Comparative Biochemistry and Physiology, Part A 166 (2013) 70–73

Contents lists available at SciVerse ScienceDirect

Comparative Biochemistry and Physiology, Part A

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

Fitting statistical models in bivariate allometry: Scaling metabolic rate to body mass in mustelid carnivores

Gary C. Packard *

Department of Biology, Colorado State University, Fort Collins, CO 80523-1878, USA

ARTICLE INFO

Article history:
Received 12 March 2013
Received in revised form 8 May 2013
Accepted 13 May 2013
Available online 18 May 2013

Keywords:
Akaike's Information Criterion
Allometry
Logarithms
Nonlinear regression
Scaling
Transformation

ABSTRACT

The ongoing debate about methods for fitting the two-parameter allometric equation $y = \mathbf{a}x^{\mathbf{b}}$ to bivariate data seemed to be resolved recently when three groups of investigators independently reported that statistical models fitted by the traditional allometric method (i.e., by back-transforming a linear model fitted to log–log transformations) typically are superior to models fitted by standard nonlinear regression. However, the narrow focus for the statistical analyses in these investigations compromised the most important of the ensuing conclusions. All the investigations focused on two-parameter power functions and excluded from consideration other simple functions that might better describe pattern in the data; and all relied on Akaike's Information Criterion instead of graphical validation to identify the better statistical model. My re-analysis of data from one of the studies (BMR vs. body mass in mustelid carnivores) revealed (1) that the best descriptor for pattern in the dataset is a straight line and not a two-parameter power function; (2) that a model with additive, normal, heteroscedastic error is superior to one with multiplicative, lognormal, heteroscedastic error; and (3) that Akaike's Information Criterion is not a generally reliable metric for discriminating between models fitted to different distributions. These findings have apparent implications for interpreting the outcomes of all three of the aforementioned studies. Future investigations of allometric variation should adopt a more holistic approach to analysis and not be wedded to the traditional allometric method.

© 2013 Elsevier Inc. All rights reserved.

1. Introduction

Three recent investigations of allometric variation in both plants and animals doubtless caught the attention of many biologists because the studies seemed, at first blush, to resolve an ongoing debate about methods for fitting simple power functions to bivariate data (Mascaro et al., 2011; Xiao et al., 2011; Ballantyne, 2013). The exponent in the allometric equation may be affected profoundly by the method used to fit the statistical model (e.g., Kerkhoff and Enquist, 2009; Packard, 2009), so the debate and its apparent resolution are matters of considerable importance to all those who are involved with describing and interpreting the scaling of biological variables with body size.

At the heart of the three studies was a comparison of two-parameter power functions fitted by standard nonlinear regression and by the traditional allometric method. A model fitted directly to untransformed data by standard nonlinear regression is assumed to have additive, normal, homoscedastic error (Kutner et al., 2004), and has the form

$$y_i = \mathbf{a}x_i^{\mathbf{b}} + \varepsilon_i \quad \varepsilon_i \sim \mathrm{N}\left(0,\sigma^2\right). \tag{1}$$

Line-fitting by traditional allometry, however, requires that a straight line first be fitted to logarithmic transformations of the data and that the resulting homoscedastic model,

$$\log(y_i) = \log(\mathbf{a}) + \mathbf{b}\ \log(x_i) + \varepsilon_i \quad \varepsilon_i \sim \mathrm{N}\left(0,\sigma^2\right), \tag{2}$$

then be back-transformed to form a new model with multiplicative, log-normal, heteroscedastic error on the original scale,

$$y_i = \mathbf{a}x_i^{\mathbf{b}} * \exp(\varepsilon_i). \tag{3}$$

The traditional method emerged early in the 20th century as a means for fitting a simple power function to observations following a curvilinear path in bivariate display (see Packard, 2012a), but emphasis has been placed more recently on the heteroscedastic error in the final model (e.g., Kerkhoff and Enquist, 2009; White et al., 2012). The traditional allometric method currently is in widespread use despite numerous expressions of concern over the years regarding reliability of the procedure (e.g., Thompson, 1943; Zar, 1968; Glass, 1969; Smith, 1984; Verwust, 1991; Packard, 2011a,b).

