

# Allometric equations for predicting body mass of dinosaurs

G. C. Packard<sup>1</sup>, T. J. Boardman<sup>2</sup> & G. F. Birchard<sup>3</sup>

<sup>1</sup> Department of Biology, Colorado State University, Fort Collins, CO, USA

<sup>2</sup> Department of Statistics, Colorado State University, Fort Collins, CO, USA

<sup>3</sup> Department of Environmental Science and Policy, George Mason University, Fairfax, VA, USA

## Keywords

allometry; scaling; logarithms; dinosaurs.

## Correspondence

G. C. Packard, 865 Three Corner Gate Road,  
Livermore, CO 80536, USA.

Email: packard@amar.colostate.edu

Editor: Andrew Kitchener

Received 27 February 2009; revised 14 April  
2009; accepted 5 May 2009

doi:10.1111/j.1469-7998.2009.00594.x

## Abstract

We use data from the literature to compare two statistical procedures for estimating mass (or size) of quadrupedal dinosaurs and other extraordinarily large animals in extinct lineages. Both methods entail extrapolation from allometric equations fitted to data for a reference group of contemporary animals having a body form similar to that of the dinosaurs. The first method is the familiar one of fitting a straight line to logarithmic transformations, followed by back-transformation of the resulting equation to a two-parameter power function in the arithmetic scale. The second procedure entails fitting a two-parameter power function directly to arithmetic data for the extant forms by nonlinear regression. In the example presented here, the summed circumferences for humerus plus femur for 33 species of quadrupedal mammals was the predictor variable in the reference sample and body mass was the response variable. The allometric equation obtained by back-transformation from logarithms was not a good fit to the largest species in the reference sample and presumably led to grossly inaccurate estimates for body mass of several large dinosaurs. In contrast, the allometric equation obtained by nonlinear regression described data in the reference sample quite well, and it presumably resulted in better estimates for body mass of the dinosaurs. The problem with the traditional analysis can be traced to change in the relationship between predictor and response variables attending transformation, thereby causing measurements for large animals not to be weighted appropriately in fitting models by least squares regression. Extrapolations from statistical models obtained by back-transformation from lines fitted to logarithms are unlikely to yield reliable predictions for body size in extinct animals. Numerous reports on the biology of dinosaurs, including recent studies of growth, may need to be reconsidered in light of our findings.

## Introduction

Biologists frequently estimate the mass (or size) of dinosaurs and other large animals in extinct lineages by extrapolating from allometric equations fitted to data for contemporary species having a similar body form (Farlow, Dodson & Chinsamy, 1995). Common practice is to take some dimension from a skeletal element in extant species spanning a wide range in body size and then to use this dimension as the independent variable in a regression model for predicting body mass (Smith, 2002). Values for predictor and response variables for species in the reference sample first are converted to logarithms, after which a straight line is fitted by ordinary least squares (Smith, 1984, 2002). Back-transformation of the resulting equation yields a two-parameter power function (i.e. an equation having the form  $Y = aX^b$ ) that subsequently is used to predict mass of extinct animals from measurements taken from their skeletal remains (Anderson, Hall-Martin & Russell, 1985). Such predictions for

body mass may provide valuable insights concerning physiology, natural history and population dynamics of the extinct animals (Farlow *et al.*, 1995), so the statistical model needs to be as accurate as possible.

Unfortunately, the aforementioned (traditional) approach to allometric analysis is subject to several sources of bias. Some of these sources have come to light only recently and, thus, are largely unknown to workers in the field. First, the two-parameter power function, that is the basis for the traditional allometric analysis may not provide a good fit to the data (Zar, 1968), in which case parameter estimates may be inaccurate and misleading (Packard & Boardman, 2008, 2009b). Second, logarithmic transformation profoundly alters the relationship between the predictor and response variables (Emerson & Stoto, 1983; Jansson, 1985; Osborne, 2002), so that influential outliers in the arithmetic domain (Anscombe, 1973; Stevens, 1984; Osborne & Overbay, 2004) may go undetected, remain in the dataset and cause parameter estimates in the underlying statistical model to be

biased and misleading (Packard & Boardman, 2008; Packard, 2009). Third, the statistical model obtained by back-transformation from logarithms is one that predicts geometric means for the response variable instead of arithmetic means (Smith, 1993; Hayes & Shonkwiler, 2006). And fourth, a straight line fitted to logs may undergo distortional rotation owing to the fact that squared residuals in the logarithmic domain are not equivalent for large and small values of the untransformed response variable (Zar, 1968; Jansson, 1985; McCuen, Leahy & Johnson, 1990; Pandey & Nguyen, 1999; Packard & Birchard, 2008), thereby introducing bias when parameters of a power function are estimated by back-transformation (Packard & Boardman, 2009a). This last source of bias, which is especially insidious, has received scant attention from biologists.

Our goal in the present report is to illustrate the nature and severity of problems with the traditional method for fitting predictive equations to allometric data and to offer an alternative approach that conforms better than the traditional method with current statistical practice. We make no judgment concerning the relative merits of statistical versus nonstatistical procedures for predicting body size in extinct lineages (Henderson, 1999; Seebacher, 2001; Gunga *et al.*, 2007). Our study is based on a reanalysis of the important and influential dataset compiled by Anderson *et al.* (1985). Their reference sample was comprised of 33 species of quadrupedal mammals ranging in size from a microtine rodent weighing 47 g to an elephant weighing 5897 kg. A two-parameter equation estimated by back-transformation from a line fitted to logarithms was used in the original analysis to predict body mass of several quadrupedal dinosaurs (Anderson *et al.*, 1985). Although Anderson *et al.* (1985) followed what was at the time accepted practice, their estimates for body mass of these dinosaurs are much higher than warranted owing largely to the distortional effects of fitting lines to logarithms.

