Applied Vegetation Science **18** (2015) 724–732

Estimating species richness and biomass of tropical dry forests using LIDAR during leaf-on and leaf-off canopy conditions

Jose Luis Hernández-Stefanoni, Kristofer D. Johnson, Bruce D. Cook, Juan Manuel Dupuy, Richard Birdsey, Alicia Peduzzi & Fernando Tun-Dzul

Keywords
Above-ground biomass; Canopy conditions; Forest structure, Species richness; LIDAR; Topography; Tropical dry forest

Abbreviations
AGB = above-ground biomass; LIDAR = light detection and ranging; REDD+ = Reducing Emissions from Deforestation and forest Degradation plus enhancing forest carbon stocks; TDF = Tropical dry forests; RMSE = root mean square error.

Received 19 January 2015
Accepted 30 June 2015
Co-ordinating Editor: Sarah Goslee

Hernández-Stefanoni, J.L.
(Corresponding author, jl_stefanoni@cicy.mx)[1],
Johnson, K.D. (kristoferdjohnson@fs.fed.us)[2],
Cook, B.D. (bruce.cook@nasa.gov)[3],
Dupuy, J.M. (jmdupuy@cicy.mx)[1],
Birdsey, R. (rbirdsey@fs.fed.us)[2],
Peduzzi, A. (apeduzzi@fs.fed.us)[2],
Tun-Dzul, F. (fjtun@cicy.mx)[1]

[1]Centro de Investigación Científica de Yucatán A.C, Unidad de Recursos Naturales, Calle 43 # 130, Colonia Chuburná de Hidalgo, C.P. 97200 Mérida, Yucatán, México;
[2]U.S. Forest Service, Northern Research Station, Newtown Square, PA 19073, USA;
[3]NASA Goddard Space Flight Center, Greenbelt, MD 20771, USA

Abstract

Questions: Is the accuracy of predictions of above-ground biomass (AGB) and plant species richness of tropical dry forests from LIDAR data compromised during leaf-off canopy period, when most of the vegetation is leafless, compared to the leaf-on period? How does topographic position affect prediction accuracy of AGB for leaf-off and leaf-on canopy conditions?

Location: Tropical dry forest, Yucatan Peninsula, Mexico.

Methods: We evaluated the accuracy of predictions using both leaf-on and leaf-off LIDAR estimates of biomass and species richness, and assessed the adequacy of both LIDAR data sets for characterizing these vegetation attributes in tropical dry forests using multiple regression analysis and ANOVA. The performance of the models was assessed by leave-one-out cross-validation. We also investigated differences in vegetation structure between two topographic conditions using PCA and ANOSIM. Finally, we evaluated the influence of topography on the accuracy of biomass estimates from LIDAR using multiple regression analysis and ANOVA.

Results: A higher overall accuracy was obtained with leaf-on vs leaf-off conditions for AGB (root mean square error (RMSE) = 21.6 vs 25.7 ton·ha^{-1}), as well as for species richness (RMSE = 5.5 vs 5.8 species, respectively). However, no significant differences in mean dissimilarities between biomass estimates from LIDAR and *in situ* biomass estimates comparing the two canopy conditions were found ($F_{1,39} = 0.03$, $P = 0.87$). In addition, no significant differences in dissimilarities of AGB estimation were found between flat and hilly areas ($F_{1,39} = 1.36$, $P = 0.25$).

Conclusions: Our results suggest that estimates of species richness and AGB from LIDAR are not significantly influenced by canopy conditions or slope, indicating that both leaf-on and leaf-off models are appropriate for these variables regardless of topographic position in these tropical dry forests.

Introduction

Information about the spatial distribution of above-ground biomass (AGB) and species richness is critical to obtain accurate estimates of carbon stocks and biodiversity at broad scales. Such estimates are essential for deforestation reduction strategies such as a REDD+ (Reducing Emissions from Deforestation and forest Degradation plus enhancing forest carbon stocks), and to design effective strategies for selecting natural protected areas. Reliable maps of AGB and species richness at the landscape level are useful tools for many different management objectives. LIDAR (Light Detection and Ranging) offers an advantage over most other remote sensing technologies, thanks to its ability to penetrate tropical forest canopies and detect three-dimensional forest structure. Therefore, it is a useful tool for map-

Applied Vegetation Science
Doi: 10.1111/avsc.12190

ping vegetation structure variables, such as height, forest volume and biomass, since, unlike optical images, LIDAR is free from saturation at high biomass values (Næsset & Okland 2002; Drake et al. 2003). Yet, only a few studies have estimated plant species richness with LIDAR (Leutner et al. 2012; Simonson et al. 2012; Hernández-Stefanoni et al. 2014).

Despite the benefits of using LIDAR to estimate AGB and species richness, there is a concern about the accuracy of estimations when using data from different acquisition periods. Gaveau & Hill (2003) reported an under-estimation of tree height under leaf-off conditions for boreal forest. This was due to a deeper penetration of laser pulses into the tree crowns compared to leaf-on conditions. Additionally, leaf-off conditions result in a larger number of laser pulses being reflected from the ground and understorey vegetation, and a lower probability that some of them reach the actual canopy top (Hill & Broughton 2009). However, LIDAR data from leaf-on canopy conditions are not always available for large geographic regions, including those where tropical dry forests are found, since leaf-on conditions coincide with high cloud cover associated with the rainy season. LIDAR data can be more readily obtained during leaf-off canopy conditions, when cloud-free atmospheric conditions are common. For example, NASA has recently acquired long strips of LIDAR during the dry season of 2013 throughout Mexico (see www.gliht.gsfc.nasa.gov). Therefore, there is a need to assess the accuracy of LIDAR data sets in estimating forest structure parameters for both canopy conditions from tropical dry forests (TDF). These forests are especially important because they are the most extensive land-cover type in the tropics and one of the most diverse and least protected terrestrial ecosystems (Miles et al. 2006). Over half of the TDF are found in the Americas, and Mexico harbours 38% of the TDF in the continent. In particular, the Yucatan Peninsula has one of the largest areas in Mexico with this type of forest (Portillo-Quintero & Sánchez-Azofeifa 2010).

