

Model Selection and Logarithmic Transformation in Allometric Analysis

Gary C. Packard^{1,*}

Thomas J. Boardman²

¹Department of Biology, Colorado State University, Fort Collins, Colorado 80523-1878; ²Department of Statistics, Colorado State University, Fort Collins, Colorado 80523-1877

Accepted 10/16/2007; Electronically Published 5/30/2008

ABSTRACT

The standard approach to most allometric research is to gather data on a biological function and a measure of body size, convert the data to logarithms, display the new values in a bivariate plot, and then fit a straight line to the transformations by the method of least squares. The slope of the fitted line provides an estimate for the allometric (or scaling) exponent, which often is interpreted in the context of underlying principles of structural and functional design. However, interpretations of this sort are based on the implicit assumption that the original data conform with a power function having an intercept of 0 on a plot with arithmetic coordinates. Whenever this assumption is not satisfied, the resulting estimate for the allometric exponent may be seriously biased and misleading. The problem of identifying an appropriate function is compounded by the logarithmic transformations, which alter the relationship between the original variables and frequently conceal the presence of outliers having an undue influence on properties of the fitted equation, including the estimate for the allometric exponent. Much of the current controversy in allometric research probably can be traced to substantive biases introduced by investigators who followed standard practice. We illustrate such biases with examples taken from the literature and outline a general methodology by which the biases can be minimized in future research.

Introduction

Analyses of interspecific allometry (or scaling) have become so commonplace that a day seldom passes without a report of

such research appearing on the pages of some biological journal or another (Web of Science). The primary goal in much of this work is to characterize patterns of variation in morphology, physiology, and/or ecology in organisms spanning a broad range in body size in an attempt to identify underlying principles in the design of biological systems (Brown et al. 2004 and references therein). Unfortunately, many of the patterns identified by this research are inaccurate and misleading, and these mischaracterizations likely contribute to the ongoing debate about ways in which animals are constructed (Agutter and Wheatley 2004; Glazier 2005; West and Brown 2005; White and Seymour 2005a; da Silva et al. 2006).

Allometric relationships typically are represented by a simple, two-parameter power function,

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

where Y is the morphological, physiological, or ecological variable of interest and X is the measure of body size. The parameters a and b are better known as the allometric coefficient and the allometric exponent, respectively (Gould 1966; Peters 1983; Calder 1984; Schmidt-Nielsen 1984). It is unclear whether this equation emerged from strictly theoretical considerations, whether it was adopted because of the relative ease in manipulating logarithmic transformations of empirical data conforming with such a function, or whether the answer lies in some combination of these factors (Kleiber 1961; Gould 1966; Hayes and Shonkwiler 2001). Regardless of the origin of the equation, however, values for the so-called dependent and independent variables seldom are examined in their original arithmetic scale but instead are immediately transformed to their logarithms (Smith 1980, 1984; Peters 1983; Calder 1984; Schmidt-Nielsen 1984), at which point the allometric equation is expressed as

$$\log Y = \log a + b \log X.$$

A straight line is then fitted to the transformations, usually by the method of least squares. If the computed value for the coefficient of determination, R^2 , is sufficiently high, the fit is accepted, and discussion generally turns to the putative importance of the estimate for the allometric exponent b (Smith 1980, 1984). The value of this exponent is at the center of the aforementioned debate on functional design (Agutter and Wheatley 2004; Glazier 2005; West and Brown 2005; White and Seymour 2005a; da Silva et al. 2006).

This standard approach to allometric research is based on the critical assumption that the original data conform with a power function having an intercept of 0 in a plot with arith-

* Corresponding author. Present address: 865 Three Corner Gate Road, Livermore, Colorado 80536; e-mail: packard@lamar.colostate.edu.

metric coordinates (Zar 1968). However, few investigators take time to perform graphical analyses of data displayed on the original arithmetic scale (Smith 1980, 1984), so they cannot determine whether the assumption actually is satisfied. The assumption may be wrong (e.g., Gould 1966; Calder 1984), and its failure can introduce a substantial bias into an analysis. The fundamental problem of identifying an appropriate function is compounded by the logarithmic transformations, which alter the relationship between the original variables (Peters 1983; Jansson 1985; Smith 1993; Osborne 2002) and may thereby conceal the presence of statistical outliers that also bias estimates for the parameters a and b in equation (1) (Anscombe 1973; Stevens 1984; Osborne and Overbay 2004). Finally, fitting a straight line to logarithmic transformations of the data by the least squares method invokes a statistical model with multiplicative error,

$$Y = aX^b 10^\varepsilon,$$

whereas a least squares fit to the original data invokes a model with additive error,

$$Y = aX^b + \varepsilon$$

(Glass 1969; Manaster and Manaster 1975; Smith 1993). Differences in the way by which the models are fitted to data typically lead to additional bias when an equation fitted to logarithmic transformations is used to estimate parameters of a two-parameter power function (Glass 1969; Miller 1984; Jansson 1985; McCuen et al. 1990; Smith 1993; Pandey and Nguyen 1999; Hayes and Shonkwiler 2001, 2006).

Here we use two hypothetical examples and three drawn from the literature to illustrate how the standard allometric analysis using logarithmic transformations can lead investigators to conclusions that are not well supported by the original data. In so doing, we also outline a general approach to allometric analysis that conforms better than the standard approach with recommendations of applied statisticians. Our emphasis is on fitting the most appropriate mathematical function to a data set and on identifying and removing influential outliers. In addition to the alternative forms of the two-parameter power function already mentioned, we also consider a three-parameter model with additive error,

$$Y = Y_0 + aX^b + \varepsilon,$$

because the fit by a three-parameter model is not constrained to pass through the origin and consequently affords potential benefits with regard to residuals (Manaster and Manaster 1975). Straight lines are treated as special cases of two- and three-parameter power functions. We made no attempt to correct the analyses for phylogeny (Martin et al. 2005) or for shortcomings of least squares regression (Warton et al. 2006), nor did we address issues pertaining to sample size (White and Seymour 2005b). Although all these are legitimate concerns,

we believe that those on which we focus here are even more fundamental (Zar 1968; Manaster and Manaster 1975; Hayes and Shonkwiler 2001).

Hypothetical Example 1: Fitting a Power Function to Linear Data

We generated three simplistic sets of hypothetical values to illustrate the problem that results from use of an inappropriate function to describe data in allometric research (Fig. 1). Because straight lines sometimes fit empirical data better than power functions do (Gould 1966; Calder 1984), each of these data sets conformed with a model of a straight line having a slope of 7 in an arithmetic plot (Fig. 1A, 1C, 1E). Whereas a line fitted by ordinary least squares to the first set of values passes through the origin of the graph (Fig. 1A), the second and third sets have intercepts of 20 and 40 (Fig. 1C, 1E, respectively). Values for R^2 are 1, and the allometric exponents b also are 1, irrespective of the intercept, because X is implicitly raised to the power of 1 in every case (Fig. 1A, 1C, 1E).

Log transformations of the original data also are described quite well by straight lines (Fig. 1B, 1D, 1F), as is indicated by the fact that values for R^2 approach or equal 1. The allometric exponent equals 1 in the case of the data set having an intercept of 0 on the original scale (Fig. 1B). However, the estimate for the allometric exponent is 0.95 in the analysis of log-transformed data for the second example (Fig. 1D), and it is 0.91 in the analysis of transformed values in the third example (Fig. 1F).