## Methods

Data for the current study were taken from table 3 in Anderson *et al.* (1985), where values were presented for the heaviest specimen for each of the 33 species represented in the reference sample. Values for circumference of a humerus were summed with corresponding values for circumference of a femur to yield the predictor variable for ensuing regression analyses. Body mass was the response variable.

We plotted arithmetic values and  $\log_{10}$  transformations in bivariate scatterplots and then inspected the graphs for patterns and trends (Anscombe, 1973). A straight line was fitted to the log transformations by ordinary least squares, whereas a two-parameter power function was fitted to arithmetic values by nonlinear regression (Motulsky & Christopoulos, 2004). A three-parameter power function ( $Y = Y_0 + aX^b$ ) also was fitted to arithmetic data, but it did not yield a demonstrably better model (based on the PRESS statistic; Kutner, Nachtsheim & Neter, 2004). Calculations were performed in SigmaPlot 10.0, which uses the Marquardt–Levenberg algorithm to fit nonlinear functions by

an iterative process that minimizes the sum of squares for residuals (Marquardt, 1963). We examined residuals for normality and homoscedasticity (Kutner *et al.*, 2004), and used plots of standardized residuals against predicted values to assess the goodness of fit of the models to the data. Studentized deleted residuals were used to identify potential outliers (Kutner *et al.*, 2004), and Cook's distance was used to assess the influence of individual data points on parameters in the fitted equations (Kutner *et al.*, 2004).

Finally, the equation obtained by back-transformation of the model fitted to logs was displayed together with the equation estimated by nonlinear regression on a graph with data in the arithmetic scale. Such a display provides the critical validation that is often omitted from allometric analyses, and is particularly important when predictions from the fitted model lie outside the range of values for the reference sample (Snee, 1977).