Akaike's Information Criterion (AIC), which was used in the three studies to assess goodness of fit of the alternative models, usually was lower for models fitted by traditional allometry than for those fitted by standard nonlinear regression. This finding led the authors to conclude

---

* Tel.: +1 970 566 3478.
E-mail address: Gary.Packard@ColoState.edu.

1095-6433/$ – see front matter © 2013 Elsevier Inc. All rights reserved.
http://dx.doi.org/10.1016/j.cbpa.2013.05.013

that models fitted by traditional allometry are generally superior to those fitted by standard nonlinear regression and, therefore, that the former typically provide good descriptions for pattern in the data. Other workers engaged in research on scaling were thereby encouraged to use the traditional approach to line-fitting because of its apparent reliability and simplicity. However, these putative validations of the traditional method are misleading, because none of the studies went far enough either to clarify the issues fully or to demonstrate the general utility of the traditional method. Unresolved issues concern (1) the adequacy of fitted models for describing patterns of variation in bivariate data and (2) how best to model random variation (residuals) in the response variable.

2. Methods

For purposes of illustration, I took data for basal metabolic rate (BMR) and body mass of 13 species of mustelid carnivores from Table 1 in Muñoz-Garcia and Williams (2005) and submitted them to new analyses (Fig. 1A). Although the dataset is small (and consequently less than ideal for use in a case-study like this), it is the same dataset that was examined by Ballantyne (2013) in the most recent of the aforementioned inquiries. Inasmuch as all three of the investigations followed the same basic protocol, my re-analysis provides new perspective on analytical procedures used in all the studies.

I took a more holistic approach than was followed by Ballantyne (and by Mascaro et al. and Xiao et al.) in that a variety of models with different forms for error were fitted to the data. First, a two-parameter power function with multiplicative, lognormal, heteroscedastic error [3] was “fitted” to arithmetic data by back-transforming the model for a straight line fitted to logarithmic transformations [2]. Next, I fitted models for straight lines (with and without intercepts) and for two- and three-parameter power functions directly to untransformed data by the method of least squares. The several least-squares models were of two types: one set of models assumed additive, normal, homoscedastic error (as in linear and nonlinear regression in standard statistical software) whereas the other set assumed additive, normal, heteroscedastic error. The functional form for variance in least-squares models with heteroscedastic error was

$$\mathrm{var}(\varepsilon_i) = \sigma^2 * (f(x_i, \beta))^{2\theta} \tag{4}$$

where β was a vector of parameters in a function, *f*, relating BMR (y) to mass (x) and θ was an additional parameter in the fitted model (Pinheiro and Bates, 2000, p. 210 ff; Ritz and Streibig, 2008, p. 74 ff; Zuur et al., 2009, p. 78 ff). Thus, models with normal, heteroscedastic error were of the general form

$$y_i = f(x_i, \beta) + \varepsilon_i \quad \varepsilon_i \sim \mathrm{N}\left(0, \left(\sigma^2 * (f(x_i, \beta))^{2\theta}\right)\right) \tag{5}$$

with variance modeled as a power of the mean for y. This model reduces to the standard linear or nonlinear (homoscedastic) model [1] when θ equals zero, because functional form for variance then becomes σ^2. Mascaro et al. (2011) also fitted models with additive, normal, heteroscedastic error, but they considered only two-parameter power functions. All the analyses were run using the nlrwr package in R 2.13.2 (Ritz and Streibig, 2008).

Table 1
Comparison of statistical models fitted to untransformed data for metabolic rate vs. body mass for 13 species of mustelid carnivores.

