

CORRESPONDENCE



Seeing the forest beyond the trees

ABSTRACT

In a recent paper (Mitchard *et al.* 2014, *Global Ecology and Biogeography*, **23**, 935–946) a new map of forest biomass based on a geostatistical model of field data for the Amazon (and surrounding forests) was presented and contrasted with two earlier maps based on remote-sensing data Saatchi *et al.* (2011; RS1) and Baccini *et al.* (2012; RS2). Mitchard *et al.* concluded that both the earlier remote-sensing based maps were incorrect because they did not conform to Mitchard *et al.* interpretation of the field-based results. In making their case, however, they misrepresented the fundamental nature of primary field and remote-sensing data and committed critical errors in their assumptions about the accuracy of research plots, the interpolation methodology and the statistical analysis. By ignoring the large uncertainty associated with ground estimates of biomass and the significant under-sampling and spatial bias of research plots, Mitchard *et al.* reported erroneous trends and artificial patterns of biomass over Amazonia. Because of these misrepresentations and methodological flaws, we find their critique of the satellite-derived maps to be invalid.

Keywords

Allometry, biomass, lidar, remote sensing, spatial modelling, tree height, tropical forests, wood density.

INTRODUCTION

The estimation of carbon stocks in tropical forests is challenging for several reasons: (1)

diversity in the structure, wood density and dynamics of tropical forests leads to complex and variable allometry (Chave *et al.*, 2004); (2) natural and anthropogenic disturbances at various spatial and temporal scales add to forest heterogeneity (Espírito-Santo *et al.*, 2014); and (3) there is no strong relationship between environmental (climate and soil) variables and forest biomass for predicting regional variations. Therefore, to meet this challenge, ground and remote-sensing observations have been combined to provide estimates of biomass distribution in the tropics at regional to continental scales (Saatchi *et al.*, 2007, 2011; Mascaro *et al.*, 2011; Baccini *et al.*, 2012; Asner *et al.*, 2013). On a per-hectare basis, the ground data (generally consisting of all tree diameters above a threshold, a sampling of tree heights and species identification that permits the inference of wood densities) are more comprehensive than remote-sensing data that generally measure aggregate structure such as canopy height. In contrast, airborne or satellite remote-sensing data are far more extensive, including millions of measurements over regional or continental scales compared with hundreds for research plots. Both are measures of physical properties that are not forest biomass (Clark and Kellner, 2012). Both efforts rely on statistical techniques to estimate biomass, using single-tree allometry in the case of field plots and plot-aggregate allometry in the case of satellite data.

Here, we show that Mitchard *et al.* (2014) misrepresented what they measured in the plots and committed significant methodological errors in extrapolating biomass estimates from plots to the whole of Amazonia and comparing their results with the satellite-derived maps of Saatchi *et al.* (2011) (RS1) and Baccini *et al.* (2012) (RS2).

THE FALLACY OF GROUND TRUTH

Mitchard *et al.* (2014) used 413 plots covering a total of 404.6 ha to sample more than 650 million ha of forests in Amazonia and the Choco region, west of the Andes. The census data were taken between 1956 and 2013 with

more than a third of the plots last censused before 1995. The quality of structure measurements and botanical information for more than half of the plots outside RAINFOR (Malhi *et al.* 2006) and the TEAM network (<http://www.teamnetwork.org/>) is unknown. Mitchard *et al.* (2014) argued that their biomass estimation from research plots and the derived maps are more accurate than satellite derived maps. The reader is left with the tacit impression that the data of Mitchard *et al.* (2014) must be correct because they come from the ground, while the RS1 and RS2 data are from satellites. This argument is unsurprisingly compelling because the human brain is hard-wired to accept results from physical contact rather than from distant measurements. Liberman & Trope (2008) note that for the human brain, 'Remote locations should bring to mind the distant rather than the near future, other people rather than oneself, and unlikely rather than likely events.' In other words, 'to see the forest we need to step back, whereas to see the trees we need to get closer' (Liberman and Trope, 2008).

