



Tropical forest biomass estimation and the fallacy of misplaced concreteness

David B. Clark & James R. Kellner

Keywords

Accuracy; Allometry; Costa Rica; Fallacy of misplaced concreteness; Forest biomass estimation; La Selva; Precision; Remote sensing; Tropical rain forest

Received 1 May 2012

Accepted 31 July 2012

Co-ordinating Editor: Michael Palmer

Abstract

Despite the importance of measuring tropical forest biomass, the accuracy of biomass estimates is poorly constrained due to fundamental weaknesses in the design and implementation of field studies. We identify these issues and propose a radical paradigm shift to advance tropical forest biomass research to a firmer theoretical and empirical basis.

Clark, D.B. (corresponding author, dbclark50@yahoo.com): Department of Biology, University of Missouri-St. Louis, St. Louis, MO, 63121, USA

Clark, D.B.: mail 1384 Lindenwood Grove, Colorado Springs, CO, 80907, USA

Kellner, J.M. (jkellner@umd.edu): Department of Geographical Sciences, University of Maryland, College Park, MD, 20742, USA

Introduction

Tropical forests store an estimated 193–229 Pg of carbon in above-ground biomass (Saatchi et al. 2011; Baccini et al. 2012), or roughly 20 times the annual emissions from combustion and land-use change (Friedlingstein et al. 2010). The quantity of biomass, usually measured in units of mass·ha⁻¹ (such as Pg or Mg·ha⁻¹) varies among continents, regions and landscapes because of climate, disturbance history, geochemistry and idiosyncratic site factors (Laurance et al. 1999; Baker et al. 2004; Letcher & Chazdon 2009). Major research efforts have been directed towards quantifying stocks of carbon in tropical forest (Asner et al. 2010; Saatchi et al. 2011; Baccini et al. 2012). Given the twin pressures of land-use change and deforestation in the humid tropics, the importance of such estimates is obvious. In this essay we will argue that in fact little progress is being made towards improving the accuracy of these estimates due to a series of fundamental weaknesses in the design and implementation of field studies. We first identify these issues, and then outline a radical paradigm shift to put tropical forest biomass research on a firmer theoretical and empirical basis.

The starting point for this discussion is the fact that no technique of remote sensing and no conventional forest inventory study directly measures biomass. Direct measurement requires physical comparison to an object traceable to an internationally accepted standard, providing an unambiguous and well-defined relationship' (Woodhouse et al. 2012). In the case of forest inventory studies, the most commonly measured variable is trunk diameter, which can be measured directly. In contrast, remote sensing instruments measure absorbed, transmitted or reflected electromagnetic radiation. These measurements may be correlated with biomass, but mass, the amount of 'stuff' in a given volume of matter, can only be directly measured by comparison to a known standard. How then do we obtain conventional estimates of forest biomass? Figure 1 illustrates the process. Measurements of tree diameter and/or estimates of height are converted into units of biomass using allometric scaling equations (ASEs; Brown 1997; Chave et al. 2005) and then summed to obtain a stand-level estimate that can be extrapolated to a landscape or continent (Asner et al. 2009, 2010; Lewis et al. 2009; Dubayah et al. 2010; Pan et al. 2011; Saatchi et al. 2011). We refer to this stand-level prediction as