The second and third of these hypothetical examples were biased by implicitly fitting a power function with an intercept of 0 (Fig. 1D, 1F) to data conforming with an equation for a straight line with a nonzero intercept (Fig. 1C, 1E). The bias in these examples probably would have gone undetected had we not examined bivariate plots of data on an arithmetic scale (Fig. 1C, 1E) because of the deceptively high values for R^2 in the log-log regressions (Fig. 1D, 1F). Although R^2 provides an indication for how much of the variation in $\log Y$ can be explained by variation in $\log X$, it is not a reliable index to the strength of the relation between Y and X in their original scale, nor does it validate the underlying mathematical function (McCuen et al. 1990).

Analyses of untransformed and log-transformed data produced the same estimates for the allometric exponent in the first of the hypothetical examples (Fig. 1A, 1B), but this exceptional outcome can be traced to the 0 intercept for the untransformed values, coupled with the unrealistic R^2 of 1 (Glass 1969; Jansson 1985). In this special case, the equations for the two-parameter power function and the straight line assume the same general form (i.e., $Y = aX^1$).

Fitting an Inferior Function: Head Width in Varanid Lizards

We draw on a study of interspecific variation in morphology of varanid lizards from Western Australia (Thompson and

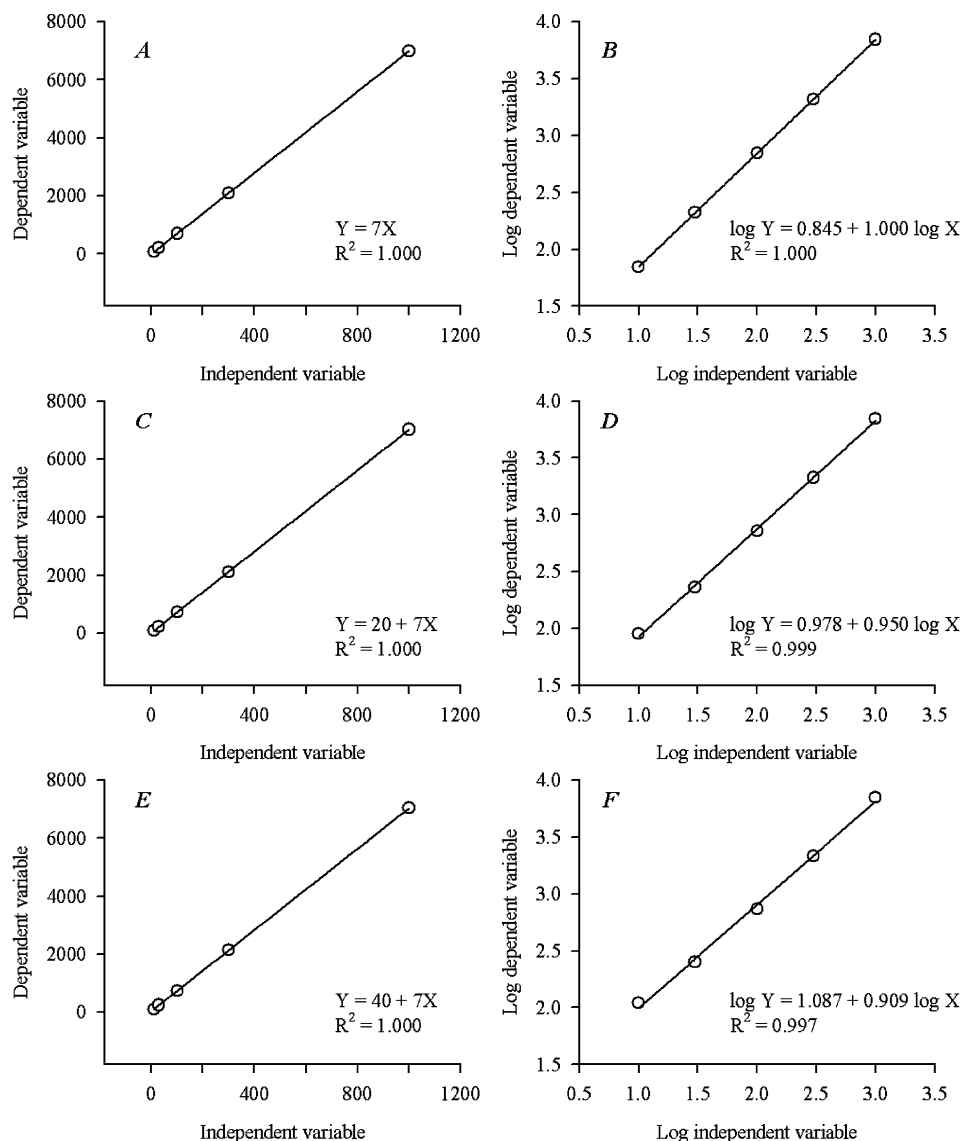


Figure 1. Three sets of simplistic, hypothetical data illustrate the importance of the assumption that data conform with a power function having an intercept of 0. The data sets differ only in the intercept for straight lines fitted to values displayed in arithmetic plots (*left*). Scatterplots in the right-hand column display the corresponding values after transformation to their common logarithms. Lines were fitted by least squares using SigmaPlot 10. A, B, Intercept on original scale = 0. C, D, Intercept = 20. E, F, Intercept = 40. The allometric exponent is 1 in all three examples using untransformed data (*left*), but corresponding estimates for that exponent decline from 1 to 0.91 in analyses of logarithms (*right*).

Withers 1997) for data to illustrate the potential for introducing bias into an allometric analysis by fitting a two-parameter power function to values that are better described by a straight line. Varanids are excellent candidates for allometric research because members of the family share a similar, evolutionarily conservative morphology and vary appreciably in body size (Greer 1989; Pianka 1995).

Means for head width (which is an index to size of the gape) and body length (i.e., thorax plus abdomen) for 18 species/subspecies of varanids first were displayed in a scatterplot with arithmetic coordinates and then studied visually for patterns

and trends (Fig. 2A). The plot reveals that width of the head tends to increase linearly with body size, and linear regression by least squares (using SigmaPlot 10.0) yielded a statistically significant fit ($F_{1,16} = 575$, $P < 0.001$, $R^2 = 0.973$). The data satisfied assumptions concerning normality of residuals (Kolmogorov-Smirnov test) and homoscedasticity (Spearman rank correlation between absolute values of residuals and observed values for body size; see Kutner et al. 2004), and no pattern is apparent in a plot of Studentized residuals against body size (Fig. 2B). The highest value for a Studentized deleted residual (which is computed by dividing a raw residual by the

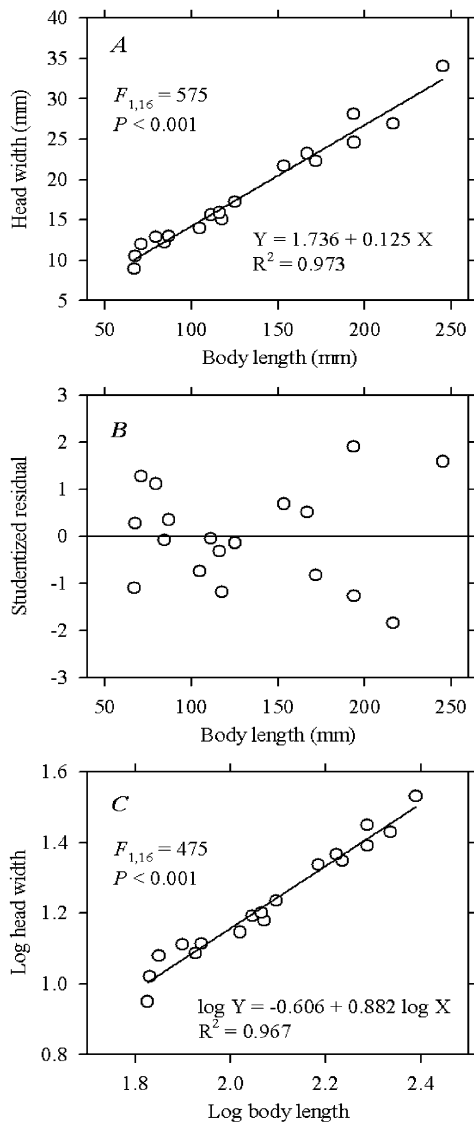


Figure 2. A, Means for head width and body length (thorax plus abdomen) for 18 species/subspecies of varanid lizards (Thompson and Withers 1997) in their original units. Statistics are from the regression ANOVA, in which a straight line was fitted by the method of least squares. B, Studentized residuals from the analysis in A plotted against body length. These residuals were computed by dividing each raw residual by an estimate for its standard deviation (Kutner et al. 2004). Approximately equal numbers of residuals are positive and negative, and no size-related pattern is apparent. C, Bivariate plot of data after transformation to common logarithms. Statistics are from the regression ANOVA, in which a straight line was fitted by the method of least squares.