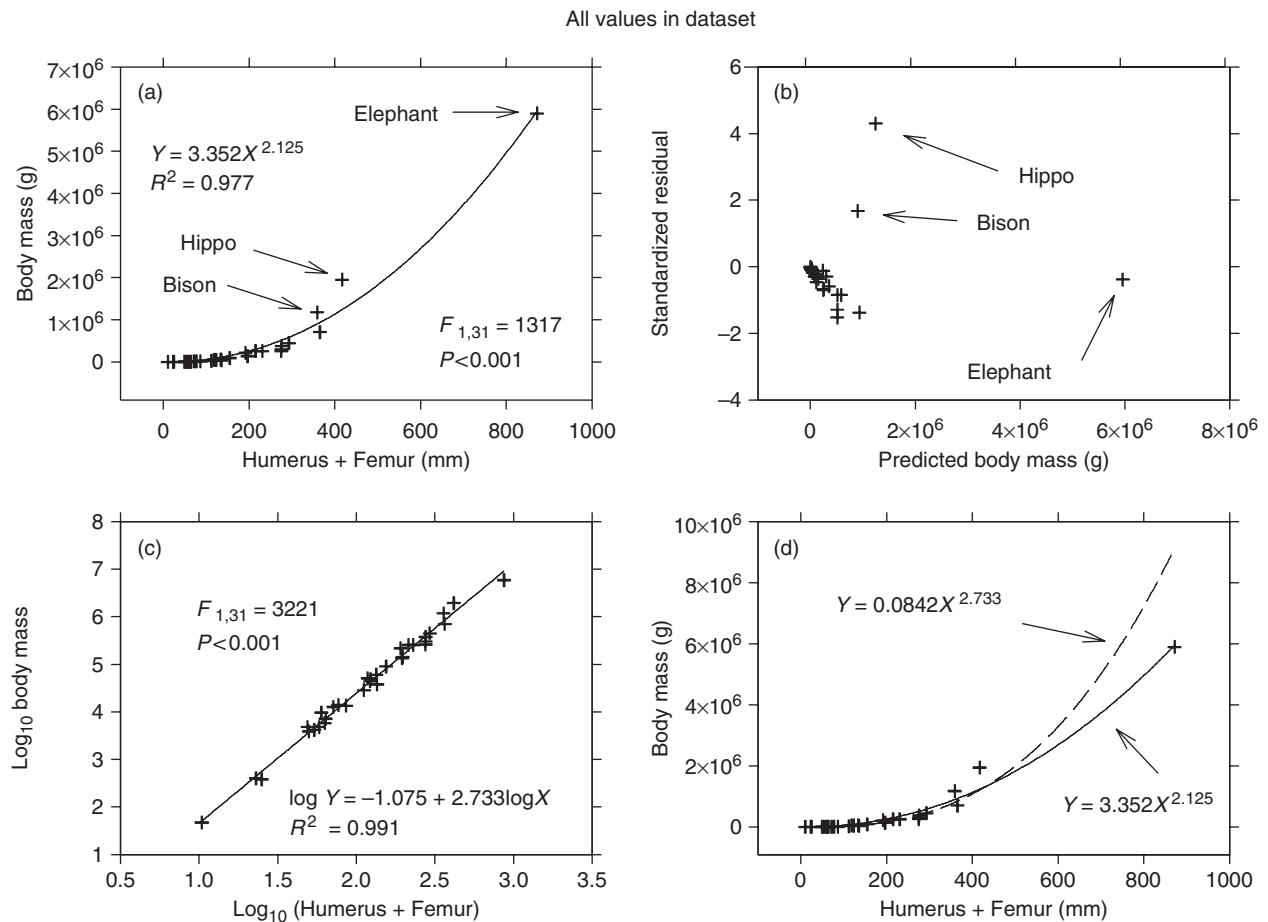
## Results

### All data

The nonlinear model fitted to arithmetic values for 33 species was statistically significant and yielded a high  $R^2$  (Fig. 1a). However, the analysis failed tests for normality of residuals and homoscedasticity ( $P < 0.001$ ), owing largely to the influence of the hippo (Fig. 1a and b). The Studentized deleted residual for the hippo (8.99) was the most extreme value for any species in the dataset, and Cook's distance (4.30) pointed to undue influence of the hippo on parameters in the fitted equation. These diagnostics, together with the plot of standardized residuals against predicted values (Fig. 1b), indicate that the hippo probably was an influential outlier (Anscombe, 1973; Stevens, 1984; Osborne & Overbay, 2004). Whereas the display of residuals clearly points to a problem with the distribution of the data, it does not reveal the funnel-shaped pattern that would be expected of data exhibiting multiplicative error (Kutner *et al.*, 2004), which is the kind of error that is implicit in the use of logarithmic transformations (Smith, 1993; Hayes & Shonkwiler, 2006).

The straight line fitted to logarithmic transformations was also statistically significant, and it, too, was characterized by a high  $R^2$  (Fig. 1c). In this instance, however, tests for normality and homoscedasticity were satisfied ( $P \geq 0.268$ ), and nothing untoward appeared in the plot of residuals (not shown). The hippo does not appear at all unusual in this analysis, with a Studentized deleted residual of only 2 and a Cook's distance of 0.17. In other words, the log transformation caused the value for the hippo to be merged with other values in the dataset.

The two-parameter power function resulting from back-transformation is a good fit to arithmetic values for animals of small and intermediate size but is not a suitable fit for the largest species in the sample (Fig. 1d). In contrast, the equation fitted by nonlinear regression is a generally good fit to data spanning the entire range in body size (Fig. 1d), despite the fact that the underlying analysis failed tests for



**Figure 1** (a) Body masses for 33 species of quadrupedal mammals are plotted against values for circumference of humerus plus femur (Anderson *et al.*, 1985). The summary statistics are from a nonlinear regression. (b) Standardized residuals from the analysis summarized in (a) are plotted against predicted values for body mass. The hippo is a possible outlier. (c) Log<sub>10</sub> transformations for body mass are plotted against log<sub>10</sub> transformations for humerus plus femur. Note that the largest animal in the sample appears to have been drawn by transformation toward the center of the distribution, whereas the smallest animal seems to have been moved farther from the center. The hippo does not appear at all unusual in this graphic. The summary statistics are from ordinary least squares regression. (d) Equations obtained by nonlinear regression (solid line) and back-transformation (dashed line) are shown against the backdrop of values expressed on the scale of measurement.

two of the assumptions of parametric statistics (i.e. normality of residuals and constancy of variances). These assumptions certainly are important when an objective of the study is to compute confidence limits for parameters in the allometric equation, but failure of the assumptions does not invalidate parameter estimates themselves (Myers, 1986; Finney, 1989; Asselman, 2000; Cox *et al.*, 2008). The parameter estimates are of primary concern here, and better estimates seem to be forthcoming from nonlinear regression than from the traditional method for fitting an allometric equation (Fig. 1d).