Topography and soil properties can affect vegetation attributes of tropical forests such as height, basal area, stem density, species richness and species composition (Clark & Clark 2000). These environmental variables also influence soil water availability, the main factor limiting plant growth and survival in TDF (Oliveira-Filho et al. 1998). For example, Dupuy et al. (2012a) found a decrease in tree height and basal area, as well as an increase in stem density and number of species on hills compared with flat areas in a TDF of the Yucatan Peninsula. Stem density in turn affects the percentage of LIDAR returns reflected by the canopy, thereby influencing LIDAR-derived estimates of tree height, an important vegetation attribute related to AGB (Chave et al. 2005). Specifically, under-estimation of tree height (and hence AGB) is expected in areas with low stem density where openings in the canopy may be more frequent (Brubaker et al. 2014).

Slope can also influence tree height (and AGB) estimates from LIDAR data, because tree height is calculated by subtracting the estimated ground elevation value from the elevation value of the top vegetation layer of a canopy, and ground elevation estimates have higher error on steep slopes than in flat areas (Brubaker et al. 2014). Consequently, a lower accuracy of LIDAR-derived elevation, and hence also tree height estimates, is expected for areas with steeper slopes (Hodgson et al. 2005). Indeed, higher tree height error estimates have been found for terrains with steep slopes compared to those on gentle slopes and flat areas (Van Leeuwen & Nieuwenhuis 2010; Brubaker et al. 2014).

The objectives of this research were twofold: (1) to evaluate the accuracy of predictions using leaf-on vs leaf-off LIDAR estimates of AGB and species richness derived from plot measurements, in order to assess whether both LIDAR data sets are adequate for characterizing these vegetation attributes of a TDF; and (2) to determine if topographic condition affects the accuracy of LIDAR estimates of AGB for leaf-off and leaf-on canopy conditions in a hilly landscape. We predicted that: (1) LIDAR estimates of AGB and species richness would be better for leaf-on than for leaf-off canopy conditions, and that (2) LIDAR estimates of AGB would be better in flat areas compared to hills, especially during leaf-off conditions.

Methods

Study area

The study area is located within a private reserve in the centre of the Yucatan Peninsula (89°32′–89°34′ W, 20°04′–20°06′ N; Fig. 1). The vegetation is seasonally dry tropical forest and the climate is tropical warm, with summer rain and a dry season from Nov to Apr. The mean annual temperature is about 26 °C, and mean annual precipitation ranges between 1000 and 1100 mm. The landscape consists of low (60–160 m a.s.l.) Cenozoic limestone hills with moderate slope, alternating with flat areas (Olmsted et al. 1999). The study area is dominated by seasonally dry semi-deciduous tropical forests (50–75% of species drop their leaves during the dry season) of different ages of abandonment after traditional slash-and-burn agriculture, ranging from 3 to >100 yr, although the forest in the reserve is predominantly 50–100 yr old, and the youngest stands are 17 yr old. The forest has a relatively low canopy stature (9–13 m) with a few prominent trees attaining 15–18 m.

