Comment on 'A first map of tropical Africa's above-ground biomass derived from satellite imagery'

COMMENT

Comment on 'A first map of tropical Africa's above-ground biomass derived from satellite imagery'

E T A Mitchard[1], S S Saatchi[2], S L Lewis[3], T R Feldpausch[3], F F Gerard[4], I H Woodhouse[1] and P Meir[1]

[1] School of GeoSciences, University of Edinburgh, Drummond Street, Edinburgh EH8 9XP, UK

[2] NASA Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109, USA

[3] Earth and Biosphere Institute, School of Geography, University of Leeds, Leeds LS2 9JT, UK

[4] Centre for Ecology and Hydrology, Maclean Building, Benson Lane, Crowmarsh Gifford, Wallingford, Oxfordshire OX10 8BB, UK

E-mail: edward.mitchard@ed.ac.uk

Abstract

We present a critical evaluation of the above-ground biomass (AGB) map of Africa published in this journal by Baccini *et al* (2008 *Environ. Res. Lett.* **3** 045011). We first test their map against an independent dataset of 1154 scientific inventory plots from 16 African countries, and find only weak correspondence between our field plots and the AGB value given for the surrounding 1 km pixel by Baccini *et al*. Separating our field data using a continental landcover classification suggests that the Baccini *et al* map underestimates the AGB of forests and woodlands, while overestimating the AGB of savannas and grasslands. Secondly, we compare their map to $216\,000 \times 0.25$ ha spaceborne LiDAR footprints. A comparison between Lorey's height (basal-area-weighted average height) derived from the LiDAR data for 1 km pixels containing at least five LiDAR footprints again does not support the hypothesis that the Baccini *et al* map is accurate, and suggests that it significantly underestimates the AGB of higher AGB areas. We conclude that this is due to the unsuitability of some of the field data used by Baccini *et al* to create their map, and overfitting in their model, resulting in low accuracies outside the small areas from which their field data are drawn.

S Online supplementary data available from stacks.iop.org/ERL/6/049001/mmedia

1. Introduction

The ERL paper by Baccini *et al* (2008), 'A first map of tropical Africa's above-ground biomass derived from satellite imagery', was a timely attempt to combine available field and remotely sensed data to produce the first above-ground biomass (AGB) map of a significant portion of sub-Saharan Africa. The authors used passive optical remote sensing data, which generally has not been found to be very sensitive to AGB at higher biomass values (Zheng *et al* 2004, Lu 2006,
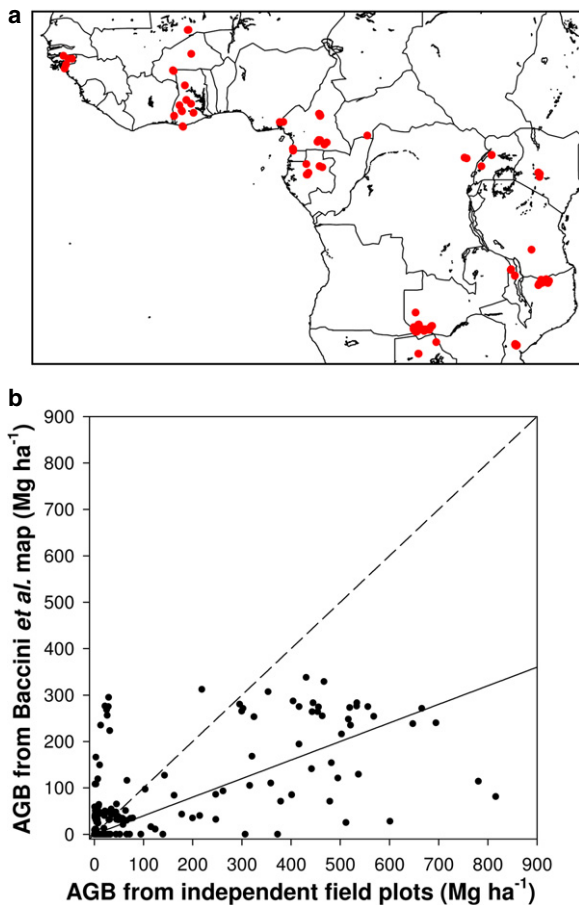
Figure 1. (a) Location of the 1154 scientific inventory plots used to compare to the Baccini *et al* AGB map; (b) AGB values from pixels of the Baccini *et al* map containing inventory plots plotted against the weighted average AGB of independent inventory plots; also shown is the best fit line (solid), and the 1:1 line (dotted).