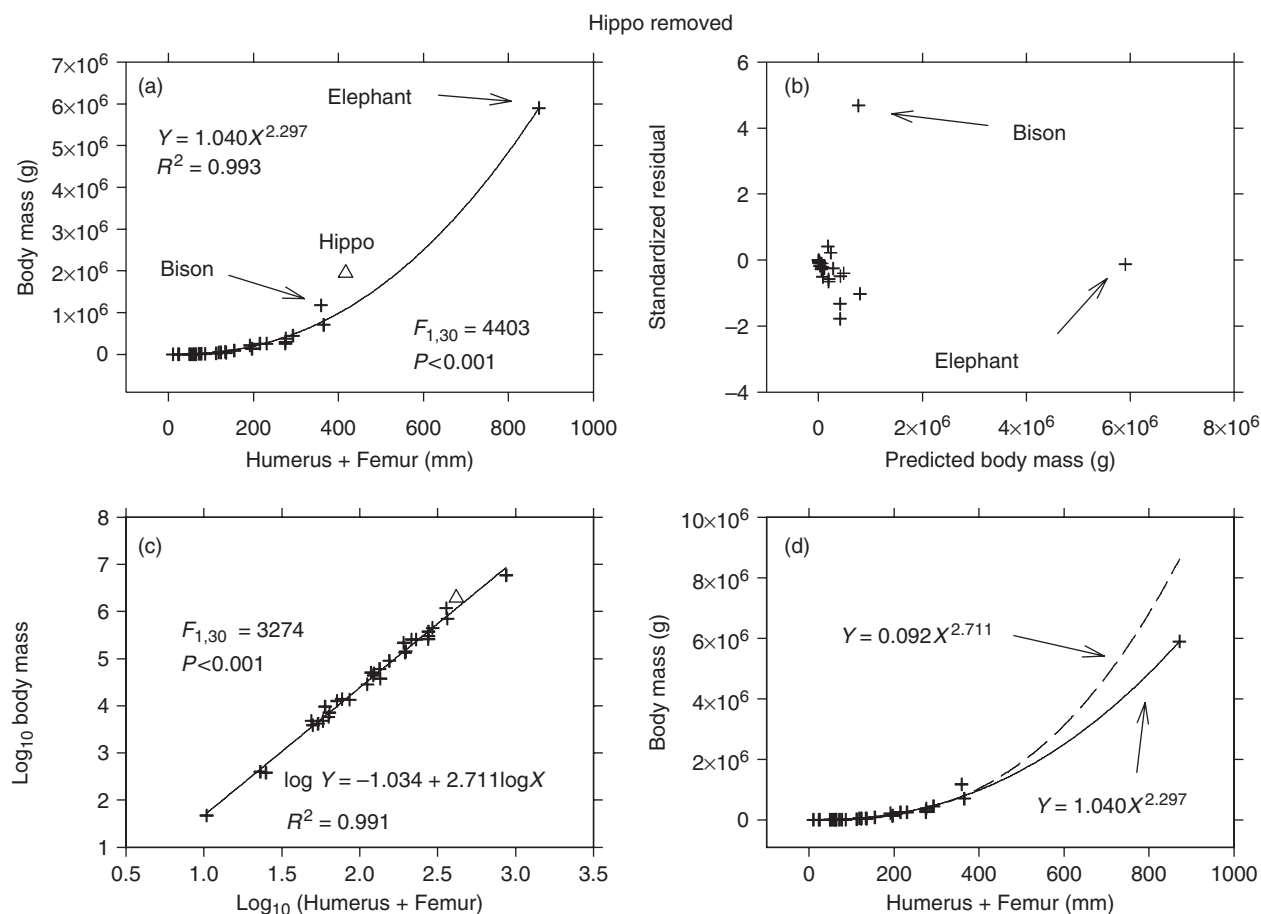
### Data exclusive of hippo

We could speculate on why the hippo, a largely aquatic animal, is heavier than expected from dimensions of the humerus and femur, but the point here is that it seems not to follow the pattern of variation exhibited by other species in

the dataset. Accordingly, we treated the hippo as a statistical outlier, deleted it from the dataset and repeated the analyses.

Nonlinear regression on arithmetic values resulted in a statistically significant model with high  $R^2$  (Fig. 2a). The analysis nonetheless failed tests for normality and homoscedasticity ( $P < 0.001$ ). The value for the bison seemed to deviate from the fitted curve (Fig. 2a), and the residual plot marked this species as another potential outlier (Fig. 2b). This perception was supported by a Studentized deleted residual of 4.69. Nevertheless, the bison was not overly influential because Cook's distance was marginally acceptable at 2.98. Removing the bison from the dataset resulted in an improved fit with a balanced distribution of residuals (not shown), but it did not result in any substantive change in the outcome of the study.

The straight line fitted to logs for data exclusive of the hippo was statistically significant, and the  $R^2$  was quite high (Fig. 2c). The analysis passed tests for normality and



**Figure 2** (a) Body masses for quadrupedal mammals exclusive of the hippo are plotted against values for circumference of humerus plus femur (Anderson *et al.*, 1985). The summary statistics are from a nonlinear regression. The open triangle here and elsewhere shows the location of the hippo. (b) Standardized residuals from the analysis summarized in (a) are plotted against predicted values for body mass. The bison is a possible outlier. (c)  $\log_{10}$  transformations for body mass are plotted against  $\log_{10}$  transformations for humerus plus femur. The bison does not appear at all unusual in this graphic. The summary statistics are from ordinary least squares regression. (d) Equations obtained by nonlinear regression (solid line) and back-transformation (dashed line) are shown against the backdrop of values expressed on the scale of measurement.

