMSE values were 177 Mg/ha, 162 Mg/ha, and 162 Mg/ha, much larger than 110 Mg/ha for biomass indices, and NRMSE values were 0.5, 0.46, and 0.46 comparing to NRMSE = 0.32 for biomass indices. Both biomass indices outperformed all the height metrics. Height metrics underestimated aboveground biomass, mainly in regions with large biomass. Both biomass indices overcome this underestimation problem.

When combining all the plots (CA + NE), both biomass indices outperformed height metrics (R^2 = 0.89 and RMSE = 68 Ma/ha and 73 Mg/ha for biomass indices vs. R^2 = 0.72, 0.55, and 0.73, RMSE = 110, 139, and 107 Mg/ha, and NRMSE = 0.46, 0.59, and 0.46 for RH75, RH50 and regression of (RH100, RH75, and RH50). Height metrics tended to underestimate AGBD at large biomass values. The NE and CA sites have significant spatial variations of vegetation characteristics, aboveground biomass, and environmental conditions. These results suggested both biomass indices show a considerable advantage in estimating aboveground biomass at large spatial scales across significant variations of aboveground biomass values.

We calculated biomass indices with a range of scaling exponent



(caption on next page)

Fig. 7. Comparison of relationships of field-measured aboveground biomass density (AGBD) with waveform-based biomass index (BI_{WF}, top row) and with foliage profile-based biomass index (BI_{FP}, second row), RH75 (third row), RH50 (fourth row), and linear regression of (RH100, RH75, and RH50) (last row) for the NE sites (first column, deciduous plots in blue), the CA site (middle column) and the combined NE and CA sites (last column, the NE plots in red). (For interpretation of the references to colour in this figure legend, the reader is referred to the web version of this article.)

values varying from 1.0 to 3.0 at 0.1 intervals to assess the optimal height scaling exponent for all sites. We elucidated the model performance (R^2 and RMSE) with different scaling exponent values and compared it to height metrics (Fig. 8). Both indices performed similarly in all the scaling exponent values we tested for the NE sites, with R^2 ranging from 0.84 to 0.79 and RMSE from 25 to 28 Mg/ha. At relatively lower height scaling exponent values (1.0–1.4), both biomass indices reached the best performance ($R^2 = 0.84$ and RMSE = 25 Mg/ha) and were similar to the best of all height metrics ($R^2 = 0.84$ and RMSE = 25 Mg/ha with the regression of RH100. RH75 and RH50). Both indices had a similar performance for the CA site, with the best performance when the scaling exponent value was >2.3. Both biomass indices performed more adequately than height metrics (higher R^2 and lower RMSE). For the combined NE and CA cases, both indices showed significant advantages over height metrics for scaling exponent values >1.8. The performance of both indices reached a plateau when the scaling exponent value was >2.4, with the highest R^2 (0.90) and lowest RMSE (64 Mg/ha). Reasonably good AGBD estimates with a general height scaling exponent value (1.8) were achieved by the model. The waveform-based biomass index correlates well with AGBD, with $R^2 = 0.83$ and RMSE = 84 Mg/ha.

5. Discussion

Our model directly used large-footprint lidar waveform measurements to estimate plot-level aboveground biomass density. We associated tree height and crown size distribution with waveforms and scaled up individual tree-based AGB and height allometric equations to plot level. Plot-level aboveground biomass density was estimated using waveform /foliage-weighted height-based allometric equations. The AGBD-height scaling exponent varies with the allometric height-stem diameter and crown volume-height relationships.

We adopted a global database of 108,753 tree measurements of stem diameter, height, and crown diameter to estimate the scaling exponent. The height scaling exponent values calculated using the global database range from 1.6 to 1.8 for all forest biomes except for boreal forests (0.9). Our analysis suggested that one general model works reasonably well across all global forest biomes except for boreal forests.

We applied the model in two distinct geographic regions—temperate deciduous/conifer forests in the northeastern US and a montane conifer forest in Sierra National Forest in California. The AGBD-height scaling exponents estimated using site-specific vegetation structure measurements for these two biomes/regions were close to each other (2.3–2.58). These scaling exponents estimated for both test sites agree well with the observations in conifer forests in the northwest US (Agee, 1981) (2.2–2.7). These results reinforced the generality of the model.

Our model produced the optimal AGBD estimates using the local AGBD-height scaling exponent value. Both biomass indices in our analysis correlated well with field-measured AGBD, reaching the best performance with a coefficient of determination $R^2 = 0.89$, RMSE = 66 Mg/ha, and normalized RMSE around 0.28 when combining both the NE and CA sites. The model with a general scaling exponent value provided reasonable AGBD estimates. The waveform-based biomass index was correlated with AGBD with RMSE = 84 Mg/ha, NRMSE = 0.36, and $R^2 = 0.83$. Our analysis implies that using the general height scaling exponent provides reasonably good estimates of AGBD. However, integrating local structure allometric relationships improved the predictive accuracy of the model.

