Available online at www.sciencedirect.com
ScienceDirect
Journal of Theoretical Biology 249 (2007) 168–177
Journal of Theoretical Biology
www.elsevier.com/locate/yjtbi

Uncertainty in allometric exponent estimation: A case study in scaling metabolic rate with body mass

Dafeng Hui*, Robert B. Jackson

Department of Biology and Nicholas School of the Environment and Earth Sciences, Duke University, Durham, NC 27708, USA

Received 21 June 2007; accepted 4 July 2007
Available online 18 July 2007

Abstract

Many factors could influence the allometric scaling exponent β estimation, but have not been explored systematically. We investigated the influences of three factors on the estimate of β based on a data set of 626 species of basal metabolic rate and mass in mammals. The influence of sampling error was tested by re-sampling with different sample sizes using a Monte Carlo method. Small random errors were introduced to measured data to examine their influence on parameter estimations. The influence of analysis method was also evaluated by applying nonlinear and linear regressions to the original data. Results showed that a relative large sample size was required to lower statistical inference errors. When sample size n was 10% of the base population size ($n = 63$), 35% of the samples supported $\beta = 2/3$, 39% supported $\beta = 3/4$, and 15% rejected $\beta = 0.711$, even though the base population had a $\beta = 0.711$. The controversy surrounding the estimation of β in the literature could be partially attributable to such small sample sizes in many studies. Measurement errors in body mass and base metabolic rate, especially in body mass, could largely increase alpha and beta errors. Analysis methods also affected parameter estimations. Nonlinear regressions provided better estimates of the scaling exponent that were significantly higher than these commonly estimated by linear regressions. This study demonstrated the importance of the quantity and quality of data as well as analysis method in power law analysis, raising caution in interpreting power law results. Meta-data synthesis using data from independent studies seems to be a proper approach in the future, but caution should be taken to make sure that such measurements are made using similar protocols.

Keywords: Body mass; Measurement error; Metabolic rate; Regression analysis; Sampling

1. Introduction

A power law appears widely in many different research disciplines such as physics, biology, and earth and planetary sciences (Newman, 2005). Its application has also caused considerable debate. Arguments over the allometric scaling exponent β in the power law $Y = Y_0 M^{\beta}$ have intensified during recent years, especially whether β equals 3/4, 2/3 or some other number for physiological traits (e.g., Dodds et al., 2001; White and Seymour, 2003, 2005b; Kozłwski and Konarzewski, 2004, 2005; Li et al., 2004; Brown et al., 2005; West and Brown, 2005; Glaizer, 2006; Reich et al., 2006). While the physical and biological basis for power law phenomena are still not entirely clear (Feldman, 1995; Agütter and Wheatley, 2004; Farrell-Gray and Gotelli, 2005; Hulbert and Else, 2004; Suarez and Darveau, 2005; van der Meer, 2006), the controversy in β may also be induced by the empirical estimation of the power law function. β was often estimated using algorithms based on a limited number of samples of animals or plants on which the independent variable (e.g. mass) and dependent variable (e.g. basal metabolic rate (BMR)) were measured. During these processes, at least three factors could affect the estimation of β: sample size, measurement error and analysis method. However, to our knowledge how these factors influence power law analysis has not been explored systematically.

Sampling size could influence the variability of parameter estimation and varies from a few to several hundred

*Corresponding author. Department of Biological Sciences, Tennessee State University, 3500 John A. Merritt Blvd., Nashville, TN 37209, USA. Tel.: +1 615 963 5681; fax: +1 615 963 5747.
E-mail address: dfhui@hotmail.com (D. Hui).

0022-5193/$ - see front matter © 2007 Elsevier Ltd. All rights reserved.
doi:10.1016/j.jtbi.2007.07.003

species in the literature (Smith, 1984; Savage et al., 2004). But very few studies have been conducted on how sampling size might influence the scaling exponent estimation. White and Seymour (2005a) reported that a minimum sample size of 50 drawn from a data set of 571 species was required to minimize the effect of skewed data introduced by influential small or large species, and 150 samples were necessary to distinguish $\beta = 2/3$ from 3/4. Measurement errors on both mass and basal metabolic rate could also influence parameter estimates. A study of mineralization showed that small errors in the data could cause serious errors in parameter estimations (Böttcher, 2004). To date, comparisons of analysis methods in power law analysis has mostly focused on linear regressions of log-transformed data (Beauchamp and Olson, 1973; Miller, 1984; Niklas, 2004). Reduced major axis (RMA) regression was suggested to fit linear regression, but in practice, the ordinary least squares (OLS) method has typically been used. By comparing linear and non-linear approaches, Kaitaniemi (2004) also demonstrated that much of the discrepancy in the literature may be an artifact caused by the methods. However, no study in practice has used nonlinear least squares method (NLS) and compared it with other methods.

We investigated the influences of sampling size, measurement error, and analysis method on estimated scaling exponents based on a data set of 626 pairs of species-averaged BMR and mass in mammals (Savage et al., 2004). Our objectives were to: (1) demonstrate the influence of sample size on alpha and beta errors in statistical inference and estimate the minimum requirement of sample size to lower these errors below 0.05; (2) examine the influence of errors in measured data on alpha and beta errors; and (3) evaluate the performance of nonlinear regression and its influence on estimation of the scaling exponent.

2. Materials and methods

2.1. Base data set

A recently compiled data set of species-averaged BMR and body mass of mammals by Savage et al. (2004) was used in this study. To control the quality of data set, Savage et al. (2004) checked the original reports, eliminated duplicated data, and averaged values for the same species. The complete data set includes $N = 626$ species. The properties of body mass and metabolic rate distributions have been carefully studied by Savage et al. (2004). This data set is heavily biased in favor of small species, due to the fact that small species are more abundant than larger species and easy to measure. We considered this data set as a base population and calculated the parameters $Y_0 = 0.021$ and $\beta = 0.711$ using the original data as in Savage et al. (2004). All data analyses in this study were based on, or derived from, this data set.