Are the primary Mitchard *et al.* (2014) data superior because they are closer to trees? No

Before 2005, most local and regional tropical forest biomass allometry was based on measurements of tree diameter (D) only (Brown, 1997; Chambers *et al.*, 2007; Chave *et al.*, 2004), and in most cases did not include wood density (ρ) or height–diameter (H – D) relations. Following the work of Chave *et al.* (2005), Mitchard *et al.* (2014) estimated biomass using diameter, height and wood density – as did the estimates of biomass using satellite data in RS1 and RS2. The analyses of RS1 and RS2 use similar and sometimes overlapping data to Mitchard *et al.* (2014). There is no clear evidence that one set of data is superior to another. Therefore, the main difference between the Mitchard *et al.* (2014) and the satellite analyses is in the extrapolation approach. We note that the ground-based biomass estimates for all three studies have been challenged

recently by the publication of a new tropical forest allometry by Chave *et al.* (2014).

Mitchard *et al.* (2014) provided six estimates of biomass using the Chave *et al.* (2005) moist forest allometry with three or two parameters (D , H and ρ) to allow for variations in biomass estimation but did not include these differences as uncertainty in their analysis. Moreover, each allometric estimate may have additional error of 10–20% (smaller plots have a larger uncertainty) if one uses error propagation from basic measurements to model implementation (Chave *et al.*, 2004). Including these errors (e.g. 10%) along with estimates from different allometries provides a realistic variation around the mean biomass for each plot location (Fig. S1a in Supporting Information).

Mitchard *et al.* (2014) claimed that the four regional H – D models introduced by Feldpausch *et al.* (2012) improve ground estimates of biomass by ‘greatly reducing the error in the prediction of H from D compared to a pan-Amazonian model’. The models are implemented by casually assigning the plots to four manually delineated regions of Amazonia without any systematic method of stratification. The western Amazon model, for example, covers areas from highly seasonal rainfall in the south to areas with no dry season in the north, with soils varying from infertile on the east to fertile by the Andes foothills, and vegetation types varying as in floodplains, bamboo dominated and terrains with widely different geomorphology and topography. The H – D allometry by Feldpausch *et al.* (2012) has been found to introduce large bias (>20%) (Chave *et al.*, 2014) in estimates of biomass when compared with local relations (Fig. S1b) (see, e.g., Hunter *et al.*, 2013; Kearsley *et al.* 2013) and are probably simplistic approximations of H – D variations (Fig. S1c).

Wood specific gravity, the Achilles’ heel of biomass estimation

In allometric models, biomass at tree level or at aggregate plot level varies linearly with wood density (Chave *et al.*, 2005; Asner & Mascaro, 2014). However, wood density is not directly measured in the field and estimates are often extracted from published tabulated data with large uncertainty due to variations in measurement techniques, sample size, geographic concentration of samples and identification of species (Muller-Landau, 2004). The spatial variation of average wood density over Amazonia is

unknown, but is expected to be large because of geographic variations in taxonomy and phylogenetic characteristics (Chave *et al.*, 2009), as well as interspecific and inter-site variations in both soil fertility and complex processes of tree mortality (Muller-Landau, 2004). Field observations suggest that there is a significant pattern in wood density related to soil characteristics – trees with higher wood density in infertile soils of eastern Amazonia and those with lower wood density in more fertile soils of western Amazonia near the Andes foothills (ter Steege *et al.*, 2006; Quesada *et al.*, 2012). However, without systematic spatial sampling from ground or remote-sensing observations of wood traits, we will not be able to prove but only suggest a regional and large-scale pattern.

We challenge the claim of Mitchard *et al.* (2014) that research plots provide accurate estimates of variations in wood density over Amazonia. To demonstrate this, we use a larger dataset ($n = 3616$) compiled over Amazonia using plots provided by Mitchard *et al.* (2014) and additional data from other sources (Nogueira *et al.*, 2005; Saatchi *et al.*, 2011; S. Brown, Winrock International, pers. comm.). Dividing the data over the same four regions suggested by Mitchard *et al.* (2014), we show that the within-region variations in wood density are larger than the between-region variations and the regional mean values (average wood density of individual trees in plots) are less divergent (Fig. S2). Our data, although not based on a systematic sampling of Amazonia, suggest that the wood density may have larger variations within landscapes than at regional scales because of the heterogeneity in forest composition, soil characteristics, geomorphology, size-dependent tree mortality and disturbance regimes, all functioning at small scales (metres to hectares).

RESEARCH PLOTS AND THE CURSE OF SAMPLING

The Mitchard *et al.* (2014) data may be more comprehensive within individual plots, yet they are several orders of magnitude less extensive across space. In other words, they are missing the forest for 0.00001% of the trees.

