Elsevier Editorial System(tm) for Remote Sensing of Environment
Manuscript Draft

Manuscript Number:

Title: Mapping Aboveground Biomass Uncertainty of African Tropical Forests: Propagating Errors from Tree to Plot to Landscape

Article Type: Original Research Paper

Keywords: Uncertainty, error, aboveground biomass, carbon, lidar, tropical forests

Corresponding Author: Prof. Qi Chen,

Corresponding Author's Institution: University of Hawai`i at Manoa

First Author: Qi Chen

Order of Authors: Qi Chen; Gaia Vaglio Laurin, PhD; Riccardo Valentini, PhD

Letter to the Editor

August 8, 2014

Prof. Marvin Bauer
Editor-in-Chief
Remote Sensing of Environment
University of Minnesota
1530 Cleveland Avenue N.
St. Paul, MN 55108-6112

Dear Prof. Bauer,

I am submitting a manuscript entitled *Mapping Aboveground Biomass Uncertainty of African Tropical Forests: Propagating Errors from Tree to Plot to Landscape* to your prestigious journal. Characterizing the error of aboveground biomass (AGB) has become one of the central topics in global climate change science and policy. This manuscript presents a novel framework for analyzing the errors of remotely sensed AGB. The study is comprehensive in that it includes the errors from all major steps of remote sensing AGB mapping. One unique aspect of our method is that we derived the analytical formula to characterize the errors related to model residuals, model parameters, and model predictors, for both allometric model development and lidar-biomass modeling. I hope this manuscript is of interest to the remote sensing and global climate change communities. Thank you very much for your consideration.

Sincerely,

Qi Chen
Associate Professor
Department of Geography
University of Hawaii at Manoa
(Tel): 808-9563524
(Email): qichen@hawaii.edu

Highlights:

A new method of quantifying errors of remotely-sensed AGB was proposed.

Errors in the whole workflow of AGB mapping were considered.

Models need to consider errors related to residuals, parameters, and predictors.

AGB errors will be underestimated if ignoring plot-level AGB uncertainty.

Lidar-based AGB prediction error is large over African tropical forests.

Mapping Aboveground Biomass Uncertainty of African Tropical Forests: Propagating Errors from Tree to Plot to Landscape

Qi Chen [a,*], Gaia Vaglio Laurin [b], Riccardo Valentini [b,c]

[a] Department of Geography, University of Hawai`i at Mānoa, 422 Saunders Hall, 2424 Maile Way, Honolulu, HI, 96822, USA

[b] CMCC - Centro Mediterraneo sui i Cambiamenti Climatici, via Augusto Imperatore (Euro-Mediterranean Center for Climate Change), IAFENT Division, via Pacinotti 5, Viterbo 01100, Italy

[c] Department for Innovation in Biological, Agro-food and Forest Systems, Tuscia University, Viterbo 01100, Italy

[*] Corresponding author: Email: qichen@hawaii.edu; Phone: 1-808-956-3524; Fax: 1- 808-956-3512

Abstract:

Quantifying the uncertainty of the aboveground biomass (AGB) and carbon (C) stock is crucial for understanding the global C cycle and implementing the United Nations Program on Reducing Emissions from Deforestation and Forest Degradation (UN-REDD). The uncertainty analysis of remotely-sensed AGB is tricky because, if validation plots or cross-validation are used for error assessment, the AGB of validation plots does not necessarily represent the actual measurements but estimates of the true AGB. Leveraging a recently published pan-tropical destructively measured tree AGB database, this study proposed a new method of characterizing the uncertainty of the remotely-sensed AGB. The method propagates errors from tree- to landscape-level by considering errors in the whole workflow of AGB mapping process, including allometric model development, tree measurements, tree-level AGB prediction, plot-level AGB estimation, plot-level remote-sensing-based biomass model development, remote sensing feature extraction, and pixel-level AGB prediction. Applying such a method to the tree AGB mapped using airborne lidar over tropical forests in Ghana, we found that the AGB prediction error is over 20% at 1 ha spatial resolution, larger than the results reported in previous studies for other tropical forests. The discrepancy between our studies and other reflects not only our focus on African tropical forests but also the methodological differences in error analysis. This study also highlights the importance of considering the plot-level AGB estimate uncertainty when field plots are used to calibrate remote sensing based biomass models.

Keywords: Uncertainty, error, aboveground biomass, carbon, lidar, tropical forests

1. Introduction

Tropical forests contain ~50% of the aboveground carbon (C) in global vegetation (Hunter et al., 2013; Fischer et al., 2014), account for ~33% of terrestrial net primary productivity (Bonan 2008), and play a crucial role in global C cycle and climate change (Bonan 2008; Lewis et al., 2009; Bagley et al., 2014; Grace et al., 2014). Tropical forests have also been experiencing intense pressure from land use changes such as deforestation and degradation (Cochrane 2003; Fearnside and Laurance 2004; Asner et al., 2009; Berenguer et al., 2014; Hernández-Stefanoni et al., 2014). However, substantial uncertainty remains in estimating tropical forest C emissions from those human activities (Clark et al., 2011). Because land use change is a patchy process (Ometto et al., 2014), accurately mapping the spatial distribution of tropical C stock and its dynamics is vital to reduce such uncertainty (Houghton et al., 2009; Clark et al., 2011; Aguiar et al., 2012; Achard et al., 2014). Remote sensing is a promising technology to achieve this goal with its ability of providing synoptic view of the whole study area (DeFries et al., 2007; Gibbs et al., 2007; Clark et al., 2011; Lu et al., 2012; Baccini et al., 2012; Chen, 2013; Asner et al., 2013).

Considerable efforts have been devoted to map tropical forest biomass at the landscape (e.g., Dubayah et al., 2010; Clark et al., 2011; Mascaro et al., 2011; Asner et al., 2012a, 2012c; d'Oliveira et al., 2012; Kronseder et al., 2012; Ioki et al., 2014; Vaglio Laurin et al., 2014), national (Asner et al., 2012b; Asner et al., 2013), continental (Goetz et al., 2009; Baccini et al., 2008), and even cross-continental (Saatchi et al., 2011; Baccini et al., 2012) scales using remote sensing technology. However, accompanying with the sheer number of biomass mapping studies is the substantial variations among the various estimates of biomass and C stock (Houghton et al.,

2001; Mitchard et al., 2013; Ometto et al., 2014), which makes it difficult to choose a product for making forest management decision in mitigating the impacts of climate change.

Central to understanding the quality of remotely sensed biomass and C maps is to quantify the uncertainty of the estimated biomass from remote sensing based models (Wang et al., 2009; Lu et al., 2012). Root mean square errors (RMSE) is the most common statistic to characterize the error of remote sensing based biomass models (Zolkos et al., 2013) and it is calculated by comparing model prediction to "true" biomass over a sample of forest plots. One of the key distinctions of mapping biomass, compared to mapping many other vegetation attributes such as tree height and basal area, is that the ground truth biomass for calibrating a remote sensing model has rarely been directly measured (Clark and Kellner, 2012). Instead, it is estimated using allometric models with other tree- and site-level attributes, such as DBH (diameter at breast height), tree height and wood density, as predictors. Both the allometric model and tree attributes could have errors, which can be propagated to the plot-level biomass estimates and thus affect the uncertainty of the biomass estimation from a remote sensing-based model. Only a few studies (e.g., Gonzalez et al., 2010) have considered the uncertainty in plot-level biomass estimates in their biomass modeling.

Remotely-sensed biomass mapping involves the combined use of two types of models: 1) allometric models for estimating tree- and plot-level biomass using tree attributes such as DBH, tree height, and wood density, 2) models for predicting plot-level biomass using remote sensing derived variables. Both models have parameters, the errors of which could lead to errors in biomass estimation. The omission of model parameter errors will underestimate the biomass

77 prediction errors (Yanai et al., 2010). Unfortunately, the impacts of model parameter errors on
78 biomass prediction have rarely been examined in the literature.

79 The main goal of this study is to develop a methodology to map the uncertainty of tree
80 aboveground biomass (AGB) using airborne lidar over western African tropical forests in Ghana.
81 The uncertainty estimate will consider errors in the whole process of upscaling biomass from tree
82 to plot and landscape levels with a coordinated use of field measurements, allometric models,
83 lidar data, and statistical modeling. In particular, this study addresses these questions: 1) what the
84 errors associated with allometric models and lidar-based biomass models are, 2) how the errors
85 of tree measurements collected in a forest plot will be propagated to the biomass estimates at the
86 tree- and plot-level when an allometric model is used to predict biomass, 3) how the errors in
87 lidar metrics will be propagated to AGB prediction, 4) how the plot-level AGB errors affect the
88 lidar-biomass AGB modeling and prediction errors, and 5) what the major error sources in AGB
89 prediction at the tree- and pixel-level are.

90 ## 2. Study area and data

91 ### 2.1. Study area

92 Our study area transverses transects along a ~100 km latitudinal gradient in Southwest Ghana
93 close to the border with Ivory Coast (Fig. 1). These transects are along the orbits of ICESat and
94 were mapped with airborne lidar with width of ~250 m to 750 m. The first group of transects are
95 located in the Bia Conservation Area that comprises of Bia National Park (BNP, northern part)
96 and Bia Resource Reserve (BRR, southern part). The area covers the transition between two of
97 Ghana's forest types, Moist Evergreen forest in the south and Moist Semi-deciduous forests in

98 the north (Hall and Swaine 1981). BRR was logged in 1980-90, and possibly even after; it can be
99 impacted by natural (fire, elephants' damages) and illegal human-related disturbance. BNP has a
100 better protection status and no logging records, but fires, elephants' damages and illegal access
101 could occur.

102 The second group of transects is located in the Dadieso Forest Reserve (DFR), which lies south
103 of the Bia Conservation Area but north of Boin river Forest Reserve and Disue Forest Reserve.
104 The vegetation of the reserve is transitional between Moist Evergreen and Wet Evergreen types.
105 DFR was illegally logged and surrounded by communities and coffee farms; furthermore it has
106 swampier characteristics, and flooding can represent a frequent natural disturbance.

107 **2.2. Field measurements**

108 Along the ICESat orbits, the field plots were set up at the footprints of GLAS laser shots with the
109 goal of upscaling biomass from local to regional scale. The GLAS waveforms were first
110 screened to exclude the shots that are saturated or contaminated by clouds (Chen, 2010). So, the
111 plots can be considered as a quasi-transect sample of the forests.

112 The field plots have a square shape of 40 m by 40 m. For each plot, diameter at breast height
113 (DBH), height (H), and species information was collected for all trees having DBH>20 cm. For
114 trees with DBH in the 10-20 cm range, the same information was collected in subplots of 400 m^2.
115 Wood density was estimated using Chave et al. (2009). We used a total of 36 field plots in our
116 analysis (13 in BNP, 3 in BRR, and 20 in DFR).

117 **2.3. Airborne lidar data**

The study area was surveyed by an airborne campaign in March 2012 over pre-defined flight lines covering the field plots, using a Pilatus PC-6 Porter aircraft equipped with lidar and hyperspectral sensors and a digital camera for aerial photographs. The lidar sensor ALTM GEMINI (Optech Ltd.), characterized by a 1064 nm laser wavelength and able to record up to 4 range measurements, was operated between 650 and 850 m above ground level. The minimum laser density was set to 11 points/m^2. The lidar dataset was delivered as a point cloud of discrete returns, preprocessed in Terrascan (Terrasolid) software and adopted the Applanix IN-Fusion™ PPP Inertially-Aided Precise Point Positioning (IAPPP) to cope with the of absence of GPS base stations in the region. Positional error in x, y, z was lower than 0.27 m.

The raw all-returns point cloud was processed using the Toolbox for Lidar Data Filtering and Forest Studies (TIFFS) (Chen, 2007) to derive a range of metrics for AGB estimation from each plot, including: mean height, quadratic mean height, standard deviation height, height bins at 5 m intervals and 10% percentile heights. TIFFS used the ground returns identified by the data provider to generate a DTM (Digital Terrain Model) and calculated the relative height above terrain of each laser return by subtracting the corresponding DTM elevation from its original Z value. The lidar metrics were derived using the relative height of all laser points. We generated lidar metric maps of 40 m by 40 m cell size, equivalent to the field plot size.

2.4. Pan-tropical tree AGB database

Instead of using published allometric equations for estimating tree AGB, we developed an allometric model from a pan-tropical destructive tree database compiled by Chave et al. (2014) to fully characterize the tree AGB prediction errors. This database includes a total of 4004 trees

139 from 53 undisturbed and five secondary forest sites across tropics in Latin America, Southeast
140 Asia, and Africa. The tree measurements include DBH, tree height, wood density, and total
141 oven-dry AGB. This database is an improvement of the earlier compilation (Chave et al., 2005)
142 by adding more than 30 new sites and including some sites from Feldpausch et al. (2012). The
143 database can be downloaded from http://chave.ups-tlse.fr/pantropical_allometry.htm. We call
144 this database Chave14 hereinafter.