Both biomass indices performed similarly regarding AGBD estimates in the NE and CA sites. The waveform-based biomass index performed slightly better in some cases than the foliage-based biomass index. The waveform-based index is straightforward to implement, which may give it broader applicability. The foliage profile-based index involves one additional step of deriving the canopy gap probability and a logarithmic

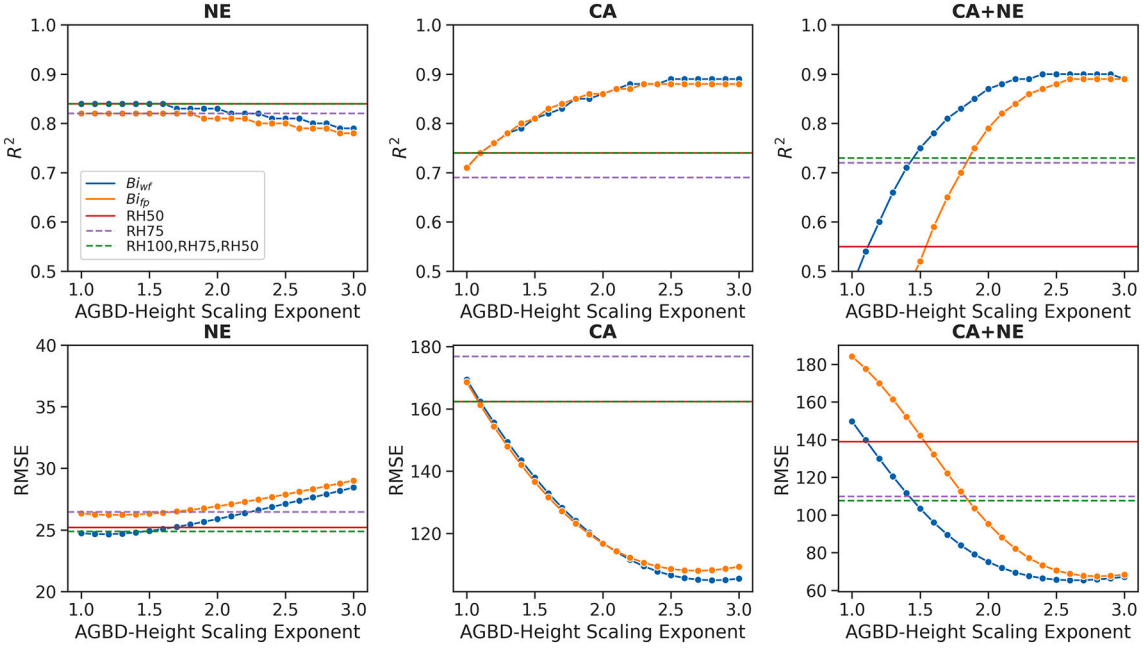


Fig. 8. Performance (Coefficient of determination - $R^2$and RMSE) of waveform-based and foliage-based biomass indices for predicting aboveground biomass with the height scaling exponent for the NE site, the CA sites, and combined NE and CA sites. Horizontal lines indicate the performance of RH50 and RH75 and the regression of RH100, RH75, and RH50.

transformation, which may introduce additional error sources. Further tests are required to evaluate their performance.

We compared the performance of biomass indices to the traditional lidar height metrics-based approach for AGBD estimation. Our analysis found that biomass indices performed similarly to the lidar height metrics-based method for low and medium-range biomass. However, both indices outperformed the height-based approach in high biomass regions. The height metrics-based approach underestimated aboveground biomass in high biomass regions (old-growth Sequoia forests), which was overcome by biomass indices. While removing the extreme large 1500 Mg/ha point will result in similar performance between biomass indices and height metrics. Theoretically, our scaling scheme provides a consistent framework for linking bottom-up demographic estimates of biomass and top-down lidar footprint estimates of biomass. A region with large biomass may not be as frequent as low to medium biomass regions. However, large biomass regions often make a more significant contribution to total biomass estimation due to their large biomass values. Old-growth Sequoia forests produce more aboveground biomass annually—and sequester more carbon than forests dominated by other species (Sillett et al., 2020). It is critical to have accurate biomass estimates for large biomass.

