Title:

Linking lidar and forest modeling to assess biomass estimation across scales and disturbance states

List of authors:

Nikolai Knapp[1], Rico Fischer[1], Andreas Huth[1, 2, 3]

Authors' affiliation:

1) Department of Ecological Modeling, Helmholtz Centre for Environmental Research (UFZ), 04318 Leipzig, Germany
2) German Centre for Integrative Biodiversity Research (iDiv), Halle-Jena-Leipzig, 04103 Leipzig, Germany
3) Institute for Environmental Systems Research, Department of Mathematics/Computer Science, University of Osnabrück, 49076 Osnabrück, Germany

Corresponding author:

Nikolai Knapp, Email: nikolai.knapp@ufz.de), Tel.: +49 3412354764

Type of paper:

Primary Research Article

Abstract

Light detection and ranging (lidar) is currently the state-of-the-art remote sensing technology for measuring the 3D structures of forests. Studies have shown that various lidar-derived metrics can be used to predict forest attributes, such as aboveground biomass. However, finding out which metric works best at which scale and under which conditions requires extensive field inventories as ground-truth data. The goal of our study was to overcome the limitations of inventory data by complementing field-derived data with virtual forest stands from a dynamic forest model. The simulated stands were used to compare 29 different lidar metrics for their utility as predictors of tropical forest biomass at different spatial scales. We used the process-based forest model FORMIND, developed a lidar simulation model, based on the Beer-Lambert law of light extinction, and applied it to a tropical forest in Panama. Simulation scenarios comprised undisturbed primary forests and stands exposed to logging and fire disturbance regimes, resulting in mosaics of different successional stages, totaling 3.7 million trees on 4,200 ha. The simulated forest was sampled with the lidar model. Several lidar metrics, in particular height metrics, showed good correlations with forest biomass, even for disturbed forest. Estimation errors (nRMSE) increased with decreasing spatial scale from < 10% (200-m scale) to > 30% (20-m scale) for the best metrics. At the often used 1-ha scale, the top-of-canopy height obtained from canopy height models with fine to relatively coarse pixel resolutions (1 to 10 m) yielded the most accurate biomass predictions, with nRMSE < 6% for undisturbed and nRMSE < 9% for disturbed forests. This study represents the first time dynamic modeling of a tropical forest has been combined with lidar remote sensing to systematically investigate lidar-to-biomass relationships for varying lidar metrics, scales and disturbance states. In the future, this approach can be used to explore the potential of remote sensing of other forest attributes, e.g., carbon dynamics, and other remote sensing systems, e.g., spaceborne lidar and radar.

Keywords: aboveground biomass; tropical forest; disturbance; lidar simulation; forest modeling; resolution; scale

1. Introduction

Due to their important role in the global carbon cycle and ongoing deforestation and degradation, tropical forests are of particular interest to biomass remote sensing. Tropical forest carbon accounting and monitoring of deforestation are important tasks in the context of REDD+ and global climate modeling. In recent years, remote sensing has led to considerable improvements in this field (Gibbs et al., 2007; De Sy et al., 2012; Pan et al., 2013). Airborne small-footprint lidar (light detection and ranging) is currently the state-of-the-art technology for measuring the 3D structure of forests (Lefsky et al., 2002b; Wulder et al., 2012; Mascaro et al., 2014). Various lidar metrics correlate well with different forest attributes. In particular, lidar-derived height metrics have commonly been used to predict forest aboveground biomass (AGB) and carbon density (ACD) (Drake et al., 2002; Asner et al., 2009; Dubayah et al., 2010; Jubanski et al., 2013; Asner & Mascaro, 2014).The major challenges in biomass estimation based on lidar data are that 1) the calibration of the prediction functions relies on field data that must be collected manually in inventory plots; and 2) there are many different metrics available using different spatial scales, and the task is to find the combination that provides accurate AGB predictions.

In inventory plots, tree diameters at breast height (DBH) are typically measured, from which AGB is calculated via known allometric equations (e.g., Chave et al., 2005, 2014; Chen 2015). Lidar data are acquired for the same inventory plots to build regression models between lidar-based structure metrics and ground-based AGB. A wide range of metrics can be calculated from lidar data. To date, no standard approach for AGB estimation from lidar has been established and different studies have applied different metrics (Chen 2013; Lu et al. 2014). Several publications have compared metrics among each other for different forest types (e.g., Lefsky et al., 1999, 2002a; Dubayah et al., 2010; Jubanski et al., 2013). However, there has not been a comparison of a wide range of metrics on a single tropical forest dataset. Lidar metrics can generally be divided into metrics which are based on the full 3D point cloud of lidar returns and metrics which are based on canopy height models (CHM), i.e., the rasterized canopy surfaces which are derived from the uppermost returns of the point clouds (Chen 2013). The full 3D point cloud contains more information about the vertical canopy structure than the corresponding CHM. On the other hand, the vertical distribution of lidar returns also depends on technical properties of the specific sensor, making point-cloud-based metrics less robust and comparable between different studies than CHM-based metrics (Næsset, 2009; Asner & Mascaro, 2014). Many commonly used metrics can be calculated based on both types of data. Those metrics include mean heights (Lefsky et al., 2002a; Asner & Mascaro, 2014), relative height quantiles (the heights below which a certain percentage of returns or pixels falls) (Patenaude et al., 2004; Dubayah et al., 2010; Meyer et al., 2013), and metrics of heterogeneity such as the standard deviation of heights or the Shannon diversity index of the height profiles (Stark et al., 2012). Other metrics, such as the ratio of above ground returns to total returns or fractional canopy cover above a certain height, that can be derived either from point clouds or CHMs describe relative vegetation cover.

An important aspect of AGB prediction from remote sensing is spatial resolution. Resolution means, first, spatial resolution of the remote sensing data from which different metrics are calculated and, second, the spatial resolution of the output map, i.e., the grain size of the units for which the metrics are calculated to produce an AGB prediction. The resolution of the data is determined by the sensor's technical specifications and the capacities to store and process data. The resolution of the mapping units is influenced by the desired estimation accuracy and the desired spatial detail of the mapped product. Köhler & Huth (2010), Mascaro et al. (2011b) and Chen et al. (2016) showed how errors in AGB estimations from mean lidar heights decreased with increasing grain sizes and that a grain of approximately 1 ha is required to achieve errors of < 10%.

Fitting any of the described lidar metrics to measured AGB relies on field inventory data. Forest inventory plots are limited in number, size and structural variety. The collection of inventory data is costly and laborious and most studies in the past made use of tens to a few hundred plots (Fassnacht et al., 2014). Those plots are often located in old growth forests. Hence, available data sets might not cover the full structural complexity of forests over their entire successional range (noteworthy exceptions are e.g., Dubayah et al. 2010, Poorter et al. 2016). For lidar-to-AGB-calibration, a broad range of different forest succession states that cover the range of all possible AGB stocks and associated forest structures is preferable. To overcome this limitation, we propose a new approach in which we complement in situ measurements with simulated forest stands (Fig. 1). We used an individual-based forest model (FORMIND, Fischer et al., 2016) to simulate a large virtual inventory dataset, covering the full range of succession stages by including forest disturbances in the simulations. The model was parameterized to represent the well-studied lowland tropical rainforest of Barro Colorado Island, Panama (Condit et al., 2001; Kazmierczak et al., 2014). We developed a lidar model to sample lidar data of simulated forest stands.

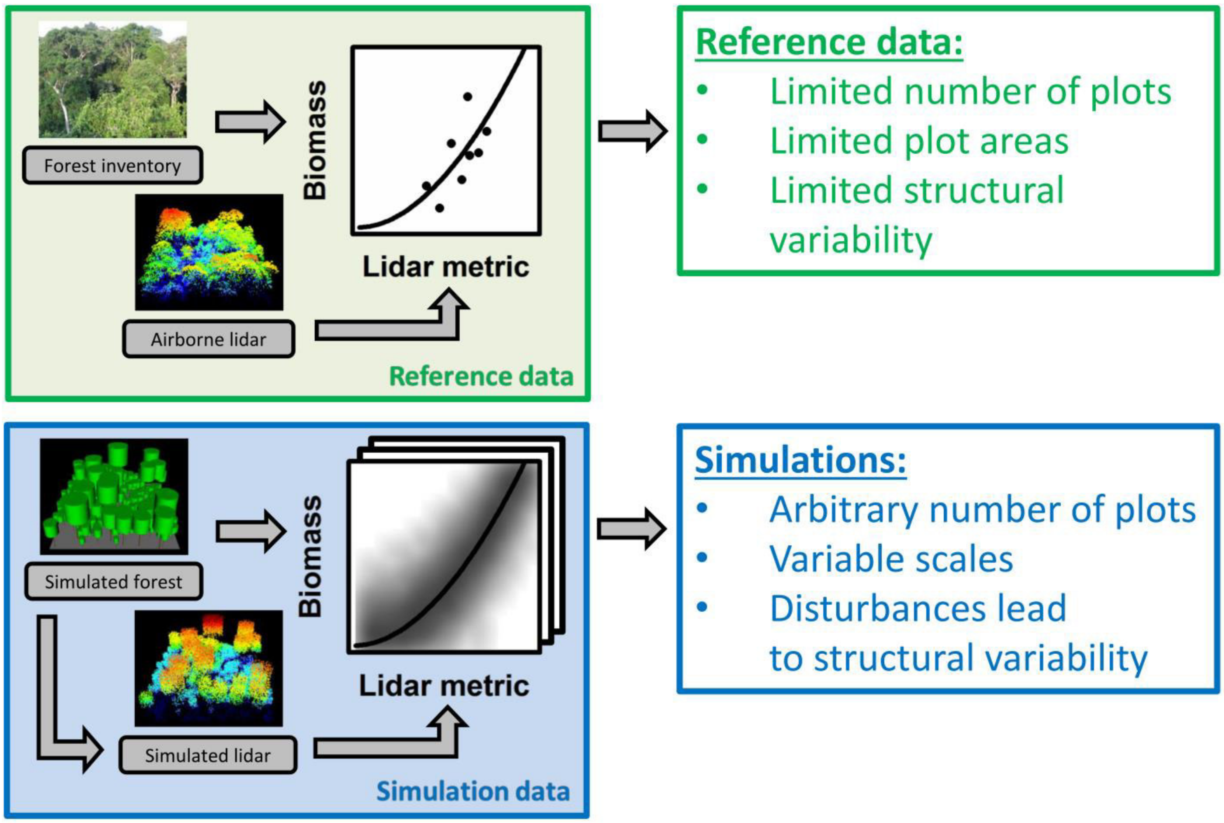

Fig. 1: Workflow of the study. Reference data from field inventories and an airborne lidar campaign were used to parameterize and calibrate a forest model and a lidar model. With the models, large quantities of simulated inventory and simulated lidar data were generated, allowing for a systematic analysis of lidar-to-biomass relationships under different disturbance regimes and for various spatial scales.

The research goals of this study were 1) to establish a lidar simulation model that is able to produce synthetic lidar-like data for dynamic forest model output; 2) to test a wide variety of lidar metrics for their ability to predict AGB of a tropical rainforest at various spatial scales; and 3) to investigate the influence of disturbances on the lidar-to-biomass relationships.