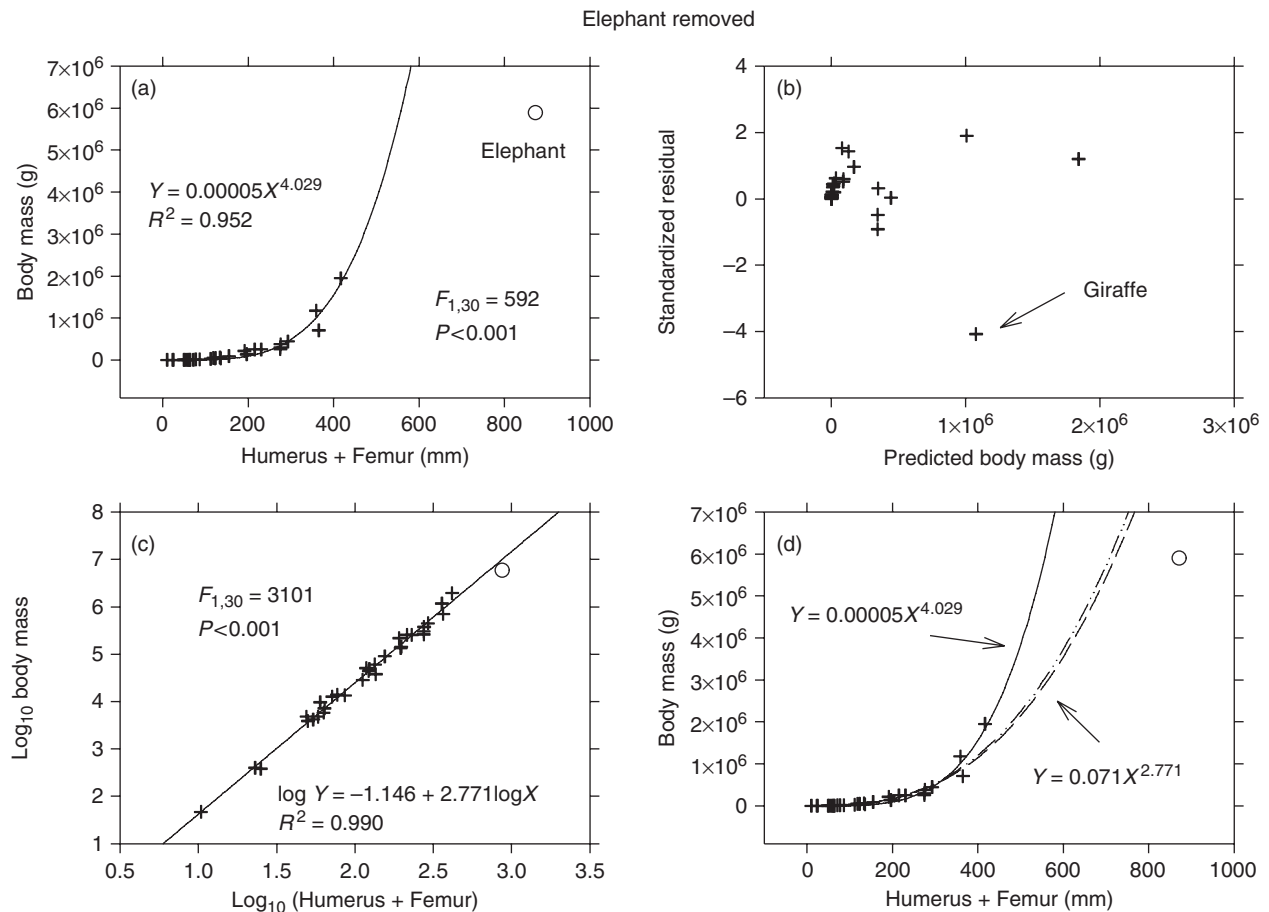
constancy of variances ( $P \geq 0.933$ ), and the plot of residuals (not shown) gave no hint of an outlier. Again, log transformation caused a potential outlier (the bison) to be merged with values for the rest of the sample (Fig. 2c).

The two-parameter power function resulting from back-transformation is a good fit to arithmetic values for animals of small and intermediate size but is not a suitable fit for the largest species in the sample (Fig. 2d). In contrast, the equation fitted by nonlinear regression is a generally good fit to data spanning the entire range in body size (Fig. 2d), despite the fact that the underlying analysis failed tests for two of the assumptions of parametric statistics. Once again, failure of the analysis to satisfy assumptions of normality and homoscedasticity compromises the computation of confidence intervals for parameters in the fitted model but it does not invalidate the parameter estimates themselves (Myers, 1986; Finney, 1989; Asselman, 2000; Cox *et al.*, 2008).

### Data exclusive of elephant

The initial analysis of arithmetic values by nonlinear regression pointed to the hippo as a statistical outlier (Fig. 1b). However, the Studentized deleted residual for the elephant also was suspicious ( $-6.18$ ), and Cook's distance (1028.61) indicated that the elephant had a major influence on parameters in the two-parameter power function (Fig. 1a). Is it possible that the extraordinary leverage exerted by the data point for the elephant caused the fitted line to depart from a pattern characterizing other data and thereby give the erroneous impression that the hippo was an outlier? To address this question, we deleted the elephant and performed new analyses on values for the remaining 32 species.

The two-parameter power function fitted to arithmetic values by nonlinear regression was statistically significant, with variation in the predictor variable explaining *c.* 95% of the variation in the response variable (Fig. 3a). However, the



**Figure 3** (a) Body masses for quadrupedal mammals exclusive of the elephant are plotted against values for circumference of humerus plus femur (Anderson *et al.*, 1985). The summary statistics are from a nonlinear regression. The open circle here and elsewhere shows the location of the elephant. (b) Standardized residuals from the analysis summarized in (a) are plotted against predicted values for body mass. The giraffe is a possible outlier. (c)  $\log_{10}$  transformations for body mass are plotted against  $\log_{10}$  transformations for humerus plus femur for species exclusive of the elephant. The summary statistics are from ordinary least squares regression. The giraffe does not appear at all unusual in this graphic. (d) Equations obtained by nonlinear regression (solid line) and back-transformation (dashed line) are shown against the backdrop of values expressed on the scale of measurement. The dash-dot line depicts the equation obtained by fitting a generalized linear model using Proc Glimmix in SAS 9.2.

analysis failed tests for normality of residuals and constancy of variances ( $P < 0.001$ ), and the giraffe appeared in this analysis to be a possible outlier (Fig. 3b). In contrast, the linear regression on log transformations yielded a statistically significant model (Fig. 3c) that met assumptions of normality and homoscedasticity ( $P \geq 0.137$ ), and residuals were satisfactory (not shown).

Graphing the equations obtained by nonlinear regression and back-transformation on a plot with arithmetic values revealed once again that both methods yielded good fits to data for animals of small and intermediate size (Fig. 3d). However, the model estimated by back-transformation did not describe the hippo, whereas that from nonlinear regression did (Fig. 3d). Indeed, the only problem with the nonlinear regression was the slightly discrepant value for the giraffe, yet this discrepancy was not nearly as serious as was deviation of the hippo from the model estimated by back-transformation (Fig. 3d).