The model had two components—biomass indices and coefficients. Biomass indices directly connected lidar waveform measurements with the dominant structural variables—tree height and crown size distribution directly contributed to plot-level AGBD. This study mainly focused on model development and the initial assessment of the correlation of biomass indices with field-measured AGBD. Other parameters in the coefficients k_{FW} and $k_{FP,}$ such as tree taper, F, the wood-specific gravity, ρ, clumping factor, γ, and foliage volume density, F_a, also affect AGBD. In our study sites, the lack of specific field measurements of these parameters prevents us from thoroughly analyzing the impact of those variables on aboveground biomass. However, we are evaluating this model in the tropical forest in a separate study, including some of those variables, such as wood-specific gravity in our analysis.

Our initial assessment demonstrated that the biomass indices account for about 81%–89% of AGBD variability (Fig. 7). Therefore, the other 11–20% AGBD variability was contributed by the parameters in the coefficients k_{FW} and k_{FP} components, which were not universal. Tree taper, $F = 0.6$ for broadleaf species, and $F = 0.333$ for perfect conical shape (Chave et al., 2005). Wood specific gravity varies with species and plant functional types. The clumping factor, γ, varies with tree structure characteristics (shape, size, tree count density, and within-crown foliage density) and foliage volume density, F_a may vary with species and stand age.

We found the R^2 value in the montane CA conifer forest sites (0.89) was slightly larger than the one (0.81) in temperate NE mixed forests sites in this study. Likely, roughly 19% for the NE sites and 11% for the CA sites of AGBD uncertainties were contributed by the parameters in the coefficient term of the model. Conifer forests dominated the CA sites. However, the NE sites were mixed with coniferous and deciduous forests, and had more diverse species than the CA sites. The species and plant function diversities may result in more considerable variations in wood-specific gravity and wood taper for the NE plots than CA plots. Larger variations in the coefficient terms in the NE sites may explain the larger uncertainties of AGBD estimates using biomass indices only. It is critical to characterize the variability of these variables to improve the predictive accuracy of the model.

In addition, lidar saturation might occur in dense forests, as previously discussed (Ni-Meister et al., 2010a). All the NE stands were dense forests except Shelterwood, making waveforms less sensitive to different vegetation structure conditions and AGBD changes. As discussed before, one waveform in each C2-CT, C2-NE, C2-SE, and C2-SW plot had only ground returns and no vegetation returns. Maybe, in reality, these footprints had no vegetation. However, it was also possible that vegetation was too dense in these footprints for the ground detection algorithm to detect ground returns correctly. Very dense forests could make the lidar waveforms less sensitive to vegetation structure and less accurate AGBD estimates using lidar waveform measurements. The time gaps between the field data (2003, 2007) and the LVIS data (2009) might also contribute to the large AGBD uncertainties in the NE sites.

Future work will include testing the sensitivity of those parameters in coefficient terms and identifying the most critical structure parameters. However, many of these parameters were correlated, which poses a challenge for sensitivity analysis. More field data are required to assess the performance of the model fully.

6. Conclusion

The model presented in this study directly used large-footprint lidar waveform measurements to estimate plot-level aboveground biomass density. Individual tree-based AGB and height allometric equations were scaled up to plot-level using the tree height and crown size distribution characteristics measured by waveforms. The plot-level aboveground biomass density was estimated based on a waveform/foliage profile-weighted height-based allometric equation. The height scaling exponent in the model was built on the allometric relationships of tree height with stem diameter and crown volume with tree height. Initial assessment using a global database of 108,753 tree measurements of stem diameter, height, and crown diameter demonstrated that one general model (scaling exponent value ~1.6–1.8) works reasonably well across all global forest biomes except boreal forests (scaling exponent value ~0.9).

We applied the model to two different forest biomes in two distinct geographic regions—temperate deciduous/conifer forests in the northeastern USA and a montane conifer forest in Sierra National Forest in California. We found a consistent height scaling exponent for these two distinct biomes based on field vegetation structure measurements. Still, this value was slightly different from the global data analysis results. Using the local height scaling exponent value, our model produced optimal AGBD estimates. It Using the general heigh scaling exponent also provided reasonably good AGBD estimates. Our analysis demonstrated one general allometric relationship of plot-level AGBD with large-footprint lidar waveforms. Integrating local structure allometric relationships improved the predictive accuracy of the model.

We compared the performance of biomass indices to those using the traditional lidar height metrics-based approach for AGBD estimation. Our analysis showed that both indices outperformed the height-based approach in high biomass regions. The lidar height metrics-based approach underestimated aboveground biomass in high biomass regions, which was overcome by biomass indices. Our model could be applied on a large scale with various vegetation structure conditions to obtain high-accuracy aboveground biomass estimates.

This model directly links plot-level aboveground biomass using large-footprint lidar waveform measurements. This approach was appealing as it eliminated the need for site-based calibration efforts to estimate aboveground biomass using lidar waveform measurement. It reduced the number of steps (waveforms to RH values and RH values to AGBD with site-based calibration) for estimating AGBD and mitigating the uncertainties caused at each step. This model can potentially curtail the site-specific calibration effort for aboveground biomass estimates using medium-range lidar full-waveform measurements.

