ESTIMATING FOREST ATTRIBUTES IN INDUSTRIAL *Pinus taeda* L. FOREST PLANTATIONS IN BRAZIL USING SIMULATED NASA'S GEDI SPACEBORNE LIDAR DATA

Carlos Alberto Silva[1,2], *Laura Duncanson*[1,2], *Steven Hancock*[2]; *Carine Klauberg*[3], *Andrew T. Hudak*[4], *Ralph Dubayah*[2]

[1]Biosciences Laboratory, NASA Goddard Space Flight Center, Greenbelt, MD 20707, USA, carlos_engflorestal@outlook.com ; [2]Department of Geographical Sciences, University of Maryland, College Park, Maryland, MD 20740, USA; lauraiduncanson@gmail.com; hancock@umd.edu; dubayah@umd.edu. [3]Federal University of São João Del Rei – UFSJ, Sete Lagoas, MG, Brazil - 35701-970; carine_klauberg@hotmail.com and [4]US Forest Service (USDA), Rocky Mountain Research Station, RMRS, 1221 South Main Street, Moscow, ID, USA – 83843, ahudak@fs.fed.us

ABSTRACT

Remote sensing technologies can dramatically increase the efficiency of plantation management by reducing or replacing time-consuming field sampling. In this study, we evaluated the capability of the NASA's Global Ecosystem Dynamic Investigation (GEDI) spaceborne lidar system for estimating forest attributes at footprint level in industrial *Pinus teada* L. forest plantations in Southern Brazil. In the field, 100 field plots were measured and top canopy height (HMAX; m) and timber volume (V; m^3/ha) were computed. GEDI-derived metrics were simulated using airborne lidar (ALS) data. We used multiple linear regression for modeling HMAX and V from GEDI-like metrics, and we found that models defined as a function of only three GEDI-like metrics (RH98: canopy height at 98 percentiles of energy, COV: canopy cover; FHD: foliage height diversity) had a very strong and unbiased predictive power. The promising results presented herein show that GEDI, during its lifetime time of two years, may provide an appropriate technology to assist forest managers towards more cost effective and efficient forest inventory in industrial pine forest plantations.

Keywords — spaceborne lidar, forest attributes, stand modeling, pine plantations.

1. INTRODUCTION

NASA's Global Ecosystem Dynamic Investigation (GEDI) will produce the first high resolution laser ranging observations of 3-D forest structure while orbiting the Earth aboard the Japanese Experimental Module's Exposed Facility on the International Space Station (ISS). GEDI is a three-lasers, full waveform recording lidar and will operate for a 2-year minimum period, producing ~12 billion cloud-free land surface observations. Two of the lasers are full power, and one is split into two beams, producing a total of four beams. GEDI was lunched on December 5th, 2018 and is currently at the ISS (Figure 1a-b) in the process of installation and tests. In this study, we evaluated the use of GEDI spaceborne lidar technology for estimating top canopy height (HMAX) and timber volume (V) at the footprint level in *Pinus taeda* L. forest plantations in Southern Brazil. Our justification is that improving management practices in industrial forest plantations may increase production efficiencies, thereby reducing pressures on native tropical forests for meeting global wood and pulp needs.

2. MATERIAL AND METHODS

2.1 Study area

The study area consisted of *P. taeda* stands located within the Telêmaco Borba municipality in the state of Paraná,

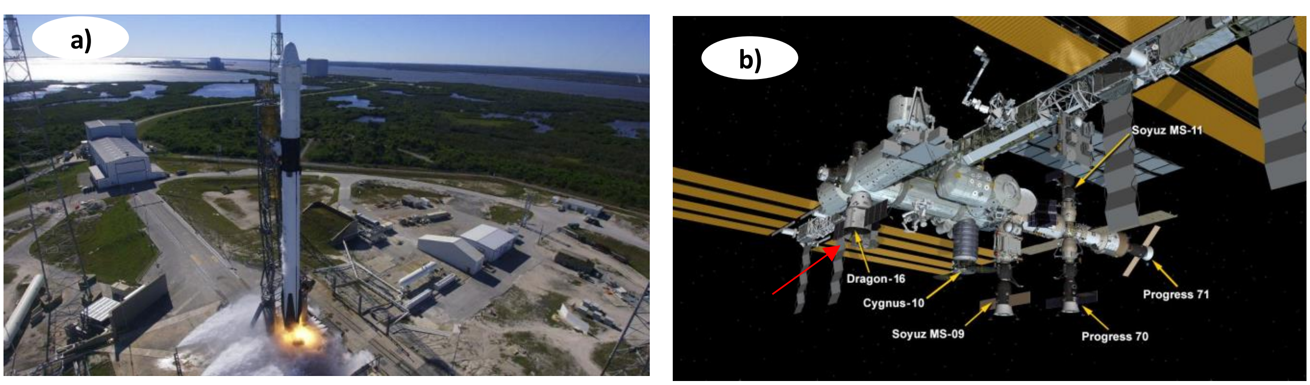

Figure 1. GEDI ready for launch at 1:15pm (EST, USA) at Kennedy Space Center, December 5th, 2019 a); SpaceX's Dragon cargo spacecraft installed on the Earth-facing side of the ISS b). Source (https://twitter.com/gedi_knights)