The traditional method for fitting an allometric equation predicts geometric means, which are always smaller than corresponding arithmetic means (Smith, 1993; Hayes & Shonkwiler, 2006). Consequently, the underestimate for mass of the hippo (Fig. 3d) could represent nothing more than this form of transformation bias. To address this possibility, we also examined the data with Proc Glimmix in SAS version 9.2. Glimmix fits a generalized linear model to allometric data using a log link and a gamma function (Lane, 2002; Cox *et al.*, 2008). Such an analysis predicts logarithms for means of  $Y$  instead of means for the logarithms, so back-transformation from the fitted equation does not introduce the kind of transformation bias that commonly plagues the traditional approach (Lane, 2002; Cox *et al.*, 2008).

The equation fitted by Proc Glimmix,

$$\hat{Y} = 0.071X^{2.779}$$

differs only slightly from the equation fitted by the traditional method,

$$\hat{Y} = 0.071X^{2.771}$$

Plotting the equation obtained by fitting a generalized linear model shows that the equation provides only a marginal improvement over the traditional method (Fig. 3d).

## Discussion

### The reference data

The dataset compiled by Anderson *et al.* (1985) was comprised of values for the largest specimen for each of the 33 species in the sample. Thus, none of the reference values is likely to be representative for any of the species in question (Smith, 2002), thereby casting doubt on the reliability of any allometric equation fitted to the data. However, we do not wish to dwell on limitations of the reference sample. Our objective here is to use the data to illustrate the pitfalls awaiting investigators who use the traditional method to fit an equation for predicting body size of extinct species and also to illustrate the importance of validating statistical models in the arithmetic scale. Additionally, we show more specifically that the allometric equation reported by Anderson *et al.* (1985) is not a good descriptor for their reference data and, thus, that the equation cannot be a reliable predictor for body mass of dinosaurs or other large animals in extinct lineages.

### The question of outliers

Identifying outliers is a tricky business. If you can identify one outlier, you may be able to identify a second, and the second can lead to a third and so on until the dataset has been so altered that it bears little resemblance to the original. Conclusions and predictions may suffer accordingly. On the other hand, failure to identify outliers can be equally devastating (Anscombe, 1973; Stevens, 1984; Osborne & Overbay, 2004). Outliers often are hidden from view when data are expressed in logarithmic form (Figs 1c and 2c), yet their presence can cause equations fitted in the traditional way to be extraordinarily misleading (Packard & Boardman, 2008; Packard, 2009). Thus, a search for outliers is both warranted and necessary, but the search also must be conservative and judicious.

The problem of potential outliers was evident in the current study in the form of the hippo, the bison and possibly the elephant. A case can be made on statistical grounds to focus on the entire dataset, or to exclude values for the hippo, for the hippo and bison or for the elephant. Deletion of potential outliers elicited apparent changes in parameters in power functions fitted by nonlinear regression but not in those estimated by back-transformation from lines fitted to logarithms (Figs 1d, 2d and 3d). This is because relatively small species in the sample were the primary determinants of parameters in the equations esti-

mated by back-transformation; information on large species like the hippo and elephant was largely irrelevant. Consequently, models obtained by back-transformation consistently failed to describe mass for the largest species (Figs 1d, 2d and 3d), regardless of which of the largest forms was retained in the dataset and which was deleted. Failure of the traditional method to incorporate information from large species probably is commonplace in allometric research (Glass, 1969; Hui & Jackson, 2007; Packard & Birchard, 2008; Packard, 2009; Packard & Boardman, 2009a) and calls into question the general utility of allometric equations that have been estimated by back-transformation from logarithms.

In contrast, information for the largest species was incorporated into all the equations fitted by nonlinear regression. As a result, parameters in the fitted equations varied appreciably among the several analyses, depending on which of the largest species was included and which was excluded (Figs 1a, 2a and 3a). A decision to treat any one of the largest forms as an outlier clearly would have a major impact on the ensuing analysis, so such a decision would need to be carefully considered and fully justified. However, all the nonlinear regressions afforded better fits to the respective datasets than did any of the equations resulting from application of the traditional method (Figs 1d, 2d and 3d).

### Predicting body mass of dinosaurs

We used predictive equations from both nonlinear regression and back-transformation to estimate body masses for the same sample of quadrupedal dinosaurs that was the focus for the investigation by Anderson *et al.* (1985). The equations were from analyses of the full dataset (Fig. 1) because Anderson *et al.* (1985) used the full dataset in their study. Our equation from back-transformation differs somewhat from theirs, owing apparently to an error by them in data entry or computation. Alexander (1989) seems to have detected this same error, because the equation reported by him is identical to ours.

Predictions based on skeletal measurements falling within the range of values for mammals in the reference sample are most important for assessing the reliability of the alternative statistical models of concern here. The two smallest dinosaurs in the sample compiled by Anderson *et al.* (1985) are therefore of special interest, because both have dimensions for humerus plus femur within that range. Predictions from the equation fitted by nonlinear regression are substantially lower – by 21% for *Styracosaurus* and 27% for *Diplodocus* – than those from the equation obtained by back-transformation from logarithms (Table 1). The mass predicted for the elephant by nonlinear regression also is much lower than the mass predicted by back-transformation, but the former estimate is very near to the value actually recorded for this species by Anderson *et al.* (1985) (Table 1). If data for the elephant are reliable, then failure of the traditional method to yield a plausible prediction for mass of this animal indicates that the procedure did not yield a valid predictive