This model could potentially serve as a general and robust approach for monitoring forest carbon stocks using large-footprint lidar waveform measurements, such as the GEDI mission at the continental and global scales. Demography-based terrestrial ecosystem model simulates carbon stock and flux using ecosystem demography (tree size and height distribution). The model could be a framework for integrating demography terrestrial ecosystem models and ecosystem demography measurements from GEDI full-waveform measurements to improve global carbon stock and flux estimates.

Nomenclature

Roman alphabet

a	Coefficient of aboveground biomass and height relationship
b	coefficient of tree height and stem diameter relationship
D	*Stem diameter* - diameter at breast height
F	Tree tapering factor
F_a	Foliage volume density (m2/m3)
G	Leaf orientation function
H	Tree height
$L_e(z)$	Cumulative effective leaf area index from the canopy top to height z
n	Total number of height class
$P(0,z)$	Canopy gap probability at height z for nadir pointing angle
P_{gap}	Averaged canopy gap probability
r	Averaged horizontal crown radius
$R_v(z)$	Accumulated laser energy return from the canopy top to height Z
$R_v(0)$	Accumulated laser energy return for the whole canopy layer
R_g	Laser energy returns from the ground
z_1	Lower boundary of vegetation canopy
z_2	Upper boundary of vegetation canopy
z	Height in the canopy

Greek alphabet

α	Crown volume and tree height scaling factor and a is the coefficient.
β	Height and stem diameter scaling factor,
$\lambda(z)$	The accumulated tree crown density from the canopy top to tree height z.
$d\lambda_i(z_i)$	tree count density at height interval from z_i to z_{i+1},
$\frac{d\lambda(z)}{dz}$	The tree crown count density distribution function, i.e., tree crown count density per unit vertical interval, in the unit of $1/m^3$
ρ	Wood density (g/cm^3)
γ	Clumping factor ρ_v Volume backscattering coefficient of a canopy element
ρ_g	Backscattering coefficient of the ground

CRediT authorship contribution statement

Wenge Ni-Meister: Conceptualization, Methodology, Formal analysis, Visualization, Writing – original draft, Writing – review & editing. **Alejandro Rojas:** Software, Validation, Formal analysis, Visualization, Writing – review & editing. **Shihyan Lee:** Conceptualization, Methodology, Software, Formal analysis, Writing – original draft.

Declaration of Competing Interest

The authors declare that they have no known competing financial interests or personal relationships that could have appeared to influence the work reported in this paper.

Acknowledgments

We thank Jucker et al. (2017) for sharing their compiled global field data at the individual tree level, which helped assess our model on a global scale. We are grateful to the research groups from Boston University led by A. H. Strahler, the University of Maryland led by R. Dubayah, and NASA Goddard Space Flight Center led by J. Ranson and G. Sun for collecting and sharing the field data used in this study. The final derivation for both biomass indices was the result from a fruitful discussion with David L. B. Jupp, who carefully checked the derivation process and helped finalize the AGBD formula for multi-cohort canopies. Special thanks go to Nancy Kiang for editing and clarifying the manuscript. We are also thankful to three anonymous reviewers for spending their time reviewing this manuscript and for their valuable comments, which helped improve the quality of this paper tremendously. This work was funded by NASA, United States under contract # 80NSSC21K0194.

References

Agee, J.K., 1981. Biomass of Coniferous Understory Trees in Crater Lake National Park, Oregon. National Park Service, Cooperative Park Studies Unit College of Forest Resources University of Washington Seattle, Washington, p. 98195.

Albani, M., Medvigy, D., Hurtt, G.C., Moorcroft, P.R., 2006. The contributions of land-use change, CO2 fertilization, and climate variability to the eastern US carbon sink. Glob. Chang. Biol. 12, 2370–2390.

Anderson, J., Martin, M.E., Smith, M.L., Dubayah, R.O., Hofton, M.A., Hyde, P., Peterson, B.E., Blair, J.B., Knox, R.G., 2006. The use of waveform lidar to measure northern temperate mixed conifer and deciduous forest structure in New Hampshire. Remote Sens. Environ. 105, 248–261.

Anderson, J.E., Plourde, L.C., Martin, M.E., Braswell, B.H., Smith, M.-L., Dubayah, R.O., Hofton, M.A., Blair, J.B., 2008. Integrating waveform lidar with hyperspectral imagery for inventory of a northern temperate forest. Remote Sens. Environ. 112, 1856–1870.

Bar-On, Y.M., Phillips, R., Milo, R., 2018. The biomass distribution on earth. Proc. Natl. Acad. Sci. U. S. A. 115, 6506–6511.