By referring to their plot network as an ‘inventory’, Mitchard *et al.* (2014) conflate measurement protocol in the field (i.e. wherein all trees are inventoried) with strategic planning to sample biomass and other forest properties as conducted by national forest inventories (e.g. McRoberts *et al.*,

2005). Their research plot network, although designed and used for ecological studies, is not suitable for biomass inventory because: (1) more than half of the plots are inherited from different groups, increasing the likelihood of measurement errors; (2) plots are spatially clustered near roads, rivers and research stations for easy access; and (3) the plots are haphazardly located, yet falsely depicted on the map to convey a widespread distribution over Amazonia. The coordinates have large uncertainty (c. 10–50 km) (Supporting Information in Mitchard *et al.*, 2014) because of lack of GPS recordings, particularly in older plots. Locations provided in their paper do not always match with data provided in RAINFOR publications or websites (Baker *et al.*, 2004; Malhi *et al.*, 2006; <http://www.rainfor.org/>). The comparison of the biomass estimates from the opportunistic plot collection that is only broadly constrained in space (at times as poorly as c. 10–50 km) and time (the years 1956–2013) leads to uncertainties ignored by Mitchard *et al.* (2014) when comparing with the satellite estimates that are tightly constrained (2005 ± 3 years and < 100 m for Geoscience Laser Altimeter System (GLAS) lidar observations; Lefsky, 2010).

To compensate for the sparse sampling of their plot collection, Mitchard *et al.* (2014) opted to average ‘field plots within $20 \text{ km} \times 20 \text{ km}$ boxes and compared the mean biomass values for these boxes to the mean AGB [aboveground biomass] of RS1 and RS2’. In the process, however, they committed several methodological errors.

First, they found 107 unique points ($20 \text{ km} \times 20 \text{ km}$) with an average of 3.9 (1–14) plots for each box. We could not reproduce the same number of unique points with their data. We found 109 unique points at $80 \text{ km} \times 80 \text{ km}$ (with a similar average of 3.8 plots per box) or 189 unique points for $5 \text{ km} \times 5 \text{ km}$ boxes (with an average of 2.1 plots). This difference has large implications for trend analysis and map comparison because 80-km boxes are one degree of magnitude larger than typical landscape scales (< 10 km).

Second, Mitchard *et al.* (2014) ignored the sampling problem and treated these average biomass values as the true mean for each box. The large spatial variability of AGB suggests that an average of at least 9–15 1-ha plots randomly located in the 20-km box areas are required to estimate the mean biomass at each point with 20% error (Chave *et al.*, 2003) (Fig. S3 shows examples of biomass spatial heterogeneity using airborne lidar data). Unlike Mitchard *et al.* (2014), we

addressed the sampling problem in our map (RS1) and provided the uncertainty of using five GLAS lidar shots (> 0.25 ha each) systematically sampling the 1-km map units in developing the RS1 biomass map (Saatchi *et al.*, 2011).

Third, we followed their approach and performed the trend analysis with unique points (109 or 189) derived from their research plots along three directions (N–S, E–W and NE–SW) using the ordinary least squares (OLS) regression model [$y = \alpha + X\beta + \epsilon$; with X being the explanatory variable (e.g. latitude or longitude or diagonal distance), y being AGB, and ϵ representing a geometric error term]. The reproduced results show that β has similar significance levels to Fig. 2 in Mitchard *et al.* (2014) (Table S1). However, if AGB can be fully explained by X using OLS, the residual of OLS regression should be white noise, otherwise any significance test based upon OLS is erroneous (Lennon, 2000). Our analysis shows that the OLS residual error is spatially correlated (Moran's I test in Table S1), confirming the existence of spatial autocorrelation (Fig. S4a), even after accounting for the changes in the proposed explanatory variable (Fig. S4b). Such spatial autocorrelation can be modelled as non-zero covariance in the regression residual, under the assumption of covariance stationarity (which is the same assumption underlying ordinary or universal kriging). The so-called geostatistical regression (GR) (Johnson & Hoeting, 2011) utilizing the generalized least squares (GLS) method shows that none of the trends provided by Mitchard *et al.* (2014) is significant under this approach (Seber & Lee, 2012). By ignoring the presence of spatial correlation in the data, they have effectively overestimated the number of samples, and reported artificially low P -values (Duffy *et al.*, 2007). The transformed residuals in the GR method are no longer dependent on distance, suggesting that the results of our GLS approach are valid (Fig. S4b).

