

## ***Interactive comment on “Detecting tropical forest biomass dynamics from repeated airborne Lidar measurements” by V. Meyer et al.***

**Anonymous Referee #1**

Received and published: 29 March 2013

**General comments** This paper presents work in developing a method to estimate biomass change in a tropical forest using two lidar surveys over the same area separated by a decade, calibrated against field-derived (census) estimates of biomass from the same area. This is definitely an interesting and useful piece of work – estimating tropical biomass change is difficult due to the measurements required and lidar holds the potential for doing this on quite large scales, rapidly and with quantifiable uncertainty. The essence of the problem is that the signal (biomass change, expressed as change in canopy height distributions) may be really quite small against a background of potentially large variance in absolute canopy height and within any given forest region due to many other environmental (eg slope, soil, hydrology) and ecological (eg species composition) properties.

C643

The paper presents interesting results, and could potentially be a sound, valuable piece of work. However, significant problems exist with the analysis as it stands. The key weakness of the work is that it purports to be an exploration of lidar as a tool for measuring biomass change in tropical forests. However the paper itself focuses very little on the issue of how to do this from the point of view of the lidar data, and what the associated problems, uncertainties and limitations might be and focuses instead very largely on the resulting biomass change values. In some respects, the absolute values of change are not important; to demonstrate how lidar can be used to quantify change and how one might for eg combine lidar surveys from different instruments and survey properties (which is def. a very useful aim and one that is currently poorly addressed), the focus should be on how the estimates of biomass change are derived, and how robust and general the method is under the assumptions made. If this were done more completely, the uncertainty resulting from the lidar estimates of height change could then be fed through to the estimated biomass change. As it is, uncertainty due to the lidar methods is only considered qualitatively and very briefly at that, mainly in section 4.1.

All of the analysis of the lidar data is presented in the supplementary information, and only biomass results are presented in the main body of text. For the stated aims of the paper, this emphasis should effectively be reversed and the issue of how, how robust, how uncertain should be addressed in the main text.

These issues should be relatively straightforward to address as most of the material is there, albeit in the wrong place or requiring further depth to address the particular stated aims.

I make specific comments expanding these general points below.

**Specific comments** The lidar data are key to this whole study, particularly the ability to compare the LVIS and DRL datasets collected a decade apart. In this case, making sure these two datasets are estimating the same things, as far as possible, and quan-

C644

tifying differences that arise for reasons other than actual changes in height between the two dates, is key to robust, useful results. At present, the analysis does not fulfil this requirement.

All changes in biomass rest on this comparison and we know that these two lidar datasets are likely different in a number of ways (footprint size, geo-location, energy recorded, height, gain, scan angle etc etc). We also know that these things affect the resulting estimates of canopy height derived from the resulting lidar points – this has been shown in various model and measurement studies of DRL systems in particular (eg Naesset et al 2009; Hopkinson, 2007 who actually proposed ways of minimising differences between surveys; Disney et al., 2010 etc etc). The authors make little or no mention of these impacts, the differences between the two surveys, and their possible effects. All detail of the lidar acquisitions is in the supplementary information & the descriptions are also incomplete. For example, what altitude were the LVIS data collected at? What scan angles were the DRL collected at? What were the instrument characteristics of the DRL in terms of the threshold for pulse collection (if known)?

As an example, the geo-location accuracy of the two lidar datasets and the ground measurements, will probably be a major source of error, particularly if it is more than a few m (which is very likely) & particularly for any analysis at that order of scale. Clearly, analyses at larger scales reduces this effect but at scales of the LVIS footprint this is inevitably going to dominate. This suggests that all the analysis here at much less than 0.5-1 ha would be dominated simply by this aspect, so is there any point in doing it any finer? The authors seem to come to this conclusion through the paper, but it ought to be clear at the outset, so they could focus on the 1ha and larger and just state this. The authors acknowledge these location issues with the field data, but not the lidar data. So what is the accuracy of the GPS location of the 3 datasets? And what is the quantitative (not qualitative) impact of this? They could estimate this from the DRL for eg which has sufficient resolution to identify tree crowns which ought to be identifiable from the field data.

