



New perspectives for estimating body condition from mass/length data: the scaled mass index as an alternative method

Jordi Peig and Andy J. Green

J. Peig (*jpeig@ub.edu*), Dept of Animal Biology, Univ. of Barcelona, Avinguda Diagonal, ES-08028 Barcelona, Spain. – A. J. Green, Dept of Wetland Ecology, Estación Biológica de Doñana, C/ Américo Vespucio s/n, ES-41092 Sevilla, Spain.

Body condition is assumed to influence an animal's health and fitness. Various non-destructive methods based on body mass and a measure of body length have been used as condition indices (CIs), but the dominant method amongst ecologists is currently the calculation of residuals from an ordinary least squares (OLS) regression of body mass against length. Recent studies of energy reserves in small mammals and starlings claimed to validate this method, although we argue that they did not include the most appropriate tests since they compared the CI with the absolute size of energy reserves. We present a novel CI (the 'scaled mass index') based on the central principle of scaling, with important methodological, biological and conceptual advantages. Through a reanalysis of data from small mammals, starlings and snakes, we show that the scaled mass index is a better indicator of the relative size of energy reserves and other body components than OLS residuals, performing better in all seven species and in 19 out of 20 analyses. We also present an empirical and theoretical comparison of the scaled mass index and OLS residuals as CIs. We argue that the scaled mass index is a useful new tool for ecologists.

Many studies of animal ecology rely on non-destructive methods to estimate the body condition of different individuals in a population, relying on measures of body mass and linear measures of body size to calculate a condition index (CI) (Green 2001, Stevenson and Woods 2006). Many authors fail to define what they mean by 'body condition', but others define it as a measure of the energetic (or nutritional state) of an individual animal, especially the relative size of energy reserves such as fat and protein (Krebs and Singleton 1993, Gosler 1996, Schulte-Hostedde et al. 2001). CIs are morphological, biochemical or physiological metrics indicative of the health or quality of individual animals and assumed to be related to fitness. We are concerned from hereon only with CIs based on mass data and morphometric measures (see Stevenson and Woods 2006 for review of other nondestructive CIs). For the purposes of this paper, we define an animal's body condition as the energy capital accumulated in the body as a result of feeding, which we assume to be an indicator of an animal's health and quality. It is vital for ecologists to seek a CI capable of separating effects of structural size of the body (*sensu* Green 2001) from the size of the energy capital, since both aspects can be expected to have major consequences for fitness, survival rates, habitat use, etc. It is also essential to bear in mind that the correct interpretation of CIs depends on the species and ecological framework under study, and that CIs will not reflect energy reserves in all studies. For example, the size of water stores may be the most important determinant of CI variation and even of

fitness in some desert reptiles (Nagy et al. 2002). Likewise, extreme mass for a given size may indicate disease in some circumstances (e.g. human obesity).

A popular CI amongst ecologists is currently the residual from an ordinary least squares (OLS) regression of body mass (Y variable) against a linear morphometric measure assumed to represent size (X variable) (Jakob et al. 1996, Hayes and Shonkwiler 2001). Various reasons why this method may violate key assumptions and produce misleading results were previously pointed out by Green (2001). Others (Smith 1992, García-Berthou 2001, Freckleton 2002, Hayes and Shonkwiler 2006) have pointed out the dangers in calculating residuals and then relating them to ecological predictor variables, whether discontinuous (e.g. sex, age, habitat type) or continuous (e.g. fecundity, strength of secondary sexual characters), in further analyses. Nevertheless, later studies have defended the use of OLS residuals by demonstrating significant correlations between them and the absolute size of fat stores (Ardia 2005, Schulte-Hostedde et al. 2005). Recent authors now rely on these latter studies to justify their choice of OLS residuals as a CI, without contrasting them with alternative indices or validating them using their own data on energy reserves (Husak 2006, O'Dwyer et al. 2006, Toigo et al. 2006, Liker et al. 2008).

Schulte-Hostedde et al. (2005) and Ardia (2005) found that OLS residuals of mass on a linear measure of size (L from hereon) usually showed higher correlations with the absolute size of fat stores (in g) than residuals from an

alternative regression technique (standardized major axis [SMA] regression). OLS residuals were also highly correlated with absolute values of lean dry mass (largely protein), and both results have been interpreted as confirming their value as CIs (Schulte-Hostedde et al. 2005). However, we question this reasoning because the absolute mass of fat or any other body component is not the most appropriate measure of ‘true condition’ since it does not control for body size (a vital requirement for a CI). The absolute size of fat stores or protein tends to covary with body size (Table 1). Other things being equal (i.e. under isometry in which shape and body composition does not change with size), larger animals will have more absolute fat. OLS residuals are systematically biased towards larger individuals, partly

because OLS methods underestimate the true slope between mass and length (Seim and Sæther 1983, Green 2001, Arnold and Green 2007). It is therefore no surprise that OLS residuals correlate well with absolute fat and/or protein. OLS residuals are also sometimes validated based on correlations with components as percentages of body mass (Krebs and Singleton 1993, Weatherhead and Brown 1996), although this approach has been criticized because it assumes an isometric relationship between different body components (Kotiaho 1999). In order to validate a CI, it should be compared with the relative amount of energy stores while taking into account the scaling relationships between overall body size, the size of body components and linear measures of size.

Table 1. Correlations between body composition (absolute or relative [%] amount) with body mass (M) and linear body measurements (L) for each species (significant differences at $p < 0.1$ are in bold).

	M				L			
	r	p-value	r	p-value	r	p-value	r	p-value
Chipmunks			body mass		body length		skull length	
Fat mass	0.440	0.041	0.265	0.233	0.226	0.312	0.014	0.950
Lean dry mass	0.868	<0.001	0.387	0.076	0.365	0.095	0.442	0.039
Water mass	0.882	<0.001	0.406	0.060	0.075	0.741	0.644	0.001
% fat	0.183	0.416	0.174	0.438	0.192	0.391	-0.114	0.612
% lean dry tissue	-0.393	0.071	-0.083	0.713	0.109	0.629	-0.176	0.434
% water	-0.141	0.531	0.025	0.911	-0.402	0.063	0.354	0.106
Deer mice			body mass		body length		foot length	
Fat mass	0.238	0.017	-0.025	0.807	0.025	0.806	-0.167	0.097
Lean dry mass	0.563	<0.001	0.280	0.005	0.435	<0.001	0.310	0.002
Water mass	0.898	<0.001	0.210	0.036	0.445	<0.001	0.121	0.232
% fat	-0.007	0.944	-0.123	0.223	-0.115	0.254	-0.232	0.020
% lean dry tissue	-0.088	0.386	0.135	0.179	0.134	0.185	0.174	0.083
% water	-0.020	0.841	-0.148	0.143	-0.067	0.508	-0.268	0.007
Meadow voles			body mass		body length		foot length	
Fat mass	0.385	0.024	0.243	0.166	-0.091	0.608	0.335	0.053
Lean dry mass	0.933	<0.001	0.760	<0.001	-0.214	0.224	0.317	0.067
Water mass	0.989	<0.001	0.751	<0.001	-0.255	0.146	0.544	0.001
% fat	-0.108	0.543	-0.134	0.449	0.073	0.681	0.184	0.297
% lean dry tissue	-0.221	0.209	-0.051	0.777	0.103	0.562	-0.487	0.003
% water	0.489	0.003	0.349	0.043	-0.262	0.134	0.685	<0.001
Red-backed voles			body mass		body length		foot length	
Fat mass	0.191	0.078	-0.001	0.991	-0.098	0.372	0.031	0.774
Lean dry mass	0.891	<0.001	0.645	<0.001	0.180	0.098	0.628	<0.001
Water mass	0.984	<0.001	0.680	<0.001	0.101	0.354	0.611	<0.001
% fat	-0.192	0.077	-0.284	0.008	-0.122	0.264	-0.196	0.070
% lean dry tissue	-0.147	0.176	-0.050	0.648	0.126	0.248	0.128	0.240
% water	0.244	0.024	0.120	0.269	-0.077	0.483	0.275	0.010
Wood rats			body mass		body length		skull length	
Fat mass	0.358	0.004	0.251	0.049	0.168	0.191	0.024	0.854
Lean dry mass	0.965	<0.001	0.810	<0.001	0.812	<0.001	0.618	<0.001
Water mass	0.992	<0.001	0.819	<0.001	0.781	<0.001	0.605	<0.001
% fat	-0.005	0.972	-0.029	0.822	-0.130	0.312	-0.228	0.075
% lean dry tissue	0.162	0.208	0.197	0.124	0.313	0.013	0.264	0.038
% water	-0.089	0.492	-0.025	0.846	-0.107	0.407	0.075	0.561
Starlings			body mass		tarsus length		wing length	
Fat mass	0.827	<0.001	0.189	0.061	0.449	<0.001	0.474	<0.001
% fat	0.654	<0.001	0.096	0.344	0.372	<0.001	0.357	<0.001
Water snakes			body mass		snout-vent length		tail length	
Fat mass	0.881	<0.001	0.695	<0.001	0.414	0.029	0.704	<0.001
Protein mass	0.988	<0.001	0.905	<0.001	0.520	0.005	0.913	<0.001
Water mass	0.998	<0.001	0.924	<0.001	0.499	0.007	0.926	<0.001
Ash mass	0.984	<0.001	0.952	<0.001	0.492	0.008	0.949	<0.001
% fat	0.426	0.024	0.259	0.184	-0.037	0.851	0.222	0.256
% protein	-0.112	0.571	-0.068	0.730	0.154	0.433	-0.028	0.887
% water	0.246	0.208	0.322	0.095	0.256	0.188	0.340	0.077
% ash	-0.073	0.712	0.143	0.468	0.221	0.258	0.173	0.378

The two principal aims of this paper are 1) to present a new non-destructive CI (the ‘scaled mass index’) based on mass and morphometric measurements, and 2) to show that this index is a more reliable indicator of true condition, using data on the size of energy stores previously presented by Schulte-Hostedde et al. (2005), Ardia (2005) and Weatherhead and Brown (1996) and comparing the scaled mass index with the performance of OLS residuals. We also explain the methodological and conceptual advantages of the scaled mass index over the OLS residual index (see also Green 2001 for a list of key CI assumptions that are often violated by OLS residuals).