Doi: 10.1111/avsc.12190

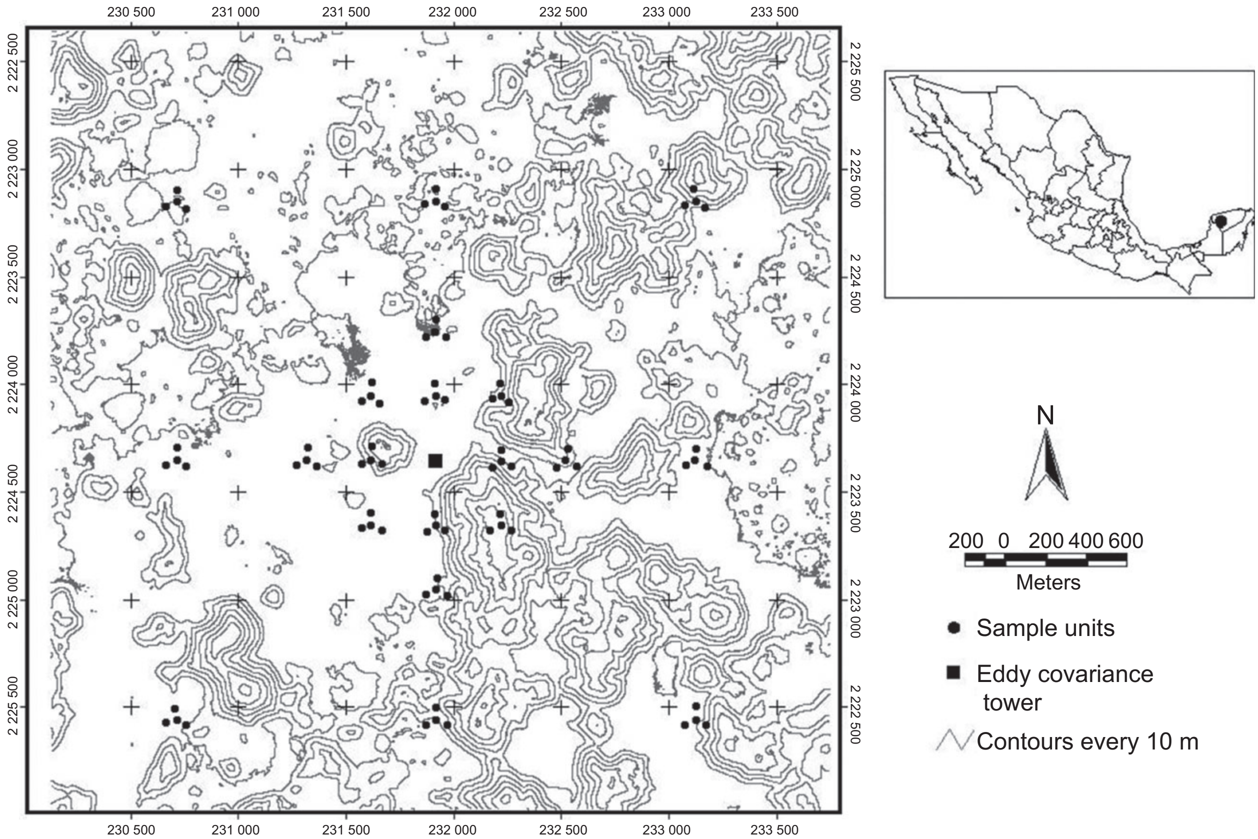

Fig. 1. Location of the study site and field samples.

Field data

Field data were recorded from a systematic plant survey conducted during the rainy season of 2013 in an area of 9 km^2. Twenty clusters of plots were located systematically around an eddy covariance flux tower: 12 within the central 1 km^2, and the rest in the remaining 8 km^2 (Fig. 1). The cluster plot design was based on Mexico's National Forest and Soil Inventory field data layout (Comisión Nacional Forestal 2012). Each cluster of plots consists of three circular 400-m^2 plots (11.28-m radius) at fixed distances (38.6 m) and azimuths (0, 120 and 240°) from a central 1000-m^2 plot (17.84-m radius) within a circular area of 1 ha. All plots were located on the ground with a Garmin GPS unit. The precision of x and y coordinates of centre plots was estimated from different measurements of the position, and the mean location errors were <3 m. In each plot, all woody plants >7.5 cm DBH (at 1.3 m) were sampled, and the diameter of all stems and the height of all individuals measured. Species richness, *i.e.* the number of woody plant species per sample site was used as a measure of local or α species diversity for the cluster of four plots (2200 m^2). AGB was calculated for the 1000-m^2 plots from tree diameter (and height) using three allometric equations developed for tropical forests of Mexico: one for trees ≥10 cm DBH (Cairns et al. 2003; modified by Urquiza-Haas et al. 2007), another for trees <10 cm DBH (Hughes et al. 1999) and the other for lianas (Chave et al. 2003). These are the best plot designs for improving prediction accuracy of species richness and biomass for this type of forest (Hernández-Stefanoni et al. 2014).

Data collection and processing of LIDAR

The airborne LIDAR data were acquired during two different periods: Aug 2012 under leaf-on canopy conditions and Apr 2013 under leaf-off conditions. Both data sets were acquired using the same LIDAR sensor, an airborne laser scanner RIEGL-QV-480, with similar sensor settings and flight parameters (Table 1). The leaf-on data set was obtained by a private contractor, CartoData (http:/www.cartodata.com/), whereas the leaf-off data set was obtained by the G-LiHT airborne imager of NASA (Cook et al. 2013). Although differences in sensor settings and flight parameters can affect the spatial and vertical distribution of laser return points (Magnussen et al. 2010), and hence the accuracy of predicted forest structure parameters such as a tree height, basal area, volume and biomass (Goodwin

Table 1. Leaf-off and leaf-on LIDAR acquisition specifications.

	Leaf-on LIDAR	Leaf-off LIDAR
Flight Dates	Aug 2012	Apr 2013
Sensor	RIEGL VQ480	RIEGL VQ480
Scan Angle (±degrees of nadir)	15	30
Average Flight Height (m)	396.2	335.0
Pulse Rate (Khz)	200	300
Swath Overlap (%)	50	60
Beam Divergence (mrad)	0.3	0.3
Pulse Density (m^{-2})	5	6
Returns per Pulse	5	7

et al. 2006), in this study, pulse density was very similar between the two LIDAR data sets. Moreover, Jakubowski et al. (2013) concluded that when pulse density is >1 m^{-2} the accuracy of forest structure metrics remains relatively high. Since both LIDAR data sets complied with this requirement, no thinning of the leaf-on data set was performed, and both data sets were used at their acquired pulse densities.

The Fusion 'clipdata' routine (McGaughey 2012) was used to extract all returns from the boundaries of the plots in both leaf-on and leaf-off data sets. Using the *x y* coordinates and the radius for each field plot, the point clouds were clipped to correspond with the area of plots (400 and 1000 m^2). Before applying the clipping process, the data were normalized to the ground surface, in order to express the returns in terms of heights above the ground instead of elevation above sea level. We used a 1-m^2 resolution digital terrain model (DTM) from leaf-on canopy conditions to normalize both data sets. We chose the leaf-on DTM because it had a lower RMSE (0.51) than the leaf-off DTM (0.66) when compared with nine samples of high precision GPS measures on the ground provided by CartoData. This is probably because the leaf-on DTM had a lower scan angle, which is known to improve ground detection under dense foliage (Lovell et al. 2005). Next, a set of 62 LIDAR metrics was calculated within each of the four plots (three 400-m^2 plots and one 1000-m^2 plot) within a cluster, both for leaf-on and for leaf-off data sets. Additionally, for species richness analyses, the mean and SD values for a cluster were calculated from metrics of the four plots.