southern Brazil. Trees were planted using a 3.0 × 2.0 m or 2.5 × 2.5 m grid configuration, resulting in an average tree density of 1667 or 2000 trees ha^{-1}, respectively. The plantations are managed by Klabin S.A., a pulp and paper company.

2.2 Field Data

A total of 100 rectangular field plots, each approximately 600 m^2, were randomly established and measured across *P. taeda* stands. As such, the field plots represent well the study area, and they capture the entire structural variability within stands with ages ranging from 3 to 35 years old. In each sample plot, individual trees were measured for dbh (diameter at breast height) at 1.30 m and a random subsample (15%) of trees for tree height (Ht; m). For those trees in the plots that were not directly measured for Ht, the inventory team of Klabin S.A. estimated their heights using hypsometric equations. The HMAX was defined as the maximum tree height within a plot area and the V (m^3/ha) as the sum of all individual timber volume within the plot, and then scaled to a hectare. The V for each individual tree was computed using a fifth-degree polynomial equation [2]

2.3 GEDI simulated data

In this study, GEDI-like full waveform data were simulated using ALS data via the GEDI simulator tool described by Hancock et al [3] (Figure 2 and 4). GEDI coverage beams (which are half the laser energy of the power beams) acquired during the day have higher errors than power beam or coverage beams acquired at night. After simulating the GEDI waveforms, a suite of canopy metrics was computed based on the cumulative waveform energy (i.e., 10%, 25%, 50%, 75%, 98%, and 100%; RH10, RH25, R50, RH75, RH98, RH100; COV: canopy cover (%); FHD – foliage height diversity).

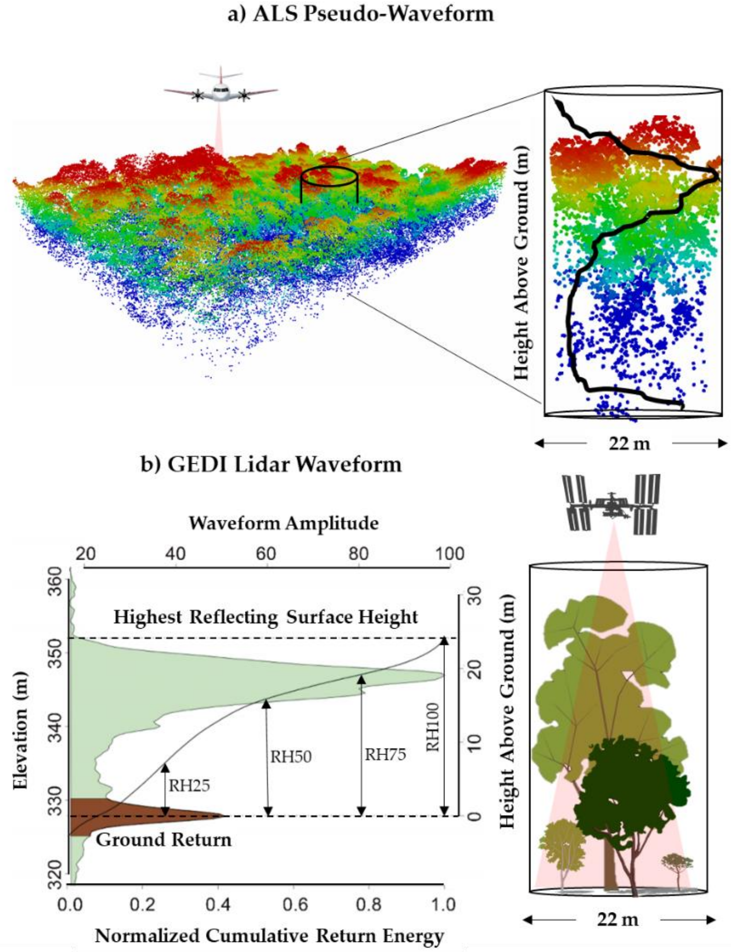

Figure 2. ALS-derived pseudo-waveform (vertical black line) and (b) GEDI-like waveform and canopy metrics.

