



Consequences of metabolic scaling and log-scale allometry on means, variances and parameters from type I and type II linear regression models

M. S. DHANOA¹, R. SANDERSON², S. LOPEZ³, E. KEBREAB⁴ AND J. FRANCE^{5*}

¹Rothamsted Research, North Wyke, Okehampton, EX20 2SB, UK

²Institute of Biological, Environmental and Rural Sciences, Aberystwyth University, Gogerddan, Aberystwyth, SY23 3EB, UK

³Instituto de Ganadería de Montaña (IGM) CSIC-ULE
Departamento de Producción Animal, Universidad de León, E-24071 León, Spain

⁴Agricultural Sustainability Institute, Department of Animal Science,
University of California, Davis, CA 95616, USA

⁵Centre for Nutrition Modelling, Department of Animal Biosciences,
University of Guelph, Guelph ON, N1G 2W1, Canada

*jfrance@uoguelph.ca

Date of receipt: 29. 01. 2016

Date of acceptance: 10. 06. 2016

ABSTRACT

The slope bias when the predictor variable suffers from measurement errors is investigated. The presence of measurement errors can undermine the least squares linear regression parameter estimates, which in turn will have consequences if slope-based meaningful functions are calculated and used. Methods to determine suitable regression model choice are outlined. Also, the consequences of data size shrinkage due to scaling by metabolic weight in energy balance studies are illustrated. A problem arises when the assumed value of the metabolic index (b) changes. In the literature, this index varies from 0.62 to 0.75 for calculation of metabolic weight (MW) from live weight (LW) *i.e.* $MW=LW^b$. The estimates of regression parameters vary according to the assumed value of the metabolic index b and that will impact further on intercept and slope based calculations. Similar problems occur when allometry functions are linearized using logarithmic transformation. Disproportional shrinkage of data size introduces scale bias which can introduce inaccuracies in further use of the regression parameters. Both of these issues have potential difficulties when using databases where data size is unevenly distributed.

Key words: Allometry, energy balance, measurement errors, metabolic scaling, scaling index, type I model, type II model

INTRODUCTION

Measurement errors in the predictor variable lead to slope ($\hat{\beta}$) attenuation or negative slope bias when using linear type I or ordinary least squares regression (OLS). This bias is proportional to simple correlation ($1/r_{xy}$). Mitigation of this problem is achieved by using a type II regression instead, which

accounts for these errors in the x -variable. With assumptions about the measurement error variances, two options of type II model are relevant: (1) major axis regression (MA) when error variances of the y - and x -variables are deemed to be equal; and (2) reduced major axis regression (RMA) when the

error variances of the y - and x -variables are proportional to their underlying true variances. RMA is related to the usual “ y -on- x and x -on- y ” regression lines. Application of type II regression models is demonstrated for metabolic scaling of data and estimation of allometric relationships.

The power model, $y = ax^b$, is the most commonly used and debated model in allometry studies. Historically, the allometry function is linearized using logarithmic transformation. Justification for this transformation tends to be the equalisation of variance and also allometry index estimation is much simpler on the log-scale using linear regression. However, parameter estimation of allometric functions on the log-scale has the additional problem of unequal shrinkage of data size, which in turn introduces bias in relation to the original scale. To alleviate this problem, Guest (1961) recommended use of weighted linear regression as mentioned below. To accommodate changes in the variance of the predictor variable, type II regression methods are needed. The objectives of this study were to analyse changes in the variables when scaling factors or logarithmic transformations are introduced and the influence of the type of regression used on the estimation of biologically meaningful parameters from scaled or log-transformed variables. The analysis is based on two case studies (energy balance and allometry) with fundamental relevance in the field of animal nutrition and metabolism.