145 ## 3. Methods

146 In this section, we will introduce how we analyze errors for each step of biomass mapping
147 process.

148 ### 3.1. Errors of tree-level AGB prediction

149 We first developed a pan-tropical allometric model from the Chave14 tree database using an
150 approach different from log-transformation (Chave et al., 2014). An allometric model is to
151 predict AGB using other easily measurable tree attributes such as DBH, tree height, and wood
152 density (denoted as *x* as a whole). The model is usually calibrated from a sample of trees for
153 which AGB has been measured via destructive sampling and *x* has been obtained by direct
154 measurements or estimation. A statistical model is fitted between AGB and *x* of the sampled
155 trees:

156 $$E(AGB_{tree,i} \mid x_i) = f_{tree}(\beta, x_i) \quad (1)$$

157 $$\mathrm{var}(AGB_{tree,i} \mid x_i) = \sigma_{\varepsilon,tree,i}^{\ 2} \quad (2)$$

where *E*() and var() represent the expectation and variance of a variable; f_{tree} is the allometric equation with parameter β to predict tree AGB; $\sigma_{\varepsilon,tree,i}^2$ is the expected variance of AGB for trees with attributes x_i; *i* is the index of a tree from the sample and it varies between 1 and 4004 (the number of trees in the Chave14 tree database) in this study. Note that β and x_i could be scalars or vectors $[\beta_1, \beta_2, ...]$ and $[x1_i, x2_i, ...]$, depending on whether the model uses multiple parameters and multiple tree attributes for biomass prediction.

One issue of developing AGB allometric models is that $\sigma_{\varepsilon,tree,i}^2$ often increases with tree size, implying that the assumption of homoscedasticity (constant variance across the range of biomass variation) is violated if OLS (ordinary least squares) method is used for model fitting. A simple remedy is to transform both the response variable (AGB) and predictors (*x*) to the logarithmic scale and then fit the model using OLS (Parresol, 1999; Basuki et al., 2009; Djomo et al., 2010; Clifford et al., 2013; Chave et al., 2014). The problem with log-transformation, however, is that the "naive" back-transformation of the log-scale AGB estimates is biased. A user has to choose among different bias correction methods (Finney, 1941; Baskerville, 1972; Sprugel, 1983; Snowdon, 1991). A more severe problem is that, if one wants to assess the impacts of model parameter β errors on AGB prediction, the log-scale parameter errors have to be converted to the original scale as well, a mathematical issue seldom addressed in the literature.

To circumvent the problems with the log-transformation, we developed the allometric model using a Generalized Least Squares (GLS) based approach that can model AGB without bias and easily express AGB and model parameter errors (Brown et al., 1989).

178 $$AGB_{tree,i} = f_{tree}(\beta, x_i) + \varepsilon_{tree,i} \quad (3)$$

179 $$\varepsilon_{tree,i} = g(\beta, x_i, \theta)\xi \quad (4)$$

180 where $\varepsilon_{tree,i}$ is the residual between the measured and estimated AGB for tree *i*; $g(\beta, x_i, \theta)$ is a
181 function of x_i with parameters β and θ; ξ is a standard normal variable. The common forms for
182 *g* are linear, proportional, or power functions of $f(\beta, x)$. By inspection of residuals of an OLS
183 model (see Fig. 2), we used the proportional model:

184 $$g(\beta, x_i, \theta) = \theta f_{tree}(\beta, x_i) \quad (5)$$

185 Note that, since ξ is a standard normal variable, the standard deviation of the residuals $\varepsilon_{tree,i}$ is:

186 $$\sigma_{\varepsilon,tree,i} = \theta f_{tree}(\beta, x_i) \quad (6)$$

187 We used the same model form as the one in Chave et al. (2014) for $f_{tree}(\beta, x_i)$:

188 $$f_{tree}(\beta, x_i) = \beta_1(\rho_i D_i^2 H_i)^{\beta_2} \quad (7)$$

189 where ρ_i, D_i, and H_i are wood density, DBH, and tree height, respectively, for tree *i* in
190 Chave14. With the choice of such a model, $\beta = [\beta_1, \beta_2]$ and $x_i = [\rho_i, D_i, H_i]$. The detailed steps
191 of fitting the allometric model using GLS are described in Appendix 2. The estimated model
192 parameters are denoted as β_1, β_2, and θ.

193 An important distinction has to be made between the AGB estimation errors of 1) a tree from the
194 sample for model development (i.e., from the Chave14 tree database) and 2) a *new* tree (i.e., not
195 in the Chave14 database) for which AGB is to be predicted. The AGB error of the former is
196 $\sigma_{\varepsilon,tree,i}$. However, the error of the latter is called tree AGB *prediction error*, which is always
197 larger than $\sigma_{\varepsilon,tree,i}$ due to 1) errors in model parameters, and 2) errors in the measurements of tree
198 attributes *x*. The prediction error is what we should be concerned about when the allometric
199 model is applied to a new place for biomass estimation. The next two subsections 3.1.1 and 3.1.2
200 discuss the AGB prediction errors of the allometric model.

201 **3.1.1. Errors of AGB prediction associated with model parameters**

202 When the estimated parameter β_1 and β_2 are plugged into Eq. 7 to estimate the AGB of a *new*
203 tree (let its *x* measurements be ρ, D, and H):

204 $$AGB_{tree} = \beta_1(\rho D^2 H)^{\beta_2} \quad (8),$$

205 the estimated AGB_{tree} is uncertain due to the errors in the estimates of parameters. The
206 advantage of using GLS is that it can directly estimate model parameter errors. The parameter
207 errors can be explicitly expressed in the form of variance-covariance matrix $\text{cov}(\beta_j, \beta_k)$, where
208 $j, k \in [1, 2]$ for Eq. (8). The AGB_{tree} error caused by parameter errors is denoted as $\sigma_{f,tree}$. The
209 formula to calculate $\text{cov}(\beta_j, \beta_k)$ and $\sigma_{f,tree}$ are described in Appendix 3.

Note that AGB_{tree} essentially represents the *mean* (or expectation of) biomass of trees with the same attributes x. $\sigma_{f,tree}$ characterizes the "standard error" of the *mean*, not the error of *individual* tree AGB. If AGB_{tree} is used as the AGB estimate at the individual level, its error is larger:

$$\sigma_{tree}^2 = \sigma_{\varepsilon,tree}^2 + \sigma_{f,tree}^2 \quad (9)$$

where $\sigma_{\varepsilon,tree}^2$ is calculated using Eq. 6 for the new tree; σ_{tree} is the tree AGB prediction error of the allometric model.

3.1.2. Errors of AGB prediction associated with the measurements of tree attributes

The σ_{tree} shown in Eq. 9 is derived based on the assumption that no errors exist in the measurements of *new* tree attributes *x*. In fact, tree attributes *x* can be measured or estimated with errors. Because both σ_{ε} and σ_f are dependent on *x*, the measurement errors in *x* will be propagated to the tree AGB estimates. We found that, when the *x* errors are considered, the tree AGB prediction error σ_{tree} is as follows:

$$\sigma_{tree}^2 = \sigma_{\varepsilon,tree}^2 + \sigma_{f,tree}^2 + (\theta^2 + 1)\,\mathrm{var}(f_{tree}(\beta, x)) + \mathrm{var}(\sigma_{f,tree}) \quad (10)$$

where $\mathrm{var}(f_{tree}(\beta, x))$ can be solved using first-order Taylor Series expansion (Gertner et al., 1995) and $\mathrm{var}(\sigma_{f,tree})$ can be estimated using Monte-Carlo simulation (Gonzalez et al., 2010). See Appendix 4 for the derivation of Eq. 10 and the calculation of $\mathrm{var}(f_{tree}(\beta, x))$. Note that

when tree measurements x have no errors, both $\mathrm{var}(f_{tree}(\beta, x))$ and $\mathrm{var}(\sigma_{f,tree})$ are equal to zero, making Eq. 10 equivalent to Eq. 9. We denote $(\theta^2+1)\mathrm{var}(f_{tree}(\beta, x)) + \mathrm{var}(\sigma_{f,tree})$ as $\sigma_{x,tree}{}^2$ to indicate that these components of tree AGB prediction errors are caused by errors in tree attributes *x*.

3.2. Errors of plot-level AGB estimated using field data

Eqs. 8-10 can predict tree AGB and errors for the field plots used for developing remote sensing based biomass models. The predicted plot-level AGB density AGB_{plot} is the sum of individual tree AGB_{tree} divided by the plot area:

$$AGB_{plot,j} = \sum_{i=1}^{n_j} AGB_{tree,i} / s_j \quad (11)$$

where $AGB_{plot,j}$ and s_j are the AGB density and area of plot *j*; n_j is the number of trees in plot *j*; $AGB_{tree,i}$ is calculated using Eq. 8 for tree *i* in plot *j*.

The standard deviation of $AGB_{plot,j}$ is:

$$\sigma_{AGB_{plot,j}} = \sqrt{\sum_{i=1}^{n_j} \sigma_{tree,i}{}^2} / s_j \quad (12)$$

where σ_{tree} is the tree AGB error calculated using either Eq. 9 or 10.

3.3. Errors of AGB estimates from lidar-based biomass model

Using the plot-level AGB estimates $AGB_{plot,j}$ and lidar metrics (denoted as z_j for plot *j*), a lidar-based biomass prediction model can be developed:

$$AGB_{plot,j} = f_{plot}(\phi, z_j) + \varepsilon_{plot,j} \quad (13)$$

where ϕ is the lidar-biomass model parameter(s); $\varepsilon_{plot,j}$ is the residual term of the model; ϕ and z_j could be scalars or vectors, depending on the model form and the number of lidar metrics used as predictors; *j* is the index of the plots used for model development. Since we have a total of 36 field plots, *j* varies between 1 and 36.

We chose the multiplicative power models (Chen et al., 2012; Vaglio Laurin et al., 2014) for biomass estimation:

$$f_{plot}(\phi, z_j) = \phi_0 z1_j^{\phi_1} z2_j^{\phi_2} ... zm_j^{\phi_m} \quad (14)$$

where $z1$, $z2$, ..., and zm are different lidar metrics. With such a model form, $\phi = [\phi_0, \phi_1, \phi_2, ..., \phi_m]$ and $z_j = [z1_j, z2_j, ..., zm_j]$.

We found that $\varepsilon_{plot,j}$ shows heteroscedasticity pattern when OLS was used for model fitting. Therefore, we modeled the residuals directly:

$$\varepsilon_{plot,j} = k f_{plot}(\phi, z_j)\xi \quad (15)$$

where *k* is a parameter, ξ is a standard normal variable. Correspondingly, the standard deviation of the lidar biomass model residual (i.e., $\varepsilon_{plot,j}$) is:

$$\sigma_{\varepsilon,plot,j} = k f_{plot}(\phi, z_j) \quad (16)$$

Eqs. 13-16 can be solved similarly using the GLS method introduced in Appendix 2. However, a key difference exists between developing tree-level allometric models and developing plot-level remote sensing based biomass model. For the tree-level allometric model, the AGB of a sampled tree is destructively measured in the field. In contrast, for the lidar-based biomass model, the AGB of a field plot is not an actual measurement but an estimate. The estimate $AGB_{plot,j}$ is different from the true plot-level AGB density $AGB_{plot,j}$. If the uncertainty of $AGB_{plot,j}$ is not considered in model fitting, the standard deviation of the model residuals (i.e., $\sigma_{\varepsilon,plot,j}$) will be underestimated. To address this issue, we modified the GLS method as shown in Appendix 5. The estimated parameters are denoted as ϕ and k.

Once the lidar-based biomass model has been calibrated, Eq. 14 can be used to predict AGB density AGB_{pix} for pixels that have equivalent areas to the field plots. As a reminder, similar to the tree-level allometric model, Eq. 14 is to predict the expected (or mean) biomass of pixels with the same lidar metrics z. This predicted mean has a standard error caused by the errors in the estimated model parameters ϕ. We denote this error as $\sigma_{f,pix}$.

When lidar metrics z has no errors, the predicted AGB at the pixel-level has an error of:

$$\sigma_{pix}^{\ 2} = \sigma_{\varepsilon,pix}^{\ 2} + \sigma_{f,pix}^{\ 2} \quad (17)$$

where $\sigma_{\varepsilon,pix}$ is calculated using Eq. 16 for the pixel; σ_{pix} is the AGB prediction error of the lidar-based biomass model for a pixel.