2. Material & Methods

2.1 Study area

The study focused on the tropical forest on Barro Colorado Island (BCI), Panama (9.15° N, 79.85° W). BCI is a 15 km^2 island located in Lake Gatun, an artificial water body created by the construction of the Panama Canal (Condit et al., 2001). It is covered with semi-deciduous tropical lowland rainforest, the minimum forest age is estimated to range from 300 to 1500 years (Bohlman & O'Brien, 2006; Meyer et al., 2013; Lobo & Dalling, 2014). The climate is characterized by average daily maximum and minimum temperatures of 30.8 and 23.4 °C and an annual precipitation sum of approximately 2600 mm, with a dry season from January to April (Condit et al., 2001). A 50-ha rainforest observation plot is located on the central plateau of the island, with terrain altitudes varying between 120 and 160 m above sea level (Lobo & Dalling, 2014). Since the establishment of the plot in the early 1980s, each tree in the 1000 m × 500 m area with a DBH ≥ 1 cm has been measured during censuses in five year intervals (Condit, 1998; Hubbell et al., 1999, 2005). Estimates of the mean canopy height are 24.6 ± 8.2 m, and those of the mean AGB are 281 ± 20 t/ha (Chave et al., 2003).

2.2 Lidar data

An airborne discrete point cloud lidar dataset was collected on BCI in August 2009 with a multi-pulse scanning laser altimeter (Optech ALTM Gemini system; BLOM Sistemas Geoespaciales SLU, Madrid, Spain, Lobo & Dalling, 2014). The terrain elevation was subtracted from the point cloud to obtain the relative height above ground. Point densities ranged from 0 to 60 m^{-2} with a median of 10 m^{-2} and a 5^{th}-percentile of 4 m^{-2}. To avoid locally varying point densities, caused by flight swath overlaps, the point clouds were thinned by random subsampling of 4 returns in each square meter. A 1-m resolution canopy height model (CHM) was derived from the highest returns in each square meter. Data processing was performed using LAStools (Isenburg, 2011) and R (R Development Core Team, 2014).

2.3 Lidar model description

The purpose of the lidar model is the simulation of a lidar scan of a given forest stand. More specifically, it generates point clouds of discrete returns as usually produced by small-footprint lidar systems. As input, a tree list has to be provided. The list can either be real forest inventory data or data generated by a forest model (Fig. 2a). The basic elements of the model are trees, lidar pulses and lidar returns. Trees are characterized by their position (X- and Y-coordinate), height, crown length, crown radius, crown shape and leaf area index (LAI). The model operates in a 3D space represented by an array of cuboid voxels. Each vertical column of voxels represents one modeled lidar pulse. Lidar returns are points in 3D space, characterized by their X-, Y- and Z-coordinates.

From the tree list, a voxel representation of the entire forest is created. Thus, voxels that could potentially produce a lidar return, because they belong to a tree crown or the ground, are distinguished from empty space voxels. The voxel forest is then scanned with a virtual lidar. The simulation follows a probabilistic approach. Instead of explicitly simulating the branches and foliage and their interaction with laser beams within the tree crowns, the model assumes that the tree crown space is a homogeneous, turbid medium filled with a certain leaf area density (LAD). The probability of having a lidar return from a certain point decreases as the distance the laser beam has to travel through the medium before reaching the point increases. This relationship is analogous to the Beer-Lambert light-extinction law (Campbell & Norman, 2012). Thus, the probability for a lidar return *P* for each tree and ground voxel (Fig. 2c) can be calculated as a function of cumulative leaf area index *LAI* above the voxel (Fig. 2b).

$$\boldsymbol{P(LAI) = P_0 \cdot e^{-k \cdot LAI}} \tag{1}$$

P_0 in Eq. (1) represents the probability of obtaining a return from the very upper voxel, where the laser beam hits a tree or the ground for the first time. The parameter k is the exponential extinction coefficient, which determines how fast the return probability decreases after entering the crown space. The decision regarding whether each voxel will contain a return is taken stochastically, based on the calculated return probability. Ultimately, this leads to a discrete point cloud (Fig. 2d). The voxel resolution was set to 0.5 m × 0.5 m along the horizontal direction and 1 m along the vertical direction. The parameters P_0 and k were calibrated such that simulated point cloud profiles derived for subareas of the 50-h inventory data set matched the airborne lidar profiles of those subareas (details see supplements). The resulting value for $k = 0.2$ can be confirmed by literature (Campbell & Norman 2012, Jones 2013). For P_0 we found 0.2 to be a good value, leading to simulated point densities that were similar to the airborne reference point cloud. P_0 being smaller than 1 can be interpreted by the heterogeneity of leafs, branches and empty space within the tree crown. This means that a laser beam entering the idealized cylindrical tree crown does not necessarily trigger a return in the first voxel.

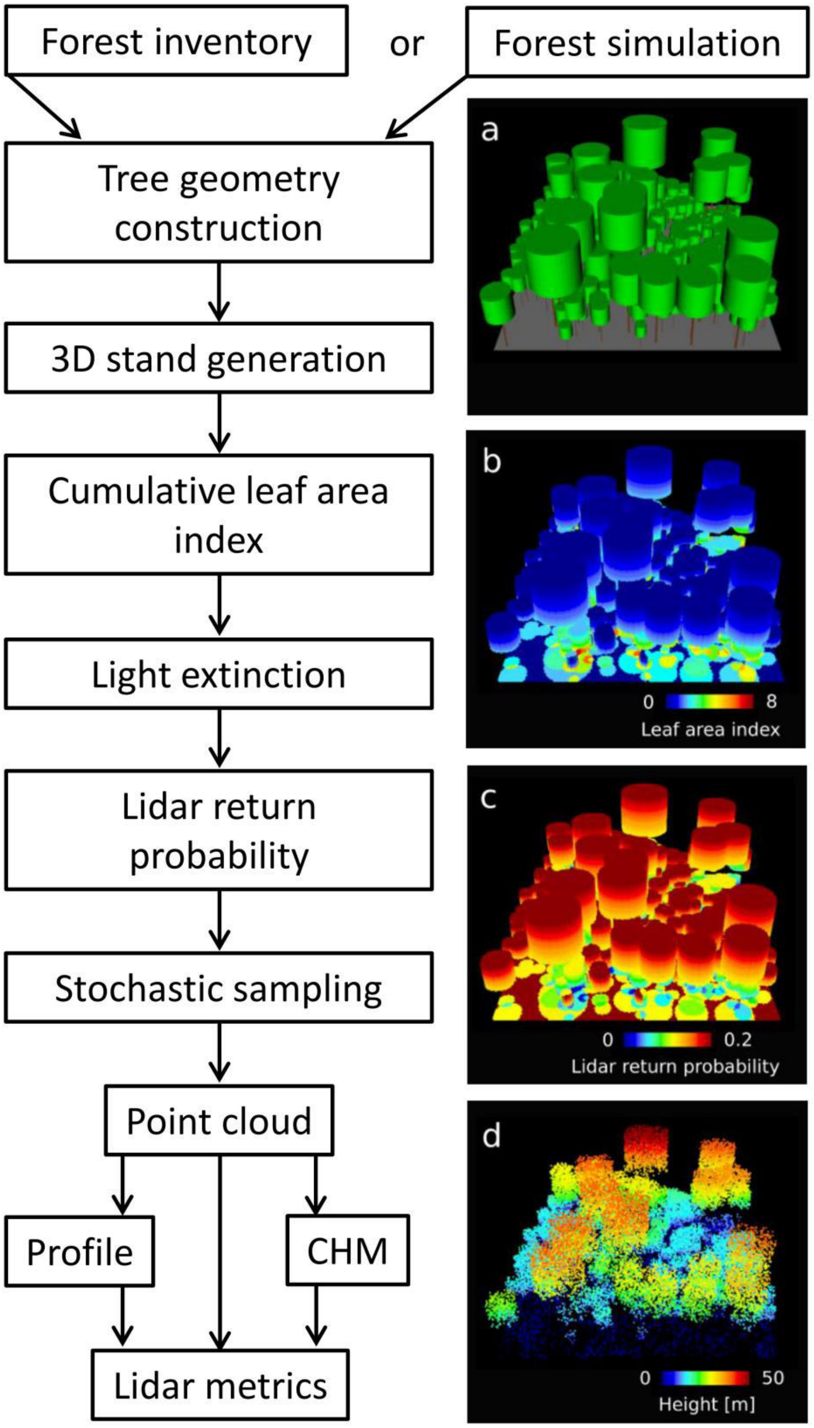

Fig. 2: Principle of the lidar model. Inputs to the workflow can either be forest model output or field inventory data. The pictures on the right side show intermediate products: a) Visualization of a forest stand; b) voxel representation with colors indicating the cumulative leaf area index; c) voxel representation with colors indicating the probability of containing a lidar return; d) simulated lidar point cloud with colors indicating height above ground.

2.4 Forest model description

FORMIND belongs to the group of forest gap models (Botkin et al., 1972; Shugart, 1984; Bugmann, 2001). As such, the model simulates the processes of establishment, growth, competition and mortality of trees on spatial patches with the dimensions of a typical treefall gap (20 m × 20 m). By combining many patches, large forest areas of hundreds of hectares can be simulated. FORMIND is an individual-based model (IBM) in which the individuals represent trees that belong to different plant functional types (PFTs). One PFT may contain several species with similar ecological traits. FORMIND has been applied to many tropical forest sites and has proven capable of accurately reproducing patterns observed in these complex ecosystems (Fischer et al., 2016). The individual-based model architecture allows for the inclusion of disturbances such as logging or forest fires in a structurally realistic way. A detailed description of FORMIND including the modules for logging and fire disturbance can be found in Fischer et al. (2016). The supplements contain descriptions of the parameterization of the lidar model and the forest model (Tab. S1). Before using the forest model output for remote sensing analyses, the structural validity of the simulated old growth stands was confirmed by visually comparing biomass stocks (Fig. S1) and stem size distributions (Fig. S2) of all PFTs to the values obtained from the inventory data.

2.5 Simulation experiment

Using FORMIND, we simulated the development of a 16 ha (400 m × 400 m) area of the BCI forest over several thousands of years and stored the results at 20-yr intervals. The simulations were repeated with different disturbance regimes. The first run comprised 2000 yr without any external disturbance, simulating only natural gap dynamics. In the second run, forest fires were introduced as a source of spatially heterogeneous disturbance to clear parts of the area regularly and enable natural succession and regrowth. Fire occurrence was drawn from a Poisson distribution such that the mean interval between two fire events was 25 yr. Fire size at each fire event was drawn from an exponential distribution, such that on average 50% of the total area was affected. More information on the fire module used is provided in Fischer (2013) and Fischer et al. (2016). The third scenario included selective logging. At a logging cycle of 99 yr, all trees with DBH > 30 cm were felled and removed. More information on the logging module used is provided in Huth et al. (2004). For all three runs, the first 200 yr were discarded as spin-up. For each of the remaining simulation years, a virtual lidar campaign using the lidar model was conducted. The disturbance frequencies and intensities were not intended to represent realistic disturbances scenarios in the study region. The intention was to sample many stands at each stage along the full successional range, using the disturbance modules to regularly set the forest back to an early stage. The selective logging acts on the whole area, while the fires move in a spatially explicit way through the simulated area, causing mosaics of unaffected forest next to cleared areas where succession starts over. Such patchy landscapes are typical for many forest regions, although the reasons for the structures may be as diverse as clear cuts, wind blowdowns, fires or natural areas without vegetation, e.g., grasslands or water bodies. Thus, these simulations produce landscapes that can be used as general examples of heterogeneous landscapes.