2.2. Influence of sample size

The influence of sampling error was tested by re-sampling with different sample sizes from the base population using a Monte Carlo approach. Sample size was selected as $n = 1/20$, $1/10$, $1/5$, $1/3$, $1/2$ of population size N. For each sample size, n observations were randomly selected from the population as one sample and this process was repeated 500 times. For each sample, parameters Y_0 and β were estimated using the OLS method on log-transformed data as in Savage et al. (2004). Quartiles of parameter estimate b of scaling exponent β from 500 samples under different sample sizes were displayed.

Similarly, the relationship of probability for accepting a certain null hypothesis $\beta = \beta_0$ (i.e., $\beta_0 = 3/4$, $2/3$ or 0.711) with sample size was constructed. The sample size n increased from 6, 7,…, to 500. For each sample size n, 500 samples were drawn from the base population. For each sample, the frequency and probability that the null hypothesis $\beta = 3/4$, $\beta = 2/3$, or $\beta = 0.711$ was accepted were calculated.

2.3. Influence of measurement error

The influence of measurement error on parameter estimations was analyzed by adding small random errors to the measured data. Virtual data sets were created using body mass in the base data set, with given parameters $Y_0 = 0.021$ and $\beta = 0.711$, assuming certain measurement errors in BMR and/or mass as described below.

There were three steps in this analysis. (1) *BMR calculation*. Given parameters Y_0 and β, BMR was calculated based on body mass of each species in the base data set; (2) *Measurement error determination*. Error distributions of BMR and body mass were assumed to follow normal distributions with mean $\mu = 0$ and standard deviations σ equal to small percentages of the BMR and observed mass values; given the error distributions of BMR and mass, random errors were drawn from the normal distributions for each observation and added to the calculated BMR and measured mass. (3) *Parameter estimation*. For each virtual data set, we estimated parameters using the OLS method as in Savage et al. (2004).

The effects of measurement error on the parameter estimation were examined using different combinations of measurement error settings (i.e., different σ values for BMR and mass). The probability for accepting $\beta = 3/4$, $2/3$, or 0.711 was estimated by generating 500 samples for each combination of measurement errors of BMR and mass.

Probability was also estimated for accepting the null hypothesis $\beta = 3/4$, $2/3$, or 0.711, given a population with $\beta = 3/4$ or $2/3$ instead of $\beta = 0.711$, when the standard deviations of BMR and mass were set as 20% and 40% of their respective values.

2.4. Influence of analysis method

The power law is a simple, nonlinear regression equation. At least two approaches can be used to estimate its parameters. One is nonlinear regression analysis, minimizing an objective function by an iterative process. Certain optimization algorithms, such as Gauss–Newton method, can be applied for this iteration. Another is linear regression analysis on the transformed data. Taking the logarithms of both sides of the equation, we obtain a linear regression equation, $Y' = Y_0' + bM'$, where $Y' = \log(Y)$, $M' = \log(M)$, and $Y_0' = \log(Y_0)$. In this case b is the estimate of β. Thus, linear regression can be applied to obtain the parameter estimates $Y_0 = \exp(Y_0')$ and b.

The influence of analysis methods was evaluated by re-analyzing the base data set using nonlinear and linear regression methods. NLS methods, both non-weighted (NLS) and weighted nonlinear (WNL), were applied with the Gauss–Newton optimization algorithm. To determine which weight was the best for WNL, we used Furnival's Index of Fit (Furnival, 1961). The weight was set as the reciprocal of body mass raised to various powers: $1/M^{1/2}$, $1/M$, $1/M^{3/2}$ and $1/M^2$. Since $1/M^{1/2}$ provided the smallest FI (results not shown), we selected $1/M^{1/2}$ as weight in this study. For the linear regression, OLS, major axis regression (MAR), and RMA methods were conducted. The methods were evaluated quantitatively by examining the bias and mean square error. Bias and the root mean square error (RMSE) were calculated as

$$\text{bias} = \frac{\sum_i (Y_i - \hat{Y}_i)}{n}, \quad \text{bias\%} = \frac{\sum_i (Y_i - \hat{Y}_i)/n}{\sum_i \hat{Y}_i/n},$$

$$\text{RMSE} = \sqrt{\frac{\sum_i (Y_i - \hat{Y}_i)^2}{n-2}},$$

$$\text{RMSE\%} = \sqrt{\frac{\sum_i (Y_i - \hat{Y}_i)^2/(n-2)}{\sum_i \hat{Y}_i/n}}.$$

All analyses were conducted using SAS (Hui and Jiang, 1996; SAS Institute Inc. Cary, NC, USA). SAS IML was used as a program tool and for linear regression analysis (OLS). NLIN procedure was used for nonlinear regression analysis. NLP procedure was used for MAR and RMA with appropriate minimization criterion. We also used TableCurve to explore the influence of different methods of both linear and nonlinear regression analyses (TableCurve 2D V5.01, SYSTAT software Inc., 2002).

3. Results

3.1. Influence of sample size

The range of the parameter estimate b decreased as sample size n increased (Table 1). When the sample size was small, e.g., $n = 31$ (i.e., 1/20 of the population size $N = 626$), b ranged from 0.578 to 0.795 with a median of 0.708. When n was increased to 1/2 of the population size ($n = 313$), the range of b was much smaller and it ranged between 0.685 and 0.733 with a median of 0.711.

The probability of accepting a certain null hypothesis $\beta = \beta_0$ was also strongly influenced by sample size (Fig. 1). When $n = 6$, the probability to accept the true null hypothesis $\beta = 0.711$ was 0.83, and the alpha error (i.e., the probability of rejecting a true null hypothesis) was therefore 0.17. As n increased, the probability to accept $\beta = 0.711$ increased linearly. When n was larger than 380 (i.e., more than 61% of the population size), the alpha error became smaller than 0.05.

