

RESEARCH

Open Access



# Combined effect of pulse density and grid cell size on predicting and mapping aboveground carbon in fast-growing Eucalyptus forest plantation using airborne LiDAR data

Carlos Alberto Silva<sup>1,2\*</sup>, Andrew Thomas Hudak<sup>2</sup>, Carine Klauberg<sup>2</sup>, Lee Alexandre Vierling<sup>1</sup>, Carlos Gonzalez-Benecke<sup>3</sup>, Samuel de Padua Chaves Carvalho<sup>4</sup>, Luiz Carlos Estraviz Rodriguez<sup>5</sup> and Adrián Cardil<sup>6</sup>

## Abstract

**Background:** LiDAR remote sensing is a rapidly evolving technology for quantifying a variety of forest attributes, including aboveground carbon (AGC). Pulse density influences the acquisition cost of LiDAR, and grid cell size influences AGC prediction using plot-based methods; however, little work has evaluated the effects of LiDAR pulse density and cell size for predicting and mapping AGC in fast-growing Eucalyptus forest plantations. The aim of this study was to evaluate the effect of LiDAR pulse density and grid cell size on AGC prediction accuracy at plot and stand-levels using airborne LiDAR and field data. We used the Random Forest (RF) machine learning algorithm to model AGC using LiDAR-derived metrics from LiDAR collections of 5 and 10 pulses  $m^{-2}$  (RF5 and RF10) and grid cell sizes of 5, 10, 15 and 20 m.

**Results:** The results show that LiDAR pulse density of 5 pulses  $m^{-2}$  provides metrics with similar prediction accuracy for AGC as when using a dataset with 10 pulses  $m^{-2}$  in these fast-growing plantations. Relative root mean square errors (RMSEs) for the RF5 and RF10 were 6.14 and 6.01%, respectively. Equivalence tests showed that the predicted AGC from the training and validation models were equivalent to the observed AGC measurements. The grid cell sizes for mapping ranging from 5 to 20 also did not significantly affect the prediction accuracy of AGC at stand level in this system.

**Conclusion:** LiDAR measurements can be used to predict and map AGC across variable-age Eucalyptus plantations with adequate levels of precision and accuracy using 5 pulses  $m^{-2}$  and a grid cell size of 5 m. The promising results for AGC modeling in this study will allow for greater confidence in comparing AGC estimates with varying LiDAR sampling densities for Eucalyptus plantations and assist in decision making towards more cost effective and efficient forest inventory.

**Keywords:** Carbon modeling, Remote sensing, Modeling, Forest inventory, Random forest

## Background

Atmospheric carbon dioxide concentration [ $CO_2$ ] has increased by 40% since pre-industrial times, contributing

greatly to climate change [1]. Managing the exchange of  $CO_2$  and other greenhouse gases between the biosphere and the atmosphere is an important strategy for mitigating climate change [2]. Forest ecosystems play a key role in the global carbon cycle [3–6], since carbon is exchanged naturally between forests and the atmosphere through photosynthesis, respiration, decomposition and combustion [7]. Forest management can therefore alter

\*Correspondence: carlos\_englorestal@outlook.com

<sup>1</sup> Department of Natural Resources and Society, College of Natural Resources, University of Idaho, (UI), 875 Perimeter Drive, Moscow, ID 83843, USA

Full list of author information is available at the end of the article

the amount and magnitude of CO<sub>2</sub> exchange between forests and the atmosphere, thus serving as an effective and economical means to help mitigate climate change [8].

Forest plantations cover approximately 1% of the tropics (40–50 million ha) and have great capacity to store carbon [9], both during growth and in the form of durable forest products after harvest. Forest plantations play a role in terms of carbon in reducing further degradation and deforestation of natural forests, as well as providing alternative to the fossil fuels. *Eucalyptus* spp. are fast-growing with desirable wood qualities and are, therefore, a preferred species in plantations; they are widely grown in the tropics and subtropics, especially in Brazil, India, and China [10]. At present, Eucalyptus is grown on more than 20 million ha of plantation land around the world [11]. Eucalyptus is the dominant hardwood plantation species in Brazil, where it has been planted on more than 3.1 million hectares, accounting for approximately 57% of the country's total reforested area [12]. Most Eucalyptus plantations are managed in short rotations (6–8 years) and the mean annual increment is approximately 40 m<sup>3</sup> ha<sup>-1</sup> year<sup>-1</sup> roundwood, ranging from 25 to 60 m<sup>3</sup> ha<sup>-1</sup> year<sup>-1</sup> depending on the level of environmental stress [13]. Quantifying the substantial roles of fast-growing Eucalyptus plantation on AGC stores, as sources of carbon emissions and as carbon sinks, has become key to understanding the global carbon cycle [6, 14].

Forest inventory in Eucalyptus plantations is usually conducted annually to monitor forest growth [13]. The aboveground carbon (AGC) production of *Eucalyptus* spp. is extremely high compared to natural forests especially when trees are grown for timber production and therefore contributes strongly to the reduction of atmospheric CO<sub>2</sub> [9, 14, 15]. Remote sensing of forest AGC has received increased attention during the last decade due to its relevance to global carbon cycle modelling and to international programs aimed at reducing greenhouse gas emissions [15, 16]. LiDAR (Light Detection and Ranging) is a powerful remote sensing technology for predicting forest attributes [17], since it enables precise mapping of the landscape distribution of forest attributes at high spatial resolution and in a relatively short time compared to conventional methods [15]. In particular, many studies showed good relationships between parameters derived from airborne LiDAR data and forest measures such as canopy height [18–20], basal area [21, 22], stem volume [23–26] and aboveground carbon [7, 15, 27–31].

While airborne LiDAR is increasingly used to map forest attributes at the landscape level, multitemporal LiDAR data acquisition over large areas is expensive [32]. Many factors influence the cost of LiDAR data, including normal cost variables, such as project size, location, and

deliverables, as well as market variables, such as competition amongst LiDAR vendors. One of the most important variables affecting the cost of acquisition of LiDAR data is pulse density [33, 34]. Pulse density is defined by the number of pulses sent by the sensor per m<sup>2</sup> (pulses m<sup>-2</sup>), and as pulse density increases the acquisition cost increases as well [32].

There has been a wide interest in understanding how reducing pulse density affects the accuracy of inventory information derived from LiDAR data [34–37]. For instance, a study conducted in a Douglas-fir plantation in South Island, New Zealand, found that the precision of models to predict forest attributes, such as mean height, volume and mean diameter at breast height (DBH), remained stable until densities of 10 pulses m<sup>-2</sup> were culled to 2–3 pulses m<sup>-2</sup> [38]. Moreover, they also found that for the scenario where the DTM created from high pulse density was used to height normalize a downscaled point clouds for corresponding LiDAR-derived metrics, little change in the precision of regression models was found until pulse densities of 0.2–0.04 pulses m<sup>-2</sup> were reached [38]. In mixed conifer forest in Washington state, USA, model precision was more affected by sample plot size than pulse density [39]. In mixed conifer-hardwood in Canada, pulse density could be reduced from 3.2 to 0.5 pulses m<sup>-2</sup> with little effect on the quality of inventory results [40].

Besides the effect of pulse density on forest attribute prediction, there is also interest in understanding how cell size combined with pulse density affects the prediction of AGC at the stand level. Most LiDAR studies have been mapping forest attributes at stand level with grid cell size similar to the size of the sample plots used for calibrating the models [19, 22, 37], which is the general recommendation [42]. However, the shapes of the inventory plots are usually different from the cell shapes utilized in the prediction maps [43] that are normally square. It seems that at high pulse density, it would be advantageous to map forest attributes with higher spatial precision; i.e., at a finer grid resolution than the plot size, especially in Eucalyptus plantation, where the trees are planted in a grid and in most cases are hybrid clones. However, it is not clear how different grid cell sizes could influence stand level forest attributes estimates due to the degree of variability within stands and how LiDAR predictions are derived within a cell [43]. For instance, a study evaluating the influence of six prediction cell sizes (2, 4, 10, 25, 50, and 100 m squares) and two prediction methods (parametric vs. non-parametric) on LiDAR-derived stand-level estimates of total volume found not trend of smaller cell sizes producing lower values of volume, while larger cell sizes tended to give higher estimates [43].

Understanding the combined influence of pulse density and grid cell sizes on AGC prediction and mapping

allows us to understand how well AGC estimates can be compared over time, as LiDAR technology and flight characteristics vary. Although the effects of LiDAR pulse density on the prediction accuracy of forest attributes has been widely investigated in boreal and temperate forests [33–36, 38–40], no studies have been conducted in fast-growing plantations with the goal of assessing the combined effect of pulse density and grid cell size on predicting and mapping aboveground carbon. Therefore, the aim of this study was first to predict AGC in eight fast-growing Eucalyptus plantations in southeast Brazil using metrics derived from LiDAR with pulse densities of 5 and 10 pulses  $m^{-2}$ ; and second, to assess the combined effect of these two pulse densities and grid cell sizes on AGC predictions at plot and stand levels. We hypothesized that due to the high homogeneity of this type of plantation, LiDAR data with a density of 5 pulses  $m^{-2}$  would provide AGC prediction accuracies similar to those from LiDAR data with a density of 10 pulses  $m^{-2}$ , at grid cell sizes ranging from 5 to 20 m.

## Methods

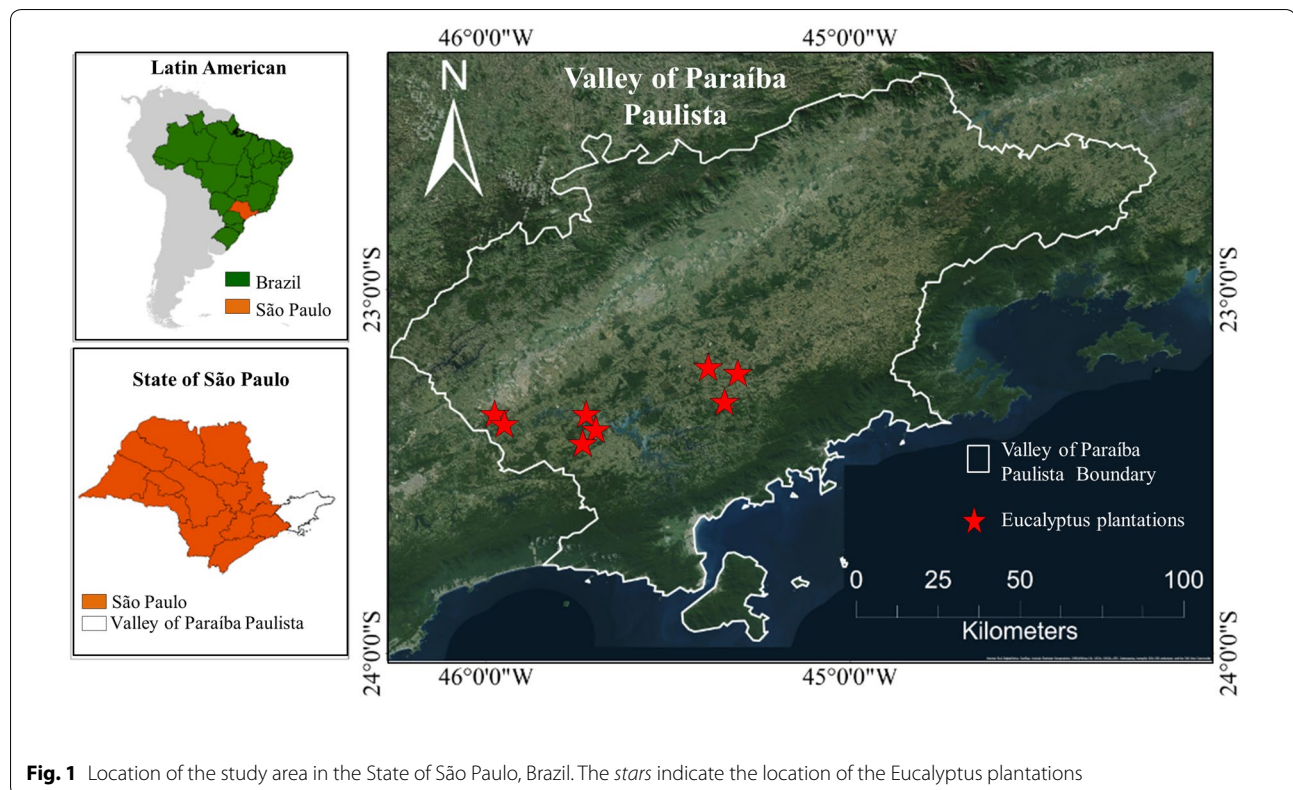
### Study area description

The study area consisted of eight farms located within the Paraíba Valley in the state of São Paulo, Brazil (Fig. 1). The climate of the region is characterized as humid