MATERIALS AND METHODS

Type II regression models

The type II model is also known as Deming regression (Deming, 1943) with the measurement error variance ratio $\lambda_{\text{Deming}} = 1/\lambda_{\text{ML}}$ (see below). This model is necessary in cases where meaningful interpretation of regression coefficients and their functions is required. Type II regression allows incorporation of predictor measurement errors to

estimate the maximum likelihood (ML) slope (β_{ML}) with the model defined (Deming, 1943; Kendal and Stuart, 1966; Madansky, 1959; Dhanoa, *et al.*, 2010a) as:

$$\hat{\beta}_{\text{ML}} = \frac{\hat{\sigma}_y^2 - \lambda_{\text{ML}} \hat{\sigma}_x^2 + \sqrt{(\hat{\sigma}_y^2 - \lambda_{\text{ML}} \hat{\sigma}_x^2)^2 + 4\lambda_{\text{ML}} \hat{\sigma}_{xy}^2}}{2\hat{\sigma}_{xy}}$$

where $\lambda_{\text{ML}} = \hat{\sigma}_e^2 / \hat{\sigma}_\delta^2$ *i.e.* the ratio of the measurement error variances of the y - and x -variables, $\hat{\sigma}_x^2$ and $\hat{\sigma}_y^2$ are the sample variances of the x - and y -variables, respectively, and $\hat{\sigma}_{xy}^2$ is the sample covariance between the x - and y -variables. Mandel (1964) showed that the above slope estimate is related to the ordinary least squares (OLS) slope estimate as:

$$\hat{\beta}_{\text{ML}} = \frac{\hat{\beta}_{\text{OLS};y,x}}{1 - (\hat{\sigma}_\delta^2 / s_x^2)}$$

Snedecor and Cochran (1980) give this formula in the alternative form:

$$\hat{\beta}_{\text{ML}} = \hat{\beta}_{\text{OLS};y,x} \frac{\hat{s}_x^2}{\hat{s}_x^2 - \hat{\sigma}_\delta^2}$$

For the special case MA, set $\lambda_{\text{ML}}=1$ in the above maximum likelihood (ML) equation. For alternative forms of the major axis slope, β_{MA} , see Dhanoa *et al.* (2010a).

In the ML equation it is assumed that the variance of the data values is constant. In cases where this assumption is not tenable, Ripley and Thompson (1987) proposed a reiterated weighted procedure referred to as functional relationship estimation by maximum likelihood (FREML; Royal Society of Chemistry, 2002).

Metabolic scaling of data

In animal energy balance studies, it is customary to scale data with metabolic weight (MW) where $\text{MW} = \text{LW}^b$ (LW denotes live weight and b metabolic index). Unfortunately such scaling results in nonlinear effects on data value size and variance (Dhanoa *et al.*, 2015) and has consequences for regression parameter estimates. OLS regression may no longer be appropriate.

Allometry on the log-scale

The general allometry function $y = ax^b$ (Huxley, 1924), where a is the scaling constant and b the metabolic index, is usually expressed as a linear function on the log-scale:

$$\log(y_i) = \log(a) + b\log(x_i) + \varepsilon_i$$

Here, subscript i denotes the i th item of samples x and y , and ε_i are assumed to be normally distributed additive errors. This equation holds for any base of logarithm but in this paper \log_e is assumed. In the absence of measurement error estimates OLS, MA or RMA regression models are used to derive estimates of the metabolic index b . In the case of OLS regression, slope (and hence the intercept) are biased estimates which may be corrected as explained above. Also in the case of OLS regression, unequal shrinkage of small to large values causes scale mismatch which requires that weights are necessary to approximate nonlinear estimates. These weights (w_i) were proposed by Guest (1961) and were shown to be $w_i = w_i y_i^2$ where $w_i = 1/\text{var}(y_i)$. On the importance of weights, Guest (1961) stated 'even if the original observations were all of equal weight, the transformed observations must be weighted, for otherwise the estimates will be inefficient and the standard deviation calculations completely incorrect'.

RESULTS

Metabolic scaling effects on energy balance analysis

In the particular case of energy balance studies, a further complication arises from scaling the energy components (*i.e.* energy retention (ER) and metabolisable energy intake (MEI)) by the metabolic weight (LW^b) of the animals. Due to lack of consensus, various values for the metabolic index (b), ranging from 0.62 to 0.75 have been used. These values result in respective metabolic weights that follow a nonlinear trend over the range of b for both the mean and variance (Fig.1). Rate of metabolic-weight-value shrinkage from $b = 0.75$ down to $b = 0.62$ is faster for variance than for the mean.