C645

The authors say: “In contrast, the DRL sensor provides an accurate estimate of ground elevation and vegetation height.” – how accurate and how do you know? This will depend to some extent on the particular sensor and survey characteristics again. The LVIS correction also depends very much on the accuracy of co-location as the DRL-derived DEM is used to correct the LVIS data. As a result, errors in this will propagate through to all AGB estimates. The authors note that “. . .it is difficult to quantify the improvements made by these corrections at the plot level on AGB estimations because there are only a few outliers in this part of the island.” But this is key to the subsequent results as the authors are interested in the absolute differences between the two. As a result they cannot at present quantify errors arising from uncertainty in the two lidar datasets.

Another issue is the use of the height metrics and regression models derived from them. This is key in going from lidar to biomass, but again, all this information and analysis of this is in the supplementary information, not the main paper. This is the heart of the method to estimate biomass (and hence biomass change) from lidar, so again is absolutely critical to the results and conclusion, but is rather hidden away. The assumptions made here, and their robustness and generality (or otherwise) must be analysed critically and in full in the main paper. If this is at the expense of the discussion of the resulting biomass change values, then those could go in the supplementary information – after all, those results are essentially the demonstration of the method, and so are only useful in the light of the method’s accuracy.

A key question here is: are the differences in biomass between the dates is due to differences in the observation methods and metrics derived from the lidar data, or genuine changes? In order to answer this question (and as noted above), the methods ought to be made as similar as possible in order to determine these potentially rather small changes against a background large variation in height. However, rather than provide a common metric/framework the authors find empirical best fits between lidar height and AGB, and this is different for the different lidar datasets (eqns 1 and 2 from

C646

the main paper). Not only are different numbers of parameters used (5 for LVIS, 4 for DRL) but the form appears to be different (quadratic for LVIS, linear for DRL). There is no justification given for these different forms, and the decision is based purely on RMSE of fit. But the form of the required relationship is known (eqns ES1 and ES2) and so why is this not used to inform the regression model form (which makes sense in eqn2 but not eqn1)? This immediately means that i) the models are different for the two different dates; and ii) the results are entirely dependent on the particular local calibration, and for each date.

The authors note: "Consequently, we used DRL and LVIS metrics independently in the determination of regression models for AGB estimations." – one way around this would be to generate a pseudo-LVIS dataset from the DRL by aggregation at the same scale, and comparing at that level (this is done at 1m scale, rather than 20m scale). This would seem a fairer comparison in many ways.

In section 4.1 the authors note the issues of uncertainty in choice of metric, but then do not present any quantitative analysis. As a result they are not able to propagate sources of uncertainty through from the height change estimates to the resulting biomass change estimates, despite their claims to do so (see abstract and p1962).

Figure S2 "In all cases, the correlations among the height metrics are strong" – yet in all cases but 1  $r^2 < 0.6$  and in 7 out of 9  $< 0.5$ . So what do you mean by strong? Fig S2 is critical to the results but again is in the supplementary info and needs more detailed analysis.

Re comparison with Mascaro et al 2011 - present study is locally calibrated and uses many params so is this "realistic" (p1972). What does this mean?

"analysis and used these errors when analyzing the Lidar estimation of biomass and biomass change" – so what are these errors? I presumed the SDs quoted in the main text were due to variation within the dataset not errors propagating from the biomass regression. Is this not the case?

C647

Technical comments Are changes significant at each scale? P1966 line 27 - is this statsig? P1967 line 17 - don't use amplitude you mean range or variance Table 1 6dp?? Higuchi 1994 missing Fig S4 RMSE not RSME. And why not just remove all erroneous trees from the census rather than replacing them with mean measurements? They are wrong so leave them out.

p1974: uncertainty in detecting small changes of biomass . . . is minimized in the BCI dataset "because of intensive effort in the field". What does this mean?

When analysing in terms of forest age - why not see if the analysis detects these differently, independent of a priori stratification? As it is this is imposed on analysis so if hypothesis is changes in these different age stands look different to lidar then analysis ought to show this maybe?

---

Interactive comment on Biogeosciences Discuss., 10, 1957, 2013.

C648