2.6 Lidar-based biomass prediction

We analyzed forest plots measuring 20, 33, 50, 100 or 200 m (side length). At each spatial scale, a range of 29 different lidar metrics (Tab. 1) were tested for their suitability as single predictors of AGB. Metrics were either derived from point clouds (PC) or canopy height models (CHM). CHMs were constructed from point clouds by rasterizing the highest lidar returns in each pixel of a given pixel size.

Point-cloud-based metrics comprised the mean canopy profile height (MCH), which is the mean height of all lidar returns, and the quadratic mean canopy profile height (QMCH), where high returns receive a larger weighting than low returns. For a given point cloud profile p_{PC} that consists of lidar return counts at height bins h_i, MCH and QMCH can be calculated from Eq. (2) and (3), respectively.

$$\boldsymbol{MCH} = \frac{\sum_{i=1}^{i_{max}}(\boldsymbol{p}_{PC,i} \cdot \boldsymbol{h}_i)}{\sum_{i=1}^{i_{max}} \boldsymbol{p}_{PC,i}} \tag{2}$$

$$\boldsymbol{QMCH} = \sqrt{\frac{\sum_{i=1}^{i_{max}}(\boldsymbol{p}_{PC,i} \cdot \boldsymbol{h}_i^{\,2})}{\sum_{i=1}^{i_{max}} \boldsymbol{p}_{PC,i}}} \tag{3}$$

where $p_{PC,i}$ is the lidar return counts in height bin h_i. A metric similar to MCH can be derived from the vertical CHM profile instead of the point cloud profile. This metric corresponds to the mean of all pixel values of the CHM, and is commonly referred to as the mean top-of-canopy height (TCH, Eq. (4)).

$$\boldsymbol{TCH} = \frac{\sum_{i=1}^{i_{max}}(\boldsymbol{p}_{CHM,i} \cdot \boldsymbol{h}_i)}{\sum_{i=1}^{i_{max}} \boldsymbol{p}_{CHM,i}} \tag{4}$$

Because a CHM can be derived from a point cloud at variable pixel resolutions, by taking the height of the highest return that falls into each pixel, TCH always depends on the pixel size used. We calculated TCH from CHMs with pixel side lengths of 1, 5, 10, 20, 33, 50 and 100 m. Note that, once the pixel size equals the plot size for which AGB is calculated, TCH is equal to the maximal height in the plot, which is also referred to as H_{max} or RH100 in the literature. Another method for measuring forest height from lidar data is by using relative height quantiles of either the point cloud or the CHM. These quantiles represent the heights below which a certain percentage of the returns or CHM pixels fall. We calculated RH25, RH50 and RH75 for the point clouds and 1-m resolution CHMs.

Other metrics, however, capture the vertical heterogeneity of the forest. Those metrics include the standard deviation (SD) of heights (point-cloud- or CHM-based), the coefficient of variation (CV, Eq. (5) and (6)), the skewness of the vertical point cloud profile (Eq. (7), where N is the total number of points and h_i is the height of each point i), the Shannon Index (Eq. (8), where i_{max} is the number of height layers and p_i is the count of points in the layer i) as a measure of entropy of the profile and the P:H ratio (Eq. (9), where i_{max} is the number of height layers, p_i is the count of points in the layer i and h_i is height of layer *i*), which describes the height of the densest part of the point cloud (peak in the profile) relative to the maximal height (Marvin et al., 2014).

$$\boldsymbol{CV_{PC}} = \frac{\boldsymbol{SD_{PC}}}{\boldsymbol{MCH}} \quad (5)$$

$$\boldsymbol{CV_{CHM}} = \frac{\boldsymbol{SD_{CHM}}}{\boldsymbol{TCH}} \quad (6)$$

$$\boldsymbol{Skewness} = \frac{\boldsymbol{1}}{\boldsymbol{N}} \cdot \sum_{i=1}^{N} \left(\frac{\boldsymbol{h_i - MCH}}{\boldsymbol{SD_{PC}}}\right)^3 \quad (7)$$

$$\boldsymbol{Shannon\ Index} = -\sum_{i=1}^{i_{max}} \boldsymbol{p_i} \cdot \ln(\boldsymbol{p_i}) \quad (8)$$

$$\boldsymbol{P{:}H\ ratio} = \frac{\boldsymbol{h}\left(\max_{i \in [1,\, i_{max}]} (\boldsymbol{p_i})\right)}{\max_{i \in [1,\, i_{max}]} (\boldsymbol{h_i})} \quad (9)$$

Furthermore, we calculated vegetation density metrics. Based on the point clouds, the count of aboveground returns divided by either the count of ground returns N_{AGR}/N_{GR} or the count of total returns N_{AGR}/N_{TR} was calculated. Based on the CHMs, the fractional canopy cover (FCC) was derived by defining different height thresholds below which a CHM-pixel was considered a canopy gap. We calculated FCC0, FCC10 and FCC20 using the forest floor, 10 m and 20 m as height thresholds, respectively.

Tab. 1: List of the lidar metrics and the underlying data (PC = point cloud, CHM = canopy height model). CHM usually refers to 1-m resolution rasters, except for TCH where various resolutions were tested.

Lidar metric	Description	Data
MCH	Mean canopy profile height	PC
QMCH	Quadratic mean canopy profile height	PC
TCH	Mean top-of-canopy height (at variable CHM pixel resolutions), e.g., TCH5 is based on 5-m pixels	CHM
RH	Relative height quantile, e.g., RH50 is the 50-percentile of heights	PC or CHM
SD	Standard deviation of heights	PC or CHM
CV	Coefficient of variation of heights (normalized SD)	PC or CHM
Skewness	Skewness of the vertical profile	PC
Shannon Index	Entropy of the vertical profile	PC
P:H ratio	Relative height of the peak in the vertical profile	PC
N_{AGR}/N_{GR}	Ratio of aboveground returns to ground returns	PC
N_{AGR}/N_{TR}	Ratio of aboveground returns to total returns	PC
FCC	Fractional canopy cover, e.g., FCC10 is the relative share of pixels higher than 10 m	CHM

Each lidar metric *LM* was fit to the dependent variable *AGB* using a power law model (Eq. (10)) and maximum likelihood estimation in R.

$$\boldsymbol{AGB = a \cdot LM^{b}} \tag{10}$$

If possible, such relationships were derived for plots with side lengths of 20, 33, 50, 100 and 200 m. Relationships could not be derived in cases where pixel size exceeded plot size or where the maximum likelihood estimation did not provide a parameter *b* different from zero. The AGB-prediction accuracy for the different power law functions was quantified as the normalized root mean square error (*nRMSE*) [%]. The measure was calculated as the RMSE of *n AGB* predictions against *n* observations, normalized by the mean observed *AGB* (Eq. (11)).

$$\boldsymbol{nRMSE = \sqrt{\frac{\sum_{i=1}^{n}(predAGB_i - obsAGB_i)^2}{n}} \cdot \frac{1}{\overline{obsAGB}}} \tag{11}$$

The power law parameters and additional statistics (mean, RMSE, bias, R^2, slope and intercept of linear fits between predictions and observations) for all metrics, scales and datasets (672 models) can be found in Tab. S2.

3. Results

3.1 Forest and lidar simulation results

The forest simulations could reproduce AGB succession over time for the four PFTs. An overshoot of total AGB around a forest age of 100 yr was observed (Fig. S1). The duration of the primary succession and the biomass overshoot are consistent with observations by Mascaro et al. (2012). Furthermore, the stem size distributions for all four PFTs matched well between the model and reference data (Fig. S2). The AGB distributions of reference data and undisturbed and disturbed FORMIND runs can be found in Fig. 3, and for the undisturbed case, the simulated distributions are in good agreement with previously reported distributions based on field data (Chave et al., 2003). At all scales the range of AGB in undisturbed simulations was smaller than the observed range of AGB in the field reference data. In the disturbance scenarios, the range of AGB values increased. At the small 20 m × 20 m scale, the real forest contained extremely high local AGB values (max. 2022 t/ha) caused by single large trees. Such extreme values were not reached in the simulations.

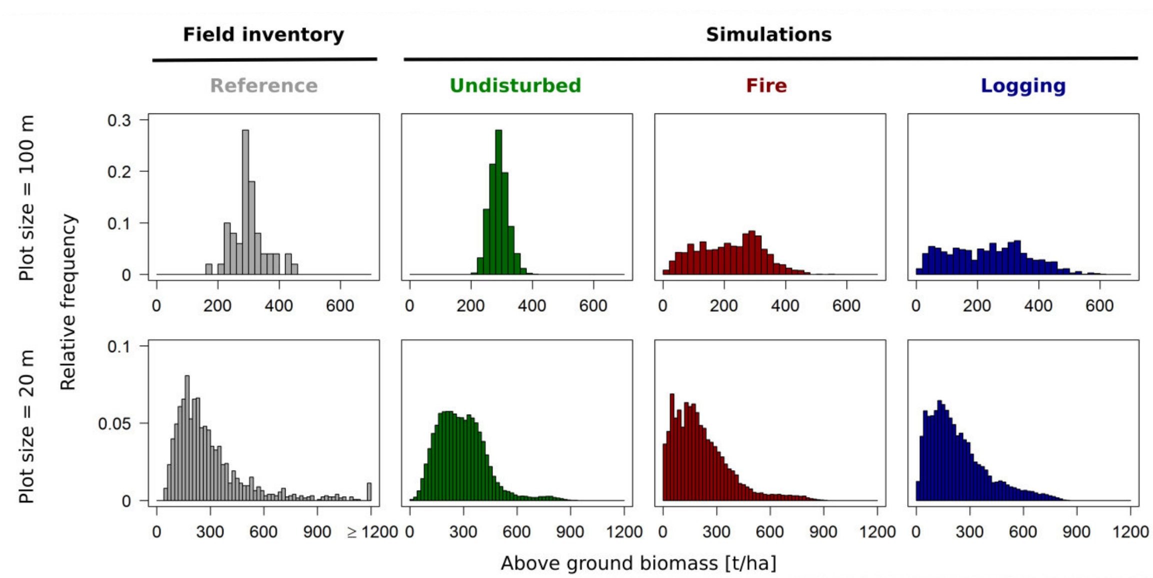

Fig. 3: Relative frequency distributions of aboveground biomass (AGB). Columns represent the BCI field data (50 ha) and output of FORMIND simulations from different disturbance scenarios (1,400 ha each). Rows represent different spatial resolutions. Notice the different axis scaling in each row.

Using the lidar simulation approach, synthetic lidar data were generated for the simulated forest stands. Lidar simulation outputs, such as the vertical point cloud profile (Fig. 4) and CHMs, closely resembled their airborne equivalents. In the supplements we present how alternative assumptions about the tree geometry affect the simulated lidar profiles and metrics (Fig. S15 to S18).