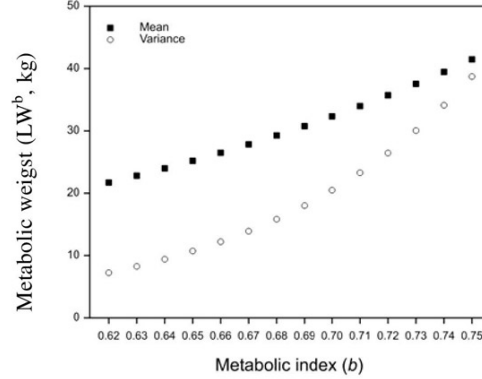


Fig. 1. Data from a calorimeter study using thirty-one growing steers with mean live weight of 142.4 kg (Sanderson *et al.*, 1995)

One of the main purposes of energy balance studies is to derive estimates of efficiency of energy utilisation for growth (k_g or k_j) and metabolisable energy requirement for maintenance (M_m). Mostly, the type I model or OLS regression is used to fit a linear model to relate ER and MEI:

$$ER = (MEI - M_m) \times k_g = k_g \times MEI - k_g \times M_m$$

so that k_g and $k_g \times M_m$ are the slope and the intercept, respectively, of the model. In OLS, metabolic scaling effects on MEI variance will not be accounted for as predictor variables are assumed error free which results in attenuation of the least-squares slope estimate (Dhanao *et al.*, 2000; 2007; 2010b). However, with the type II regression model, predictor measurement errors are modelled and slope attenuation is avoided (McArdle, 1988; Dhanao and Sanderson, 2010).

Uncertainty from both ER or MEI and LW^b gets propagated into the ratios ER/LW^b and MEI/LW^b according to the formula (Ku, 1966; Wikipedia, 2016):

$$\sigma_{P/Q}^2 \approx (P/Q)^2 \left[(\sigma_P/P)^2 + (\sigma_Q/Q)^2 - 2 \frac{\sigma_{PQ}}{PQ} \right]$$

where $P = ER$ (or MEI) and $Q = LW^b$ and

$$\sigma_Q^2 = (LW^b)^2 b \left(\frac{\sigma_{LW}}{LW} \right)^2$$

Decreasing values of b cause shrinkage of the LW^b scale which in turn causes reverse shrinkage of ER and MEI data scaled by LW^b (Table 1).

Table 1. Influence of metabolic scaling index (b) on mean scaled daily ER and MEI (Sanderson *et al.*, 1995) and their correlation (r)

b	ER (MJ/kg LW ^{b})		MEI (MJ/kg LW ^{b})		$r_{ER,MEI}$
	Mean	SD	Mean	SD	
0.62	0.1463	0.12553	1.2937	0.20631	0.6970
0.63	0.1392	0.11949	1.2310	0.19501	0.7016
0.64	0.1325	0.11375	1.1714	0.18434	0.7063
0.65	0.1261	0.10828	1.1146	0.17428	0.7109
0.66	0.1200	0.10308	1.0606	0.16478	0.7156
0.67	0.1142	0.09813	1.0092	0.15582	0.7203
0.68	0.1087	0.09341	0.9603	0.14736	0.7250
0.69	0.1035	0.08893	0.9138	0.13937	0.7298
0.70	0.0985	0.08466	0.8695	0.13183	0.7345
0.71	0.0938	0.08059	0.8274	0.12471	0.7392
0.72	0.0892	0.07673	0.7873	0.11800	0.7440
0.73	0.0849	0.07304	0.7492	0.11166	0.7487
0.74	0.0808	0.06954	0.7129	0.10567	0.7535
0.75	0.0769	0.06621	0.6784	0.10002	0.7582
1.00	0.0224	0.01945	0.1964	0.02684	0.8577