Fourth, Mitchard *et al.* (2014) rely on interpolating 412 research plots over 650 million ha of Amazonian forests with large stretches (> 100 million ha) without a single plot. The performance of their interpolation is strongly dependent on the noise in the data, the spatial autocorrelation, the sampling pattern and the method of interpolation. By using a log–log axis for semivariogram analysis, Mitchard *et al.* (2014) misinterpreted the autocorrelation among plots. Using a linear axis, our analysis shows that spatial autocorrelation does exist

and extends to more than 2000 km with large variations at local scales (presence of a non-zero nugget) (Fig. S4a). Mitchard *et al.* (2014) consequently ignored any statistically rigorous kriging analysis and simply applied an inverse distance-based kernel function for spatial interpolation (Isaaks and Srivastava, 1989). In the presence of spatial autocorrelation the use of the inverse distance approach that ignores the uncertainty in the data has no sound statistical basis and can provide misleading interpolated surfaces (Zimmerman *et al.*, 1999). In addition, the performance of the interpolation also deteriorates significantly when sampling patterns are clustered instead of random, unless the sampling is designed initially to minimize the maximum or average kriging prediction-error variance (Brus & Heuvelink, 2007), which is not the case with the research plots. Without any optimized spatial sampling used in the forest inventory techniques (Mandallaz, 2007) the interpolated surfaces of the AGB offered by Mitchard *et al.* (2014) are erroneous and do not provide any further information than the original plots.

MISLEADING ASSESSMENT OF BIOMASS MAPS

Mitchard *et al.* (2014) claim to have documented a gradient in biomass from the south-west to the north-east in Amazonia. However, not only do they fail to rigorously prove this point, as demonstrated above, but they present a false narrative of the distribution of Amazonian biomass and what the maps present.

The literature on the spatial pattern of forest biomass in Amazonia is extremely diverse and sometimes contains contradictory results (Clark & Clark, 2000; DeWalt & Chave, 2004; Slik *et al.*, 2010; Quesada *et al.*, 2012). Forest biomass is a synthesis of several ecological and biological processes modulated with climate, soil and disturbance. If there is some non-randomness in these processes, then what are the controlling factors and at what scale do they operate? We showed that research plots, without any statistical sampling design, are not suitable for providing reliable answers to these important questions.

In the absence of rigorously designed (and extensive) ground-based forest inventories, remote-sensing techniques with spatially resolved, systematic and repeated measurements of forest structure are the best alternative to inventory sampling (Asner *et al.*, 2013; Neigh *et al.*, 2013). The GLAS lidar measure-

ments taken along the ICESAT orbital tracks provide systematic samples of forest structure that are three orders of magnitude denser than the opportunistic research plots (Fig. S5). Using a simple statistical aggregation of the height samples measured by GLAS lidar at 50 km × 50 km grids (with over a thousand samples in undisturbed forest pixels), we provide spatial patterns of forest structure over Amazonia (Fig. S5), showing the distribution of tall trees, potential gradients and large spatial variability.

We converted the GLAS lidar samples to biomass from a single equation derived from plots scattered in Amazonia close to lidar measurements with knowledge about the uncertainty of using a single allometry or average wood density (Saatchi *et al.*, 2011). However, using a single model based on canopy height over Amazonia had the advantage of not introducing any additional artificial and spatially correlated errors in the biomass map from unknown variations in wood density (Chave *et al.*, 2009). Nevertheless, contrary to the claims of Mitchard *et al.* (2014) the biomass map (RS1), with its regional bias, preserves the potential known patterns in Amazonia. To show this, we used a soil map of Amazonia (Fig. S6) as the basic stratification of forest types, calculated the mean forest height and biomass for each stratum and coloured the map to highlight regional patterns (Fig. S7).

IN CONCLUSION

Mitchard *et al.* (2014) have conducted an analysis with numerous and serious technical flaws. Instead, they rely on human psychology and create false impression that measures by touch (e.g. field plots) are superior to remotely sensed measures (from air and space). We know this is a flawed argument. Yet it is innately compelling because the human brain is wired to accept it. These instincts and arguments affected their methodology so that they erroneously used few plots for inferences about distant forests in time and place and comparison with satellite observations, that are extensive in space and constrained in time. By doing this, they made 'predictions, evaluations, and choices with respect to [their] construal of objects rather than the objects themselves'. (Lieberman and Trope, 2008).