Similar to the derivations for tree-level AGB prediction errors (see Appendix 4), we found that, when the lidar metrics z of a pixel has errors, the error ${\sigma_{pix}}^2$ is:

$$ {\sigma_{pix}}^2 = {\sigma_{\varepsilon,pix}}^2 + {\sigma_{f,pix}}^2 + (k^2+1)\,\mathrm{var}(f_{plot}(\phi,z)) + \mathrm{var}(\sigma_{f,pix}) \quad (18) $$

where $\mathrm{var}(f_{plot}(\phi,z))$ again can be solved using Taylor Series expansion and $\mathrm{var}(\sigma_{f,pix})$ can be estimated using Monte-Carlo simulation. When lidar metrics *z* have no errors, both $\mathrm{var}(f_{plot}(\phi,z))$ and $\mathrm{var}(\sigma_{f,pix})$ are equal to zeros, which implies that the above two equations are equivalent in such a situation. The sum of the last two terms of the right-hand-side of Eq. 18, $(k^2+1)\,\mathrm{var}(f_{plot}(\phi,z)) + \mathrm{var}(\sigma_{f,pix})$, are caused by errors in lidar metrics *z* and thus are denoted as ${\sigma_{z,pix}}^2$.

3.3.1. Selection of lidar metrics for model fitting

Considerable correlations exist among the individual lidar metrics. We used a two-step procedure to select the most relevant ones for biomass prediction. First, we did a log-transformation of both the response and predictor variables, which turned the multiplicative power model to a linear model. Then forward stepwise regression was used to select the significant lidar metrics. The inference of the model parameters at the log-scale are not exactly the same as the one in the original scale. To avoid the exclusion of useful lidar metrics, we set the p-values for variables entering and removing from the model to be 0.1 and 0.2, respectively,

295 higher than the common values of 0.05 and 0.1. Second, we used the lidar metrics selected by
296 stepwise regression to develop the multiplicative power model using nonlinear model fitting
297 technique. The p-values of the individual variables in the multiplicative power model were
298 checked and, any insignificant variables (p-value > 0.05) were removed and the models are
299 refitted. The second step can be iterated until all variables are significant (Vaglio Laurin et al.,
300 2014). After the lidar metrics were selected, they were used in the GLS method for parameter
301 estimation and error characterization (see Eqs. 13-18).

302 **3.4. Evaluating the impacts of field measurement errors and lidar metric errors**

303 The uncertainty analysis of the AGB maps derived from field measurements and remote sensing
304 data requires information about the errors of tree attributes x and lidar metrics z. We do not
305 have direct measurements of these errors, so we rely on the previous studies to set up the error
306 estimates for our analysis. Throughout the text, we refer the standard deviation of an estimate or
307 measurements as *absolute error* while the standard deviation divided by the estimate or
308 measurement mean as *relative error*. For wood density, we followed Chave et al. (2004) and set
309 the relative errors to 10%. For tree height, Chave et al. (2004) used a relative error of 10%,
310 which is higher than 7.8% of temperate forest used by Phillips et al. (2000). However, Hunter et
311 al. (2013) found the precision of tropical tree height measurement could range from 3% to 20%
312 of the tree height. Here, we adopted the upper limit, 20%, in our analysis. Chave et al. (2004)
313 used a sum of two normal distributions to characterize the DBH error. We simplified this to set
314 the relative error of DBH to be 5%, which is higher than 2-3% used in temperate forests (Phillips
315 et al., 2000; Gonzalez et al., 2012; Ahmed et al., 2013) and considers the difficulty of measuring

316 DBH in tropical forests. We will introduce the errors of lidar metrics *z* in the next section after
317 we have determined the remote sensing biomass model.

318 ## 4. Results and Discussion

319 ### 4.1. Errors of tree-level AGB prediction

320 The GLS estimates of the allometric model parameters are: $\beta_1 = 0.0704$, $\beta_2 = 0.9701$, and
321 $\theta = 0.3777$. Therefore, the pan-tropical allometric model for predicting tree AGB is:

322 $$AGB_{tree} = 0.0704(\rho D^2 H)^{0.9701} \quad (19)$$

323 where ρ is wood specific gravity (unitless), *D* is DBH in cm, *H* is tree height in meter, and *AGB*
324 is in kg.

325 The equation for $\sigma_{\varepsilon,tree}$ is:

326 $$\sigma_{\varepsilon,tree} = 0.3777 AGB_{tree} \quad (20)$$

327 This indicates that the relative AGB error caused by $\sigma_{\varepsilon,tree}$ is 37.77%. With a database of fewer
328 trees, Chave et al. (2004) reported a smaller relative $\sigma_{\varepsilon,tree}$ error of 31.3%. The larger $\sigma_{\varepsilon,tree}$ here
329 is mainly related to the difference between the Chave14 database and the one used in Chave et al.
330 (2004).

Besides AGB_{tree} and $\sigma_{\varepsilon,tree}$, we also calculated the model parameter related AGB error $\sigma_{f,tree}$ and tree attributes *x* related AGB error $\sigma_{x,tree}$ for all trees (n=1191) of the 36 field plots in Ghana. The equation to calculate $\sigma_{f,tree}$ is presented as Eq. A16 in Appendix 6. We found that, when the model parameter related $\sigma_{f,tree}$ is considered in addition to $\sigma_{\varepsilon,tree}$, the relative AGB error increases slightly from 37.770% to 37.778%. If we consider all error sources including those related to tree attributes *x*, the relative error of tree AGB prediction increases from 37.8% (rounded from 37.778%) to 50.0%. This is slightly higher than the 47% value reported in Chave et al. (2004), mostly caused by an increase in our $\sigma_{\varepsilon,tree}$.

The relative AGB prediction error caused by $\sigma_{f,tree}$ alone is very small ($\sigma_{f,tree} / AGB_{tree}$ = 0.74%), mainly because the sample size is large (*n*=4004) for the Chave14 tree database. As shown in Eq. A6 in Appendix 3, the variance-covariance of model parameters are inversely related to the sample size. We did simulations to reduce the sample size for model development from 4004 to 400 and 40. The results showed that the relative AGB prediction error related to $\sigma_{f,tree}$ increases from 0.74% to 2.52% and 5.69%, respectively. Thus, using a large tree database such as Chave14 for calibrating the allometric model is valuable to reduce AGB prediction error related to model parameters.

For AGB prediction at the tree-level, the relative contributions of the individual error sources are: $\sigma_{\varepsilon,tree}^2 / \sigma_{tree}^2$ = 57.586% for allometric model residuals, $\sigma_{f,tree}^2 / \sigma_{tree}^2$ = 0.024% for model parameters, and $\sigma_{x,tree}^2 / \sigma_{tree}^2$ = 42.39% for tree attributes *x*. Therefore, for our field plots in

Ghana, the AGB prediction error were mainly related to model residuals ($\sigma_{\varepsilon,tree}$) and tree attributes x ($\sigma_{x,tree}$).

4.2. Errors of plot-level AGB estimates derived from field data

The previous section shows that the tree-level AGB error is 50.0% when all error sources are considered. At the plot-level, the corresponding AGB density AGB_{plot} error reduces to 19.5%. The cause of the smaller relative AGB error at the plot level is mathematical: the plot-level AGB standard deviation $\sigma_{AGB_{plot,j}}$ involves the square root of sum (Eq. 12) while AGB mean $AGB_{plot,j}$ directly uses the sum in the numerator (Eq. 11). Hence, the relative error, $\sigma_{AGB_{plot,j}}$ divided by $AGB_{plot,j}$, usually has a negative relationship with the number of trees within a plot (Chave et al., 2004).

4.3. Errors of pixel-level AGB estimates from lidar-based biomass model

Based on the feature selection method introduced in 3.3.1, the lidar-based biomass model developed from the 36 field plots in Ghana is:

$$AGB_{pix} = 14.55 h_{mean}^{1.27} h_{20to25}^{0.38} \quad (21)$$

where h_{mean} is the mean height of laser points; h_{20to25} is the proportion of laser points within 20 to 25 m. The lidar metrics were calculated using all returns including those on the ground. This model was developed with the assumption of spatial independence among residuals. The experimental variogram calculated from the model residuals did not show obvious spatial

368 autocorrelation, which implies that the above model can be used to predict pixel-level AGB and
369 its errors without bias (Barrett et al., 2001).

370 The estimated k = 0.479. Therefore, the equation for $\sigma_{\varepsilon,pix}$ is:

371 $$\sigma_{\varepsilon,pix} = 0.479 AGB_{pix} \quad (22)$$

372 This means that, at the plot- or pixel-level, the relative AGB prediction error due to lidar-
373 biomass model residuals is 47.9%.

374 Using the 40 m maps of lidar metrics h_{mean} and h_{20to25}, we produced maps of AGB_{pix} and $\sigma_{\varepsilon,pix}$
375 of the same spatial resolution. Besides, we calculated the pixel-level model parameter related
376 AGB error $\sigma_{f,pix}$ and lidar metrics related AGB error $\sigma_{z,pix}$ (Fig. 3). The equations for these two
377 errors are not given, but one can derive them following the examples in Appendix 3 and 4. We
378 calculated $\sigma_{z,pix}$ by assuming the relative errors of both h_{mean} and h_{20to25} to be 10%. The errors in
379 lidar metrics are mainly caused by extracting ground elevation from laser points. Asner et al.
380 (2012b) used an error estimate of 5% for lidar height metrics. We set the error estimate
381 arbitrarily higher, simply for the purpose of testing how much AGB prediction errors can be
382 caused by lidar metrics errors in a worse scenario.

383 We found that the relative AGB prediction error caused by $\sigma_{f,pix}$ and $\sigma_{z,pix}$ are 16.2% and
384 16.0%, respectively ($\sigma_{f,pix} / AGB_{pix}$ =16.2%, $\sigma_{z,pix} / AGB_{pix}$ =16.0%). Recall that if only $\sigma_{\varepsilon,pix}$ is
385 considered, the relative AGB prediction error is 47.9%. When both $\sigma_{\varepsilon,pix}$ and $\sigma_{f,pix}$ were

386 considered, the error increases to 51.7%. If we further considered lidar metric related AGB errors
387 $\sigma_{z,pix}$, the total relative AGB prediction error increased to 54.1%.

388 The contributions of individual error sources at the pixel level are: $\sigma_{\varepsilon,pix}^2 / \sigma_{pix}^2 = 80.541\%$ for
389 lidar-biomass model residuals, $\sigma_{f,pix}^2 / \sigma_{pix}^2 = 10.644\%$ for model parameters, and
390 $\sigma_{z,pix}^2 / \sigma_{pix}^2 = 8.814\%$ for lidar metrics *z*. So, when the lidar-biomass model was applied to the
391 study area, the dominant error source is the AGB error related to model residuals (i.e., $\sigma_{\varepsilon,pix}$).
392 Note that the error related to model parameters, $\sigma_{f,pix}$, is no more negligible as in the tree
393 allometric model. This is because the sample size for developing the lidar-biomass model (n=36)
394 is much smaller than the sample size for developing the allometric model (n=4004).

395 **4.4. Analysis of the remotely sensed AGB prediction errors**

396 When all the error sources are considered, the previous section indicated that the relative AGB
397 prediction error at the pixel-level, σ_{pix} / AGB_{pix}, is 54.1%, which seems high. Next, we discuss
398 different error sources and see whether the pixel-level relative AGB prediction error can be
399 reduced and, if yes, how.

400 First, the large pixel-level AGB errors are partially caused by the fact that we considered the
401 errors in measuring or estimating tree attributes *x* and the lidar metrics *z* of field plots. If we
402 assume that tree attributes were measured without errors, the pixel-level relative AGB prediction
403 error reduces from 54.1% to 51.9%. If we further assume the lidar metrics are error-free as well,
404 the AGB error reduces to 49.4%. Overall, the errors in tree measurements and lidar metrics

405 contributed only a small portion of the total errors, especially considering that we intentionally
406 set the errors in tree measurements and lidar metrics larger than those from most previous studies.
407 Therefore, it is unlikely that improving the tree measurements and lidar data processing skills
408 upon the state-of-the-art will dramatically reduce the pixel-level AGB relative error.

409 Second, the allometric model could be a potential problem. The allometric model used in this
410 study was developed from the pan-tropical tree database, but it was applied locally to our study
411 site in Ghana. Compared to a local allometric model that represents the true relationship between
412 AGB and tree attributes *x* in our study site, the pan-tropical model could be biased (which means
413 an allometric model different from Eq. 19) or have different precision (i.e., σ_{tree}) for tree AGB
414 prediction.