subtropical, with dry winters and hot summers [41]. Mean annual precipitation is  $\sim 1200$  mm; mean temperature ranges from a minimum of 17.1 °C in the coldest month (July) to a maximum of 23.9 °C in the hottest month (February). The topography in the selected plantations is complex with high relief, ranging from 578 to 1310 m in elevation. The plantations contain hybrid clones of two Eucalyptus species, *Eucalyptus grandis* W. Hill ex Maid and *Eucalyptus urophylla* S.T. Blake. The plantations are managed by Fibria Celulose S/A, a pulp company. Stand age across the farms was variable and ranged from 3 to 8 years. All the trees were planted in 3 m between rows and 2 m tree within a row (1667 trees per ha).

### Field data collection

A total of 136 circular plots of 400  $m^2$  each were established across the eight farms (Site ID, Table 1) for stand measurement during the months of January and March of 2012. All plots were georeferenced with a geodetic GPS with differential correction capability (Trimble Pro-XR). Sample plot centers were geolocated with a horizontal error of up to 10 cm. In each plot, all trees were measured for diameter at breast height (DBH; cm) and a random subsample (15%) of trees for maximum height (Ht; m). For trees in the plot that were not directly



**Fig. 1** Location of the study area in the State of São Paulo, Brazil. The stars indicate the location of the Eucalyptus plantations

**Table 1 Descriptive values of the biometric parameters of the network of plots inventoried in the study area (January of 2012)**

| Sites ID | Area (ha) | DBH (cm) |      | Ht (m) |      | AGC (Mg.ha <sup>-1</sup> ) |       | Age (years) | N plots |
|----------|-----------|----------|------|--------|------|----------------------------|-------|-------------|---------|
|          |           | Mean     | SD   | Mean   | SD   | Mean                       | SD    |             |         |
| F987     | 39.53     | 8.82     | 1.57 | 11.65  | 2.55 | 15.76                      | 3.75  | 2.3         | 14      |
| F986     | 94.16     | 12.73    | 0.76 | 18.60  | 1.45 | 45.45                      | 6.19  | 3.3         | 21      |
| F849     | 138.96    | 14.14    | 0.60 | 22.16  | 1.54 | 62.71                      | 6.69  | 4.7         | 26      |
| F950     | 86.72     | 13.73    | 0.80 | 21.46  | 1.39 | 62.97                      | 9.81  | 5.5         | 17      |
| F184     | 58.34     | 14.57    | 0.70 | 23.74  | 1.19 | 64.59                      | 6.81  | 5.9         | 14      |
| F166     | 84.35     | 14.59    | 0.95 | 24.17  | 1.38 | 69.76                      | 9.77  | 6.1         | 17      |
| F948     | 79.33     | 13.70    | 1.20 | 23.29  | 1.68 | 60.42                      | 13.12 | 6.8         | 12      |
| F634     | 84.80     | 15.26    | 1.20 | 25.24  | 2.35 | 73.75                      | 9.87  | 8.0         | 15      |

SD standard deviation

measured for Ht, the inventory team of Fibria Celulose S/A company predicted heights from hypsometric models, which use DBH as the predictor of Ht. For mapping validation purpose, trees were sub-sampled in square sub-plots of 25 m<sup>2</sup> (5 × 5 m), 100 m<sup>2</sup> (10 × 10 m) and 225 m<sup>2</sup> (15 × 15 m) located at the center of each plot. The total aboveground carbon—AGC (Mg.tree<sup>-1</sup>) for the plots and sub-plots was obtained through the allometric model according to [15], employing as independent variables the logarithm of DBH and the Ht, and as dependent variables the AGC, as shown in the equation below:

$$\ln(\text{AGC}) = -2.87 + 1.95 \times \ln(\text{DBH}) + 0.44 \times \ln(\text{Ht})$$

where: DBH is tree diameter over bark at breast height (1.37 m) in cm; Ht is tree height (m). The AGC model has adjusted coefficients of determination (Adj.R<sup>2</sup>) of 0.97, absolute and relative root mean squared errors (RMSE) of 4.57 kg tree<sup>-1</sup> and 12.38%, respectively. The ln (AGC) was back transformed to natural scale, and multiplied by a correction factor of 1.03 [exp (0.5 × MSE)] to remove the bias added by the log transformation. The summed AGC stock of all trees within the plots and sub-plots was then divided by the plot and sub-plot area to calculate the AGC at plot-level in Mg.ha<sup>-1</sup>. The summary statistics of AGC stocks (Mg.ha<sup>-1</sup>) measured at plot-level in the farms evaluated are presented in Table 1.

#### LiDAR data acquisition and data processing

LiDAR data were obtained with a Riegl LMS-Q680I sensor mounted on a Piper Seneca II aircraft during the months of January and March of 2012. Data were acquired subsequently with pulse densities of 5 and 10 pulses m<sup>-2</sup>; the characteristics and precision of the LiDAR data are listed in Table 2. LiDAR data processing consisted of several steps that ingested the LiDAR point cloud data and provided two major outputs: the digital

terrain model (DTM) and the LiDAR-derived canopy structure metrics.

LiDAR data processing for both pulse densities occurred with the following sequence of steps using FUSION/LDV toolkit software [44]. (I) Ground returns were classified using a filtering algorithm adapted from [45] based on linear prediction [46]. (II) Digital Terrain Models (DTMs) of 1 m spatial resolution were then developed using the classified ground returns. (III) Normalized LiDAR point clouds were obtained by subtracting the DTM elevation from each LiDAR return. (IV) Normalized point clouds were subset within each of the 136 sample plots, and then structure metrics were computed for each plot using only all returns above 2 m to remove the returns not belonging to tree crowns (e.g., hits on near-ground shrubs, forbs, grasses, etc.). (V) To evaluate the effect of grid cell size on the AGC prediction at stand level, we computed LiDAR metrics at grid cell sizes of 5, 10, 15, and 20 m. From the point cloud, it is possible to generate many LiDAR metrics, however, we generated only those metrics that have been frequently used as candidate predictors for forest attribute modeling in other studies [15, 18–22, 29, 47]. Therefore, a total of 27 LiDAR metrics calculated from all returns were considered for AGC modelling at the plot and stand levels (Table 3).

**Table 2 LiDAR flight characteristics**

| Parameter              | Value                           |
|------------------------|---------------------------------|
| Average flight height  | 422.94 m                        |
| Pulse density          | 5 and 10 pulses m <sup>-2</sup> |
| Pulse frequency        | 400 kHz                         |
| Scan angle             | ±45°                            |
| Laser wavelength       | 1055 nm                         |
| Average aircraft speed | 57 m/s (205.20 km/h)            |
| Horizontal precision   | 0.1–0.15 m (1.0 sigma)          |



**Table 3** LiDAR-derived canopy height metrics considered as candidate variables for predictive AGC models [43]

| Metric | Description  |
|--------|--|
| HMIN   | Minimum canopy height                                  |
| HMAX   | Maximum canopy height                                  |
| HMEAN  | Mean canopy height                                     |
| HMAD   | Median canopy height                                   |
| HSD    | Standard deviation of canopy height                    |
| HSKEW  | Skewness of canopy height                              |
| HKURT  | Kurtosis of canopy height                              |
| HVAR   | Variance of canopy height                              |
| HCV    | Coefficient of variation canopy height                 |
| HMODE  | Mode canopy height                                     |
| H01TH  | 1st percentile of canopy height                        |
| H05TH  | 5th percentile of canopy height                        |
| H10TH  | 10th percentile of canopy height                       |
| H20TH  | 20th percentile of canopy height                       |
| H25TH  | 25th percentile of canopy height                       |
| H30TH  | 30th percentile of canopy height                       |
| H40TH  | 40th percentile of canopy height                       |
| H50TH  | 50th percentile of canopy height                       |
| H60TH  | 60th percentile of canopy height                       |
| H70TH  | 70th percentile of canopy height                       |
| H75TH  | 75th percentile of canopy height                       |
| H80TH  | 80th percentile of canopy height                       |
| H90TH  | 90th percentile of canopy height                       |
| H95TH  | 95th percentile of canopy height                       |
| H99TH  | 99th percentile of canopy height                       |
| CR     | Canopy relief ratio [(HMEAN-HMIN)/(HMAX-HMIN)]         |
| COV    | Canopy cover (percentage of first return above 1.30 m) |

### Random forest modeling

We used the random forest package [48, 49] in R [50] to create random forest (RF) models for predicting AGC at the plot and stand levels from the LiDAR dataset of 5 and 10 pulses  $m^{-2}$  (RF5 and RF10). In the analysis, we defined the number of classification trees (ntree) as 1000, and for the number of variables randomly (mtry) sampled as candidates at each split, we used the default value, which for regression is defined as  $p/3$ , where  $p$  is the number of covariates. For the remaining parameters, we used the default values, which are all defined in [50] as well.

The selection of the best LiDAR metrics to be included in the models was performed in two steps. First, even though highly correlated variables won't cause multicollinearity issues in RF, Pearson's correlation ( $r$ ) was used to identify highly correlated predictor variables ( $r > 0.85$ ) as presented in [29] and [15, 47]. If a given group (2 or more) of LiDAR metrics were highly correlated, we retained only one metric by excluding the others that were most highly correlated with the retained metrics.

