The Linnean Society of London
BIOLOGICAL Journal of the Linnean Society

Biological Journal of the Linnean Society, 2014, **111**, 230–233. With 1 figure

COMMENT

Is logarithmic transformation necessary in allometry? Ten, one-hundred, *one-thousand-times yes*

JOSEPH MASCARO[1]*, CREIGHTON M. LITTON[2], R. FLINT HUGHES[3], AMANDA UOWOLO[3] and STEFAN A. SCHNITZER[4,5]

[1]*Department of Global Ecology, Carnegie Institution for Science, Stanford, CA 94305, USA*
[2]*Department of Natural Resources and Environmental Management, University of Hawai'i at Mânoa, Honolulu, HI 96822, USA*
[3]*Institute of Pacific Islands Forestry, USDA Forest Service, Hilo, HI 96720, USA*
[4]*Department of Biological Sciences, University of Wisconsin-Milwaukee, Milwaukee, WI 53211, USA*
[5]*Smithsonian Tropical Research Institute, Apartado 2072, Balboa, Republic of Panama*

Received 11 July 2013; revised 15 August 2013; accepted for publication 16 August 2013

In a recent analysis, Packard (2013) re-examined several allometric model-fitting techniques for *Metrosideros polymorpha* (Mascaro *et al*., 2011), a tropical tree endemic to Hawaii, asking: is logarithmic transformation necessary in allometry? Packard (2013) used three fitting techniques (Table 1) to arrive at the theoretically ideal power-law relationship between aboveground tree biomass and stem diameter:

$$Y = \mathrm{a}X^{\mathrm{b}}, \tag{1}$$

where Y is the aboveground tree biomass, X is the stem diamter, and a and b are constants to be estimated empirically. The power-law equation has been a stalwart in allometry for nearly a century (Huxley, 1932; Baskerville, 1972; Jenkins, Birdsey & Pan, 2001; Niklas, 2006). Packard (2013) concluded that 'the traditional allometric method is not well suited for fitting statistical models to data expressed in the arithmetic scale'. 'Traditional' in this context refers to linear fitting to logarithmically transformed biomass and diameter data, and back-transformation to a power-law form (see 'method 2' in Table 1). Contrasted were two nonlinear fitting techniques, the first assuming homoscedastic errors ('method 1') and the next assuming heteroscedastic errors ('method 3').

Packard sees the approach to allometry described by Mascaro *et al*. (2011) as misguided. 'Where did Mascaro *et al*. [and others] go wrong?' he asks. Packard has previously suggested that the logarithmic transformation leads to biased results (Packard & Boardman, 2008; but see Kerkhoff & Enquist, 2009). In the current article, however, Packard seems determined to banish the logarithm from allometry. Readers of his current paper are left with the tacit impression that Mascaro *et al*. (2011) considered only the time-tested traditional method. In fact, we considered the exact same three power-law fitting techniques that Packard employs, comparing the performance of each method against one another. We came to very different conclusions, however, and to understand this friendly disagreement we must revisit Packard's most recent analysis.

Packard begins by applying nonlinear fitting of the power-law model assuming homoscedastic error structure ('method 1' in Table 1), noting that 'the mean function generally follows the path of the observations, albeit the line departs slightly from that path for plants with a stem diameter of 8–12 cm'; however, this is not the case. The model is biased for all but one

*Corresponding author. E-mail: jmascaro@stanford.edu

 © 2013 The Linnean Society of London,

Table 1. Fitting techniques used to produce power-law models relating plant biomass and stem diameter for *Metrosideros polymorpha* by Mascaro *et al*. (2011), and subsequently reanalysed by Packard (2013)

Method 1. Nonlinear fitting using the assumption of homoscedastic errors (i.e. the default of most statistical packages commonly used by ecologists and foresters, including R, SAS, and JMP):

$$Y_i = \mathbf{a}X_i^{\mathbf{b}} + \varepsilon_i \qquad \varepsilon_I \sim N(0,\ \boldsymbol{\sigma}^2)$$

Method 2. Linear fitting to logarithmically transformed biomass and diameter data, followed by back-transformation of the fitted model. In this case, the arithmetic error in the logarithms must be adjusted via a correction factor [$e^{(\mathrm{MSE}/2)}$], where MSE is the mean squared error of the fitted linear model (*sensu Baskerville, 1972*), a step that Packard omits:

$$\ln(Y_\mathrm{i}) = \ln(\mathbf{a}) + \mathbf{b}\ln(X_\mathrm{i}) + \varepsilon_i \quad \varepsilon_I \sim N(0,\ \boldsymbol{\sigma}^2)$$

where back-transforming gives the equation: $Y = \mathbf{a}X^{\mathbf{b}} * \exp(\mathrm{MSE}/2)$

or $Y_i = \mathbf{a}X_i^{\mathbf{b}} * \exp(\varepsilon_i) \qquad \varepsilon_I \sim N(0,\ \boldsymbol{\sigma}^2)$