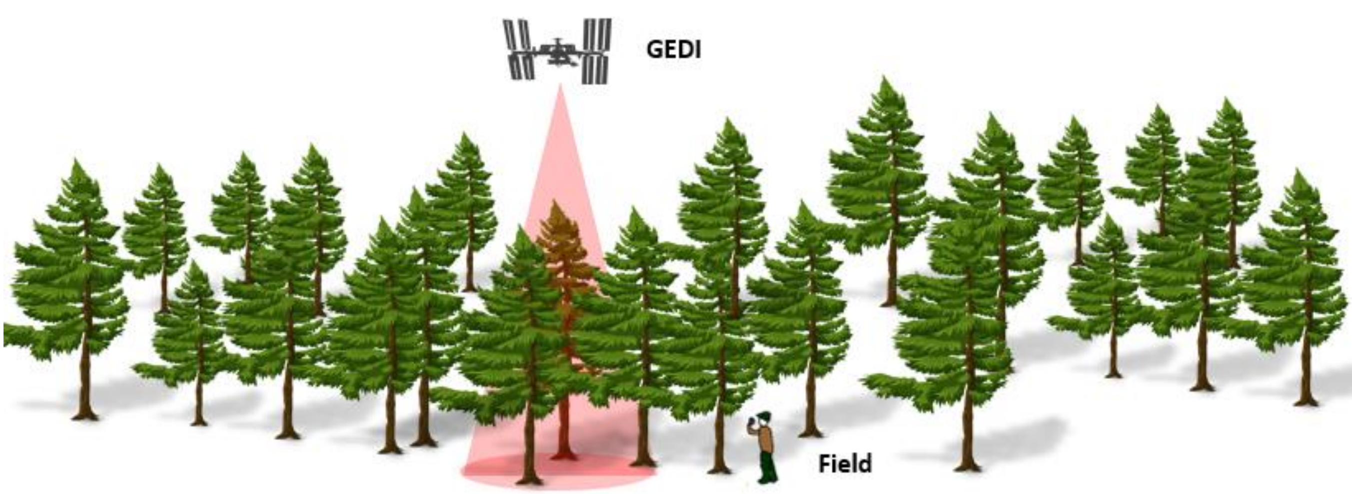

Figure 3. Illustration of GEDI and field data collection in pine forest plantation

2.4 Forest attribute modeling

We used the *lm* linear model function in R statistical software [4] to train multiple linear regression models for estimating HMAX and V from GEDI-like metrics. Shapiro – Wilk [5] and Breusch–Pagan [6] tests were used to evaluate the normality and heteroscedasticity of the models. The precision and accuracy of estimates for each model were evaluated in terms of adjusted coefficient of determination (adj.R^2), absolute and relative root mean square error (RMSE) and bias (eqs.1 and 2). Also, the performance of the models for estimating HMAX and V were validated using a leave-one-out cross-validation (LOOCV) strategy.

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

$$\mathrm{Bias} = \frac{1}{n}\sum_{i=1}^{n}(\hat{y}_i - y_i) \quad \text{(eq.2)}$$

where n is the number of observations, y_i is the reference value for observation i, and $\hat{y}_i$ is the estimated (HAMX or V) value for observation i. The relative RMSE (%) and bias (%) were calculated by dividing the respective absolute values (eqs. 1 and 2) by the mean of observed values.

3. RESULTS

Table 1 and Figure 4 show the performance of the multivariate linear regression models created to estimate HMAX and V of *Pinus teada* based on the GEDI-like metrics. Models defined as a function of only three GEDI-like metrics (RH98, COV and FHD) for estimating HMAX and V had a very strong and unbiased predictive power, with overall model adj. $R^2 > 0.95$ and RMSE (%) < 15 % for both training and LOOCV derived models. Also, results from the Shapiro–Wilk test and Breusch –Pagan test reveal that the data were normally distributed, and heteroscedasticity had a 0.05 per cent level of significance.

Estimates of HMAX and V were well balanced overall, being slightly overpredicted during early and advanced stand ages and slightly underpredicted at intermediate ages (Figure 4). These differences may reflect varying site indices and management practices across the plantations.

4. DISCUSSION

Traditional forest inventory approaches are not effective in terms of cost, especially in *Pinus taeda* forest plantations, where there is a need to monitor annual forest growth and properties over large areas. Based on our results herein and a 2-year lifetime, GEDI should provide an important technology to enhance conventional forest inventory and support operational timber procurement

Table 1. GEDI-derived models for predicting canopy top height (HMAX) and timber volume (V) in pine plantations.