Chave, J., Andalo, C., Brown, S., Cairns, M.A., Chambers, J.Q., Eamus, D., Folster, H., Fromard, F., Higuchi, N., Kira, T., Lescure, J.P., Nelson, B.W., Ogawa, H., Puig, H., Riera, B., Yamakura, T., 2005. Tree allometry and improved estimation of carbon stocks and balance in tropical forests. Oecologia 145, 87–99.

Chave, J., Réjou-Méchain, M., Búrquez, A., Chidumayo, E., Colgan, M., Delitti, W.C., Duque, A., Eid, T., Fearnside, P., Goodman, R., Henry, M., Martínez-Yrízar, A., Mugasha, W., Muller-Landau, H., Mencuccini, M., Nelson, B., Ngomanda, A., Nogueira, E., Ortiz-Malavassi, E., Pélissier, R., Ploton, P., Ryan, C., Saldarriaga, J., Vieilledent, G., 2014. Improved allometric models to estimate the aboveground biomass of tropical trees. Glob. Chang. Biol. 20, 3177–3190.

Choi, S., Ni, X.L., Shi, Y.L., Ganguly, S., Zhang, G., Duong, H.V., Lefsky, M.A., Simard, M., Saatchi, S.S., Lee, S., Ni-Meister, W., Piao, S.L., Cao, C.X., Nemani, R.R., Myneni, R. B., 2013. Allometric scaling and resource limitations model of tree heights: part 2. Site based testing of the model. Remote Sens. 5, 202–223.

Coops, N.C., Tompalski, P., Goodbody, T.R.H., Queinnec, M., Luther, J.E., Bolton, D.K., White, J.C., Wulder, M.A., van Lier, O.R., Hermosilla, T., 2021. Modelling lidar-derived estimates of forest attributes over space and time: A review of approaches and future trends. Remote Sens. Environ. 260.

Drake, J.B., Dubayah, R.O., Clark, D.B., Knox, R.G., Blair, J.B., Hofton, M.A., Chazdon, R. L., Weishampel, J.F., Prince, S.D., 2002. Estimation of tropical forest structural characteristics using large-footprint lidar. Remote Sens. Environ. 79, 305–319.

Dubayah, R.O., Sheldon, S.L., Clark, D.B., Hofton, M.A., Blair, J.B., Hurtt, G.C., Chazdon, R.L., 2010. Estimation of tropical forest height and biomass dynamics using lidar remote sensing at La Selva, Costa Rica. J. Geophys. Res. Biogeosci. 115.

Dubayah, R., Blair, J.B., Goetz, S., Fatoyinbo, L., Hansen, M., Healey, S., Hofton, M., Hurtt, G., Kellner, J., Luthcke, S., Armston, J., Tang, H., Duncanson, L., Hancock, S., Jantz, P., Marselis, S., Patterson, P.L., Qi, W., Silva, C., 2020. The global ecosystem dynamics investigation: high-resolution laser ranging of the Earth's forests and topography. Sci. Remote Sens. 1.

Duncanson, L.I., Dubayah, R.O., Enquist, B.J., 2015. Assessing the general patterns of forest structure: quantifying tree and forest allometric scaling relationships in the United States. Glob. Ecol. Biogeogr. 24, 1465–1475.

Duncanson, L., Neuenschwander, A., Hancock, S., Thomas, N., Fatoyinbo, T., Simard, M., Silva, C.A., Armston, J., Luthcke, S.B., Hofton, M., Kellner, J.R., Dubayah, R., 2020. Biomass estimation from simulated GEDI, ICESat-2 and NISAR across environmental gradients in Sonoma County, California. Remote Sens. Environ. 242.

Feldpausch, T.R., Banin, L., Phillips, O.L., Baker, T.R., Lewis, S.L., Quesada, C.A., Affum-Baffoe, K., Arets, E.J.M.M., Berry, N.J., Bird, M., Brondizio, E.S., de Camargo, P., Chave, J., Djagbletey, G., Domingues, T.F., Drescher, M., Fearnside, P.M., França, M. B., Fyllas, N.M., Lopez-Gonzalez, G., Hladik, A., Higuchi, N., Hunter, M.O., Iida, Y., Salim, K.A., Kassim, A.R., Keller, M., Kemp, J., King, D.A., Lovett, J.C., Marimon, B. S., Marimon-Junior, B.H., Lenza, E., Marshall, A.R., Metcalfe, D.J., Mitchard, E.T.A., Moran, E.F., Nelson, B.W., Nilus, R., Nogueira, E.M., Palace, M., Patiño, S., Peh, K.S. H., Raventos, M.T., Reitsma, J.M., Saiz, G., Schrodt, F., Sonké, B., Taedoumg, H.E., Tan, S., White, L., Wöll, H., Lloyd, J., 2011. Height-diameter allometry of tropical forest trees. Biogeosciences 8, 1081–1106.

