Journal of Theoretical Biology 257 (2009) 519–521

Contents lists available at ScienceDirect

Journal of Theoretical Biology

journal homepage: www.elsevier.com/locate/yjtbi

Letter to Editor

Multiplicative by nature: Why logarithmic transformation is necessary in allometry

While we agree with Packard (2008) on the importance of selecting appropriate statistical models of allometric scaling relationships, we strongly disagree with his conclusion that standard methods for fitting allometric models produce results that are "biased and misleading". His argument, which is elaborated in a series of related publications (Packard and Boardman, 2008; Packard and Birchard, 2008) is based on questionable and misleading assumptions. In short, he reaches the wrong conclusion about log-transformation in allometry because most biological phenomena, including growth, reproduction, sensation, and metabolism are inherently multiplicative, rather than additive, and it is proportional rather than absolute variation that matters, especially across the orders of magnitude spanned by most allometric analyses.

Our difference with Packard is based on the conflict between arithmetic and geometric (or multiplicative) assumptions about biological variation (Galton, 1879; Gingerich, 2000). Packard (2008) correctly shows that using standard regression methods on log-transformed data makes the implicit assumption that errors are multiplicative rather than additive. This standard approach in allometry is often seen as introducing bias (Zar, 1968; Smith, 1993), because the resulting parameters do not actually minimize the arithmetic deviations between the original, untransformed data and the model. That is, standard allometric models actually predict the geometric (rather than arithmetic) mean response. However, in biology (and many other physical and social sciences) the assumed multiplicative error is often the appropriate model of variation. As eloquently pointed out by Galton (1879), the normal, additive error model, "… asserts that the existence of giants, whose height is more than double the mean height of their race, implies the possibility of dwarfs, whose stature is less than nothing at all". In contrast, the multiplicative error model asserts that double-sized giants occur as frequently as half-sized (rather than zero-sized) dwarves. That is, where the additive error model assumes that equivalent deviations differ by equal *amounts*, the multiplicative error model assumes that they differ by equal *proportions* (Gingerich, 2000). We argue here that the multiplicative nature of biological processes and the focus of allometric study on phenomena that span orders of magnitude, make the multiplicative error model an appropriate feature, rather than a defect, of standard allometric analyses.

As was emphasized by Julian Huxley's pioneering work (Huxley, 1932) allometry is the study of *relative* biological variation. Packard asserts that log-transformation is inappropriate for allometry because it distorts the relationship between the variables, overweights observations of small magnitude, and hides outliers. In fact, the log-transformation is used in allometry specifically to uniformly depict relative variation in the relation between variables, to avoid overweighting large-magnitude observations, and to normalize sample variance and reduce the influence of outliers (Huxley, 1932; Peters, 1983; Calder, 1984; Niklas, 1994). Thus, the supposed weaknesses Packard sees in log-transformation are the very strengths that necessitate its use. To understand why Packard's assertion is problematic, consider the observation by William Calder (1984) that ordinary measurements of metabolic rate vary by about 20% in mammals from shrews and elephants, corresponding to about 0.175 units on a (base-10) logarithmic scale. That is, shrew metabolic rate varies by only about 0.035 W, while elephant metabolic rate can vary by 284 W. Thus, despite the fact that variation in both animals is proportionally the same, elephants appear to be much more variable on an arithmetic scale. For an elephant, a 1 W variation in metabolic rate is miniscule, a barely detectable noise. In contrast, for a shrew, a single-watt change in metabolic rate is more than a 5-fold adjustment, dwarfing its ordinary metabolic rate. Assuming that the absolute variation observed in elephants applies to shrews, or *vice versa*, is clearly absurd, yet this is the implicit assumption of applying an additive error model on an arithmetic scale, whether using a linear or non-linear fit. For data spanning orders of magnitude, logarithmic scales allow us to assess relative variability in a meaningful way, and linear regression on log-transformed variables ensures that *relative* residual variation is treated uniformly across scales in the estimation of parameters.

If we discard the assumption that the arithmetic scale (and additive error model) have some sort of conceptual priority (see below), we can use relative variation as a basis for comparing the standard allometric model to the alternatives presented by Packard in his examples. He first examines the eye mass–brain mass allometry in 24 species of terrestrial Carnivora (Burton, 2006), concluding that log-transformation hides the true nature of the bears as outliers. However, for the reasons detailed above, Packard's method of visual inspection on an arithmetic plot is clearly biased toward identifying points at the large end of the size spectrum; they are (arithmetically) more variable. Dealing with outliers is always a tricky business, but given the two order of magnitude range of brain sizes (from ~5 to 500 g) and the relatively small number of observations, we need a more reliable method than simple visual inspection, especially since the bear species have the two largest brains in the dataset. Preferably, we would like to see some biologically relevant reason for excluding the bears. Packard offers none, so we are left to compare a model with bears to one without.