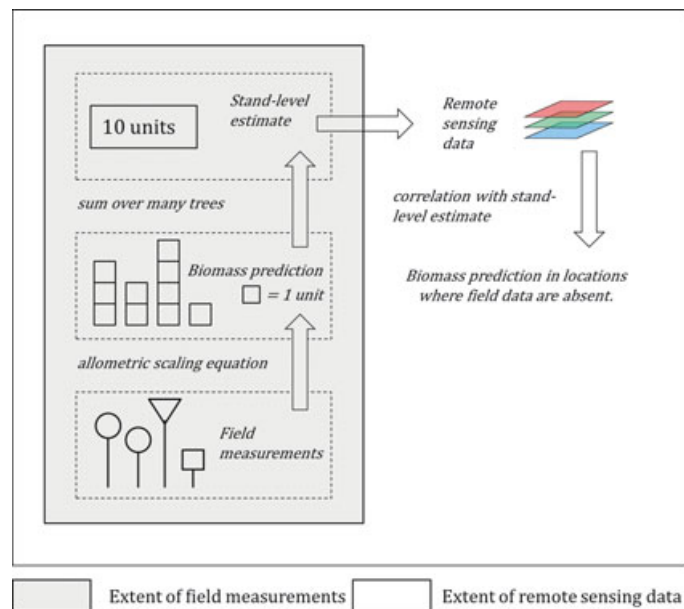


Fig. 1. Venn diagram showing a typical conceptual framework to estimate above-ground biomass in tropical forests. Field measurements of tree size within a fixed area (shown in grey) can be combined with allometric models to predict above-ground biomass. These predictions can be summed over many individual trees to obtain stand-level EAGB. Remote sensing variables such as canopy height or foliar chemical properties in the same fixed area can be regressed against EAGB, and the regression equation can be used to predict EAGB in locations with remote sensing data where field measurements are absent (area in white).

estimated above-ground biomass (EAGB), and to biomass quantified by direct harvest simply as biomass.

Pan-tropical estimates of forest biomass will necessarily incorporate data from remote sensing, since only remotely sensed data offer the potential to sample a significant fraction of the immense area of this biome. Using data from airborne or satellite systems, characteristics of forest structure or condition can be related to EAGB, and then extrapolated to large areas (Fig. 1). Because predicting EAGB using remote sensing begins with stand-level EAGB, the accuracy of remotely sensed predictions cannot surpass the accuracy of field measurements and ASEs used to derive EAGB. In contrast to these indirect procedures, plot-level biomass can be measured directly by harvesting vegetation and weighing the component parts. This is laborious and has been done in only a handful of studies that we are aware of (e.g. Araujo et al. 1999; Chambers et al. 2001).

We distinguish measurements from estimates. Measurements are obtained directly (e.g. tree diameter), whereas estimates are based on indirect approximations, usually involving calibrating equations. All estimates and measurements include sources of uncertainty, and a listing of all sources of uncertainty and their magnitudes constitutes an error budget. For both estimates and measurements, we will emphasize the distinction between accuracy and precision. Accuracy is the difference between a measurement

or estimate and 'truth'; precision is the variance among repeated measurements or estimates, irrespective of accuracy (Zar 1996). Assessment of both accuracy and precision requires replication. A technique could, for example, be highly accurate (mean of replicates close to the true value) but imprecise (high variance among replicates), or highly precise (low variance of replicates) but inaccurate (e.g. measured with an instrument calibrated with an incorrect standard). Note that evaluation of accuracy requires some method for directly assessing the variable being measured. For biomass in particular, direct measurements of mass are necessary to assess accuracy.

This distinction between accuracy and precision is fundamentally important. For many policy and management issues, a measure of forest biomass that is precise but of unknown accuracy may be sufficient. It is arguable, for example, that a remotely sensed metric of forest structure, one that should in general be correlated with forest carbon content, might be sufficient for the needs of monitoring of the sort envisioned by the Reduced Emissions from Deforestation and Degradation initiative (REDD+; UNFCCC 2012). For global carbon budgets and climate change science, however, accuracy is essential. Determining the impacts of tropical forest biomass carbon on global climate requires that we know accurately how much carbon is there. For example, projections of increases in atmospheric CO₂ due to tropical deforestation and temperature-induced

biome shifts are based on assumptions about absolute quantities of carbon in tropical forests. In this case accuracy of estimates is the issue, not precision.

There are at least three sources of uncertainty that are common to both measurements and estimates of biomass, and the extrapolation of these numbers to large areas. One is sampling design. The most common design is opportunistic (i.e. to sample near roads or rivers where it is cheapest), with unreplicated plots. Ideally, samples would be sited randomly, with replication, and stratified by environmental gradients that are thought or known to affect forest structure, such as climate, soil nutrients and disturbance history (Laurance et al. 1999, 2010; Chazdon 2003; Malhi et al. 2006). Inevitably, the total area in such intensively studied plots is miniscule relative to the total extent of tropical forest, and non-random plot locations may impact extrapolation of biomass estimates or measurements if locations are not representative of conditions throughout large areas. The magnitude of this effect has been vigorously debated and there is no clear consensus for tropical forests (Körner 2003; Fisher et al. 2008; Chambers et al. 2009; Lloyd et al. 2009; Rutishauser et al. 2010). Because tropical biomass estimates come from extrapolations far beyond the area represented by ground samples, biogeographic variation in allometric relations, wood density and soil fertility come into play, but these have yet to be systematically incorporated as sources of uncertainty.