Second, we identified the most important metrics based on the Model Improvement Ratio (MIR), a standardized measure of variable importance [51, 52]. MIR scores are derived by dividing raw variable importance scores (output from RF) by the maximum variable importance score, so that MIR values range from 0 to 1. MIR scores allow for variable importance comparisons among different RF models. For each of the two AGC models, we ran 1000 iterations of RF that included the  $<0.85$  correlated preliminary set of LiDAR metrics to create distributions of MIR for each metric. Running 1000 iterations of RF produced consistent MIR distributions and avoided unnecessary processing time. To create parsimonious models, we reserved metrics for final RF models that were consistent and exhibited the highest mean MIR values.

The accuracy of estimates for each model was evaluated in terms of percentage of variation explained (Pseudo- $R^2$ ) in the RF models, Root Mean Square Error (RMSE), and Bias (both absolute and relative) computed by the linear relationship between observed and predicted AGC:

$$Pseudo - R^2 = \left[ 1 - \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{\sum_{i=1}^n (\frac{y}{n} - y_i)^2} \right] \quad (1)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{n}} \quad (2)$$

$$Bias = \frac{1}{n} \sum_{i=1}^n (y_i - \hat{y}_i) \quad (3)$$

where  $n$  is the number of plots,  $y_i$  is the observed value for plot  $i$ , and  $\hat{y}_i$  is the predicted value for plot  $i$ . Moreover, relative RMSE and biases were calculated by dividing the absolute values (Eqs. 1, 2) by the mean of the observed AGC. We defined acceptable model precision and accuracy as a relative RMSE and Bias of  $\leq 15\%$  to have a model precision and accuracy higher than or equal to the conventional forest inventory standard in fast-growing Eucalyptus forest plantations in Brazil [14, 47].

The performance of the RF models to predict AGC was also evaluated by means of a leave-one-out cross-validation (LOOCV) strategy [e.g. 22, 49]. Statistical equivalence tests were employed [53, 54] to assess whether the AGC predictions were statistically similar at (i.e., equivalent;  $p$  value  $>0.05$ ) to the field-based AGC stocks. Following [55], we employed a regression-based equivalence test for intercept equality to 0 (i.e., the mean of predicted AGC is equal to the mean of the field-based AGC) and slope equality to 1 (i.e., if the pairwise, predicted and observed AGC are equal, the regression will have a slope of 1). A description of equivalence tests can be also found

in the ‘equivalence’ package in R [56], and examples of equivalence plots in LiDAR studies can be found in [15, 29, 45, 53, 55]. We also computed RMSE, Bias and the adjusted Coefficient of determination ( $Adj.R^2$ ) from the observed and predicted AGC from LOOCV. Finally, the effects of pulse density of the AGC predictions at the plot-level were investigated by the comparison of RMSE and Bias statistics across pulse densities.

**Assessing effects of cell size on the AGC prediction and mapping at the stand level**

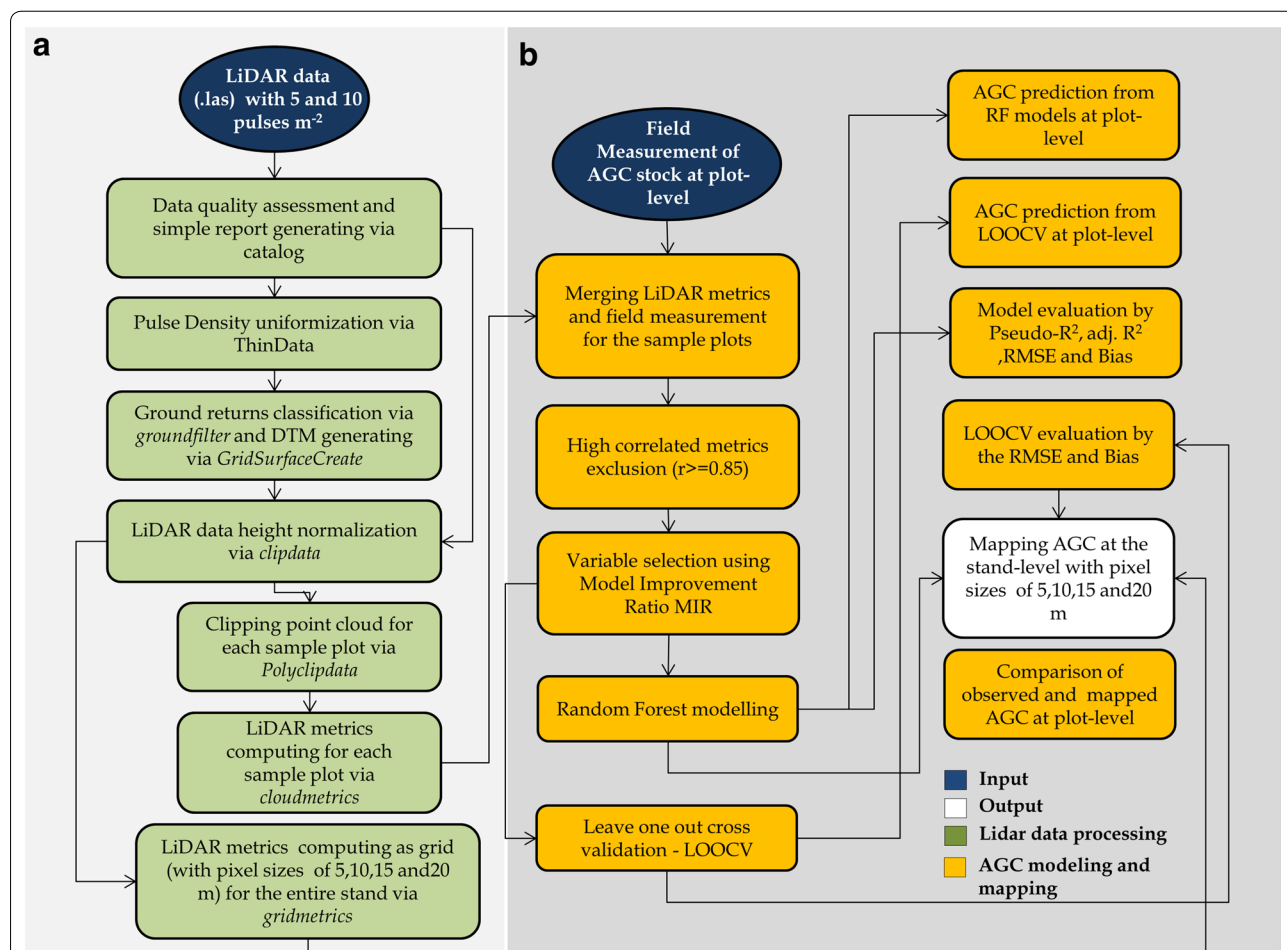
As the study area is characterized by fast-growing plantations, with tree hybrid clones equally spaced in the ground, we assumed that AGC maps could be generated with high spatial resolution, even though the model is based on a sample unit of 400 m<sup>2</sup>. Therefore, we applied the predictive model across the landscape to map AGC at

spatial resolutions of 5, 10, 15, and 20 m, the last of which equates to the sample plot size of 400 m<sup>2</sup>. To evaluate the effect of grid cell size on the AGC map prediction, we extracted the AGC for each study site, and boxplots were generated to compare variability of AGC at the stand level containing various grid cell sizes. Moreover, from the maps we extracted predicted AGC within the plot and sub-plot boundaries to validate the maps by the comparison with reference AGC through an equivalence test [53–57]. An overview of the methodology is outlined in Fig. 2.

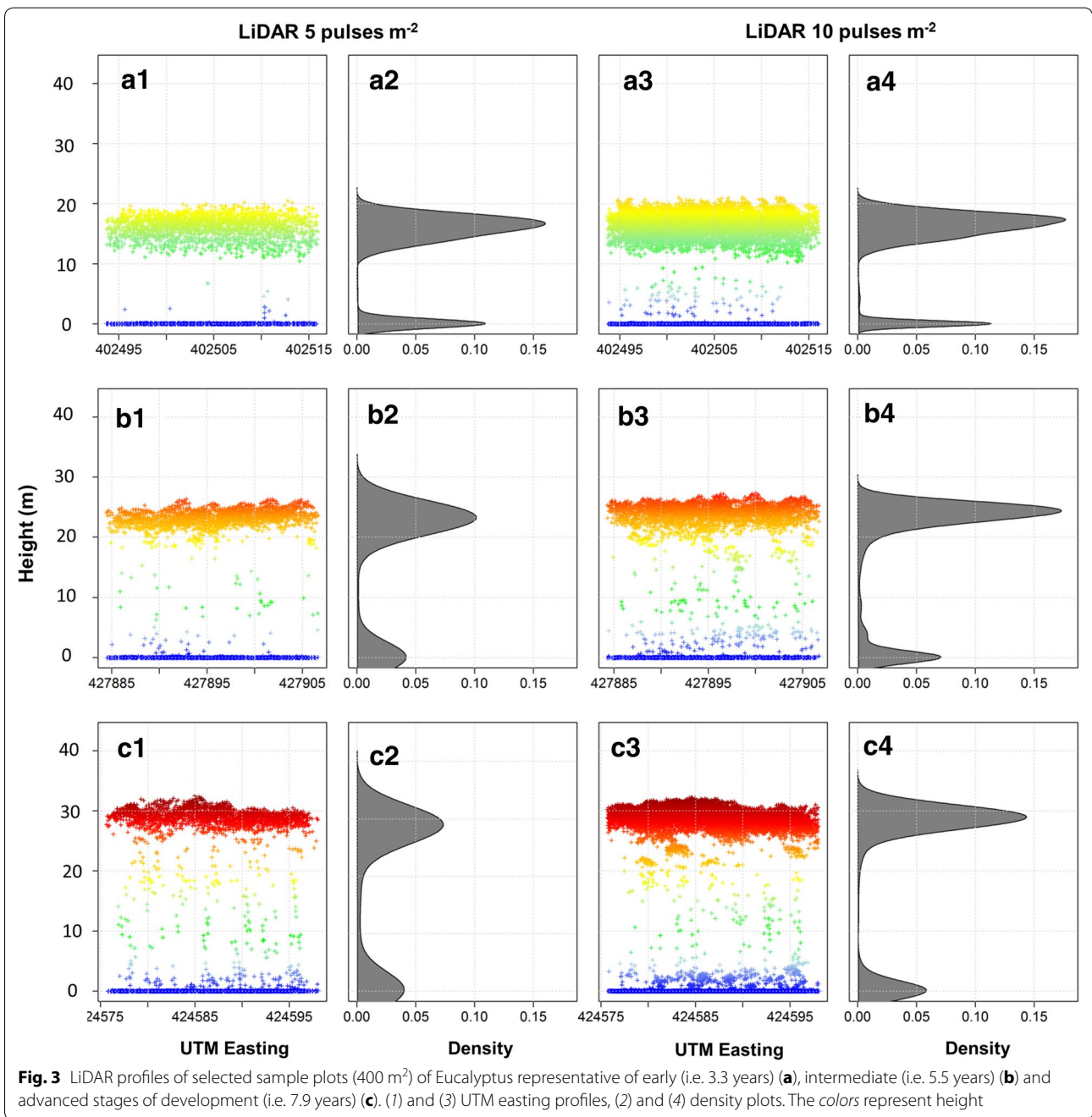
**Results**

**Canopy profile of Eucalyptus plantation across age and LiDAR pulse density**

The canopy profiles of Eucalyptus plantations across age and pulse density are represented in Fig. 3.