standard error of the estimate for a regression that excludes the value in question; Kutner et al. 2004) was only 2.10, indicating that none of the data points is likely to be a statistical outlier. The maximum value for Cook's distance (which is a measure of the influence of an individual data point on estimates for slope and intercept; Kutner et al. 2004) was an acceptable 0.55 (where a value under 1 identifies a point having

little influence on slope and intercept for the equation and a value greater than 4 identifies a point having major influence). The model seems, therefore, to be appropriate on both visual and statistical grounds, and the resulting estimate for the allometric exponent is 1 (because $\hat{Y} = 1.736 + 0.125X^1$).

We also fitted two- and three-parameter power functions to the data because power functions converge on straight lines as the allometric exponent b approaches 1. For example, a two-parameter power function assumes the form of a straight line when the intercept is 0 and the allometric exponent is 1, and a three-parameter power function assumes the form of a straight line when the intercept is other than 0 and the allometric exponent is 1. Thus, there may be instances where fitting a linear equation and a power function will lead to the same outcome.

The two-parameter function fitted by the method of least squares (using SigmaPlot 10.0) yielded a seemingly good fit ($\hat{Y} = 0.214X^{0.912}$, $F_{1,16} = 545$, $P < 0.001$, $R^2 = 0.972$), and so, too, did the three-parameter function ($\hat{Y} = 4.753 + 0.032X^{1.231}$, $F_{2,15} = 286$, $P < 0.001$, $R^2 = 0.974$). Both these analyses also satisfied assumptions concerning normality of residuals and homoscedasticity, and no pattern was apparent in plots of residuals against body size (not shown). However, the predicted residual error sum of squares (PRESS), which is superior to both R^2 and the residual sum of squares for assessing goodness of fit (Allen 1971), was 34.62 for the two-parameter power function and 36.09 for the three-parameter power function. In contrast, the PRESS for the straight line was 31.63. Lower values for PRESS scores identify better models (Kutner et al. 2004), so the straight line affords an 8.6% improvement over the two-parameter power function and a 12.4% improvement over the three-parameter function. Thus, whereas any one of the equations describes the data adequately, the straight line is the best of the three.

We next transformed the data to their logarithms (despite the fact that such a transformation was not needed to meet assumptions concerning normality of residuals or homoscedasticity), displayed the transformations in a scatterplot, and fitted a straight line by the method of least squares (Fig. 2C). Assumptions of normality and homoscedasticity again were satisfied, and the regression was statistically significant ($F_{1,16} = 475$, $P < 0.001$, $R^2 = 0.967$). The resulting estimate for the allometric exponent was 0.882, which agrees with the value reported by Thompson and Withers (1997). However, the estimate based on analyses of log transformations differs substantially from the value of 1 derived from the analysis of untransformed data (Fig. 2A).

The exponent of 0.88 obtained by Thompson and Withers (1997) led them to conclude that head width varies allometrically (instead of isometrically) with body size among varanids from Western Australia. The linear regression reported here for the untransformed data (Fig. 2A) also supports the idea that variation is allometric, but the cause for allometry stems in this instance from the nonzero intercept instead of an allometric exponent differing from 1 (Gould

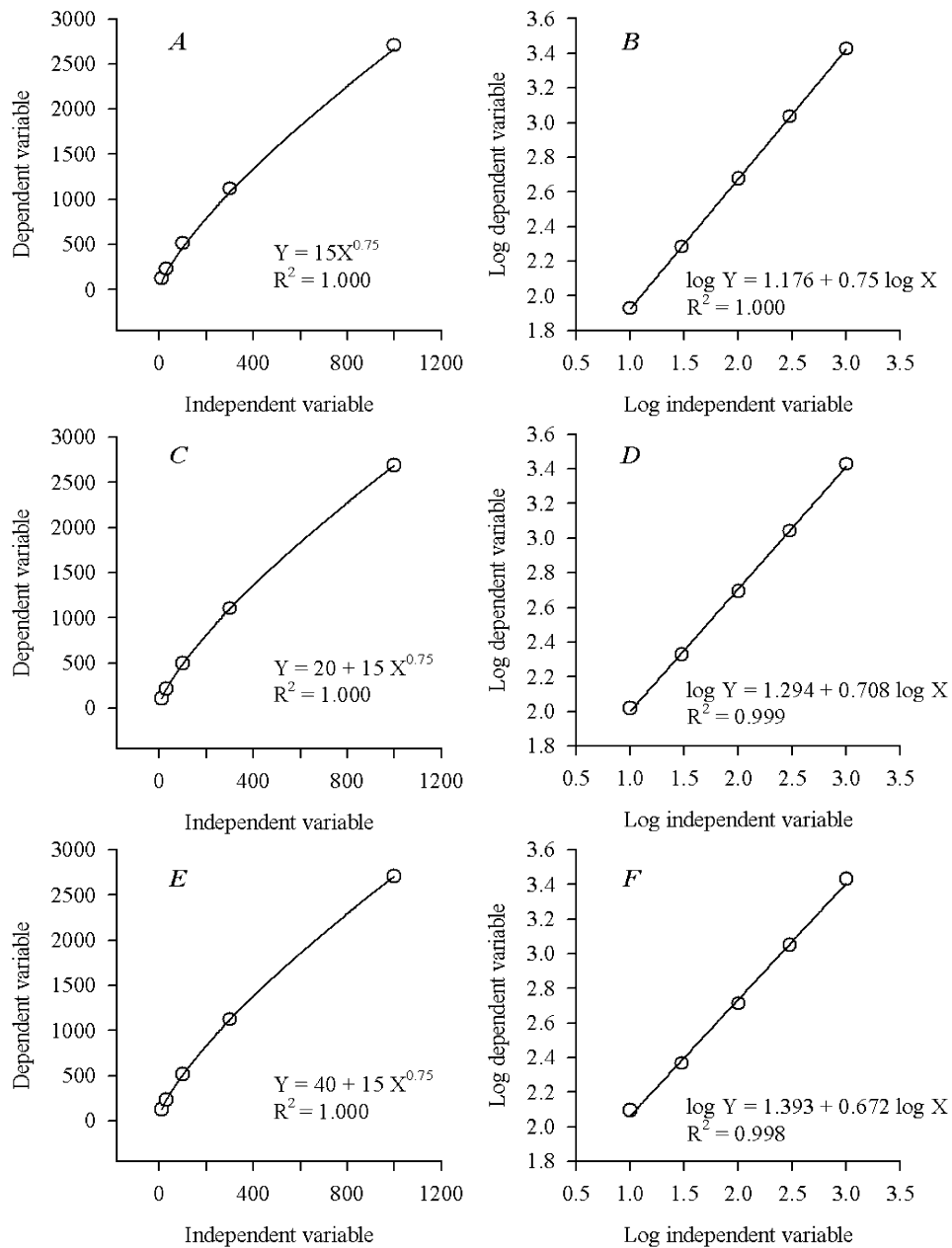


Figure 3. Three additional sets of simplistic, hypothetical data further illustrate the importance of the assumption that data conform with a power function having an intercept of 0. The data sets differ only in the intercept for power functions fitted to values displayed in arithmetic plots (*left*). Scatterplots in the right-hand column display the corresponding values after transformation to their common logarithms. Lines were fitted by least squares using SigmaPlot 10. A, B, Intercept on original scale = 0. C, D, Intercept = 20. E, F, Intercept = 40. The allometric exponent is 0.75 in all three examples using untransformed data (*left*), but corresponding estimates for that exponent decline from 0.75 (the “3/4-power law”) to 0.67 (the “surface law”) in analyses of logarithms (*right*).