Feldpausch, T.R., Lloyd, J., Lewis, S.L., Brienen, R.J.W., Gloor, M., Monteagudo Mendoza, A., Lopez-Gonzalez, G., Banin, L., Abu Salim, K., Affum-Baffoe, K., Alexiades, M., Almeida, S., Amaral, I., Andrade, A., Aragão, L.E.O.C., Araujo Murakami, A., Arets, E.J.M.M., Arroyo, L., Aymard, C.G.A., Baker, T.R., Bánki, O.S., Berry, N.J., Cardozo, N., Chave, J., Comiskey, J.A., Alvarez, E., de Oliveira, A., Di Fiore, A., Djagbletey, G., Domingues, T.F., Erwin, T.L., Fearnside, P.M., França, M.B., Freitas, M.A., Higuchi, N., Iida, Y., Jiménez, E., Kassim, A.R., Killeen, T.J., Laurance, W.F., Lovett, J.C., Malhi, Y., Marimon, B.S., Marimon-Junior, B.H., Lenza, E., Marshall, A.R., Mendoza, C., Metcalfe, D.J., Mitchard, E.T.A., Neill, D.A., Nelson, B.W., Nilus, R., Nogueira, E.M., Parada, A., Peh, K.S.H., Pena Cruz, A., Peñuela, M.C., Pitman, N.C.A., Prieto, A., Quesada, C.A., Ramírez, F., Ramírez-Angulo, H., Reitsma, J.M., Rudas, A., Saiz, G., Salomão, R.P., Schwarz, M., Silva, N.,

Silva-Espejo, J.E., Silveira, M., Sonké, B., Stropp, J., Taedoumg, H.E., Tan, S., ter Steege, H., Terborgh, J., Torello-Raventos, M., van der Heijden, G.M.F., Vásquez, R., Vilanova, E., Vos, V.A., White, L., Willcock, S., Woell, H., Phillips, O.L., 2012. Tree height integrated into pantropical forest biomass estimates. Biogeosciences 9, 3381–3403.

Fischer, F.J., Labrière, N., Vincent, G., Hérault, B., Alonso, A., Memiaghe, H., Bissiengou, P., Kenfack, D., Saatchi, S., Chave, J., 2020. A simulation method to infer tree allometry and forest structure from airborne laser scanning and forest inventories. Remote Sens. Environ. 251.

Fisher, R.A., Koven, C.D., 2020. Perspectives on the future of land surface models and the challenges of representing complex terrestrial systems. J. Adv. Model. Earth Syst. 12.

Goldstein, A., Turner, W.R., Spawn, S.A., Anderson-Teixeira, K.J., Cook-Patton, S., Fargione, J., Gibbs, H.K., Griscom, B., Hewson, J.H., Howard, J.F., Ledezma, J.C., Page, S., Koh, L.P., Rockström, J., Sanderman, J., Hole, D.G., 2020. Protecting irrecoverable carbon in Earth's ecosystems. Nat. Clim. Chang. 10, 287–295.

Goodman, R.C., Philips, O.L., Baker, T.R., 2014. The importance of crown dimensions to improve tropical tree biomass estimates. Ecol. Appl. 24 (4), 680–698.

Henry, M., Besnard, A., Asante, W.A., Eshun, J., Adu-Bredu, S., Valentini, R., Bernoux, M., Saint-André, L., 2010. Wood density, phytomass variations within and among trees, and allometric equations in a tropical rainforest of Africa. For. Ecol. Manag. 260, 1375–1388.

Ishii, H., Ford, E.D., Sprugel, D.G., 2003. Comparative crown form and branching pattern of four coexisting tree species in an old-growth Pseudotsuga-Tsuga forest. Eur. J. For. Res. Hokkaido Univ. (Jpn.) 6 (2), 99–109.

Jucker, T., Caspersen, J., Chave, J., Antin, C., Barbier, N., Bongers, F., Dalponte, M., van Ewijk, K.Y., Forrester, D.I., Haeni, M., Higgins, S.I., Holdaway, R.J., Iida, Y., Lorimer, C., Marshall, P.L., Momo, S., Moncrieff, G.R., Ploton, P., Poorter, L., Rahman, K.A., Schlund, M., Sonke, B., Sterck, F.J., Trugman, A.T., Usoltsev, V.A., Vanderwel, M.C., Waldner, P., Wedeux, B.M., Wirth, C., Woll, H., Woods, M., Xiang, W., Zimmermann, N.E., Coomes, D.A., 2017. Allometric equations for integrating remote sensing imagery into forest monitoring programmes. Glob. Chang. Biol. 23, 177–190.