ACKNOWLEDGEMENTS

This work was carried out at the Jet Propulsion Laboratory, California Institute of

Technology, and University of California Los Angeles, USA, under a contract by National Aeronautics and Space Administration (NASA).

SASSAN SAATCHI^{1,2*}, JOSEPH MASCARO³,
LIANG XU², MICHAEL KELLER⁴,
YAN YANG², PAUL DUFFY⁵,
FERNANDO ESPÍRITO-SANTO¹,
ALESSANDRO BACCINI⁶,
JEFFERY CHAMBERS⁷ AND DAVID SCHIMEL¹

¹NASA Jet Propulsion Laboratory, California Institute of Technology, Pasadena, CA 91109, USA,

²Institute of Environment and Sustainability, University of California, Los Angeles, CA 90095, USA,

³American Association for the Advancement of Science, 1300 New York Ave NW, Washington, DC 20001, USA,

⁴USDA Forest Service, International Institute of Tropical Forestry, San Juan, Puerto Rico,

⁵Neptune and Company, Inc. Lakewood, CO, 80215, USA,

⁶The Woods Hole Research Center, Falmouth, MA 02540, USA,

⁷Department of Geography, University of California, Berkeley, CA 94720, USA

*Correspondence: Sassan Saatchi, Jet Propulsion Laboratory, California Institute of Technology, 4800 Oak Grove Drive, Pasadena, CA 91109, USA.
E-mail: saatchi@jpl.nasa.gov
(http://carbon.jpl.nasa.gov/)

