

Traditional allometric analysis fails to provide a valid predictive model for mammalian metabolic rates

Gary C. Packard^{1,*} and Geoffrey F. Birchard²

¹Department of Biology, Colorado State University, Fort Collins, CO 80523, USA and ²Department of Environmental Science and Policy, George Mason University, Fairfax, VA 22030, USA

*Author for correspondence (e-mail: packard@lamar.colostate.edu)

Accepted 23 September 2008

SUMMARY

The field of biological allometry was energized by the publication in 1997 of a theoretical model purporting to explain 3/4-power scaling of metabolic rate with body mass in mammals. This 3/4-power scaling exponent, which was first reported by Max Kleiber in 1932, has been derived repeatedly in empirical research by independent investigators and has come to be known as 'Kleiber's Law'. The exponent was estimated in virtually every instance, however, by fitting a straight line to logarithmic transformations of data and by then re-expressing the resulting equation in the arithmetic scale. Because this traditional method may yield inaccurate and misleading estimates for parameters in the allometric equation, we re-examined the comprehensive data set that led Savage and colleagues to reaffirm the view that the metabolic rate of mammals scales to the 3/4-power of body mass. We found that a straight line fitted to logged data for the basal metabolic rate (BMR) of mammals ranging in size from a 2.4 g shrew to a 3672 kg elephant does not satisfy assumptions underlying the analysis and that the allometric equation obtained by back-transformation underestimates BMR for the largest species in the sample. Thus, the concept of 3/4-power scaling of metabolic rate to body mass is not well supported because the underlying statistical model does not apply to mammalian species spanning the full range in body size. Our findings have important implications with respect to methods and results of other studies that used the traditional approach to allometric analysis.

Key words: allometry, scaling, logarithms, transformations, metabolic rates.

INTRODUCTION

The field of biological allometry was energized by the publication in 1997 of a theoretical model purporting to explain 3/4-power scaling of metabolic rate with body mass in mammals (West et al., 1997). The model generated a wave of new interest in the discipline (Agutter and Wheatley, 2004; Glazier, 2005; da Silva et al., 2006), and it also re-opened the long-simmering debate about the 'true' value for the scaling factor in the allometric equation (Hoppeler and Weibel, 2005). However, this debate about the value for the scaling exponent may be premature, because investigators on all sides of the issue have for years unknowingly used an unreliable procedure to estimate parameters in the equation (see below). This problem has gone largely unrecognized because allometric equations typically have not been validated in the scale of measurement.

The relationship between metabolic rate and body mass is usually assumed to follow a simple, two-parameter power function:

$$Y = aX^b, \quad (1)$$

where Y is metabolic rate, X is body mass, and the parameters a and b are the scaling (allometric) coefficient and the scaling exponent, respectively (Agutter and Wheatley, 2004; Glazier, 2005; da Silva et al., 2006). It is unclear whether this expression emerged from strictly theoretical considerations; or whether it was adopted in the era prior to digital computers because of the relative ease in manipulating logarithmic transformations of empirical data conforming with such an equation; or whether the answer lies in some combination of these and other factors (Kleiber, 1961; Gould, 1966; Heusner, 1987). Regardless of the origin of the function, however, values for the predictor and response variables are seldom

examined in the original arithmetic scale (Smith, 1984) but, instead, are immediately transformed to their logarithms, at which point the expression assumes the form:

$$\log Y = \log a + b \log X \quad (2)$$

(Smith, 1984; Agutter and Wheatley, 2004; Glazier, 2005; da Silva et al., 2006). A straight line commonly is fitted to the data by the method of ordinary least squares (Glazier, 2005), after which parameters in the allometric equation are estimated by back-transformation to the arithmetic scale (Smith, 1984).

The aforementioned ('traditional') approach to allometric analysis has not changed appreciably since the time of Kleiber, Benedict and Brody (Kleiber, 1932; Benedict, 1938; Brody, 1945). The approach nonetheless is beset by a variety of problems, most of which result from the use of logarithmic transformations. First, transformation profoundly alters the relationship between predictor and response variables (Emerson and Stoto, 1983; Jansson, 1985; Osborne, 2002), so influential outliers may go undetected, remain in the data set, and bias parameter estimates in the fitted statistical model (Packard and Boardman, 2008a). Second, the two-parameter power function (Eqn 1) underlying the traditional allometric analysis may not provide a good fit to the data (Zar, 1968; Albrecht and Gelvin, 1987; Albrecht, 1988; Packard and Boardman, 2008a; Packard and Boardman, 2008c), in which case parameter estimates again may be inaccurate and misleading. Third, the statistical model obtained by back-transformation from logarithms is one that predicts geometric means for Y instead of arithmetic means (Miller, 1984; Smith, 1993; Hayes and Shonkwiler, 2006). And fourth, a straight line fitted to logged values may undergo distortional rotation owing

to the fact that squared residuals are not equivalent for large and small values of the original response variable (Zar, 1968; Jansson, 1985; McCuen et al., 1990; Pandey and Nguyen, 1999; Packard and Boardman, 2008b).

We re-examined data for the basal metabolic rate (BMR) of 626 species of mammals (Savage et al., 2004) to illustrate how applying the traditional method for allometric analysis can result in biased and misleading estimates for parameters in a two-parameter allometric equation. Species represented in this comprehensive sample varied in size from a 2.4 g shrew to a 3672 kg elephant. We focused on a modified data set created by binning logged values for body mass [Appendix 2 in the study by Savage and colleagues (Savage et al., 2004)], because the resulting estimate of 0.737 for the allometric exponent is regarded by many workers as providing strong support for the concept of 3/4-power scaling (Brown et al., 2004; Farrell-Gray and Gotelli, 2005; West and Brown, 2005). The dimension for each of the bins was 0.1 in logged units for body mass, and each bin yielded one representative ('average') value irrespective of the number of species assigned to it. Binning was undertaken by Savage and colleagues (Savage et al., 2004) to prevent the preponderance of small species in the full sample from exerting undue influence on estimates for parameters in the allometric equation; but the use of binned values also facilitates graphical analysis by avoiding the visual clutter that would accompany the display of more than 600 values in a single plot. Preliminary examination of values for the full 626 species, coupled with the results of an independent study of the same data set (Hui and Jackson, 2007), indicates that none of our conclusions was affected by using binned values.