Beta errors (i.e., the probabilities of accepting the false null hypothesis $\beta = 3/4$ or $2/3$) were very high when n was as low as 6, but decreased sharply as n increased (Fig. 1). When $n = 63$, beta errors for $\beta = 2/3$ and $3/4$ were 0.35 and 0.39, respectively (Table 2). When sample sizes were larger than 170 and 185 (i.e., 27% and 30% of population size), the beta errors for $\beta = 2/3$ and $\beta = 3/4$ became less than 0.05.

3.2. Influence of measurement error

The probability of accepting the true null hypothesis $\beta = 0.711$ decreased (i.e., alpha error increased) as the measurement errors of BMR and mass increased (Fig. 2a). When the measurement error of BMR was small, the probability decreased sharply as the measurement error of mass increased. As the measurement error of BMR increased, the probability decreased slowly at first, but more quickly as it approached zero. When measurement errors of both BMR and mass were high, the probability tended towards zero, and alpha error became very high.

The probability of accepting $\beta = 3/4$ (i.e., beta error) (Fig. 2b) showed a similar pattern as the beta error to

Table 1
Distribution of parameter estimation b of 500 samples drawn from a population with $\beta = 0.711$ at different sample sizes

Quantile	Sample size				
	20	31	63	125	313
100% Max	0.814	0.795	0.776	0.753	0.733
95%	0.771	0.763	0.746	0.737	0.724
75%	0.737	0.731	0.725	0.722	0.718
50% Median	0.709	0.708	0.710	0.711	0.711
25%	0.679	0.681	0.695	0.700	0.706
5%	0.624	0.641	0.667	0.683	0.697
0% Min	0.559	0.578	0.582	0.663	0.685

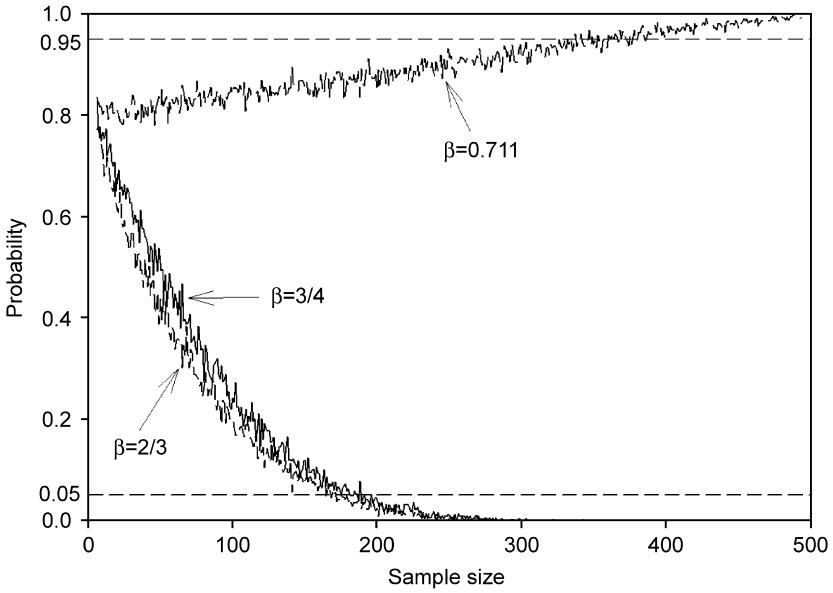

Fig. 1. Influence of sample size on the probability of accepting the null hypothesis $\beta = 0.711$, $\beta = 2/3$, and $\beta = 3/4$. For $\beta = 0.711$, 1-probability is the β error (i.e., reject a true null hypothesis $\beta = 0.711$). For $\beta = 2/3$ or $\beta = 3/4$, probability is the α error (i.e., accept a false null hypothesis $\beta = 2/3$ or $\beta = 3/4$).

Table 2
Frequency and probability of accepting the null hypothesis $\beta = 2/3$, 0.711 or 3/4 of 500 samples drawn from a population with $\beta = 0.711$ at sample size $n = 63$

	Acceptance of null hypothesis		
	$\beta = 2/3$	$\beta = 0.711$	$\beta = 3/4$
Frequency	177	427	195
Probability	0.35	0.85	0.39

accept $\beta = 2/3$ (Fig. 2c). The measurement error of body mass showed a larger influence on the probability than the measurement error of BMR. Interestingly, beta errors also decreased when measurement error of mass was very high.

When the measurement errors were 20% and 40% of their mean values of body mass and BMR, respectively, the alpha error of rejecting the true null hypothesis $\beta = 2/3$ was 0.14 (i.e., 1−0.86), and the beta error of accepting the false null hypothesis $\beta = 0.711$ was 0.17 (Table 3). When the samples were drawn from a population with $\beta = 3/4$, the alpha error of rejecting the correct null hypothesis $\beta = 3/4$ was 0.15 (i.e., 1−0.85), the beta error of accepting $\beta = 2/3$ was 0.08, and the beta error of accepting $\beta = 0.711$ reached 0.49. Both alpha and beta errors were quite large.

3.3. Influence of analysis method

Parameter estimates for b by nonlinear regression were significantly higher than those by linear regression methods (Table 4). Among the five methods applied here, WNL had the smallest bias, % bias, RMSE, and % RMSE. Only MAR marginally accepted the null hypothesis $\beta = 3/4$. All other methods showed that neither $\beta = 2/3$ nor $\beta = 3/4$ should be accepted.

Estimated BMR of regression curves were compared to the measured values. Estimated BMR by linear regressions were corrected using correction factor CF = exp(MSE/2) for bias (Miller, 1984). Compared to measured values, curves back transformed from linear regressions obviously underestimated BMR when body mass was large (Fig. 3). Even when body mass was small, nonlinear regression curves fit the data better than linear regression methods did.