Two general categories of LIDAR metrics were used as the predictor variables in the models for estimating the spatial distribution of species richness and biomass for both leaf-on and leaf-off canopy conditions. The first category was based on height statistics and included mean, maximum and minimum elevation, the variability of return heights (variance, coefficient of variation), statistics to quantify location (percentiles 1, 5, 10… 100 and L-moments), among others. The second category included canopy density metrics and was used to evaluate the amount of vegetation cover. A threshold of 1.5 m was used as a minimum height above the ground to reduce noise within the near-ground cloud of returns caused by low vegetation and uneven ground surface, and a canopy threshold height of 4.0 m was used to compute LIDAR canopy density metrics. For a detailed description of the LIDAR metrics see McGaughey (2012).

Data analysis

Multiple regression analysis was used to model the relationship between the response variables (AGB and species richness) and explanatory variables (LIDAR metrics for the 1000-m^2 plot for AGB and the mean and SD values of these metrics calculated using the cluster of four plots for species richness). The independent variables were formally tested for normality and homoscedasticity, while the response variables were transformed as needed with $1/x$, log10 (x), log10 ($x + 1$) and sqrt (x) to meet linearity assumptions (Zar 1999). None of the LIDAR metrics (predictors) required transformation once the distributions of response variables were normalized. All multiple regression analyses were carried out using forward selection. Since multicollinearity between predictor variables can cause problems in multivariable modelling, we verified that the explanatory variables considered for the analysis were either uncorrelated or expressed only small collinearity (Zar 1999).

The performance of the different models was assessed by leave-one-out cross-validation. In this procedure, one observation is sequentially removed from the data set, and the remaining sampling plots are used to fit the model. Then, coefficients obtained are applied to this datum in order to produce a predicted value. The cross-validation yields a list of estimated values of AGB and species richness paired to those obtained from the observed sampling plots. Predicted values were also back-transformed to original values as needed and corrected for bias introduced during the back-transformation process using a method suggested by Miller (1984). The predicted and observed values of AGB and species richness were compared using the coefficient of determination (R^2) and the root mean square error (RMSE).

We also investigated the differences in vegetation structure (abundance, number of species, basal area, plant cover, average and canopy height) between two topographic conditions (flat areas with a slope <10% and hills with a slope >10%) using PCA and analysis of similarities (ANOSIM; Clarke 1993). All six vegetation structure variables were grouped, standardized and ordinated using PCA to identify the major gradients and patterns of variation in vegetation structure among sampling sites. We further used ANOSIM to test for significant differences in vegetation structure between hills and flat areas. This procedure is a non-parametric analogue of multivariate ANOVA that tests variability within and between groups (hills and flat areas), but is free from parametric assumptions of multivariate normality and homogeneity of variances. PCA and ANOSIM analysis were performed using v6 of the PRIMER-E statistical package (Clarke & Gorley 2005).

Finally, to assess the effects of topography (hills vs flat areas) and canopy condition (leaf-on vs leaf-off) on AGB estimates from LIDAR, mean differences were calculated between biomass estimates obtained after applying leave-one-out cross validation and *in situ* biomass estimates for each sampled plot, and compared using two-factor (topography and canopy condition) ANOVA. This analysis was not performed for species richness because this variable was assessed for the whole cluster of four plots (2200 m^2), and most clusters contained a mixture of topographic conditions. Therefore, differences between *in situ* counts of species richness and LIDAR-derived estimates from cross-validation comparing leaf-on and leaf-off conditions were assessed using a paired *t*-test.

Results

The fitted regression models to estimate AGB and species richness explained 88% and 84% of variability for AGB using leaf-on and leaf-off LIDAR data, respectively, and 67% and 59% for species richness using leaf-on and leaf-off LIDAR data, respectively. AGB was explained by point density measures and height metrics for the 1000-m^2 sampled area, whereas species richness was mainly explained by the SD of the LIDAR metrics in the four plots of 2200 m^2 (Appendix S1).

We found a strong correspondence between observed values of AGB and those predicted from LIDAR metrics (Fig. 2), although leaf-on estimates were slightly more accurate (R^2 = 0.84; RMSE = 21.6 Mg·ha^{-1}) than leaf-off estimates (R^2 = 0.77; RMSE = 25.7 Mg·ha^{-1}).

Species richness showed a similar pattern to AGB, but a moderate association with LIDAR data (Fig. 3). The R^2 values increased from 0.42 to 0.49, respectively, for leaf-off and leaf-on LIDAR data sets, whereas the RMSE values decreased from 5.8 for leaf-off to 5.5 for leaf-on conditions.

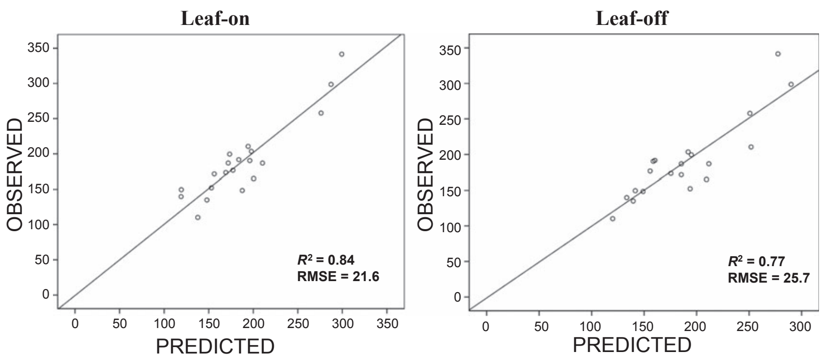

Fig. 2. Results of cross-validation analyses used to compare the performance of observed and predicted values of above-ground biomass from LiDAR using leaf-on and leaf-off canopy conditions.