415 Although we do not have access to a "true" local allometric model, we can indirectly assess the
416 impacts of allometric models on the AGB prediction errors by changing the pan-tropical
417 allometric model itself. We randomly drew a sample of 100 trees (~2.5% of all 4004 trees) from
418 the Chave14 tree database, developed an allometric model from the sampled trees with the same
419 model form as Eq. 7, used the allometric model to predict biomass for our field plots in Ghana,
420 developed a lidar-based biomass model, and finally mapped the pixel-level AGB and errors. We
421 repeated this randomized process for 100 times. In the situation of assuming that both tree
422 attributes and lidar metrics for the field plots in Ghana have no errors, the pixel-level relative
423 AGB error is 49.5% on average, which is close to 49.4%, the results corresponding to the case of
424 developing an allometric model using the whole pan-tropical database. The variation of AGB
425 errors among these simulations is small (standard deviation: 1.3%; minimum: 46.3%; maximum:

426 52.3%). These results suggest that replacing the pan-tropical allometric model with a local
427 allometric model is unlikely to substantially change the pixel-level relative AGB prediction error.

428 We have to emphasize that the simulations conducted above were based on fitting allometric
429 models with the same model form (i.e., Eq. 7). If different model forms were used (e.g., Brown
430 1997), the results could be different. Eq. 7 incorporates three major tree attributes for biomass
431 prediction: wood density, DBH, and tree height. The model is biologically meaningful because
432 the latter two characterize the wood volume, which can be converted to wood biomass when
433 combined with wood density. Chave et al. (2014) demonstrated that the omission of tree height
434 from the AGB predictors led to worse model fitting. It is possible that adding other tree attributes
435 (e.g., crown depth and crown size) as predictors might lead to better allometric model fitting and
436 smaller σ_{tree}. However, such tree attributes are not readily available over large spatial scales. Eq.
437 7 is among the best, if not the best, in terms of performance and feasibility, so we do not expect
438 that much room exists to reduce the pixel-level AGB prediction errors by changing the model
439 form of Eq. 7.

440 Another error source is the lidar-biomass model parameter related error $\sigma_{f,pix}$. As indicated in
441 Eq. A6 of Appendix 3, $\sigma_{f,pix}$ is inversely related to the number of field plots. Recall that the
442 pixel-level AGB relative error is 49.4% if no errors exist in tree measurements and lidar metrics.
443 We created bootstrap samples from the 36 field plots and found that the error reduces slightly
444 from 49.4% to 47.3%, 46.7%, and 46.6% when bootstrap sample sizes of field plots are 100, 200,
445 and 500, respectively. Thus, collecting a large number of field plots does not have a large impact
446 on $\sigma_{f,pix}$ either.

447 The last remaining major error is $\sigma_{\varepsilon,pix}$, which contributes about 80% of the total pixel-level
448 AGB prediction error ($\sigma_{\varepsilon,pix}^2 / \sigma_{pix}^2 = 80.541\%$). The corresponding relative AGB prediction
449 error, $\sigma_{\varepsilon,pix} / AGB_{pix}$, is equal to 47.9%. $\sigma_{\varepsilon,pix} / AGB_{pix}$ is also known as the relative RMSE when
450 RMSE is calculated using calibration plots instead of separate validation or test plots (Zolkos et
451 al., 2013). Asner et al. (2014) published a database of 754 field plots from Colombia, Hawaii,
452 Madagascar, and Panama, and Peru. We fitted a simple power model to predict AGB for the
453 plots from Madagascar because of its geographic proximity to our study site. The calculated
454 relative RMSE is 43.6%, which is similar to our estimate of 47.9%. In a separate study we did in
455 Sierra Leone (Vaglio Laurin et al., 2014), the relative RMSE is 42.2%. These datasets in Africa
456 (Sierra Leone, Ghana, and Madagascar) suggest that the lidar-biomass models had relative
457 RMSEs of 40-50%, which are among the highest compared to the relative RMSEs reported for
458 other tropical forests (Zolkos et al., 2013). However, the number of studies of using lidar for
459 biomass mapping in Africa is still limited. It is unclear whether in general it is more difficult to
460 use lidar for mapping AGB of African tropical forests compared to other tropical forests and, if
461 yes, what factors have caused the large model residuals (i.e., large $\sigma_{\varepsilon,pix}$ values).

462 One way to reduce $\sigma_{\varepsilon,pix}$ is to use larger field plots. It is well known that larger plot size can
463 alleviate the "edge effects" of field plots and improve the model fitting (Asner et al., 2010, Frazer
464 et al., 2011; Mascaro et al., 2011, Zolkos et al., 2013, Asner and Mascaro, 2014). Asner et al.
465 (2010) and Mascaro et al. (2011) proposed that the lidar-biomass modeling error, $\sigma_{\varepsilon,pix}$, is
466 inversely related to the square root of plot area. Using such a relationship, the relative AGB
467 prediction error associated with $\sigma_{\varepsilon,pix}$ can reduce from 47.9% to 27.1% and 19.2% if we increase

our plot size from 0.16 ha to 0.5 ha and 1 ha, respectively. Of course, these results were derived based on their equation, not based on actual data because we did not measure the individual tree locations using GPS and cannot change our field plot size for analysis.

Another way that can reduce the overall AGB prediction error is to aggregate the 40-m AGB pixels to coarser resolutions:

$$AGB_{pix,c,j} = \frac{\sum_{i=1}^{n_k} AGB_{pix,i}}{n_k} \quad (23)$$

$$\mathrm{var}(AGB_{pix,c,j}) = \frac{\sum_{i=1}^{n_k} \sigma^2_{pix,\mathrm{i}}}{{n_k}^2} \quad (24)$$

where $AGB_{pix,c,j}$ is the AGB for pixel *j* at a coarser spatial resolution, $AGB_{pix,i}$ and $\sigma^2_{pix,i}$ are the predicted AGB density and error for pixel *i* at the original 40 m cell size, n_k refers to the number of 40 m pixels included in a coarser pixel. We aggregated AGB from 40 m to 80 m, 120 m, and 200 m pixel sizes, for which n_k equals to 4, 9, and 25, respectively. We found that the average pixel-level AGB errors (σ_{pix} / AGB_{pix}) reduce from 54.1% (assuming both tree attributes and lidar metrics have errors) to 38.3%, 31.3%, and 23.9% when the resolution is 80 m (0.64 ha), 120 m (1.44 ha), and 200 m (4 ha), respectively.

Using either of the above approach, the final relative AGB prediction error (σ_{pix} / AGB_{pix}) is larger than 20% at 1 ha spatial scale. Several recent studies (Zolkos et al., 2013; Asner and Mascaro, 2014) stated that the uncertainty of lidar-based AGB mapping can reach 10% at the 1

485 ha spatial resolution. The divergent results between ours and those are not only related to our
486 focus on African tropical forests but also because we are using a different approach for
487 uncertainty analysis.

488 **4.5. Difference between our error analysis method and conventional ones**

489 A common practice to evaluate remotely sensed AGB prediction error is to set aside validation
490 field plots or use cross-validation for accuracy assessment. Such methods are legitimate if the
491 AGB of the validation plots are directly measured. Unfortunately, directly measuring tree AGB
492 for forest plots is extremely labor-intensive and thus rarely carried out. It has been increasingly
493 recognized that this is one of the main problems of conventional methods for characterizing
494 AGB uncertainty (Clark and Kellner, 2012; Asner and Mascaro, 2014). Error propagation
495 method (Chave et al., 2004; Gonzalez et al., 2010), as used in this study, represents a
496 fundamentally different approach to characterize the errors of the remotely sensed AGB. If the
497 conventional error assessment method using validation plots are considered to be retrospective,
498 the error propagation method we are using is forward-looking. The key of the conventional error
499 assessment method is the true AGB measurements of the validation field plots while the key of
500 the error propagation method is the true AGB measurements of the trees used for developing the
501 allometric method. So, it is impossible for us to propose our method if destructive tree AGB
502 databases such as Chave14 were not released publicly.

503 **4.6. Impacts of plot-level AGB uncertainty on lidar-based AGB modeling and prediction**

504 In section 4.2, we showed that the plot-level AGB uncertainty is 19.5% when the plot-level AGB
505 is upscaled from tree AGB values (for the case that all the tree-level errors related to allometric

model residuals, allometric model parameters, and tree measurements are considered). Unlike our approach (see Appendix 5), most previous studies did not consider the plot-level AGB uncertainty in remote sensing biomass modeling. If this type of error is not considered in our study, the relative AGB error related to the lidar-biomass model residuals ($\sigma_{\varepsilon,pix}/AGB_{pix}$) will be 41.94% instead of 47.9%; the total pixel-level relative AGB prediction error (σ_{pix}/AGB_{pix}) will be 48.7% instead of 54.1%. So, our study shows that the AGB estimation and prediction errors can be underestimated by a moderate amount if the plot-level AGB uncertainty is ignored in the lidar-biomass modeling.

5. Conclusions

The major contribution of this study is to introduce an analytical framework of characterizing AGB prediction errors by considering errors in the whole workflow of AGB mapping, including allometric model development, tree measurements, tree-level AGB prediction, plot-level AGB estimation, plot-level lidar-biomass model development, plot-level lidar metrics generation, and pixel-level AGB prediction. Our implementation of model development, for both allometric and lidar-biomass models, and error characterization using the Generalized Least Squares method is novel in that it enables us to estimate three important AGB prediction errors simultaneously: 1) errors related to model residuals, 2) errors related to predictors (tree measurements or lidar metrics), and 3) errors related to model parameters. Our method is not only useful for the remote sensing mapping community but also for researchers who use solely forest plots for AGB estimation (e.g., using large-area forest inventory plots for estimating mean AGB and its errors over an area). The application of our error analysis framework to tropical forests sites in Ghana

indicates that the AGB prediction error is larger than 20% at 1 ha spatial resolution, much larger than the 10% goal proposed in previous studies. This calls for more studies in AGB estimation and mapping over tropical forests in Africa. This study also highlights the importance of considering the plot-level AGB estimate uncertainty when field plots are used to calibrate remote sensing based biomass models.

Acknowledgements

We acknowledge the ERC grant Africa GHG #247349. The Matlab codes for GLS model fitting and AGB error analysis are available from the lead-author upon request.

Appendix:

1. Basic formula

For a function *f* with two predictors *x* and *y*, using first-order Taylor series expansion, we can calculate the variance of *f* caused by variance in *x* and *y*.

$$\mathrm{var}(f(x,y)) \approx \left(\frac{\partial f}{\partial x}\right)^2 \mathrm{var}(x) + \left(\frac{\partial f}{\partial y}\right)^2 \mathrm{var}(y) + 2\frac{\partial f}{\partial x}\frac{\partial f}{\partial y}\mathrm{cov}(x,y) \quad \text{(A1)}$$

This equation has similar form if there are more than 2 predictors.

The variance of the product of two variables *x* and *y* is:

$$\mathrm{var}(xy) = (E(x))^2\,\mathrm{var}(y) + (E(y))^2\,\mathrm{var}(x) + \mathrm{var}(x)\,\mathrm{var}(y) \quad \text{(A2)}$$

2. Generalized Least Squares (GLS) method for model fitting

Eqs. 3-7 can be solved iteratively as follows:

1) Fit a nonlinear model for Eqs. 3 and 7 to obtain an initial estimate of the allometry model parameter, $\beta^{(0)}$, by assuming the residuals $\varepsilon_{tree,i}$ are zero.

2) Obtain an initial maximum likelihood estimate of the residual model parameter $\theta^{(0)}$ by minimizing the objective function L_{tree}:

$$L_{tree} = \sum_{i=1}^{n_{tree}} \left(\frac{(AGB_{tree,i} - f_{tree}(\beta^{(0)}, x_i))^2}{2\times(\theta f_{tree}(\beta^{(0)}, x_i))^2} + \log(\theta f_{tree}(\beta^{(0)}, x_i))\right) \quad \text{(A3)}$$

where n_{tree} is the number of trees that are used to develop the allometry model (for the Chave et al., 2014 database, n_{tree} = 4004). After $\theta^{(0)}$ is obtained, it will be used along with $\beta^{(0)}$ in Eq. 6 to estimate the expected standard deviation of residuals $\sigma^{(0)}_{\varepsilon,tree,i}$ for every tree.

3) Define a weight function w_i that is inversely related to $\sigma^{(0)}_{\varepsilon,tree,i}$ as follows:

$$w_i = \left(\frac{1}{\sigma^{(0)}_{\varepsilon,tree,i}}\right)^2 = \left(\frac{1}{\theta^{(0)} f(\beta^{(0)}, x_i)}\right)^2 \quad \text{(A4)}$$

and solve Eqs. 3 and 7 using the nonlinear weighted least squares method (Brown et al., 1989; Struz 2010) to get updated estimates of the parameter, $\beta^{(1)}$. The weighted least squares (WLS) method essentially assigns lower weights to the trees of larger AGB variance.