So how do the models compare biologically? The standard allometric model tells us that the two bear species do indeed have relatively small eyes. Given their brain size, we would expect them to be substantially (~4 times) heavier, similar to the observed eye mass of a hyena or a leopard (Burton, 2006). This conclusion leads to additional biological questions. For example, is relative eye size related to feeding strategy or phylogenetic affinity? Indeed, from these data, the six species within the

0022-5193/$ - see front matter
doi:10.1016/j.jtbi.2008.12.026

monophyletic Arctoidea (the two bears, raccoon, coati, kinkajou, and weasel) all have eyes that are smaller than expected from the standard allometric model. In contrast, among the other 18 species, which are from the Canidae (which is sister to the Arctoidea, (Flynn et al., 2005) or the more distant Feliformia, only three species (two foxes and a cougar) have negative residuals. Analysis of covariance suggests that the Arctoidea do indeed have a lower scaling coefficient ($F = 42.8$, $p < 10^{-5}$) but share common exponent ($F = 0.30$, $p = 0.59$) with the other clades. Thus, it appears that the residual variation around the model may carry a phylogenetic signal. Of course, establishing the importance of this signal requires further analysis (and probably more data), but the hypothesis itself represents utility for the standard allometric model.

According to Packard's model based on absolute, additive error, bears truly would be outliers. We would expect their eyes to be more than ten times larger, between 50 and 70 g, or about the size of an elephant's. Thus bears are simply "different" and can only be eliminated from the analysis, even though doing so reduces the size range of our data and eliminates $\frac{1}{3}$ of the Arctoidea from an already small and phylogenetically spotty dataset. Furthermore, while three of the four remaining Arctoidea still have eyes that are smaller than expected, they are joined by six other species of Canidae and Feliformia. Thus, the phylognetic signal suggested by the standard model is effectively eliminated along with the bears. We should also point out that the eyes of domestic cats are ~4 times heavier than expected from Packard's model. That is, in relative terms, their deviation from Packard's model is similar in magnitude to the deviation of the bear species from the standard allometric model. Yet domestic cats would never be construed as outliers by Packard's method, simply because they are relatively small. By focusing on the arithmetic domain and eliminating biologically valid data, Packard mischaracterizes the pattern of covariation between brains and eyes and obscures potentially important phylogenetic signals in the comparative biology of the Carnivora. In contrast, log-transformation demonstrates that *relative* variability in eye size is homoscedastic across orders of magnitude in brain size (Packard, 2008, Fig. 1A) and that although bears and their relatives have small eyes, they are in fact still comparable to other carnivores.

Packard's second example focuses on the basal metabolic rate in mustelids (Muñoz-Garcia and Williams, 2005) and especially highlights how the scale-sensitivity of arithmetic plots can hide important variation at small magnitudes. Packard shows that the standard allometric function provides a poor prediction for the largest species. Indeed, the prediction is about 5000 kJ/d too low for ~30 kg sea otter, about a 2-fold deviation. On the other hand, the nonlinear fit (which assumes an additive error model) is right on for the otter. Indeed, because of its large-size, the otter has a very strong influence on the additive model. If the otter is withheld from the analysis, a nonlinear fit with additive errors results in the model $Y = 1.58X^{0.773}$, which is much closer to the standard allometric model derived from the all of the data, $Y = 2.29X^{0.736}$ than it is to Packard's fit, $Y = 0.03X^{1.470}$. This disproportionate influence does not mean that the otter should be removed from the data, it simply points to the need to change the statistical model to one that considers relative rather than absolute errors—as in the standard allometric analysis. Furthermore, by examining the additive fit on the logarithmic plot (Packard, 2008, Fig. 2A), it becomes clear that while the arithmetic deviations from the Packard's model are indeed small, estimates of basal metabolic rate for the three smallest species of mustelid are all approximately 100-fold too low! In contrast, the data are all within about 2-fold of the "biased and misleading" standard allometric model. If the desire is to model variation in metabolic rate across *all* of the species, and not just for the largest one, we must reassess where the bias lay.

Packard's conclusions depend on the assumption that additive variation should always be the default standard for parameter estimation and data exploration. We have shown that this assumption is especially dangerous when examining patterns across many orders of magnitude, because the same relative deviation (e.g. 10% or 5-fold) will be a much smaller absolute (arithmetic) deviation for small observations than large observations. That is, for biological processes that span orders of magnitude, variation in the arithmetic domain is *scale-dependent*, and using an additive error model, whether with a linear or nonlinear fit, will be biased in its treatment of *relative* deviations. While the biological relevance of proportional, rather than absolute variation was pointed out almost 130 years ago (Galton, 1879), formal comparisons of arithmetic and geometric error models are almost non-existent (Gingerich, 2000). However, while the two models are difficult to distinguish when the variance is small relative to the mean, recent empirical and theoretical work suggests that geometric error should often be the default in biology (Gingerich, 2000; Graham et al., 2003). In any case, the voluminous literature full of homoscedastic logarithmic plots, including the papers from which Packard drew his examples (Burton, 2006; Muñoz-Garcia and Williams, 2005), underscores the general utility of a multiplicative error model for many allometric and other biological regression problems.

