Bias and saturation in estimated LAD and LAI using a rangefinder ground LiDAR

Diogo Martins Rosa [1]
Bruce Walker Nelson [1]
Danilo Roberti Alves de Almeida [2]
Scott Christopher Stark [3]
Dalton Freitas do Valle [1]
Juliana Schietti de Almeida [1]

[1] Brazil´s National Institute for Amazon Research - INPA
Av. André Araújo, 2.936 - Petrópolis - 69067-375 - Manaus - AM - Brazil
{mrosa.diogo, bnelsonbr, daltonvalle, jujuschietti}@gmail.com

[2] University of São Paulo - USP/ESALQ
Caixa Postal 96 - 13416-000 - Piracicaba - SP - Brazil
daniloflorestas@gmail.com

[3] Michigan State University - Department of Forestry,
East Lansing - MI - 48824 - USA
scott.c.stark@gmail.com

Abstract. Leaf area index (LAI) and canopy openness were estimated at 2m intervals along 900m of transect in each of four forest types comprising different parts of a topographic sequence in a Central Amazon *terra firme* forest. Both metrics were obtained with a 1000 Hz rangefinder type LiDAR, using the distance to last return for each pulse. Pulses were directed vertically upward from the forest floor by an operator walking at constant speed. We avoided hours near noon, when direct sunlight could enter the detector. We estimated the leaf area density (LAD) of each voxel by the MacArthur-Horn method. We obtained LAI from the summed LADs of each voxel stack. Three causes of bias in LAD estimation are discussed. Histograms of LAI showed saturation, i.e. an upper limit of detectability, for all four forest types. Consequently, mean LAI did not differ between forest types, even if all 2m sections of transect were assumed independent (n=450). Corrections to these problems are being prepared, led by authors SCS and DRAA. Canopy openness did differ between the four forest types. This suggests that LAI does differ but our unadjusted methods did not permit detection of this difference.

Keywords: Forest canopy, canopy voxels, leaf area density.

1. Introduction

The leaf area index (LAI) is an important biophysical variable for tropical forest studies, but is difficult to obtain. Direct measurement is time consuming and has been undertaken in very few places (McWilliams et al., 1993; Clark et al., 2008). The most widely employed indirect methods employ hemispherical cameras or the LAI-2000 instrument. These are limited to ~45 minute windows of operation in the early morning and late afternoon. The LAI-2000 requires two instruments, one on a nearby tower for synchronized above-canopy measurements of the rapidly changing incoming radiant intensity during these time windows.

A recent and under-utilized alternative is a rangefinder-type LiDAR (Parker et al., 2004; Stark et al., 2012, 2015). Leaf area density (LAD) is estimated in vertical profiles of canopy voxels, and these are summed to obtain LAI of each voxel stack. LiDAR pulses are directed vertically upward from the ground by an operator walking at a constant pace. Pulses enter the base of each voxel. A certain fraction of these are returned to the sensor, registered as separate distance measurements. The remainder are assumed to continue upward to sample the next voxel in the stack. By accounting for all returns and their distances from the sensor, the fraction of pulses that enter, that return from, and that pass through each voxel can be calculated and used to estimate LAD of that voxel.

Using a portable tower in tropical rain forest at La Selva Biological Station in Costa Rica, Clark et al. (2008) directly measured LAD in 575 voxels (each 8.4 m^3) comprising 45 voxel stacks, providing 45 LAI measurements. This is a useful data set for detecting bias, error and departure from assumptions of indirect methods. For example, their LAD values had an exponentially decreasing frequency histogram (Figure 1), indicating that leaves were highly clustered in the canopy, with most voxels being empty or having few leaves, and a few voxels having high leaf density. This violates an assumption of the LAI-2000 instrument´s inversion algorithm. They also showed that the LAI obtained above 45 footprints of 4.6m^2 followed a normal, symmetric frequency distribution (Figure 1).