Method 3. Nonlinear fitting using the assumption of heteroscedastic errors:

$$Y_i = \mathbf{a}X_i^{\mathbf{b}} + \varepsilon_i \qquad \varepsilon_i \sim N[0,\ (\boldsymbol{\sigma}^2 * X_i^{\mathbf{k}})]$$

tree < 14 cm diameter at breast height, reaching a bias of 400%, as we showed originally for these small-diameter individuals. Packard cannot observe the bias because he examines only the arithmetic scale (Fig. 1a–d).

Packard then presents what he calls a 'back-transformed OLS' model as the 'traditional method'; however, he plots the model without applying the correction factor for back-transformation of the regression error (*sensu* Baskerville, 1972). Plotting the uncorrected model (shown by the dashed line in Fig. 1) and implying that this is what we originally presented (shown by the solid line in Fig. 1) misrepresents our original article. Packard also considers generalized linear modelling as an alternative to the correction-factor approach (shown by the dotted line in Fig. 1), but this too is a poorer fit than the back-transformed model that we originally presented. Once the correction factor is applied, as in recent practice (Chave *et al*., 2005; Schnitzer, DeWalt & Chave, 2006), the model produces much better predictions (Fig. 1e–h).

Next, Packard plots the (incorrectly constructed) traditional method on the geometric scale, arguing that it 'does not follow the path of the observations', but fails to do so for the other methods. When the models are properly compared, as in Figure 1 here, it becomes clear that methods 2 and 3 fit the data much better than method 1. This is impossible to determine from Packard's analysis because panels C, D, K, and L are absent from his analysis.

Most disappointingly, Packard ignores the central conclusion of our original article that nonlinear fitting with the assumption of homoscedastic errors may lead to biases in excess of 100%, particularly for small-diameter individuals (which have low leverage in the fitting routine), and that nonlinear fitting assuming heteroscedastic errors can mitigate this problem. Our article was the first published use of method 3 in plant allometry that we know of, and we even noted specifically that this method 'may be more reliable' than the traditional method.

Packard's last point deserves added attention. Looking at the data in geometric space, Packard argues that the curvilinearity obviates logarithmic transformation because the 'transformation failed to linearize the distribution'. Indeed, the data are curvilinear in geometric space. We did not discuss this issue in our original article, and we appreciate Packard calling attention to it; however, Packard jumps to a conclusion that is favourable with his argument without looking at the whole picture. Power-law equations, irrespective of model-fitting technique, are linear in geometric space. Thus, the key question is ecological rather than statistical in this case: why do the *Metrosideros* data not perfectly follow theoretical allometric scaling at very small diameters?

Metrosideros was among several trees that displayed a small degree of curvilinearity in geometric space, in all cases very near the lowest diameter range sampled (i.e. 0.2–3.0 cm in diameter, depending on the species; Mascaro *et al*., 2011: fig. 1). The likely explanation for this curvilinearity is that consistent forest sampling protocol is not constrained to perfect allometric scaling. For both general forest sampling and biomass harvests, the position of the standard point-of-measurement for diameter at breast height is typically 1.3 m from the ground, with exceptions for malformations such as buttresses (e.g. '1.3 m or above buttress'). With shorter and shorter trees, this position moves closer to the meristem of the tree, resulting in a lower diameter estimate than what would be expected around the 'bole' of very small trees. This is easy to see in the abstract: before a tree reaches 1.3 m in height and enters a field census or harvest data set, it effectively has an apparent diameter of zero and positive biomass. This departure from the power law in no way contradicts power-law allometric