We do not address in our study subsidiary issues like phylogenetic independence of measurements for different species (Garland et al., 2005) or assumptions of least-squares regression (Warton et al., 2006). Instead, our treatment follows the same general approach that was used by Savage and colleagues (Savage et al., 2004), thereby enabling us to make detailed comparisons of their findings with our own. We carried all calculations to six decimal places before rounding to three.

METHODS AND RESULTS

Preliminary analyses

Data for 52 samples expressed in both logarithmic and arithmetic scales first were displayed in bivariate scatterplots and examined for patterns and trends (Anscombe, 1973). This step was followed by preliminary statistical analyses (Packard and Boardman, 2008a) wherein a straight line was fitted to logarithmic transformations by ordinary least squares and a two-parameter power function was fitted to values in the arithmetic scale by non-linear regression (Motulsky and Christopoulos, 2004). Computations were performed in SigmaPlot (version 10.0 from Systat Software, Inc., San Jose, CA, USA), which uses the Marquardt–Levenberg algorithm to fit non-linear functions by an iterative process that minimizes the sum of squares for residuals (Marquardt, 1963). Allometric exponents estimated by the two procedures were quite different: 0.737 from the slope of the straight line fitted to logarithms ($R^2=0.986$) and 0.909 from the two-parameter power function fitted to values in the arithmetic scale ($R^2=0.997$). However, logarithmic transformation failed to linearize the data (e.g. a plot of residuals for logBMR against predictions yielded a pattern in the form of an inverted parabola) (see also Kozłowski and Konarzewski, 2005), and the equation resulting from back-transformation seriously underestimated metabolic rate for the elephant, the largest species in the sample. In addition, neither of the analyses passed the test for constancy of

variances (Spearman rank correlation between absolute values for residuals and observed values for the predictor variable) (Kutner et al., 2004).

The unbalanced distribution of arithmetic values for body mass was also of concern. The elephant (which was the only species represented in that bin) was nearly an order of magnitude heavier than the next species in the sample. This observation, coupled with the decidedly different estimates for allometric exponents, raised the possibility that the elephant was an unduly influential outlier in the non-linear regression (Anscombe, 1973; Stevens, 1984; Osborne and Overbay, 2004). We subsequently discovered that Cook's Distance, which is a sensitive measure of the influence of a data point on parameters in the fitted model (Kutner et al., 2004), was an extraordinary 4600 for the elephant. Any data point for which Cook's Distance exceeds 4 is likely to exert undue influence, so we treated the elephant as a statistical outlier and removed it from the data set.

Fitting the allometric equation

A straight line fitted to the remaining 51 logged values yielded a statistically significant equation (Fig. 1A), even though a plot of residuals against predicted values indicated once again that a straight line was not an appropriate model (Fig. 1B). Nonetheless, R^2 was extraordinarily high (Fig. 1A), and the estimate of 0.728 for the allometric exponent is similar to the estimate from examination of the full data set (i.e. exclusion of the elephant had little effect on the outcome). The analysis passed the test for normality ($P=0.288$ by Kolmogorov–Smirnov test) (Kutner et al., 2004) but it failed the one for constancy of variances ($P=0.002$). Consequently, confidence limits for the slope and intercept are unlikely to be reliable (Myers, 1986; Finney, 1989).

A plot of data in the original scale shows an unbalanced distribution for body mass, but the imbalance is not as extreme as it was when the elephant was included in the sample (Fig. 1C). A two-parameter power function fitted to the data by non-linear regression yielded an estimate of 0.686 for the allometric exponent (Fig. 1C). Although the model fitted by non-linear regression was statistically significant and R^2 was high (Fig. 1C), the analysis failed tests for both normality of residuals and homogeneity of variances ($P<0.001$). A plot of residuals shows that variances were related to body mass in a manner that commonly is associated with multiplicative error (Fig. 1D). Thus, reliable confidence limits cannot be computed for the parameter estimates a and b (Myers, 1986; Finney, 1989). However, the parameter estimates themselves may be better than those obtained by back-transformation (Myers, 1986), and the predictive equation is efficient in terms of minimizing residual variance (Asselman, 2000).

Validating the allometric equation

Next, we back-transformed the equation for the line fitted to logarithms and displayed the resulting function on bivariate graphs together with the function obtained by non-linear regression (Fig. 2A,B). The line from back-transformation is a good descriptor for values in the logarithmic scale (Fig. 2A), but it fails to predict values for large animals in the arithmetic scale (Fig. 2B). In contrast, the non-linear regression predicts consistently higher values for the response variable in the logarithmic domain (Fig. 2A) while performing much better than the alternative model for large animals in the arithmetic scale (Fig. 2B).