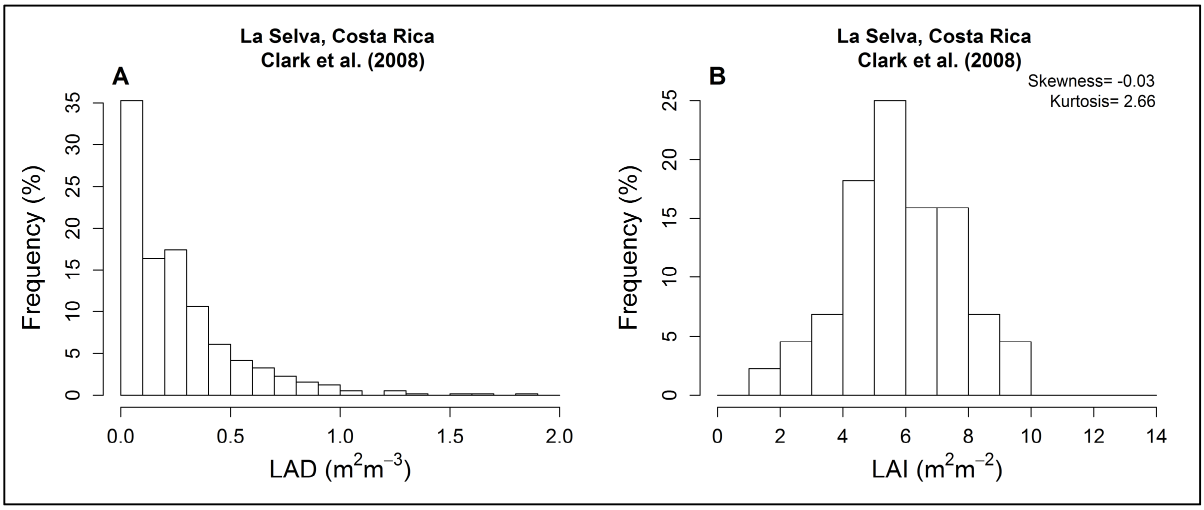

Figure 1. (A) Frequency histogram of LAD measured directly at La Selva Costa Rica for 575 voxels and (B) frequency histogram with normal distribution, for the corresponding 44 LAI values of the voxel stacks. All leaves were harvested and scanned in 44 vertical voxel stacks each having a footprint of 4.6 m^2. A single high outlier (with LAI = 12.9) was excluded. The number of measurements does not allow use of narrower classes. Data from Clark et al. (2008)

Stark et al. (submitted) used artificially constructed canopies of known LAI for the purpose detecting several types of bias in LiDAR-derived LAD and LAI. Here we report expected errors in estimates of voxel LAD, and two intriguing errors evident in the frequency histograms of LAI. We also compare the ranked frequency curves of voxel columns ordered by their canopy openness, for four forest types, as a non-saturating proxy for LAI differences between these forest types.

2. Methods

We estimated LAD, LAI and canopy openness at 2m intervals along 900m of transect in each of four forest types in a Central Amazon *terra firme* (Plateau, Upper slope, Lower slope on white sand and Riparian on white sand). We used the distance to last return for each pulse of a 1,000 Hz rangefinder-type LiDAR, Riegl model LD90-3100VHS-FLP (Horn, Austria). The instrument was held 1m above the ground and aimed vertically while transiting each forest transect at a constant slow walking pace. The pulse beam cross section was 4x11cm at 25m distance. This provided 2,300 pulses for last return detections, per linear meter of transect.

LAI is the sum of all LAD estimated for each of the voxels that comprise a voxel stack over a small ground footprint. The LAD of each voxel in a stack was estimated using the MacArthur-Horn method (Equation 1) (MacArthur and Horn, 1969; Parker et al., 2004; Stark et al., 2012), which assumes that the fraction of input pulses entering at the base of a voxel that return to the instrument is related exponentially to the LAD of that voxel (Equation 1; Figure 2).

$$\text{LAD} = -\ln(pulses.out\ /pulses.in) \times \frac{1}{D} \times \frac{1}{K} \tag{1}$$

Where D is the voxel height (D=1m) and K = calibration constant. K has been found to be about 0.83 when average LAI of a long transect was calibrated to directly measured LAI of 5.7 (McWilliams et al., 1993), near Manaus in the Central Amazon (SC Stark, unpublished data). Here we use K=1 for convenience, providing an “effective LAD” and “effective LAI”, hereafter referred to simply as LAD and LAI. Using 2m long voxels along track, we obtained 450 estimates of LAI and 450 estimates of canopy openness for each of the four forest types. Canopy openness was based on the number of non-returning pulses (sky shots) for each voxel column.