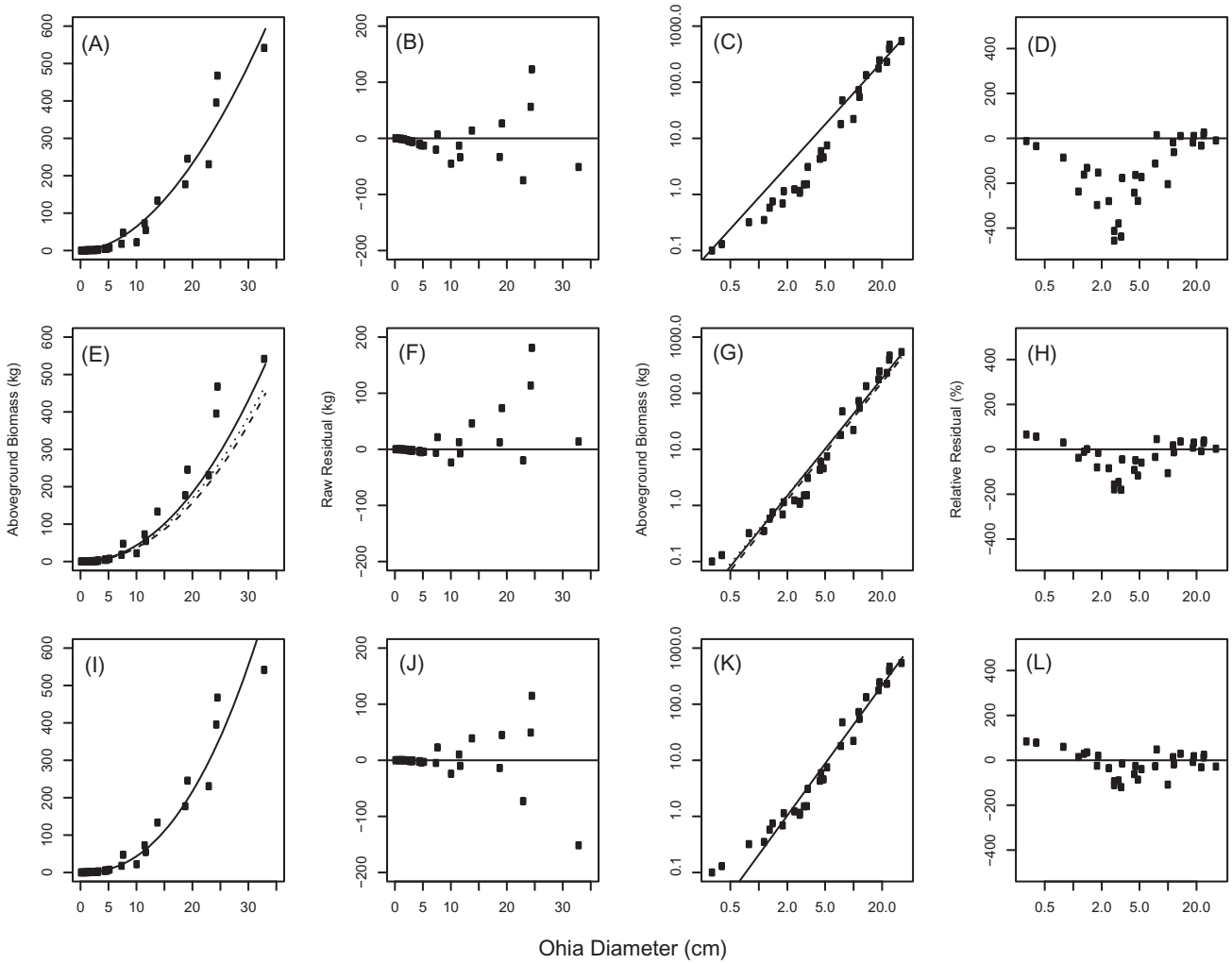

Figure 1. Three fitting techniques for the power-law model architecture of *Metrosideros polymorpha*. A–D, nonlinear fitting to the untransformed data with the assumption of homoscedastic errors ('method 1' in Table 1), which Packard (2013) argued 'followed the path of the observations'. E–H, linear fitting to the log-transformed diameter and biomass data, followed by back-transformation ('method 2' or the traditional method). The relative residual reflects the fraction (percentage) of the residual compared with the observed value. The dashed line represents the equation Packard (2013) plotted for this fitting technique, for which he excluded the requisite correction factor (i.e. Baskerville, 1972), which is necessary to properly back-transform the error structure with this fitting method. The dotted line represents a generalized linear modelling result offered by Packard as a substitute for the correction factor technique that we originally used. I–L, nonlinear fitting to the untransformed data with the assumption of heteroscedastic errors ('method 3'). Mascaro *et al*. (2011) considered all three fitting techniques and concluded that methods 2 and 3 (second and third rows) were preferred within the confines of power-law architecture.

scaling, however. Instead, it reflects a decoupling of the measurement of the 'apparent' diameter of a tree from its 'real' diameter. For purposes of scaling theory, the allometrically constrained diameter of interest will be closer to the ground for very small trees.

An analogy can be made to curvilinearity in geometric space detected by Chave *et al*. (2005) and Muller-Landau *et al*. (2006) in biomass–diameter relationships for the largest tropical trees. At very high diameters (and particularly with old age), tropical trees experience crown breakage, bole rot, and other bruising associated with longevity. Most ailments result in lower tree biomass, but little if any effect on apparent diameter, and thus biomass begins to decline relative to the power-law expectation. Again, allometric scaling theory has not been broken in this case. Rather, the 'apparent' diameter has been decoupled from the 'real' diameter: within the tree's bole, the bundle of xylem cells that governs water flow (and thus biomass) is also in decline.