Our analysis was focused on the parameter estimates in the power law function. We also fit different two-parameter functions using TableCurve. Of 57 implemented in TableCurve, 36 functions fit the data well with R^2 larger than 0.50. Among these functions, the best functions included power, logarithmic, and linear functions, with R^2 larger than 0.98.

4. Discussion

Three interesting findings from this study demonstrated the uncertainties in allometric scaling exponent estimation and raised some cautions in interpreting power law results. One was that quite large sample size was required (up to 61% of population size) in order to obtain a reliable estimate of the allometric scaling exponent for the population. When the sample size was small, as in most studies, it was quite possible to erroneously accept a false null hypothesis (i.e., $\beta = 3/4$ or $\beta = 2/3$). Secondly, nonlinear regression methods generally had the least bias and model error but produced estimates of the scaling exponent that were significantly higher than 2/3 or 3/4 and higher

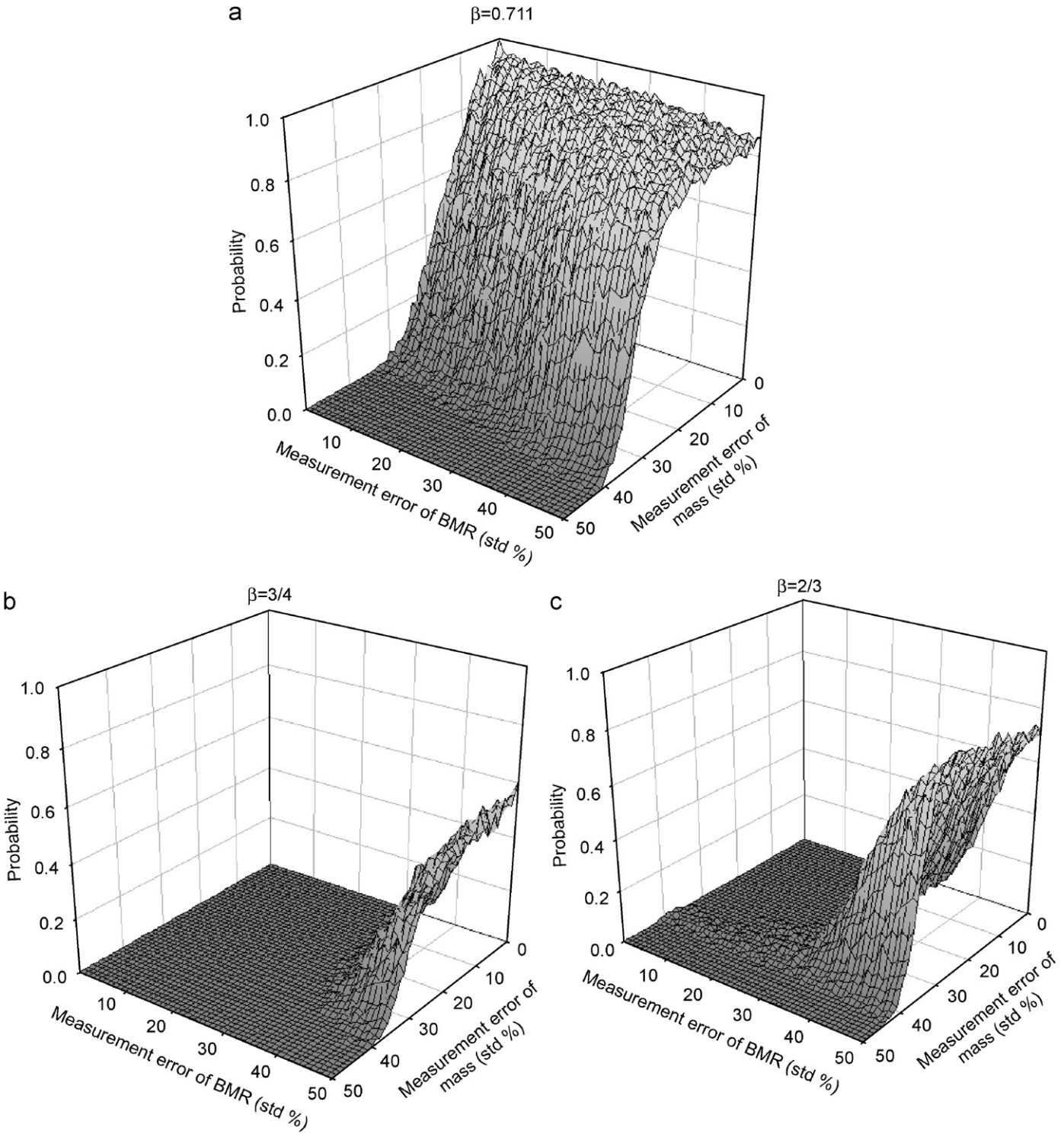

Fig. 2. Influence of measurement errors on the probability of accepting the null hypothesis $\beta = 0.711$ (a), $\beta = 3/4$ (b) and $\beta = 2/3$ (c) when 500 samples were drawn from a population with $\beta = 0.711$. For $\beta = 0.711$, 1-probability is the β error (i.e., reject a true null hypothesis $\beta = 0.711$). For $\beta = 2/3$ or $\beta = 3/4$, probability is the α error (i.e., accept a false null hypothesis $\beta = 2/3$ or $\beta = 3/4$).

than estimates from linear regression methods. Thirdly, measurement errors of both metabolic rate and body mass further increased alpha and beta errors; this fact poses additional challenges for interpreting power law results and highlights the importance of quality control of measurements. Based on these results, it is difficult to conclude with strong support that a universal scaling exponent β of either 3/4 or 2/3 exists. Considering the large uncertainty in power law analysis, future research efforts that focus on theoretical development would complement ongoing

Table 3
Probability (and frequency) of accepting the null hypothesis $\beta = 2/3$, 3/4, and 0.711 from 500 samples drawn from a population with parameter $\beta = 2/3$ or $\beta = 3/4$

Population parameter	Acceptance of null hypothesis		
	$\beta = 2/3$	$\beta = 3/4$	$\beta = 0.711$
$\beta = 2/3$	0.86 (429)	0.01 (4)	0.17 (83)
$\beta = 3/4$	0.08 (41)	0.85 (424)	0.49 (244)

Measurement errors (i.e., standard deviations) were 20% and 40% of the mass and basal metabolic rates, respectively.

efforts that just validate or disprove a scaling exponent β using empirical methods.