GOFC-GOLD 2009, Mitchard *et al* 2009). Still, Baccini *et al* report a high accuracy, with the map explaining 82% of the variance in AGB for 10% of field plots held back for validation, with a root mean squared error (RMSE) of 50.5 Mg ha^{-1}. They then perform a test against spaceborne LiDAR height metrics from across the whole spatial extent of the map, and report an r^2 of 0.90 in a regression between mean LiDAR derived height and AGB (averaged over 10 Mg ha^{-1} AGB classes). We tested the Bacinni *et al* results against independent and spatially extensive field data, and newly calculated spaceborne LiDAR results, and found little support for the accuracy of the map (figures 1, 2, table 1). Our conclusion is that this is due to the low accuracy and limited spatial extent of the field data used to train and validate the Random Forest model used to produce the AGB map.

2. Test against field data

We first test the accuracy of the Baccini *et al* map directly using AGB derived from 1154 scientific inventory plots from 16 African countries, ranging in size from 0.1–10 ha (mean plot size 0.32 ha, mean 1.5 ha inventoried per 1 km pixel,

figure 1; for plot details, see the supplementary material, available online at stacks.iop.org/ERL/6/049001/mmedia). In order to ensure sufficient sampling within each 1 km pixel, small plots (<0.5 ha) are included in this analysis only if the 1 km pixel in which they are located contains at least 0.5 ha of inventory plots. If multiple field plots occurred within one pixel, we calculated a mean AGB value, weighted by the square root of plot size. There are on average 4.8 field plots per 1 km pixel, so we compared field plots and the AGB map in a total of 239 pixels. The plots were collected from 1995 to 2010, with a mean julian date corresponding to July 2005 (compared to the remote sensing data in the Baccini *et al* map from 2000 to end 2003).

We find a significant, but very weak correlation, between our field plot AGB values and those in the Baccini *et al* map: a linear regression gave $r^2 = 0.28$, $p < 0.001$ (F-test), slope of 0.37, and RMSE of 145 Mg ha^{-1} (figure 1(b)). In this the best fit line had an intercept and slope significantly different from 0 and 1 respectively ($p < 0.01$). Errors range from an overestimate of 295 Mg ha^{-1} to an underestimate of -734 Mg ha^{-1}; the Baccini *et al* map has a much smaller range of AGB values than our field plots, with all higher AGB plots underestimated. When the plots are grouped by landcover type, using the Global Land Cover 2000 (GLC 2000) dataset (Mayaux *et al* 2004), the AGB of forest and woodland classes are underestimated by the Baccini *et al* map by ~50%, while shrubland/grassland classes are mostly overestimated (table 1, only landcover classes where we had at least 10 ha of field plots covering at least ten 1 km pixels were considered).