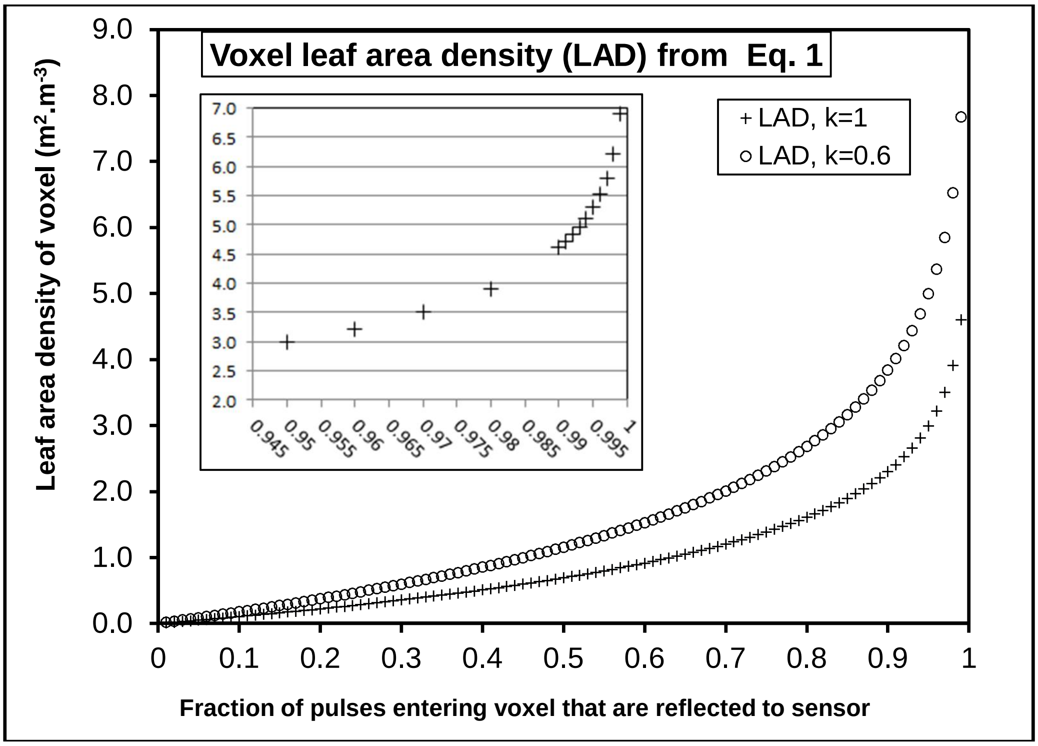

Figure 2. Leaf area density estimate of a voxel increases exponentially with the fraction of pulses entering at the base of a voxel that return to the sensor as a last-return distance measurement. Inset shows rapid increase in estimated LAD, when more than 95% of pulses return, for K = 1.

3. Results and Discussion

Three types of expected bias in the estimates of upper canopy voxel LAD, and two types of error evident in frequency histograms of LAI are reported. We begin with voxel LAD bias.

- First, for K=1 in equation 1, and for LAD greater than about 3.0, small changes in the fraction of input pulses returned to the sensor correspond to very large differences in voxel LAD (Figure 2). This will occur most frequently in partially occluded upper canopy voxels. As the LiDAR paints a narrow beam along the length of an upper canopy voxel that is mostly occluded it may see only see a few square cm of that voxel base. If this small area happens to be occupied by a leaf, or by a dense leaf cluster, 95% or more of the pulses reaching this small area may be returned, leading to an overestimate of LAD.
- Second, if 100% of the pulses striking the small visible portion of a mostly occluded upper canopy voxel are returned to the sensor, this will give log of zero in eq 1. This is

not permitted, so the voxel is registered as NA, a code reserved for the empty space above the top of the canopy. This will therefore cause underestimate of LAD and LAI, because a voxel containing leaves is coded as completely void of leaves.