Both field and remotely sensed estimates of biomass depend on ASEs to convert diameters, or diameters and heights, to biomass. Sources of uncertainty introduced through ASEs include species diversity and wood density in addition to the underlying strength of the ASE (Chave et al. 2004). For example, species diversity in the humid tropics is enormous, yet wood density is known for only a minority of species. Chave et al. (2009) significantly advanced our knowledge of wood density by compiling data for 8412 taxa, but this is dwarfed by the estimated total of 40 000–50 000 tropical tree species (Hamilton et al. 2010). Even if an ASE is developed using local taxa and environmental conditions, the relationship will always include residual error (i.e. the relationship between tree size and biomass does not have an $r^2 = 1$). Graphs of biomass against tree size show that biomass increases logarithmically with stem diameter (e.g. Brown 1997), and that the spread of points around the ASE regression line increases with tree size. This indicates that large-diameter trees can take on a wider range of biomass values than smaller trees. Because conventional approaches usually work with logarithms and maximum likelihood parameter estimates, this form of uncertainty is almost completely ignored (Mascaro et al. 2011 is an exception).

A third, and to date unmeasured, source of uncertainty is the possibility of positive bias in published ASEs due to

the harvest of trees with perfect form. In most cases there is insufficient documentation to determine whether harvested trees were selected at random. Given the massive amount of work involved, it seems to us parsimonious to assume in these cases that individuals were selected to have idealized form (i.e. entire crowns, cylindrical solid trunks and no obvious defects). Many trees in the real world are in fact not perfect and have a variety of conditions that decrease their biomass relative to perfect specimens (e.g. hollow trunks, missing major sections of crown and residual trunk irregularities due to prior damage). The magnitude of this potential bias has never been measured by plot-level harvest.

Finally, biomass allometries differ greatly among the life forms commonly assessed in forest inventories, and the data for trees are much more extensive than those for lianas (woody vines) and palms. Significant portions of biomass are normally not sampled at all in forest inventory studies (stems below a lower size class limit, epiphytes) and can only be estimated by fudge factors applied to stand-level data.

An additional source of uncertainty for estimating carbon stocks is the factor used to convert biomass to carbon. This factor is commonly assumed to be a constant between 0.48 and 0.50, but recent work suggests this assumption leads to a positive bias in carbon stock estimates (Martin & Thomas 2011).

Because remotely sensed estimates of biomass begin with plot-level EAGB, they incorporate all of the sources of uncertainty described above in addition to their own, instrument-specific errors. For example, light detection and ranging (LiDAR) estimates of canopy height depend completely on the ability to detect ground elevation. Reliably detecting ground elevation under closed-canopy tropical rain forest can be difficult (Dubayah et al. 2010), and independent measurements to validate derived terrain elevation maps are generally not available. Analyses that integrate multiple data types have their own distinct sources of error (e.g. Saatchi et al. 2011; Baccini et al. 2012).

Relating remotely sensed data to EAGB requires that the two data sources are co-registered. Uncertainty in the location of either data source introduces uncertainty into the relationship between EAGB and remotely sensed data. This is a significant issue in closed-canopy tropical forests, where obtaining differentially corrected GPS data for plot locations can be difficult or impossible. There are methods to assess the effects of geolocation error on the relation between field and remotely sensed data (e.g. Blair & Hofton 1999), but these are almost never used or reported.

In contrast to sources of uncertainty in EAGB, there are also sources of uncertainty in direct measurements of biomass by felling and weighing individuals in plots. Sawdust and attached vegetation like epiphytes must be collected

and weighed; a standard procedure for conversion of field weight to dry weight must be followed; procedures to handle trunks and crowns that are only partially within the plot must be implemented. Scales and balances should be calibrated to standard and traceable masses. It is clear that there is very little overlap in the components of error budgets of EAGB by forest inventory and/or remote sensing and the error budgets of direct measurements of forest biomass through stand-level harvest.