4.1. Influence of sample size

A surprising finding in this study was that more than 61% of the total base population size needed to be sampled in order to reduce both alpha error and beta error below 0.05, based on the largest data set of mammalian metabolic rate and body mass compiled so far. Since most power law analysis to date has focused on whether a specific $\beta = \beta_0$

Table 4
Comparison of analysis methods

Method	Objective function (Q)	Calculation of b	Estimated b	95% confidence interval	Bias	Bias %	RMSE	RMSE %
Direct methods using original data (nonlinear regression)								
NLS	$\sum_i (Y_i - Y_0 M_i^b)^2$	1. Nonlinear least squares method, iteration	0.865	0.856–0.874	0.74	0.07	9.66	2.95
WNL	$\sum_i w_i (Y_i - Y_0 M_i^b)^2$	2. Weighted least squares method, iteration $w_i = 1/M_i^{1/2}$.	0.842	0.832–0.851	0.20	0.02	9.91	2.95
Indirect methods using transformed data (linear regression)								
OLS	$\sum_i (Y_i' - [Y_0' + bM_i'])^2$	$b = \frac{s_{Y'M'}}{s_{Y'}^2}$	0.711	0.699–0.724	2.54	0.28	47.1	15.8
MAR	$\sum_i \frac{(Y_i' - [Y_0' + bM_i'])^2}{1+b^2}$	$b = \frac{1}{2s_{Y'M'}}\left\{s_{Y'}^2 - s_{M'}^2 + \left[(s_{Y'}^2 - s_{M'}^2)^2 + 4s_{Y'M'}^2\right]^{1/2}\right\}$	0.724	0.696–0.752	2.19	0.24	43.0	14.1
RMA	$\sum_i \frac{(Y_i' - [Y_0' + bM_i'])^2}{b}$	$b = \begin{cases} \frac{s_{Y'}}{s_{M'}} & \text{if } s_{Y'M'} > 0 \\ -\frac{s_{Y'}}{s_{M'}} & \text{if } s_{Y'M'} < 0 \end{cases}$	0.729	0.710–0.748	1.15	0.11	36.1	11.2

$Y' = \ln(Y)$, $M' = \ln(M)$, and $Y_0' = \ln(Y_0)$; NLS, nonlinear least squares method; WNL, weighted nonlinear least squares method; OLS, ordinary least squares method; MAR, major axis regression; RMA, reduced major axis regression. RMSE, root mean square error.

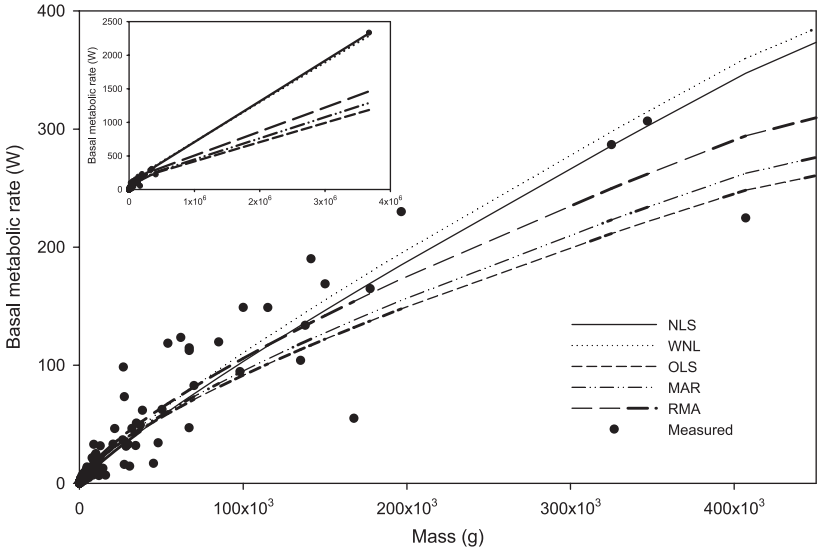

Fig. 3. Comparisons of modeled curves and observed basal metabolic rates. The inset shows all data points. Due to the large scale on the x axis, some of these data points are overlapped when x is small. NLS, nonlinear least squares method; WNL, weighted nonlinear least squares method; OLS, ordinary least squares method; MAR, major axis regression; RMA, reduced major axis regression.

should be accepted, beta error seems more problematic and should be minimized. To lower the beta error of accepting a false hypothesis of $\beta = 2/3$ or $\beta = 3/4$ below 0.05, at least 27% and 30% of the base population should be sampled. Otherwise, both $\beta = 2/3$ and $\beta = 3/4$ could be accepted (when sampled from a population of $\beta = 0.711$). These results may help the scientific community understand why there are so many different allometric exponents estimated, and controversial findings reported in the literature, as most researches typically focuses on a small number of species (Dodds et al., 2001). Smith (1984) compiled a table of 60 power-law functions, and showed that the sample size ranged primarily from 7 to 69, with only two studies with sample sizes that were ~100. With such small sampling sizes, the chance of a beta error is very large. Some recent synthesis studies have used relatively more data compiled from the literature, but still more data are needed.