- Third, if an upper canopy voxel is fully occluded along the entire narrow line painted by the LiDAR beam, it will also be registered as NA, also causing underestimate.

Two problems are evident in the LAI histogram. These are hiatuses and saturation, the latter causing a skewed frequency distribution.

- Hiatus in LAI range. Figure 3 shows he frequency histogram for 450 LAI measurements along 900m of plateau forest transect. Three breaks can be seen in the fine scale histogram and are more clearly visible in a normal QQ plot of the same data (Figure 4). These hiatuses are likely due to the loss of precision in the estimate of LAD in the upper canopy, i.e. intervals exist between the mathematically possible values of LAD when only a few pulses reach an upper voxel.
- Saturation. The histogram is truncated at the upper end, so that effective LAI saturates at about 8.0. The highest LAI class, which is above the last hiatus, has a very high frequency, also indicating saturation. A normal distribution would have very low frequencies at the upper end, but this highest LAI class holds 11% of the 450 LAI measurements. Note that a tail of low frequencies for very high LAI classes does exist in the real world data of Clark et al. (2008), which are normally distributed and not saturated.

All four forest types showed hiatuses and saturation in their LAI histograms. We tested for normality of the 450 LAI values of each forest type (Table 1). When subsampled to be compatible with the 44 LAI values from La Selva real-world data, the LiDAR-derived LAI class frequencies are reported by the Shapiro-Wilks test to be from a non-normal population, for at least 46% of the subsamples (based on 10,000 random draws of size 44).

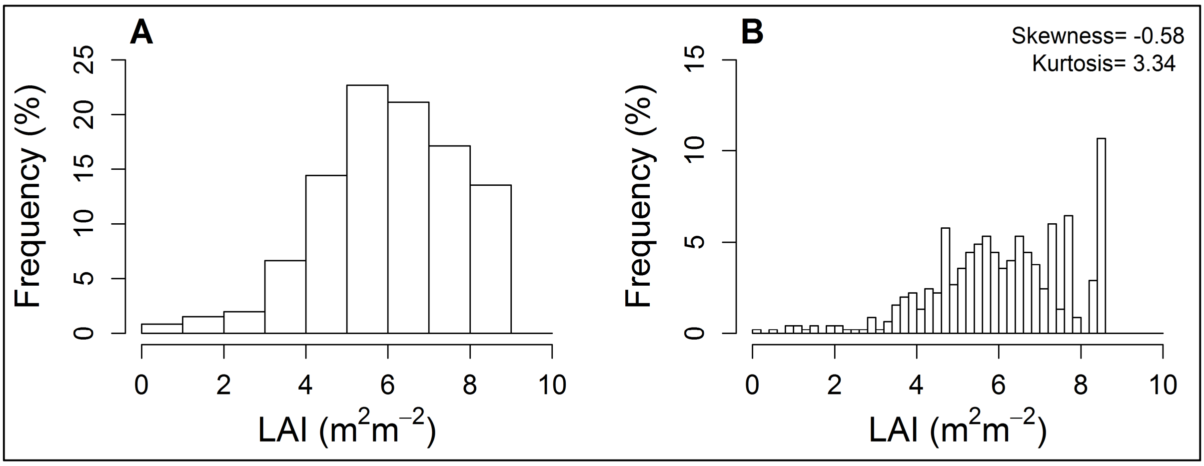

Figure 3. Frequency histogram of 450 small-footprint LAI measurements in Central Amazon evergreen forest on well drained clay-soil plateau. Along-track voxel dimension= 2m and K=1. Two class widths are shown, 1.0 LAI units (A) and 0.2 LAI units (B), the former for comparison to the real world data from La Selva, Costa Rica. Data was not normally distributed, but should have been so. This was the case in all forest types at our study area (Table 1).