Material and methods

Theoretical framework for condition indices

To be successful, a CI should be developed from a scaling perspective. The problem of developing indirect measurements of body condition based on morphology is rooted in the general problem of relative growth (Kotiaho 1999). The two fundamental dimensions of growth are time and mass, or alternatively, time and a linear dimension (Calder 1984). Thus, the increase in mass and the increase in length are parts of the same phenomenon (Thompson 1961), i.e. both body mass and linear body measurements are indicators of size per se. In many species, mass may be a better predictor of true structural size than any individual linear measure. The central challenge of a CI is to control for growth effects in body size, and implicitly, the size of distinct body components. Many biological traits are functionally related to body size by the power function $Y = \alpha X^\beta$ (Y being the biological trait to be predicted, X the body size, and α and β being constants) (Hoppeler and Weibel 2005). β , known as the ‘allometric or scaling exponent’, determines the dimensional balance or scaling between Y and X . α and β are estimated empirically by linear regression of Y against X following log transformation of the power equation (e.g. $\ln Y_i = \ln a + b \ln x_i$, where a and b are parameter estimates for α and β , and y and x the observations of the random variables Y and X respectively). Since Huxley (1932), the power equation has become the most accepted model describing the relationship between different morphological measurements. The empirical estimation of β is essential to establish the structural and functional consequences of growth on mass-length relationships (Fairbairn 1997).

If body shape and the proportions of different body components varying in density remained constant with growth or size (i.e. ‘isometry’), the scaling exponent β that relates body mass (M) to a linear body measurement (L) would be 3, since mass would be perfectly correlated with body volume (V) which in turn would be proportional to the cube of L ($M \cong V \propto L^3$). In practice, however, the relative proportion of different components and the relative length of different body parts tends to change as the overall size increases (i.e. ‘allometry’). For this reason, the scaling exponent between M and L for vertebrates tends to deviate from 3, although it usually lies between 2.5 and 3.2 (Green 2001).

The Thorpe–Lleonart model: relationship with OLS residuals from log-transformed variables

Thorpe (1975) proposed a method to standardize one measure of body size (Y) with respect to another (X) so as to take account of scaling relationships. This method was then interpreted and elegantly rearranged into a comprehensive power function (see Lleonart et al. 2000 for mathematical development) in the following form:

$$\text{Thorpe-Lleonart (TL) model of scaling: } Y_i^* = y_i \left[\frac{X_0}{X_i} \right]^b \quad (1)$$

Y_i^* being the predicted value of Y for the individual i after correcting for the underlying scaling relationship between Y and X ; x_i and y_i being the observed values of X and Y for the individual i , b the slope from the ordinary least squares (OLS) regression on log transformed Y and X variables, and X_0 an arbitrary X value (e.g. the arithmetic mean for the study population).

Because the TL model is based on OLS to compute the parameter b , the resulting Y^* scores are strongly correlated with the residual errors from OLS regression on log variables in the linear equation $\ln y_i = \ln a + b \ln x_i + \varepsilon_i$ (Supplementary material Appendix 1). However, the TL model is a power function developed from the stochastic allometric equation $Y = \alpha X^\beta e^{\varepsilon_i}$, which by a mathematical rearrangement makes the model independent of the parameter α (see Lleonart et al. 2000 for details).

Calculating the scaled mass index of body condition

The standardization technique proposed by Lleonart et al. (2000) was originally conceived for comparative morphology, both for standardizing a specific trait from different individuals to the same body size and for adjusting their shape to that they would have at size $X = X_0$ when accounting for allometry. When applied to mass-length data (i.e. taking Y as mass and X as length), this method generates standardized mass values that are strongly correlated, but not identical, to the OLS residuals of log transformed mass on log length. (Supplementary material Appendix 1). However, the TL method has three key advantages that should not be overlooked. Firstly, standardized Y^* values retain the original units of the Y . Secondly, the TL method standardizes the Y values in a two step process that better accounts for the scaling (i.e. power law) between mass and length measurements, essentially because it uses a multiplicative error function instead of an additive one. Thirdly, this method creates standardized Y values that can readily be compared between different study populations of the same species (Lleonart et al. 2000).

To our knowledge, the TL method has never previously been applied to body condition studies. We modified the TL model by computing the slope b with an alternative regression method, for reasons explained below. There are three main methods of bivariate line-fitting: OLS regression, major axis (MA) regression, and standardised major axis (SMA) regression (also known as RMA or reduced major axis). Assuming that the information on body size is partitioned between the y - (mass) and x - (length) axes of a biplot, there is no universal solution for finding the line of

best fit (McArdle 1988). Recently, Warton et al. (2006) provided useful criteria for selecting the most appropriate method, based on a priori assumptions and the purpose of the analysis. Thus, we propose the SMA regression between body mass 'M' and a length measure 'L' because (see Warton et al. 2006 for details): (1) natural variability due to growth affects both size variables (M and L) within the structural relationship (i.e. both are indicators of body size); (2) M and L variables are not measured on comparable scales, thus the magnitude of natural variability and measurement error may vary between variables (a reason for preferring SMA to MA); (3) individuals have been randomly sampled, and the observed pairs of M-L data may deviate from the underlying true association between them; (4) the measurement error in L is unlikely to be negligible and may even be higher than that in M (Krebs and Singleton 1993, Green 2001) (in Supplementary material Appendix 2, we show that CIs developed with OLS methods are highly sensitive to the existence of measurement error in L); and (5) the aim of the regression is not to predict the value M of an unsampled individual given its L value, but to estimate the scaling exponent between two interdependent variables (LaBarbera 1989, Fairbairn 1997). The term 'prediction' has a particular meaning for CIs which does not match the classical sense in statistics. The generation of OLS residuals as an index of condition does not represent an attempt to predict Y from X (both are known) but rather to score the residual variability from the fitted line in terms of variations in body mass (y-axis). Since this can also be done by other line-fitting methods such as SMA regression, the selection of the OLS method should not be justified on the basis of a need for 'classical' prediction (LaBarbera 1989).

Thus, our proposed 'scaled mass index' of body condition (\hat{M}_i) can be computed as follows:

$$\text{scaled mass index: } \hat{M}_i = M_i \left[\frac{L_0}{L_i} \right]^{b_{\text{SMA}}} \quad (2)$$

where M_i and L_i are the body mass and the linear body measurement of individual i respectively; b_{SMA} is the scaling exponent estimated by the SMA regression of M on L; L_0 is an arbitrary value of L (e.g. the arithmetic mean value for the study population); and \hat{M}_i is the predicted body mass for individual i when the linear body measure is standardized to L_0 . The scaling exponent b_{SMA} can be calculated indirectly by dividing the slope from an OLS regression (b_{OLS}) by the Pearson's correlation coefficient r (LaBarbera 1989), or directly using online software (Bohonak and van der Linde 2004). We prefer single linear measurements of body size to linear combinations computed by principal component analysis PCA (Supplementary material Appendix 3). We recommend the use of that single L variable which has the strongest correlation with M on a log-log scale, since this is likely to be the L that best explains that fraction of mass associated with structural size.

The following procedure can be used to compute the scaled mass index (Fig. 1):

(Step 1) Bivariate plots. As for any scaling study, plotting the M and L observations helps to discern which set of individuals provide the most reliable value of b_{SMA} (i.e. the best estimate of β for the power function $M = \alpha L^\beta$). b_{SMA} values close to ~ 3 on a log-log scale may be used as a guideline because the assumption of isometric growth is a fair approximation for many, but not all, species (Stevenson and Woods 2006). Statistical and/or biological considerations should be taken into account at this point, since removing outliers leads to a better refit of the regression line, but can dramatically reduce the sample size and the power of the regression model. If outliers are removed, the reasons why some individuals have been excluded prior to step 2 should be explained in detail.

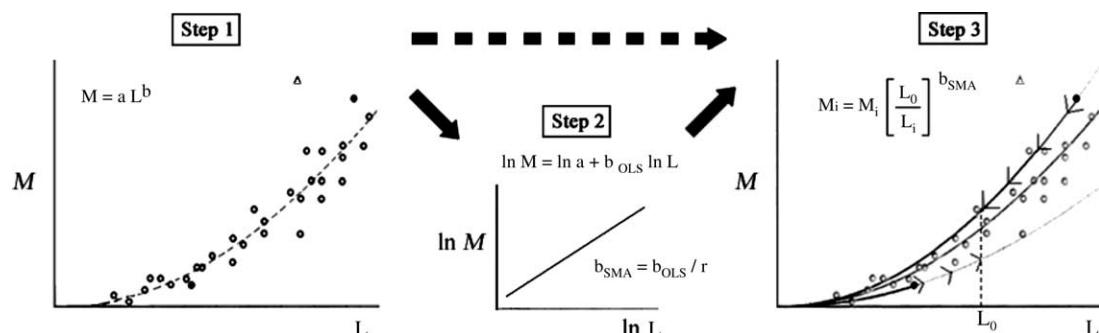


Figure 1. Stepwise procedure to compute the scaled mass index (\hat{M}_i) of body condition from body mass (M) and a linear body measurement (L). In step 1, a bivariate plot of M versus L is performed to identify points that may distort the expected relationship (e.g. the point represented by a white triangle). This outlier may be excluded in the next step. In step 2, the best fit line to the remaining points is obtained by the standardised major axis (SMA) regression on ln-transformed data. The slope is the b_{SMA} value used in Eq. 2. Alternatively, b_{SMA} can be computed by dividing the regression coefficient b_{OLS} from the OLS regression on ln-variables by the Pearson's r correlation coefficient. In step 3, the scaled index is calculated for each individual (including outliers if desired) following Eq. 2 (also shown on the figure). The arithmetic mean of L is a suitable value for L_0 , and $M_i - L_i$ variables represent the raw data for each individual i . This index adjusts the mass of all individuals to that which they would have at length L_0 (represented by the vertical dashed line). For example, data points situated on their individual curves shown in step 3 are effectively transformed to the points where this line intersects the dashed vertical line. The various curves in step 3 share a common slope in the linearized function of step 2. The scaled index can also be expected to adjust the whole body composition of each individual to that which it would have at the new length L_0 , according to allometry.