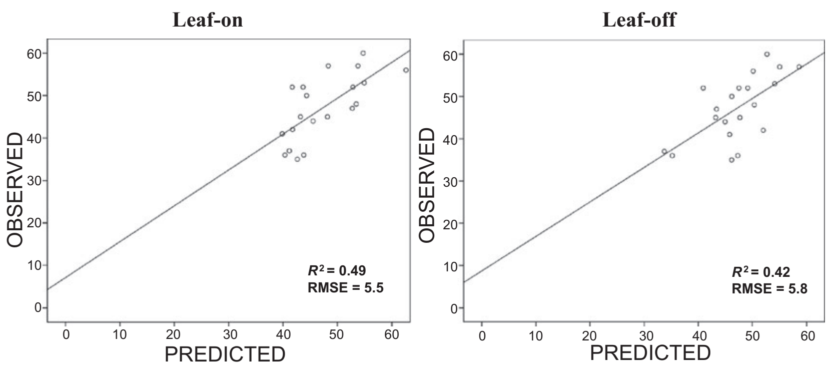

Fig. 3. Results of cross-validation analyses used to compare the performance of observed and predicted values of species richness from LiDAR using leaf-on and leaf-off canopy conditions.

Applied Vegetation Science
Doi: 10.1111/avsc.12190

The ANOSIM revealed significant differences in vegetation structure between hills and flat areas (R = 0.194, $P = 0.017$). The first two components of the PCA accounted for 64.2% of the variation in vegetation structure (PC1: 39.6%; PC2: 22.8%) and showed that hills tended to have higher abundance, plant cover and species richness than flat areas (Appendix S2).

We found no significant differences between the residuals of LIDAR estimates and *in situ* biomass estimates comparing leaf-on and leaf-off canopy conditions ($F_{1,39} = 0.03$, $P = 0.87$), nor comparing flat areas and hills ($F_{1,39} = 1.36$, $P = 0.25$; Table 2). Similarly, no significant differences were found in species richness LIDAR estimations between leaf-on and leaf-off canopy conditions ($t = -0.21$, $df = 19$, $P = 0.84$).

Discussion

As predicted, we found slightly better estimates of AGB and species richness from LIDAR data (as indicated by higher R^2 values and lower RMSE values in the cross-validation procedure) for leaf-on compared to leaf-off canopy conditions. This result was expected, considering that a smaller proportion of the laser pulses would be reflected from the upper canopy during leaf-off conditions due to a decrease in leaf biomass (Gaveau & Hill 2003). In this study, leaf-off models of AGB and species richness had higher errors in plots, with RMSE values above 4.1 Mg·ha^{-1} and 0.3 species, respectively. However, the differences in mean dissimilarities between LIDAR-derived estimates of AGB or species richness and *in situ* estimates were not significant, indicating that the accuracy of predictions of these variables for leaf-off canopy conditions in this TDF are comparable with the accuracy of estimates using leaf-on LIDAR data sets, as also found in previous studies on AGB in mixed temperate forests (Naesset 2005; Wasser et al. 2013). This can be attributed to laser pulse returns reflecting from a proportion of coniferous trees in mixed forests, or to the fact that some deciduous species in both mixed and TDF reflect a sufficient number of returns even during leaf-off canopy conditions (Naesset 2005). Thus, our analysis shows that a leaf-off model may substitute a leaf-on model with satisfactory results, especially for AGB, for which the leaf-off accuracy ($R^2 = 0.77$) was comparable to other regional tropical forest models ($R^2 = 0.56$–0.80; Asner et al. 2012). This result is relevant for LIDAR-derived estimates of forest structure in seasonally dry tropical forests, where it is difficult to acquire LIDAR data during leaf-on conditions, since they coincide with the rainy season when most of the forest area is covered with clouds.

Table 2. Differences between LIDAR estimates and *in situ* biomass estimates comparing two topographic positions and two canopy conditions. No significant differences were found.

	Mean Differences (Mg·ha^{-1})	SE
Leaf-off, <10% Slope	−6.26	6.12
Leaf-off, >10% Slope	6.40	9.41
Leaf-on, <10% Slope	−1.48	8.62
Leaf-on, >10% Slope	4.08	5.91

Species richness and AGB were strongly related to different categories of LIDAR data. Specifically, species richness was mainly related to the SD of LIDAR elevation and canopy density metrics, whereas AGB was mainly related to canopy density and elevation metrics. TDF have a relatively simple forest structure, characterized by lower stature and fewer canopy strata than their humid counterparts, likely reflecting harsher environmental conditions, especially drought (Murphy & Lugo 1986). Moreover, resprouting is common in TDF, especially when subjected to recurring disturbance, resulting in dominance of a few species with high resprouting capacity and a consequent decoupling of forest structure and species richness (McLaren & McDonald 2003; Dupuy et al. 2012a). Thus, although LIDAR is able to capture the 3-D structure of forests, the relatively simple structure of the sampled forest stand, dominated by a few species with high resprouting capacity (Dupuy et al. 2012a), may not adequately reflect variation in species richness. On the other hand, species richness seems to be promoted by habitat heterogeneity, as indicated by the positive coefficient of the SD of LIDAR metrics. In other words, higher variation of canopy and sub-canopy LIDAR returns reflects higher variability of tree height (possibly linked to forest succession), topography and forest floor conditions that provide opportunities for different species to be present within the sample area. This result coincides with several studies aimed at predicting species richness from different remotely sensed data (Gillespie 2005; Rocchini 2007; Hernández-Stefanoni et al. 2012), which showed that species richness is related to the structural and compositional complexity of habitats. This highlights the importance of modelling ecological heterogeneity to estimate species diversity over space, as well as the importance of remotely sensing signals that measure this heterogeneity (Rocchini et al. 2015). However, the accuracy of prediction of species richness using LIDAR data in this study was admittedly moderate ($R^2 = 0.42$–0.49). Higher prediction accuracy of species richness (close to 80%) has been obtained in TDF using image texture from high-resolution satellite images, possibly because a high spatial resolution allows analysis of the internal spatial variation within a plot (Gallardo-Cruz et al. 2012). Accuracy of species richness estimates may thus be improved by combining the internal spatial variation within a plot using optical sensors with high spatial resolution, and habitat heterogeneity in vegetation structure among plots using LIDAR-derived canopy density and elevation metrics.