Table 1. LAI frequency histogram parameters (mean, standard deviation, coefficient of variation, skewness, kurtosis and "% non-normal"). "% non-normal" is the percent of ten thousand runs of Shapiro-Wilks test for normality, each run having 44 randomly selected voxel stacks drawn from the full set of columns (full n= 450).

Forest types	Voxel length (m)	LAI mean	LAI SD	LAI CV	Skewness	Kurtosis	% non-normal
Clark et al. (2008)	-	5.8	1.9	0.32	-0.03	2.66	-
Plateau	2	6.0	1.7	0.28	-0.58	3.34	49
Upper slope	2	5.9	1.9	0.32	-0.58	3.04	51
Lower slope	2	5.8	1.7	0.30	-0.45	2.84	46
Riparian	2	5.9	1.7	0.28	-0.30	2.54	44

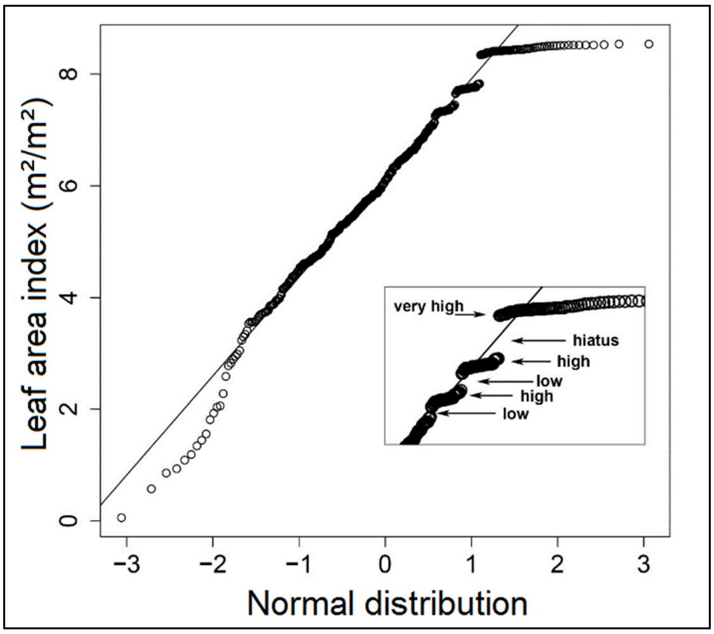

Figure 4. Q-Q plot of 450 ranked observed LAI values (*y*-axis) against their expected standardized distance from the mean in a normal distribution (*x*-axis and line). The inset is the upper portion expanded and the arrows indicate intervals of LAI having alternating hiatuses and high numbers of observations, confirming the patterns seen in Figure 3B.

Probably because LAI is saturated it was not possible to detect differences in mean LAI between any of the four forest types, even if all 450 transect segments of 2m length were assumed to be independent samples when estimating their means. However, filtering the highest LAD values and other adjustments (Stark et al., submitted) may resolve this issue, bringing out differences in the mean or the mode LAI values of these four forests. Discarding the highest ~5% of all voxels does make the LAI histograms normal.

Meanwhile, we draw attention to the potential of sky shots for separating the forest types based on their canopy openness. For this, we ranked the 450 voxel columns for each forest type, by their sky shots (Figure 5). While no test has yet been applied, visual inspection suggests that the forests differ. The upper slope forest had the most open canopy when comparing the most open voxel columns of each forest types (ranks 1 to 30, Figure 5A). The white sand forest of lower slopes had the most open canopy further along in the ranking (Figure 5B).