There are five possible explanations for this discrepancy if the Baccini *et al* map is accurate; however regressions with subsets of our field data do not support any of these hypotheses. In all the following regressions the best fit lines are significant ($p < 0.01$), and intercepts and slopes are significantly different from 0 and 1 respectively ($p < 0.05$). First, this could be caused by our field plots having a larger AGB range than the AGB map. This is not the case, as excluding pixels with an average AGB > 338 Mg ha^{-1} (the maximum in the Baccini *et al* dataset) gives an r^2 of 0.12, slope of 0.36, and an RMSE of 79.6 Mg ha^{-1} ($n = 204$ pixels): as would be expected the RMSE is reduced by removing the high AGB plots, but the overall accuracy (based on the r^2 and slope) actually decreases. Second, the non-normal distribution of biomass for very small plots (Chave *et al* 2003) may drive the poor fit. This is not the case, as if we limit our field data to pixels that have a total plot area $\geqslant$1 ha (1% coverage of the 1 km^2 pixel), although the r^2 increases to 0.32, the slope does not change at 0.38 ($n = 128$), and the RMSE of 169.5 Mg ha^{-1} is higher than for the whole dataset; additionally we have four plots of 10 ha in size from eastern Democratic Republic of Congo—these have an average AGB value of 463 Mg ha^{-1}, but the two Baccini *et al* pixels in which they fall (of which these plots sample 15%), are given AGB values of 273 and 283 Mg ha^{-1}. Third, our independent validation compares field-measured values from small plots to 1 km pixels (mean plot size = 0.3 ha, mean 4.8 plots per 1 km pixel); such plots may not sample the whole pixel sufficiently to accurately estimate its AGB. However, we do not think this third hypothesis can explain the extent of the poor correlation,
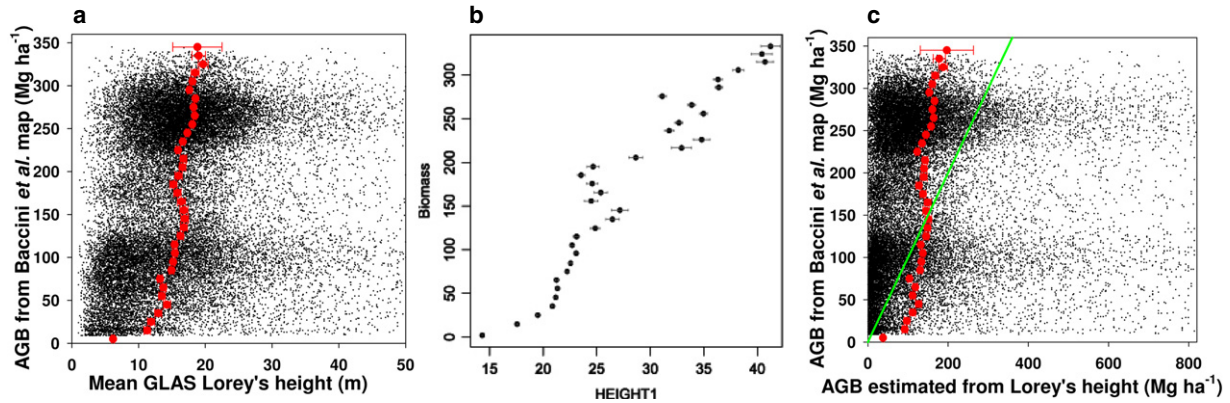
Figure 2. (a) AGB from 35 034 pixels from the Baccini *et al* map plotted against the mean Lorey's height from that pixel, each estimated from ⩾5 ICESat GLAS footprints; also shown in red is the mean height estimated for each 10 Mg ha^{-1}, with error bars showing standard errors; (b) reprint of figure 7(a) from Baccini *et al*, showing the mean maximum height for each pixel averaged into 10 Mg ha^{-1} biomass classes, with errors bars showing standard errors; (c) as for (a), but with Lorey's height values transformed into estimates of AGB (using an equation derived from African plot data), and with a 1:1 line shown in green.

Table 1. Difference between Baccini *et al* map and field plots by landcover class. (Note: the average AGB value from the Baccini *et al* map and field plots are displayed, separated by the class from the GLC 2000 land cover class Mayaux *et al* 2004.)