**Fig. 2** Procedure for predicting and mapping AGC from LiDAR and inventory plot data in Eucalyptus fast-growing plantations. Pearson’s correlation ( $r$ ); Model Improvement Ratio (MIR); Leave-one-out cross-validation (LOOCV); Adjusted coefficient of determination ( $adj.R^2$ ); Percentage of variation explained (Pseudo- $R^2$ ); Root mean squared error (RMSE). **a** The light gray panel corresponds to the LiDAR data processing step and **b** the gray panel correspond to the AGC modeling step



LiDAR-derived height increased from early (Fig. 3a) to intermediate (Fig. 3b) and advanced ages (Fig. 3c). The variation in height was affected more by stand age than by pulse density. From a probabilistic perspective, the canopy height distribution remained the same using datasets with 5 and 10 pulses m<sup>-2</sup>. The majority of LiDAR returns came from the upper canopy and the ground. However, as the stands approached harvest age (e.g. ~7–8 years), the number of LiDAR returns in the

0–5 m strata increased, indicating that the understory plant community increased in cover and height as the stand ages.

#### LiDAR metrics selection

Pearson’s correlation test (*r*) applied to the 27 candidate LiDAR metrics determined that 18 metrics were highly correlated (*r* > 0.85) for models using the 5 and 10 pulses m<sup>-2</sup>. For each dataset, we excluded all but one of the

highly correlated metrics (H99TH), and the remaining not highly correlated metrics were used to build prospective models after MIR analysis. The not highly correlated LiDAR metrics derived from the 5 and 10 pulses m<sup>-2</sup> dataset were HVAR, HCV, HSKE, HKUR, H01TH, H05TH, H10TH, H99TH and COV. In both correlation matrices for the LiDAR metrics derived from the 5 and 10 pulses m<sup>-2</sup>, there were positively correlated metrics, such as H99TH and HVAR ( $r = 0.65$ ;  $r = 0.66$ ;  $p$  value < 0.001), and negatively correlated metrics, such as HCV and COV ( $r = -0.34$ ;  $r = -0.30$ ;  $p$  value < 0.001), as shown in Table 4.

The remaining, not highly correlated LiDAR-derived metrics from 5 and 10 pulses m<sup>-2</sup> were correlated against each other and metrics such as H99TH and HVAR where highly stable across pulse density, while COV and H01TH showed to be less stable. Pearson's correlation between not highly correlated LiDAR-derived metrics from 5 and 10 pulses m<sup>-2</sup> are shown in the Table 5.

The most important LiDAR metrics for predicting AGC, both from the 5 and 10 pulses m<sup>-2</sup> datasets, were H99TH and HVAR. In addition to these two metrics, the H10TH and HSKE were also the most consistently

important metrics with higher MIR (Table 6). The variation in metrics selection was less affected by pulse density, and the LiDAR metrics selected according to MIR were the same for both pulse densities.

**RF model fit and cross validation**

Both RF models predicting AGC at plot level explained 83% of the variation in AGC. Greater pulse density did not improve model fit, and both RF5 and RF10 models resulted in a very low RMSE and Bias, indicating high model prediction accuracy. In addition to the RMSE and Bias,  $r$  and  $adj.R^2$  did not significantly differ between the RF5 and RF10 models (Table 7).

The LOOCV analysis indicated that the RF5 and RF10 models were stable resulting RMSE of 13.61 and 13.38% and Bias of -0.09 and -0.27% for the models derived from the 5 and 10 pulses m<sup>-2</sup> datasets, respectively. Moreover, the equivalence plots of observed versus predicted AGC both from the two models showed that predicted and observed AGC were equivalent ( $p$  values > 0.05 for intercepts and  $p$  values < 0.05 for slopes (Fig. 4a, b).

**Table 4 Pearson's correlation among the derived metrics from LiDAR pulse density of 5 (white) and 10 (gray) pulses m<sup>-2</sup>**

| r     | HVAR     | HCV      | HKUR     | H01TH    | H05TH    | H10TH    | H99TH   | COV      |
|-------|----------|----------|----------|----------|----------|----------|---------|----------|
| HVAR  | -        | 0.84***  | -0.50*** | -0.44*** | -0.50*** | -0.30*** | 0.66*** | -0.21*   |
| HCV   | 0.86***  | -        | -0.73*** | -0.58*** | -0.79*** | -0.66*** | 0.25**  | -0.30*** |
| HKUR  | -0.53*** | -0.73*** | -        | 0.50***  | 0.83***  | 0.70***  | 0.06    | 0.10     |
| H01TH | -0.48*** | -0.63*** | 0.51***  | -        | 0.53***  | 0.26***  | -0.13   | -0.29*** |
| H05TH | -0.57*** | -0.80*** | 0.85***  | 0.57***  | -        | 0.75***  | 0.17*   | 0.07**   |
| H10TH | -0.37*** | -0.67*** | 0.71***  | 0.26**   | 0.74***  | -        | 0.44*** | 0.40***  |
| H99TH | 0.65***  | 0.29***  | 0.02     | -0.18*   | 0.10     | 0.39***  | -       | 0.49***  |
| COV   | -0.30*** | -0.34*** | 0.12     | 0.08     | 0.25**   | 0.18*    | 0.45*** | -        |

The italic values indicates the values of  $r$  for metrics derived from 10 pulse m<sup>-2</sup>

\*\*\*  $p$  value < 0.001; \*\*  $p$  value < 0.01; \*  $p$  value < 0.05;  $p$  value < 0.1

**Table 5 Pearson's correlation between selected metrics from LiDAR pulse densities of 5 and 10 pulses m<sup>-2</sup>**

| r   | LiDAR-derived metrics (10 pulses m <sup>-2</sup> ) |          |         |         |         |         |         |      |
|---|--|----------|---------|---------|---------|---------|---------|------|
|   | HVAR   | HCV      | HKUR    | H01TH   | H05TH   | H10TH   | H99TH   | COV  |
| LiDAR-derived metrics (5 pulses m <sup>-2</sup> ) |  |          |         |         |         |         |         |      |
| HVAR  | 0.93***  |          |         |         |         |         |         |      |
| HCV   | 0.76***  | 0.87***  |         |         |         |         |         |      |
| HKUR  | -0.38***   | -0.55*** | 0.69*** |         |         |         |         |      |
| H01TH   | -0.38***   | -0.46*** | 0.40*** | 0.60*** |         |         |         |      |
| H05TH   | -0.37***   | -0.60*** | 0.60*** | 0.36*** | 0.68*** |         |         |      |
| H10TH   | -0.25**  | -0.55*** | 0.61*** | 0.17    | 0.61*** | 0.90*** |         |      |
| H99TH   | 0.64***  | 0.28***  | 0.03    | -0.17*  | 0.11    | 0.40*** | 1.00*** |      |
| COV   | 0.14   | -0.09*** | 0.13    | -0.19*  | 0.18*   | 0.45*** | 0.50*** | 0.13 |

\*\*\*  $p$  value < 0.001; \*\*  $p$  value < 0.01; \*  $p$  value < 0.05;  $p$  value < 0.1



**Table 6 Mean of model improvement ratio (MIR) among derived metrics from LiDAR pulse density of 5 and 10 pulses m<sup>-2</sup> (N = 1000)**

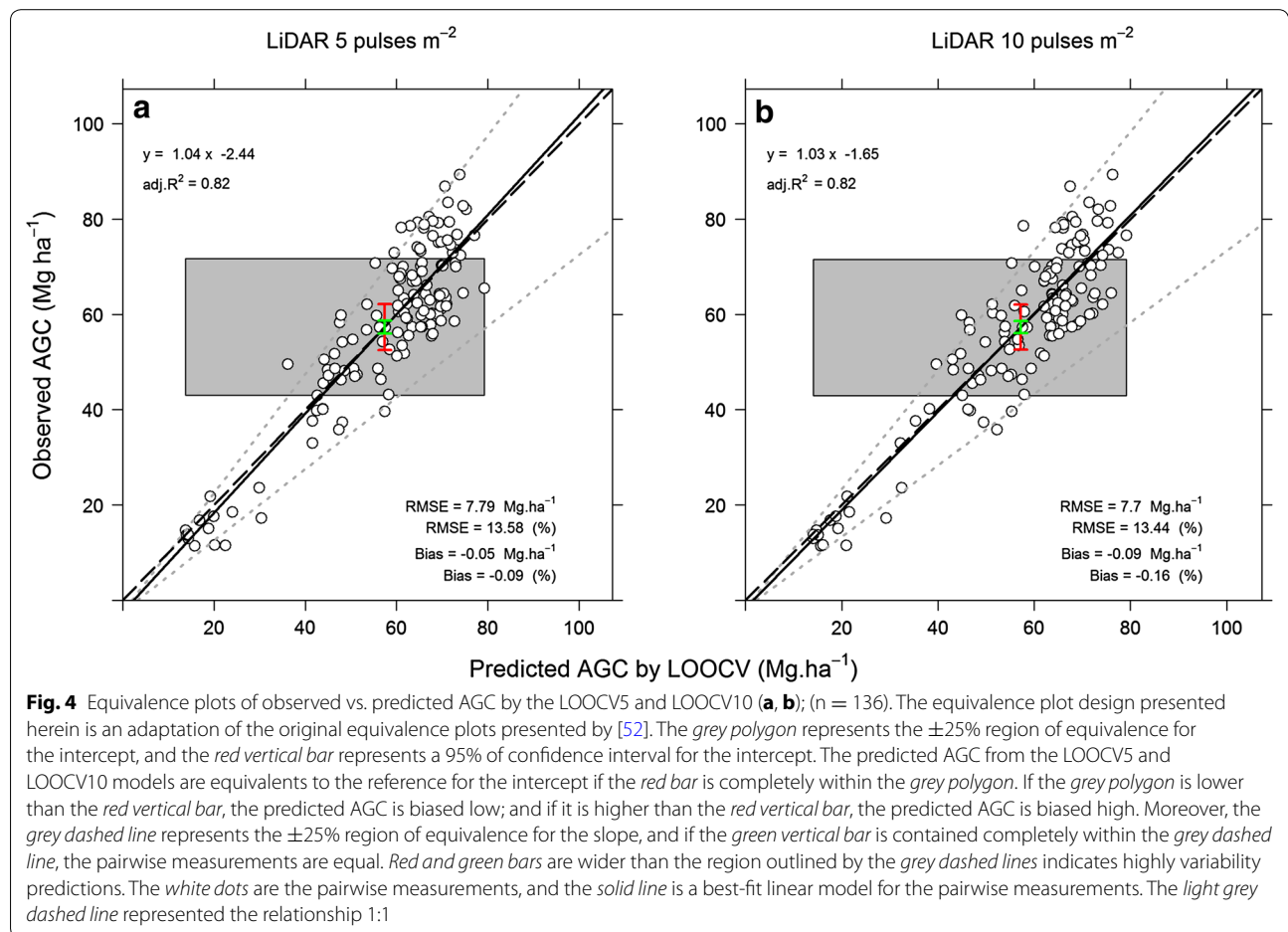
| Pulse density (pulse m <sup>-2</sup> ) | LiDAR-derived metrics |       |             |             |             |      |      |      |
|--|-----------------------|-------|-------------|-------------|-------------|------|------|------|
|  | H01TH                 | H05TH | H10TH       | H99TH       | HVAR        | HCV  | HKUR | COV  |
| 5                                      | 0.04                  | 0.13  | <i>0.47</i> | <i>1.00</i> | <i>0.68</i> | 0.21 | 0.17 | 0.02 |
| 10                                     | 0.03                  | 0.11  | <i>0.33</i> | <i>1.00</i> | <i>0.52</i> | 0.13 | 0.07 | 0.07 |