Doi: 10.1111/avsc.12190

In contrast to species richness, AGB was mainly related to canopy density and elevation metrics, indicating that an increase in the horizontal and vertical distribution of plant AGB is strongly correlated with a higher density and elevation of LIDAR returns. The ability to predict AGB is thus based on the structural attributes of vegetation, i.e. canopy cover, tree diameter, basal area and tree height, which are indicators of forest structural complexity at the stand scale (McElhinny et al. 2005). In turn, forest structural complexity is related to stand biomass, since it fosters different biomass levels (Eaton & Lawrence 2009); for example, an increase in vertical canopy structure and basal area in TDF is related to an increase in biomass across forest succession (Dupuy et al. 2012a,b).

The PCA and ANOSIM analyses revealed a clear effect of topographic condition (flat areas with a slope <10% vs hills with a slope >10%) on forest structure in this TDF landscape. This result concurs with those of Dupuy et al. (2012a) for the same study area, where canopy height and basal area were found to be smaller on hills than in flat areas, whereas stem density and number of species showed the opposite trend. This general pattern of decreasing tree height and increasing stem density with a corresponding increase in slope has been found in different ecosystems, ranging from tropical to coniferous forests (Clark & Clark 2000; Saremi et al. 2014).

Because tree height is calculated from LIDAR data by subtracting the estimated ground elevation value from the elevation value of the top forest canopy layer, and ground elevation estimates have higher error on steep slopes than in flat areas (Brubaker et al. 2014), we predicted that LIDAR estimates of AGB would be better in flat areas compared to hills. This prediction was not supported since we found no significant differences in the average error between LIDAR-derived AGB estimates and field estimates comparing the two topographic positions. However, the average error between LIDAR AGB estimates and field estimates tended to be negative in areas with gentle slopes (suggesting a slight under-estimation of AGB in these conditions) and positive (suggesting a slight over-estimation) in areas with more pronounced slopes (Table 2). Since stem density directly affects the percentage of LIDAR returns reflected by the forest canopy (Gaveau & Hill 2003), it is reasonable to expect that variation in stem density would affect LIDAR-derived height, and hence AGB estimates. The slight under-estimation of biomass in gentle slopes may thus be due to an under-estimation of tree heights, indicating that the tree tops were relatively less likely to be hit by LIDAR beams under these conditions where stem density is comparatively low. Conversely, in areas with more pronounced slopes where stem density is comparatively higher, one would expect a slight over-estimation of biomass. Moreover, the downhill portion of tree crowns and the surface area of the downhill part of the crown are larger than their uphill counterparts (Breidenbach et al. 2008). Therefore, more reflections from laser beams are expected from the downhill part of trees, resulting in a slight over-estimation of tree height, and consequently of AGB, in areas with pronounced slopes. Although our results were not significant, they show a trend that is consistent with an over-estimation of tree height in areas with high slope, as reported in previous studies (Takahashi et al. 2005; Breidenbach et al. 2008). The fact that our study region consists of relatively small (up to 180 m a.s.l.), gentle hills with moderate slope (20–55%) can help explain why the error differences we found between gentle and more pronounced slopes were not significant.

Conclusions

We developed models to predict AGB and species richness from LIDAR metrics. The accuracy of predictions for AGB was comparable to other tropical forest models, whereas the accuracy of predictions for species richness was only moderate compared to tropical forest models from high spatial resolution imagery data. Future work should therefore endeavour to combine high spatial resolution optical and LIDAR data sets to improve the accuracy of species richness estimates in this type of forest. We also showed that variation in species richness is mainly explained by habitat heterogeneity, gauged by the SD of LIDAR metrics, whereas variation in AGB was related to vegetation structure – represented by canopy density and elevation metrics.

The LIDAR-derived estimates of AGB and species richness did not differ between leaf-on and leaf-off canopy conditions, indicating that the accuracy of predictions of these variables for leaf-off canopy conditions (the best time for data acquisition due to the prevalence of cloud-free conditions) in this TDF are comparable with the accuracy of estimations using leaf-on LIDAR data sets. This result may be applicable to other seasonally dry tropical forests.

Although estimates of AGB from LIDAR were not significantly influenced by slope for either canopy condition (possibly because our study area consists of low hills with moderate slopes), they showed a trend that is consistent with the previous studies documenting an over-estimation of tree height in areas of higher slope. Since tree height is an important parameter for biomass estimation, more attention should be paid to the influence of topography on tree elevation and hence AGB estimates from LIDAR data.

Acknowledgements

We are grateful to Kaxil Kiuic A. C. for logistic support, and Filogonio May-Pat and Manuela Tamayo for assistance with fieldwork and data processing. We also thank two anonymous reviewers and the editor, whose constructive and insightful comments helped improve the clarity and impact of this paper. The study was financially supported by the Reinforcing REDD+ and South-South Cooperation Project, funded by the Mexican National Forest Commission (CONAFOR), United States Forest Service (USFS) and USAID. This study is part of the Mexican network of Intensive Carbon Monitoring Sites (MEX-SMIC), an affiliation of academic institutions, government and NGOs with various funding sources. The network currently includes five forested sites located in strategic landscapes for REDD+ activities in Mexico.

References

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

Breidenbach, J., Koch, B., Kändler, G. & Kleusberg, A. 2008. Quantifying the influence of slope, aspect, crown shape and stem density on the estimation of tree height at plot level using LIDAR and InSAR data. *International Journal of Remote Sensing* 29: 1511–1536.