Because of the unbalanced distribution for the predictor variable in the arithmetic scale (Fig. 2B), we displayed arithmetic values for BMR against body mass on a log scale in order to better visualize

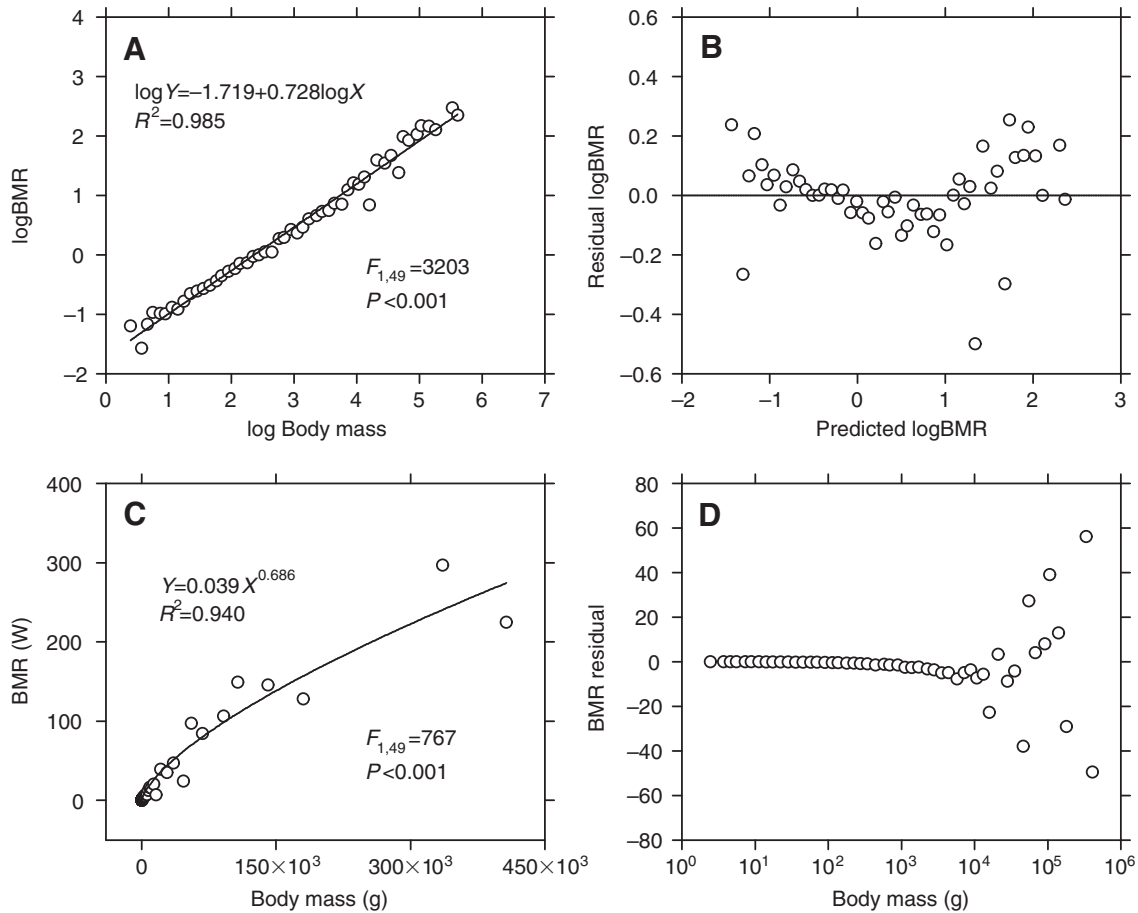


Fig. 1. Values for 51 samples exclusive of the elephant were taken from Appendix 2 in the study by Savage and colleagues (Savage et al., 2004). (A) A straight line was fitted to log-transformed data by ordinary least squares. (B) Residuals from the analysis in A are plotted against predicted values for logBMR. The parabolic pattern to the display indicates that a straight line is not an adequate function for describing values in the bivariate plot and calls into question the utility of the transformation. Departures from linearity were noted by Dodds and colleagues (Dodds et al., 2001) and Kozłowski and Konarzewski (Kozłowski and Konarzewski, 2005) in similar analyses, but both groups of investigators nonetheless continued to work with values in the logarithmic domain. (C) A two-parameter power function was fitted to arithmetic data by non-linear regression. (D) Residuals from the non-linear regression are plotted against body mass on a log scale to illustrate how variance increases with body size.

goodness of fits to data for smaller species (Fig. 2C). This display facilitated examination of BMR by expanding the apparent distribution for body mass at the low end of the scale without introducing transformation bias (Finney, 1989). The equation from back-transformation is the better predictor of BMR for animals between 10^2 and 10^4 g in body mass (Fig. 2C, inset), but clearly is the inferior function for predicting BMR of larger mammals (Fig. 2C). Thus, the equation from back-transformation is useful for predicting metabolic rates of animals weighing less than 10^4 g but the function is not valid over the full range in body size. The function fitted by non-linear regression is a better descriptor over the entire range in body size, even though it is more biased than the function from back-transformation in the mid-range for body size. (Part of the bias in the function fitted by non-linear regression is the result of forcing the line to pass through the origin and could be eliminated by fitting a three-parameter function instead. However, the scaling exponent in the three-parameter function differs only slightly from that for the two-parameter function, i.e. 0.667 vs 0.686, respectively.)

Curvilinearity in the allometric relationship

The scaling exponent for small species is predicted to be smaller than that for large species (Savage et al., 2004) – a prediction that

seems to be confirmed by the observed curvilinearity (concave upward) in the relationship between logBMR and log body mass (Fig. 1A,B). Such a curvilinear pattern of variation in log-transformed data should be cause for concern, because it calls into question the underlying allometric model (Eqn 1). Nevertheless, we examined data for large and small mammals separately to see whether the aforementioned prediction was realized. Binned data for mammals weighing less than 260 g were taken to represent small species whereas those for mammals weighing more than 260 g were taken to represent large ones (Savage et al., 2004). The elephant was omitted from the analyses because of the likelihood that it is an outlier.