1966; Packard and Boardman 1988). Concepts of size-related variation in width of the head of these lizards thus depend on which of the analyses is adopted. The analysis based on data in their original scale, however, is better than the one based on logarithmic transformations, because the latter represents the fit of an inferior model (i.e., a two-parameter power function with multiplicative error).

Hypothetical Example 2: Fitting a Power Function to Nonlinear Data

Even when a power function is an appropriate model to fit to allometric data, the standard protocol implicitly requires that the resulting regression pass through the origin of a graph in the original arithmetic scale (Fig. 3A, 3B). Any departure of

the intercept from 0 again has the potential to introduce a serious bias into the analysis (Fig. 3C–3F), but the actual intercept may be close enough to 0 that it has little influence on the estimate for the scaling exponent. However, the location of the intercept cannot be determined without doing a preliminary graphical (and possibly statistical) analysis of data in their original scale (see Anscombe 1973; Tukey 1977; Tufte 1983). Moreover, graphical analysis of data in their original state can reveal significant details that are not clearly evident from an examination of a plot of logarithmic transformations. This is especially important with regard to outliers and their influence on the ultimate outcome of an allometric analysis (Anscombe 1973; Stevens 1984; Osborne and Overbay 2004).

Fitting an Appropriate Function to Data with Outliers: Standard Metabolic Rate of Passerine Birds

This second case study is based on data compiled by Lasiewski and Dawson (1967) for their pivotal investigation of standard metabolic rate in passerine birds. This compilation has been the focus for a valuable debate concerning procedures for fitting allometric equations to empirical data (Zar 1968, 1970; Lasiewski and Dawson 1969; Manaster and Manaster 1975).

We begin as before with a bivariate plot of values in their original scale so that we can gain some initial impressions concerning the distribution of data points and the relationship between dependent and independent variables (Fig. 4A). A nonlinear regression appears to be in order, albeit measurements of metabolism for two samples of large passerines appear to be unusually high (Fig. 4A, *triangles*). Two- and three-parameter power functions next were fitted to the data by the iterative least squares procedure. Both the resulting equations were statistically significant ($P < 0.001$), but neither of them passed tests for normality of residuals or homoscedasticity.

The two noteworthy data points (Fig. 4A, *triangles*) lie well above the resulting regression lines, and the Studentized residuals for these samples are unusually high (Fig. 4B). Consequently, the values in question may be statistical outliers. This possibility is supported by consideration of the Studentized deleted residuals, which are 3.26 and 10.57 in the fit of the two-parameter model and 2.58 and 10.16 in the fit of the three-parameter function. Further inspection revealed that the data points in question were for the same species (*Corvus corvinus*) in summer and winter and that both were reported by the same laboratory. Such findings raise the possibility that the metabolic rate of this species is unusually high for a passerine of this body size or that a flaw in protocol biased the measurements. We therefore treated these values as outliers, deleted them from the data set, and then repeated the analyses.

Both the two-parameter power function ($\hat{Y} = 104.1X^{0.665}$) and the three-parameter power function ($\hat{Y} = -0.2 + 104.2X^{0.658}$) fitted to data exclusive of the two suspect values passed tests for normality of residuals and homoscedasticity. The PRESS score for the two-parameter model was 6.7% lower than that for the three-parameter function, thereby pointing

to a slightly better fit by the former than by the latter. Because the two-parameter function also is the simpler of the two, we focus on this model (Fig. 4C). Examination of Studentized deleted residuals and Cook's distance failed to reveal values that deviate appreciably from the fitted line, and no pattern was detected in the plot of Studentized residuals (Fig. 4D). Moreover, visual examination of data for birds weighing less than 0.1 kg revealed that the model is a satisfactory fit even over the lower part of the range in body size (Fig. 4E), thereby indicating that the data for large forms were not unduly influential in fitting the function. Thus, both graphical and statistical analyses indicate that this two-parameter power function is a good fit to the data.

Next we performed a standard allometric analysis on logarithmic transformations. Linear regression on transformed data for the 48 samples forming the full data set satisfied assumptions of normality of residuals and homoscedasticity despite inclusion of the two suspect values (Fig. 5A). No pattern is apparent in a plot of Studentized residuals, albeit the value for one of the suspect samples is unduly high (Fig. 5B) and a Studentized deleted residual of 3.66 marks it as an outlier. Nonetheless, data points for the log transforms themselves are distributed fairly symmetrically around the least squares line, and the two suspected outliers give the appearance of being no more deviant than other data in the sample (Fig. 5A). The fact that the putative outliers did not stand out in the plot of log transformations presumably accounts for the failure by Lasiewski and Dawson (1967) to detect them in the course of the original investigation. Regardless, the allometric exponent estimated from this function is 0.724, a value similar to that expected from the 3/4-power law.

Because we previously identified the two data points for a large corvid as outliers, we again removed these values from the data set and fitted a new regression to log-transformed values by the least squares method (Fig. 5C). This treatment yielded a statistically significant model (Fig. 5C) that satisfied assumptions concerning normality of residuals and constancy of variances, and no pattern is apparent in a plot of Studentized residuals (Fig. 5D). The allometric exponent estimated from this treatment is 0.68 (Fig. 5C), a value that is very similar to the one estimated by fitting a nonlinear function to untransformed data (Fig. 4C).

Finally, we back-transformed the linear model fit to log transformations (less the putative outliers) and displayed the line in a scatterplot together with the original data (Fig. 4F). The line compares quite favorably with the one fitted to untransformed values (Fig. 4C) because a two-parameter power function is an appropriate descriptor and a least squares equation provides a tight fit to the original data (cf. Fig. 3A, 3B).