(Step 2) The line of best fit to the $\ln M - \ln L$ scattergram. The scaling exponent of the power function is estimated by calculating the regression coefficient ' b_{SMA} ' from the linearized power equation $\ln M = \ln a + b \ln L$, i.e. by fitting an SMA regression line to \ln -transformed data.

(Step 3) Selecting L_0 and computing the scaled mass index. The scaled mass index computes the mass each individual would have at a fixed body size as indicated by L_0 . The constant L_0 denotes a particular point along the M–L relationship, i.e. a specific body size along the power curve. The arithmetic mean or median of the linear body measure (L) for all sampled individuals make suitable L_0 values, but we emphasize that any value within the range of L observations can be used. Finally, according to the above Eq. 2, the scaled mass index (\hat{M}_i) of body condition (i.e. the body mass predicted when $L=L_0$) is calculated for each individual. If desired, the scaled index can even be calculated for any outliers excluded when estimating b_{SMA} during the step 1 and 2.

Exactly the same method can be used to standardize the amount of fat or other body components for a fixed size, i.e. taking M_i in Eq. 2 to be the mass of the component. Such standardisation is recommended because body components (e.g. fat, protein, water, etc.) are generally correlated with body size. Apart from energy stores, a minimum amount of protein and fat is required for structural maintenance to permit life, and these fractions are inevitably correlated with structural size. Note, if the aim is to validate the scaled mass index using data on the amount of fat or other components, it is essential to use the same L variable and L_0 value for both the scaled component mass and the scaled mass index.

In Supplementary material Appendix 4, we provide an example of the application of the scaled mass index to an ecological dataset, and show how it reveals a significant change in condition in response to a change in food supply in a fish population. An important advantage which practitioners of CIs may find useful is that, unlike OLS residuals, the scaled mass index can be easily compared between different populations and studies with adequate sample sizes (Supplementary material Appendix 4). This is largely because parameters b_{SMA} and L_0 in Eq. 2 from one population can be applied to another. In contrast, for any given study the sum of OLS residuals $\sum R_{i-\log} = 0$, making direct comparison of CIs between studies impossible.

Data used to validate CIs and to illustrate the advantages of the scaled mass index

We used data from Schulte-Hostedde et al. (2001, 2005) for adult individuals of five small mammal species, Ardia (2005) for chicks of European starlings *Sturnus vulgaris*, and Weatherhead and Brown (1996) for water snakes *Nerodia sipedon*. The following data were available: body mass, body length, skull length, and skull width from yellow-pine chipmunks *Tamias amoenus* ($n=22$); body mass, body length, foot length and ear length from deer mice *Peromyscus maniculatus* ($n=100$), meadow voles *Microtus pennsylvanicus* ($n=34$), and red-backed voles *Clethrionomys gapperi* ($n=86$); body mass, skull length and foot length from bushy-tailed wood rats *Neotoma cinerea* ($n=62$); body mass, tarsus length, wing length and head–bill length

from starlings ($n=99$); and finally, body mass, snout–vent length, tail length and total body length from snakes ($n=28$). Body mass was also divided into fat, water and lean dry mass for mammals, whilst fat mass was determined for starling, and fat, water, protein and ash mass for snakes.

The relationship between different components and body mass or different linear measures of size was initially assessed using Pearson correlation coefficients (Table 1). The scaled mass index of body condition (\hat{M}_i) was then computed following Eq. 2. We did not remove outliers during step 1 of our procedure, so our results are directly comparable with those obtained previously with the same datasets. For each species we selected the length measurement (L) with the highest correlation with body mass (M) after log transformation as the most appropriate to compute CIs. Thus we selected body length as L for all the mammals ($r=0.394$, $p=0.064$ for chipmunks; $r=0.285$, $p=0.004$ for deer mice; $r=0.773$, $p<0.001$ for meadow voles; $r=0.705$, $p<0.001$ for red-backed voles; $r=0.816$, $p<0.001$ for wood rats), head–bill length as L for starlings ($r=0.428$, $p<0.001$), and snout–vent length for snakes ($r=0.962$, $p<0.001$). L_0 was taken as the arithmetic mean of L .

Similarly, for each species, body components (fat, lean dry mass, etc) were scaled against the same L variable and the same L_0 value using Eq. 2. This provides the standardized or ‘scaled component mass’, an unbiased measure of composition which controls for scaling relationships and enables validation of non-destructive CIs calculated using M and L . Using the same L variable, OLS residuals of M against L (after log transformation) were also calculated as an alternative CI for comparison.

We correlated scaled component mass with both the scaled mass index and the OLS residual index to assess which CI best explained individual body composition, comparing our results with the analyses based on absolute component mass presented by Schulte-Hostedde et al. (2001) and Ardia (2005). For comparison, we also compared the correlations between OLS residuals for body mass and OLS residuals for body components with the correlations between scaled mass index and scaled components. For further comparison, we also correlated the OLS residual index and the scaled mass index with the percentage of body mass represented by each component. In the case of OLS residuals, this method has been used by others when validating OLS residuals (Krebs and Singleton 1993, Weatherhead and Brown 1996).

Apart from the above species, we also used published data from Cade et al. (2008) on fish (walleyes *Stizostedion vitreum*) ($n=6856$) to illustrate the application of the scaled mass index to an ecological dataset, as well as some of its methodological and conceptual advantages over the OLS residual index (Supplementary material Appendix 4). All statistical analyses were performed with SPSS ver. 15.

Results and discussion

The relationship between body components and body size

The size of different organs or other structures often scale differently with an increase in total body size (Huxley 1932,

Calder 1984, Kotiaho 1999). In the absence of isometry, using correlations with ratios or % of different body components to validate nondestructive CIs will be misleading. Correlating CIs with absolute size of body components in an attempt to validate the former is also misleading and implicitly assumes that no size variation exists. A given amount of energy reserves (e.g. 5 g of fat) has distinct implications for animals that differ in size (Kotiaho 1999). Thus, a CI cannot be validated in an unbiased manner without taking account of the scaling between body components and body size (Kotiaho 1999).

To illustrate these problems, we correlated body components (absolute and % values) with body mass (M) and linear size measurements (L) (Table 1). Absolute values tended to be positively correlated with both M and L variables. Since both M and L variables are potentially indicative of body size, this suggests that body components were generally dependent on total size. For example, absolute lean dry mass correlated positively with almost every M and L in mammals. Fat mass was significantly correlated with at least one L variable in five vertebrate species (at $p < 0.1$). Results for % values were not consistent (Table 1), confirming the existence of a diversity of scaling relationships between distinct body components and distinct measurements of body size. For example, % fat was negatively correlated with at least one L measure in three mammals, had no significant relationship with M or L in two other mammals, and was positively correlated with both M and L in starlings. This illustrates that energy stores are related to structural size, but in a complex manner in which scaling relationships vary according to the variables selected, as well as between species (and potentially populations).

Hence, the scaling relationship between different size measures and components must be taken into account in order to produce reliable CIs, yet neither ratios nor OLS residuals can do this satisfactorily. For example, in a given population, larger animals may have more fat (absolute value) and at the same time have less % fat (a case of negative allometry in the size of fat stores) when compared with smaller individuals. We should not assume that larger animals are in better condition because they have a higher absolute fat value, nor that smaller animals are in better condition because they have more % fat. It is likely that every body size has an optimal amount of energy reserves for that specific size, and to estimate the condition of a given individual we should consider how it compares with individuals of the same size, i.e. a CI should remove the covariation between body size and body components. This is precisely the aim of the scaled mass index. As the allometric relationships depend on the variables considered, the mass of components needs to be standardised respect to the same L variable and L_0 value used for the CI. Although any body length value could be used as L_0 , we recommend those values which falls in the middle range of L (e.g. arithmetic mean, geometric mean, or median), since confidence intervals tend to be narrower in the middle of the fitted $\ln M - \ln L$ line than at its extremes.

Residuals are often generated from untransformed data (Schulte-Hostedde et al. 2001, Ardia 2005), yet log transformation is essential to convert the power law into a linear relationship (Krebs and Singleton 1993, Weather-

head and Brown 1996, but see Material and methods). In fact, use of residuals as a CI makes immediate and often erroneous assumptions about the scaling relationship between energy reserves and size. As pointed out by Kotiaho (1999), use of residuals from a regression of untransformed M–L data as a CI assumes that one unit residual equals the same condition irrespective of L, i.e. that the underlying relationship between energy reserves and L is negatively allometric. On the other hand, use of residuals from a regression of log transformed variables assumes that one unit residual is relatively the same for all individuals, i.e. that the underlying scaling relationship between energy reserves and L is isometric.

Validating CIs as predictors of true condition

First, we compared the SMA and OLS slopes of regression of $\ln M$ against $\ln L$, and confirmed that there were significant differences for all species (Table 2), with OLS slopes consistently underestimating the scaling relationships. We found that the scaled fat mass was more closely related to the scaled mass index than to OLS residuals of $\ln M$ on $\ln L$ in six out of seven species (the exception being deer mice, Table 3). In five of six species, the correlation coefficients between scaled fat mass and the scaled mass index were also higher than the original r values between OLS residuals and absolute fat cited by Schulte-Hostedde et al. (2001) and Ardia (2005) as evidence that such residuals provide a suitable CI. Likewise, the scaled mass index was consistently better correlated with other standardized components (lean dry mass, water, protein and ash) than OLS residuals (Table 3), and these correlations were also higher than those between OLS residuals and absolute components as reported by the aforementioned authors. This indicates that the scaling relationship between different components is properly accounted for by Eq. 2, and that the scaled mass index explains more of the variance in individual body components than OLS residuals.

In deer mice, the results may appear at first glance to support the validity of OLS residuals, since we found a stronger correlation between scaled fat and OLS residuals

Table 2. Computed slopes from an OLS (model I) and SMA (model II) regression of body mass against a linear body measurement (both log-transformed) in seven vertebrate species. Confidence intervals were calculated following Sokal and Rohlf (1995). p-values of t-tests for differences in slopes are also shown.