Straight lines fitted to transformed values for both small and large species yielded statistically significant equations with high values for R^2 (Fig. 3A,B). The analysis of values for small species passed tests for normality ($P=0.065$) and homogeneity of variances ($P=0.059$). The analysis for large species, however, passed the test for normality ($P=0.298$) but not that for homoscedasticity ($P=0.016$). Scaling exponents for small and large species were 0.678 and 0.797, respectively, which seemingly confirmed expectation (Savage et al., 2004).

Non-linear regression on the two sets of values in the original scale also yielded statistically significant fits and high values for

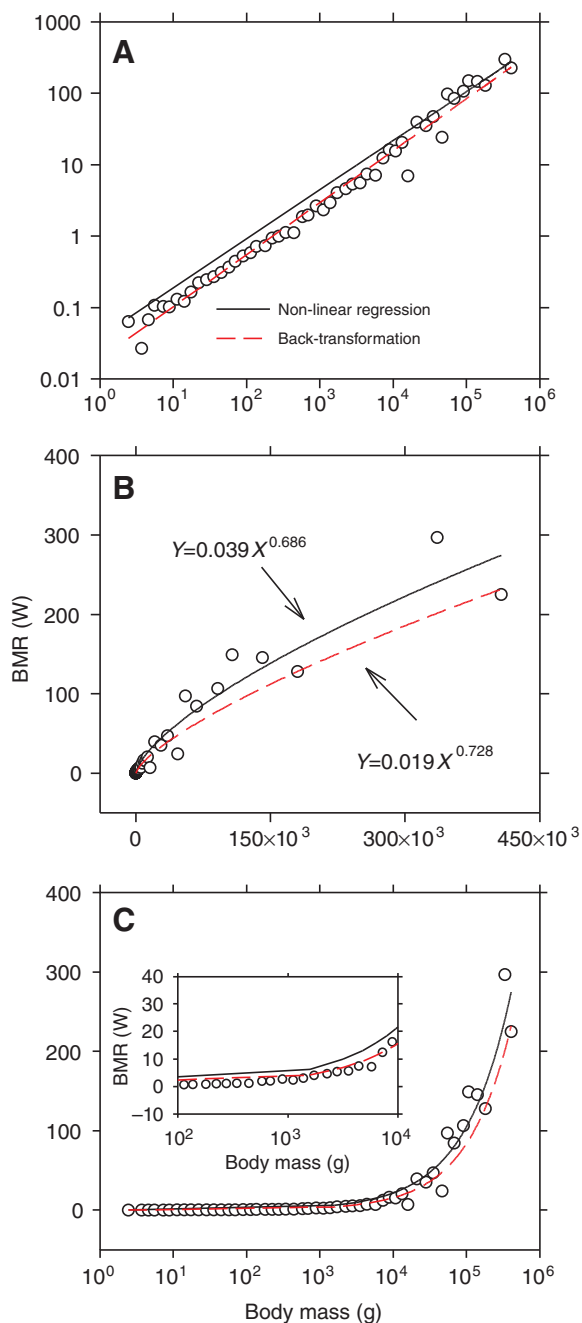


Fig. 2. Validation of alternative regression models derived from analyses summarized in Fig. 1. (A) Equations fitted to data by non-linear regression (solid black line) and by back-transformation from logarithms (dashed red line) are displayed together in a double logarithmic plot. (B) Equations fitted to data by non-linear regression (solid black line) and by back-transformation from logarithms (dashed red line) are displayed together in an arithmetic plot. (C) Arithmetic values for BMR are displayed against values for mass on a log scale. The upward trajectory for the lines is a consequence of the semi-log plot. The inset expands that part of the scale for animals weighing between 100 and 10,000 g.

R^2 (Fig. 3C,D). The analysis on the subset of values for small species met assumptions for both normality ($P=0.922$) and homogeneity of variances ($P=0.381$), whereas the treatment of large animals failed the test for normality ($P=0.008$) as well as that for homoscedasticity ($P<0.001$). Scaling exponents were estimated to be 0.656 and 0.686, respectively, for small and large species.

The alternative methods for fitting the allometric equation to data for small mammals yielded functions that are reasonably good visual fits to values in the original scale (Fig. 3E). The non-linear function fitted to values for large mammals is also a good fit graphically (Fig. 3F) but the equation obtained by back-transformation seriously underestimates BMR for species with masses between 75 and 150 kg (Fig. 3F).

Scaling exponents estimated for large and small species by the traditional method are quite different (Fig. 3A,B) whereas exponents estimated by non-linear regression are quite similar (Fig. 3C,D). Indeed, the 95% confidence interval for the exponent estimated by non-linear regression for small mammals (i.e. the group for which such limits can be reliably computed) is 0.615–0.697, which includes the exponent estimated for large mammals (Fig. 3D) as well as the one for all species exclusive of the elephant (Fig. 1C). Thus, analyses of different subsets of the data by non-linear regression lead to a common estimate for a scaling exponent (in the range 0.656–0.686) – not to the different exponents predicted by the aforementioned theoretical model (Savage et al., 2004).

DISCUSSION

Cause for the log bias

Why do the two-parameter functions estimated by back-transformation differ so much from those estimated by non-linear regression? We suggest that the disparity in this instance is largely the result of fitting straight lines to logarithmic transformations, which are inherently non-linear and which consequently change the relationship between predictor and response variables (Emerson and Stoto, 1983; Jansson, 1985; Osborne, 2002). Logarithmic transformation results in an overall compression of distributions for the variables, but the compression is greater at the high ends of the scales than at the low ends (Emerson and Stoto, 1983; Jansson, 1985; Osborne, 2002). This disproportionate compression causes small values for the variables to have a large effect on parameters in the fitted equation and large values to have a small effect (Glass, 1969; Jansson, 1985; Packard and Boardman, 2008b).