Brubaker, K.M., Johnson, S.E., Brinks, J. & Leites, L.P. 2014. Estimating canopy height of deciduous forests at a regional scale with leaf-off, low point density LIDAR. *Canadian Journal of Remote Sensing* 40: 123–134.

Cairns, M.A., Olmsted, I., Granados, J. & Argaez, J. 2003. Composition and aboveground tree biomass of a dry semi-evergreen forest on Mexico's Yucatan Peninsula. *Forest Ecology and Management* 186: 125–132.

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

Chave, J., Andalo, C., Brown, S., Cairns, M.A., Chambers, J.Q., Eamus, D., Folster, H., Fromard, F., Higuchi, N. (. . .) &, Yakamura, T. 2005. Tree allometry and improved estimation of carbon stocks and balance in tropical forests. *Oecologia* 145: 87–99.

Clark, D.B. & Clark, D.A. 2000. Landscape-scale variation in forest structure and biomass in a tropical rain forest. *Forest Ecology and Management* 137: 185–198.

Clarke, K.R. 1993. Non-parametric multivariate analyses of changes in community structure. *Australian Journal of Ecology* 18: 117–143.

Clarke, K.R. & Gorley, R.N. 2005. *Multivariate ecological research (PRIMER). v. 6.1*. PRIMER-E Ltd, Plymouth, UK.

Comisión Nacional Forestal. 2012. *Manuales Para el Levantamiento de Datos en Campo del Inventario Nacional Forestal y de Suelos*. Comision Nacional Forestal, Zapopan, Jal., Mexico.

Cook, B.D., Corp, L.A., Nelson, R.F., Middleton, E.M., Morton, D.C., McCorkel, J.T., Masek, J.G., Ranson, K.J., Ly, V. & Montesano, P.M. 2013. NASA Goddard's LIDAR, Hyperspectral and Thermal (G-LiHT) airborne imager. *Remote Sensing* 5: 4045–4066.

Drake, J.B., Knox, R.G., Dubayah, R.O., Clark, D.B., Condit, R., Blair, J.B. & Hofton, M. 2003. Above-ground biomass estimation in closed canopy Neotropical forests using LIDAR remote sensing: factors affecting the generality of relationships. *Global Ecology and Biogeography* 12: 147–159.

Dupuy, J.M., Hernández-Stefanoni, J.L., Hernández-Juarez, R.A., Tetetla-Rangel, E., López-Martínez, J.O., Leyequién-Abarca, E., Tun-Dzul, F.J. & May-Pat, F. 2012a. Patterns and correlates of tropical dry forest structure and composition in a highly replicated chronosequence in Yucatan, Mexico. *Biotropica* 44: 151–162.

Dupuy, J.M., Hernández-Stefanoni, J.L., Hernández-Juárez, R.A., Tun-Dzul, F. & May Pat, F. 2012b. Efectos del cambio de uso del suelo en la biomasa y diversidad de plantas leñosas en un paisaje de bosque tropical seco en Yucatán. *Investigación Ambiental Ciencia y Política Pública* 4: 130–140.

Eaton, J.M. & Lawrence, D. 2009. Loss of carbon sequestration potential after decades of shifting cultivation in the southern Yucatán. *Forest Ecology and Management* 258: 949–958.

Gallardo-Cruz, J.A., Meave, J.A., González, E.J., Lebrija-Trejos, E., Romero-Romero, M.A., Pérez-García, E.A., Gallardo-Cruz, R., Hernández-Stefanoni, J.L. & Martorell, C. 2012. Predicting tropical dry forest successional attributes from space: is the key hidden in image texture? *PLoS ONE* 7: e30506.

Gaveau, D.L.A. & Hill, R.A. 2003. Quantifying canopy height underestimation by laser pulse penetration in small-footprint airborne laser scanning data. *Canadian Journal of Remote Sensing* 29: 650–657.

Gillespie, T.W. 2005. Predicting woody-plant species richness in tropical dry forests: a case study from south Florida, USA. *Ecological Applications* 15: 27–37.

Goodwin, N.R., Coops, N.C. & Culvenor, D.S. 2006. Assessment of forest structure with airborne LIDAR and the effects of platform altitude. *Remote Sensing of Environment* 103: 140–152.

Hernández-Stefanoni, J.L., Gallardo-Cruz, J.A., Meave, J.A., Rocchini, D., Bello-Pineda, J. & López-Martínez, J.O. 2012. Modelling α- and β-diversity in a tropical forest from remotely sensed and spatial data. *International Journal of Applied Earth Observation and Geoinformation* 19: 359–368.

Hernández-Stefanoni, J.L., Dupuy, J.M., Johnson, J.D., Birdsey, R., Tun-Dzul, F., Peduzzi, A., Caamal-Sosa, J.P., Sánchez-Santos, G. & López-Merlín, D. 2014. Improving species diversity and biomass estimates of tropical dry forests using airborne LIDAR. *Remote Sensing* 6: 4741–4763.

Hill, R.A. & Broughton, R.K. 2009. Mapping the understory of deciduous woodland from leaf-on and leaf-off airborne LIDAR data: a case study in lowland Britain. *ISPRS Journal of Photogrammetry and Remote Sensing* 64: 223–233.

Hodgson, M.E., Jensen, J., Raber, G., Tullis, J., Davis, B.A., Thompson, G. & Schuckman, K. 2005. An evaluation of LIDAR-derived elevation and terrain slope in leaf-off conditions. *Photogrammetric Engineering and Remote Sensing* 71: 817–823.