Species	Model	Slope [95% CI]	p-value
Chipmunk	OLS	0.9 [-0.1, 1.9]	<0.001
	SMA	2.3 [1.3, 3.3]	
Deer mouse	OLS	0.4 [0.1, 0.7]	<0.001
	SMA	1.4 [1.1, 1.7]	
Meadow vole	OLS	2.3 [1.6, 3.0]	<0.001
	SMA	2.9 [2.3, 3.6]	
Red-backed vole	OLS	1.4 [1.1, 1.7]	<0.001
	SMA	2.0 [1.7, 2.3]	
Woodrat	OLS	3.0 [2.4, 3.5]	<0.001
	SMA	3.6 [3.1, 4.2]	
Starling	OLS	1.9 [1.4, 2.4]	<0.001
	SMA	3.2 [2.6, 3.7]	
Water snake	OLS	3.0 [2.6, 3.3]	<0.05
	SMA	3.1 [2.7, 3.5]	

Table 3. Correlations between the scaled mass index or OLS residuals (from a regression of $\ln M$ on $\ln L$) with body components. Both indices were correlated with component mass scaled according to Eq. 2. For comparison, correlations between OLS residuals and component mass as a percentage of body mass are also presented. Correlations with the absolute mass of body components were already reported by Schulte-Hostedde et al. (2001) and Ardia (2005). Significant differences at $p < 0.05$ are in bold. LDM = lean dry mass.

	Scaled	Scaled mass index		OLS residual index		%	OLS residual index	
		r	p	r	p		r	p
Chipmunk	fat	0.603	0.003	0.302	0.173	fat	0.118	0.602
	LDM	0.888	< 0.001	0.699	< 0.001	LDM	-0.402	0.064
	water	0.906	< 0.001	0.712	< 0.001	water	-0.167	0.457
Deer mouse	fat	-0.151	0.134	0.263	0.008	fat	0.017	0.870
	LDM	0.729	< 0.001	0.415	< 0.001	LDM	-0.127	0.207
	water	0.933	< 0.001	0.688	< 0.001	water	0.015	0.883
Meadow vole	fat	0.164	0.353	-0.010	0.956	fat	-0.083	0.640
	LDM	0.869	< 0.001	0.781	< 0.001	LDM	-0.222	0.206
	water	0.980	< 0.001	0.914	< 0.001	water	0.306	0.078
Red-backed vole	fat	0.444	< 0.001	0.196	0.071	fat	0.032	0.770
	LDM	0.823	< 0.001	0.714	< 0.001	LDM	-0.167	0.124
	water	0.975	< 0.001	0.893	< 0.001	water	0.218	0.044
Wood rat	fat	0.323	0.011	0.200	0.119	fat	0.038	0.769
	LDM	0.876	< 0.001	0.820	< 0.001	LDM	-0.029	0.823
	water	0.976	< 0.001	0.929	< 0.001	water	-0.094	0.465
Starling	fat	0.841	< 0.001	0.684	< 0.001	fat	0.545	< 0.001
	Water snake	fat	0.748	< 0.001	0.488	0.008	fat	0.539
	protein	0.625	< 0.001	0.497	0.007	protein	-0.215	0.271
	water	0.961	< 0.001	0.713	< 0.001	water	-0.223	0.254
	ash	0.418	< 0.001	0.250	0.199	ash	-0.614	0.001

than between scaled fat and the scaled mass index (Table 3). However, OLS residuals do not adequately control for structural size, thus the correlation between them and scaled fat is likely to be misleading. For deer mice, scaled lean dry mass or water had a much higher correlation with the scaled mass index than with OLS residuals (Table 3). This suggests that, in this species, being heavier for a given length has more to do with having additional proteins or water than with fat.

For comparison, we also carried out correlations between OLS residuals (after log transformation) and relative component mass calculated as a ratio (% of total mass), in a way calculated by other researchers (Krebs and Singleton 1993, Weatherhead and Brown 1996). In all seven species and in 19 of 20 analyses, the correlation between individual body components (as % of total mass) and OLS residuals was weaker than that between the scaled component and the scaled mass index (Table 3). Components as % of body mass make no allowance for scaling relationships. Hence it is no surprise that correlations between scaled mass index and components as % of mass also fared badly. In 10 analyses, such correlations were higher than those between OLS residuals and components as % of mass, and in 10 they were lower (results not shown).

On occasions, the OLS residual index has been also validated relative to residual components from a OLS regression (Redfern et al. 2000), which in principle appears more reasonable than using absolute or percent values of body components. However, these correlations were still weaker than our suggested approach for validation. Thus, in 15 of 20 analyses the correlations between the scaled mass index and scaled components led to higher r values than those between the OLS residual index and OLS component

residuals (after log transformation), and in one analysis there was an equal r value (results not shown).

Comparing scores between CIs

When using the same values of M and L, the scaled mass index and OLS residuals were significantly correlated but to a varying degree ($r^2 = 0.636\text{--}0.963$, Fig. 5-1 in Supplementary material Appendix 5). Given that OLS residuals fail to control effectively for the effect of growth on body size and the scaling relationship between M and L, and that the scaled mass index is a more reliable indicator of body composition, the differences between these two CIs are indicative of the potential for bias when using OLS residuals. The CIs proposed by Schulte-Hostedde et al. (2001, 2005) and Ardia (2005) were OLS residuals based on different L variables than those used in the present paper, and had a weaker relationship with the scaled mass indices of Table 3 ($r^2 = 0.255\text{--}0.863$, Fig. 5-2 in Supplementary material Appendix 5).

Conclusions

Given the consistency of the method with statistical and biological principles, and its successful validation using data on body composition from seven species, the scaled mass index can be considered as an improvement on existing CIs based on mass and length data. This new CI should be very useful for a broad range of studies in animal ecology, conservation biology and wildlife management. Published studies based on use of OLS residuals should be interpreted with caution unless their results are robust enough to be repeatable using the scaled mass index (Green 2001).

Numerous CIs have been developed and are applied in different scientific disciplines (reviewed by Stevenson and Woods 2006), and traditions have often been established in which different methods are used in different fields such as fisheries biology, mammalogy, ornithology or human biology. This is not because the accuracy of different CIs has been shown to vary according to the organism group, but rather because scientists tend to specialize and to repeat the methods of their peers. The diversity of CIs available should remind us that no method is ideal and that successfully estimating condition is a difficult challenge. We recommend that users of CIs adopt a healthy skepticism and should not place a blind faith in any individual method, especially if they lack data on body composition to validate the CIs. No CI should be assumed to accurately reflect 'true condition' without such validation. Nevertheless, when developing a CI it is vital to recognize the importance of the scaling relationship between different measures of body size (including mass) which may vary between species and populations. Unlike OLS residuals, mass/length ratios and many other CIs, the scaled mass index is based on this central principle and is thus more likely to be reliable.

Acknowledgements – We are extremely grateful to A. I. Schulte-Hostedde, D. R. Ardia, P. J. Weatherhead and G. P. Brown for allowing us to reanalyze their data. We also thank J. L. Parra, J. D. Rodríguez-Teijeiro and J. Nadal for valuable comments on earlier versions of this manuscript. JP was supported by a FPU fellowship from the Spanish MEC. (AP99-dni). This work was partially funded by the Spanish (project no. BOS 2000-0567) and Catalan Government (project no. 2001SGR-00090 and 2001ACOM-00009).

References

- Ardia, D. R. 2005. Super size me: an experimental test for the factors affecting lipid content and the ability of residual body mass to predict lipid stores in nestling European starlings. – *Funct. Ecol.* 19: 414–420.
- Arnold, T. W. and Green, A. J. 2007. On the allometric relationship between size and composition of avian eggs: a reassessment. – *Condor* 109: 705–714.
- Bonhak, A. J. and van der Linde, K. 2004. RMA: software for reduced major axis regression, Java version. <www.kimvdlinde.com/professional/rma.html>, accessed 28 March 2008.
- Cade, B. S. et al. 2008. Estimating fish body condition with quantile regression. – *N. Am. J. Fish Manage.* 28: 349–359.
- Calder, W. A. 1984. Size, function and life history. – Harvard Univ. Press.
- Fairbairn, D. J. 1997. Allometry for sexual size dimorphism: patterns and process in the coevolution of body size in males and females. – *Annu. Rev. Ecol. Evol. Syst.* 28: 659–687.
- Freckleton, R. P. 2002. On the misuse of residuals in ecology: regression of residuals vs multiple regression. – *J. Anim. Ecol.* 71: 542–545.
- García-Berthou, E. 2001. On the misuse of residuals in ecology: testing regression residuals vs the analysis of covariance. – *J. Anim. Ecol.* 70: 708–711.
- Gosler, A. G. 1996. Environemnatl and social determinants of winter fat storage in the great tit *Parus major*. – *J. Anim. Ecol.* 65: 1–17.
- Green, A. J. 2001. Mass/length residuals: measures of body condition or generators of spurious results? – *Ecology* 82: 1473–1483.
- Hayes, J. P. and Shonkwiler, J. S. 2001. Morphometric indicators of body condition: worthwhile or wishful thinking? – In: Speakman, J. R. (ed.), *Body composition analysis of animals. A handbook of non-destructive methods*. Cambridge Univ. Press, pp. 8–38.
- Hayes, J. P. and Shonkwiler, J. S. 2006. Allometry, antilog transformations, and the perils of prediction on the original scale. – *Physiol. Biochem. Zool.* 79: 665–674.
- Hoppeler, H. and Weibel, E. R. 2005. Scaling functions to body size: theories and facts. – *J. Exp. Biol.* 208: 1573–1574.
- Husak, J. F. 2006. Does speed help you survive? A test with collared lizards of different ages. – *Funct. Ecol.* 20: 174–179.
- Huxley, J. S. 1932. Problems of relative growth. – Dover Publications.
- Jakob, E. M. et al. 1996. Estimating fitness: a comparison of body condition indices. – *Oikos* 77: 61–67.
- Kotiaho, J. S. 1999. Estimating fitness: comparison of body condition indices revisited. – *Oikos* 87: 399–400.
- Krebs, C. J. and Singleton, G. R. 1993. Indices of condition for small mammals. – *Aust. J. Zool.* 41: 317–323.
- LaBarbera, M. 1989. Analyzing body size as a factor in ecology and evolution. – *Annu. Rev. Ecol. Evol. Syst.* 20: 97–117.
- Liker, A. et al. 2008. Lean birds in the city: body size and condition of house sparrows along the urbanization gradient. – *J. Anim. Ecol.* 77: 789–795.
- Lleonart, J. et al. 2000. Removing allometric effects of body size in morphological analysis. – *J. Theor. Biol.* 205: 85–93.
- McArdle, B. H. 1988. The structural relationship: regression in biology. – *Can. J. Zool.* 66: 2329–2339.
- Nagy, K. A. et al. 2002. A condition index for the desert tortoise (*Gopherus agassizii*). – *Chelonian Conserv. Biol.* 4: 425–429.
- O'Dwyer, T. W. et al. 2006. Prolactin, body condition and the cost of good parenting: an interyear study in a long-lived seabird, Gould's petrel (*Pterodroma leucoptera*). – *Funct. Ecol.* 6: 806–811.
- Redfern, C. P. F. et al. 2000. Fat and body condition in migrating redwings *Turdus iliacus*. – *J. Avian Biol.* 31: 197–205.
- Seim, E. and Sæther, B. E. 1983. On rethinking allometry: which regression model to use? – *J. Theor. Biol.* 104: 161–168.
- Schulte-Hostedde, A. I. et al. 2001. Evaluating body condition in small mammals. – *Can. J. Zool.* 79: 1021–1029.
- Schulte-Hostedde, A. I. et al. 2005. Restitution of mass-size residuals: validating body condition indices. – *Ecology* 86: 155–163.
- Smith, R. J. 1992. Logarithmic transformation bias in allometry. – *Am. J. Phys. Anthropol.* 90: 215–228.
- Sokal, R. R. and Rohlf, F. J. 1995. *Biometry*. – W. H. Freeman.
- Stevenson, R. D. and Woods, W. A. 2006. Condition indices for conservation: new uses for evolving tools. – *Integr. Comp. Biol.* 46: 1169–1190.
- Thompson, D. W. 1961. On growth and form. – Cambridge Univ. Press.