The core of the problem is the ground data: we do not measure biomass, so the accuracy of biomass estimates is not known. Researchers loosely refer to EAGB as 'biomass' simply because it is easier to estimate biomass than to measure it. In short, we do it because we can (see example in Clark & Clark 2000).

The central theme of this essay is that estimated forest biomass is not equal to measured forest biomass. As a scientific community we have repeatedly committed what has been termed 'the fallacy of misplaced concreteness' (Whitehead 1926), which is the error of 'mistaking the abstract for the concrete.' We confound the abstract (an estimate of biomass) with the concrete (a measurement of biomass).

To understand how fundamental this problem is, it is useful to consider the normal process of scientific research and how that could, and we argue should, be applied to measurement of forest biomass. Table 1 illustrates this sequence with ground data of known accuracy and remotely sensed data: (1) develop a model, in our case this could be an ASE relating some remotely sensed metric to EAGB; (2) gather data of known accuracy to test the model by harvesting individuals and directly measuring their size and biomass; (3) compare the predictions of the model to the data; (4) repeat (1) to (3) until the desired degrees of accuracy and precision are attained or the model is discarded or modified. To our knowledge this has never been done in any tropical forest, and consequently, the accuracy of all current tropical forest biomass estimates is unknown.

Table 1. Sequential steps to estimate forest biomass using remotely sensed data and harvested trees.

1. Select calibration plots and assess with remote sensing instruments
2. Measure biomass in calibration plots by total harvest
3. Develop a relationship between remotely-sensed data and harvested biomass
4. Select test plots and assess with remote sensing instruments
5. Predict biomass in test plots using relationship developed in (3)
6. Harvest test plots to measure biomass
7. Compare predicted biomass with measured biomass
8. Iterate to obtain desired level of accuracy and precision

A radical paradigm shift

Future studies should move beyond individual-based ASEs to focus on stand-level harvests that relate biomass directly to remotely sensed data. Plots should sample across the range of environmental gradients that are known or suspected to influence forest biomass, and plot sizes and orientations should be selected to facilitate integration with remote observations. All components of forest carbon should be harvested, including small stems, epiphytes, lianas and hemiepiphytes, as well as soil carbon. In short, to quantify and improve the accuracy of tropical forest biomass estimates, we must first measure biomass at the relevant spatial scale. Plots, not individual trees, are the link between field inventories and remotely sensed data. With plot-level harvests, we are no longer concerned about the problems of applying potentially biased ASEs, and all pools of carbon can be directly measured. If the environmental conditions of biomass harvest plots are carefully defined, non-random plot locations may become less of an issue, as long as plots are spread along the environmental gradients of interest. As plot-level harvest data accumulate it will become possible to quantitatively evaluate the magnitude of variation along different environmental axes. Knowledge of the extent of this variation can be used to prioritize future plot harvest locations and potentially to evaluate the importance of opportunistic vs random sampling.

Establishing a global database of biomass harvest plots with associated remotely sensed data will require a significant research commitment. Such a project, however, is neither technically difficult nor even particularly expensive, given the importance of accurately quantifying tropical forest carbon storage above and below ground. In our opinion, the main obstacle towards such a project is not strictly scientific. Cutting down patches of forest – chopping the vegetation into small pieces and determining volume and mass, drying thousands of pieces of wood in a battery of drying ovens – is laborious; there is nothing glamorous or high-tech about the research that is needed. Unless, however, we develop field measurements of plot-level biomass and its relation to various types of remote sensing, there will be scant empirical grounds for improving the absolute accuracy of the biomass estimates that make up the vast majority of publications on tropical forest biomass.

Remotely sensed estimates of above-ground forest biomass are key for understanding and managing the global carbon cycle and planetary ecology. The critical link between remote sensing estimates of forest biomass and measured forest biomass on the ground has yet to be made. It is curious in the extreme that to preserve and manage tropical forests in the face of multiple global

threats we are advocating cutting down a tiny bit more forest. But the simple fact is that without plot-level measurements of forest biomass linked to remotely sensed observations, we will never be able to objectively assess and improve the accuracy of tropical forest biomass estimates.

Acknowledgements

D.B. Clark was supported during this work by NSF/LTREB 0841872 and NASA TE08-0037. We thank Greg Asner, Deborah A. Clark, Michael Huston, Deborah Lawrence and Ralph Dubayah for critical reviews of earlier drafts of this manuscript, and Michael Palmer for encouragement to develop our ideas into this paper.