In summary, the original analysis by Lasiewski and Dawson (1967) made use of logarithmic transformations that concealed the presence of influential outliers. By plotting data in their original arithmetic scale and performing preliminary statistical analyses, we were able to identify these outliers and to remove them from the data set. When the outliers were removed, anal-

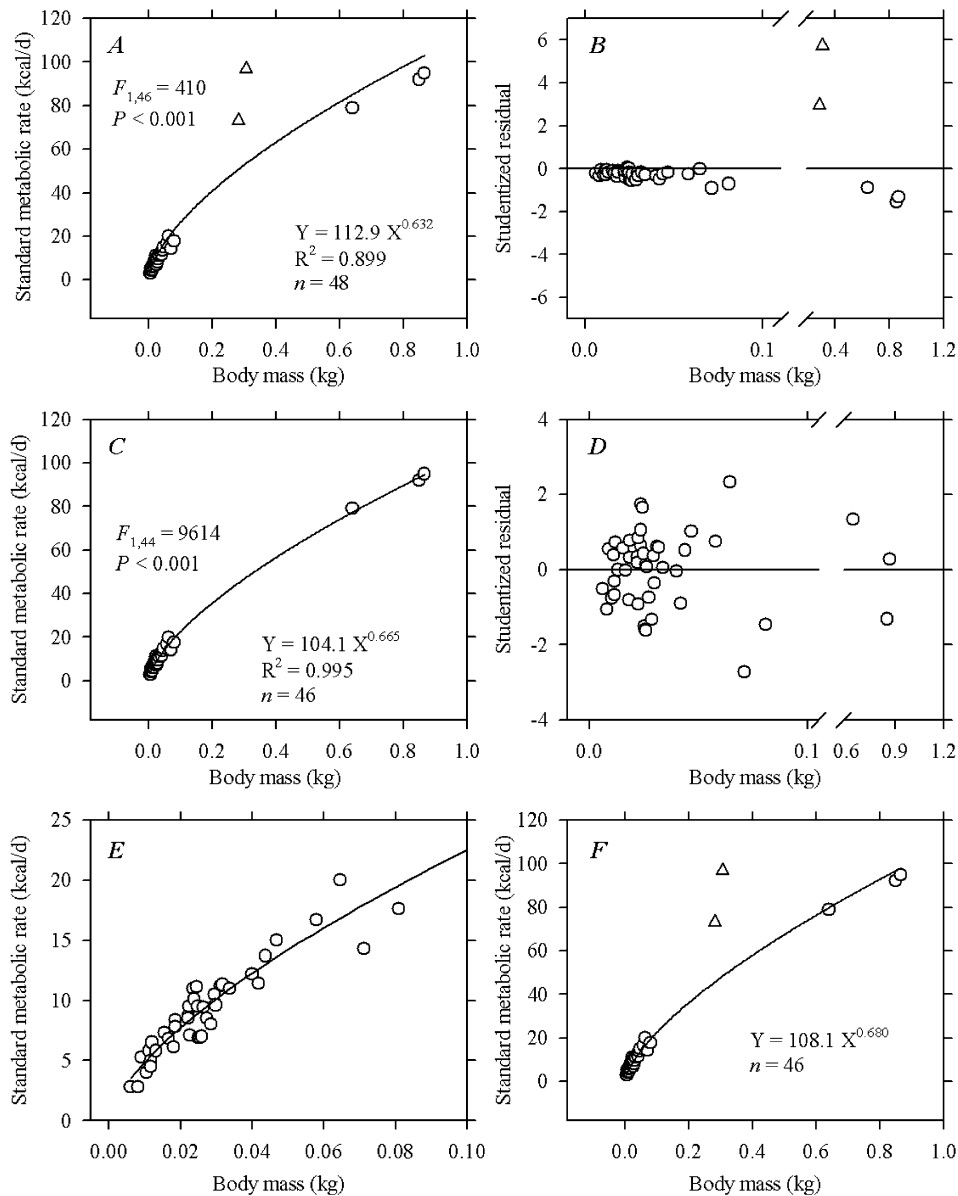


Figure 4. *A*, Means for standard metabolic rate and body mass of 48 samples of passerine birds (Lasiewski and Dawson 1967) in their original units. Triangles designate putative outliers. Statistics are from the regression ANOVA on the full data set. *B*, Studentized residuals from the analysis in *A* reveal that values for the putative outliers are suspiciously high. *C*, Means for metabolic rate and mass after outliers were removed from the data set. Statistics are from the regression ANOVA on the resulting sample of 46. *D*, Studentized residuals from the analysis in *C* are distributed randomly with respect to body size, and all of them are within acceptable limits. *E*, The scales for the abscissa and ordinate have been expanded to illustrate the goodness of the fit of the two-parameter model from *C* to data for birds weighing less than 0.1 kg. *F*, The equation from fitting a straight line to log-transformed values for 46 samples (see Fig. 5C) was back-transformed and the resulting function was plotted with data in the original scale. The putative outliers are shown for reference but were not used in the computations.

yses of untransformed and log-transformed values led to the same overall conclusion despite the fact that the study of logarithms was based on a statistical model with multiplicative error instead of one with the more appropriate additive error (see residuals in Fig. 4C for a distribution characteristic of data with additive error).

Fitting an Inferior Function to Data with Outliers: Basal Metabolic Rates of Artiodactyl Mammals

Data for metabolic rates and body mass of 20 species of artiodactyls were taken from the extensive table compiled by Savage et al. (2004). If a species was represented in this table by more

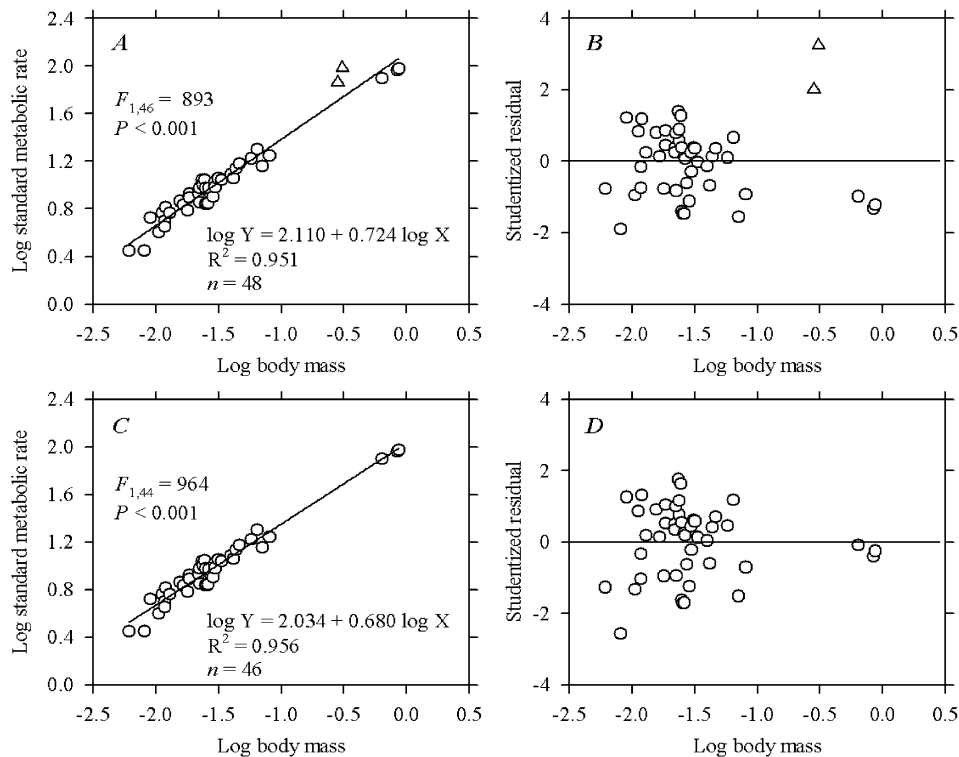


Figure 5. A, Means for standard metabolic rate and body mass for 48 samples of passerine birds (Lasiewski and Dawson 1967) were transformed to their common logarithms and displayed in a bivariate plot. Triangles designate putative outliers. Statistics are from the regression ANOVA on the full data set. B, Studentized residuals from the analysis in A reveal that both the putative outliers are displaced upward but that only one is so deviant as to be a major concern. C, Log-transformed values for samples exclusive of the putative outliers. Statistics are from the regression ANOVA on 46 samples. D, Studentized residuals from the analysis in C reveal no clear pattern of variation with respect to body size, and all of them are within acceptable limits.

than one set of values, we arbitrarily took the first entry to use in this example.