- Thorpe, R. S. 1975. Quantitative handling of characters useful in snake systematics with particular reference to intraspecific variation in the ringed snake *Natrix natrix* (L.). – Biol. J. Linn. Soc. 7: 27–43.
- Toigo, C. et al. 2006. How does environmental variation influence body mass, body size, and body condition? Roe deer as a case study. – Ecography 29: 301–308.
- Warton, D. I. et al. 2006. Bivariate line-fitting methods for allometry. – Biol. Rev. 81: 259–291.
- Weatherhead, P. J. and Brown, G. P. 1996. Measurement versus estimation of body condition in snakes. – Can. J. Zool. 74: 1617–1621.

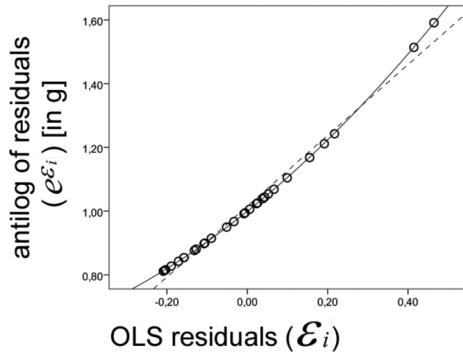
Supplementary material (available online as Appendix O17643 at <www.oikos.ekol.lu.se/appendix>). Appendix 1. Similarities between the Thorpe–Lleonart equation and residuals from OLS regression on log transformed variables. Appendix 2. On the consequences of measurement errors in body mass (M) and body length (L) for condition indices dependent on OLS methods. Appendix 3. Drawbacks to the use of PCA as a body size indicator when producing condition indices. Appendix 4. An example applying the scaled mass index to the effect of a change in habitat quality on body condition. Appendix 5. The relationship between the OLS residual and the scaled mass indices of condition.

Appendix 1.

Similarities between the Thorpe–Lleonart equation and residuals from OLS regression

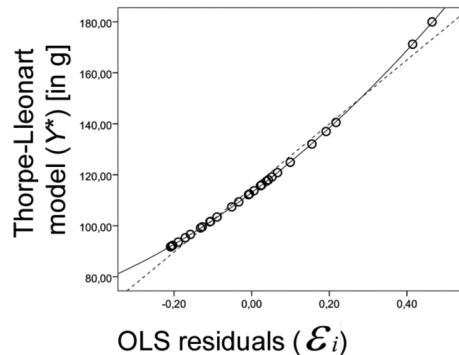
A.

LINEAR: $r^2 = 0.987 (P < 0.001)$
EXPONENTIAL: $r^2 = 1.000 (P < 0.001)$



B.

LINEAR: $r^2 = 0.987 (P < 0.001)$
EXPONENTIAL: $r^2 = 1.000 (P < 0.001)$



C.

LINEAR: $r^2 = 1.000 (P < 0.001)$

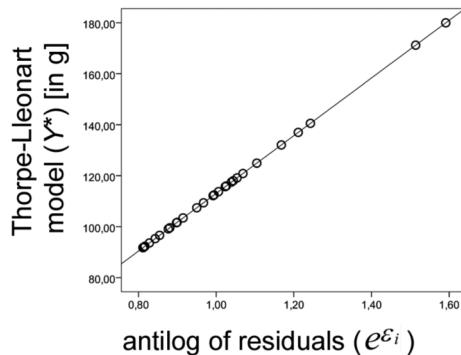


Figure 1–1. Example in water snakes of the linear and exponential functions relating values computed by the Thorpe–Lleonart model, OLS residuals from OLS regression on ln-transformed variables, or back-transformed (antilog) residuals. See Table 1–1 for the corresponding data for several individual snakes.

The scaled mass index is based on two main statistical principals: the mathematical basis of the Thorpe–Leonart (TL) model (itself reliant on OLS regression), combined with the use of standardised major axis regression as the method of line-fitting. The starting point of the Thorpe–Leonart model is a modified power equation which predicts the whole body mass when X equals a specific arbitrary value (X_0) via the stochastic allometric equation $Y = aX^b e^\varepsilon$, which contains a multiplicative error term (e^{ε_i}). In this case ε_i are roughly equivalent to the residual errors from the linear OLS regression equation $\ln Y = \ln a + b \ln X + \varepsilon$ (but see Hayes and Shonkwiler 2006).

From the allometric equation $Y = aX^b e^\varepsilon$, if \hat{y} denotes the predicted value for Y (i.e. $\hat{y} = a x^b$), then the observed y values become $y_i = \hat{y}_i e^{\varepsilon_i}$. Similarly, the linear regression equation can be expressed as $\ln y_i = \ln \hat{y}_i + \varepsilon_i$, which after backtransformation (i.e. taking antilogs) leads also to $y_i = \hat{y}_i e^{\varepsilon_i}$. This explains why a $e^{\varepsilon_i} - \varepsilon_i$ plot produces a perfect exponential fit (Fig. 1-1A). An identical fit is provided when plotting the predicted Y^* scores from the TL model against the OLS residuals (Fig. 1-1B). Likewise, the Y^* scores are perfectly correlated with the antilogs of the OLS residuals from regression of $\ln Y$ against $\ln X$ (Fig. 1-1C).

Despite this close statistical relationship between the TL model and the OLS residual index, they are not mathematically or biologically equivalent (Table 1-1). The error terms from the simple linear equation on raw Y–X variables $Y = a + b X + \varepsilon$ (for convenience indicated from hereon as ε_{raw}), on log variables $\ln Y = \ln a + b \ln X + \varepsilon$ (indicated from hereon as ε_{\log}), and from the stochastic power equation $Y = a X^\beta e^{\varepsilon_i}$ (indicated from hereon as $e^{\varepsilon-\text{power}}$) are not identical to each other (i.e. $\varepsilon_{\text{raw}} \neq \varepsilon_{\log} \neq e^{\varepsilon-\text{power}}$) nor to the Y^* scores from the TL model of Eq. 1.

The use of $R_{i-\text{raw}}$ values (additive error, where R indicates residuals) gives CI scores the same units as the independent Y variable (body mass in g), but relies on ad hoc models without incorporating the scaling principle between mass and length. In contrast, the use of $R_{i-\log}$ values invokes an underlying allometric model with multiplicative error $Y = a X^\beta e^{\varepsilon_i}$ (Packard 2009), but CI scores are no longer in readily comprehensible units. This is because $R_{i-\log}$ values are additive errors in the logged model with the same units as the Y^* ($= \ln Y$) variable. Residuals may seem attractive in a condition context because they provide positive and negative scores which imply ‘good’ and ‘bad’ condition respectively (Table 1-1). However, $R_{i-\log}$ scores can not be interpreted in a more quantitative way unless they are mathematically modified. For instance, according to these scores, individual no. 4 from Table 1-1 was relatively ‘heavier’ than individual no. 3 (0.46 vs 0.04, in ‘ $\ln g$ ’ units). However, further consideration may be confusing, since direct comparison of the scores might imply that the condition of snake no. 4 is approximately 12.1 times ($= 0.464 / 0.038$) greater than that of no. 3 (i.e. that no. 4 is 12.1 times ‘relatively heavier or fatter’). This would be very misleading from a biological perspective, given the scale and units of measurement. Furthermore, such comparisons (ratios between individual scores) are even more

non-sensical when there are negative values in the denominator (e.g. when comparing the $R_{i-\log}$ score for snakes no. 3 and no. 1 in Table 1-1).

An attempt can be made to recover the original scale by the back-transformation of $R_{i-\log}$ values by taking anti-logs (i.e. $e^{R_{i-\log}}$, the fourth column in Table 1-1). Such back-transformed results may then suggest that snake no. 4 is in 1.53 times ($= 1.59 / 1.04$) better condition than no. 3, which is biologically more reasonable. However, such back-transformation is rarely used in the literature and requires an additional calculation compared to the TL model. Note also that justifying an allometric fit by such backtransformation is not tenable because the multiplicative error term of the stochastic power equation is necessarily adimensional (i.e. while $e^{R_{i-\log}}$ would have units, the error term is unitless: $e^\varepsilon = Y/a X^\beta = Y/\hat{Y}$, then $[g]/[g] = [\emptyset]$). Furthermore, these back-transformed $e^{R_{i-\log}}$ and TL scores are different (compare columns 4 and 6 in Table 1-1), and only the TL model computes the whole body mass for a given body length.