GLC 2000 land cover class	Number of pixels sampled	Total number of field plots within sampled pixels	Total area of field plots within sampled pixels (ha)	Mean AGB from Baccini *et al* (Mg ha^{-1})	Mean AGB from independent field plots (Mg ha^{-1})	Difference (Mg ha^{-1})
Closed evergreen lowland forest	29	38	75.1	202.5	445.1	−242.6
Submontane forest (900–1500 m)	11	16	20.2	210.0	438.1	−228.1
Deciduous woodland	66	308	76.7	20.2	36.5	−16.3
Deciduous shrubland with sparse trees	44	196	42.1	14.8	4.3	10.4
Open deciduous shrubland	28	371	46.0	0.2	4.1	−3.9
Closed grassland	16	99	32.5	27.0	3.9	23.0

as the error increases when we only consider the 70 pixels that have more than five field plots located within them (mean 13 plots per 1 km pixel for this reduced dataset): $r^2 = 0.14$, slope $= 0.42$, RMSE $= 170.2$ Mg ha^{-1}. Fourth, there is a potential for changes in the landcover of our field plots (which range from 1995 to 2010, compared to 2000 to 2003 for the remote sensing data of Baccini *et al*); however considering only field plots from 2000 to 2003 the results do not improve, with $r^2 = 0.36$, slope $= 0.39$, and RMSE $= 291$ Mg ha^{-1} ($n = 38$). Finally, our fifth hypothesis relates to differences in the choice of the allometric equations relating measured tree parameters to AGB. The allometrics chosen differ between our datasets and those of Baccini *et al*: Baccini *et al* use the Brown *et al* (2005) equation involving diameter alone for the majority of their field plots, whereas we use a combination of the Chave *et al* (2005) 'moist forest' equation involving diameter, height and wood density for our forest plots, and for savanna/woodland vegetation we used either the Chave *et al* (2005) 'dry forest' equation or locally derived area-specific equations (see the supplementary material, available online at stacks.iop.org/ERL/6/049001/mmedia). We tested the impact of equation choice for 87 of our field plots where we have access to the raw stem data (from savannas and tall forests in Cameroon, Gabon and Uganda, with similar characteristics to the field data used by Baccini *et al*): the plots had a mean AGB of 363 Mg ha^{-1} using a combination of the Chave *et al*

(2005) dry and moist equations, and 376 Mg ha^{-1} using the Brown *et al* (2005) equation involving diameter alone. A direct comparison of individual plot values using both allometric equations provides little support for bias at low or high biomass values: linear RMA regression gave intercept $= 3.2$, slope $= 0.98$, $r^2 = 0.97$, $p < 0.0001$.

3. Test against LiDAR data

In our second test, we examine the spaceborne LiDAR verification as performed by Baccini *et al*. To do this we used a dataset of ICESat Geoscience Laser Altimeter System (GLAS) footprints collected between 2003 and 2007, which were processed to provide estimates of Lorey's height, a basal-area-weighted average height that can be estimated accurately from GLAS waveforms (Lefsky *et al* 2005, Lefsky 2010). We use Lorey's height, as opposed to maximum height and HOME as used by Baccini *et al*, because the evidence suggests it can more accurately be determined from GLAS data, and because it has a stronger relationship with AGB (Lefsky *et al* 2005, Lefsky 2010). After cloud- and terrain-filtering (Saatchi *et al* 2011), we selected pixels that contained ⩾5 GLAS footprints (Baccini *et al*'s used three; each footprint covers 0.20–0.25 ha), giving a total of 35 034 test pixels, using 215 733 GLAS footprints. We find only a very weak

relationship between Lorey's height and Baccini *et al*'s AGB values when considering individual pixel values (figure 2(a), linear regression: slope $= 0.02$, $r^2 = 0.045$). We also averaged Lorey's height values in 10 Mg ha^{-1} bins, replicating the display method used in figure 7 of the Baccini *et al* study, and reproduced here as figure 2(b). We could not replicate their strong relationship between mean height and AGB, instead finding just a weak trend towards increasing height with increasing AGB up to $\sim$80 Mg ha^{-1}, and no relationship thereafter. We extend this analysis further by using Lorey's height (H_L)–AGB relationship derived from plot data in Africa (AGB $= 0.3542$ ($H_L^{2.0528}$), $n = 75$, $r^2 = 0.85$, $p < 0.001$, Saatchi *et al* 2011); again this suggests the Baccini *et al* map underestimates AGB in higher AGB areas, and has a low accuracy throughout the range of AGB values (figure 2(c), linear regression: slope $= 0.05$, $r^2 = 0.01$, RMSE $= 173$ Mg ha^{-1}).