In accord with our standard practice, the data first were displayed in a bivariate plot with arithmetic coordinates (Fig. 6A). Casual inspection of the resulting graph leaves the distinct impression that most of the data would fall along a line conforming with a nonlinear function but that two values (for pigs *Sus scrofa* and camels *Camelus dromedarius*) would depart somewhat from that pattern (Fig. 5A, triangles). On the basis of this subjective impression, we fitted two- and three-parameter power functions to the data by the iterative least squares procedure. PRESS scores indicated that the three-parameter function was a somewhat better fit than the two-parameter function, so we focus on the former. However, both the models failed the test for homoscedasticity.

A plot of Studentized residuals from the three-parameter power function reveals a clearly discrepant value for camels (Fig. 6B, inverted triangle), indicating that the data point for these animals does not conform well with the model describing the rest of the data. This contention is supported by a Studentized deleted residual for camels of -4.41 . We know nothing about the circumstances under which data for this species were gathered and consequently have no specific reason for rejecting

the finding. On the other hand, the data point appears not to lie along the curve that would describe the bulk of the data, and examination of residuals indicates that something is amiss when this value is included in the analysis. We consequently treated the data point as an outlier, removed it from the data set, and repeated the analysis using the three-parameter power function.

The ANOVA from this second analysis pointed to a significant regression effect ($F_{2,16} = 188, P < 0.001, R^2 = 0.959$), and the model satisfied assumptions of normality of residuals and homoscedasticity. Consequently, we might have stopped at this point (Fig. 6C). However, the other odd value (for *S. scrofa*) still was a matter of some concern (Fig. 6C, triangle), because its Studentized deleted residual now was suspect at -5.67 . Although Cook's distance (0.42) indicates that the value for *S. scrofa* had little influence on the location of the best-fit line (because the value for body mass was near the middle of the distribution for that variable), we nonetheless deleted the data point for this species and again repeated the analysis using a three-parameter power function.

The third analysis also produced a statistically significant model ($F_{2,15} = 566, P < 0.001, R^2 = 0.987$) that satisfied assumptions regarding the normal distribution of residuals and

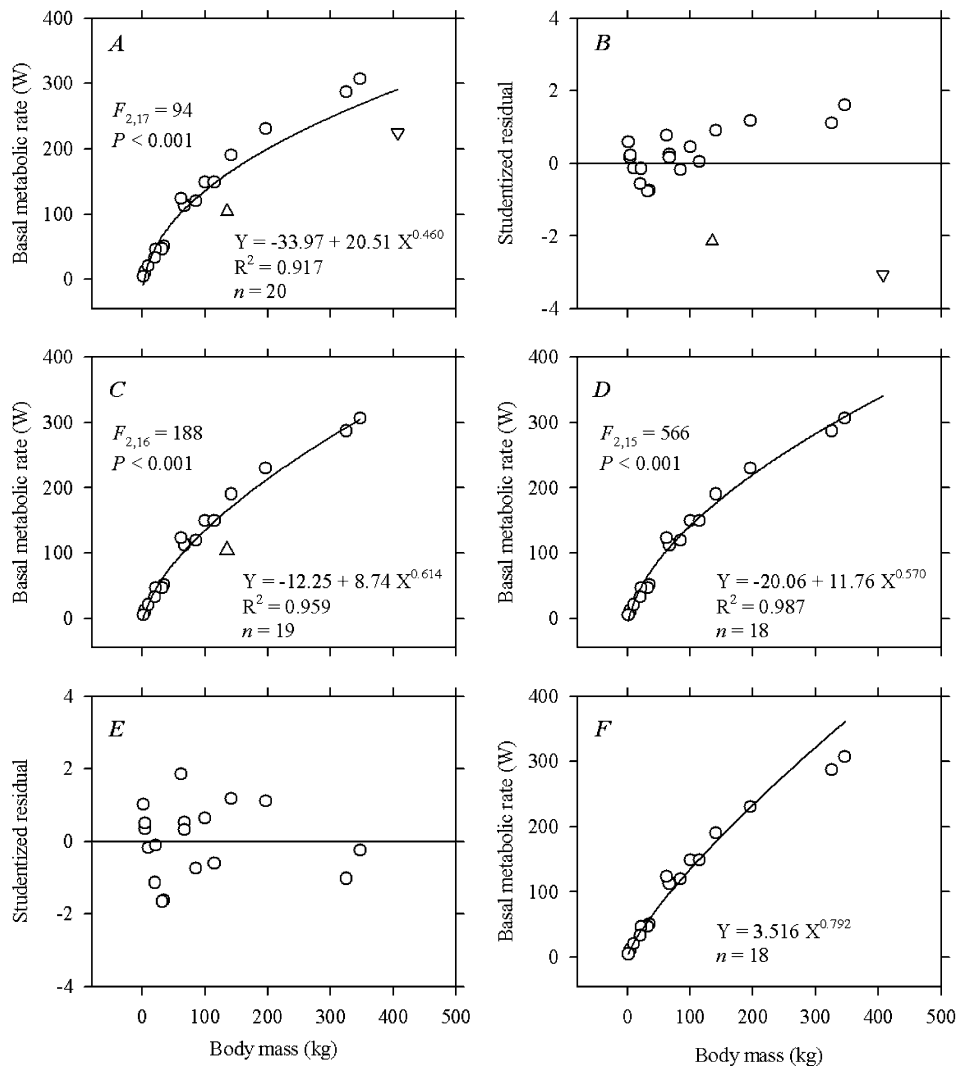


Figure 6. A, Means for basal metabolic rate and body mass for 20 species of artiodactyl mammals (Savage et al. 2004) in their original units. Triangle = pigs; inverted triangle = camels. Statistics are from the regression ANOVA on the full data set. B, Studentized residuals from the analysis in A reveal that values for the putative outliers (triangles) lie below those for other species; only the value for camels is low enough to be declared an outlier. C, Means for basal metabolic rate and body mass after removing data for camels. Statistics are from the regression ANOVA on the remaining 19 species, including pigs (triangle). D, Means and statistics after removing the remaining outlier. The regression ANOVA is based on $n = 18$. E, Studentized residuals from the analysis in D are within acceptable limits and reveal no clear pattern of variation with body size. F, The equation from fitting a straight line to log-transformed values for 18 samples (see Fig. 7C) was back-transformed and the resulting function was plotted with data in the original scale.

constancy of variances (Fig. 6D). However, no evidence of outliers was apparent in regression diagnostics or measures of influence for this third function, and the analysis yielded the added benefit of reducing the PRESS score to less than half that for the second analysis (2,712 vs. 7,697, respectively). In addition, no clear pattern is apparent in a plot of Studentized residuals (Fig. 6E), so we conclude that we have achieved a good fit to the data (Fig. 6D). The equation has the form

$$\hat{Y} = -20.062 + 11.762X^{0.570}.$$

The PRESS score for the three-parameter function (2,712) also

is substantially lower than that from a two-parameter power function fitted to the same values (3,435), so the three-parameter function is clearly the better descriptor of the data.