Ketterings, Q.M., Coe, R., van Noordwijk, M., Ambagau, Y., Palme, C.A., 2001a. Reducing uncertainty in the use of allometric biomass equations for predicting above-ground tree biomass in mixed secondary forests. For. Ecol. Manag. 146, 199–209.

Ketterings, Q.M., Coe, R., van Noordwijk, M., Ambagau, Y., Palm, C.A., 2001b. Reducing uncertainty in the use of allometric biomass equations for predicting above-ground tree biomass in mixed secondary forests. For. Ecol. Manag. 146, 199–209.

Lefsky, M.A., Cohen, W.B., Acker, S.A., Parker, G.G., Spies, T.A., Harding, D., 1999. Lidar remote sensing of the canopy structure and biophysical properties of Douglas-fir western hemlock forests. Remote Sens. Environ. 70, 339–361.

Lefsky, M.A., Harding, D.J., Keller, M., Cohen, W.B., Carabajal, C.C., Espirito-Santo, F.D., Hunter, M.O., de Oliveira, R., 2005. Estimates of forest canopy height and aboveground biomass using ICESat. Geophys. Res. Lett. 32.

Margolis, H.A., Nelson, R.F., Montesano, P.M., Beaudoin, A., Sun, G., Andersen, H.-E., Wulder, M.A., 2015. Combining satellite lidar, airborne lidar, and ground plots to estimate the amount and distribution of aboveground biomass in the boreal forest of North America. Can. J. For. Res. 45, 838–855.

Nelson, R., 2010. Model effects on GLAS-based regional estimates of forest biomass and carbon. Int. J. Remote Sens. 31, 1359–1372.

Nelson, R., Margolis, H., Montesano, P., Sun, G., Cook, B., Corp, L., Andersen, H.-E., deJong, B., Pellat, F.P., Fickel, T., Kauffman, J., Prisley, S., 2017. Lidar-based estimates of aboveground biomass in the continental US and Mexico using ground, airborne, and satellite observations. Remote Sens. Environ. 188, 127–140.

Niklas, K.J., 1994. Plant Allometry: The Scaling of Form and Process. University of Chicago Press.

Ni-Meister, W., 2015. Aboveground terrestrial biomass and carbon stock estimations from multisensory remote sensing. In: Remote Sensing Handbook. CRC Press (2015), ISBN-13: 978-1482218015, ISBN-10: 1482218011.

Ni-Meister, W., Jupp, D.L.B., Dubayah, R., 2001. Modeling lidar waveforms in heterogeneous and discrete canopies. IEEE Trans. Geosci. Remote Sens. 39, 1943–1958.

Ni-Meister, W., Lee, S.Y., Strahler, A.H., Woodcock, C.E., Schaaf, C., Yao, T.A., Ranson, K. J., Sun, G.Q., Blair, J.B., 2010a. Assessing general relationships between aboveground biomass and vegetation structure parameters for improved carbon estimate from lidar remote sensing. J. Geophys. Res. Biogeosci. 115.

Ni-Meister, W., Yang, W.Z., Kiang, N.Y., 2010b. A clumped-foliage canopy radiative transfer model for a global dynamic terrestrial ecosystem model. I: theory. Agric. For. Meteorol. 150, 881–894.

Ni-Meister, W., Yang, W., Lee, S., Strahler, A.H., Zhao, F., 2018. Validating modeled lidar waveforms in forest canopies with airborne laser scanning data. Remote Sens. Environ. 204, 229–243.

Pilli, R., Anfodillo, T., Carrer, M., 2006. Towards a functional and simplified allometry for estimating forest biomass. For. Ecol. Manag. 237, 583–593.

Ploton, P., Barbier, N., Momo, S.T., Réjou-Méchain, M., Boyemba Bosela, F., Chuyong, G., Dauby, G., Droissart, V., Fayolle, A., Goodman, R.C., Henry, M., Kamdem, N.G., Katembo Mukirania, J., Kenfack, D., Libalah, M., Ngomanda, A., Rossi, V., Sonké, B., Texier, N., Thomas, D., Zebaze, D., Couteron, P., Berger, U., Pélissier, R., 2015. Closing a gap in tropical forest biomass estimation: accounting for crown mass variation in pantropical allometries. Biogeosci. Discuss. 12, 19711–19750. https://doi.org/10.5194/bgd-12-19711-2015.

Reichstein, M., Carvalhais, N., 2019. Aspects of forest biomass in the earth system: its role and major unknowns. Surv. Geophys. 40, 693–707.

Saarela, S., Wästlund, A., Holmström, E., Mensah, A.A., Holm, S., Nilsson, M., Fridman, J., Ståhl, G., 2020. Mapping aboveground biomass and its prediction uncertainty using LiDAR and field data, accounting for tree-level allometric and LiDAR model errors. For. Ecosyst. 7, 43.