Estimating the whole body mass in the original scale from $R_{i-\log}$ values requires adding the arithmetic mean of the dependent variable, i.e. $\ln Y + R_{i-\log}$ (or $Y^* + R_{i-\log}$), and then backtransforming the sum via antilogs. Such a process provides the predicted body mass for the geometric mean of body length. Following with our example from Table 1-1, the predicted body mass for snakes no. 3 and no. 4 at body length of 56.25 cm (= antilog of $\ln X$ or \bar{X}) would be $\hat{Y}_3 = 110.41\text{g}$ and $\hat{Y}_4 = 169.09\text{g}$ respectively (i.e. no. 4 is 1.53 times ‘heavier’ than no. 3). Note that the ratio (1.53) remains the same as for $e^{R_{i-\log}}$.

However, the ‘antilogs of $[Y^* + R_{i-\log}]$ ’ are still not equivalent to the results of the TL model (compare columns 5 and 6 in Table 1-1) because the former results are based around the geometric mean of body mass, not the arithmetic mean.. Thus, according to the results of the TL model, snake no. 4 was again 1.53 times heavier than snake no. 3 when both were standardized to a length of 57.43 cm (Table 1-1). However, this value is in fact slightly different (1.531424 for antilogs, 1.531430 for Y^*) owing to the subtle differences in the way these different values are computed, as explained above.

Y^* scores produced by the TL model are directly developed from the allometric power equation, a nonlinear model likely to provide better fit to the true M–L relationship. Unlike OLS residuals, the CI provided by the TL model avoids the need for tiresome further calculations to provide results in the original scale, and single Y^* values are easier to grasp from a condition perspective, since they predict the whole body mass for a given body length. Our scaled mass index provides different results to the TL model because it relies on line-fitting by SMA regression. For example, according to the scaled mass index, snake no. 4 was 1.67 times heavier than snake no. 3 when both were standardized to a length of 57.43 cm (Table 1-1).

Table 1-1. Examples in water snakes (total n = 28) of individual scores which could be potentially used as condition indices. $R_{i\text{-raw}}$, residuals from an OLS regression of body mass (M) against snout-vent length (L); $R_{i\text{-log}}$, residuals from an OLS regression of $\ln M$ against $\ln L$; $e^{R_{i\text{-log}}}$, antilogs of $R_{i\text{-log}}$; Antilog of $[\ln Y + R_{i\text{-log}}]$, antilogarithms of the sum of $R_{i\text{-log}}$ and the arithmetic mean value of $\ln M$; Y^* , values computed by the Thorpe–Leonart model following equation 1; \hat{M}_i , values computed by the scaled mass index following Eq. 2 ($L_0 = 57.43$ cm).

No. individual snake	$R_{i\text{-raw}}$	$R_{i\text{-log}}$	$e^{R_{i\text{-log}}}$	Antilog of $[\ln Y + R_{i\text{-log}}]$	Y^*	\hat{M}_i
	(g)	(\log_e g)	(g)	(g)	(g)	(g)
1	-73.93	-0.19	0.83	87.93	93.59	90.18
2	-15.03	0.10	1.10	117.37	124.92	123.90
3	5.54	0.04	1.04	110.41	117.52	112.43
4	78.63	0.46	1.59	169.09	179.97	188.30
5	6.11	0.02	1.02	108.73	115.73	117.66
6	-17.30	-0.20	0.82	86.72	92.30	93.47

References

- Hayes, J. P. and Shonkiler, J. S. 2006. Allometry, antilog transformations, and the perils of prediction on the original scale. – *Physiol. Biochem. Zool.* 79: 665–674.
 Packard, G. C. 2009. On the use of logarithmic transformations in allometric analyses. – *J. Theor. Biol.* 257: 515–518.

Appendix 2

On the consequences of measurement errors in body mass (M) and body length (L) for condition indices dependent on OLS methods

The OLS method assumes that there is no natural variability or measurement error (which also includes the sampling error) in the X (or L) variable (Warton et al. 2006). Apart from the dubious assumption about the inexistence of natural variability in body length, the OLS assumption is clearly violated since length measures of vertebrates are also subject to considerable measurement error (Yezerinac et al. 1992). We did not have access to datasets with repeated measurements of mass or length from the same animal, but here we consider the importance of the potential effects of measurement error. The Thorpe–Leonart (TL) model based on OLS regression produces standardized mass values that are strongly correlated with OLS residuals from the simple linear regression performed on ln-transformed variables (Appendix 1). Thus, we used the TL model (see Eq. 1 in the main text) to assess the consequences of measurement error for body mass (M) and length measurements (e.g. body length or tarsus length) for the reliability of OLS methods.

For example, the Thorpe–Leonart equations for meadow voles, starlings and water snakes calculated for data from Schulte-Hostedde et al. (2001), Ardia (2005) and Weatherhead and Brown (1996) respectively can be written as:

$$\text{meadow voles: } M_i^* = M_i (109.147 / \text{body length}_i)^{(2.291)}$$

$$\text{starlings: } M_i^* = M_i (44.338 / \text{head_bill length}_i)^{(1.927)}$$

$$\text{water snakes: } M_i^* = M_i (57.436 / \text{snout_vent length}_i)^{(2.989)}$$

For example, X_0 in equation 1 from the main text for voles is taken as the arithmetic mean for body length = 109.147 mm.

These formulas produce values that are very strongly correlated with the OLS residuals from ln-body mass against ln-body length for meadow voles, ln-head_bill length for starlings or ln-snout_vent length for snakes. The coefficient of variation (CV) is independent of the measurement scale and we assumed an arbitrary measurement error of CV = 5% when taking biometric measurements of a random meadow vole whose true values are body mass $M = 36.20\text{g}$ and body length $L = 109.00\text{mm}$. Assuming a normal distribution for both variables, a 5% measurement error in body mass would imply a standard deviation of $SD_M = 1.81$, and consequently 95 % confidence intervals for measured mass of [32.65, 39.75] (in g). If the confidence interval extremes were two repeated measures of body mass, the two pairs of M–L measurements for the vole would be 32.65 g – 109 mm (A) and 39.75 g – 109 mm (B). According to the above equation, the mass predicted by the TL model would be: $M_i^* = 32.75\text{g}$ (A) and $M_i^* = 39.87\text{g}$ (B); i.e. a $CV_{M^*} = 9.80\%$ (respect to the true value of body mass).

Repeating this exercise for a 5% measurement error in body length would give an $SD_L = 5.45\text{mm}$ with 95 % confidence intervals of [98.32, 119.68] (in mm). If the confidence interval extremes were two repeated measures of body length, the two pairs of repeated M–L observations for the vole would be 36.20 g – 98.32 mm (C) and 36.20 g – 119.68 mm (D). According to the

above equation, the predicted mass would now be $M_i^* = 45.99\text{ g}$ (C) and $M_i^* = 29.31\text{ g}$ (D); i.e. $CV_{M^*} = 23.04\%$.

Similarly, in a random starling whose true values are body mass $M = 78.30\text{ g}$ and head_bill length $L = 45.02\text{ mm}$, a 5% measurement error in body mass would cause $CV_{M^*} = 9.52\%$, and 5% measurement error in head_bill length would cause $CV_{M^*} = 18.68\%$. Finally, for a water snake whose body mass is $M = 154\text{ g}$ and snout_vent length $L = 61.60\text{ cm}$, the error in body mass would imply a $CV_{M^*} = 15.63\%$ and in snout_vent length a $CV_{M^*} = 48.27\%$. These results are summarized in Table 2-1 of this Appendix:

Table 2-1. Coefficient of variation in predicted body mass (Thorpe–Leonart model based on OLS methods) in three different species caused by an equal proportion of measurement error when measuring mass or length.

	Meadow vole	Starlings	Water snake
5% measurement error in:			
Body mass (M or Y)	9.80%	9.52%	15.63%
Length measurement (L or X)	23.04%	18.68%	48.27%

Clearly, a given measurement error in L has much greater consequences for condition indices dependent on OLS methods than the same error in M.

This simulation is consistent with empirical results obtained by Krebs and Singleton (1993) in small mammals, which showed the important effect of measurement error in length measures compared with body mass measures. Linear body measures are often taken with calipers or rulers, and the final value depends on the technical skill of the investigator. In contrast, body mass is usually measured with spring scales or analytical balances without direct intervention of the observer, with accuracy being more dependent on the calibration of the tool. Thus, measurement error in L may be substantially higher than that in M, and a 5% measurement error is not unlikely when measuring lengths. In fact, Yezerinac et al. (1992) found that as much as 10–30% of the total variance in length in fifteen skeletal characters in passerines was due to the inability to make precise measurements, and concluded that, the smaller the morphological trait being measured, the greater the effect of measurement error. Thus, the measurement error in both variables should not be neglected when selecting the method (model I [OLS] or model II [SMA] regression) for fitting a line to the M–L dataset.

In conclusion, measurement error exists for linear measures of body size (traditionally used as the X value for the calculation of OLS residuals) and should not be ignored. OLS methods produce condition indices that are especially sensitive to measurement error in L. SMA methods are more appropriate because they recognize the existence of measurement error and natural variability in L (Green 2001, Warton et al. 2006).