Our conclusion on sample size requirements is different from that of White and Seymour (2005a) who used a similar approach but focused on the minimum sample size to distinguish between the exponents of 2/3 and 3/4. While both studies showed that the standard deviation of the estimated exponent decreased as sample size increased, we suggest that this evidence (and the change of model exponent with sample size) could not be used to determine the minimum sample size. Theoretically, if the scaling exponent is an unbiased estimate of the population exponent, then the mean estimated exponent should be equal to the population exponent, even when the sample size is small. The standard deviation of the estimate (i.e., standard error) is a function of sample size n in a form of $\sigma/\sqrt{n}$ as shown in Fig. 2a in White and Seymour (2005a) and cannot be used to determine a minimum sample size. Our analysis of relationships of alpha and beta errors with sample size provides direct and strong evidence of the minimum sample size required.

If our conclusion that 30% of species was required to lower the beta error of accepting a false $\beta = 2/3$ or 3/4 below 0.05 applies, at least 1389 mammal species should be sampled, considering that more than 4629 mammal species have been recognized in the world (Wilson and Reeder, 1993). Meta-data syntheses that integrate data from multiple individual studies to increase sample size seem to be a correct approach in empirical allometric scaling analysis (White and Seymour, 2003; Savage et al., 2004).

4.2. Influence of measurement error

Measurement errors are ubiquitous in biological research due to heterogeneity in space and time, growth conditions, precision of measurement equipment, and human error. Such measurement errors could increase the difficulty in accepting a true null hypothesis or rejecting a false null hypothesis. As demonstrated here, measurement errors caused serious alpha and beta errors. Measurement error of body mass apparently had more influence on the parameter estimations than that of BMR; therefore more attention should be given to obtaining accurate measurements of body mass. Interestingly, when both measurement errors were high, there was a huge alpha error, but beta errors decreased, probably due to the increased variations in parameter estimation that decreased power in data analysis. Therefore, the utmost care should be taken when the experiments are carried out and all experimental steps should be examined for possible improvements of measurement accuracy (Böttcher, 2004). When data from multiple studies are compiled, researchers should make sure that the measurements in different studies were made using similar protocols.

Another source of measurement errors could be the systematic errors in the measurement of BMR. As pointed out in White and Seymour (2003), the conditions for BMR measurement are quite strict and often not well met, the measurements of BMR would tend to be overestimated in many cases. One statistical method to separate this systematic error could be to add an intercept in the power law function. With this intercept added, the estimate of allometric scaling exponent could be more deviated from 2/3 or 3/4.

4.3. Influence of analysis method

Both linear and nonlinear regressions have been extensively used in ecological studies (e.g., Luo et al., 2001; Jackson et al., 2002; Hui and Jackson, 2006). It is somewhat surprising, therefore, that nonlinear regression has not been applied in power law analysis, because power law is a typical nonlinear regression equation. When the power law was first proposed, it was difficult to use NLS methods, since iterative procedures are needed to solve the equations but are difficult to perform without computers. Today, however, a power law equation can be fit in seconds using computer programs such as SAS. Although methods could influence the parameter estimations, as demonstrated by Kaitaniemi (2004) and in this study, to our knowledge, all calculations and discussions on analysis methods in power law analysis so far have focused on linear regression of the transformed data or on a few non-parametric methods (Beauchamp and Olson, 1973; Dodds et al., 2001; Niklas, 2004; Martin et al., 2005).

The reasons given for using log-transformed data, as summarized by Niklas and Enquist (2002), are to reduce the problem of working with outliers, and to comply with the statistical assumptions of normality and homoscedasticity. For the log-transformed metabolic rate and mass data, OLS, MAR and RMA regression can be used. RMA is recommended in allometric scaling analysis, especially when the variables of interest are biologically interdependent and are subjected to unknown measurement error, and when functional rather than predictive relationships are sought (Dodds et al., 2001; Niklas and Enquist, 2002). Dodds et al., (2001) demonstrated that parameter estimates of b varied slightly with linear regression methods, similar to what we showed here. Niklas (2004) also realized that

the different linear regression methods can profoundly influence the numerical values of scaling components, but the differences became small when the correlation of log-transformed metabolic rate and body mass was higher (Niklas and Enquist, 2002).

Measurement error on body mass and BMR also influence the performance of the methods (McArdle et al., 1990; Farrell-Gray and Gotelli, 2005). McArdle et al. (1990) concludes that OLS regression should be used in preference to RMA regression with the measurement error in the independent variable is less than one-third of that in the dependent variable. It does not matter which method to choose when the correlation coefficient is large (e.g., $r > 0.90$). Farrell-Gray and Gotelli (2005) concludes that the inappropriate use of OLS regression biases the estimates of b away from 3/4. However, their simulation analysis was based on an evenly spaced sample of body sizes from 1 to 4000 kg, and this introduces an additional bias: a more appropriate body mass distribution produces a similar pattern, but less extreme results.

Log-transformation, however, introduces a systematic bias into calculations because the largest values are compressed on the logarithmic scale, and a correction factor is needed to reduce this bias (Beauchamp and Olson, 1973; Miller, 1984; Niklas, 2004). Even adjusted for bias, the linear regression methods still could not fit our data well when the independent variable was large (Fig. 3; also see Miller, 1984). These differences were reflected on the parameter estimate b. As we showed here, nonlinear regressions provided much higher estimates of the scaling components, which were significantly larger than both 2/3 and 3/4.

The large differences found between direct nonlinear methods and indirect regression methods were due to the differences in objective functions (see Table 4). Linear regressions minimized $\sum_i w_i(Y'_i - [Y'_0 + bM'_i])^2$ ($w_i = 1$ for OLS, $w_i = 1/(1+b^2)$ for MAR, and $w_i = 1/b$ for RMA), but NLS methods minimized $\sum_i w_i(Y_i - [Y_0 + M_i^b])^2$ ($w_i = 1$ for NLS and $w_i = 1/\mathrm{M}^{1/2}$ for WNL). It is quite clear here that only NLS and WNL worked directly to minimize the difference of estimation BMR (i.e., $Y_0 M_i^b$) and observed BMR (i.e., Y_i). Due to this reason, NLS tends to fit the large data well. Possible heteroscedasticity could be reduced by WNL regression analysis. Since WNL provided the least bias and RMSE, WNLS method should be considered in power law analysis.