MIR > 0.3 are highlighted in italics

**Table 7 Random Forest models fitted for predicting AGC from LiDAR pulse densities of 5 and 10 pulses m<sup>-2</sup>**

| Pulse density | LiDAR derived metrics | Pseudo-R <sup>2</sup> | RMSE                |      | Bias                |       |
|---------------|-----------------------|-----------------------|---------------------|------|---------------------|-------|
|               |                       |                       | Mg ha <sup>-1</sup> | %    | Mg ha <sup>-1</sup> | %     |
| 5             | H99TH + HVAR + H10TH  | 81.79                 | 3.52                | 6.14 | -0.03               | -0.06 |
| 10            |                       | 82.17                 | 3.44                | 6.01 | -0.01               | -0.01 |

Pearson's correlation coefficient (r); root mean square error (RMSE)



**Fig. 4** Equivalence plots of observed vs. predicted AGC by the LOOCV5 and LOOCV10 (a, b); (n = 136). The equivalence plot design presented herein is an adaptation of the original equivalence plots presented by [52]. The grey polygon represents the ±25% region of equivalence for the intercept, and the red vertical bar represents a 95% of confidence interval for the intercept. The predicted AGC from the LOOCV5 and LOOCV10 models are equivalents to the reference for the intercept if the red bar is completely within the grey polygon. If the grey polygon is lower than the red vertical bar, the predicted AGC is biased low; and if it is higher than the red vertical bar, the predicted AGC is biased high. Moreover, the grey dashed line represents the ±25% region of equivalence for the slope, and if the green vertical bar is contained completely within the grey dashed line, the pairwise measurements are equal. Red and green bars are wider than the region outlined by the grey dashed lines indicates highly variability predictions. The white dots are the pairwise measurements, and the solid line is a best-fit linear model for the pairwise measurements. The light grey dashed line represented the relationship 1:1

**AGC predictions at plot-level**

There was no difference in the AGC prediction at plot level using LiDAR data of 5 or 10 pulses m<sup>-2</sup>. The RF

model AGC predictions for the 136 sample plots ranged from 12.99 to 82.37 for RF5, and 12.75 to 82.83 Mg. ha<sup>-1</sup> for RF10. In general, for both RF5 and RF10 the

AGC values were slightly overestimated during early and advanced stand ages, and slightly underestimated at intermediate ages. The Eucalyptus plantations containing younger stands showed the lowest AGC values (i.g. IDs F987 and F986), and advanced age stands contained the highest AGC stocks (i.g. IDs F166 and F634) (Fig. 5).

**Effect of the pulse density and cell size on the AGC prediction and mapping at stand level**

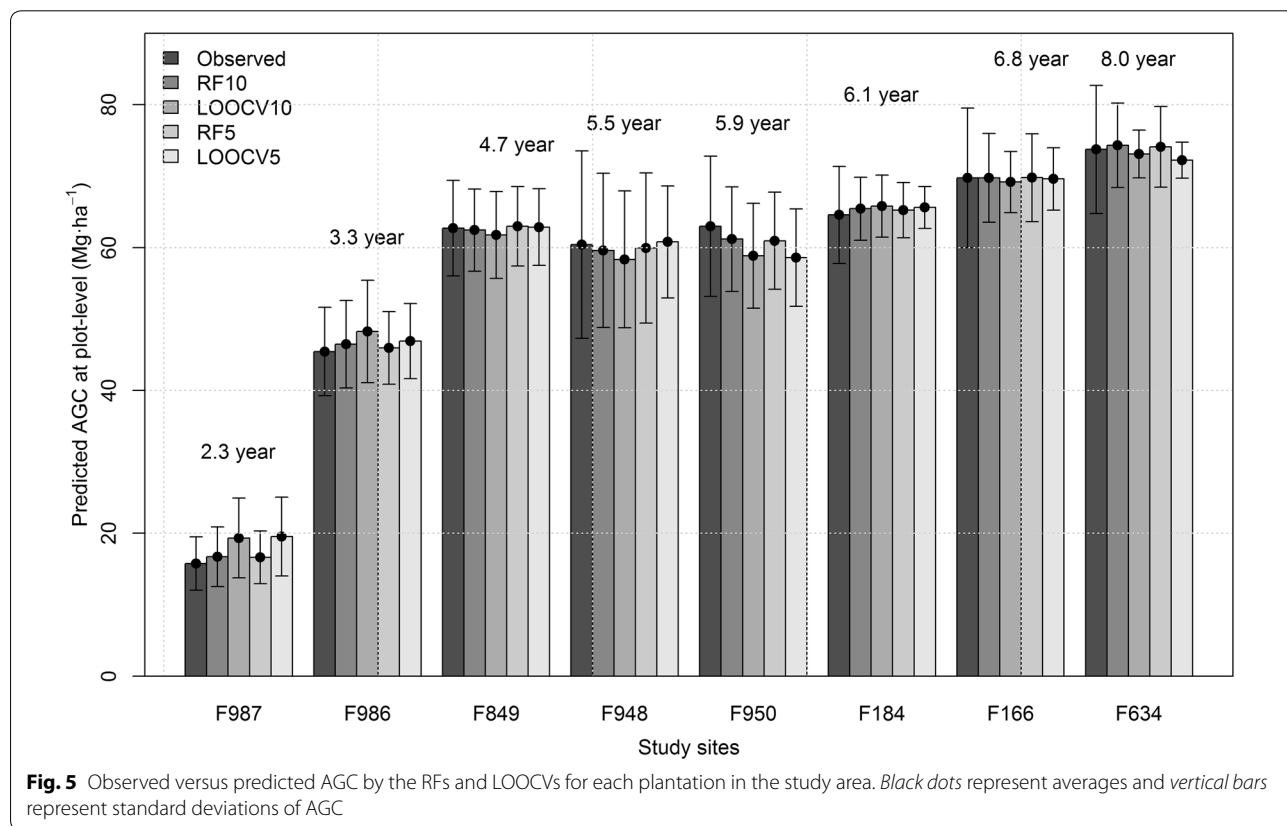
The variation of the predicted AGC at stand level, and across cell size and pulse density is presented in Fig. 6. Grid cell sizes higher than 5 m did not significantly affect AGC prediction at stand level. Even though the variability of AGC slightly decreased from 5 to 20 m cell size, we observed that in most of the sites, the AGC variability was similar in the maps with spatial resolution coarser than 5 m.

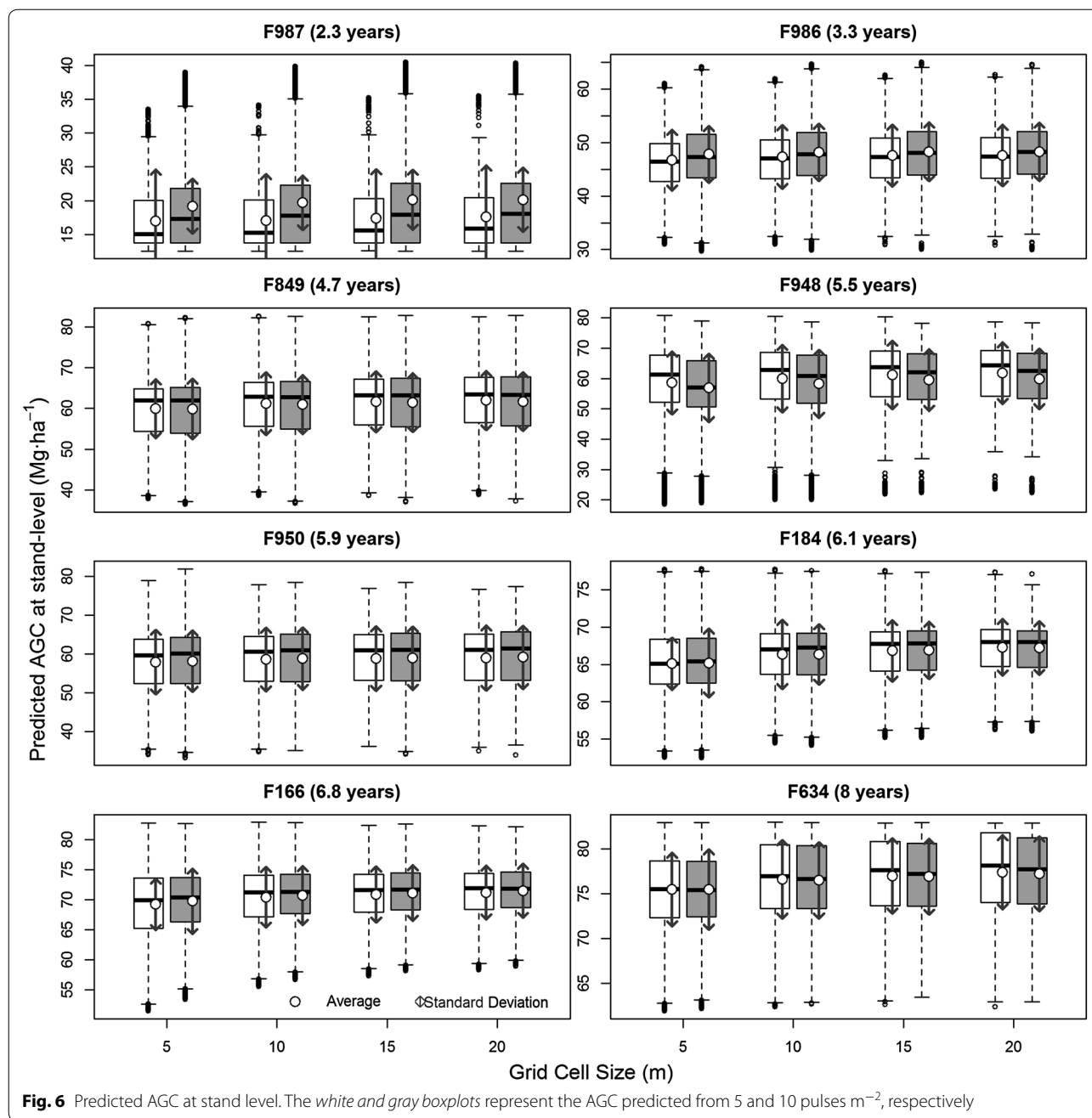
For each sample plot we extracted the predicted AGC from the maps. Equivalence plots indicated that the observed and predicted AGC were equivalent for the maps with all grid cell sizes (p values >0.05 for intercepts and p values <0.05 for slopes). RMSEs and Bias ranged from 11.01 to 12.3% and -0.37 to 0.87%, and showed

lower values in the maps with cell size of 5 m (Fig. 7). Although the observed and predicted AGC values were equivalent at spatial resolution of 5 m for both 5 and 10 pulses m<sup>-2</sup>, the maps are overestimating AGC in young stands (e.g. “F987”) and underestimating AGC in older stands (e.g. “F634”).