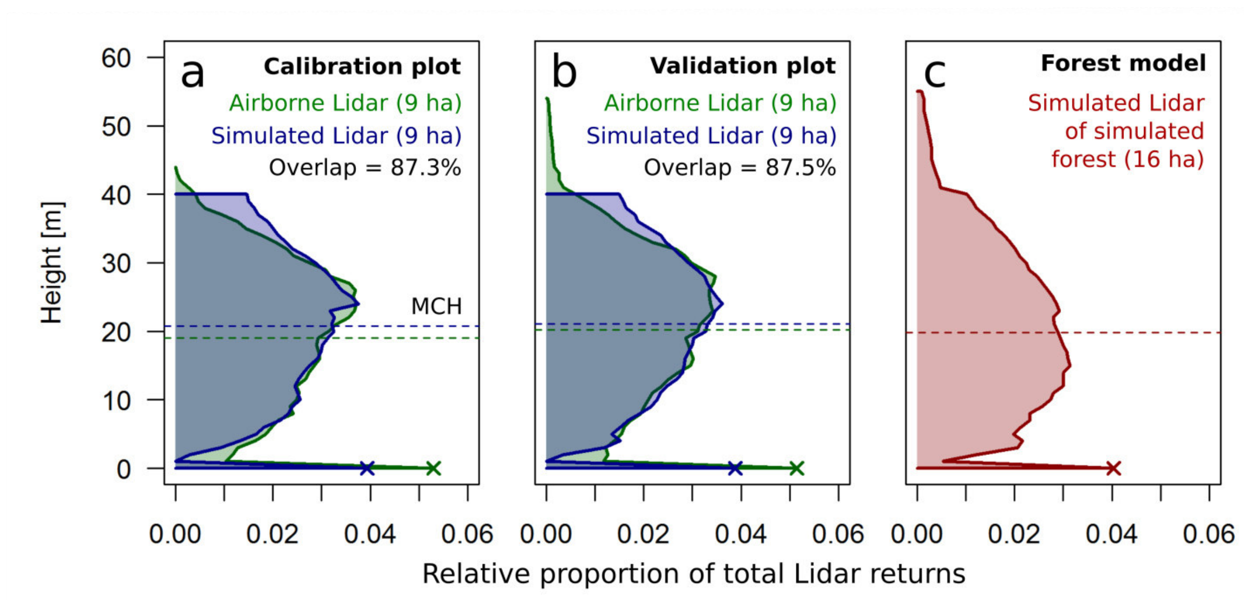

Fig. 4: Vertical lidar profiles of a) the 9 ha in the southwestern corner of the BCI megaplot, airborne and simulated based on inventory data; b) the same for the 9 ha in the northeastern corner of the BCI megaplot; and c) the simulated lidar profile of 16 ha simulated forest in FORMIND in the old growth stage (age 500 yr). Dashed lines mark the mean canopy profile height (MCH), and '×' symbols mark the ground return peaks.

3.2 Biomass prediction from top-of-canopy height

Based on the simulated stands, we analyzed 4,200 ha of forest (3.7 million trees with DBH ≥ 3 cm) with respect to the relationships between forest height (TCH) and biomass (AGB). We generated undisturbed (1,400 ha), fire-disturbed (1,400 ha) and logging-disturbed (1,400 ha) stands. Fig. 5 shows the relationships observed for different plot sizes (20 to 100 m) assuming a fine resolution (pixel size = 1 m). The disturbed stands (fire and logging were pooled) cover a wider range of TCH and AGB values than the undisturbed stands. The fitted relationships for undisturbed and disturbed forest stands are similar. The scattering around the regression lines decreases with increasing plot size. If we decrease the pixel resolution from 1 to 10 m (Fig. 6), we observe a change in the TCH-to-AGB relationship. Curves become flatter because averaging over lidar point height maxima in 10 m × 10 m pixels leads to higher TCH-values than averaging over the lidar point height maxima in all 1 m × 1 m pixels. Thus, the coarser the pixel resolution is, the higher the TCH value for a given stand becomes. For the 1-m and the 10-m pixel resolution, we observe similar relations for disturbed and undisturbed forests, respectively. More extensive analyses and graphics that consider the BCI reference data and treat the different disturbance regimes separately can be found in the supplementary material (Fig. S4 and following).

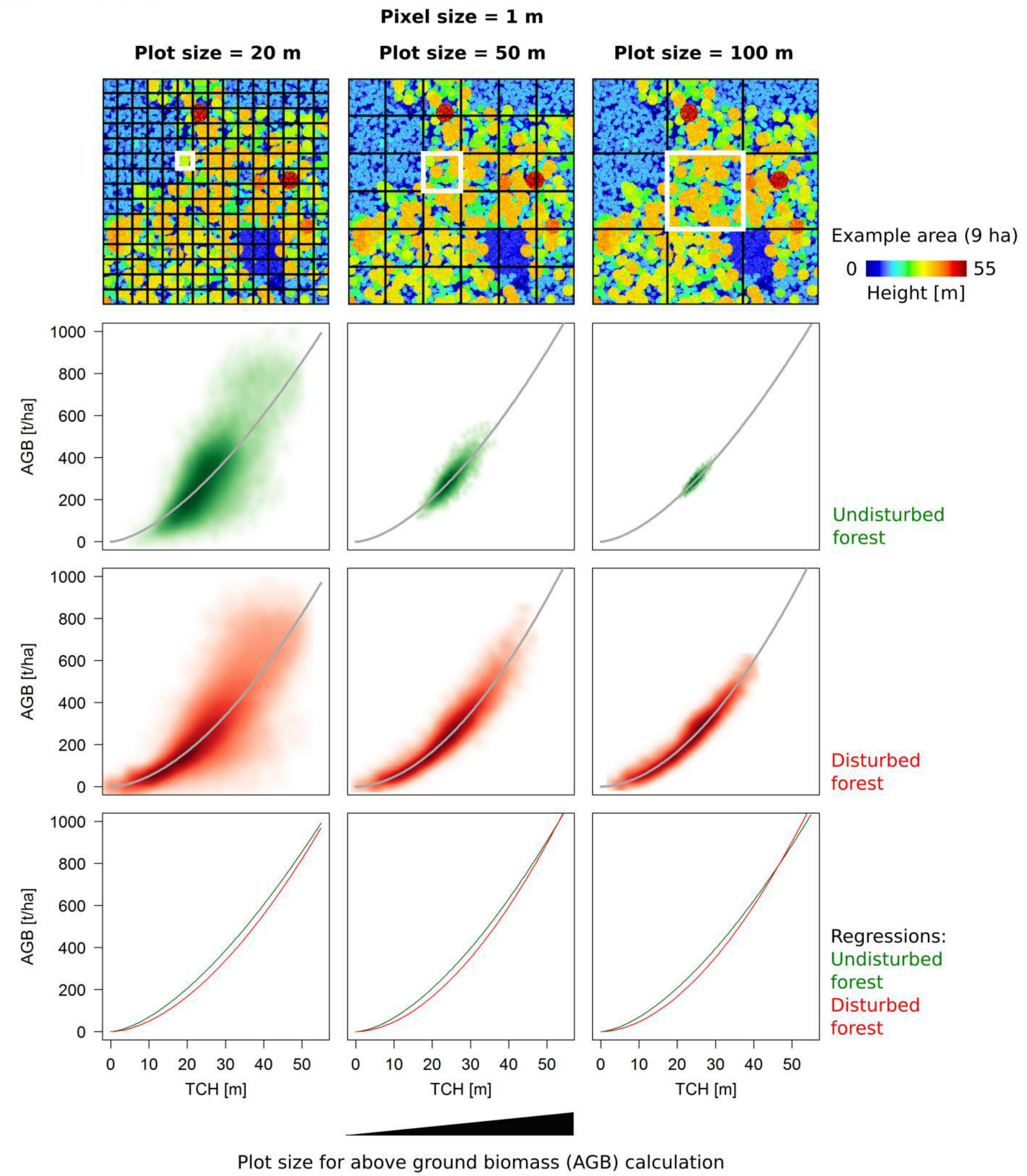

Fig. 5: Aboveground biomass (AGB) as a function of top-of-canopy height (TCH) from 1-m pixel resolution (CHM) for different plot sizes. All data was derived from FORMIND and lidar simulations. 1) The first row demonstrates the sampling approach. Shown is a scene of 9 ha simulated forest with different stages of succession. The following rows show the TCH-to-AGB relationship with each record representing one 20-m, 50-m or 100-m plot, respectively, for 2) 1,400 ha of undisturbed simulated forest (green), 3) 1,400 ha of fire-disturbed and 1,400 ha of regularly logged simulated forest (red) and 4) the curves of the best power law fits.

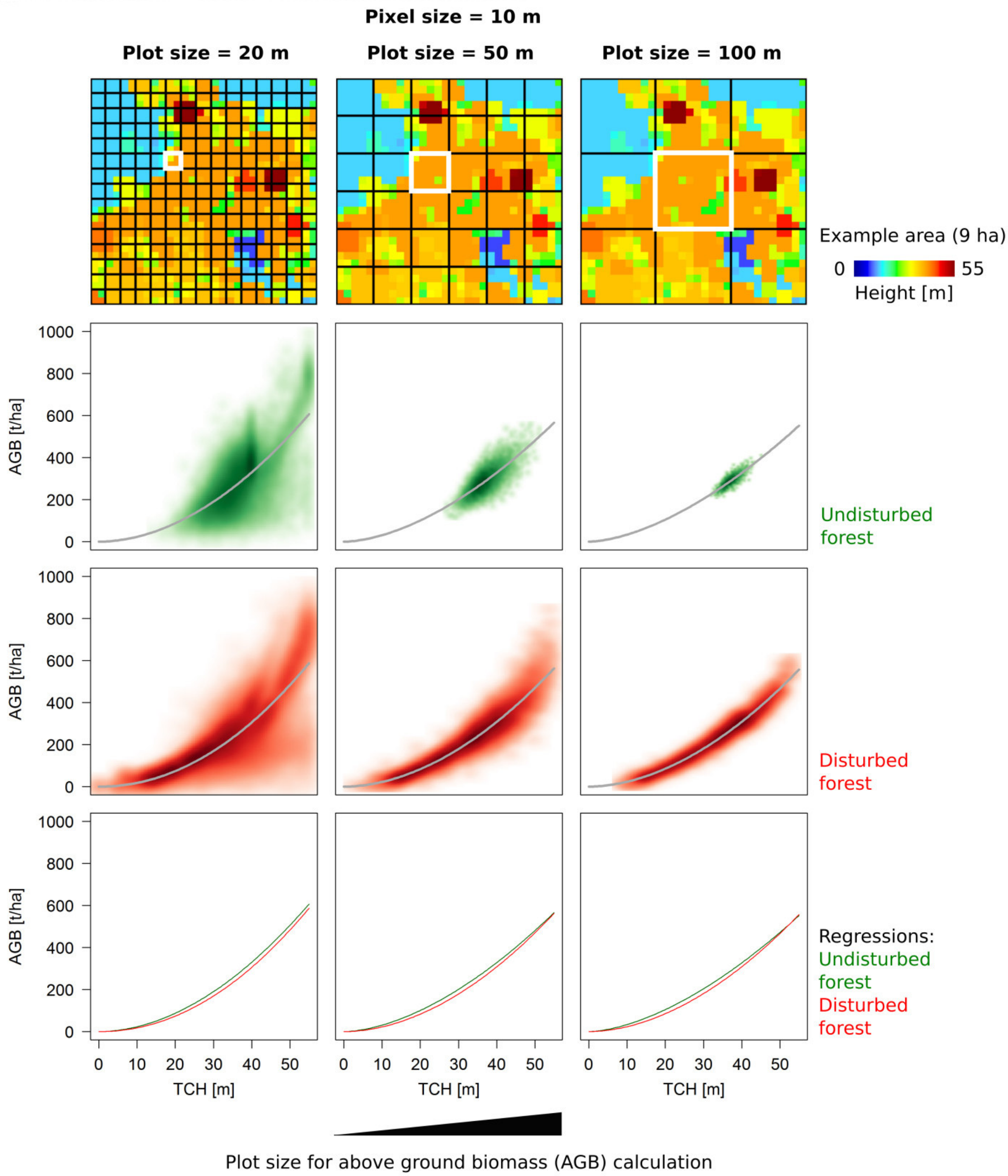

Fig. 6: Aboveground biomass (AGB) as a function of top-of-canopy height (TCH) from 10-m pixel resolution (CHM) for different plot sizes. All data was derived from FORMIND and lidar simulations. 1) The first row demonstrates the sampling approach. Shown is a scene of 9 ha simulated forest with different stages of succession. The following rows show the TCH-to-AGB relationship with each record representing one 20-m, 50-m or 100-m plot, respectively, for 2) 1,400 ha of undisturbed simulated forest (green), 3) 1,400 ha of fire-disturbed and 1,400 ha of regularly logged simulated forest (red) and 4) the curves of the best power law fits.