We are unable to explain this discrepancy between our GLAS analysis and that of Baccini *et al*, though one factor could be that the metric derived from the raw GLAS waveform that we used (Lorey's height) is different from the metric they used (an estimate of canopy height, and the ratio of HOME to height). As Lorey's height is an average height weighted by basal area, its value will always be lower than maximum height for the same forest. However, it should be more sensitive to AGB than any estimate of height alone, and yet it does not appear to increase with AGB here. The result we report here does appear to concur with the results of the field data comparison, that is, that the Baccini *et al* map appears to have a low accuracy, in contrast to those reported within the paper. The possible causes of this low accuracy are fourfold: (i) the quality of the field data, which were mostly not scientific plots; (ii) the field data were not collected at a similar time to the remote sensing data; (iii) some of the 'field data' points used by Baccini *et al* are derived from a landcover map, itself derived from remote sensing; (iv) the field data were from a very limited spatial distribution, and not from across the continent. These issues are discussed below.

4. Discussion of Baccini *et al*'s field data

We fully sympathize with the difficulties faced by Baccini *et al* in obtaining sufficient numbers of high quality field plots across a continent, as this is extremely challenging. However, the field data used by Baccini *et al* are unlikely to be suitable for developing an accurate AGB map, as in addition to likely high randomly distributed inaccuracies, they are also likely to have consistent biases. We shall specifically examine the three datasets Baccini *et al* used in detail in order to highlight the potential problems with these types of data.

(1) The commercial forest inventory plots in the Republic of Congo (collected 2001–3) relied on measuring the diameters of just 1% of stems >40 cm diameter, 0.5% of stems 20–40 cm, and 0.2% of 'commercial species only' 2–20 cm. This very low proportion of diameters measured is likely to lead to inaccurate AGB estimates and, unless the small proportion chosen for measurement is strictly

random (with regards to both species and diameters), will lead to biased estimates. Additionally, logging companies, until very recently, have not collected data to estimate biomass stocks, but to assess the approximate density and size-class distribution of timber trees. Therefore: (i) the plot sizes and tree diameters may be inaccurate (indeed it is not specified whether or not the trees were measured here, often in such commercial inventories trees are placed in broad DBH classes rather than measured to the nearest mm); and (ii) the trees to be measured were unlikely to be a strict random subset of all the trees present. Though Réjou-Méchain *et al* (2011) did not find that commercial forestry inventories have a strong bias towards commercial species, as is often assumed, the above problems are still sufficient to result in large errors in AGB estimates. Baccini *et al* only used these data when at least three biomass plots were located within the same 1 km pixel. However, this averaging step will only reduce noise in the dataset; it will not correct for any systematic biases introduced by the methodology. This dataset makes up 65% of the pixels used by Baccini *et al* for training and validation.

(2) The dataset used by Baccini *et al* from Cameroon involved measuring the diameters of all stems greater than 10 cm DBH for 3 ha × 1 ha plots within each of 61 pixels. Unfortunately the diameters were only recorded as being within 10 cm bands rather than measured to the nearest millimetre, as is normal for scientific inventory plots: this will reduce accuracy. The biomass results for these plots appear very low for 'dense humid forest' from South-Central Cameroon (mean c. 100 Mg ha^{-1}, maximum 220 Mg ha^{-1}, based on Baccini *et al*'s figure 3). Other field plots published from within this area all have AGB values >400 Mg ha^{-1} (Djuikouo *et al* 2010, Lewis *et al* 2009). Furthermore the plots were measured in 1994, while the remote sensing dataset is from 2000 to 2003: Baccini *et al* use Landsat TM scenes from 1990 and 2000 to exclude plots that have undergone 'forest cover change' over this period, but quite significant changes will not necessarily be visible in TM data (GOFC-GOLD 2009). The accuracy of this dataset is therefore hard to assess, but it makes up only 4% of the pixels used in the Baccini *et al* study.