**Table 1** Predicted body mass for quadrupedal dinosaurs and the elephant

Species	Humerus + femur (mm)	Mass (kg) predicted by equation obtained by back-transformation from logs	Mass (kg) predicted by equation fitted by nonlinear regression
<i>Styracosaurus albertensis</i>	658	4200 (3900)	3300
<i>Diplodocus</i> sp.	725	5500 (5000)	4000
Elephant	872	9200 (8300)	5900
<i>Opisthocoelicaudia skarzynskii</i>	1245	24 000 (22 000)	13 000
<i>Lourinhasaurus</i> (= <i>Apatosaurus</i> ) <i>alenquerensis</i>	1332	29 000 (26 000)	15 000
<i>Giraffatitan</i> (= <i>Brachiosaurus</i> ) <i>brancai</i>	1384	32 000 (29 000)	16 000
<i>A. louisae</i>	1474	38 000 (35 000)	18 000

One quadruped that was listed in the original table is not included here owing to uncertainty about measurements of its humerus and femur. Predictions inside parentheses were taken directly from table 4 in Anderson *et al.* (1985) or were computed with their allometric equation 1(b). Equation 1 in Anderson *et al.* (1985) apparently is incorrect (see also Alexander, 1989), which accounts for differences between estimates based on the traditional procedure. All values for mass were rounded to conform with the original report. Actual mass recorded for the elephant was 5896.7 kg (Anderson *et al.*, 1985).

equation. If data for the elephant are judged to be unreliable, however, then a different comparison is in order, namely, one involving predictions from equations fitted to data exclusive of the elephant (Fig. 3). The problems attending interpolation from the equation estimated by back-transformation nonetheless remain (Fig. 3d).

Our estimates for mass of the four largest dinosaurs in the sample assembled by Anderson *et al.* (1985) are also substantially lower than those based on back-transformation from logarithmic form (Table 1). Moreover, the difference between predictions based on nonlinear regression and those based on back-transformation from logarithms increases with apparent size of the animals in question, owing to the divergence of the fitted curves with increasing values for the predictor variable (Fig. 1d). Indeed, our prediction of 18 000 kg for mass of the *Apatosaurus* is less than half the mass estimated by the traditional procedure and only slightly more than half the mass estimated by Anderson *et al.* (1985). We cannot say unequivocally that one model is better than the other for predicting mass of extraordinarily large dinosaurs because both sets of estimates were based on extrapolations beyond the range of data in the reference sample (Smith, 2002). As noted in the preceding paragraph, however, predictions based on interpolation from the nonlinear regression yielded a better outcome than did interpolation from the equation estimated in the traditional manner, so we expect that the nonlinear fit also yielded more reliable estimates when extrapolating to the limits of body size.

Given the generally poor performance by the equation obtained by back-transformation from logarithmic form, we conclude that estimates by Anderson *et al.* (1985) were not supported adequately by the dataset on which the estimates were based. The preceding conclusion has important implications because masses of numerous dinosaurs – both quadrupedal and bipedal – have been estimated by investigators who applied some version of the equation derived by Anderson *et al.* (1985). For example, the equation has been used in recent years to predict asymptotic body mass for growing dinosaurs (e.g. Erickson *et al.*, 2004;

Bybee, Lee & Lamm, 2006; Lee & Werning, 2008; Lehman & Woodward, 2008). If the estimates for asymptotic mass are too high (Table 1), the models for growth are themselves suspect. In particular, estimates for maximum rates of growth in dinosaurs and sizes of animals at reproductive maturity may be excessive and misleading. Although the studies of growth have suggested that the process in dinosaurs was fundamentally different from that of contemporary lizards and crocodilians (Erickson, 2005), the question may warrant reexamination in light of our findings.

### Why the difference between methods for fitting models?

Investigators in other disciplines have championed the use of generalized linear models to fit equations to allometric data (Lane, 2002; Cox *et al.*, 2008). Such statistical routines have an advantage over the traditional procedure for allometric analysis in that a generalized linear model predicts logs for means of the response variable *Y* instead of means for the logs (Lane, 2002; Cox *et al.*, 2008). Consequently, back-transformation of the resulting equation yields a model that predicts arithmetic means instead of geometric means and thereby obviates the use of correction factors to try to remove transformation bias (Smith, 1993; Hayes & Shonkwiler, 2006). However, even when a generalized linear model was fitted to data in the current investigation, the resulting equation was inadequate (Fig. 3d).

If the problem with the traditional procedure cannot be traced to the bias resulting from predicting geometric means vis-à-vis arithmetic means, what is the explanation? We suggest that the explanation lies in performing least squares regression on log transformations, because the traditional method and the generalized linear model are both procedures for fitting models to logarithms. Transformations greatly alter the relationship between predictor and response variables so that least squares regression (of any kind) exaggerates the importance of small values in the arithmetic scale and simultaneously underweight large values (Jansson, 1985; Packard & Birchard, 2008; Packard & Boardman,



2009a). Small values for the original measurements have an undue influence on parameters in the straight line fitted to logs, and large values have only minimal influence (Jansson, 1985; Packard & Birchard, 2008). This is why deleting the hippo and elephant from the dataset used here had very little effect on parameters in the lines fitted to logarithms. When the equations for the lines were back-transformed to the original scale, the bias was carried along into the resultant two-parameter power functions.

## Acknowledgments

We thank J.R. zumBrunnen for helping us to create and then decipher programming code for Proc Glimmix in SAS. Two anonymous referees made several helpful suggestions that enabled us to improve the paper considerably.