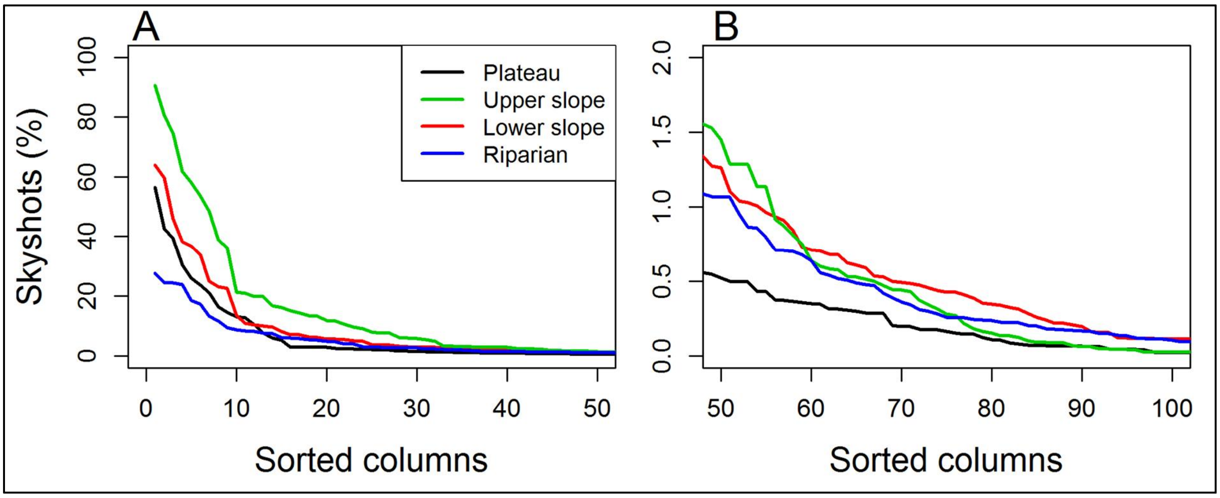

Figure 5. Proportion of sky shots (canopy openness) for the four types of *terra firme* forest, in each 2m section of transect. The *x*-axis has ranked 2m sections by decreasing number of sky shots. Sections ranked 1-50 are shown in A and those ranked 50-100 are shown in B.

4. Conclusions

Bias in LAD estimates of canopy voxels and saturation in LAI histograms derived from stacked voxels, impede detection of mean LAI differences between four different Central Amazon forests. Corrections to these problems are under consideration (Stark et al., submitted). Canopy openness (proportion of sky shots) derived from the same LiDAR is a potential non-saturating proxy for detecting LAI differences among these same four forest types.

Acknowledgements

Ground LiDAR fieldwork and data analysis were supported by the Fundação de Amparo à Pesquisa do Estado do Amazonas (FAPEAM) as part of the Green Ocean Amazon (GOAmazon) funding call (013/2013 - FAPEAM). The National Institute for Amazon Research (INPA) and its Large-Scale Biosphere-Atmosphere Project (LBA) maintain the study site and provided logistical support. Danilo Almeida acknowledges support from São Paulo Research Foundation (FAPESP) (grant 2016/05219-9).

References

Clark, D. A. D. B., Olivas, P. C., Oberbauer, S. F., Clark, D. A. D. B., Ryan, M. G., 2008. First direct landscape-scale measurement of tropical rain forest Leaf Area Index, a key driver of global primary productivity. **Ecology letters**, v. 11, n. 2, p. 163-172, 2008.

McWilliam, A. L.; Roberts, J. M.; Cabral, O. M. R.; Leitao, M. V. B. R.; de Costa, A. C. L.; Maitelli, G. T.; Zamparoni, C. A. G. P. Leaf area index and above-ground biomass of terra firme rain forest and adjacent clearings in Amazonia. **Functional Ecology**, v. 7, n. 3, p. 310-317, 1993.

Parker, G. G.; Harding, D. J.; Berger, M. L. A portable LIDAR system for rapid determination of forest canopy structure. **Journal of Applied Ecology**, v. 41, n. 4, p. 755-767, 2004.

Stark, S. C.; Leitold, V.; Wu, J. L.; Hunter, M. O.; de Castilho, C. V.; Costa, F. R.; ... & Keller, M. Amazon forest carbon dynamics predicted by profiles of canopy leaf area and light environment. **Ecology Letters**, v. 15, n. 12, p. 1406-1414, 2012.

Stark, S. C., Enquist, B. J., Saleska, S. R., Leitold, V., Schietti, J., Longo, M., ... & Oliveira, R. C. Linking canopy leaf area and light environments with tree size distributions to explain Amazon forest demography. **Ecology letters**, v.18, n. 7, p. 636-645, 2015.