By way of example (Jansson, 1985), consider a straight line fitted to logarithms of 0.9 and 1.1 at one level for X and to 1.9 and 2.1 at a higher level for X . Predictions for logs of the response variable Y are 1.0 and 2.0, respectively, with all residuals having absolute values of 0.1. Such a balanced distribution of residuals indicates that logged values for Y were weighted equally in fitting the line by ordinary least squares.

Back-transformation of the predicted values yields a geometric (not arithmetic) mean of 10 at the first level for X and 100 at the second level for X . Observed values corresponding to the prediction of 10 are 7.9 and 12.6 whereas those corresponding to the prediction of 100 are 79.4 and 125.9. Thus, absolute values for residuals expressed on the scale of measurement are 2.1, 2.6, 20.6 and 25.9, despite the fact that all residuals were ± 0.1 in the log domain. This difference between the scales is important because the fitted line minimizes the sum of the squared residuals regardless of the scale in which the data are expressed (Zar, 1968). Whereas the squares for the residuals in the log domain are identical (i.e. 0.01), the square for the largest of the values in the arithmetic domain (i.e. 670.8) is more than two orders of magnitude larger than that for the smallest value (i.e. 4.4).

At each level for X , the smaller of the two measurements lies below the fitted line and the larger one lies above it. Consequently, the smaller values in the arithmetic scale have a disproportionate influence on the elevation of the line fitted to logarithms because residuals for large and small values are identical in the logarithmic

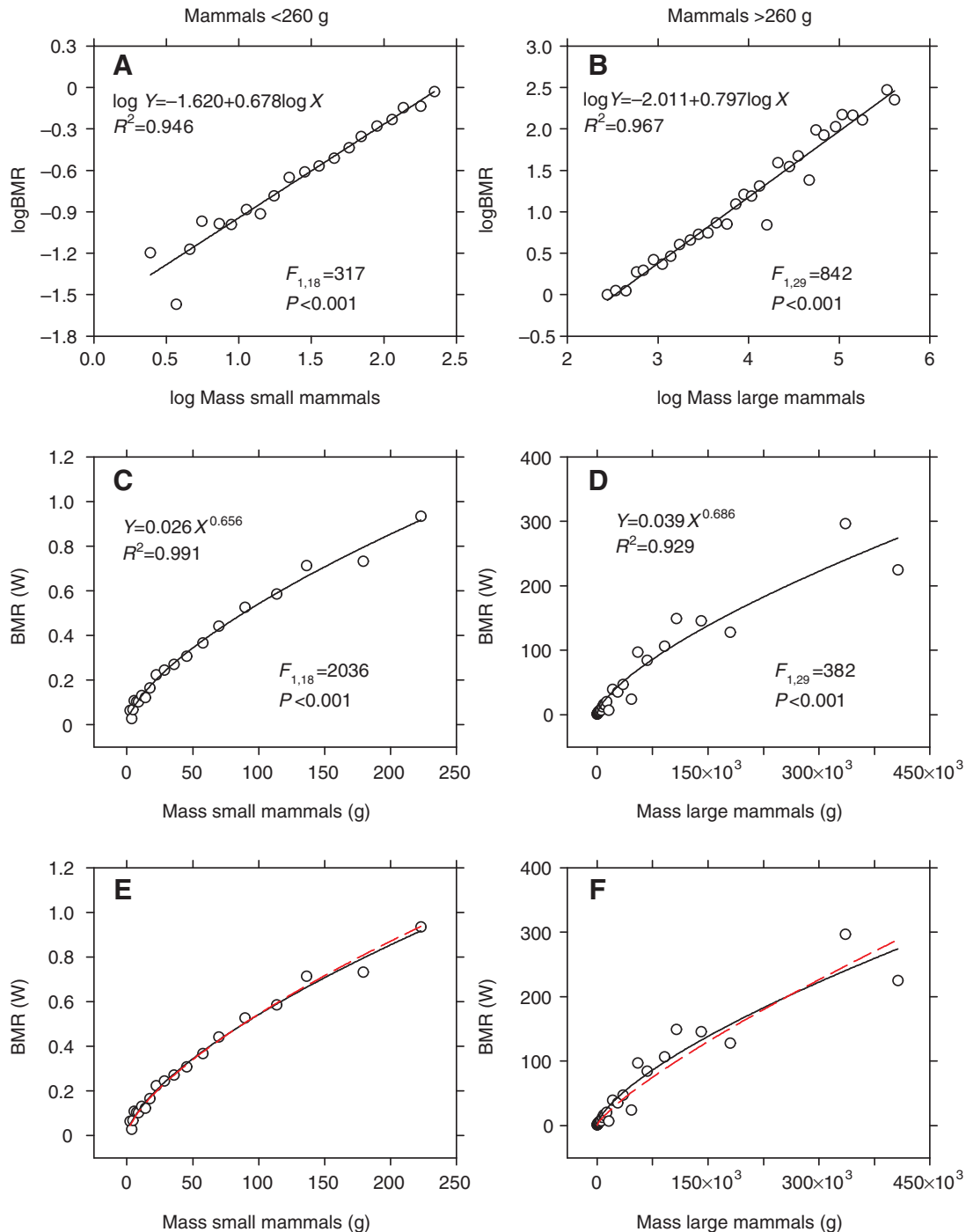


Fig. 3. (A) A straight line was fitted by ordinary least squares to log-transformed data for mammals weighing less than 260 g. (B) A straight line was fitted by ordinary least squares to log-transformed data for mammals weighing more than 260 g. (C) A two-parameter power function was fitted by non-linear regression to values for metabolic rate and body mass for mammals weighing less than 260 g. (D) A two-parameter power function was fitted by non-linear regression to values for metabolic rate and body mass for mammals weighing more than 260 g. (E) Equations fitted to data for small mammals by non-linear regression (solid black line) and back-transformation (dashed red line) are displayed against the backdrop of values in the original scale. (F) Equations fitted to data for large mammals by non-linear regression (solid black line) and back-transformation (dashed red line) are displayed against the backdrop of values in the original scale.

domain. Additionally, the smaller of the two measurements lying above (or below) the fitted line is associated with the lower level for X and the larger with the higher level. Thus, the smaller value in the arithmetic scale has a disproportionate influence on the slope of the line fitted to logarithms (again, because residuals are identical in the log domain).