4) Repeat steps 2) and 3) until convergence. The final estimated model parameters are denoted as β and θ.

3. Standard error of predicted AGB mean

The standard error $\sigma_{f,tree}$ of the predicted AGB mean for trees of *x* measurements can be calculated from:

$$\sigma_{f,tree}{}^2 = \sum_{j=1}^{p}\sum_{k=1}^{p}\frac{\partial f}{\partial \beta_j}\frac{\partial f}{\partial \beta_k}\mathrm{cov}(\beta_j,\beta_k) = \left(\frac{\partial f}{\partial \beta_1}\right)^2\mathrm{cov}(\beta_1,\beta_1) + 2\frac{\partial f}{\partial \beta_1}\frac{\partial f}{\partial \beta_2}\mathrm{cov}(\beta_1,\beta_2) + \left(\frac{\partial f}{\partial \beta_2}\right)^2\mathrm{cov}(\beta_2,\beta_2) \quad \text{(A5).}$$

$$= ((\rho_i D_i^2 H_i)^{\beta_2})^2\,\mathrm{cov}(\beta_1,\beta_1) + 2((\rho_i D_i^2 H_i)^{\beta_2} * \beta_1 (\rho_i D_i^2 H_i)^{\beta_2}\ln(\rho_i D_i^2 H_i))\,\mathrm{cov}(\beta_1,\beta_2) + (\beta_1(\rho_i D_i^2 H_i)^{\beta_2}\ln(\rho_i D_i^2 H_i))^2\,\widetilde{\mathrm{cov}(\beta_2,\beta_2)}$$

where the variance-covariance matrix $\mathrm{cov}(\beta_j, \beta_k)$ ($j,k \in [1,2]$) characterize the errors of the parameters β and it is estimated from the sample data used for developing the allometry model:

$$\mathrm{cov}(\beta_j, \beta_k) = \frac{S}{n-p} C_{jk}^{-1}$$
$$S = \sum_{i=1}^{n} w_i (AGB_{tree,i} - f_{tree}(\beta, x_i))^2 \quad \text{(A6)}$$

where *n* is the number of trees that are used to develop the allometry model, $f(\beta, x_i)$ is the allometric model with parameter β, w_i is defined in Eq. A4, *C* is the $p \times p$ (*p* = the number of allometry model parameters β, which is 2 in our case) square matrix with its element calculated as below,

$$C_{jk} = \sum_{i=1}^{n} w_i \frac{\partial f}{\partial \beta_j} \frac{\partial f}{\partial \beta_k} \quad \text{(A7)}$$

4. Error of predicting tree AGB when a tree has measurement errors in *x*

When both errors σ_f and σ_ε considered, the AGB of a tree is:

$$AGB_{tree} = f_{tree}(\beta, x) + \sigma_{\varepsilon,tree}\xi_\varepsilon + \sigma_{f,tree}\xi_f \quad \text{(A8)}$$

where both ξ_ε and ξ_f are independent standard random variables, so $E(\xi_\varepsilon) = 0, \mathrm{var}(\xi_\varepsilon) = 1$ and $E(\xi_f) = 0, \mathrm{var}(\xi_f) = 1$.

The three terms at the right hand side of the above equation are independent, so we have (according to Eq. A1):

$$\mathrm{var}(AGB_{tree}) = \mathrm{var}(f_{tree}(\beta, x)) + \mathrm{var}(\sigma_{\varepsilon,tree}\xi_\varepsilon) + \mathrm{var}(\sigma_{f,tree}\xi_f) \quad \text{(A9)}$$

Note that, according to Eqs. A1 and A2, the second term $\mathrm{var}(\sigma_{\varepsilon,tree}\xi_\varepsilon)$ of the above equation is:

$$\mathrm{var}(\sigma_{\varepsilon,tree}\xi_\varepsilon) = \mathrm{var}(\theta f_{tree}(\beta, x)\xi_\varepsilon) = \theta^2 \mathrm{var}(f_{tree}(\beta, x)\xi_\varepsilon)$$
$$= \theta^2 (\mathrm{var}(f_{tree}(\beta, x))\mathrm{var}(\xi_\varepsilon) + \mathrm{var}(f_{tree}(\beta, x))(E(\xi_\varepsilon))^2 + \mathrm{var}(\xi_\varepsilon)(E(f_{tree}(\beta, x)))^2) \quad \text{(A10)}$$
$$= \theta^2 (\mathrm{var}(f_{tree}(\beta, x)) + (f_{tree}(\beta, x))^2)$$

and the third term $\mathrm{var}(\sigma_{f,tree}\xi_f)$ is:

$$\mathrm{var}(\sigma_{f,tree}\xi_f) = \mathrm{var}(\sigma_{f,tree})\mathrm{var}(\xi_f) + \mathrm{var}(\sigma_{f,tree})(E(\xi_f))^2 + \mathrm{var}(\xi_f)(E(\sigma_{f,tree}))^2$$
$$= \mathrm{var}(\sigma_{f,tree}) + \sigma_{f,tree}^2 \quad \text{(A11)}$$

where $\mathrm{var}(\sigma_{f,tree})$ is the variance of $\sigma_{f,tree}$ caused by *x* errors.

In sum,

$$\begin{aligned}
&\mathrm{var}(AGB_{tree}) = \mathrm{var}(f_{tree}(\beta,x)) + \theta^2(\mathrm{var}(f_{tree}(\beta,x)) + (f_{tree}(\beta,x))^2) + \mathrm{var}(\sigma_{f,tree}) + {\sigma_{f,tree}}^2 \\
&= (\theta f_{tree}(\beta,x))^2 + (\theta^2+1)\,\mathrm{var}(f_{tree}(\beta,x)) + \mathrm{var}(\sigma_{f,tree}) + {\sigma_{f,tree}}^2 \\
&= {\sigma_{\varepsilon,tree}}^2 + {\sigma_{f,tree}}^2 + (\theta^2+1)\,\mathrm{var}(f_{tree}(\beta,x)) + \mathrm{var}(\sigma_{f,tree})
\end{aligned} \quad \text{(A12)}$$

$\mathrm{var}(f_{tree}(\beta,x))$ can be calculated as follows:

Recall $f_{tree}(\beta,x) = \beta_1(\rho D^2 H)^{\beta_2}$, and let $k = \rho D^2 H$.

Then, we have:

$$\begin{aligned}
&\mathrm{var}(f_{tree}(\beta,x)) = (\frac{\partial f}{\partial k})^2\,\mathrm{var}(k) \\
&= (\frac{\partial f}{\partial k})^2[(\frac{\partial k}{\partial \rho}\sigma_\rho)^2 + (\frac{\partial k}{\partial D}\sigma_D)^2 + (\frac{\partial k}{\partial H}\sigma_H)^2 + 2\frac{\partial k}{\partial \rho}\frac{\partial k}{\partial D}\mathrm{cov}(\rho,D) + 2\frac{\partial k}{\partial \rho}\frac{\partial k}{\partial H}\mathrm{cov}(\rho,H) + 2\frac{\partial k}{\partial D}\frac{\partial k}{\partial H}\mathrm{cov}(D,H)] \\
&= (\beta_1\beta_2(\rho D^2 H)^{\beta_2-1})^2[(D^2 H\sigma_\rho)^2 + (2\rho DH\sigma_D)^2 + (\rho D^2\sigma_H)^2 + \\
&+2(D^2 H)(2\rho DH)(cor(\rho,D)\sigma_\rho\sigma_D) + 2(D^2 H)(\rho D^2)(cor(\rho,H)\sigma_\rho\sigma_H) + 2(2\rho DH)(\rho D^2)(cor(D,H)\sigma_D\sigma_H)]
\end{aligned} \quad \text{(A13)}$$

where $cor()$ is the correlation coefficient between two tree attributes, which was estimated from the Chave14 dataset. σ_ρ, σ_D, and σ_H are the standard deviation of wood density, DBH, and tree height, respectively.

Since the equation of $\sigma_{f,tree}$ is complicated (see Eq. A5), we chose Monte Carlo simulation to estimate this uncertainty.

5. Revision of the GLS method for predicting plot-level biomass using lidar

To consider the uncertainty in $AGB_{plot,j}$ in the GLS method, we modified the objective function presented in Appendix A3 as follows:

$$L_{plot} = \sum_{j=1}^{n_{plot}} (\frac{(AGB_{plot,j} - f_{plot}(\phi^{(0)}, z_j))^2 + \mathrm{var}(AGB_{plot,j})}{2\times(kf_{plot}(\phi^{(0)}, z_j))^2} + \log(kf_{plot}(\phi^{(0)}, z_j))) \quad \text{(A14)}$$

where n_{plot} is the number of field plots used for developing lidar-based biomass model (n_{plot} = 36 in our study).

Compare Eqs. A3 and A14, and one will see that the difference is the addition of the term $\mathrm{var}(AGB_{plot,j})$, which is the uncertainty of $AGB_{plot,j}$.

6. Equation for estimating $\sigma_{f,tree}$ caused by allometric model parameter errors

The variance-covariance matrix of the estimated parameters of the allometric model are:

$$\begin{bmatrix} \mathrm{var}(\beta_1) & \mathrm{cov}(\beta_1,\beta_2) \\ \mathrm{cov}(\beta_1,\beta_2) & \mathrm{var}(\beta_2) \end{bmatrix} = \begin{bmatrix} 2.4656\text{e-}06 & -4.217\text{e-}06 \\ -4.217\text{e-}06 & 7.7686\text{e-}06 \end{bmatrix} \quad \text{(A15)}$$

623 The estimated parameters are: $\beta_1 = 0.0704$, $\beta_2 = 0.9701$. According to Eq. A5, we know that:

$$\begin{aligned}\sigma_{f,tree}^{\ 2} = &((\rho D^2 H)^{0.9701})^2 * 2.4656\text{e-}06 \\ 624\quad &+2((\rho D^2 H)^{0.9701} * 0.0704(\rho D^2 H)^{0.9701} \ln(\rho D^2 H)) * (-4.217\text{e-}06) \\ &+(0.0704(\rho D^2 H)^{0.9701} \ln(\rho D^2 H))^2 * 7.7686\text{e-}06\end{aligned} \quad \text{(A16)}$$

References:

1. Achard, F., Beuchle, R., Mayaux, P., Stibig, H. J., Bodart, C., Brink, A., Carboni, S., Desclée, B., Donnay, F., Eva, H., Lupi, A., Raši, R., Seliger, R., & Simonetti, D. (2014). Determination of tropical deforestation rates and related carbon losses from 1990 to 2010. *Global Change Biology*, 20(8), 2540-2554.

Ahmed, R., Siqueira, P., Hensley, S., & Bergen, K. (2013). Uncertainty of forest biomass estimates in north temperate forests due to allometry: Implications for remote sensing. *Remote Sensing*, 5(6), 3007-3036.

Aguiar, A. P. D., Ometto, J. P., Nobre, C., Lapola, D. M., Almeida, C., Vieira, I. C., Soares, J., Alvala, R., Saatchi, S., Valeriano, D., & Castilla-Rubio, J. C. (2012). Modeling the spatial and temporal heterogeneity of deforestation-driven carbon emissions: the INPE-EM framework applied to the Brazilian Amazon. *Global Change Biology*, 18(11), 3346-3366.

Asner, G. P., Rudel, T. K., Aide, T. M., Defries, R., & Emerson, R. (2009). A contemporary assessment of change in humid tropical forests. *Conservation Biology*, 23(6), 1386-1395.

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*, 107(38), 16738-16742.

Asner, G. P., Clark, J. K., Mascaro, J., Vaudry, R., Chadwick, K. D., Vieilledent, G., Rasamoelina M., Balaji, A., Kennedy-Bowdoin, T., Maatoug, L., Colgan, M., & Knapp, D. E. (2012a). Human and environmental controls over aboveground carbon storage in Madagascar. *Carbon Balance and Management*, 7(2).

Asner, G. P., Clark, J. K., Mascaro, J., Galindo García, G. A., Chadwick, K. D., Navarrete Encinales, D. A., Paez-Acosta, G., Cabrera Montenegro, E., Kennedy-Bowdoin, T., Duque, A., Balaji, A., von Hildebrand, P., Maatoug, L., Philips Bernal, J., Knapp, D., García Dávila, M., Jacobson, J., & Ordóñez, M. F. (2012b). High-resolution mapping of forest carbon stocks in the Colombian Amazon. *Biogeosciences*, 9, 2683–2696.