Response Variables	Models	Adj.R^2	RMSE		Bias	
			relative	absolute	relative	absolute
HMAX	=-4.98 + 1.02RH98 -3.03COV + 1.75FHD	0.97	9.04%	2.25m	$6.02\cdot10^{-15}$%	$1.50\cdot10^{-15}$m
V	=477+ 30.8RH98 -122.23COV – 141.40FHD	0.97	11.55%	52.38m^3/ha	$-6.05\cdot10^{-15}$%	$-2.75\cdot10^{-14}m^3$/ha

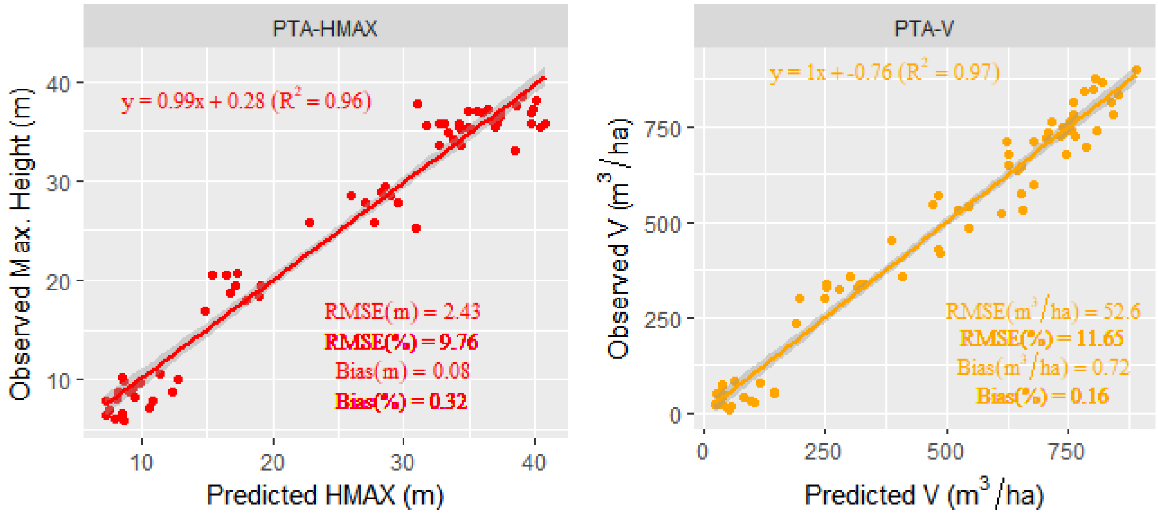

Figure 4. Leave-one out cross validation for GEDI-derived models for predicting HMAX (left) and V (right).

planning, and hence could be of high interest to forestry organizations.

Herein, the GEDI-derived models produced equivalent results compared to previous studies using ALS in *Pinus taeda* plantations in Brazil [2]. However, as GEDI will collect data using eight linear tracks of lidar shots with a ~22m footprint and 60m post spacing, the fusion of GEDI data with wall-to-wall remote sensing data, such as multispectral imagery from Sentinel-2A and Landsat 8-OLI, will be required to estimate forest attributes and map them at the stand and landscape levels.

5. CONCLUSIONS

In this study, we demonstrated that three GEDI-like metrics and multiple linear regression can be used for estimating canopy top height and timber volume at the footprint level in *Pinus taeda* forest plantations in Southeast Brazil. We hope that these promising results for forest attribute modeling will stimulate further research and applications of GEDI not just in pine plantations in Brazil but other forest plantations elsewhere.

6. REFERENCES

[1] GEDI 2018. Available in <https://gedi.umd.edu/mission/technology/> Accessed in October 2018

[2] Silva, C.A.; Klauberg, C.; Hudak, A.T.; Vierling, L.A.; Jaafar, W.S.W.M.; Mohan, M.; Garcia, M.; Ferraz, A.; Cardil, A.; Saatchi, S. Predicting Stem Total and Assortment Volumes in an Industrial Pinus taeda L. Forest Plantation Using Airborne Laser Scanning Data and Random Forest. Forests 2017, 8, 254.

[3] Hancock, S., Armston, J., Tang, H., Duncanson, L., Hofton, M., Blair, J.B., Disney, M., Marselis, S., Kellner, J.R., Fatoyinbo, L., Dubayah, R. (2017). Bridging the gap between spaceborne, airborne and terrestrial lidar: The GEDI simulator. Silvilaser 2017, Blacksburg, Virginia, USA

[4] R Development Core Team. 2018 R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing. http://www.Rproject.org. (accessed on 20 April, 2018).

[5] Shapiro, S.S. and Wilk, M.B. 1965 An analysis of variance test for normality (complete samples). Biometrika 52, 591–611.

[6] Breusch, T.S. and Pagan, A.R. 1979 A simple test for heteroscedasticity and random coefficient variation. Econometrica 47, 1287–1294.