The general trends were that the nRMSE of the TCH-based AGB predictions increased with decreasing plot size and with increasing pixel size (Fig. 7). The prediction accuracy at each scale was better for the undisturbed forest dataset than for the disturbed forest dataset, indicated by generally lower nRMSE for each plot size and pixel size combination for the undisturbed forest as compared to the disturbed forest (Fig. 7). For the disturbed dataset and large plot sizes (100 and 200 m), we observed slightly better prediction accuracies at medium pixel resolutions (5 and 10 m) than at fine pixel resolutions (1 and 2 m). The analysis shows that to achieve, a plot-level biomass estimation error < 10%, plot sizes of ≥ 100 m are required. At such plot sizes, any pixel size would be sufficient to predict AGB for undisturbed forests with the desired accuracy, but for disturbed forests, the errors exceed 10% and increase strongly at pixel sizes ≥ 20 m.

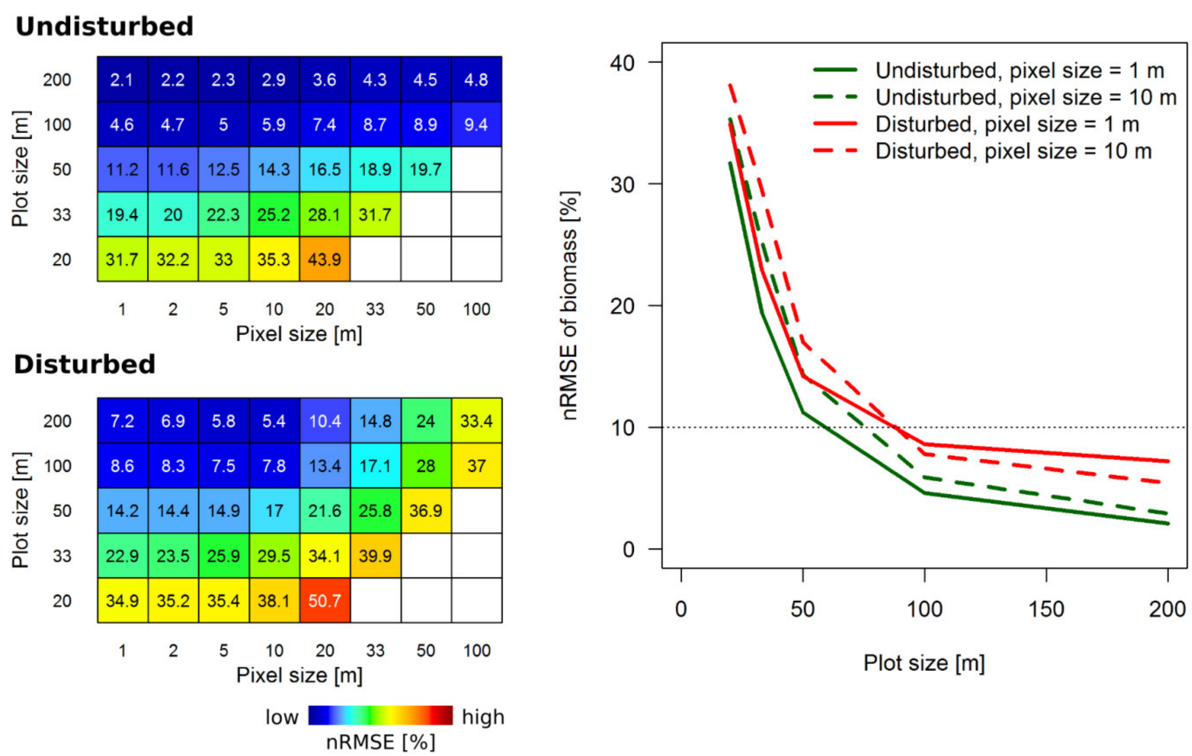

Fig. 7: Normalized root mean square errors (nRMSE) [%] of power law models that describe the relationship between aboveground biomass (AGB) and top-of-canopy height (TCH) at different plot scales and different pixel resolutions for undisturbed and disturbed simulated forest. For pixel sizes of 1 and 10 m, the decrease in nRMSE with increasing plot size is shown on the right side.

3.3 Biomass prediction based on various lidar metrics

In addition to TCH, we analyzed 21 other metrics concerning their capability to predict biomass using power law equations. For this analysis, we no longer distinguished between the different disturbance regimes and pooled all forest stands. Fig. 8 shows nRMSE values for all lidar metrics, for which it was possible to fit a power law model, at the plot scales of 100 and 20 m. From left to right, the metrics are sorted by increasing nRMSE at the 100-m plot size. The figure shows that the best ten metrics are all measures of forest height. Vegetation density metrics (e.g., N_{AGR}/N_{GR} and FCC) and vertical heterogeneity metrics (e.g., SD and Shannon Index) were less accurate AGB predictors than height metrics. The best predictions at large plot scales were achieved by TCH (10 m) and TCH (5 m), whereas at small plot scales RH75, MCH, QMCH and TCH (1 m) were the most accurate predictors. We could not find any relationship between AGB and CV of height, profile skewness or P:H ratio. The Shannon Index of the profiles only showed a relationship with AGB for plot sizes ≥ 50 m. Scatter plots of a selection of metrics against AGB can be found in Fig. S12, nRMSE values for all metrics at all plot scales are displayed in Fig. S13 and detailed statistics and the coefficients of all fit power laws are listed in Tab. S2.

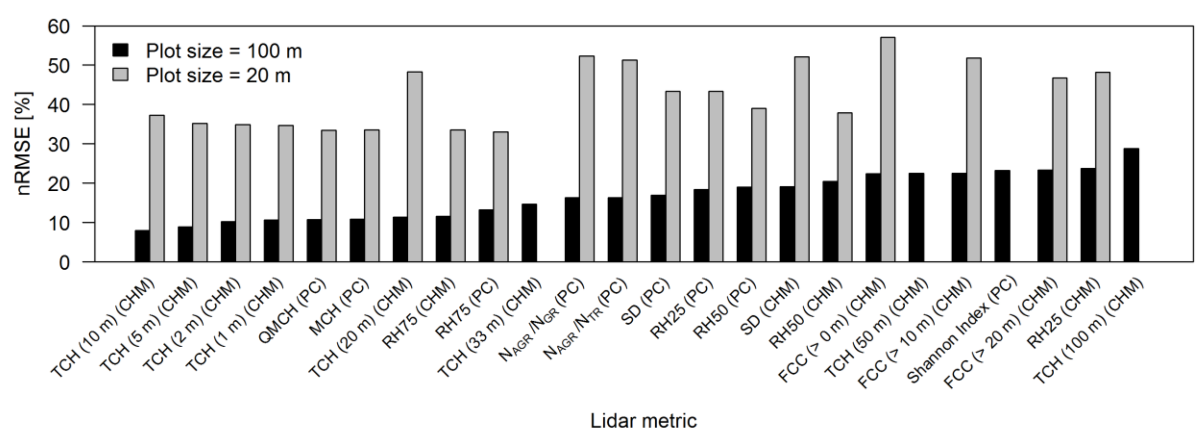

Fig. 8: Normalized root mean square errors (nRMSE) [%] of power law models that describe the relationship between aboveground biomass (AGB) and various lidar metrics (for explanations of the abbreviations, please refer to the main text and Tab. 1) at plot scales of 100 and 20 m, respectively. From left to right, the metrics are sorted by increasing nRMSE at the 100-m plot size. Whether certain metrics were derived from point clouds (PC) or from canopy-height-models (CHM) is indicated in brackets. This analysis was based on pooled (undisturbed and disturbed) simulated forest data and lidar simulations. Missing bars indicate that no power law model could be fit at the 20-m plot size.

4. Discussion

This study demonstrated a new approach for simulating 3D lidar point clouds of forest stands and for investigating structural lidar metrics for their relationship with AGB of a tropical forest using forest simulations. We explored the accuracy of AGB predictions based on various lidar metrics, spatial scales and considering undisturbed and disturbed forest plots.

4.1 Lidar simulations

Unlike other lidar simulation approaches that use detailed radiative transfer theory (Sun et al., 1993; Ni-Meister et al., 2001; Kotchenova et al., 2003; Goodwin et al., 2007) or explicit 3D models of trees and ray tracing (Disney et al., 2010; Endo et al., 2012), our method requires only a minimal parameter set to efficiently compute synthetic lidar point clouds for large areas. Under simple assumptions, e.g., one DBH-to-height and DBH-to-crown-diameter allometry, a constant crown length proportion, cylindrical crowns shapes and a homogeneous leaf area density within crowns, the lidar model was able to reproduce the vertical lidar profiles of different 9-ha subplots within the 50-ha BCI megaplot to an overlap of 87%. An extinction factor k_{NIR} of approximately 0.2 was suggested by empirical measurements (Jones, 2013) and theoretical considerations (Campbell & Norman, 2012; Tang et al., 2012) and could be confirmed by our inverse modeling tests.

Airborne and simulated profiles for the 9-ha subplots matched well in general. They diverged most in the upper canopy, where the DBH-to-height allometry led to an overestimation of high trees. Frequencies of ground returns of simulated profiles were approximately 25% lower than for the airborne data, which could be adjusted by choosing another lidar return probability P_0 for ground voxels. Because the exact size of the ground return peak does not affect most of the lidar metrics, we did not treat ground voxels differently than canopy voxels in this study. It should also be noted that simulated lidar profiles (inventory- and FORMIND-based) contain only returns from trees and ground. Non-woody vegetation such as shrubs and lianas may contribute to the airborne lidar profiles, particularly near ground, whereas they are absent in the simulations.