REFERENCES

- Asner, G.P. & Mascaró, J. (2014) Mapping tropical forest carbon: calibrating plot estimates to a simple LiDAR metric. *Remote Sensing of Environment*, **140**, 614–624.
- Asner, G.P., Mascaró, J., Anderson, C., Knapp, D.E., Martin, R.E., Kennedy-Bowdoin, T., van Breugel, M., Davies, S., Hall, J.S., Muller-Landau, H.C., Potvin, C., Sousa, W., Wright, J. & Bermingham, E. (2013) High-fidelity national carbon mapping for resource management and REDD+. *Carbon Balance and Management*, **8**(7), doi: 10.1186/1750-0680-8-7.
- Baccini, A., Goetz, S.J., Walker, W.S., Laporte, N.T., Sun, M., Sulla-Menashe, D., Hackler, J., Beck, P.S.A., Dubayah, R., Friedl, M.A., Samanta, S. & Houghton, R. (2012) Estimated carbon dioxide emissions from tropical deforestation improved by carbon-density maps. *Nature Climate Change*, **2**, 182–185.
- Baker, T.R., Phillips, O.L., Malhi, Y., Almeida, S., Arroyo, L., Di Fiore, A., Erwin, T., Killeen, T.J., Laurance, S.G., Laurance, W.F., Lewis, S.L., Lloyd, J., Monteagudo, A., Neill, D.A., Patiño, S., Pitman, N.C.A.M., Silva, J.N. & Vásquez Martínez, R. (2004) Variation in wood density determines spatial patterns in Amazonian forest biomass. *Global Change Biology*, **10**, 545–562.
- Brus, D.J. & Heuvelink, G. (2007) Optimization of sample patterns for universal kriging of environmental variables. *Geoderma*, **138**, 86–95.
- Chambers, J.Q., Asner, G.P., Morton, D.C., Anderson, L.O., Saatchi, S.S., Espírito-Santo, F.D., Palace, M. & Souza Jr, C. (2007) Regional ecosystem structure and function: ecological insights from remote sensing of tropical forests. *Trends in Ecology & Evolution*, **22**(8), 414–423.
- Chave, J., Condit, R., Lao, S., Caspersen, J.P., 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**(2), 240–252.
- Chave, J., Condit, R., Aguilar, S., Hernandez, A., Lao, S. & Perez, R. (2004) Error propagation and scaling for tropical forest biomass estimates. *Philosophical Transactions of the Royal Society B: Biological Sciences*, **359**, 409–420.
- Chave, J., Andalo, C., Brown, S., Cairns, M.A., Chambers, J.Q., Eamus, D., Fölster, H., Fromard, F., Higuchi, N., Kira, T., Lescure, J.P., Nelson, B.W., Ogawa, H., Puig, H., Riéra, B. & Yamakura, T. (2005) Tree allometry and improved estimation of carbon stocks and balance in tropical forests. *Oecologia*, **145**, 87–99.
- Chave, J., Coomes, D., Jansen, S., Lewis, S.L., Swenson, N.G. & Zanne, A.E. (2009) Towards a worldwide wood economics spectrum. *Ecology Letters*, **12**, 351–366.
- Chave, J., Réjou-Méchain, M., Búrquez, A. et al. (2014) Improved allometric models to estimate the aboveground biomass of tropical trees. *Global Change Biology*.
- 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.
- Clark, D.B. & Kellner, J.R. (2012) Tropical forest biomass estimation and the fallacy of misplaced concreteness. *Journal of Vegetation Science*, **23**(6), 1191–1196.
- DeWalt, S.J. & Chave, J. (2004) Structure and biomass of four lowland Neotropical forests. *Biotropica*, **36**, 7–19.
- Duffy, P.A., Epting, J., Graham, J.M., Rupp, T.S. & McGuire, A.D. (2007) Analysis of Alaskan burn severity patterns using remotely sensed data. *International Journal of Wildland Fire*, **16**, 277–284.
- Feldpausch, T.R., Lloyd, J., Lewis, S.L. et al. (2012) Tree height integrated into pantropical forest biomass estimates. *Biogeosciences*, **9**, 3381–3403.
- Hunter, M.O., Keller, M., Vitoria, D. & Morton, D.C. (2013) Tree height and tropical forest biomass estimation. *Biogeosciences Discussions*, **10**(6), 10491–10529.
- IBGE (1988) Mapa de vegetação do Brasil, escala 1:5,000,000, Fundação Instituto Brasileiro de Geografia e Estatística, Rio de Janeiro.
- Isaaks, E.H. & Srivastava, R.M. (1989) *Applied geostatistics*. Oxford University Press, Oxford.
- Johnson, D.S. & Hoeting, J.A. (2011) Bayesian multimodel inference for Geostatistical regression models. *PLoS ONE*, **6**, e25677.
- Kearsley, E., de Haulleville, T., Hufkens, K., Kidimbu, A., Toirambe, B., Baert, G., Huygens, D., Kebede, Y., Defourny, P., Bogaert, J., Beeckman, H., Steppe, K., Boeckx, P. & Verbeeck, H. (2013) Conventional tree height–diameter relationships significantly overestimate aboveground carbon stocks in the Central Congo Basin. *Nature Communications*, **4**.
- Lefsky, M.A. (2010) A global forest canopy height map from the Moderate Resolution Imaging Spectroradiometer and the Geoscience Laser Altimeter System. *Geophysical Research Letters*, **37**(15).
- Lennon, J.J. (2000) Red shifts and red herings in geospatial ecology. *Ecography*, **23**, 101–113.
- Liberman, N. & Trope, Y. (2008) The psychology of transcending the here and now. *Science*, **322**, 1201–1205.
- Malhi, Y., Wood, D., Baker, T.R. et al. (2006) The regional variation of aboveground live biomass in old-growth Amazonian forests. *Global Change Biology*, **12**, 1107–1138.
- Mandallaz, D. (2007) *Sampling techniques for forest inventories*. Chapman and Hall, Boca Raton, FL.
- Mascaró, J., Asner, G.P., Muller-Landau, H.C., van Breugel, M., Hall, J. & Dahlin, K. (2011) Controls over aboveground forest carbon density on Barro Colorado Island, Panama. *Biogeosciences*, **8**(6), 1615–1629.
- McRoberts, R.E., Holden, G.R., Nelson, M.D., Liknes, G.C. & Gormanson, D.D. (2005) Using satellite imagery as ancillary data for increasing the precision of estimates for the Forest Inventory and Analysis program of the USDA Forest Service. *Canadian Journal of Forest Research*, **35**(12), 2968–2980.