Predictive equation	AIC	ΔAIC
Linear model w/normal, homoscedastic error, no intercept: $\hat{y} = 0.279\,x$	213.3	36.1
Linear model w/normal, homoscedastic error, intercept: $\hat{y} = -235.4 + 0.293\,x$	214.4	37.2
2-Parameter power w/normal, homoscedastic error: $\hat{y} = 0.00253\,x^{1.470}$	203.9	26.7
3-Parameter power w/normal, homoscedastic error: $\hat{y} = 224.2 + 0.000769\,x^{1.584}$	204.4	27.2
Back-transformed OLS w/multiplicative, log-normal, heteroscedastic error: $\hat{y} = 2.292\,x^{0.736}$	183.0	5.8
Linear model w/normal, heteroscedastic error, no intercept: $\hat{y} = 0.255\,x$	190.2	13.0
Linear model w/normal, heteroscedastic error, intercept: $\hat{y} = 75.2 + 0.218\,x$	177.2	0
2-Parameter power w/normal, heteroscedastic error: $\hat{y} = 4.932\,x^{0.628}$	186.0	8.8
3-Parameter power w/normal, heteroscedastic error: $\hat{y} = 73.9 + 0.239\,x^{0.989}$	179.2	2.0

ΔAICs are based on the fit of a full linear model with normal, heteroscedastic error. AICs reported by Ballantyne (2013) are in substantial agreement with those reported here despite his making an adjustment for sample size. Note the reductions in AIC when non-zero intercepts are incorporated into models with normal, heteroscedastic error.

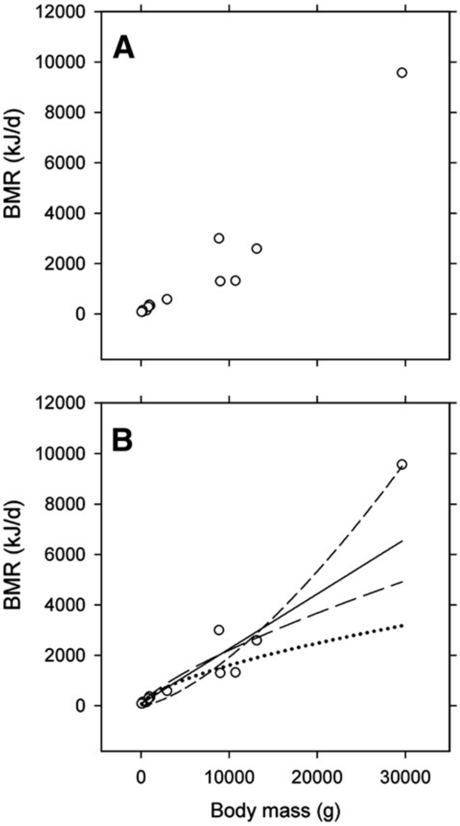

Fig. 1. (A) Data for basal metabolic rate (BMR) and body mass of 13 species of mustelid carnivores are displayed in a simple bivariate plot to show the overall distribution for the observations without introducing the visual distraction of several regression lines. Scatter in the observations appears to increase with body size, thereby pointing to heteroscedasticity in the response variable. (B) The solid line is the mean function for the straight line with non-zero intercept and additive, normal, heteroscedastic error. The short-dashed line with upward curvature is the mean function for a 2-parameter power equation fitted by standard nonlinear regression (i.e., additive, normal, homoscedastic error). The dotted line with downward curvature is the 2-parameter function with additive, normal, heteroscedastic error. The long-dashed line is the function estimated by the traditional allometric method (i.e., multiplicative, lognormal, heteroscedastic error).

I also used AIC to compare the different candidate models (Burnham and Anderson, 2002). AIC for the model estimated by the traditional allometric method was computed using R-code written by Jack Weiss at University of North Carolina (http://www.unc.edu/courses/2010spring/ecol/562/001/docs/lectures/lecture6.htm#lecture6). AICs were not adjusted for sample size because of doubts about the value of such corrections (Richards, 2005). The lowest AIC identifies the candidate model that conveys the most information about the relationship between predictor and response, and provides a reference for assessing the goodness

Download English Version:

https://daneshyari.com/en/article/10818880

Download Persian Version:

https://daneshyari.com/article/10818880

Daneshyari.com