(3) Baccini *et al*'s dataset from Uganda is possibly the least accurate. Again, very little description of these plots is given in the paper, however the referenced Drichi (2003) 'National biomass study' from the Uganda Forest Department presents a landcover map of Uganda, with the country divided into vegetation classes using SPOT remote sensing data from 1990 to 1994, with data from a field campaign involving 4000 small forest inventory plots being used to give each vegetation class an average AGB value. However, the actual field plots were not used for this study, but instead Baccini *et al* interpolated AGB values for their pixels from this 'high resolution land cover type map', i.e. the proportion of each landcover class within each 1 km MODIS pixel was multiplied by its AGB value in order to give a weighted mean AGB value for that

pixel. Landsat TM data was then used to select <0.2% of the ∼236 000 pixel dataset (442 pixels are used, selected using undefined criteria). The use of optical remote sensing data to define the original landcover classes could explain the high accuracies reported by Baccini *et al*, as similar spectral information is used both to define and later separate biomass values; this will inevitably lead to higher accuracies than when truly independent field data is used. Equally, the use of a single average AGB value for each landcover class introduces pseudoreplication, as multiple pixels containing the same landcover class will be given identical AGB values (derived from the same plot data), but are treated as independent data points by the analysis. This dataset provides almost all the savanna and woodland training points used in the Baccini *et al* map, which is the landcover of 91% of the total area predicted (Mayaux *et al* 2004).

5. Discussion

Given the likely quality of the field data, it is surprising that the model Baccini *et al* develop appears to be so accurate against their test data. For example, it performs well against 10% of data held back for testing (training: 96% variance explained and RMSE 23.5 Mg ha^{-1} versus testing: 82% and 50.5 Mg ha^{-1}). This apparent contradiction, with the model performing well against the three datasets included in Baccini *et al*, but not in the field data we compiled, may be because of the circularity of using a landcover map partially derived from remote sensing data to derive the Uganda dataset, the pseudoreplication inherent in the Uganda dataset, and the small biomass range in the Baccini *et al* dataset compared to our dataset. However, an alternative explanation may be that the complex Random Forest model developed using a suite of MODIS variables to relate to AGB is not invariant across the continent. This is a significant danger: the Baccini field plots are located in three relatively small areas from approximately 1°N–4°N, and most vegetation types, or ecoregions, were not sampled. In general using Random Forest (or other non-parametric models) with limited and uneven spatial sampling of variables, results in overfitting the training data and produces large predictive errors outside the training regions (Genuer *et al* 2008). We suggest that their model may work relatively well for these three regions containing training data, while being poor in other regions, if, as is conceptually likely, the complex interactions of reflectance data that correspond to different AGB values within their model are not invariant across the full extent of the predicted AGB map.

6. Conclusion

In conclusion, we present evidence that the Baccini *et al* biomass map of Africa has large errors, with discrepancies between their map and independent scientific inventory plots resulting in an RMSE of 145 Mg ha^{-1}, and field data averaged by vegetation class suggesting that the AGB values for forest areas are underestimated, and for savanna areas mostly overestimated. Three major lessons should be taken from this analysis, to avoid these types of errors in the future: these

apply equally to all studies that attempt to use point data to extrapolate an ecological variable across a landscape. The first lesson is that care must be taken to use good quality, unbiased field data: if there are sufficient plots then it is not necessary for the individual field data points to have a high accuracy, but if they have inherent biases then the resulting map will not be valid. The second lesson is that field data must be drawn from across the spatial extent and ecological variability of the prediction area; due to logistical constraints an even spatial distribution of plots is rarely possible. However, if plots are unevenly distributed then this must be considered in the analysis, and ideally a map showing an estimated distribution of accuracy should be included. Finally, accuracy assessments should be done against truly independent datasets, not a small random subset of the input data, which may suffer from the same biases or be related in other ways than just the parameter of interest.