## References

- Alexander, R.M. (1989). *Dynamics of dinosaurs and other extinct giants*. New York: Columbia University Press.
- Anderson, J.F., Hall-Martin, A. & Russell, D.A. (1985). Long-bone circumference and weight in mammals, birds and dinosaurs. *J. Zool. (Lond.)* **207**, 53–61.
- 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.
- Bybee, P.J., Lee, A.H. & Lamm, E.-T. (2006). Sizing the Jurassic theropod dinosaur *Allosaurus*: assessing growth strategy and evolution of ontogenetic scaling of limbs. *J. Morphol.* **267**, 347–359.
- Cox, N.J., Warburton, J., Armstrong, A. & Holliday, V.J. (2008). Fitting concentration and load rating curves with generalized linear models. *Earth Surf. Process. Landforms* **33**, 25–39.
- Emerson, J.D. & Stoto, M.A. (1983). Transforming data. In *Understanding robust and exploratory data analysis*: 97–128. Hoaglin, D.C., Mosteller, F. & Tukey, J.W. (Eds). New York: Wiley.
- Erickson, G.M. (2005). Assessing dinosaur growth patterns: a microscopic revolution. *Trends Ecol. Evol.* **20**, 677–684.
- Erickson, G.M., Makovicky, P.J., Currie, P.J., Norell, M.A., Yerby, S.A. & Brochu, C.A. (2004). Gigantism and comparative life-history parameters of tyrannosaurid dinosaurs. *Nature* **430**, 772–775.
- Farlow, J.O., Dodson, P. & Chinsamy, A. (1995). Dinosaur biology. *Annu. Rev. Ecol. Syst.* **26**, 445–471.
- Finney, D.J. (1989). Was this in your statistics textbook? V. Transformation of data. *Exp. Agric.* **25**, 165–175.
- Glass, N.R. (1969). Discussion of calculation of power function with special reference to respiratory metabolism in fish. *J. Fish. Res. Board Can.* **26**, 2643–2650.
- Gunga, H.-C., Suthau, T., Bellmann, A., Friedrich, A., Schwanebeck, T., Stoinski, S., Trippel, T., Kirsch, K. & Hellwich, O. (2007). Body mass estimations for *Plateosaurus engelhardti* using laser scanning and 3D reconstruction methods. *Naturwissenschaften* **94**, 623–630.
- Hayes, J.P. & Shonkwiler, J.S. (2006). Allometry, antilog transformations, and the perils of prediction on the original scale. *Physiol. Biochem. Zool.* **79**, 665–674.
- Henderson, D.M. (1999). Estimating the masses and centers of mass of extinct animals by 3-D mathematical slicing. *Paleobiology* **25**, 88–106.
- Hui, D. & 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. *Geograph. Ann.* **67A**, 61–70.
- Kutner, M.H., Nachtsheim, C.J. & Neter, J. (2004). *Applied linear regression models*, 4th edn. Boston: McGraw-Hill Irwin.
- Lane, P.W. (2002). Generalized linear models in soil science. *Eur. J. Soil Sci.* **53**, 241–251.
- Lee, A.H. & Werning, S. (2008). Sexual maturity in growing dinosaurs does not fit reptilian growth models. *Proc. Natl. Acad. Sci. USA* **105**, 582–587.
- Lehman, T.M. & Woodward, H.N. (2008). Modeling growth rates for sauropod dinosaurs. *Paleobiology* **34**, 264–281.
- Marquardt, D.W. (1963). An algorithm for least-squares estimation of nonlinear parameters. *J. Soc. Ind. Appl. Math.* **11**, 431–441.
- McCuen, R.H., Leahy, R.B. & Johnson, P.A. (1990). Problems with logarithmic transformations in regression. *J. Hydraul. Eng.* **116**, 414–428.
- Motulsky, H. & 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**, <http://PAREonline.net/getvn.asp?v=8&n=6>
- Osborne, J.W. & Overbay, A. (2004). The power of outliers (and why researchers should ALWAYS check for them). *Pract. Assess. Res. Eval.* **9**, <http://PAREonline.net/getvn.asp?v=9&n=6>
- Packard, G.C. (2009). On the use of logarithmic transformations in allometric analyses. *J. Theor. Biol.* **257**, 515–518.
- Packard, G.C. & Birchard, G.F. (2008). Traditional allometric analysis fails to provide a valid predictive model for mammalian metabolic rates. *J. Exp. Biol.* **211**, 3581–3587.
- Packard, G.C. & Boardman, T.J. (2008). Model selection and logarithmic transformation in allometric analysis. *Physiol. Biochem. Zool.* **81**, 496–507.
- Packard, G.C. & Boardman, T.J. (2009a). A comparison of methods for fitting allometric equations to field metabolic rates of animals. *J. Comp. Physiol. B* **179**, 175–182.
- Packard, G.C. & Boardman, T.J. (2009b). Bias in interspecific allometry: examples from morphological scaling in varanid lizards. *Biol. J. Linn. Soc.* **96**, 296–305.



- Pandey, G.R. & Nguyen, V.-T.-V. (1999). A comparative study of regression based methods in regional flood frequency analysis. *J. Hydrol.* **225**, 92–101.
- Seebacher, F. (2001). A new method to calculate allometric length–mass relationships of dinosaurs. *J. Vertebr. Paleontol.* **21**, 51–60.
- 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.
- Smith, R.J. (2002). Estimation of body mass in paleontology. *J. Human Evol.* **43**, 271–287.
- 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.
- Zar, J.H. (1968). Calculation and miscalculation of the allometric equation as a model in biological data. *BioScience* **18**, 1118–1120.