The default status of the additive error model is perhaps based on the deeper assumption that arithmetic scales are somehow truer or more intuitive, and that log-transformation thus represents a distortion of the "real" data. However, recent research strongly challenges the assumption that arithmetic scales are more natural, finding that the initial intuition in humans across cultures is to scale numbers geometrically, especially across orders of magnitude (Dahaene et al., 2008). Interestingly, Gingerich (2000) points out that our evolved senses (e.g. eyes, ears) measure the world geometrically, based on proportional changes in stimuli (Fechner, 1860), while our constructed measurement devices (e.g. rulers, balances) generally do so arithmetically, like counting on our fingers. Thus, our deliberate, arithmetic measurements may be incomplete, requiring log-transformation to make them into proportional similarities and differences sensible to our evolved brains (Gingerich, 2000). At the very least, further formal comparisons of geometric and arithmetic error models are warranted.

Log-transformation is not simply a statistical convenience. It is indeed a non-linear transformation, but it places numbers into a geometric domain in which proportional deviations are represented consistently, independent of the scale and units of measurement. But more importantly, it is often appropriate in biology because many biological phenomena (e.g. growth, reproduction, metabolism, sensation) are fundamentally multiplicative, and likely conform more closely to a geometric error model (Galton, 1879; Gingerich, 2000). Packard is right that no statistical tools should be applied uncritically, and we agree that alternative scaling models should be considered and used when the data demand it. Further, biologists need to have a deeper quantitative understanding of logarithms and be aware, for example, that standard allometric methods predict the *geometric* mean response for a given value of the predictor variable. However, we should not try to force a geometric biological world into an arithmetic box simply because we learn to count on our fingers. Log-transformation is entirely appropriate, indeed necessary, for allometric analysis and many other problems in biology.

Acknowledgements

We would like to thank Gary Packard for inspiring us to finally construct a defence of log-transformation in allometry, a subject we have discussed for some time. We also acknowledge the opportunity provided by the journal editors to respond on this issue. Brad Hartlaub provided helpful comments and statistical advice, and four anonymous reviewers provided constructive feedback. AJK was supported by NSF grant DMS-0827208, and BJE was supported by an NSF ATB award.

References

Burton, R.F., 2006. A new look at the scaling of size in mammalian eyes. J. Zool. 269, 225–232.
Calder, W.A., 1984. Size, Function, and Life History. Harvard University Press, Cambridge.
Dahaene, S., Izard, V., Spelke, E., Pica, P., 2008. Log or linear? Distinct intuitions of the number scale in Western and Amazonian Indigene cultures. Science 320, 1217–1220.
Fechner, G.T., 1860. Elemente Der Psychophysik (Zweite Unveränderte Auflage), Erster Theil. Leipzig: Breitkopf und Härtel.
Flynn, J.J., Finarelli, J.A., Zehr, S., Hsu, J., Nedbal, M.A., 2005. Molecular phylogeny of the Carnivora (Mammalia): assessing the impact of increased sampling on resolving enigmatic relationships. Syst. Biol. 54, 317–337.
Galton, F., 1879. The geometric mean in vital and social statistics. Proc. R. Soc. London 29, 365–367.
Gingerich, P.D., 2000. Arithmetic or geometric normality of biological variation: an empirical test of theory. J. Theor. Biol. 204, 201–221.
Graham, J.H., Shumuzu, K., Emlen, J.E., Freeman, D.C., Merkel, J., 2003. Growth models and the expected distribution of fluctuating asymmetry. Biol. J. Linn. Soc. 80, 57–65.
Huxley, J.S., 1932. Problems of Relative Growth. Dial Press, New York.
Muñoz-Garcia, A., Williams, J.B., 2005. Basal metabolic rate in carnivores is associated with diet after controlling for phylogeny. Physiol. Biochem. Zool. 78, 1039–1056.
Niklas, K.J., 1994. Plant Allometry: The Scaling of Form and Process. University of Chicago Press, Chicago.
Packard, G.C., Boardman, T.J., 2008. Model selection and logarithmic transformation in allometric analysis. Physiol. Biochem. Zool. 81, 496–507.
Packard, G.C., Birchard, G.F., 2008. Traditional allometric analysis fails to provide a valid predictive model of mammalian metabolic rates. J. Exp. Biol. 211, 3581–3587.
Packard, G.C., 2008. On the use of logarithmic transformations in allometric analyses. J. Theor. Biol., in press, doi:10.1016/j.jtbi.2008.10.016.
Peters, R.H., 1983. The Ecological Implications of Body Size. Cambridge University Press, Cambridge.
Smith, R.J., 1993. Logarithmic transformation bias in allometry. Am. J. Phys. Anthropol. 90, 215–228.
Zar, J.H., 1968. Calculation and miscalculation of the allometric equation as a model in biological data. BioScience 18, 1118–1120.

Andrew J. Kerkhoff *
Department of Biology, Kenyon College, Gambier, OH 43022, USA
E-mail address: kerkhoffa@kenyon.edu

Andrew J. Kerkhoff *
Department of Mathematics, Kenyon College, Gambier, OH 43022, USA
E-mail address: kerkhoffa@kenyon.edu

Brian J. Enquist
Department of Ecology and Evolution, University of Arizona, Tucson, AZ 85701, USA

Available online 30 December 2008
31 October 2008

* Corresponding author. Tel.: +1 740 427 5734.