Table 2. Influence of metabolic scaling index (b) on k_g and M_m (MJ/kg LW ^{b})

b	Type I		Type II					
	OLS		FREML		MA		RMA	
	k_g	M_m	k_g	M_m	k_g	M_m	k_g	M_m
0.62	0.4241	0.9489	0.5268	0.9884	0.5030	1.0030	0.6085	1.0534
0.63	0.4299	0.9073	0.5329	0.9439	0.5096	0.9579	0.6128	1.0039
0.64	0.4358	0.8674	0.5391	0.9015	0.5162	0.9147	0.6170	0.9567
0.65	0.4417	0.8291	0.5453	0.8608	0.5228	0.8734	0.6213	0.9116
0.66	0.4476	0.7925	0.5516	0.8220	0.5293	0.8338	0.6255	0.8687
0.67	0.4536	0.7574	0.5580	0.7849	0.5359	0.7960	0.6298	0.8278
0.68	0.4596	0.7237	0.5644	0.7494	0.5424	0.7598	0.6339	0.7888
0.69	0.4656	0.6915	0.5708	0.7155	0.5488	0.7252	0.6381	0.7516
0.70	0.4717	0.6607	0.5773	0.6831	0.5552	0.6921	0.6422	0.7161
0.71	0.4777	0.6311	0.5838	0.6522	0.5616	0.6604	0.6462	0.6823
0.72	0.4838	0.6028	0.5904	0.6226	0.5679	0.6302	0.6502	0.6501
0.73	0.4898	0.5758	0.5970	0.5943	0.5741	0.6012	0.6542	0.6193
0.74	0.4958	0.5498	0.6036	0.5673	0.5802	0.5735	0.6581	0.5900
0.75	0.5019	0.5250	0.6103	0.5415	0.5862	0.5471	0.6619	0.5621
1.00	0.6215	0.1603	0.7656	0.1667	0.6884	0.1638	0.7246	0.1654

Any impact of the b scaling on regression parameters and their functions such as k_g and M_m are illustrated with energy balance data recorded in respiration chambers over two days (Sanderson *et al.*, 1995). Four linear regression models are relevant to energy balance data analysis, *viz.*

- i. Ordinary least squares (type I model; OLS)
- ii. FREML (type II model)
- iii. Major axis (type II model; MA)
- iv. Reduced major axis (type II model; RMA)

M_m (the MEI required for maintenance) was calculated for each model as {−intercept/slope}. Each of the b index values gives different scale to the resulting MW and also leads to different scaled ER and MEI values. To get comparable data we calculated daily metabolisable energy required for maintenance of an animal with the mean LW for each b index value using M_m from the above four regression models. These models give different estimates depending on how they deal with changes of scale and variances (Table 2).

The allometry power function on the log-scale

Empirical evidence for log-transformation: the Horvitz rule

When variance of the response variable y increases as its mean or expected value $E(y)$ increases, *i.e.*

$$\text{var } y_i \propto \{E(y_i)\}^{2(1-\lambda_{bc})}$$

then Taylor series analysis shows that variance is approximately stabilised (Atkinson, 2003) by using response as:

$$\begin{array}{ll} y_i^{\lambda_{bc}} & \lambda_{bc} \neq 0 \\ \log(y_i) & \lambda_{bc} = 0 \end{array}$$

For a Box-Cox lambda (λ_{BC})=1 no transformation is needed, λ_{BC} =0 indicates log-transformation, λ_{BC} =0.5 will be square root transformation whilst other non-zero λ_{BC} values may point to Box-Cox system of transformations (Box and Cox, 1964).

The above expression can be written as:

$$\log(s_y) = \gamma_0 + (1 - \lambda_{bc}) [E(y_i)]$$

where s_y is the standard deviation of y_i if replicate values of y are available and γ_0 is the intercept. Throughout this paper \log_e is adopted but the equations can be used with any log-base using the relationship $\log_e X / \log_e NN = \log_{NN} X$, where NN is any logarithmic base *e.g.* 10, 2 etc. and X is any data. In the absence of replicates, data may be compressed into smaller size by clustering into groups of increasing mean values.

Data size shrinkage when transforming to the log-scale

To demonstrate log-transformation shrinkage effects, we have taken unadjusted mammal species metabolic rate (MR; ml O₂ h^{−1}) and mass (g) data ($N = 469$) from the electronic appendix to the article by White *et al.* (2006) on the scaling and temperature dependence of vertebrate metabolism. Many of these data are from relatively smaller size species and that makes it unsuitable for nonlinear fitting of the allometry power function.