References

- Ardia, D. R. 2005. Super size me: an experimental test for the factors affecting lipid content and the ability of residual body mass to predict lipid stores in nestling European starlings. – *Funct. Ecol.* 19: 414–420.
- Green, A. J. 2001. Mass/length residuals: measures of body condition or generators of spurious results? – *Ecology* 82: 1473–1483.
- Krebs, C. J. and Singleton, G. R. 1993. Indices of condition for small mammals. – *Aust. J. Zool.* 41: 317–323.
- Schulte-Hostedde, A. I. et al. 2001. Evaluating body condition in small mammals. – *Can. J. Zool.* 79: 1021–1029.
- Warton, D. I. et al. 2006. Bivariate line-fitting methods for allometry. – *Biol. Rev.* 81: 259–291.
- Weatherhead, P. J. and Brown, G. P. 1996. Measurement versus estimation of body condition in snakes. – *Can. J. Zool.* 74: 1617–1621.
- Yezerinac, S. M. et al. 1992. Measurement error and morphometric studies: statistical power and observer experience. – *Syst. Biol.* 41: 471–482.

Appendix 3

Drawbacks to the use of PCA as a body size indicator when producing condition indices

The usage of PC1 from a PCA of different morphometric measurements has often been advocated as the best size indicator for use in constructing non-destructive indices of condition (Green 2001, Schulte-Hostedde et al. 2001, Blackwell 2002). However, this should not be taken as a rule of thumb (LaBarbera 1989), and there are several reasons why the use of a single linear measure (e.g. body length) can be preferable:

1) individual measures can better reflect structural size than PC1, and often correlate better with body mass or body components. This is the case for starlings (Ardia 2005) and also for red-backed voles (data reanalyzed from Schulte-Hostedde et al. [2001], $r = 0.705$ for ln-body length against ln-body mass; $r = 0.477$ for ln-PC1 [from a PCA on body length, foot length and ear length] against ln-body mass). Similarly, in the wood mouse *Apodemus sylvaticus*, the Algerian mouse *Mus spretus* and the greater white-toothed shrew *Crocidura russula*, body length is correlated more strongly with body mass than PC1 from a PCA of four linear measurements (Peig unpubl.). One reason for this is that some measurements are harder to take and more subject to measurement error than others (e.g. it is almost inevitable that the length of a small structure such as an ear or foot is measured less accurately than body length; Yezernac et al. 1992). Another reason is that some measures (e.g. tail length) may reflect plastic characters that have a poor relationship with overall structural size (Yezernac et al. 1992, Badyaev and Martin 2000). However, a PCA weights all variables equally so that combining poor indicators of structural size with good ones through a PCA can provide relatively poor information.

2) conducting a PCA involves the inherent loss of information due to the reduction of the dimensional space (Shea 1985), and even PC1 rarely explains more than 55% of the variation in linear measures (Schulte-Hostedde et al. 2001). Furthermore, it has sometimes been suggested that, in a PCA of linear measures, PC1 can be taken to represent size while PC2 represents shape (Shea 1985, Blackwell 2002). This is because the loading factors of different linear measurements on PC1 are often consistently positive. For example, Schulte-Hostedde et al. (2001) considered PC1 to represent structural size only if all linear measures have positive signs, and for that reason they rejected PC1 as a valid size measure for meadow voles. Principal components may indeed represent integrated information about size. However, they also encompass information about animal shape (Shea 1985, Leonart et al. 2000), even when the computed loading factors with linear measures are all positive. This is because different absolute values of the loading factors indicate different rates of increase of the original variables. For example, in *Crocidura russula*, PC2 from a PCA of four linear measures was correlated more strongly with body mass than PC1,

and hence is probably a better indicator of structural size (Peig unpubl.). This is despite the fact that only PC1 has positive factor loadings for all four measures.

3) another disadvantage of using PCA is that it can complicate the interpretation of scaling relationships between body mass and linear size measures. A PCA involves dimensional reduction by linear combination of variables, and creates a new dimension (e.g. PC1) which does not represent either a linear measure of physical size, nor any other known physical dimension such as volume or mass. Thus, when body mass is regressed against such variables, the computed values of the scaling exponent β are totally unpredictable and can not be compared with the value of 3 that would be expected under isometry. For example, the OLS slope between mass and PC1 was as low as 0.69 for meadow voles and as high as 68.15 for wood rats (Schulte-Hostedde et al. 2001).

4) similarly, using PCA as L in equation 2 complicates the interpretation of the values of the scaled mass index. If, for example, L is body length (mm), and L_0 is taken as the arithmetic mean of length for all individuals studied, the scaled mass index represents individual mass standardized for the mean length L_0 (mm). These values are easier to understand and to compare between studies than values based on PC1 (a parameter which can not be measured in mm).

5) in order to compute reliable principal components that can then be applied to analyses based on OLS methods, it is important that the linear measures used for a PCA have a normal distribution (Sokal and Rohlf 1995). However, this is often not the case. For example, Schulte-Hostedde et al. (2001) calculated OLS residual indices of condition for deer mice and red-backed voles using PC1 from a PCA including body length, foot length and ear length. However, foot length and ear length were discontinuous variables (rounded to the nearest mm) and did not have a normal probability distribution, even after log transformation (our own re-analysis, Kolmogorov-Smirnov Z test: $p < 0.05$). In addition, extracting principal components from length measurements can greatly distort their original relationship with body mass due to transformations of variables. For instance, Schulte-Hostedde et al. (2005) conducted a PCA on log transformed length measurements, and the resulting PC1 was again log transformed to remove heteroscedasticity in the final regression model with log body mass as an independent variable. Such transformations on transformed data complicate the interpretation of the results.

In conclusion, although any index of condition can be calculated using a principal component from a PCA, we consider the use of the single linear measure that is best correlated with body mass (after log transformation) to be the rule of thumb most likely to produce useful results that can be readily interpreted. Finally, it is worth emphasizing that the scaled mass index does not rely on a single size measurement chosen by subjective criteria, but on two size indicators (mass and length) strongly correlated as reliable estimates of true structural size.

References

- Ardia, D. R. 2005. Super size me: an experimental test for the factors affecting lipid content and the ability of residual body mass to predict lipid stores in nestling European starlings. – *Funct. Ecol.* 19: 414–420.
- Badyaev, A. V. and Martin, T. E. 2000. Individual variation in growth trajectories: phenotypic and genetic correlations in ontogeny of the house finch (*Carpodacus mexicanus*). – *J. Evol. Biol.* 13: 290–301.
- Blackwell, G. L. 2002. A potential multivariate index of condition for small mammals. – *N. Z. J. Zool.* 29: 195–203.
- Green, A. J. 2001. Mass/length residuals: measures of body condition or generators of spurious results? – *Ecology* 82: 1473–1483.
- LaBarbera, M. 1989. Analyzing body size as a factor in ecology and evolution. – *Annu. Rev. Ecol. Evol. Syst.* 20: 97–117.
- Lleonart, J. et al. 2000. Removing allometric effects of body size in morphological analysis. – *J. Theor. Biol.* 205: 85–93.
- Schulte-Hostedde et al. 2001. Evaluating body condition in small mammals. – *Can. J. Zool.* 79: 1021–1029.
- Schulte-Hostedde, A. I. et al. 2005. Restitution of mass-size residuals: validating body condition indices. – *Ecology* 86: 155–163.
- Shea, B. T. 1985. Bivariate and multivariate growth allometry: statistical and biological considerations. – *J. Zool.* 206: 367–390.
- Sokal, R. R. and Rohlf, F. J. 1995. *Biometry*. – W. H. Freeman.
- Yezerinac, S. M. et al. 1992. Measurement error and morphometric studies: statistical power and observer experience. – *Syst. Biol.* 41: 471–482.

Appendix 4

An example applying the scaled mass index to the effect of a change in habitat quality on body condition

We used data taken from Cade et al. (2008) on the body mass and body length of a piscivorous fish (the walleye *Stizostedion vitreum*, $n = 6857$) before (years 1980–1988) and after (1989–2004) the introduction of a primary prey species (the alewife *Alosa pseudoharengus*) in a lake in the USA.

M is the observed body mass and L is the observed body length of walleye. For walleye studied prior to the introduction of alewife (1980–88) ($n = 4789$), the characteristics of the scaled mass index \hat{M} (Eq. 2) are as follows:

$b_{SMA} = 3.304$, 95% CI [3.289, 3.319], $L_0 = 359.32$ mm (arithmetic mean of L). Mean $\hat{M} = 421.02$ g, 95% CI [419.65, 422.39].

For walleye studied after the introduction of alewife (1989–2004) ($n = 2067$), $b_{SMA} = 3.328$ [3.312, 3.345]. There was no significant difference in the value of b_{SMA} for the two time periods ($t_{0.05[6854]} = 1.895$, $p > 0.05$). Thus, the values for the initial study without alewife ($b_{SMA} = 3.304$, $L_0 = 359.32$ mm) can be applied in Eq. 2 to calculate the scaled mass index of walleye after alewife introduction. This gives mean $\hat{M} = 434.64$ g [432.64, 436.63]. The Scaled mass index values are significantly higher after alewife introduction than before ($t_{0.05[6854]} = -12.960$, $p < 0.001$), showing that body condition increased in response to availability of the new prey (Fig. 4-1).

Exactly the same result would be obtained if we took the specific slope value for the 1989–2004 period ($b_{SMA} = 3.328$) and L_0 provided for the 1980–1988 study period ($L_0 = 359.32$ mm). This would now give values of $\hat{M} = 433.03$ g [431.02, 435.03] for 1989–2004, which is also significantly higher than for the 1980–1988 period ($t_{0.05[6854]} = -11.660$, $p < 0.001$). Another way to obtain the same result would be to use both the b_{SMA} and L_0 calculated for the whole dataset ($b_{SMA} = 3.335$ [3.324, 3.346], $L_0 = 381.90$ mm). This value for b_{SMA} might represent a reliable ‘historical’ slope for this particular fish species. This would now give values of $\hat{M} = 513.18$ g [511.49, 514.87] for 1980–88 and $\hat{M} = 530.12$ g [527.65, 532.58] for 1989–2004 ($t_{0.05[6854]} = -10.932$, $p < 0.001$). Finally, the same conclusion would be obtained by using the same historical slope ($b_{SMA} = 3.335$ [3.324, 3.346]) and any other arbitrary L_0 value favoured by biologists for walleyes (e.g. $L_0 = 370$ mm, the median body length for the whole dataset). This would now give values of $\hat{M} = 461.76$ g [460.24, 463.28] for 1980–1988 and $\hat{M} = 477.00$ g [474.79, 479.22] for 1989–2004 ($t_{0.05[6854]} = -10.932$, $p < 0.001$). Note that different means and confidence intervals between assays are due to the different body length (L_0 values) for which the scaled mass index is calculated.