Hughes, R.F., Kauffman, J.B. & Jaramillo-Luque, V.J. 1999. Biomass, carbon, and nutrient dynamics of secondary forests in a humid tropical region of México. *Ecology* 80: 1892–1907.

Jakubowski, M.K., Guo, Q. & Kelly, M. 2013. Tradeoffs between LIDAR pulse density and forest measurement accuracy. *Remote Sensing of Environment* 130: 245–253.

Leutner, B.F., Reineking, B., Müller, J., Bachmann, M., Beierkuhnlein, C., Dech, S. & Wegmann, M. 2012. Modelling forest α-diversity and floristic composition on the added value of LIDAR plus Hyperspectral remote sensing. *Remote Sensing* 4: 2818–2845.

Lovell, J.L., Jupp, D.L.B., Newnham, G.J., Coops, N.C. & Culvenor, D.S. 2005. Simulation study for finding optimal LIDAR acquisition parameters for forest height retrieval. *Forest Ecology and Management* 214: 398–412.

Magnussen, S., Næsset, E. & Gobakken, T. 2010. Reliability of LIDAR derived predictors of forest inventory attributes: a case study with Norway spruce. *Remote Sensing of Environment* 114: 700–712.

McElhinny, C., Gibbons, P., Bracks, C. & Bauhus, J. 2005. Forest and wood stand structural complexity: its definition and measurement. *Forest Ecology and Management* 218: 1–24.

McGaughey, R.J. 2012. FUSION/LDV: Software for LIDAR data analysis and visualization. United States Department of Agriculture, Forest Service, Pacific Northwest Research Station 154 pp.

McLaren, K.P. & McDonald, M.A. 2003. Coppice regrowth in a disturbed tropical dry limestone forest in Jamaica. *Forest Ecology and Management* 180: 99–111.

Miles, L., Newton, C.A., Defries, R.S., Ravilious, C., May, I., Blyth, S., Kapos, V. & Gordon, J.E. 2006. A global overview of the conservation status of tropical dry forest. *Journal of Biogeography* 33: 491–505.

Miller, D.M. 1984. Reducing transformation bias in curve fitting. *The American Statistician* 38: 124–126.

Murphy, P.G. & Lugo, A.E. 1986. Ecology of tropical dry forest. *Annual Review of Ecology and Systematics* 17: 67–88.

Naesset, E. 2005. Assessing sensor effects and effects of leaf-off and leaf-on canopy conditions on biophysical stand properties derived from small-footprint airborne laser data. *Remote Sensing of Environment* 98: 356–370.

Næsset, E. & Okland, T. 2002. Estimating tree height and tree crown properties using airborne scanning laser in a boreal nature reserve. *Remote Sensing of Environment* 79: 105–115.

Oliveira-Filho, A.T., Curi, N., Vilela, E.A. & Carvalho, D.A. 1998. Effects of canopy gaps, topography and soils on the distribution of woody species in a central Brazilian deciduous dry forest. *Biotropica* 30: 362–375.

Olmsted, I., González-Iturbe, J.A., Granados-Castellanos, J., Durán, R. & Tun-Dzul, F. 1999. Vegetación de la Península de Yucatán. In: de García Fuentes, A., Córdoba, J. & Chico, P. (eds.), *Atlas de Procesos Territoriales de Yucatán*. pp. 183–194. Universidad Autónoma de Yucatán, Mérida, MX.

Portillo-Quintero, C.A. & Sánchez-Azofeifa, G.A. 2010. Extent and conservation of tropical dry forests in the Americas. *Biological Conservation* 143: 144–155.

Rocchini, D. 2007. Effects of spatial and spectral resolution in estimating ecosystem α-diversity by satellite imagery. *Remote Sensing of Environment* 111: 423–434.

Rocchini, D., Hernández-Stefanoni, J.L. & He, K.S. 2015. Advancing species diversity estimate by remotely sensed proxies: a conceptual review. *Ecological Informatics* 25: 22–28.

Saremi, H., Kumar, L., Turner, R. & Stone, C. 2014. Airborne LIDAR derived canopy height model reveals a significant difference in radiata pine (*Pinus radiata* D. Don) heights based on slope and aspect of sites. *Trees* 28: 733–744.

Simonson, W.D., Allen, H.D. & Coomes, D.A. 2012. Use of an airborne LIDAR system to model plant species composition and diversity of Mediterranean oak forests. *Conservation Biology* 26: 840–850.

Takahashi, T., Yamamoto, K., Senda, Y. & Tsuzuku, M. 2005. Estimating individual tree heights of sugi (*Cryptomeria japonica* D. Don) plantations in mountainous areas using small-footprint airborne LIDAR. *Journal of Forest Research* 10: 135–142.

Urquiza-Haas, T., Dolman, P.M. & Peres, C.A. 2007. Regional scale variation in forest structure and biomass in the Yucatan Peninsula, Mexico: effects of forest disturbance. *Forest Ecology and Management* 247: 80–90.

Van Leeuwen, M. & Nieuwenhuis, M. 2010. Retrieval of forest structural parameters using LIDAR remote sensing. *European Journal of Forest Research* 129: 749–770.

Wasser, L., Day, R., Chasmer, L. & Taylor, A. 2013. Influence of vegetation structure on LIDAR-derived canopy height and fractional cover in forested riparian buffers during leaf-off and leaf-on conditions. *PLoS ONE* 8: e54776.

Zar, J.H. 1999. *Biostatistical analysis*. Prentice Hall, Engelwood Cliffs, NJ, US.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendix S1. Summary statistics of multiple linear regression of above-ground biomass (AGB) and species richness from LIDAR metrics for leaf-on and leaf-off canopy conditions.

Appendix S2. PCA ordination diagram of vegetation structure (abundance, number of species, basal area, plant cover, average height and canopy height) comparing flat areas and hills.