References

- Araujo, T.M., Higuchi, N. & Andrade de Carvalho, J. Jr 1999. Comparison of formulae for biomass content determination in a tropical rain forest site in the state of Para, Brazil. *Forest Ecology and Management* 117: 43–52.
- Asner, G.P., Hughes, R.F., Varga, T.A., Knapp, D.E. & Kennedy-Bowdoin, T. 2009. Environmental and biotic controls over aboveground biomass throughout a tropical rain forest. *Ecosystems* 12: 261–278.
- Asner, G.P., Powell, G.V.N., Mascaro, J., Knapp, D.E., Clark, J.K., Jacobson, J., Kennedy-Bowdoin, T., Balaji, A., Paez-Acosta, G., Victoria, E., Secada, L., Valqui, M. & Hughes, R.F. 2010. High-resolution forest carbon stocks and emissions in the Amazon. *Proceedings of the National Academy of Sciences, USA* 107: 16738–16742.
- 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.A. 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., Patino, S., Pitman, N.C.A., Silva, J.N.M. & Vasquez Martinez, R. 2004. Variation in wood density determines spatial patterns in Amazonian forest biomass. *Global Change Biology* 10: 545–562.
- Blair, J.B. & Hofton, M.A. 1999. Modeling laser altimeter return waveforms over complex vegetation using high-resolution elevation data. *Geophysical Research Letters* 26: 2509–2512.
- Brown, S. 1997. *Estimating biomass and biomass change of tropical forests: a primer*. Forestry paper 134. FAO, Rome, IT.
- Chambers, J.Q., DosSantos, J., Ribeiro, R.J. & Higuchi, N. 2001. Tree damage, allometric relationships, and above-ground net primary production in central Amazon forest. *Forest Ecology and Management* 152: 73–84.
- Chambers, J.Q., Negron-Juarez, R.I., Hurtt, G.C., Marra, D.M. & Higuchi, N. 2009. Lack of intermediate-scale disturbance data prevents robust extrapolation of plot-level tree mortality rates for old-growth tropical forests. *Ecology Letters* 12: E22–E25.
- 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 of London Series B* 359: 409–420.
- Chave, J., Andalo, C., Brown, S., Cairns, M.A., Chambers, J.Q., Eamus, D., Folster, H., Fromard, F., Higuchi, N., Kira, T., Lescure, J.P., Nelson, B.W., Ogawa, H., Puig, H., Riera, 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.
- Chazdon, R.L. 2003. Tropical forest recovery: legacies of human impact and natural disturbances. *Perspectives in Plant Ecology, Evolution and Systematics* 6: 51–71.
- 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.
- Dubayah, R.O., Sheldon, S.L., Clark, D.B., Hofton, M.A., Blair, J.B., Hurtt, G.C. & Chazdon, R.L. 2010. Estimation of tropical forest height and biomass dynamics using lidar remote sensing at La Selva, Costa Rica. *Journal of Geophysical Research-Biogeosciences* 115. doi:10.1029/2009JG000933.
- Fisher, J.L., Hurtt, G.C., Thomas, R.Q. & Chambers, J.Q. 2008. Clustered disturbances lead to bias in large-scale estimates based on forest sample plots. *Ecology Letters* 11: 554–563.
- Friedlingstein, P., Houghton, R.A., Marland, G., Hackler, J., Boden, T.A., Conway, T.J., Canadell, J.G., Raupach, M.R., Ciais, P. & Le Quéré, C. 2010. Update on CO₂ emissions. *Nature Geoscience* 3: 811–812.
- Hamilton, A.J., Basset, Y., Benke, K.K., Grimbacher, P.S., Miller, S.E., Novotny, V., Samuelson, G.A., Stork, N.E., Weiblen, G.D. & Yen, J.D.L. 2010. Quantifying uncertainty in estimation of tropical arthropod species richness. *American Naturalist* 176: 90–95.
- Körner, C. 2003. Slow in, rapid out – carbon flux studies and Kyoto targets. *Science* 300: 1242–1243.
- Laurance, W.F., Fearnside, P.M., Laurance, S.G., Delamonica, P., Lovejoy, T.E., Rankin-de-Merona, J.M., Chambers, J.Q. & Gascon, C. 1999. Relationship between soils and Amazon forest biomass: a landscape-scale study. *Forest Ecology and Management* 118: 127–138.
- Laurance, S.G.W., Laurance, W.F., Andrade, A., Fearnside, P.M., Harms, K.E., Vicentini, A. & Luizão, R.C.C. 2010. Influence of soils and topography on Amazonian tree diversity: a landscape-scale study. *Journal of Vegetation Science* 21: 96–106.
- Letcher, S.G. & Chazdon, R.L. 2009. Rapid recovery of biomass, species richness, and species composition in a forest