Depending on the distributions of the variables, both the slope and intercept of the straight line may be affected in unexpected ways (Glass, 1969; Jansson, 1985; McCuen et al., 1990; Pandey and Nguyen, 1999; Packard and Boardman, 2008b), and these effects later are transmitted by back-transformation to the two-parameter allometric equation. This disparate influence of small and large

values is apparent in the current study in graphs of the alternative equations in both logarithmic and arithmetic scales (Fig. 2). The linear regression on transformed values was rotated in a counter-clockwise direction (Fig. 2A), and the result was a poor fit of the back-transformation to data for large animals (Fig. 2B). The general problem outlined here probably occurs commonly in data sets that include animals spanning large ranges in size (Glass, 1969; Packard and Boardman, 2008b): the traditional procedure provides good predictions for small animals but poor predictions for large ones.

Savage and colleagues used the binning procedure in an attempt to reduce the disproportionate influence of the many small species in the sample and thereby obtain a more reliable estimate for the scaling exponent (Savage et al., 2004). On the other hand, Glazier argued that binning actually caused values for large species to have too great an influence on parameters in the allometric equation owing to an increase in proportional representation for large species (Glazier, 2008). Both these suggestions, however, are based on a misunderstanding of the logarithmic transformation. First, binning by logs for body mass had the effect of expanding the scale at the lower end of the distribution and compressing it at the upper end, thereby maintaining a skew in the distribution of masses expressed in grams and causing small species to be over-represented in the 51 bins exclusive of the elephant. For example, the arithmetic mean for 51 back-transformed values for body mass is 31,364 g. A total of 42 bins (of which one was empty) were available to accommodate species with masses below the average, and 12 (of which two were empty) were available to accommodate species with masses above the average. Small species consequently continued to be 'over-represented' in the data set. Second, the line fitted to logarithms was 'transformation biased' by the undue influence of the small species (Jansson, 1985; Packard and Boardman, 2008b), leading to rotation of the line, to underestimation of the allometric coefficient from back-transformation, and to overestimation of the allometric exponent (Fig. 2B). The minor influence of large species is why deletion of the elephant from the data set had little effect on the allometric exponent estimated by the traditional method.

Logarithmic transformations

Logarithmic transformations have a long history of use in allometric analyses, so it is useful here to consider briefly the reasons for performing such transformations and to ask whether the transformations continue to have application.

(1) Logarithmic transformations were used for many years to linearize data and thereby promote graphical display and statistical analysis (Smith, 1984; Smith, 1993). This application of log transformations was based on the implicit assumption that the data conformed with a two-parameter power function (Smith, 1984). The assumption was seldom addressed, and inappropriate equations were sometimes fitted (Albrecht and Gelvin, 1987; Albrecht, 1988; Packard and Boardman, 2008a; Packard and Boardman, 2008c). However, the form of the underlying equation now is moot because the advent of sophisticated PCs and software for both graphical and statistical analysis has rendered linearization unnecessary (Motulsky and Christopoulos, 2004).

(2) Logarithmic transformations commonly produce symmetrical distributions for data that have unbalanced distributions in the original scale (Emerson and Stoto, 1983). The main problem with unbalanced distributions is that the largest animals in the sample may exert undue influence on parameters that are estimated by methods based on least squares. Indeed, the problem of imbalance was apparent in data examined here (Fig. 1C). Whereas transformation may produce a better distribution of values (Fig. 1A),

it also introduces a bias that may be more serious than the problem that the transformation was intended to correct (Jansson, 1985). The bias results from the change in error structure of the statistical model on back-transformation to the arithmetic scale (Glass, 1969; Manaster and Manaster, 1975; Miller, 1984; Jansson, 1985; Smith, 1993; Hayes and Shonkwiler, 2006) coupled with the distortion that is introduced by the non-linear relationship between logs and values in the original scale (Jansson, 1985; McCuen et al., 1990; Pandey and Nguyen, 1999; Hui and Jackson, 2007; Packard and Boardman, 2008b). A better way than transformation for addressing the problem of unbalanced distributions might be to make greater use of regression diagnostics and model validation (Anscombe, 1973; Snee, 1977; Kutner et al., 2004).

(3) Allometric data often are heteroscedastic, and transformation may create a new distribution that meets the assumption of homogeneity of variances (Emerson and Stoto, 1983; Smith, 1984). Unfortunately, few authors examine the transformed variables for heteroscedasticity, so the resulting analysis may fail to correct the problem that led to transformation in the first place; the logged data examined here provide but one example of the failure of transformation to achieve its intended goal (Fig. 1B). Even when the original data are heteroscedastic and transformation results in a homoscedastic distribution, the potential exists for introducing a transformation bias that is evident only on back-transformation to the arithmetic scale. However, the problem of heteroscedasticity probably will become moot as more investigators discover statistical procedures that accommodate different assumptions concerning distributions of data in the original scale (Lane, 2002; Cox et al., 2008).