Pursuing a simple universal law is the dream work of many researchers, and human never stop searching such a law. But we may just start to understand the basic rules that appear to govern much of biology's seemingly dazzling complexity (Bokma, 2004). Besides the factors we showed here, there are several other factors could influence allometric scaling analysis. (1) Martin et al. (2005) raised the influence of phylogeny problem. Individual taxa in the comparison may not be statistically independent because of phylogenetic relationships within the tree to which they belong. Limited divergence in body size between closely related species is one of the prime examples of phylogenetic inertia. Several phylogenetically based statistical methods have been proposed include independent contrasts, generalized least-squares models, and Monte Carlo simulations (White and Seymour, 2003; Muñoz-Garcia and Williams, 2005; Garland et al., 2005). Although this problem could be partially offset by restricting analysis to the generic level, this would reduce the sample size and have other drawbacks (Martin et al., 2005). We expect that when phylogeny information will be incorporated into data analysis, the parameter estimation would be improved in the future; (2) extreme and biased data points could influence the estimation of allometric scaling component (Smith, 1984). In the data set we used here, there are several extreme large data points. If we remove one or a few large data points, the estimate of allometric scaling component decreased remarkably; and (3) other statistical functions could fit the data equally well as power law function. Using the TableCurve, we found many linear and nonlinear statistical models can be used to fit this data set. There is no way to distinguish which model is better than another for the several best regression models. More theoretical studies on the mechanisms of the allometric scaling should be focused in the future (e.g., Demetrius, 2006). Identifying and understanding broad scale convergence in functional relationships are more fascinating, and certainly very challenging.

5. Conclusions

After a systematic analysis of sampling size, measurement error, and analysis method, we provide evidence that these three factors have large influences on allometric scaling exponents, requiring a thorough investigation. Our goal in this analysis was not to prove that whether the scaling exponent β equals 2/3 or 3/4, but to raise cautions in interpreting power law results and to provide helpful suggestions for future research. A nonlinear regression method provided a better fit to the observed data here and should be considered in power law analysis. For variables that are measured with errors, it is important to obtain accurate estimates with repeated measurements on similar individuals grown under similar conditions. Increased uniformity in measurement protocol would also facilitate meta-data analyses across studies to increase overall sample size (Meinzer, 2003).

Acknowledgements

We thank Drs. Bai-Lian Li and Yiqi Luo for their constructive comments on an earlier draft of this manuscript. We also thank Drs. Denise E. Kirschner and Craig White and an anonymous reviewer for their comments and suggestion that made this manuscript much improved. This work was supported by grants from NIGEC-DOE (through the office of Biological and Environmental Research at the Department of Energy), the National

Science Foundation (EAR0223340, DEB0444518), the Inter American Institute for Global Change Research, and the Andrew W. Mellon Foundation.

Appendix A. Nonlinear regression

Nonlinear regression model describes the relationship between a dependent (or response) variable and one or a set of independent (or predictor) variables. Compared to linear regression model, nonlinear regression model is more widely used in social sciences and natural sciences such as biology and ecology. Unlike the linear regression analysis, there is generally no analytical way to solve nonlinear regression models. Usually numerical optimization algorithms are required to determine the best-fitting parameters. Among many different methods, one often used is to minimize the sum of squared deviations (residuals), as in linear regression analysis. This is the OLS approach. In cases where there are different error variances for different errors, a sum of weighted squared residuals may be minimized, so called the WLS approach.

To solve NLS problems, one method frequently used is the Gauss-Newton algorithm. This algorithm has been implemented in many statistical software, such as SAS and SPSS. For the simple nonlinear regression models such as power law function, there is no difference in the model parameter estimates among different algorithms. Here, we briefly describe the procedure using Gauss–Newton algorithm which was used in this analysis (Glantz and Slinker, 2001; Nerlove, 2005).

The basic form for a nonlinear regression model between response y and a predictor x is given as

$$y_i = f(x_i, \theta) + e_i, \tag{A.1}$$

where y_i, x_i are the data, f is a nonlinear function involving the predictor and the parameter vector θ, and e_i is a random error. The sum of squared residuals is $S(\theta) = \sum(y_i - f(x_i, \theta))^2$ that is to be minimized. Differentiating $S(\theta)$, we obtain

$$\frac{\partial S(\theta)}{\partial \theta} = -2\sum\left[y_i - f(x_i, \theta)\right]\frac{\partial f(x_i, \theta)}{\partial \theta}. \tag{A.2}$$

Setting the partial derivatives to 0 produces estimating equations (i.e., normal equations) for the regression coefficients. Because these equations are in general nonlinear, they require solution by numerical optimization. The Gauss-Newton algorithm is an iterative procedure. An initial guess for the parameter vector θ should be provide, which we will call θ_0. This method uses the Taylor series

$$f(x_i, \theta) = f(x_i, \theta_0) + J_f(\theta_0)(\theta - \theta_0) + \cdots, \tag{A.3}$$

where $J_f(\theta_0)$ denotes the Jacobian of f at $\theta = \theta_0$.

Subsequent estimates of θ_k for the parameter vector are then produced by the recurrence relation

$$\theta_{k+1} = \theta_k - (J_f(\theta_k)^T J_f(\theta_k))^{-1} J_f(\theta_k)^T e(\theta_k). \tag{A.4}$$

To find the best estimates of θ, Eq. (A.4) is iterated until convergence is achieved.

References

Agütter, P.S., Wheatley, D.N., 2004. Metabolic scaling: consensus or controversy? Theor. Biol. Med. Modell. 1, 13.

Beauchamp, J.J., Olson, J.S., 1973. Corrections for bias in regression estimates after logarithmic transformation. Ecology 54, 1403–1407.