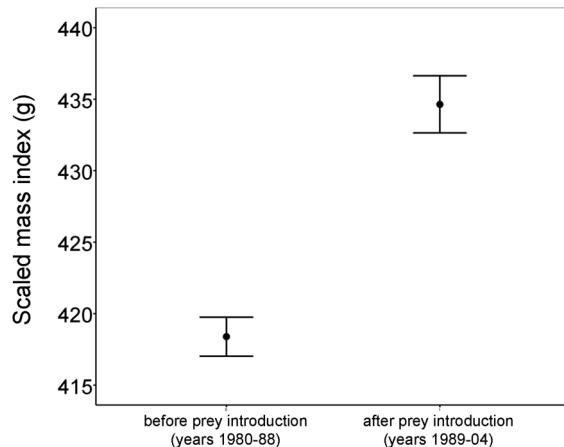


Figure 4-1. Change in body condition (scaled mass index) of walleye in response to introduction of a prey species. The data shown represent walleye body mass standardized to a length of 359.32 mm (the mean length for the 1980–1988 period).

Cade et al. (2008) used a different method (quantile regression) to study the changes in condition, and also concluded the walleye’s body condition improved following the introduction of the prey species. Like quantile regression and the scaled mass index, OLS residuals computed on the whole dataset suggest a change in condition after alewife introduction ($t_{0.05[6854]} = -91.491$, $p < 0.001$), but the residual values themselves are not as easily interpreted as the above scaled mass indices. Furthermore, the use of OLS residuals for inter-study comparisons is fraught with difficulties. For instance, if residuals are calculated separately for the periods with and without alewife, the mean residuals are necessarily zero for each period, making comparisons impossible. Furthermore, the OLS slopes of $\log M$ against $\log L$ are significantly different for the two study periods (1980–1988 $b_{OLS} = 3.260$ [3.245, 3.275], 1989–2004 $b_{OLS} = 3.307$ [3.291, 3.325], $t_{0.05[6854]} = 3.649$, $p < 0.001$), meaning that the slope value for one particular study cannot be applied to the other one to compare condition scores. For example, if the regression line estimated from the previous dataset (i.e. using the b_{OLS} value and intercept for 1980–1988) is used to recalculate OLS residuals for the 1989–2004 period, these residuals correlate with body length ($n = 2067$, $r = 0.27$, $p < 0.001$). Because the OLS residual method is an enclosed analysis, residuals are always likely to correlate significantly with body length when calculated using parameters estimated from previous datasets, even when there is relative homogeneity of slopes between studies. This indicates that OLS methods are inappropriate for these analyses, since they rely on the key assumption that the condition index (OLS residuals) is independent of X (L).

References

- Cade et al. 2008. Estimating fish body condition with quantile regression.
– N. Am. J. Fish Manage. 28: 349–359.

Appendix 5

The relationship between the OLS residual and scaled mass indices of condition

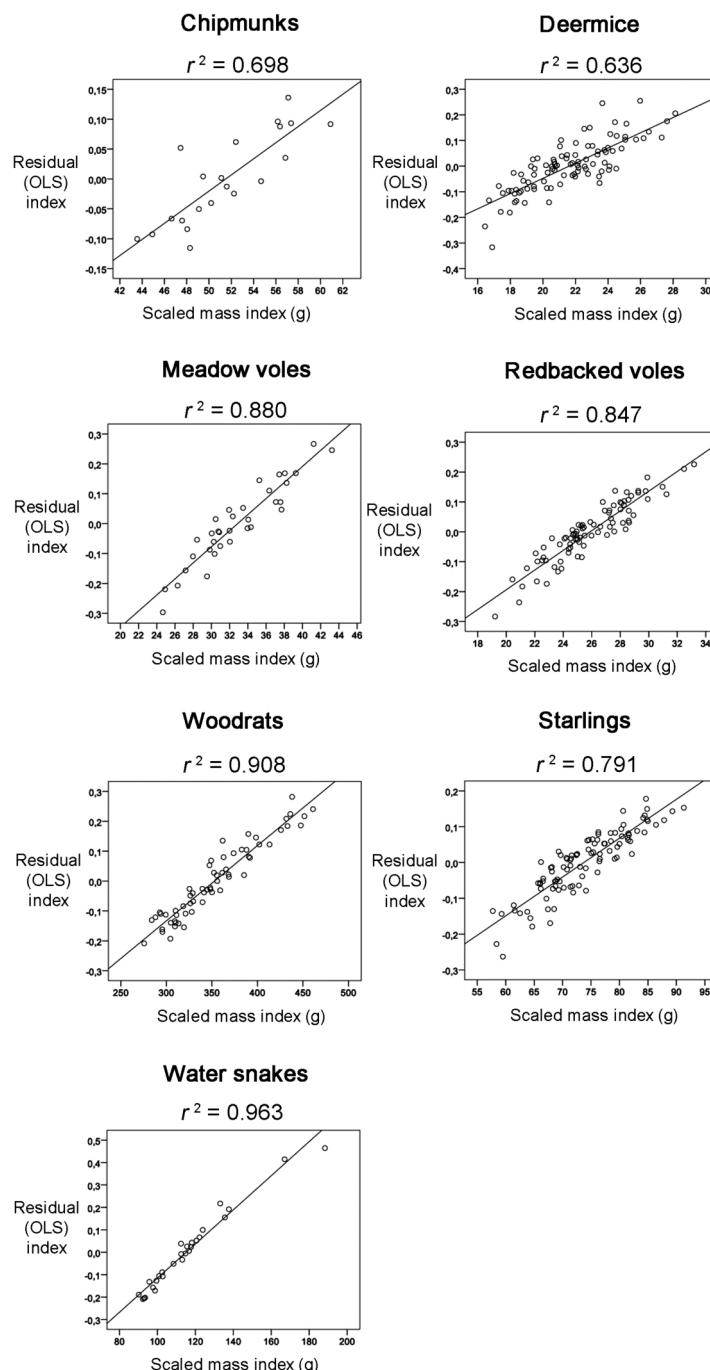


Figure 5-1. The variation (r^2) explained by the linear relationship between the Residual index (from an OLS regression of $\ln M$ on $\ln L$) and the scaled mass index of body condition for seven animal species. Both indices used the same linear body measurement (i.e. L , body length for small mammals, head-bill length for starlings, and snout-vent length for water snakes).

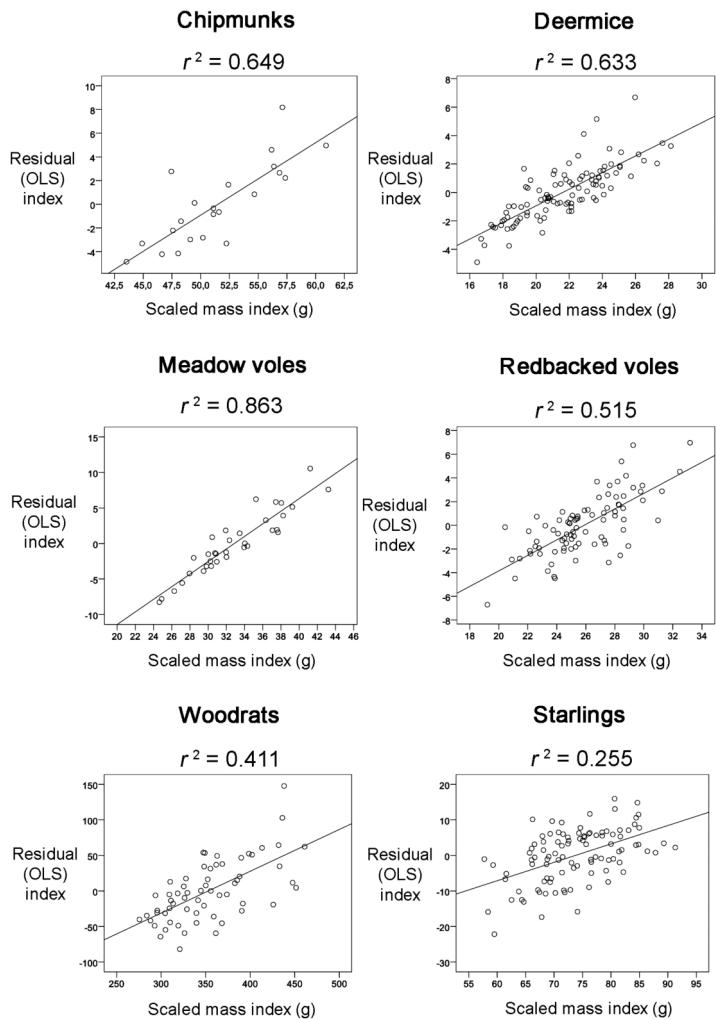


Figure 5-2. The variation (r^2) explained by the linear relationship between the original OLS residual index proposed by Shulte-Hostedde et al. (2001, 2005) and Ardia (2005) as a reliable estimate of true condition, and those we computed by the scaled mass index (Table 3). The residual index used by the above-cited authors was calculated by OLS regression of body mass against single linear body measurements or those combined by principal component analyses (X), as follows: In chipmunks, the first principal component (PC1) of a PCA performed on log-body length, log-skull width and log-skull length; in deermice, meadow voles and redbacked voles, the PC1 of a PCA on log-body length, log-foot length and log-ear length; in woodrats, the PC1 of a PCA on log-body length, log-skull length and log-ear length; and in starlings, tarsus length. The scaled mass index was computed using body length for small mammals and head-bill length for starlings.

References

- Ardia, D. R. 2005. Super size me: an experimental test for the factors affecting lipid content and the ability of residual body mass to predict lipid stores in nestling European starlings. – *Funct. Ecol.* 19: 414–420.
- Schulte-Hostedde et al. 2001. Evaluating body condition in small mammals. – *Can. J. Zool.* 79: 1021–1029.
- Schulte-Hostedde, A. I. et al. 2005. Restitution of mass-size residuals: validating body condition indices. – *Ecology* 86: 155–163.