4.2 Biomass prediction from lidar height

For the simulated BCI lidar dataset, TCH at various pixel resolutions performed better than any other lidar metric for biomass predictions. The lowest AGB prediction errors (< 10%) were found for large mapping units (plot sizes of 100 and 200 m) with TCH derived from CHMs with pixel sizes of 5 to 20 m. For the smaller mapping units of 50 m, 33 m and 20 m, the minimal achievable errors from any metric were 15%, 23% and 33%, respectively. At those scales, the high pixel resolution TCH, RH75 or point-cloud-based MCH and QMCH led to slightly smaller errors than TCH of medium pixel resolution. The finding that medium pixel resolution CHMs are sufficient to make highly accurate AGB predictions at the 1-ha scale is encouraging for spaceborne biomass mapping efforts on the global scale. The generation of high-resolution information (e.g., pixel size of 1 m) requires airborne laser scanning campaigns, whereas medium resolutions can be derived from satellites. The synthetic aperture radar satellite system TanDEM-X can provide forest heights closely correlated to TCH at a resolution of 10 m (referred to as H100 in the radar literature; Kugler et al., 2014; Lee & Fatoyinbo, 2015). Future sensors, such as GEDI (http://science.nasa.gov/missions/gedi/) and Tandem-L (https://www.tandem-l.de/), will provide data of similar horizontal resolution (20 to 50 m) and improved vertical resolution. Thus, TCH as well as MCH and RH75 of the vertical profiles are promising metrics for estimating AGB using these sensors. The analysis also showed that sensors that only provide maximum height at the coarse resolution of 100 m lead to AGB estimation errors of > 25%. It appears highly plausible that CHMs with pixels sizes around 10 m that correspond to the dimensions of the objects of interest, namely crowns of medium to large trees, which contribute most to the total AGB, are a good data source for AGB inference. High-resolution data such as

1-m pixel CHMs or the full point cloud have the advantage of providing detailed information on crown architecture and small gaps, but this information might only be additional noise in the signal for stand level AGB and may not be necessary for large-scale mapping.

4.3 The role of structural metrics

Metrics of vertical heterogeneity (e.g., standard deviation or Shannon Index) and vegetation density (e.g., N_{AGR}/N_{GR} or FCC) showed weaker relationships with AGB than most of the height metrics. Hence, these metrics might not be the optimal choice as single AGB predictors. However, considering vegetation structure in addition to mean height could potentially improve AGB estimations. Several approaches have been suggested to improve power-law-based lidar-to-AGB models by considering additional predictors. These predictors include horizontal and vertical structure indices (Tello et al., 2015) and texture metrics of the CHM (Abdullahi et al., 2016). Finally, when thinking beyond AGB stock prediction and towards the study of forest dynamics and disturbances based on remote sensing, structural metrics may become very important. The Shannon Index of the lidar profile has been previously associated with productivity and mortality (Stark et al., 2012), and gap fraction and size distribution may provide information about disturbances (Lobo & Dalling, 2014).

4.4 Prediction errors

For all tested lidar metrics, we observed the tendency for the prediction errors to decrease with increasing plot scale. This pattern has been reported and quantified previously for MCH (Asner et al., 2010; Mascaro et al., 2011b), QMCH (Chen et al. 2016) and TCH (Köhler & Huth, 2010; Asner & Mascaro, 2014) and in general for the situation in which remote sensing footprints and ground plot extents do not fully match (Réjou-Méchain et al., 2014). In our analysis, the spatial locations and extents of ground plots and remote sensing data matched perfectly, because they were based on simulations. Also there was no displacement of crowns from stem locations. Thus, our dataset is free of geolocation errors and the observed residuals in the lidar-to-AGB relationships can be attributed to the following sources of uncertainty: 1) the highly clumped biomass distribution on the ground, i.e., the majority of biomass is localized in tree trunks at specific positions with empty space in between, whereas remote sensing signals capture the tree crowns, which are spread around the trunk positions; 2) edge effects of overhanging tree crowns with trunk positions and thus biomass being located outside the focal plot area; 3) the general variability among trees with respect to their geometries and wood densities; and 4) the undergrowth vegetation that is obscured by the upper canopy and not detected by the remote sensing sensor. The error caused by 1) should decrease with increasing plot size due to the decrease in biomass variability (Fig. 3) and the decreasing influence of single large trees. The error caused by 2) should decrease with increasing core area to edge length ratio. The error caused by 3) should decrease because differences at the individual tree level average out with increasing plot size. Only errors caused by 4) can be expected to be scale-independent. Using a crown-distributed instead of a stem-localized biomass distribution as ground truth has been shown to reduce estimation errors (Mascaro et al., 2011b). However, the actual biomass distribution in a forest is expected to be closer to being stem-localized than (uniformly) crown-distributed. Thus, reducing errors by assuming crown-distributed biomass does not necessarily lead to more accurate biomass maps. Our modeling approach may allow future studies to gain a closer look at the contributions of the separate error sources by switching them off one at a time. Different lidar metrics showed different changes in errors across scales: e.g., in moving from large to small plots, the errors of TCH20, TCH33 and the Shannon Index increased much faster than for other metrics with similar errors at the 200-m scale (Fig. 8 and S13). For the Shannon Index, the relationship with AGB was entirely lost at scales smaller than 50 m.

4.5 Linking remote sensing with dynamic forest models

Despite the great potential of the proposed approach, relatively few studies have linked remote sensing and forest modeling. Applications include model initialization (Ranson et al., 2001; Hurtt et al., 2004), model parameterization (Falkowski et al., 2010), remote sensing calibration (Köhler & Huth, 2010; Palace et al., 2015), error quantification (Hurtt et al., 2010; Frazer et al., 2011) and the understanding of large-scale ecosystem patterns and processes (Shugart et al., 2015). Our study is the first to demonstrate how remote sensing simulations combined with a dynamic forest model can provide remote sensing metrics over the full range of disturbance-induced successional stages, which is particularly useful for tropical forests where available field data is limited. The lidar-to-AGB relationships can differ between disturbance types because one type (e.g., fire) might cause mosaics of surviving trees and bare ground, whereas another type (e.g., selective logging) might cause a height degradation throughout the entire study area. Horizontal heterogeneities, such as those caused by fires, are particularly problematic when lidar metrics are aggregated over larger areas. Thus, the disturbance regime of a region and the presence of the described phenomena should be taken into account when deciding which metric and resolution to choose for biomass mapping. Modeling can be one way to explore these effects in greater detail.

An important condition for combining a forest model and remote sensing is the structural realism of the model in the relevant aspects. Overall, our model was able to reproduce forest attributes and literature values well. Previous studies on BCI that linked AGB at the 1-ha scale to MCH derived from airborne lidar scans reported RMSE values of 17 t_{Carbon}/ha (Mascaro et al., 2011a) and 28.9 t_{AGB}/ha (Meyer et al., 2013) in agreement with the value of 27.1 t_{AGB}/ha we obtained for the pooled simulated dataset (Tab. S2). A noteworthy deviation between the simulation data and reference data was that for comparable AGB values the simulated TCH was higher than the airborne TCH, particularly at the upper end of the AGB and TCH ranges (described in detail in supplements). We believe that this deviation was primarily caused by the simple tree geometries used in the forest model. Using only one general DBH-to-height allometry for all trees might be suboptimal if the aim is to reproduce the natural height heterogeneity of the upper canopy at all scales. In our simulations, too many trees reached the maximum possible height of 55 m, which is an exceptional height on BCI observed for only one tree in the airborne lidar CHM. Hurtt et al. (2004) encountered a similar problem with large trees. In their case, model-derived canopy heights were restrained to a maximum, whereas observed lidar heights exceeded that limit. Therefore, one potential improvement for future model parameterizations would be to consider asymptotic instead of power law DBH-to-height allometries and allow for a certain plasticity of modeled heights and crown diameters. The sensitivity analysis about model assumptions showed that the alternative scenario using an asymptotic tree height allometry led to slight increases in R^2 and decreases in nRMSE of the stand height to biomass relationship (S16-S18). Recent advances in individual tree delineation from airborne lidar (Duncanson et al., 2014; Ferraz et al., 2016) and terrestrial laser scanning (Raumonen et al., 2013) have the potential to improve our understanding of tree allometries and the structural realism of forest models. When models are able to reproduce observed patterns in the relationship between remote sensing metrics and static biomass stocks, we can move forward using the presented methodology to explore dynamic changes of biomass stocks.

5. Conclusion

This study introduced a novel approach for coupling remote sensing simulations with a dynamic forest model to derive structure-to-biomass relationships for a tropical forest, including disturbed stands. The lidar model was validated successfully with airborne and census reference data from Barro Colorado Island. The model proved its capacity for efficient and realistic lidar point cloud simulations for large, simulated forest stands. Virtual forest inventory datasets were generated with a forest model and sampled with the lidar simulation model. The results provide a comprehensive overview of biomass estimation errors for a wide range of lidar metrics and spatial scales and may guide decisions on which metric to choose for a given remote sensing data structure (e.g., point clouds, vertical profiles, canopy height models). It was found that height-to-biomass relationships were similar for undisturbed and disturbed forest, but errors were larger for the latter. Furthermore, we found that top-of-canopy height was an accurate biomass predictor even if pixel resolutions were only 10 to 20 m. Such resolutions could be derived at large scale from spaceborne sensors.

Acknowledgments

We thank J. Dalling for providing the lidar data and the Smithsonian Tropical Research Institute for providing the census data for BCI. The BCI forest dynamics research project was founded by S.P. Hubbell and R.B. Foster and is now managed by R. Condit, S. Lao, and R. Perez under the Center for Tropical Forest Science and the Smithsonian Tropical Research in Panama. Numerous organizations have provided funding, principally the U.S. National Science Foundation, and hundreds of field workers have contributed. This study was conducted with funding by the German Federal Ministry for Economic Affairs and Energy (BMWi) under the funding reference 50EE1416. RF and AH were supported by the HGF-Helmholtz Alliance “Remote Sensing and Earth System Dynamics”. We thank two reviewers for their constructive comments on our paper.

References

Abdullahi, S., Kugler, F. & Pretzsch, H. (2016). Prediction of stem volume in complex temperate forest stands using TanDEM-X SAR data. *Remote Sensing of Environment*. 174. p.pp. 197–211.

Asner, G.P., Flint Hughes, R., Varga, T.A., Knapp, D.E. & Kennedy-Bowdoin, T. (2009). Environmental and biotic controls over aboveground biomass throughout a tropical rain forest. *Ecosystems*. 12 (2). p.pp. 261–278.

Asner, G.P. & Mascaro, J. (2014). Mapping tropical forest carbon: Calibrating plot estimates to a simple LiDAR metric. *Remote Sensing of Environment*. 140. p.pp. 614–624.

Asner, G.P., Mascaro, J., Muller-Landau, H.C., Vieilledent, G., Vaudry, R., Rasamoelina, M., Hall, J.S. & van Breugel, M. (2012). A universal airborne LiDAR approach for tropical forest carbon mapping. *Oecologia*. 168 (4). p.pp. 1147–1160.

Asner, G.P., Powell, G.V., Mascaro, J., Knapp, D.E., Clark, J.K., Jacobson, J., Kennedy-Bowdoin, T., Balaji, A., Paez-Acosta, G., Victoria, E., Secada, L., Valqui, M. & Hughes, R.F. (2010). High-resolution forest carbon stocks and emissions in the Amazon. *Proceedings of the National Academy of Sciences of the United States of America*. 107 (38). p.pp. 16738–16742.

Bohlman, S. & O'Brien, S. (2006). Allometry, adult stature and regeneration requirement of 65 tree species on Barro Colorado Island, Panama. *Journal of Tropical Ecology*. 22 (02). p.pp. 123–136.

Botkin, D.B., Janak, J.F. & Wallis, J.R. (1972). Some Ecological Consequences of a Computer Model of Forest Growth. *The Journal of Ecology*. 60 (3). p.pp. 849–872.

Bugmann, H. (2001). A review of forest gap models. *Climatic Change*. 51 (3-4). p.pp. 259–305.

Campbell, G.S. & Norman, J.M. (2012). *An Introduction to Environmental Biophysics*. Berlin Heidelberg: Springer Science & Business Media.

Chave, J., Andalo, C., Brown, S. & Cairns, M. (2005). Tree allometry and improved estimation of carbon stocks and balance in tropical forests. *Oecologia*. 145 (1). p.pp. 87–99.