Validating the model

Regardless of the means by which an allometric equation is fitted to data, it is essential that the model be validated (Snee, 1977; Emerson and Stoto, 1983; Myers, 1986; Finney, 1989; Kutner et al., 2004). A graphical display is the most effective way to verify that the fitted model actually describes the data on which it is based (Anscombe, 1973). Unfortunately, validation in the traditional allometric analysis typically is limited to a display of values in the logarithmic scale, which often has little bearing on the relationship between metabolic rate and body mass (Fig. 2A,B). A good fit of a linear function to logarithms does not imply a good fit of the re-expressed equation to data in the arithmetic scale (McCuen et al., 1990; Pattyn and Van Huele, 1998). Thus, proper validation requires that the allometric equation be shown against the backdrop of data in the scale of measurement (Emerson and Stoto, 1983; Myers, 1986; Finney, 1989).

Implications for theoretical models

The data set compiled by Savage and colleagues is widely regarded to be one of the very best (Savage et al., 2004). It comes as no surprise, therefore, that the statistical analysis performed on those data by Savage and colleagues (Savage et al., 2004) is viewed by many in the scientific community as offering strong evidence in support of the concept of 3/4-power scaling for metabolic rate on body mass in mammals (Brown et al., 2004; Farrell-Gray and Gotelli, 2005; West and Brown, 2005). However, Savage and colleagues (Savage et al., 2004) seem to have omitted three critical steps from their investigation: they apparently did not (1) examine their data for potential outliers, (2) test assumptions underlying their statistical analysis, or (3) validate the allometric model in the original scale. Our re-analysis of data exclusive of the elephant, which was an apparent outlier, revealed that a linear equation fitted to log-log

transformations failed tests for both linearity and constancy of variances, and that the two-parameter power function estimated by back-transformation did not predict metabolic rates of large animals in the sample. Consequently, the earlier estimate of 3/4-power scaling is not well supported, thereby calling into question the validity of theoretical models that purport to explain such a scaling factor in mammals (e.g. West et al., 1997; Banavar et al., 1999; Darveau et al., 2002).

The traditional approach to allometric research is to fit a straight line to logarithmic transformations and then back-transform the resulting equation to the arithmetic scale (Smith, 1984; Agutter and Wheatley, 2004; Glazier, 2005; da Silva et al., 2006). Consequently, biases of the kinds shown in the current study and elsewhere (Glass, 1969; Hui and Jackson, 2007; Packard and Boardman, 2008a; Packard and Boardman, 2008b; Packard and Boardman, 2008c) are likely to occur commonly in published research on allometry. For this reason, scaling coefficients and exponents reported in the literature should be interpreted with a healthy dose of skepticism.

We thank four anonymous referees for constructive criticism that helped us to improve this report.