- chronosequence in Northeastern Costa Rica. *Biotropica* 41: 608–617.
- Lewis, S.L., Lopez-Gonzalez, G., Sonke, B., Affum-Baffoe, K., Baker, T.R., Ojo, L.O., Phillips, O.L., Reitsma, J.M., White, L., Comiskey, J.A., Djuiouo, K.M.-N., Ewango, C.E.N., Feldpausch, T.R., Hamilton, A.C., Gloor, M., Hart, T., Hladik, A., Lloyd, J., Lovett, J.C., Makana, J.-R., Malhi, Y., Mbago, F.M., Ndangalasi, H.J., Peacock, J., Peh, K.S.H., Sheil, D., Sunderland, T., Swaine, M.D., Taplin, J., Taylor, D., Thomas, S.C., Votere, R. & Woell, H. 2009. Increasing carbon storage in intact African tropical forests. *Nature* 457: 1003–U1003.
- Lloyd, J., Gloor, E.U. & Lewis, S. 2009. Are the dynamics of forest disturbance dominated by large, infrequent events? *Ecology Letters* 12: E19–E21.
- Malhi, Y.D., Wood Baker, T.R., Wright, J., Phillips, O.L., Cochrane, T., Meir, P., Chave, J., Almeida, S., Arroyo, L., Higuchi, N., Killeen, T.J., Laurance, S.G., Laurance, W.F., Lewis, S.L., Monteagudo, A., Neill, D.A., Núñez Vargas, P., Pitman, N.C.A., Quesada, C.A., Salomão, R., Silva, J.N.M., Torres Lezama, A., Terborgh, J., Vásquez Martínez, R. & Vinceti, B. 2006. The regional variation of aboveground live biomass in old-growth Amazonian forests. *Global Change Biology* 12: 1107–1138.
- Martin, A.R. & Thomas, S.C. 2011. A reassessment of carbon content in tropical trees. 2011. *PLoS One* 6:e23533.
- Mascaro, J., Litton, C., Hughes, R.F., Uowolo, A. & Schnitzer, S. 2011. Minimizing bias in biomass allometry: model selection and log-transformation of data. *Biotropica* 43: 649–653.
- Pan, Y.D., Birdsey, R.A., Fang, J.Y., Houghton, R., Kauppi, P.E., Kurz, W.A., Phillips, O.L., Shvidenko, A., Lewis, S.L., Canadell, J.G., Ciais, P., Jackson, R.B., Pacala, S.W., McGuire, A.D., Piao, S.L., Rautiainen, A., Sitch, S. & Hayes, D. 2011. A large and persistent carbon sink in the world's forests. *Science* 333: 988–993.
- Rutishauser, E., Wagner, F., Herault, B., Nicolini, E.-A. & Blanc, L. 2010. Contrasting above-ground biomass balance in a Neotropical rain forest. *Journal of Vegetation Science* 21: 672–682.
- Saatchi, S., Harris, N.L., Lefsky, M., Brown, S., 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, USA* 108: 9899–9904.
- UNFCCC (United Nations Framework Convention on Climate Change). 2012. http://unfccc.int/methods_science/redd/items/4531.php
- Whitehead, A.N. 1926. *Science and the modern world*. Cambridge University Press, Cambridge, UK.
- Woodhouse, I.H., Mitchard, E.T.A., Brolly, M., Maniatis, D. & Ryan, C.M. 2012. Radar backscatter is not a 'direct measure' of forest biomass. *Nature Climate Change* 11: 1–2.
- Zar, J.H. 1996. *Biostatistical analysis*, 3rd ed. Prentice Hall, Upper Saddle River, NJ, US.