**Discussion**

Forest managers are seeking new management strategies that integrate forest, industry and market goals that maximize financial return while ensuring sustainability in the forest production chain [14]. The carbon credit market has created optimistic prospects for expansion of the Brazilian forest sector. Although this market seems attractive, there are few studies available with carbon stock estimates for planted forests [15]. In this study, we evaluated the combined effect of pulse density and grid size on AGC prediction at plot and stand levels in a fast-growing Eucalyptus plantation. A significant amount of work has been done to evaluate the effect of LiDAR pulse density for forest inventory modeling. However, less attention has been given to the influence of the cell size used to generate stand-level maps of forest attributes [43].

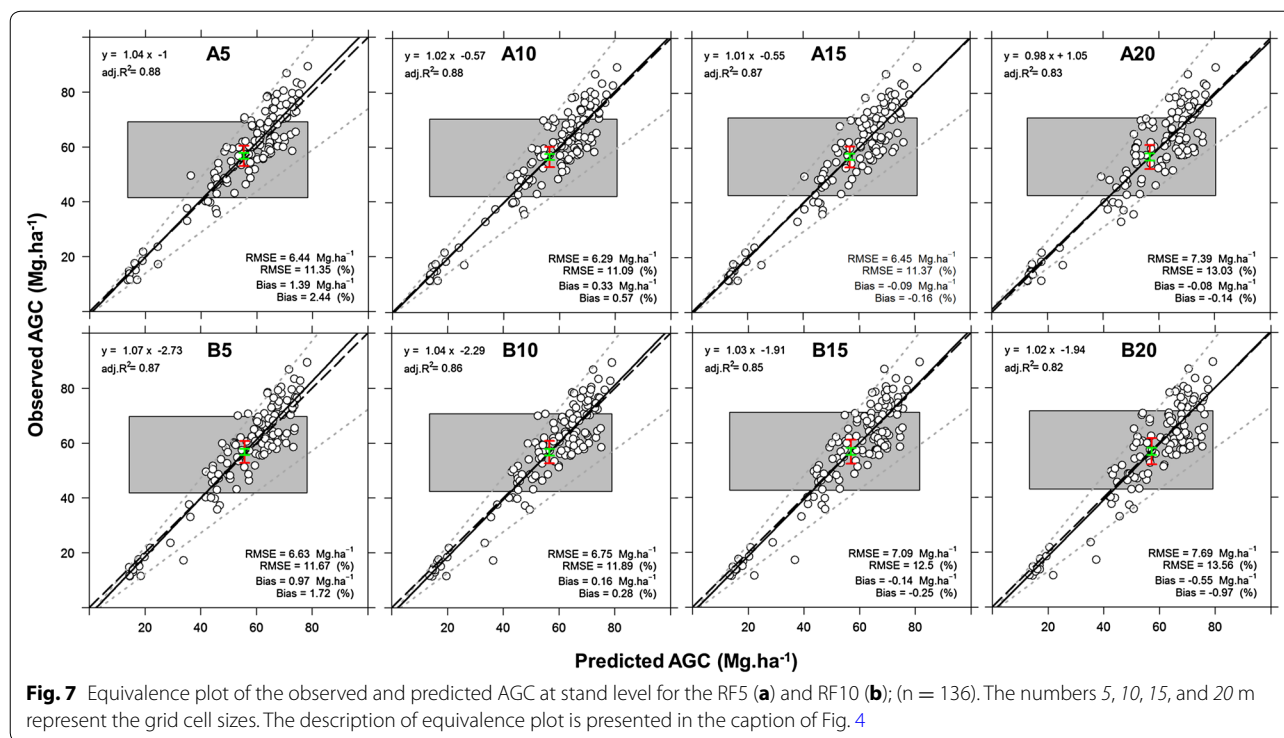




**Fig. 6** Predicted AGC at stand level. The white and gray boxplots represent the AGC predicted from 5 and 10 pulses m<sup>-2</sup>, respectively

Accurate and timely measurements of AGC are critical for understanding the structure and function of terrestrial ecosystems, as well as sustainable forest resources management and carbon accounting [29]. Eucalyptus plantations are a major carbon sink compared to others types of plantations, efficiently transferring atmospheric carbon into forest biomass and soil [10, 11]. The need to increase the accuracy and the spatial coverage of carbon accounting in these types of ecosystems has economic, politic and scientific rationales [31]. Airborne LiDAR can acquire

tree-dimensional information of all eucalyptus trees at a high spatial resolution in a short time [15], and whether there is benefit to using LiDAR for forest inventory depends in large part on characteristics specific to each forest [35]. The results presented in this current study provide a validation of plot and stand-based measurements of AGC. LiDAR has been shown to be a powerful technology for AGC prediction in Eucalyptus plantations, and our results have demonstrated that highly accurate estimates of AGC can be achieved using LiDAR data.



Our results show that a LiDAR pulse density of 5 pulses  $m^{-2}$  provides similar AGC prediction accuracy to that using a dataset with 10 pulses  $m^{-2}$  in fast-growing eucalyptus plantations. Previous studies have demonstrated similar effect of pulse density in forest attribute estimation. For example, [21] found that pulse densities could be reduced from 1.13 to 0.25 pulses  $m^{-2}$  with little effect on the quality of the forest inventory results in stands dominated by Norway spruce (*Picea abies* L. Karst.) and Scots pine (*Pinus sylvestris* L.). [35], whose objective was quantifying the effects of LiDAR pulse density and sample size on forest attributes prediction from LiDAR-derived metrics, found that model precision was more affected by sample plot sizes than by pulse density in mixed conifer forest in Washington state, USA. In a remnant forests in the rapidly urbanizing region of Charlotte, North Carolina, USA, LiDAR point density was responsible for a biomass prediction difference of 11.5% when the pulse density was down-scaled from 100 to 1% [58]. The statistical parameters adj. R<sup>2</sup>, RMSE and Bias presented in this study for the evaluation of AGC models are within the range reported by these other studies [33, 38, 47], though it is difficult to make a direct comparison due to differences in LiDAR sensors, pulse density testes, forest types and ground data collection.

LiDAR derived metrics, such as height percentiles, have been shown to be the most important LiDAR predictors of forest attributes in Eucalyptus plantations [15, 26, 47, 59]. In this study, the H99TH, H10TH, and HVAR were the most useful predictor variables for AGC prediction for the models derived from datasets of 5 and 10 pulses  $m^{-2}$ . These selected LiDAR metrics succinctly described the 3-dimensional canopy structure by capturing the majority of variation contained in the point cloud. In particular, H99TH captured canopy top height, and HVAR captured the canopy height variation of Eucalyptus stands. Also, the H99TH and HVAR were the most stable LiDAR-derived metrics when compared across pulse densities, accounting for a positive and strong linear correlation of  $r = 1$  (H99TH5  $\times$  H99TH10) and  $r = 0.93$  (HVAR5  $\times$  HVAR10).

The most accurate method of estimating AGC in Eucalyptus plantation is to physically sample it in the field. In a conventional inventory, one sample plot (300–500  $m^2$ ) per each 10 or 15 ha is normally measured with a goal of achieving a maximum acceptable RMSE of 10–15% [60]. However, this type of measurement over large areas is severely limited by cost and time. Approaches for deriving AGC information based on LiDAR data are of great utility and interest; in this study, we showed that airborne LiDAR with pulse densities of 5 pulses  $m^{-2}$  can be used



to predict AGC over large areas with RMSE and Bias equal to or less than in a conventional inventory. In this study, we achieved the RMSE goal of <15% established in the Methods section.

AGC maps are useful for forest management because they show the distribution of the AGC stock across the whole area. In theory, as spatial resolution increases, the higher is the probability to detect finer details and variability in forest stands. In this study, we found that AGC maps with spatial resolution ranging from 5 to 20 m show the same variability of AGC at stand level, indicating that in a fast-growing Eucalyptus plantation it is possible to map AGC in high detail even though the model was fitted using sample plots with larger sizes than the grid cell size used for mapping. However, LiDAR metrics computed in a grid cell of 5 m captured finer details of forest structure variability than those computed in a grid cell of 20 m; therefore, maps overestimated AGC in young stands and underestimated AGC in older stands, because the models calibrated from tree data collected within plots of 400 m<sup>2</sup> were not able to capture the forest variability at this fine resolution. The influence of the cell size used to generate forest attributes from LiDAR data, such as AGC, has not analyzed intensely at stand-level yet [43], and the range of examined cell sizes was limited [61, 62]. Whereas the differences among the cell sizes were almost negligible in [61], the effects of cell sizes on estimates was highlighted in [43], especially when comparing extreme cell sizes (2 and 100 m). These results concur with ours at resolutions between 5 and 20 m.