Schimel, D., Pavlick, R., Fisher, J.B., Asner, G.P., Saatchi, S., Townsend, P., Miller, C., Frankenberg, C., Hibbard, K., Cox, P., 2015. Observing terrestrial ecosystems and the carbon cycle from space. Glob. Chang. Biol. 21, 1762–1776.

Shenkin, A., Bentley, L.P., Oliveras, I., Salinas, N., Adu-Bredu, S., Marimon-Junior, B.H., Marimon, B.S., Peprah, T., Choque, E.L., Trujillo Rodriguez, L., Clemente Arenas, E. R., Adonteng, C., Seidu, J., Passos, F.B., Reis, S.M., Blonder, B., Silman, M., Enquist, B.J., Asner, G.P., Malhi, Y., 2020. The influence of ecosystem and phylogeny on tropical tree crown size and shape. Front. For. Glob. Chang. 3.

Sheridan, R., Popescu, S., Gatziolis, D., Morgan, C., Ku, N.-W., 2014. Modeling forest aboveground biomass and volume using airborne LiDAR metrics and forest inventory and analysis data in the Pacific northwest. Remote Sens. 7, 229–255.

Sillett, S.C., Van Pelt, R., Carroll, A.L., Campbell-Spickler, J., Antoine, M.E., 2020. Aboveground biomass dynamics and growth efficiency of Sequoia sempervirens forests. For. Ecol. Manag. 458.

Silva, C.A., Duncanson, L., Hancock, S., Neuenschwander, A., Thomas, N., Hofton, M., Fatoyinbo, L., Simard, M., Marshak, C.Z., Armston, J., Lutchke, S., Dubayah, R., 2021. Fusing simulated GEDI, ICESat-2 and NISAR data for regional aboveground biomass mapping. Remote Sens. Environ. 253.

Spriggs, R., Coomes, D., Jones, T., Caspersen, J., Vanderwel, M., 2017. An alternative approach to using LiDAR remote sensing data to predict stem diameter distributions across a temperate forest landscape. Remote Sens. 9.

Stark, S.C., Leitold, V., Wu, J.L., Hunter, M.O., de Castilho, C.V., Costa, F.R., McMahon, S.M., Parker, G.G., Shimabukuro, M.T., Lefsky, M.A., Keller, M., Alves, L. F., Schietti, J., Shimabukuro, Y.E., Brandao, D.O., Woodcock, T.K., Higuchi, N., de Camargo, P.B., de Oliveira, R.C., Saleska, S.R., Chave, J., 2012. Amazon forest carbon dynamics predicted by profiles of canopy leaf area and light environment. Ecol. Lett. 15, 1406–1414.

Stark, S.C., Enquist, B.J., Saleska, S.R., Leitold, V., Schietti, J., Longo, M., Alves, L.F., Camargo, P.B., Oliveira, R.C., 2015. Linking canopy leaf area and light environments with tree size distributions to explain Amazon forest demography. Ecol. Lett. 18, 636–645.

Swatantran, A., Dubayah, R., Roberts, D., Hofton, M., Blair, J.B., 2011. Mapping biomass and stress in the Sierra Nevada using lidar and hyperspectral data fusion. Remote Sens. Environ. 115, 2917–2930.

Taubert, F., Fischer, R., Knapp, N., Huth, A., 2021. Deriving tree size distributions of tropical forests from Lidar. Remote Sens. 13.

Tritton, L.M., Hornbeck, J.M., 1982. Biomass Equations for Majortree Species of the Northeast, Gen. Tech. Rep. NE-69, Northeast. Res. Stn., For. Serv. U.S. Dep. of Agric., Newtown Square, Pa.

Waddell, K., Hiserote, B., 2005. The PNW-FIA Integrated Database user guide: A database of forest inventory information for California, Oregon, and Washington. Pacific Northwest Research Station, Portland, Oregon, USA (Accessed 8 July 2022).

Wulder, M.A., White, J.C., Nelson, R.F., Næsset, E., Ørka, H.O., Coops, N.C., Hilker, T., Bater, C.W., Gobakken, T., 2012. Lidar sampling for large-area forest characterization: a review. Remote Sens. Environ. 121, 196–209.

Yang, W., Ni-Meister, W., Kiang, N.Y., Moorcroft, P.R., Strahler, A.H., Oliphant, A., 2010. A clumped-foliage canopy radiative transfer model for a global dynamic terrestrial ecosystem model II: comparison to measurements. Agric. For. Meteorol. 150, 895–907.

Zianis, D., Mencuccini, M., 2004. On simplifying allometric analyses of forest biomass. For. Ecol. Manag. 187, 311–332.