Bokma, F., 2004. Evidence against universal metabolic allometry. Funct. Ecol. 18, 184–187.

Böttcher, J., 2004. Uncertainties of nonlinearly estimated parameters from incubations of soil organic matter. J. Plant Nutr. Soil Sci. 167, 293–302.

Brown, J.H., West, G.B., Enquist, B.J., 2005. Yes, West, Brown and Enquist's model of allometric scaling is both mathematically correct and biologically relevant. Funct. Ecol. 19, 735–738.

Demetrius, L., 2006. The origin of allometric scaling laws in biology. J. Theor. Biol. 243, 455–467.

Dodds, P.S., Rothman, D.H., Weitz, J.S., 2001. Re-examination of the "3/4-law" of metabolism. J. Theor. Biol. 209, 9–27.

Farrell-Gray, C.C., Gotelli, N.J., 2005. Allometric exponents support a 3/4—power scaling law. Ecology 86, 2083–2087.

Feldman, H.A., 1995. On the allometric mass exponent, when it exists. J. Theor. Biol. 172, 187–197.

Furnival, G.M., 1961. An index for comparing equations used in constructing volume tables. Forest Sci. 7, 337–341.

Garland Jr., T., Bennett, A.F., Rezende, E.L., 2005. Phylogenetic approaches in comparative physiology. J. Exp. Biol. 208, 3015–3035.

Glaizer, D.S., 2006. The 3/4-power law is not universal: evolution of isometric, ontogenetic metabolic scaling in pelagic animals. Bioscience 56, 325–332.

Glantz, S.A., Slinker, B.K., 2001. Primer of Applied Regression and Analysis of Variance, second ed. McGraw-Hill Publishing Company.

Hui, D., Jackson, R.B., 2006. Geographic and interannual variability in biomass partitioning in grassland ecosystems: a synthesis of field data. New Phytol. 169, 85–93.

Hui, D., Jiang, C., 1996. Practical Statistical Analysis System (SAS) Usage. Beijing University of Aeronautics and Astronautics Press, Beijing, China.

Hulbert, A.J., Else, P.L., 2004. Basal metabolic rate: history, composition, regulation, and usefulness. Physiol. Biochem. Zool. 77, 869–876.

Jackson, R.B., Banner, J.L., Jobbágy, E.G., Pockman, W.T., Wall, DH., 2002. Ecosystem carbon loss with woody plant invasion of grasslands. Nature 418, 623–626.

Kaitaniemi, P., 2004. Testing the allometric scaling laws. J. Theor. Biol. 228, 149–153.

Kozłwski, J., Konarzewski, M., 2005. West, Brown and Enquist's model of allometric scaling again: the same questions remain. Funct. Ecol. 19, 739–743.

Li, B.-B., Gorshkov, V.G., Makarieva, A.M., 2004. Energy partitioning between different-sized organisms and ecosystem stability. Ecology 85, 1811–1813.

Luo, Y., Wan, S., Hui, D., Wallace, L., 2001. Acclimatization of soil respiration to warming in a tall grass prairie. Nature 413, 622–625.

Martin, RD., Genoud, M., Hemelrijk, C.K., 2005. Problems of allometric scaling analysis: examples from mammalian reproductive biology. J. Exp. Biol. 208, 1731–1747.

McArdle, B.H., Gaston, K.J., Lawton, J.H., 1990. Variation in the size of animal populations: patterns, problems and artefacts. J. Animal Ecol. 59, 439–454.

Meinzer, F.C., 2003. Functional convergence in plant responses to the environment. Oecologia 134, 1–11.

Miller, D.M., 1984. Reducing transformation bias in curve fitting. Am. Stat. 38, 124–126.

Muñoz-Garcia, A.M., Williams, J.B., 2005. Basal metabolic rate in carnivores is associated with diet after controlling for phylogeny. Physiol. Biochem. Zool. 78, 1039–1056.

Nerlove, M., 2005. On the numerical accuracy of Mathematica 5.0 for doing linear and nonlinear regression. Math. J. 9, 824–851.

Newman, M.E.J., 2005. Power laws, Pareto distributions and Zipf's law. Contemporary Phys. 46, 323–351.

Niklas, K.J., 2004. Plant allometry: is there a grand unifying theory? Biol. Rev. 79, 871–889.

Niklas, K.J., Enquist, B.J., 2002. Canonical rules for plant organ biomass partitioning and annual allocation. Am. J. Bot. 89, 812–819.

Reich, P., Tjoelker, M.G., Machado, J.L., Oleksyn, J., 2006. Universal scaling of respiratory metabolism, size and nitrogen in plants. Nature 439, 457–461.

Savage, V.M., Gillooly, J.F., Woodruff, W.H., West, G.B., Allen, A.P., Enquist, B.J., Brown, J.H., 2004. The predominance of quarter-power scaling in ecology. Funct. Ecol. 18, 257–282.

Smith, R.J., 1984. Allometric scaling in comparative biology: problems of concepts and method. Am. J. Physiol. 246, R152–R160.

Suarez, R.K., Darveau, C.A., 2005. Multi-level regulation and metabolic scaling. J. Exp. Biol. 208, 1627–1634.

van der Meer, J., 2006. Metabolic theories in ecology. Trends Ecol. Evol. 21, 136–140.

West, G.B., 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.

White, C.R., Seymour, R.S., 2003. Mammalian basal metabolic rate is proportional to body mass$^{2/3}$. Proc. Natl. Acad. Sci. USA 100, 4046–4049.

White, C.R., Seymour, R.S., 2005a. Sampling size and mass range effects on the allometric exponent of basal metabolic rate. Comp. Biochem. Physiol. 142A, 74–78.

White, C.R., Seymour, R.S., 2005b. Allometric scaling of mammalian metabolism. J. Exp. Biol. 208, 1611–1619.

Wilson, D.E., Reeder, D.M., 1993. Mammal Species of the World. Smithsonian Institution Press.