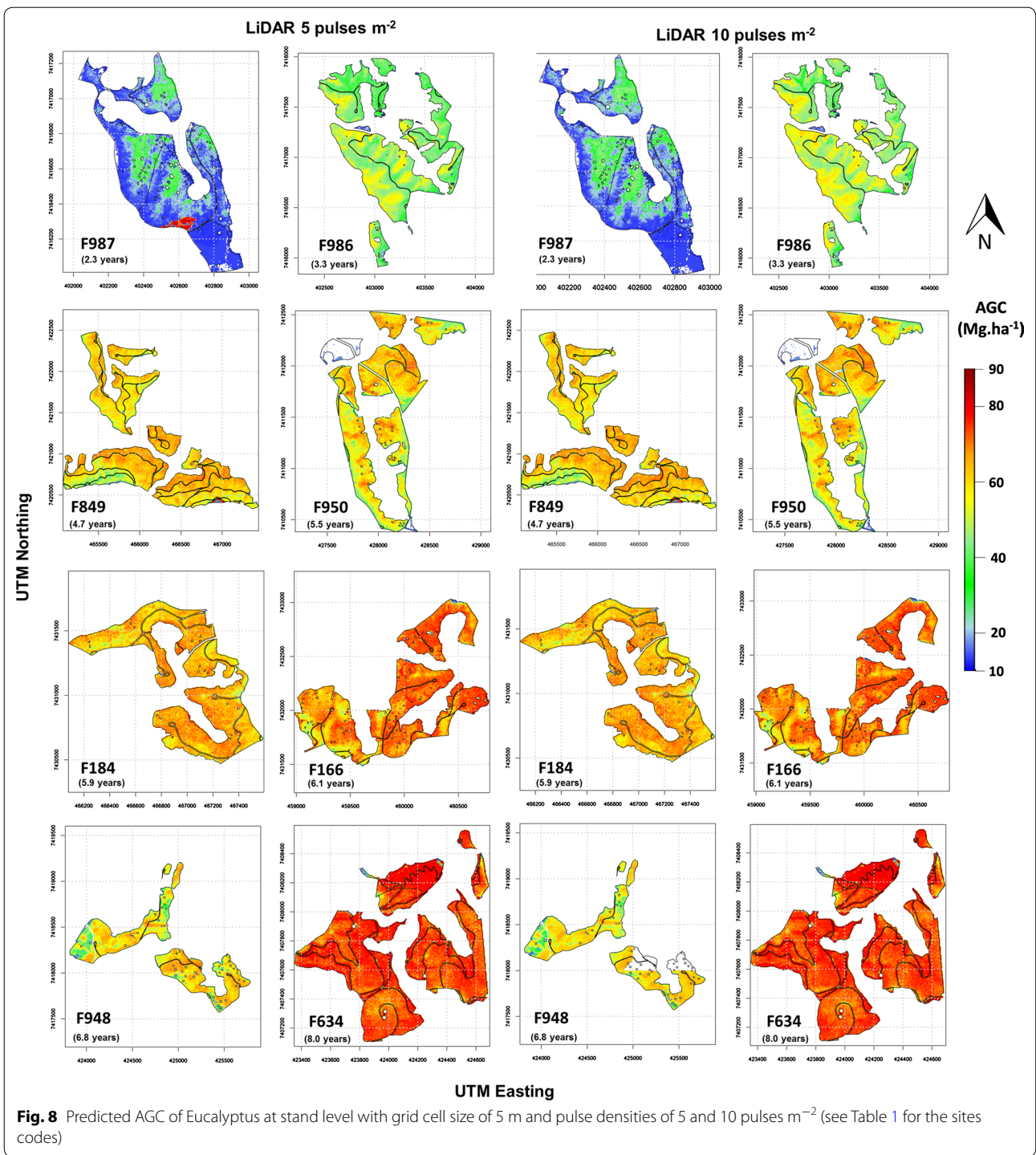
The maps with grid cell size of 5 m generated from 5 and 10 pulses m<sup>-2</sup> illustrated the broad similarities of AGC patterns across landscapes (Fig. 8). AGC stock was greatest in plantations with advanced ages (e.g. F634). AGC was more variable in the young plantations (such as F986 and F987). While the site F987 (Fig. 8a1, a2) showed relatively large variability in AGC, the map prediction was not affected by the LiDAR pulse density. Due to the large area and distance between the plantations, the LiDAR data at 5 and 10 pulses m<sup>-2</sup> at the same site were not collected on the same day. At the time of the first survey to collect LiDAR data with 5 pulses m<sup>-2</sup>, a small stand, not studied herein, 8 years old at site F987 was not completely harvested, as indicated by the red spot in Fig. 8a1. On the other hand, the red spot was not detected in Fig. 8a2, because 1 week later, at the time of the second survey to collect the LiDAR of 10 pulse m<sup>-2</sup>, the small stand was completely harvested.

Although the cost of LiDAR data acquisition was not a central objective that we evaluated in this study, it is

nonetheless an important factor to consider. As already mentioned, pulse density has a strong influence on the acquisition cost of LiDAR data, and even though the cost for using LiDAR with high or low pulse density might be lower than the cost of a conventional inventory for AGC in fast-growing Eucalyptus plantations, as presented in [32], it will still be highly expensive for a large area. Although field-based carbon estimations remain necessary for these purposes, integrating LiDAR remote sensing into carbon inventory schemes allows recovery of spatially-explicit AGC estimates across landscapes, while reducing the total costs and need for extensive field-based sampling [31]. In a conventional inventory, normally the variability of the forest attributes at stand level is not always known and therefore less studied [63]. In this study we mapped AGC at the landscape level with a spatial resolution of 5 m, so that it is now possible to compute and capture the variability of AGC at the stand level as well. Moreover, the estimates of AGC in the stands are helpful to determine how much carbon will be stored as forest products made from the timber after harvest.

## Conclusion

The effect of LiDAR pulse density and grid cell size on prediction accuracy of AGC at plot and stand level has been evaluated and analyzed in this paper. First, we found that LiDAR measurements can be used to predict AGC across variable-age Eucalyptus plantations with adequate levels of precision and accuracy using both 5 and 10 pulses m<sup>-2</sup>. Second, we found that H99TH, H10TH, and HVAR metrics were the most important LiDAR metrics for modeling AGC in this study. Third, we found that 5 pulses m<sup>-2</sup> and grid cell size of 5, 10, 15 and m did not affect the AGC predicted accuracy at plot and stand level. Finally, we demonstrated that the spatial distribution of AGC stocks can be precisely mapped at a grid cell size of 5 m using LiDAR pulse density of 5 pulses m<sup>-2</sup>, which can be used to provide key information for carbon sequestration in this type of ecosystem. Moreover, the promising results for AGC modeling in this study will allow for greater confidence in comparing AGC estimates with varying LiDAR sampling densities for Eucalyptus plantations and assist in decision making towards more cost effective and efficient forest inventory modeling. Although this study presents results for assessing AGC in fast-growing eucalyptus plantations in Brazil, the framework presented herein can serve as a useful methodology, and we hope that the promising results for AGC modeling in this study will stimulate



further research and applications not just in Eucalyptus forest plantation in southeast Brazil, but elsewhere as well.

**Abbreviations**

AGC: aboveground carbon; CO<sub>2</sub>: carbon dioxide; CNPq: National Council of Technological and Scientific Development; DBH: diameter at breast height; DTM: digital terrain model; FAPESP: State of São Paulo Research Foundation;

RF: random forest; RF5: random forest model of 5 pulses m<sup>-2</sup>; RF10: random forest model of 10 pulses m<sup>-2</sup>; GPS: global positioning system; Ht: tree height; LOOCV5: leave-one out cross validation 5 pulses m<sup>-2</sup>; LOOCV10: leave-one out cross validation 10 pulses m<sup>-2</sup>; PVE: percentage of the variance explained; RMSE: root mean squared errors; USFS: US Forest Service; 3D: tree-dimensional.

**Authors' contributions**

All the authors have made substantial contribution towards the successful completion of this manuscript. They all have been involved in designing the

study, drafting the manuscript and engaging in critical discussion. CAS, CK and ATH, contributed with the methodological framework, data processing analysis and write up. LAV, CG, LCER and AD contributed to the interpretation, quality control and revisions of the manuscript. SPC contributed with the field data survey. All authors read and approved the final manuscript.

#### Author details

<sup>1</sup> Department of Natural Resources and Society, College of Natural Resources, University of Idaho, (UI), 875 Perimeter Drive, Moscow, ID 83843, USA. <sup>2</sup> US Forest Service (USDA), Rocky Mountain Research Station, RMRS, 1221 South Main Street, Moscow, ID 83843, USA. <sup>3</sup> Department of Forest Engineering, Oregon State University, 269 Peavy Hall, Corvallis, OR 97331, USA. <sup>4</sup> College of Forestry, Federal University of Mato Grosso, Av. Fernando Correa da Costa, 2367, Boa Esperança, Cuiabá, MT 78060-900, Brazil. <sup>5</sup> Department of Forest Sciences, College of Agriculture Luiz de Queiroz (ESALQ), University of Sao Paulo (USP), Av. Pádua Dias, 11, Piracicaba, SP 13418-900, Brazil. <sup>6</sup> Tecnosylva, Parque Tecnológico de León, 24009 León, Spain.

#### Acknowledgements

The authors are very grateful for the LiDAR and field inventory data collections funded by Fibria S.A., a pulp and paper company. We thank Ben Bright and Jan Eitel for their helpful suggestions on an earlier draft of the manuscript.

#### Competing interests

The authors declare that they have no competing interests.

#### Availability of data and materials

Data will not be shared to due commercial licenses.

#### Funding

This research was funded through a Ph.D. scholarship from the National Council of Technological and Scientific Development - CNPq via the Science Without Borders Program (Process 249802/2013-9) and by the State of São Paulo Research Foundation-FAPESP (Process 12/03176-0).

#### Publisher's Note

Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Received: 8 December 2016 Accepted: 26 May 2017