We followed the examination of data in their original scale with one based on logarithmic transformations. In accord with common practice (Smith 1980, 1984), we displayed the transformations in a bivariate plot and then fitted a straight line by the least squares method (Fig. 7A). The regression was statistically significant ($F_{1,18} = 594$, $P < 0.001$, $R^2 = 0.971$), and the analysis met the assumptions of normality of residuals and homoscedasticity despite inclusion of data for the two species suspected of being outliers (Fig. 7A, triangles). The plot of

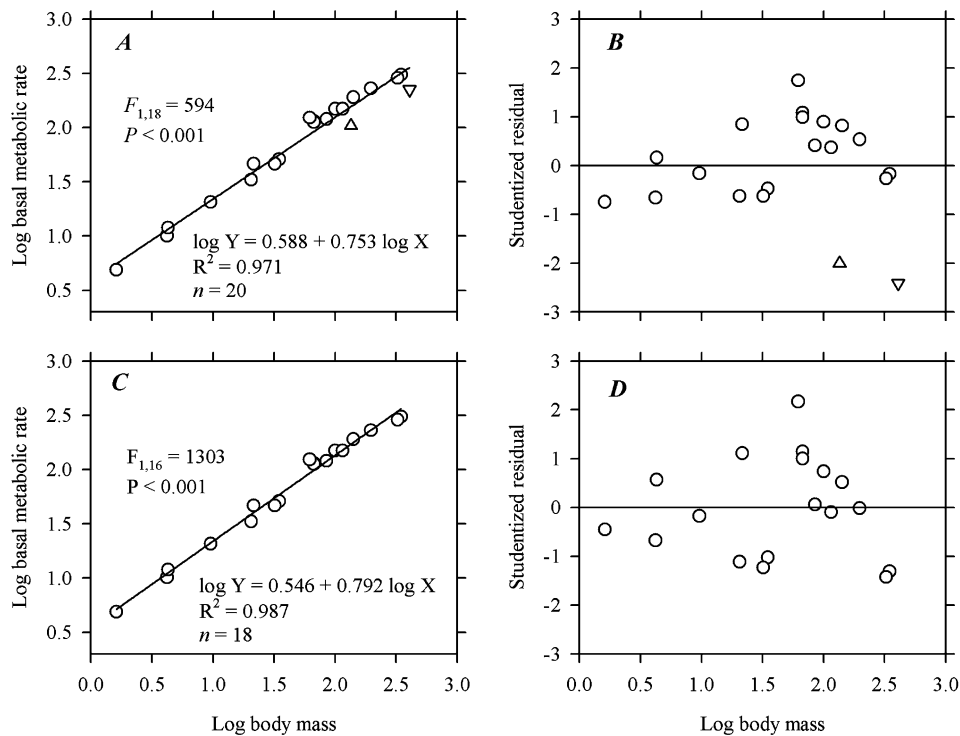


Figure 7. *A*, Means for basal metabolic rate and body mass of 20 species of artiodactyl mammals (Savage et al. 2004) were transformed to their common logarithms and displayed in a bivariate plot. Triangles designate putative outliers. Statistics are from the regression ANOVA on the full data set. *B*, Studentized residuals from the analysis in *A* reveal that values for both camels (*inverted triangle*) and pigs (*triangle*) are relatively low, but neither can be declared an outlier. *C*, Log-transformed values for samples exclusive of the putative outliers. Statistics are from the regression ANOVA on 18 samples. *D*, Studentized residuals from the analysis in *C* reveal a pattern of variation that usually is associated with heteroscedasticity, but the test for inequality of variances was not significant.

Studentized residuals reveals that the entries for camels and pigs are a bit low (Fig. 7*B*), but the Studentized deleted residuals (-2.85 and -2.21 , respectively) probably are not low enough to declare the values for these species to be outliers. Indeed, the putative outliers appear to lie no farther from the fitted line than other values in the sample (Fig. 7*A*). The slope of the least squares regression line fitted to data for the entire sample is 0.753.

Because we removed values for camels and pigs from the data set before running the final analyses on untransformed data, we also deleted log-transformed values for these species before running the final analysis on transformations (Fig. 7*C*). The regression yielded a significant fit ($F_{1,16} = 1,303$, $P < 0.001$, $R^2 = 0.988$) that met assumptions concerning the normal distribution of residuals and homoscedasticity. The plot of Studentized residuals has something of a megaphone pattern (Fig. 7*D*), which probably can be discounted owing to the fact that only one value falls slightly outside the range ± 2 . It is noteworthy that deletion of the putative outliers had the effect of increasing the estimate for the allometric exponent to 0.792 (Fig. 7*C*).

Finally, the equation relating $\log Y$ to $\log X$ for species exclusive of camels and pigs was back-transformed and then plotted with the original data (Fig. 6*F*). The line does not provide

a good visual fit because it seriously overestimates the basal metabolic rate of larger animals remaining in the sample. The relatively poor fit results from using a two-parameter power function with multiplicative error to characterize data that are described better by a three-parameter power function with additive error (Fig. 6*D*).

Logarithmic Transformations

The preceding case studies were chosen because they illustrate so clearly our points about the importance of fitting the most appropriate function and identifying and removing influential outliers. However, these studies serve also to highlight two additional points, both of which pertain more generally to the use of logarithmic transformations in allometric analyses.

First, logarithmic transformations were widely used early in the history of allometric research to linearize data and thereby facilitate graphical and statistical analysis (Peters 1983; Smith 1984, 1993). This use of transformations was warranted because methods for fitting nonlinear functions did not appear until after 1960 (Hartley 1961; Marquardt 1963; Glass 1967), and a general awareness of the problems introduced by log transforms did not emerge until even later (Glass 1969; Miller 1984; Jansson 1985; McCuen et al. 1990; Smith 1993; Pandey and Nguyen

1999; Hayes and Shonkwiler 2006). However, computer-based graphics now are widely available, and so, too, is sophisticated statistical software that can fit virtually any function to data in their original scale. Thus, linearization is no longer a sufficient rationale for making logarithmic transformations. Indeed, if transformations are not required, they should not be performed (Bartlett 1947; Finney 1989; Osborne 2002).

Second, logarithmic transformations commonly are used in contemporary allometric research to adjust (or “stabilize”) distributions when data in their original units fail to satisfy assumptions concerning normality of residuals and/or homogeneity of variances (Hoyle 1973). In the event that such transformations are needed (and assuming propriety of the underlying two-parameter power function), the ensuing regression analysis may point to some pattern of increase in values for Y with increasing values for X because of the identical ordering of values before and after transformation. However, little more can be said with any degree of confidence about the relationship between the two variables in their original state because back-transformation from logarithms usually will not yield reliable estimates for parameters in the allometric equation (Glass 1969; Miller 1984; Jansson 1985; McCuen et al. 1990; Pandey and Nguyen 1999). This outcome results in part from the differences in the statistical models that underlie analyses of data in arithmetic and logarithmic scales (Glass 1969; Miller 1984; Jansson 1985; McCuen et al. 1990; Smith 1993; Pandey and Nguyen 1999; Hayes and Shonkwiler 2006) and in part from the distortion that results from the transformation itself (Miller 1984; Jansson 1985; McCuen et al. 1990).

Overview

Applied statisticians usually advise their clients to begin analyses by plotting and examining data in the original scale (Anscombe 1973; Tukey 1977; Tufte 1983; Osborne 2002). With respect to allometry, such preliminary study enables investigators to gain an impression about the nature of the relationship between the variables (e.g., linear vs. nonlinear), to make tentative determinations about outliers, and even to distinguish between statistical models with additive or multiplicative error. Transformations should not be made before preliminary analyses have been completed, and transformations should then be performed only for cause (Bartlett 1947; Finney 1989; Osborne 2002). Most investigators are concerned with patterns of variation in the original scale, so this is the scale in which analyses should be performed, if it is at all possible (Finney 1989).

More than one model often can be fitted to a given data set, as illustrated by the example of varanid lizards, where three acceptable models were identified: a straight line with nonzero intercept, a two-parameter power function, and a three-parameter power function. Some readers may suggest, therefore, that the two-parameter power function should be embraced because of its simplicity, the apparent ease in interpreting the exponent, and historical precedent. Although this argument has some merit, it calls into question efforts to formulate general theory

that depends on a specific value for the allometric exponent, and it trivializes debate about the true value of that exponent (see Hoppeler and Weibel 2005). Moreover, current practice, arguably, is tantamount to fitting empirical data to the hypothesis that a two-parameter power function is an appropriate descriptor instead of using the data to test the hypothesis. We believe, therefore, that there is value to efforts to identify the “best” model and that this is most likely to be accomplished while working with data in their original scale.