REFERENCES

- Agutter, P. S. and Wheatley, D. N. (2004). Metabolic scaling: consensus or controversy? *Theor. Biol. Med. Model.* **1**, 13.
- Albrecht, G. H. (1988). Simple allometry coefficients: fact or fiction. *Am. Zool.* **28**, 172A.
- Albrecht, G. H. and Gelvin, B. R. (1987). The simple allometry equation reconsidered: assumptions, problems, and alternative solutions. *Am. J. Phys. Anthropol.* **72**, 174.
- Anscombe, F. J. (1973). Graphs in statistical analysis. *Am. Stat.* **27**, 17-21.
- Asselman, N. E. M. (2000). Fitting and interpretation of sediment rating curves. *J. Hydrol.* **234**, 228-248.
- Banavar, J. R., Maritan, A. and Rinaldo, A. (1999). Size and form in efficient transportation networks. *Nature* **399**, 130-132.
- Benedict, F. G. (1938). *Vital Energetics*. Washington, DC: Carnegie Institution of Washington.
- Brody, S. (1945). *Bioenergetics and Growth*. New York: Reinhold Publishing.
- Brown, J. H., Gillooly, J. F., Allen, A. P., Savage, V. M. and West, G. B. (2004). Response to forum commentary on "Toward a metabolic theory of ecology." *Ecology* **85**, 1818-1821.
- Cox, N. J., Warburton, J., Armstrong, A. and Holliday, V. J. (2008). Fitting concentration and load rating curves with generalized linear models. *Earth Surf. Process. Landforms* **33**, 25-39.
- Darveau, C. A., Suarez, R. K., Andrews, R. D. and Hochachka, P. W. (2002). Allometric cascade as a unifying principle of body mass effects on metabolism. *Nature* **417**, 166-170.
- da Silva, J. K. L., Garcia, G. J. M. and Barbosa, L. A. (2006). Allometric scaling laws of metabolism. *Phys. Life Rev.* **3**, 229-261.
- Dodds, P. S., Rothman, D. H. and Weitz, J. S. (2001). Re-examination of the "3/4-law" of metabolism. *J. Theor. Biol.* **209**, 9-27.
- Emerson, J. D. and Stoto, M. A. (1983). Transforming data. In *Understanding Robust and Exploratory Data Analysis* (ed. D. C. Hoaglin, F. Mosteller and J. W. Tukey), pp. 97-128. New York: Wiley.
- Farrell-Gray, C. C. and Gotelli, N. J. (2005). Allometric exponents support a 3/4-power scaling law. *Ecology* **86**, 2083-2087.
- Finney, D. J. (1989). Was this in your statistics textbook? V. Transformation of data. *Exp. Agric.* **25**, 165-175.
- Garland, T., Jr, Bennett, A. F. and Rezende, E. L. (2005). Phylogenetic approaches in comparative physiology. *J. Exp. Biol.* **208**, 3015-3035.
- Glass, N. R. (1969). Discussion of calculation of power function with special reference to respiratory metabolism in fish. *J. Fish. Res. Bd. Canada* **26**, 2643-2650.
- Glazier, D. S. (2005). Beyond the '3/4-power law': variation in the intra- and interspecific scaling of metabolic rate in animals. *Biol. Rev. Camb. Philos. Soc.* **80**, 611-662.
- Glazier, D. S. (2008). Effects of metabolic level on the body size scaling of metabolic rate in birds and mammals. *Proc. R. Soc. Lond., B, Biol. Sci.* **275**, 1405-1410.
- Gould, S. J. (1966). Allometry and size in ontogeny and phylogeny. *Biol. Rev. Camb. Philos. Soc.* **41**, 587-640.
- Hayes, J. P. and Shonkwiler, J. S. (2006). Allometry, antilog transformations, and the perils of prediction on the original scale. *Physiol. Biochem. Zool.* **79**, 665-674.
- Heusner, A. A. (1987). What does the power function reveal about structure and function in animals of different size? *Annu. Rev. Physiol.* **49**, 121-133.
- Hoppeler, H. and Weibel, E. R. (2005). Scaling functions to body size: theories and facts. *J. Exp. Biol.* **208**, 1573-1574.
- Hui, D. and Jackson, R. B. (2007). Uncertainty in allometric exponent estimation: a case study in scaling metabolic rate with body mass. *J. Theor. Biol.* **249**, 168-177.
- Jansson, M. (1985). A comparison of detransformed logarithmic regressions and power function regressions. *Geograf. Ann.* **67A**, 61-70.
- Kleiber, M. (1932). Body size and metabolism. *Hilgardia* **6**, 315-353.
- Kleiber, M. (1961). *The Fire of Life*. New York: Wiley.
- Kozlowski, J. and Konarzewski, M. (2005). West, Brown and Enquist's model of allometric scaling again: the same questions remain. *Funct. Ecol.* **19**, 739-743.
- Kutner, M. H., Nachtsheim, C. J. and Neter, J. (2004). *Applied Linear Regression Models*, 4th edn. Boston: McGraw-Hill Irwin.
- Lane, P. W. (2002). Generalized linear models in soil science. *European J. Soil Sci.* **53**, 241-251.
- Manaster, B. J. and Manaster, S. (1975). Techniques for estimating allometric equations. *J. Morphol.* **147**, 299-308.
- Marquardt, D. W. (1963). An algorithm for least-squares estimation of nonlinear parameters. *J. Soc. Indust. Appl. Math.* **11**, 431-441.
- McCuen, R. H., Leahy, R. B. and Johnson, P. A. (1990). Problems with logarithmic transformations in regression. *J. Hydraul. Eng.* **116**, 414-428.
- Miller, D. R. (1984). Reducing transformation bias in curve fitting. *Am. Stat.* **38**, 124-126.
- Motulsky, H. and Christopoulos, A. (2004). *Fitting Models to Biological Data Using Linear and Nonlinear Regression*. Oxford: Oxford University Press.
- Myers, R. H. (1986). *Classical and Modern Regression with Applications*. Boston: Duxbury Press.
- Osborne, J. W. (2002). Notes on the use of data transformations. *Pract. Assess. Res. Eval.* **8**, 6.
- Osborne, J. W. and Overbay, A. (2004). The power of outliers (and why researchers should ALWAYS check for them). *Pract. Assess. Res. Eval.* **9**, 6.
- Packard, G. C. and Boardman, T. J. (2008a). Model selection and logarithmic transformation in allometric analysis. *Physiol. Biochem. Zool.* **81**, 496-507.
- Packard, G. C. and Boardman, T. J. (2008b). A comparison of methods for fitting allometric equations to field metabolic rates of animals. *J. Comp. Physiol. B* (in press).
- Packard, G. C. and Boardman, T. J. (2008c). Bias in interspecific allometry: examples from morphological scaling in varanid lizards. *Biol. J. Linn. Soc.* (in press).
- Pandey, G. R. and Nguyen, V.-T.-V. (1999). A comparative study of regression based methods in regional flood frequency analysis. *J. Hydrol.* **225**, 92-101.
- Pattyn, F. and Van Huelbe, W. (1998). Power law or power flaw? *Earth Surf. Process. Landforms* **23**, 761-767.
- Savage, V. M., Gillooly, J. F., Woodruff, W. H., West, G. B., Allen, A. P., Enquist, B. J. and Brown, J. H. (2004). The predominance of quarter-power scaling in biology. *Funct. Ecol.* **18**, 257-282.
- Smith, R. J. (1984). Allometric scaling in comparative biology: problems of concept and method. *Am. J. Physiol.* **246**, R152-R160.
- Smith, R. J. (1993). Logarithmic transformation bias in allometry. *Am. J. Phys. Anthropol.* **90**, 215-228.
- Snee, R. D. (1977). Validation of regression models: methods and examples. *Technometrics* **19**, 415-428.
- Stevens, J. P. (1984). Outliers and influential data points in regression analysis. *Psychol. Bull.* **95**, 334-344.
- Warton, D. I., Wright, I. J., Falster, D. S. and Westoby, M. (2006). Bivariate line-fitting methods for allometry. *Biol. Rev. Camb. Philos. Soc.* **81**, 259-291.
- West, G. B. and Brown, J. H. (2005). The origin of allometric scaling laws in biology from genomes to ecosystems: towards a quantitative unifying theory of biological structure and organization. *J. Exp. Biol.* **208**, 1575-1592.
- West, G. B., Brown, J. H. and Enquist, B. J. (1997). A general model for the origin of allometric scaling laws in biology. *Science* **276**, 122-126.
- Zar, J. H. (1968). Calculation and miscalculation of the allometric equation as a model in biological data. *BioScience* **18**, 1118-1120.