Thus many, if not all, trees harvested from the seedling stage to very old age will be slightly sigmoidal in their biomass–diameter relationships, or curvilinear in geometric space. Contrary to Packard's implication, curvilinearity in geometric space can be an ecological and methodological phenomenon, and may have nothing to do with logarithmic transformation. In practice, curvilinearity in geometric space can be dealt with by bisecting the data to create separate models for smaller and larger trees (via methods 2 or 3), or applying polynomial models using logarithmically transformed data (requiring a correction factor, e.g. Chave *et al*., 2005). As direct nonlinear fitting with non-arithmetic errors is increasingly used in allometry or allometry-type problems (e.g. Asner *et al*., 2012), alternative sigmoidal model architectures such as the Weibull may be fitted with heteroscedastic errors. Even in these cases, the fundamental issue is not whether one uses the logarithmic transformation in model fitting, but whether one examines allometric scaling in geometric space (Glazier, 2013). The logarithmic transformation remains an obvious and reasonable tool in this effort. Packard (2013) used the logarithmic transformation for this exact purpose. In the end, he validated its role in allometry.

ACKNOWLEDGEMENTS

We thank G. C. Packard for advocating open access to allometry data that are exceedingly difficult to collect. We thank M. S. Colgan and L. M. Hallett, and three anonymous reviewers, for comments on a previous draft of this article.

REFERENCES

Asner GP, Mascaro J, Muller-Landau HC, Vieilledent G, Vaudry R, Rasamoelina M, Hall JS, van Breugel M. 2012. A universal airborne LiDAR approach for tropical forest carbon mapping. *Oecologia* **168:** 1147–1160.

Baskerville G. 1972. Use of logarithmic regression in the estimation of plant biomass. *Canadian Journal of Forest Research* **2:** 49–53.

Chave J, Andalo C, Brown S, Cairns MA, Chambers JQ, Eamus D, Folster H, Fromard F, Higuchi N, Kira T, Lescure J-P, Nelson BW, 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.

Glazier DS. 2013. Log-transformation is useful for examining proportional relationships in allometric scaling. *Journal of Theoretical Biology* **334:** 200–203.

Huxley JS. 1932. *Problems of relative growth*. New York City: Dial Press.

Jenkins JC, Birdsey RA, Pan Y. 2001. Biomass and NPP estimation for the Mid-Atlantic Region (USA) using plot-level forest inventory data. *Ecological Applications* **11:** 1174–1193.

Kerkhoff AJ, Enquist BJ. 2009. Multiplicative by nature: why logarithmic transformation is necessary in allometry. *Journal of Theoretical Biology* **257:** 519–521.

Mascaro J, Litton CM, Hughes FR, Uowolo A, Schnitzer SA. 2011. Minimizing bias in biomass allometry: model selection and log-transformation of data. *Biotropica* **43:** 649–653.

Muller-Landau HC, Condit RS, Chave J, Thomas SC, Bohlman SA, Bunyavejchewin S, Davies S, Foster R, Gunatilleke S, Gunatilleke N, Harms KE, Hart T, Hubbell SP, Itoh A, Kassim AR, LaFrankie JV, Lee HS, Losos E, Makana JR, Ohkubo T, Sukumar R, Sun IF, Supardi NMN, Tan S, Thompson J, Valencia R, Munoz GV, Wills C, Yamakura T, Chuyong G, Dattaraja HS, Esufali S, Hall P, Hernandez C, Kenfack D, Kiratiprayoon S, Suresh HS, Thomas D, Vallejo MI, Ashton P. 2006. Testing metabolic ecology theory for allometric scaling of tree size, growth and mortality in tropical forests. *Ecology Letters* **9:** 575–588.

Niklas KJ. 2006. A phyletic perspective on the allometry of plant biomass-partitioning patterns and functionally equivalent organ-categories. *New Phytologist* **171:** 27–40.

Packard GC. 2013. Is logarithmic transformation necessary in allometry? *Biological Journal of the Linnean Society* **109:** 476–486.

Packard GC, Boardman TJ. 2008. Model selection and logarithmic transformation in allometric analysis. *Physiological and Biochemical Zoology* **81:** 496–507.

Schnitzer SA, DeWalt SJ, Chave J. 2006. Censusing and measuring lianas: a quantitative comparison of the common methods. *Biotropica* **38:** 581–591.