We have focused on cases where analyses of data in their original scale and after logarithmic transformation led to different conclusions. However, we also encountered several data sets in which analyses of values before and after transformation led to similar estimates for the allometric exponent (despite differences in the error structure of the fitted models). Such outcomes occur when data actually conform with a two-parameter power function and the recorded values lay close to the fitted line (see Fig. 1A, 1B; Fig. 4A, 4B). The problem in this second instance is that the data had to be examined both ways in order to determine that the logarithmic transformation was appropriate. But if graphical and statistical analysis of data in their original scale yields a good model, there is no reason (or justification) then to perform an analysis of transformations (see Finney 1989).

Acknowledgments

We thank Martin Feder, Albert Bennett, Warren Burggren, and Raymond Huey for providing the initial impetus for this investigation. We also thank three anonymous reviewers for constructive criticisms that helped us to sharpen our arguments and avoid embarrassing mistakes. Finally, we are grateful to the authors of work that was the basis for our several case studies; these colleagues had the foresight to report their findings in sufficient detail that others could study the data at a later date using different procedures.

Literature Cited

- Agutter P.S. and D.N. Wheatley. 2004. Metabolic scaling: consensus or controversy? *Theor Biol Med Model* 1:13.
- Allen D.M. 1971. Mean square error of prediction as a criterion for selecting variables. *Technometrics* 13:469–475.
- Anscombe F.J. 1973. Graphs in statistical analysis. *Am Stat* 27: 17–21.
- Bartlett M.S. 1947. The use of transformations. *Biometrics* 3: 39–52.
- Brown J.H., J.F. Gillooly, A.P. Allen, V.M. Savage, and G.B. West. 2004. Toward a metabolic theory of ecology. *Ecology* 85: 1771–1789.
- Calder W.A., III. 1984. *Size, Function, and Life History*. Harvard University Press, Cambridge, MA.
- da Silva J.K.L., G.J.M. Garcia, and L.A. Barbosa. 2006. Allometric scaling laws of metabolism. *Phys Life Rev* 3:229–261.

- Finney D.J. 1989. Was this in your statistics textbook? V. Transformation of data. *Exp Agric* 25:165–175.
- Glass N.R. 1967. A technique for fitting nonlinear models to biological data. *Ecology* 48:1010–1013.
- . 1969. Discussion of calculation of power function with special reference to respiratory metabolism in fish. *J Fish Res Board Can* 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* 80:611–662.
- Gould S.J. 1966. Allometry and size in ontogeny and phylogeny. *Biol Rev* 41:587–640.
- Greer A.E. 1989. *The Biology and Evolution of Australian Lizards*. Surrey Beatty, Chipping Norton.
- Hartley H.O. 1961. The modified Gauss-Newton method for the fitting of non-linear regression functions by least squares. *Technometrics* 3:269–280.
- Hayes J.P. and J.S. Shonkwiler. 2001. Morphometric indicators of body condition: worthwhile or wishful thinking? Pp. 8–38 in J.R. Speakman, ed. *Body Composition Analysis of Animals: A Handbook of Non-Destructive Methods*. Cambridge University Press, Cambridge.
- . 2006. Allometry, antilog transformations, and the perils of prediction on the original scale. *Physiol Biochem Zool* 79:665–674.
- Hoppeler H. and E.R. Weibel. 2005. Scaling functions to body size: theories and facts. *J Exp Biol* 208:1573–1574.
- Hoyle M.H. 1973. Transformations: an introduction and a bibliography. *Int Stat Rev* 41:203–223.
- Jansson M. 1985. A comparison of detransformed logarithmic regressions and power function regressions. *Geogr Ann Ser A Phys Geogr* 67:61–70.
- Kleiber M. 1961. *The Fire of Life*. Wiley, New York.
- Kutner M.H., C.J. Nachtsheim, and J. Neter. 2004. *Applied Linear Regression Models*. 4th ed. McGraw-Hill Irwin, Boston.
- Lasiewski R.C. and W.R. Dawson. 1967. A re-examination of the relation between standard metabolic rate and body weight in birds. *Condor* 69:13–23.
- . 1969. Calculation and miscalculation of the equations relating avian standard metabolism to body weight. *Condor* 71:335–336.
- Manaster B.J. and S. Manaster. 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 Ind Appl Math* 11: 431–441.
- Martin R.D., M. Genoud, and C.K. Hemelrijk. 2005. Problems of allometric scaling analysis: examples from mammalian reproductive biology. *J Exp Biol* 208:1731–1747.
- McCuen R.H., R.B. Leahy, and P.A. Johnson. 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.
- Osborne J.W. 2002. Notes on the use of data transformations. *Pract Assess Res Eval* 8:6.
- Osborne J.W. and A. Overbay. 2004. The power of outliers (and why researchers should ALWAYS check for them). *Pract Assess Res Eval* 9:6.
- Packard G.C. and T.J. Boardman. 1988. The misuse of ratios, indices, and percentages in ecophysiological research. *Physiol Zool* 61:1–9.
- Pandey G.R. and V.-T.-V. Nguyen. 1999. A comparative study of regression based methods in regional flood frequency analysis. *J Hydrol (Amst)* 225:92–101.
- Peters R.H. 1983. *The Ecological Implications of Body Size*. Cambridge University Press, Cambridge.
- Pianka E.R. 1995. Evolution of body size: varanid lizards as a model system. *Am Nat* 146:398–414.
- Savage V.M., J.F. Gillooly, W.H. Woodruff, G.B. West, A.P. Allen, B.J. Enquist, and J.H. Brown. 2004. The predominance of quarter-power scaling in biology. *Funct Ecol* 18:257–282.
- Schmidt-Nielsen K. 1984. *Scaling: Why Is Animal Size So Important?* Cambridge University Press, Cambridge.
- Smith R.J. 1980. Rethinking allometry. *J Theor Biol* 87:97–111.
- . 1984. Allometric scaling in comparative biology: problems of concept and method. *Am J Physiol* 246:R152–R160.
- . 1993. Logarithmic transformation bias in allometry. *Am J Phys Anthropol* 90:215–228.
- Stevens J.P. 1984. Outliers and influential data points in regression analysis. *Psychol Bull* 95:334–344.
- Thompson G.G. and P.C. Withers. 1997. Comparative morphology of Western Australian varanid lizards (Squamata: Varanidae). *J Morphol* 233:127–152.
- Tufte E.R. 1983. *The Visual Display of Quantitative Information*. Graphics Press, Cheshire, CT.
- Tukey J.W. 1977. *Exploratory Data Analysis*. Addison-Wesley, Reading, MA.
- Warton D.I., I.J. Wright, D.S. Falster, and M. Westoby. 2006. Bivariate line-fitting methods for allometry. *Biol Rev* 81:259–291.
- West G.B. and J.H. Brown. 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.
- White C.R. and R.S. Seymour. 2005a. Allometric scaling of mammalian metabolism. *J Exp Biol* 208:1611–1619.
- . 2005b. Sample size and mass range effects on the allometric exponent of basal metabolic rate. *Comp Biochem Physiol A* 142:74–78.
- Zar J.H. 1968. Calculation and miscalculation of the allometric equation as a model in biological data. *BioScience* 18:1118–1120.
- . 1970. On the fitting of equations relating avian standard metabolism to body weight. *Condor* 72:247.