Published online: 07 June 2017

#### References

- IPCC. Climate change 2013: The physical science basis. contribution of working group I to the Fifth assessment report of the intergovernmental panel on climate change. Cambridge, UK, and New York, USA. <http://www.ipcc.ch/report/ar5/wg1/>. Accessed 10 Jan 2016.
- IPCC. Climate change 2001: The scientific basis; Cambridge: Cambridge University Press; 2016. [http://www.grida.no/climate/ipcc\\_tar/wg1/pdf/wg1\\_tar-front.pdf](http://www.grida.no/climate/ipcc_tar/wg1/pdf/wg1_tar-front.pdf). Accessed 10 Jan 2016.
- Binkley CS, Apps MJ, Dixon RK, Kauppi P, Nilsson LO. Sequestering carbon in natural forests. *Crit Rev Environ Sci Technol*. 1998;27:23–45.
- Fan S, Gloor M, Mahlman J, Pacala S, Sarmiento JL, Takahashi T, Tans P. The North American Sink. *Science*. 1815;1999:283. doi:10.1126/science.283.5409.1813q.
- Houghton RA. Interannual variability in the global carbon cycle. *J Geophys Res*. 2000;105:20121–30.
- Gálvez FB, Hudak AT, Byrne JC, Crookston NL, Keefe RF. Using climate-FVS to project landscape-level forest carbon stores for 100 years from field and LiDAR measures of initial conditions. *Carbon Balance Manage*. 2014;9:1–22. doi:10.1186/1750-0680-9-1.
- García M, Riaño D, Chuvieco E, Danson FM. Estimating biomass carbon stocks for a Mediterranean forest in central Spain using LiDAR height and intensity data. *Remote Sens Environ*. 2010;114:816–30. doi:10.1016/j.rse.2009.11.021.
- Bellassen V, Luysaert S. Carbon sequestration: managing forests in uncertain times. *Nature*. 2014;506:153–5.
- Stape JL, Binkley D, Ryan MG. Production and carbon allocation in a clonal Eucalyptus plantation with water and nutrient manipulations. *For Ecol Manag*. 2008;255:920–30. doi:10.1016/j.foreco.2007.09.085.
- Du H, Zeng F, Peng W, Wang K, Zhang H, Liu L, Song T. Carbon Storage in a Eucalyptus Plantation chronosequence in Southern China. *Forests*. 2015;6:1763–78. doi:10.3390/f6061763.
- Booth TH. Eucalypt plantations and climate change. *For Ecol Manag*. 2013;301:28–34. doi:10.1016/j.foreco.2012.04.004.
- Abá. Brazilian tree industry. 2015. [http://www.iba.org/images/shared/iba\\_2015.pdf](http://www.iba.org/images/shared/iba_2015.pdf).
- Gonçalves JLDM, Alvares CA, Higa AR, Silva LD, Alfenas AC, Stahl J, Ferraz SFDB, Lima WDP, Brancalion PHS, Hubner A, Bouillet JPD, Laclau JP, Nouvellon Y, Epron D. Integrating genetic and silvicultural strategies to minimize abiotic and biotic constraints in Brazilian eucalypt plantations. *For Ecol Manag*. 2013;301:6–27. doi:10.1016/j.foreco.2012.12.030.
- Silva CA, Klauber C, Pádua SCD, Piccolo M, Rodriguez LCE. Estoque de carbono na biomassa aérea florestal em plantações comerciais de Eucalyptus spp. *Sci For*. 2015;43:135–46.
- Silva CA, Klauber C, Carvalho SPC, Hudak A, Rodriguez LCE. Mapping aboveground carbon stocks using LiDAR data in *Eucalyptus* spp. plantations in the state of São Paulo, Brazil. *Sci For*. 2014;42:591–604.
- Laurin VG, Chen Q, Lindsell JA, Coomes DA, Del Frate F, Guerriero L, Pirotti F, Valentini R. Above ground biomass estimation in an African tropical forest with lidar and hyperspectral data. *ISPRS J Photogramm Remote Sens*. 2014;89:49–58. doi:10.1016/j.isprsjprs.2014.01.001.
- Hudak AT, Evans JS, Smith AMS. LiDAR utility for natural resource managers. *Remote Sens*. 2009;1:934–51. doi:10.3390/rs1040934.
- Næsset E. Determination of mean tree height of forest stands using airborne laser scanner data. *ISPRS J Photogramm Remote Sens*. 1997;52:49–56. doi:10.1016/S0924-2716(97)83000-6.
- Næsset E, Bjerknes KO. Estimating tree heights and number of stems in young forest stands using airborne laser scanner data. *Remote Sens Environ*. 2001;78:328–40. doi:10.1016/S0034-4257(01)00228-0.
- Næsset E, Økland T. Estimating tree height and tree crown properties using airborne scanning laser in a boreal nature reserve. *Remote Sens Environ*. 2002;79:105–15. doi:10.1016/S0034-4257(01)00243-7.
- Gobakken T, Næsset E. Weibull and percentile models for lidar-based estimation of basal area distribution. *Scand J For Res*. 2005;20:490–502. doi:10.1080/02827580500373186.
- Hudak AT, Crookston NL, Evans JS, Falkowski MJ, Smith AMS, Gessler P. Regression modeling and mapping of coniferous forest basal area and tree density from discrete-return LiDAR and multispectral satellite data. *Can J Remote Sens*. 2006;32:126–38. doi:10.5589/m06-007.
- Nelson R, Krabill W, Tonelli J. Estimating forest biomass and volume using airborne laser data. *Remote Sens Environ*. 1988;24:247–67. doi:10.1016/0034-4257(88)90028-4.
- Oderwald R, Popescu S. A simplified method of predicting percent volume in log portions. *South J Appl*. 2003;27:149–52.
- Holmgren J. Prediction of tree height, basal area and stem volume in forest stands using airborne laser scanning. *Scand J For Res*. 2004;19:543–53. doi:10.1080/02827580410019472.
- Tesfamichael SG, Van Aardt JAN, Ahmed F. Estimating plot-level tree height and volume of Eucalyptus grandis plantations using small-footprint, discrete return lidar data. *Prog Phys Geogr*. 2010;34:515–40. doi:10.1177/0309133310365596.
- Næsset E. Estimation of above- and below-ground biomass in boreal forest ecosystems. *Int Arc Photogramme Remote Sens Spat Inf Sci*. 2004;36:145–8.
- Næsset E, Gobakken T. Estimation of above- and below-ground biomass across regions of the boreal forest zone using airborne laser. *Remote Sens Environ*. 2008;112:3079–90.
- Hudak AT, Strand EK, Vierling LA, Byrne JC, Eitel JUH, Martinuzzi S, Falkowski MJ. Quantifying aboveground forest carbon pools and fluxes from repeat LiDAR surveys. *Remote Sens Environ*. 2012;123:25–40. doi:10.1016/j.rse.2012.02.023.
- Mascaro J, Detto M, Asner GP, Muller-Landau HC. Evaluating uncertainty in mapping forest carbon with airborne LiDAR. *Remote Sens Environ*. 2011;115:3770–4. doi:10.1016/j.rse.2011.07.019.

31. Patenaude G, Hill R, Milne R, Gaveau DLA, Briggs BBJ, Dawson TP. Quantifying forest above ground carbon content using LiDAR remote sensing. *Remote Sens Environ*. 2004;93:368–80. doi:10.1016/j.rse.2004.07.016.
32. Hummel S, Hudak AT, Uebler EH, Falkowski MJ, Megown KA. A Comparison of accuracy and cost of LiDAR versus stand exam data for landscape management on the Malheur National Forest. *J For*. 2011;109:267–73.
33. Watt MS, Adams T, Aracil SG, Marshall H, Watt P. The influence of LiDAR pulse density and plot size on the accuracy of New Zealand plantation stand volume equations. *N Z J For Sci*. 2013;43:1–10. doi:10.1186/1179-5395-43-15.
34. Magnussen S, Næsset E, Gobakken T. Reliability of LiDAR derived predictors of forest inventory attributes: a case study with Norway spruce. *Remote Sens Environ*. 2010;114:700–12.
35. Magnussen S, Boudewyn P. Derivations of stand heights from airborne laser scanner data with canopy-based quantile estimators. *Can J Remote Sens*. 1998;28:1016–31.
36. Ruiz L, Hermosilla T, Mauro F, Godino M. Analysis of the influence of plot size and LiDAR density on forest structure attribute estimates. *Forests*. 2014;5:936–51.
37. Leitold V, Keller M, Morton DC, Cook BD, Shimabukuro YE. Airborne lidar-based estimates of tropical forest structure in complex terrain: opportunities and trade-offs for REDD+. *Carbon Balance Manag*. 2015; doi:10.1186/s13021-015-0013-x.
38. Watt MS, Meredith A, Watt P, Gunn A. The influence of LiDAR pulse density on the precision of inventory metrics in young unthinned Douglas-fir stands during initial and subsequent LiDAR acquisitions. *N Z J For Sci*. 2014;44:18. doi:10.1186/s40490-014-0018-3.
39. Strunk J, Temesgen H, Andersen HE, Flewelling JP, Madsen L. Effects of LiDAR pulse density and sample size on a model-assisted approach to estimate forest inventory variables. *Can J Remote Sens*. 2012;38:644–54.
40. Treitz P, Lim K, Woods M, Pitt D, Nesbitt D, Etheridge D. LiDAR sampling density for forest resource inventories in Ontario. *Can J Remote Sens*. 2012;4:830–48. doi:10.3390/rs4040830.
41. Koppen W, Geiger R. *Klimate der Erde*. Gotha: Verlag Justus Perthes. Wall-map 150cm x 200cm. 1928.
42. White JC, Wulder MA, Varhola A, Vastaranta M, Coops N, Cook BD, Pitt D, Woods M. A best practices guide for generating forest inventory attributes from airborne laser scanning data using an area-based approach. *Nat For Chron*. 2013;89:5.
43. Hayashi R, Weiskittel A, Kershaw JA. Influence of prediction cell size on LiDAR-derived area-based estimates of total volume in mixed species and multi-cohort forests in north eastern North America. *Can J Remote Sens*. 2016;41:473–88.
44. Mcgaughey RJ. *FUSION/LDV: Software for LiDAR data analysis and visualization*. 3rd ed. USDA, Forest Service Pacific Northwest Research Station, Seattle; 2015.
45. Kraus K, Pfeifer N. Determination of terrain models in wooded areas with airborne laser scanner data. *ISPRS J Photogramm Remote Sens*. 1998;53:193–203.
46. Kraus K, Mikhail EM. Linear least squares interpolation. *Photogramm Eng*. 1972;38:1016–29.
47. Silva CA, Klauber C, Hudak TA, Vierling LA, Carvalho SP, Rodriguez LC. A principal component approach for predicting the stem volume in Eucalyptus plantations in Brazil using airborne LiDAR data. *Forestry*. 2016;1:412. doi:10.1093/forestry/cpw016.
48. Breiman L. Random forests. *Mach Learn*. 2001;45(1):5–32.
49. Liaw A, Wiener M. Classification and regression by random-forest. *R News*. 2002;2:18–22.
50. R Development Core Team. R: A language and environment for statistical computing. Vienna: R Foundation for Statistical Computing Access: <http://www.Rproject.org>. 2015.
51. Evans JS, Cushman SA. Gradient modeling of conifer species using random forest. *Landscape Ecol*. 2009;24:673–83.
52. Evans JS, Murphy MA, Holden ZA, Cushman SA. Modeling species distribution and change using Random Forests. In: Drew CA, Huettmann F, Wiersma Y, editors. *Predictive Modeling in Landscape Ecology*. New York: Springer; 2011. p. 139–59.
53. Robinson AP, Froese RE. Model validation using equivalence tests. *Ecol Mod*. 2004;176:349–58.
54. Robinson AP, Duursma RA, Marshall JD. A regression-based equivalence test for model validation: shifting the burden of proof. *Tree Physiol*. 2005;25:903–13.
55. Smith AMS, Falkowski MJ, Hudak AT, Evans JS, Robinson AP, Steele CM. A cross-comparison of field, spectral, and lidar estimates of forest canopy cover. *Can J Remote Sens*. 2009;35(5):447–59.
56. Robinson, A. *Equivalence: Provides Tests and Graphics for Assessing Tests of Equivalence*, Version 0.7.2. <https://cran.r-project.org/web/packages/equivalence/> (accessed on 20 January, 2016).
57. Silva CA, Hudak A, Vierling LA, Loudermilk L, O'Brien JJ, Hiers J, Jack J, Gonzalez-Benecke CA, Lee H, alkowski MJ, Khosravipour A. Imputation of individual longleaf pine (Mill.) Tree attributes from field and LiDAR Data. *Can J Remote Sens*. 2016;42:554–73.
58. Singh KK, Chen G, McCarter JB, Meentemeyer RK. Effects of LiDAR point density and landscape context on estimates of urban forest biomass. *ISPRS J Photogramm Remote Sens*. 2015;101:310–22. doi:10.1016/j.isprsjprs.2014.12.021.
59. Packalen P, Maltamo M, Mehtatalo L. ALS-based estimation of plot volume and site index in a Eucalyptus plantation with a nonlinear mixed-effect model that accounts for the clone effect. *Ann For Sci*. 2011;2011(68):1085–92.
60. Batista JLF, Couto HTZ, Silva Filho DF. *Quantificação de Recursos Florestais: árvores, Arvoredos e Florestas*. 1st edn. Oficina de Textos; 2014.
61. Gonzalez P, Asner GP, Battles JJ, Lefsky MA, Waring KM, Palace M. Forest carbon densities and uncertainties from Lidar, QuickBird, and field measurements in California. *Remote Sens Environ*. 2010;114:1561–75.
62. Brubaker KM, Johnson SE, Brinks J, Leites LP. Estimating canopy height of deciduous forests at a regional scale with leaf-off, low point density LiDAR. *Can J Remote Sens*. 2014;40:123–34.
63. Wulder MA, White JC, Bater CW, Coops NC, Hokinson CH, Chen G. Lidar plots a new large-area data collection option: context, concepts, and case study. *Can J Remote Sens*. 2012;38:600–18.

Submit your manuscript to a SpringerOpen® journal and benefit from:

- Convenient online submission
- Rigorous peer review
- Immediate publication on acceptance
- Open access: articles freely available online
- High visibility within the field
- Retaining the copyright to your article

---

Submit your next manuscript at ► [springeropen.com](http://springeropen.com)

---