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

Asner, G. P., Mascaro, J., Anderson, C., Knapp, D. E., Martin, R. E., Kennedy-Bowdoin, T., van Breugel, M., Davies, S., Hall, J., Muller-Landau, H., & Bermingham, E. (2013). High-fidelity national carbon mapping for resource management and REDD+. *Carbon Balance and Management*, 8(1).

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

Baccini, A., Laporte, N., Goetz, S. J., Sun, M., & Dong, H. (2008). A first map of tropical Africa's above-ground biomass derived from satellite imagery. *Environmental Research Letters*, 3(4), 045011.

Baccini, A., Goetz, S. J., Walker, W. S., Laporte, N. T., Sun, M., Sulla-Menashe, D., Hackler, J., Beck, P., Dubayah, R., Friedl, M., Samanta, S., & Houghton, R. A. (2012). Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. *Nature Climate Change*, 2(3), 182-185.

Barrett, D. J., Galbally, I. E., & Graetz, R. D. (2001). Quantifying uncertainty in estimates of C emissions from above-ground biomass due to historic land-use change to cropping in Australia. *Global Change Biology*, 7(8), 883-902.

Baskerville, G. L. (1972). Use of logarithmic regression in the estimation of plant biomass. *Canadian Journal of Forest Research*, 2(1), 49-53.

Bagley, J. E., Desai, A. R., Harding, K. J., Snyder, P. K., & Foley, J. A. (2014). Drought and Deforestation: Has Land Cover Change Influenced Recent Precipitation Extremes in the Amazon? *Journal of Climate*, 27(1), 345-361.

Basuki, T. M., Van Laake, P. E., Skidmore, A. K., & Hussin, Y. A. (2009). Allometric equations for estimating the above-ground biomass in tropical lowland *Dipterocarp* forests. *Forest Ecology and Management*, 257(8), 1684-1694.

Berenguer, E., Ferreira, J., Gardner, T. A., Aragão, L. E. O. C., De Camargo, P. B., Cerri, C. E., Durigan, M., De Oliveira Junior, R., Vieira, I., & Barlow, J. (2014). A large-scale field assessment of carbon stocks in human-modified tropical forests. *Global Change Biology*, DOI: 10.1111/gcb.12627.

Bonan, G.B. (2008) Forests and climate change: forcings, feedbacks, and the climate benefits of forests. *Science*, 320, 1444–1449.

Brown, S. (1997). Estimating biomass and biomass change of tropical forests: a primer (Vol. 134). Food & Agriculture Org..

Brown, S., Gillespie, A. J., & Lugo, A. E. (1989). Biomass estimation methods for tropical forests with applications to forest inventory data. *Forest Science*, 35(4), 881-902.

Chave, J., Condit, R., Aguilar, S., Hernandez, A., Lao, S., & Perez, R. (2004). Error propagation and scaling for tropical forest biomass estimates. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 359(1443), 409-420.

713 Chave, J., Andalo, C., Brown, S., Cairns, M. A., Chambers, J. Q., Eamus, D., Fölster, H.,
714 Fromard, F., Higuchi, N., & Yamakura, T. (2005). Tree allometry and improved estimation of
715 carbon stocks and balance in tropical forests. *Oecologia*, 145(1), 87-99.
716
717 Chave, J., Coomes, D., Jansen, S., Lewis, S. L., Swenson, N. G., & Zanne, A. E. (2009).
718 Towards a worldwide wood economics spectrum. *Ecology letters*, 12(4), 351-366.
719
720 1. Chave, J., Réjou-Méchain, M., Búrquez, A., Chidumayo, E., Colgan, M. S., Delitti, W. B.,
721 Duque, A., Eid, T., Fearnside, P., Goodman, R., Henry, M., Martínez-Yrízar, A., Mugasha, W.,
722 Muller-Landau, H., Mencuccini, M., Nelson, B., Ngomanda, A., Nogomanda, A., Nogueira, E.,
723 Ortiz-Malavassi, E., Pélissier, R., Ploton, P., Ryan, C., Saldarriaga, J., & Vieilledent, G. (2014).
724 Improved allometric models to estimate the aboveground biomass of tropical trees. *Global*
725 *Change Biology*, DOI:10.1111/gcb.12629.
726
727 Chen, Q. (2007). Airborne lidar data processing and information extraction. Photogrammetric
728 engineering and remote sensing, 73(2), 109-112.
729
730 Chen, Q., Vaglio Laurin, G., Battles, J. J., & Saah, D. (2012). Integration of airborne lidar and
731 vegetation types derived from aerial photography for mapping aboveground live
732 biomass. *Remote Sensing of Environment*, 121, 108-117.
733
734 Chen, Q. (2010). Retrieving vegetation height of forests and woodlands over mountainous areas
735 in the Pacific Coast region using satellite laser altimetry. *Remote Sensing of Environment*, 114(7),
736 1610-1627.
737
738 Chen, Q. (2013). Lidar remote sensing of vegetation biomass. *Remote Sensing of Natural*
739 *Resources*, CRC Press: Taylor & Francis, Group, 399-420.
740
741 Clark, D. B., & Kellner, J. R. (2012). Tropical forest biomass estimation and the fallacy of
742 misplaced concreteness. Journal of Vegetation Science, 23(6), 1191-1196.
743
744 Clark, M. L., Roberts, D. A., Ewel, J. J., & Clark, D. B. (2011). Estimation of tropical rain forest
745 aboveground biomass with small-footprint lidar and hyperspectral sensors. *Remote Sensing of*
746 *Environment*, 115(11), 2931-2942.
747
748 Clifford, D., Cressie, N., England, J. R., Roxburgh, S. H., & Paul, K. I. (2013). Correction
749 factors for unbiased, efficient estimation and prediction of biomass from log–log allometric
750 models. *Forest Ecology and Management*, 310, 375-381.
751
752 Cochrane MA (2003) Fire science for rainforests. *Nature*, 421, 913–919.
753
754 DeFries, R., Achard, F., Brown, S., Herold, M., Murdiyarso, D., & Schlamadinger, B. (2007).
755 Earth observations for estimating greenhouse gas emissions from deforestation in developing
756 countries. *Environmental Science & Policy*, 10(4), 385-394.

Djomo, A. N., Ibrahima, A., Saborowski, J., & Gravenhorst, G. (2010). Allometric equations for biomass estimations in Cameroon and pan moist tropical equations including biomass data from Africa. *Forest Ecology and Management*, 260(10), 1873-1885.

d'Oliveira, M. V., Reutebuch, S. E., McGaughey, R. J., & Andersen, H. E. (2012). Estimating forest biomass and identifying low-intensity logging areas using airborne scanning lidar in Antimary State Forest, Acre State, Western Brazilian Amazon. *Remote Sensing of Environment*, 124, 479-491.

Dubayah, R. O., Sheldon, S. L., Clark, D. B., Hofton, M. A., Blair, J. B., & 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, G: *Biogeosciences*, 115, GE00E09

Fearnside, P. M., & Laurance, W. F. (2004). Tropical deforestation and greenhouse-gas emissions. *Ecological Applications*, 14(4), 982-986.

Feldpausch TR, Lloyd J, Lewis SL et al. (2012) Tree height integrated into pantropical biomass forest estimates. *Biogeosciences*, 9, 3381–3403.

Finney, D. J. (1941). On the distribution of a variate whose logarithm is normally distributed. *Journal of the Royal Statistical Society*, 7(2), 155-161.

Fischer, R., Armstrong, A., Shugart, H. H., & Huth, A. (2014). Simulating the impacts of reduced rainfall on carbon stocks and net ecosystem exchange in a tropical forest. *Environmental Modelling & Software*, 52, 200-206.

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), 636-649.

Gertner, G., Cao, X., & Zhu, H. (1995). A quality assessment of a Weibull based growth projection system. *Forest Ecology and Management*, 71(3), 235-250.

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*, doi:10.1088/1748-9326/2/4/045023.

Goetz, S. J., Baccini, A., Laporte, N. T., Johns, T., Walker, W., Kellndorfer, J., Houghton, R., & Sun, M. (2009). Mapping and monitoring carbon stocks with satellite observations: a comparison of methods. *Carbon Balance and Management*,4(1).

Gonzalez, P., Asner, G. P., Battles, J. J., Lefsky, M. A., Waring, K. M., & Palace, M. (2010). Forest carbon densities and uncertainties from Lidar, QuickBird, and field measurements in California. *Remote Sensing of Environment*, 114(7), 1561-1575.

Grace, J., Mitchard, E., & Gloor, E. (2014). Perturbations in the carbon budget of the tropics. *Global Change Biology*, DOI: 10.1111/gcb.12600.

Hernández-Stefanoni, J. L., Dupuy, J. M., Johnson, K. D., Birdsey, R., Tun-Dzul, F., Peduzzi, A., Caamal-Sosa, J., Sánchez-Santos, G., & López-Merlín, D. (2014). Improving Species Diversity and Biomass Estimates of Tropical Dry Forests Using Airborne LiDAR. *Remote Sensing*, 6(6), 4741-4763.

Houghton, R. A., Lawrence, K. T., Hackler, J. L., & Brown, S. (2001). The spatial distribution of forest biomass in the Brazilian Amazon: a comparison of estimates. *Global Change Biology*, 7(7), 731-746.

Houghton, R. A., Hall, F., & Goetz, S. J. (2009). Importance of biomass in the global carbon cycle. *Journal of Geophysical Research: Biogeosciences*, 114, G00E03, doi:10.1029/2009JG000935.

Hunter, M. O., Keller, M., Vitoria, D., & Morton, D. C. (2013). Tree height and tropical forest biomass estimation. *Biogeosciences*, 10, 8385-8399.

Ioki, K., Tsuyuki, S., Hirata, Y., Phua, M. H., Wong, W. V. C., Ling, Z. Y., Saito, H., & Takao, G. (2014). Estimating above-ground biomass of tropical rainforest of different degradation levels in Northern Borneo using airborne LiDAR. *Forest Ecology and Management*, 328, 335-341.

Kronseder, K., Ballhorn, U., Böhm, V., & Siegert, F. (2012). Above ground biomass estimation across forest types at different degradation levels in Central Kalimantan using LiDAR data. *International Journal of Applied Earth Observation and Geoinformation*, 18, 37-48.

Lewis, S. L., Lopez-Gonzalez, G., Sonké, B., Affum-Baffoe, K., Baker, T. R., Ojo, L. O., Phillips, O.L., Reitsma, J.M., White, L., Comiskey, J.A., Djuikouo K, M-N., Ewango, C.E.N., Feldpausch, T.R., Hamilton, A.L., Gloor, M., Hart, T., Hladik, A., Lloyd, J., ... & Comiskey, J. A. (2009). Increasing carbon storage in intact African tropical forests. *Nature*, 457(7232), 1003-1006.

Lu, D., Chen, Q., Wang, G., Moran, E., Batistella, M., Zhang, M., Vaglio Laurin, G., & Saah, D. (2012). Aboveground forest biomass estimation with Landsat and LiDAR data and uncertainty analysis of the estimates. *International Journal of Forestry Research*, Article ID 436537.

Malhi, Y., Aragão, L. E., Galbraith, D., Huntingford, C., Fisher, R., Zelazowski, P., Sitch, S., McSweeney, C., & Meir, P. (2009). Exploring the likelihood and mechanism of a climate-change-induced dieback of the Amazon rainforest. *Proceedings of the National Academy of Sciences*, 106(49), 20610-20615.

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

Mitchard, E. T., Saatchi, S. S., Baccini, A., Asner, G. P., Goetz, S. J., Harris, N., & Brown, S. (2013). Uncertainty in the spatial distribution of tropical forest biomass: a comparison of pan-tropical maps. *Carbon Balance and Management*, 8(10), 1-13.

Ometto, J. P., Aguiar, A. P., Assis, T., Soler, L., Valle, P., Tejada, G., Lapola, D., & Meir, P. (2014). Amazon forest biomass density maps: tackling the uncertainty in carbon emission estimates. *Climatic Change*, 124(3), 545-560.

Parresol, B. R. (1999). Assessing tree and stand biomass: a review with examples and critical comparisons. *Forest Science*, 45(4), 573-593.

Phillips, D. L., Brown, S. L., Schroeder, P. E., & Birdsey, R. A. (2000). Toward error analysis of large-scale forest carbon budgets. *Global Ecology and Biogeography*, 9(4), 305-313.

Saatchi, S. S., Harris, N. L., Brown, S., Lefsky, M., Mitchard, E. T., Salas, W., Zutta, B., Buermann, W., Lewis, S., Hagen, S., Petrova, S., White, L., Silman, M., & Morel, A. (2011). Benchmark map of forest carbon stocks in tropical regions across three continents. *Proceedings of the National Academy of Sciences*,108(24), 9899-9904.