References

Baccini A, Laporte N, Goetz S J, Sun M and Dong H 2008 A first map of tropical Africa's above-ground biomass derived from satellite imagery *Environ. Res. Lett.* **3** 045011

Brown S, Pearson T, Moore N, Parveen A, Ambagis S and Shoch D 2005 Impact of selective logging on the carbon stock of tropical forests: Republic of Congo as a case study *Technical Report 6* (Winrock International http://www.winrock.org)

Chave J, Condit R, Lao S, Caspersen J P, Foster R B and Hubbell S P 2003 Spatial and temporal variation of biomass in a tropical forest: results from a large census plot in Panama *J. Ecol.* **91** 240–52

Chave J *et al* 2005 Tree allometry and improved estimation of carbon stocks and balance in tropical forests *Oecologia* **145** 87–99

Djuikouo M N K, Doucet J-L, Nguembou C K, Lewis S L and Sonké B 2010 Diversity and aboveground biomass in three tropical forest types in the Dja Biosphere Reserve, Cameroon *Afr. J. Ecol.* **48** 1053–1063

Drichi P 2003 *National Biomass Study Technical Report* (Forest Department PO Box 1613, Kampala, Uganda)

Genuer R, Poggi J-M and Tuleau C 2008 Random Forests: some methodological insights *Research report INRIA Saclay, RR-6729* (http://hal.inria.fr/inria-00340725/fr/)

GOFC-GOLD 2009 Reducing greenhouse gas emissions from deforestation and degradation in developing countries: a sourcebook of methods and procedures for monitoring, measuring and reporting *GOFC-GOLD Report, Version COP14-2* (Alberta)

Lefsky M A 2010 A global forest canopy height map from the Moderate Resolution Imaging Spectroradiometer and the Geoscience Laser Altimeter System *Geophys. Res. Lett.* **37** 043622

Lefsky M A, Harding D J, Keller M, Cohen W B, Carabajal C C, Espirito-Santo F D B, Hunter M O and de Oliveira R 2005 Estimates of forest canopy height and aboveground biomass using ICESat *Geophys. Res. Lett.* **32** L22S02

Lewis S L *et al* 2009 Increasing carbon storage in intact African tropical forests *Nature* **457** 1003–6

Lu D 2006 The potential and challenge of remote sensing-based biomass estimation *Int. J. Remote Sens.* **27** 1297–32

Mayaux P, Bartholome E, Fritz S and Belward A 2004 A new land-cover map of Africa for the year 2000 *J. Biogeogr.* **31** 1–17

Mitchard E T A, Saatchi S S, Gerard F F, Lewis S L and Meir P 2009 Measuring woody encroachment from 1982–2006 along a forest-savanna boundary in central Africa *Earth Interact.* **18** 8

Réjou-Méchain M, Fayolle A, Nasi R, Gourlet-Fleury S, Doucet J-L, Gally M, Hubert D, Pasquier A and Billand A 2011 Detecting large-scale diversity patterns in tropical trees: Can we trust commercial forest inventories? *For. Ecol. Manage.* **261** 187–194

Saatchi S S *et al* 2011 Benchmark map of forest carbon stocks in tropical regions across three continents *Proc. Natl Acad. Sci. USA* at press (doi:10.1073/pnas.1019576108)

Zheng D, Rademacherb J, Chena J, Crowc T, Bresee M, Moined J L and Ryu S 2004 Estimating aboveground biomass using Landsat 7 ETM + data across a managed landscape in northern Wisconsin, USA *Remote Sens. Environ.* **93** 402–11