- Mitchard, E.T., Feldpausch, T.R., Brien, R.J. *et al.* (2014) Markedly divergent estimates of Amazon forest carbon density from ground plots and satellites. *Global Ecology and Biogeography*, **23**, 935–946.
- Muller-Landau, H.C. (2004) Interspecific and inter-site variation in wood specific gravity of tropical trees. *Biotropica*, **36**(1), 20–32.
- Neigh, C.S., Nelson, R.F., Ranson, K.J., Margolis, H.A., Montesano, P.M., Sun, G., Kharuke, V., Næsset, E., Wulder, M.A. & Andersen, H.E. (2013) Taking stock of circumboreal forest carbon with ground measurements, airborne and spaceborne LiDAR. *Remote Sensing of Environment*, **137**, 274–287.
- Nogueira, E.M., Nelson, B.W. & Fearnside, P.M. (2005) Wood density in dense forest in central Amazonia, Brazil. *Forest Ecology and Management*, **208**(1), 261–286.
- Quesada, C.A., Phillips, O.L., Schwarz, M. *et al.* (2012) Basinwide variations in Amazon forest structure and function are mediated by both soils and climate. *Biogeosciences*, **9**, 2203–2246.
- Saatchi, S.S., Houghton, R.A., Alvalá, R., Soares, J.V. & Yu, Y. (2007) Distribution of aboveground live biomass in the Amazon basin. *Global Change Biology*, **13**, 816–837.
- Saatchi, S.S., Harris, N.L., Brown, S., Lefsky, M., Mitchard, E.T.A., Salas, W., Zutta, B.R., Buermann, W., Lewis, S.L., Hagen, S., Petrova, S., White, L., Silman, M. & Morel, A. (2011) Benchmark map of forest carbon stocks in tropical regions across three continents. *Proceedings of the National Academy of Sciences of the United States of America*, **108**, 9899–9904.
- Espírito-Santo, F.D., Gloor, M., Keller, M. *et al.* (2014) Size and frequency of natural forest disturbances and the Amazon forest carbon balance. *Nature Communications*, **5**.
- Seber, G.A. & Lee, A.J. (2012) *Linear regression analysis*. John Wiley & Sons, Hoboken, NJ.
- Slik, J.W.F., Aiba, S.I., Brearley, F.Q., Cannon, C.H., Forshed, O., Kitayama, K., Nagamasu, H., Nilus, R., Payne, J., Paoli, G., Poulsen, A., Raes, N., Sheil, D., Sidiyasa, K., Suzuki, E. & van Valkenburg, J.L. (2010) Environmental correlates of tree biomass, basal area, wood specific gravity and stem density gradients in Borneo's tropical forests. *Global Ecology and Biogeography*, **19**, 50–60.
- ter Steege, H., Pitman, N.C., Phillips, O.L., Chave, J., Sabatier, D., Duque, A., Molino, J.F., Prévost, M.F., Spichiger, R., Castellanos, H., von Hildebrand, P. & Vásquez, R. (2006) Continental-scale patterns of canopy tree composition and function across Amazonia. *Nature*, **443**(7110), 444–447.
- Zimmerman, D., Pavlik, C., Ruggles, A. & Armstrong, M.P. (1999) An experimental comparison of ordinary and universal kriging and inverse distance weighting. *Mathematical Geology*, **31**, 375–390.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article at the publisher's web-site.

Figure S1 Uncertainty in ground estimates of biomass from regional versus local height–diameter allometry.

Figure S2 Distribution of wood density over four regions in Amazonia derived from a large plot-level data compiled from different sources.

Figure S3 Spatial heterogeneity of forest biomass at landscape scale represented by three sites of approximately 2.5 km × 2.5 km.

Figure S4 (a) Empirical semivariogram of average aboveground biomass derived from 189 unique observations. (b) Correlogram of ordinary least squares (OLS) regression residual and transformed residual of generalized least squares (GLS) for comparison of OLS and GLS models.

Figure S5 Systematic sampling of more than 1.2 million Geoscience Laser Altimeter System (GLAS) lidar shots over Amazonia. (a) GLAS sampling along the satellite orbits and clusters over time in a selected region of central Amazonia and location of research plots. (b) Top canopy height from GLAS lidar shots gridded at 0.5° × 0.5° showing variations of maximum top canopy height. (c) Mean top canopy height over the Amazon Basin.

Figure S6 Soil map of Amazonia derived from the soil and terrain database for Latin America and the Caribbean (SOTERLAC) and the Brazilian Institute of Geography and Statistics (IBGE) soil map of Brazil.

Figure S7 Distribution of 90th percentile and mean biomass in forests, aggregated based on soil types of Amazonia provided in Fig. S6.

Table S1 Regression tests using ordinary least squares and geostatistical regression on the 189 unique plots and 109 available unique points aggregated from Mitchard *et al.* provided plots.

Editor: Jeremy Kerr

doi: 10.1111/geb.12256