Snowdon, P. (1991). A ratio estimator for bias correction in logarithmic regressions. *Canadian Journal of Forest Research*, 21(5), 720-724.

Sprugel, D. G. (1983). Correcting for bias in log-transformed allometric equations. *Ecology*, 64(1), 209-210.

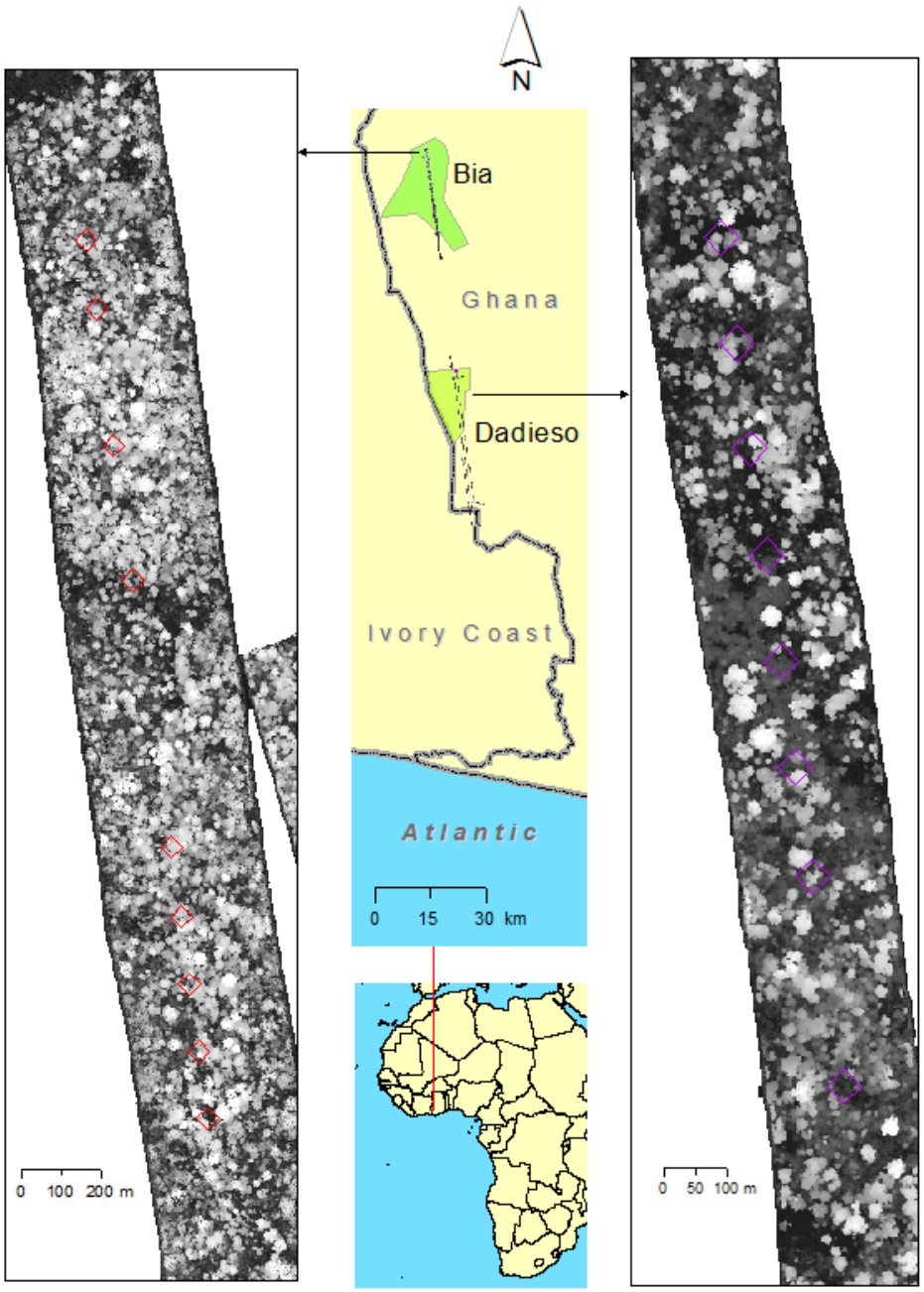
Vaglio Laurin, G., Chen, Q., Lindsell, J. A., Coomes, D. A., Frate, F. D., Guerriero, L., Pirotti, F., & Valentini, R. (2014). Above ground biomass estimation in an African tropical forest with lidar and hyperspectral data. *ISPRS Journal of Photogrammetry and Remote Sensing*, 89, 49-58.

Wang, G., Oyana, T., Zhang, M., Adu-Prah, S., Zeng, S., Lin, H., & Se, J. (2009). Mapping and spatial uncertainty analysis of forest vegetation carbon by combining national forest inventory data and satellite images. *Forest Ecology and Management*, 258(7), 1275-1283.

Yanai, R. D., Battles, J. J., Richardson, A. D., Blodgett, C. A., Wood, D. M., & Rastetter, E. B. (2010). Estimating uncertainty in ecosystem budget calculations. *Ecosystems*, 13(2), 239-248.

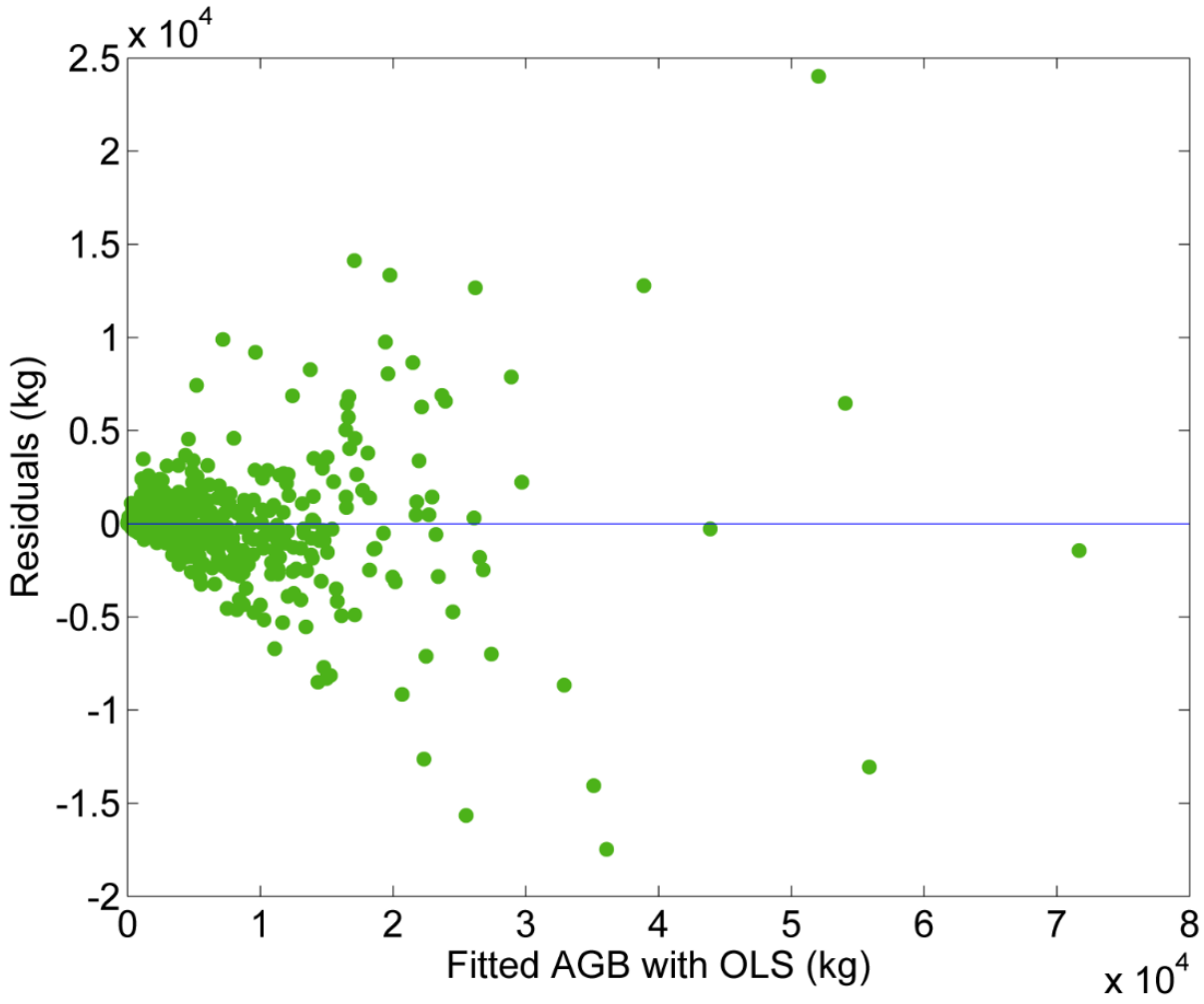
Zolkos, S. G., Goetz, S. J., & Dubayah, R. (2013). A meta-analysis of terrestrial aboveground biomass estimation using lidar remote sensing. *Remote Sensing of Environment*, 128, 289-298.

888 Fig. 1 The study area (transects in the middle upper figure), lidar transects (left and right figures
889 as examples), and 40 by 40 m field plots (see squares over the lidar transects as examples).



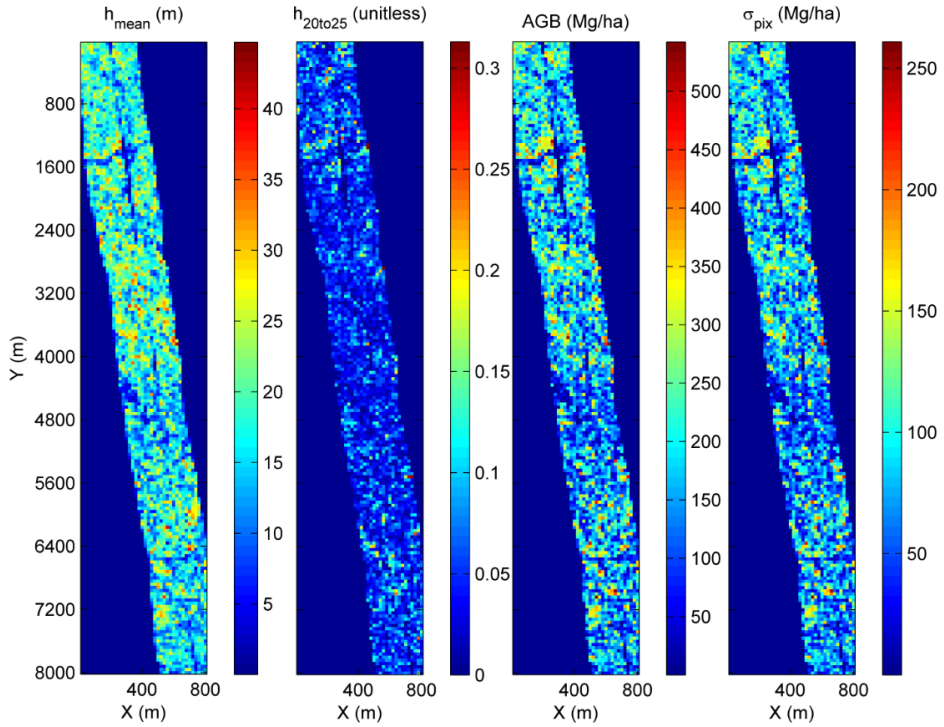
890

891 Fig. 2 Residuals of AGB fitted using an OLS approach that show heteroscedasticity.
892



893
894

Fig 3. Maps of lidar metrics, AGB, and errors. h_{mean} and h_{20to25} are two lidar metrics. AGB is the mapped aboveground biomass, σ_{pix} is the pixel-level AGB error, $\sigma_{\varepsilon,pix}$, $\sigma_{f,pix}$, and $\sigma_{z,pix}$ are errors related to lidar-biomass model residuals, model parameters, and lidar metrics, respectively.