Chave, J., Condit, R., Lao, S., Caspersen, J.P., Foster, R.B. & Hubbell, S.P. (2003). Spatial and temporal variation of biomass in a tropical forest: results from a large census plot in Panama. *Journal of Ecology*. 91 (2). p.pp. 240–252.

Chave, J., Réjou-Méchain, M., Búrquez, A., Chidumayo, E., Colgan, M.S., Delitti, W.B.C., Duque, A., Eid, T., Fearnside, P.M., Goodman, R.C., Henry, M., Martínez-Yrízar, A., Mugasha, W.A., Muller-Landau, H.C., Mencuccini, M., Nelson, B.W., Ngomanda, A., Nogueira, E.M., Ortiz-Malavassi, E., Pélissier, R., Ploton, P., Ryan, C.M., Saldarriaga, J.G. & Vieilledent, G. (2014). Improved allometric models to estimate the aboveground biomass of tropical trees. *Global Change Biology*. 20 (10). p.pp. 3177–3190.

Chen, Q. (2013). Lidar remote sensing of vegetation biomass. In: *Remote sensing of natural resources*. pp. 399–420.

Chen, Q. (2015). Modeling aboveground tree woody biomass using national-scale allometric methods and airborne lidar. *ISPRS Journal of Photogrammetry and Remote Sensing*. 106. p.pp. 95–106.

Chen, Q., McRoberts, R.E., Wang, C. & Radtke, P.J. (2016). Forest aboveground biomass mapping and estimation across multiple spatial scales using model-based inference. *Remote Sensing of Environment*. 184. p.pp. 350–360.

Condit, R. (1998). *Tropical forest census plots*. Berlin, Germany and George Town, Texas: Springer-Verlag and R. G. Landes Company.

Condit, R., Robinson, W.D., Ibáñez, R., Aguilar, S., Sanjur, A., Martínez, R., Stallard, R.F., García, T., Angehr, G.R., Petit, L., Wright, S.J., Robinson, T.R. & Heckadon, S. (2001). The Status of the Panama Canal Watershed and Its Biodiversity at the Beginning of the 21st Century. *BioScience*. 51 (5). p.pp. 389–398.

Disney, M.I., Kalogirou, V., Lewis, P., Prieto-Blanco, A., Hancock, S. & Pfeifer, M. (2010). Simulating the impact of discrete-return lidar system and survey characteristics over young conifer and broadleaf forests. *Remote Sensing of Environment*. 114 (7). p.pp. 1546–1560.

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. (2002). Estimation of tropical forest structural characteristics, using large-footprint lidar. *Remote Sensing of Environment*. 79 (2-3). p.pp. 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. *Journal of Geophysical Research: Biogeosciences*. 115 (2). p.pp. 1–17.

Duncanson, L.I., Cook, B.D., Hurtt, G.C. & Dubayah, R.O. (2014). An efficient, multi-layered crown delineation algorithm for mapping individual tree structure across multiple ecosystems. *Remote Sensing of Environment*. 154. p.pp. 378–386.

Endo, T., Sawada, Y., Kobayashi, T. & Sawada, H. (2012). Developing a 3D Waveform Lidar Simulator for Forest. In: *ISPRS - International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences*. 2012, Melbourne, Australia, pp. 399–402.

Falkowski, M.J., Hudak, A.T., Crookston, N.L., Gessler, P.E., Uebler, E.H. & Smith, A.M.S. (2010). Landscape-scale parameterization of a tree-level forest growth model: a k-nearest neighbor imputation approach incorporating LiDAR data. *Canadian Journal of Forest Research*. 40 (2). p.pp. 184–199.

Fassnacht, F.E., Hartig, F., Latifi, H., Berger, C., Hernández, J., Corvalán, P. & Koch, B. (2014). Importance of sample size, data type and prediction method for remote sensing-based estimations of aboveground forest biomass. *Remote Sensing of Environment*. 154 (1). p.pp. 102–114.

Ferraz, A., Saatchi, S., Mallet, C. & Meyer, V. (2016). Lidar detection of individual tree size in tropical forests. *Remote Sensing of Environment*. 183. p.pp. 318–333.

Fischer, R. (2013). *Modellierung der Dynamik afrikanischer Tropenwälder. Analyse des Einflusses von Störungen auf tropische Wälder mit Hilfe des Waldmodells FORMIND.* Dissertation, Universität Osnabrück, Germany.

Fischer, R., Bohn, F., Dantas de Paula, M., Dislich, C., Groeneveld, J., Gutiérrez, A.G., Kazmierczak, M., Knapp, N., Lehmann, S., Paulick, S., Pütz, S., Rödig, E., Taubert, F., Köhler, P. & Huth, A. (2016). Lessons learned from applying a forest gap model to understand ecosystem and carbon dynamics of complex tropical forests. *Ecological Modelling*. 326. p.pp. 124–133.

Frazer, G.W., Magnussen, S., Wulder, M.A. & Niemann, K.O. (2011). Simulated impact of sample plot size and co-registration error on the accuracy and uncertainty of LiDAR-derived estimates of forest stand biomass. *Remote Sensing of Environment*. 115 (2). p.pp. 636–649.

Gibbs, H.K., Brown, S., Niles, J.O. & Foley, J.A. (2007). Monitoring and estimating tropical forest carbon stocks: making REDD a reality. *Environmental Research Letters*. 2 (2007). p.p. 045023.

Goodwin, N.R., Coops, N.C. & Culvenor, D.S. (2007). Development of a simulation model to predict LiDAR interception in forested environments. *Remote Sensing of Environment*. 111 (4). p.pp. 481–492.

Hubbell, S., Foster, R., O'Brien, S., Harms, K., Condit, R., Wechsler, B., Wright, S. & Loo de Lao, S. (1999). Light-gap disturbances, recruitment limitation, and tree diversity in a neotropical forest. *Science*. 283 (5401). p.pp. 554–557.

Hubbell, S.P., Condit, R. & Foster, R.B. (2005). *Barro Colorado Forest Census Plot Data*. [Online]. 2005. Available from: http://ctfs.si.edu/webatlas/datasets/bci.

Hurtt, G.C., Dubayah, R., Drake, J., Moorcroft, P.R., Pacala, S.W., Blair, J.B. & Fearon, M.G. (2004). Beyond Potential Vegetation: Combining Lidar Data and a Height-Structured Model for Carbon Studies. *Ecological Applications*. 14 (3). p.pp. 873–883.

Hurtt, G.C., Fisk, J., Thomas, R.Q., Dubayah, R., Moorcroft, P.R. & Shugart, H.H. (2010). Linking models and data on vegetation structure. *Journal of Geophysical Research*. 115. p.pp. 1–11.

Huth, A., Drechsler, M. & Köhler, P. (2004). Multicriteria evaluation of simulated logging scenarios in a tropical rain forest. *Journal of Environmental Management*. 71 (4). p.pp. 321–333.

Isenburg, M. (2011). *LAStools - efficient tools for LiDAR processing*.

Jones, H.G. (2013). *Plants and Microclimate - A Quantitative Approach to Environmental Plant Physiology*. Cambridge: Cambridge University Press.

Jubanski, J., Ballhorn, U., Kronseder, K., Franke, J. & Siegert, F. (2013). Detection of large above-ground biomass variability in lowland forest ecosystems by airborne LiDAR. *Biogeosciences*. 10 (6). p.pp. 3917–3930.

Kazmierczak, M., Wiegand, T. & Huth, A. (2014). A neutral vs. non-neutral parametrizations of a physiological forest gap model. *Ecological Modelling*. 288. p.pp. 94–102.

Köhler, P. & Huth, A. (2010). Towards ground-truthing of spaceborne estimates of above-ground life biomass and leaf area index in tropical rain forests. *Biogeosciences*. 7 (8). p.pp. 2531–2543.

Kotchenova, S.Y., Shabanov, N. V, knyazikhin, Y., Davis, A.B., Dubayah, R.O. & Myneni, R.B. (2003). Modeling lidar waveforms with time-dependent stochastic radiative transfer theory for remote estimations of forest structure. *Journal Of Geophysical Research-Atmospheres*. 108 (D15). p.p. 4484.

Kugler, F., Schulze, D., Hajnsek, I., Pretzsch, H. & Papathanassiou, K.P. (2014). TanDEM-X Pol-InSAR performance for forest height estimation. *IEEE Transactions on Geoscience and Remote Sensing*. 52 (10). p.pp. 6404–6422.

Lee, S. & Fatoyinbo, T.E. (2015). TanDEM-X Pol-InSAR Inversion for Mangrove Canopy Height Estimation. *Ieee Journal of Selected Topics in Applied Earth Observations and Remote Sensing*. 8 (7). p.pp. 1–11.

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 Sensing of Environment*. 70 (3). p.pp. 339–361.

Lefsky, M.A., Cohen, W.B., Harding, D.J., Parker, G.G., Acker, S.A. & Gower, S.T. (2002a). Lidar remote sensing of above-ground biomass in three biomes. *Global Ecology and Biogeography*. 11 (5). p.pp. 393–399.

Lefsky, M.A., Cohen, W.B., Parker, G.G. & Harding, D.J. (2002b). Lidar remote sensing for ecosystem studies. *BioScience*. 52 (1). p.pp. 19–30.

Lobo, E. & Dalling, J.W. (2014). Spatial scale and sampling resolution affect measures of gap disturbance in a lowland tropical forest: implications for understanding forest regeneration and carbon storage. *Proceedings. Biological sciences / The Royal Society*. 281. p.p. 20133218.

Lu, D., Chen, Q., Wang, G., Liu, L., Li, G. & Moran, E. (2014). A survey of remote sensing-based aboveground biomass estimation methods in forest ecosystems. *International Journal of Digital Earth*. (December). p.pp. 1–43.

Marvin, D.C., Asner, G.P., Knapp, D.E., Anderson, C.B., Martin, R.E., Sinca, F. & Tupayachi, R. (2014). Amazonian landscapes and the bias in field studies of forest structure and biomass. *Proceedings of the National Academy of Sciences of the United States of America*. 111 (48). p.pp. E5224–32.

Mascaro, J., Asner, G.P., Davies, S., Dehgan, A. & Saatchi, S. (2014). These are the days of lasers in the jungle. *Carbon balance and management*. 9 (1). p.p. 7.

Mascaro, J., Asner, G.P., Dent, D.H., DeWalt, S.J. & Denslow, J.S. (2012). Scale-dependence of aboveground carbon accumulation in secondary forests of Panama: A test of the intermediate peak hypothesis. *Forest Ecology and Management*. 276. p.pp. 62–70.

Mascaro, J., Asner, G.P., Muller-Landau, H.C., Van Breugel, M., Hall, J. & Dahlin, K. (2011a). Controls over aboveground forest carbon density on Barro Colorado Island, Panama. *Biogeosciences*. 8 (6). p.pp. 1615–1629.

Mascaro, J., Detto, M., Asner, G.P. & Muller-Landau, H.C. (2011b). Evaluating uncertainty in mapping forest carbon with airborne LiDAR. *Remote Sensing of Environment*. 115 (12). p.pp. 3770–3774.

Meyer, V., Saatchi, S.S., Chave, J., Dalling, J.W., Bohlman, S., Fricker, G.A., Robinson, C., Neumann, M. & Hubbell, S. (2013). Detecting tropical forest biomass dynamics from repeated airborne lidar measurements. *Biogeosciences*. 10 (8). p.pp. 5421–5438.