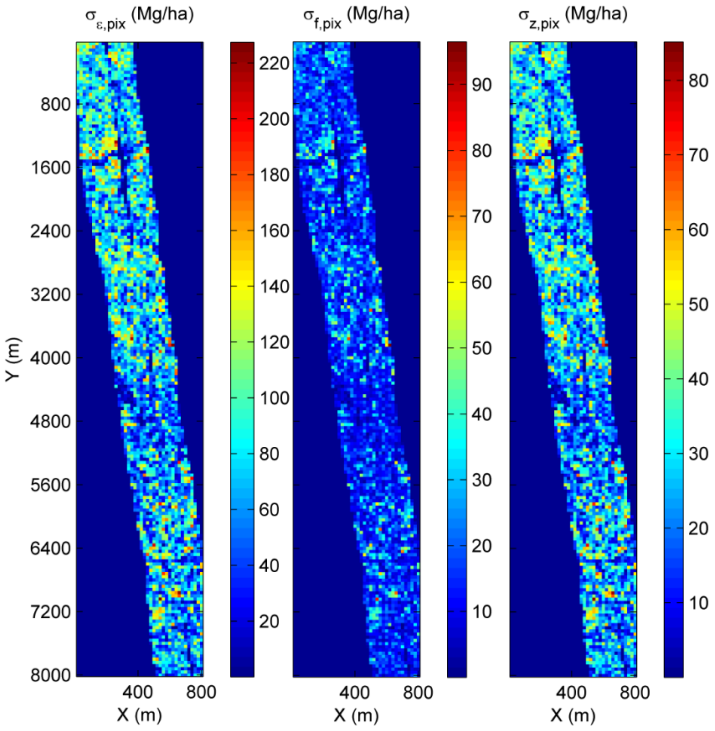


Figure1

Click here to download high resolution image

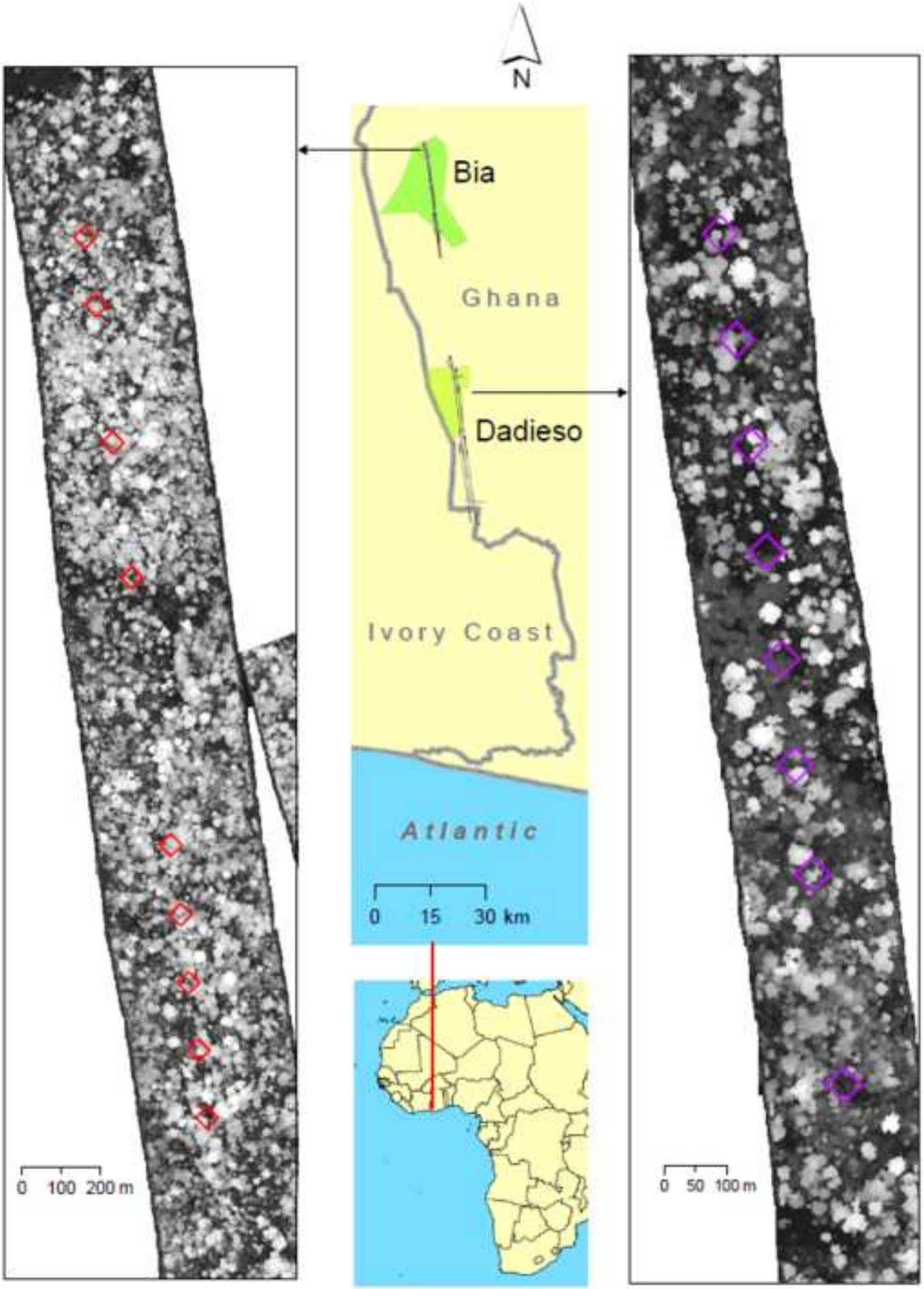


Figure2

Click here to download high resolution image

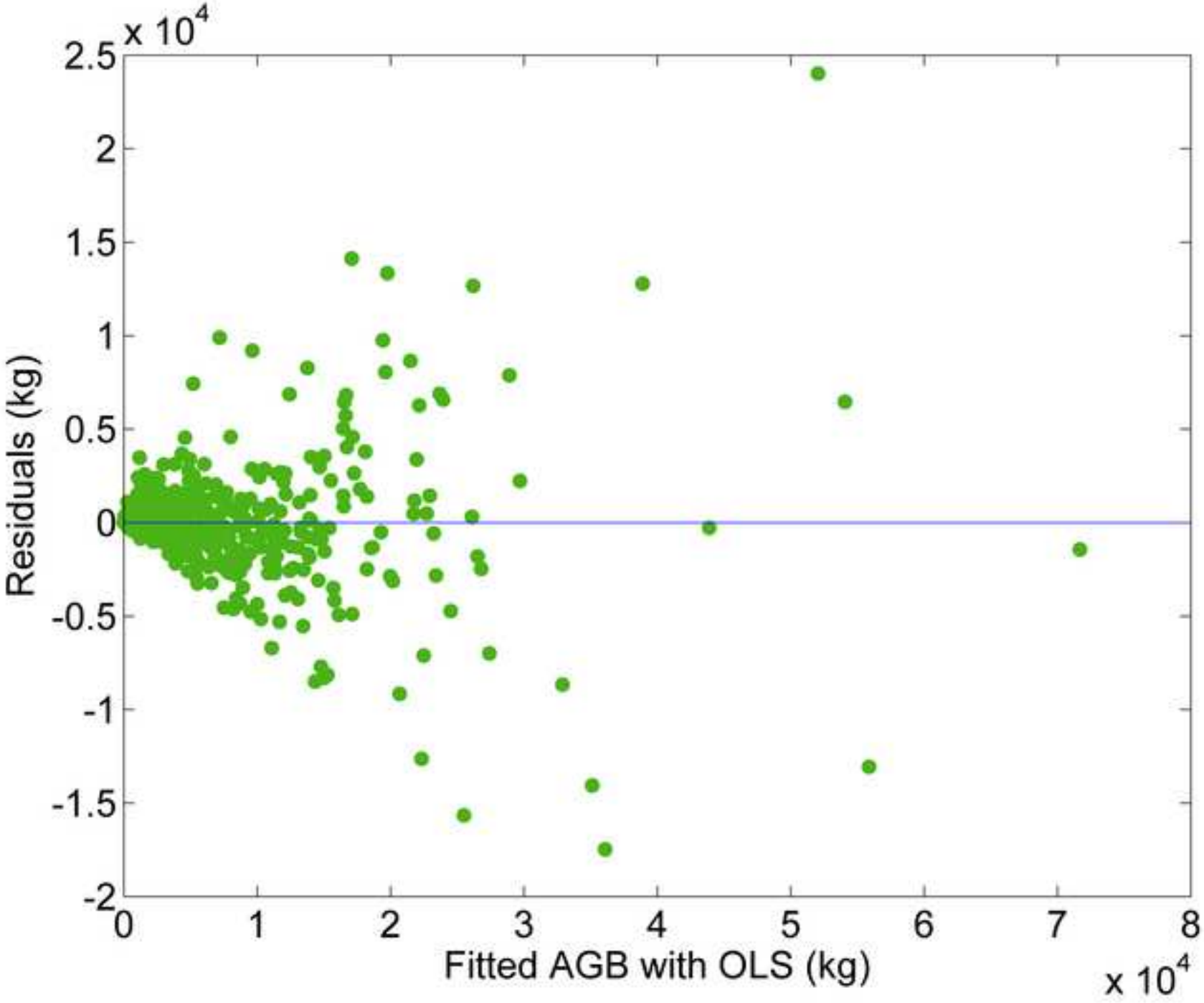


Figure3_part1

Click here to download high resolution image

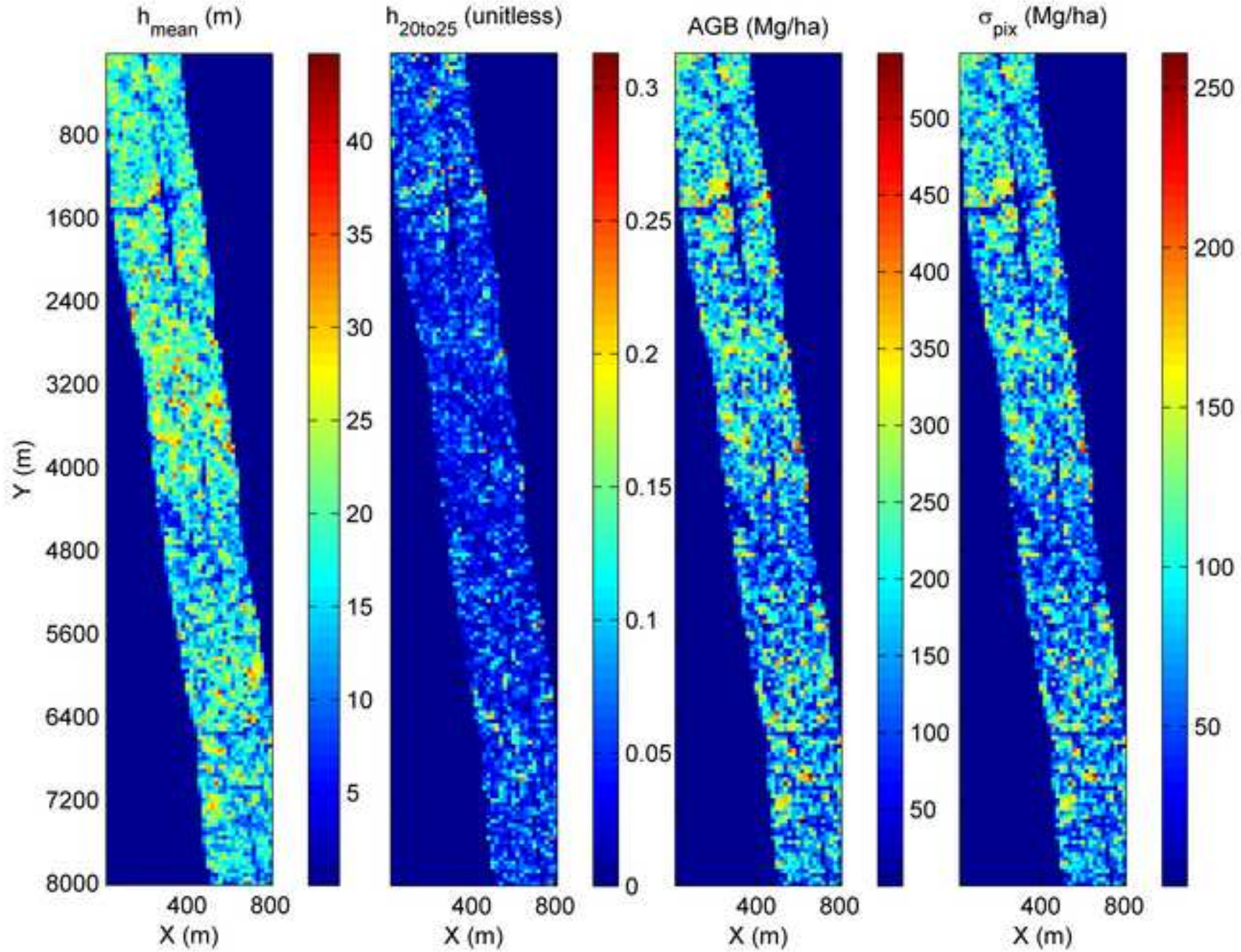


Figure3_part2

Click here to download high resolution image