Næsset, E. (2009). Effects of different sensors, flying altitudes, and pulse repetition frequencies on forest canopy metrics and biophysical stand properties derived from small-footprint airborne laser data. *Remote Sensing of Environment*. 113 (1). p.pp. 148–159.

Ni-Meister, W., Jupp, D.L.B. & Dubayah, R. (2001). Modeling lidar waveforms in heterogeneous and discrete canopies. *IEEE Transactions on Geoscience and Remote Sensing*. 39 (9). p.pp. 1943–1958.

Palace, M.W., Sullivan, F.B., Ducey, M.J., Treuhaft, R.N., Herrick, C., Shimbo, J.Z. & Mota-E-Silva, J. (2015). Estimating forest structure in a tropical forest using field measurements, a synthetic model and discrete return lidar data. *Remote Sensing of Environment*. 161. p.pp. 1–11.

Pan, Y., Birdsey, R.A., Phillips, O.L. & Jackson, R.B. (2013). The Structure, Distribution, and Biomass of the World's Forests. *Annual Review of Ecology, Evolution, and Systematics*. 44 (1). p.pp. 593–622.

Patenaude, G., Hill, R.A., Milne, R., Gaveau, D.L.A., Briggs, B.B.J. & Dawson, T.P. (2004). Quantifying forest above ground carbon content using LiDAR remote sensing. *Remote Sensing of Environment*. 93 (3). p.pp. 368–380.

Poorter, L., Bongers, F., Aide, T.M., Almeyda Zambrano, A.M., Balvanera, P., Becknell, J.M., Boukili, V., Brancalion, P.H.S., Broadbent, E.N., Chazdon, R.L., Craven, D., de Almeida-Cortez, J.S., Cabral, G.A.L., de Jong, B.H.J., Denslow, J.S., Dent, D.H., DeWalt, S.J., Dupuy, J.M., Durán, S.M., Espírito-Santo, M.M., Fandino, M.C., César, R.G., Hall, J.S., Hernandez-Stefanoni, J.L., Jakovac, C.C., Junqueira, A.B., Kennard, D., Letcher, S.G., Licona, J.-C., Lohbeck, M., Marín-Spiotta, E., Martínez-Ramos, M., Massoca, P., Meave, J.A., Mesquita, R., Mora, F., Muñoz, R., Muscarella, R., Nunes, Y.R.F., Ochoa-Gaona, S., de Oliveira, A.A., Orihuela-Belmonte, E., Peña-Claros, M., Pérez-García, E.A., Piotto, D., Powers, J.S., Rodríguez-Velázquez, J., Romero-Pérez, I.E., Ruíz, J., Saldarriaga, J.G., Sanchez-Azofeifa, A., Schwartz, N.B., Steininger, M.K., Swenson, N.G., Toledo, M., Uriarte, M., van Breugel, M., van der Wal, H., Veloso, M.D.M., Vester, H.F.M., Vicentini, A., Vieira, I.C.G., Bentos, T.V., Williamson, G.B. & Rozendaal, D.M.A. (2016). Biomass resilience of Neotropical secondary forests. *Nature*. 530 (7589). p.pp. 211–214.

R Development Core Team (2014). *R: A Language and Environment for Statistical Computing*.

Ranson, K.J., Sun, G., Knox, R.G., Levine, E.R., Weishampel, J.F. & Fifer, S.T. (2001). Northern forest ecosystem dynamics using coupled models and remote sensing. *Remote Sensing of Environment*. 75 (2). p.pp. 291–302.

Raumonen, P., Kaasalainen, M., Åkerblom, M., Kaasalainen, S., Kaartinen, H., Vastaranta, M., Holopainen, M., Disney, M. & Lewis, P. (2013). Fast Automatic Precision Tree Models from Terrestrial Laser Scanner Data. *Remote Sensing*. 5 (2). p.pp. 491–520.

Réjou-Méchain, M., Muller-Landau, H.C., Detto, M., Thomas, S.C., Le Toan, T., Saatchi, S.S., Barreto-Silva, J.S., Bourg, N.A., Bunyavejchewin, S., Butt, N., Brockelman, W.Y., Cao, M., Cárdenas, D., Chiang, J.M., Chuyong, G.B., Clay, K., Condit, R., Dattaraja, H.S., Davies, S.J., Duque, A., Esufali, S., Ewango, C., Fernando, R.H.S., Fletcher, C.D., N. Gunatilleke, I.A.U., Hao, Z., Harms, K.E., Hart, T.B., Hérault, B., Howe, R.W., Hubbell, S.P., Johnson, D.J., Kenfack, D., Larson, A.J., Lin, L., Lin, Y., Lutz, J.A., Makana, J.R., Malhi, Y., Marthews, T.R., Mcewan, R.W., Mcmahon, S.M., Mcshea, W.J., Muscarella, R., Nathalang, A., Noor, N.S.M., Nytch, C.J., Oliveira, A.A., Phillips, R.P., Pongpattananurak, N., Punchi-Manage, R., Salim, R., Schurman, J., Sukumar, R., Suresh, H.S., Suwanvecho, U., Thomas, D.W., Thompson, J., Uríarte, M., Valencia, R., Vicentini, A., Wolf, A.T., Yap, S., Yuan, Z., Zartman, C.E., Zimmerman, J.K. & Chave, J. (2014). Local spatial structure of forest biomass and its consequences for remote sensing of carbon stocks. *Biogeosciences*. 11 (23). p.pp. 6827–6840.

Shugart, H.H. (1984). *A theory of forest dynamics: the ecological implications of forest succession models*. New York, USA: Springer.

Shugart, H.H., Asner, G.P., Fischer, R., Huth, A., Knapp, N., Le Toan, T. & Shuman, J.K. (2015). Computer and remote-sensing infrastructure to enhance large-scale testing of individual-based forest models. *Frontiers in Ecology and the Environment*. 13 (9). p.pp. 503–511.

Stark, S.C., Leitold, V., Wu, J.L., Hunter, M.O., de Castilho, C. V., Costa, F.R.C., Mcmahon, S.M., Parker, G.G., Shimabukuro, M.T., Lefsky, M.A., Keller, M., Alves, L.F., Schietti, J., Shimabukuro, Y.E., Brand??o, D.O., Woodcock, T.K., Higuchi, N., de Camargo, P.B., de Oliveira, R.C. & Saleska, S.R. (2012). Amazon forest carbon dynamics predicted by profiles of canopy leaf area and light environment. *Ecology Letters*. 15 (12). p.pp. 1406–1414.

Sun, G., Ranson, K.J. & Park, C. (1993). Modeling Lidar returns from vegetation canopies. *IEEE Transactions on Geoscience and Remote Sensing*. 3504 (1990). p.pp. 67–74.

De Sy, V., Herold, M., Achard, F., Asner, G.P., Held, A., Kellndorfer, J. & Verbesselt, J. (2012). Synergies of multiple remote sensing data sources for REDD+ monitoring. *Current Opinion in Environmental Sustainability*. 4 (6). p.pp. 696–706.

Tang, H., Dubayah, R., Swatantran, A., Hofton, M., Sheldon, S., Clark, D.B. & Blair, B. (2012). Retrieval of vertical LAI profiles over tropical rain forests using waveform lidar at La Selva, Costa Rica. *Remote Sensing of Environment*. 124. p.pp. 242–250.

Tello, M., Cazcarra-Bes, V., Pardini, M. & Papathanassiou, K. (2015). Structural classification of forest by means of L-band tomographic SAR. *International Geoscience and Remote Sensing Symposium (IGARSS)*. p.pp. 5288–5291.

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 Sensing of Environment*. 121. p.pp. 196–209.

LIST OF FIGURE CAPTIONS

Fig. 1: Workflow of the study. Reference data from field inventories and an airborne lidar campaign were used to parameterize and calibrate a forest model and a lidar model. With the models, large quantities of simulated inventory and simulated lidar data were generated, allowing for a systematic analysis of lidar-to-biomass relationships under different disturbance regimes and for various spatial scales.

Fig. 2: Principle of the lidar model. Inputs to the workflow can either be forest model output or field inventory data. The pictures on the right side show intermediate products: a) Visualization of a forest stand; b) voxel representation with colors indicating the cumulative leaf area index; c) voxel representation with colors indicating the probability of containing a lidar return; d) simulated lidar point cloud with colors indicating height above ground.

Fig. 3: Relative frequency distributions of aboveground biomass (AGB). Columns represent the BCI field data (50 ha) and output of FORMIND simulations from different disturbance scenarios (1,400 ha each). Rows represent different spatial resolutions. Notice the different axis scaling in each row.

Fig. 4: Vertical lidar profiles of a) the 9 ha in the southwestern corner of the BCI megaplot, airborne and simulated based on inventory data; b) the same for the 9 ha in the northeastern corner of the BCI megaplot; and c) the simulated lidar profile of 16 ha simulated forest in FORMIND in the old growth stage (age 500 yr). Dashed lines mark the mean canopy profile height (MCH), and '×' symbols mark the ground return peaks.

Fig. 5: Aboveground biomass (AGB) as a function of top-of-canopy height (TCH) from 1-m pixel resolution (CHM) for different plot sizes. All data was derived from FORMIND and lidar simulations. 1) The first row demonstrates the sampling approach. Shown is a scene of 9 ha simulated forest with different stages of succession. The following rows show the TCH-to-AGB relationship with each record representing one 20-m, 50-m or 100-m plot, respectively, for 2) 1,400 ha of undisturbed simulated forest (green), 3) 1,400 ha of fire-disturbed and 1,400 ha of regularly logged simulated forest (red) and 4) the curves of the best power law fits.

Fig. 6: Aboveground biomass (AGB) as a function of top-of-canopy height (TCH) from 10-m pixel resolution (CHM) for different plot sizes. All data was derived from FORMIND and lidar simulations. 1) The first row demonstrates the sampling approach. Shown is a scene of 9 ha simulated forest with different stages of succession. The following rows show the TCH-to-AGB relationship with each record representing one 20-m, 50-m or 100-m plot, respectively, for 2) 1,400 ha of undisturbed simulated forest (green), 3) 1,400 ha of fire-disturbed and 1,400 ha of regularly logged simulated forest (red) and 4) the curves of the best power law fits.

Fig. 7: Normalized root mean square errors (nRMSE) [%] of power law models that describe the relationship between aboveground biomass (AGB) and top-of-canopy height (TCH) at different plot scales and different pixel resolutions for undisturbed and disturbed simulated forest. For pixel sizes of 1 and 10 m, the decrease in nRMSE with increasing plot size is shown on the right side.

Fig. 8: Normalized root mean square errors (nRMSE) [%] of power law models that describe the relationship between aboveground biomass (AGB) and various lidar metrics (for explanations of the abbreviations, please refer to the main text and Tab. 1) at plot scales of 100 and 20 m, respectively. From left to right, the metrics are sorted by increasing nRMSE at the 100-m plot size. Whether certain metrics were derived from point clouds (PC) or from canopy-height-models (CHM) is indicated in brackets. This analysis was based on pooled (undisturbed and disturbed) simulated forest data and lidar simulations. Missing bars indicate that no power law model could be fit at the 20-m plot size.

LIST OF TABLE CAPTIONS

Tab. 1: List of the lidar metrics and the underlying data (PC = point cloud, CHM = canopy height model). CHM usually refers to 1-m resolution rasters, except for TCH where various resolutions were tested.