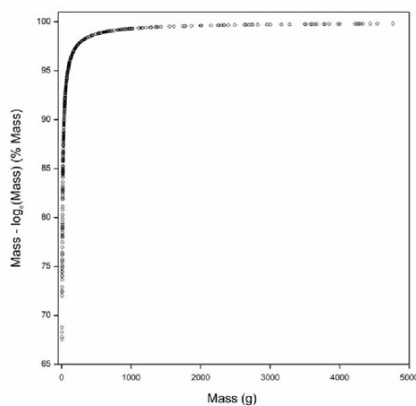
Skew data distribution does not always imply a log normal distribution. An overlapping mixture of normal distributions can cause skewness and multimodality. For example, Box-Cox transformation of the mammal species data (both MR and mass) gave $\lambda_{BC} = -0.2$ instead of 0 which would indicate a log normal distribution (only amphibians gave $\lambda_{BC} = 0$). Therefore, a multimodal distribution is indicated. Given no species misclassification, the distribution mixture is most likely due to lack of data over the range where these two components overlap. Further analysis of these data on a log-scale showed a double-normal model fitted well with either common variance or separate variance (Table 3). These results may be used to partition data into two component distributions before analysis. However, for a single species one expects a common allometric index. More data are required to correct distribution fragmentation in the White *et al.* (2006) databases.

Table 3. Parameters of the fitted double normal distribution with component means (Mean 1; Mean 2) with standard deviations (SD 1; SD 2) respectively and the proportion of the larger mean component (p)

		λ_{BC}	Mean 1	Mean 2	SD 1	SD 2	p
Mammals	MR	-0.2	6.793	4.213	0.9583	0.9583	0.2468
	Mass	-0.2	6.644	3.584	1.8112	0.8906	0.4908
Birds	MR	-0.5	5.630	3.820	0.8045	0.4235	0.3254
	Mass	-0.3	6.038	3.066	0.8453	0.8453	0.2635
Reptiles	MR	-0.1	3.481	0.290	1.3378	1.3378	0.4019
	Mass	-0.1	5.090	2.335	1.2145	1.2145	0.4399
Fish	MR	0.1	1.894	-2.058	1.4050	1.4050	0.8292
	Mass	0.1	5.186	1.093	1.4777	1.4777	0.7444
Amphibians	MR	0.0	0.422	-1.586	1.1540	0.2818	0.7043
	Mass	0.0	2.660	0.048	1.6440	0.2625	0.8694

Here, we will concentrate on the problem of data size shrinkage and its consequences when analysis is performed on the log-scale. Using mammal mass data on the log-scale as percentage of the original mass observations, the differential size reduction can be seen in Fig. 2. It is evident that larger size observations are shrunk to a greater extent relative to the smaller size observations, leading to changes in the variance. The effect on log-scale variance change is accommodated by type II regression models, *e.g.* RMA, but not by OLS.

Further, we use quantile regression (Koenker, 2005) of $\log_e(\text{MR})$ on $\log_e(\text{Mass})$ for mammals to

**Fig. 2.** Shrinkage of White *et al.* (2006) mammal mass data (≤ 5000 g) on transformation to the \log_e -scale

see if data distribution on the log-scale is free from unexpected serious stratification. For this purpose 10th, 25th, 50th, 75th and 90th percentile regression lines are selected as shown in Fig. 3. Despite observation crowding at the lower end (original scale), slopes are generally parallel from the 10th (lower line) to the 90th (upper line) percentiles. A normal distribution with mean 4.850 (± 0.0678) and standard deviation 1.468 (± 0.0480) was fitted to the log-scale data for MR. However, lack of fit as measured by deviance (distributed as χ^2) was 69.86 on 19 degrees of freedom, which was significant ($P < 0.05$). When examined further using a probability plot, some observations did not fall within the 95% confidence interval. Also, the Box-Cox λ was slightly less than zero. As shown above (Table 3), the best fit model was an overlapping double-normal with unequal variances for the two components. Such features can cause variation or bias in the slope estimates if sub-samples of these data are used for regression analysis. Quantile regression also shows how the slope estimate (allometric index) varies across the entire range of the percentiles as shown in Fig. 4. These changes in the slope estimate will cause variation in the estimates of the y-axis intercepts. If one is using regression parameter estimates for further calculations, these variations are worth considering.

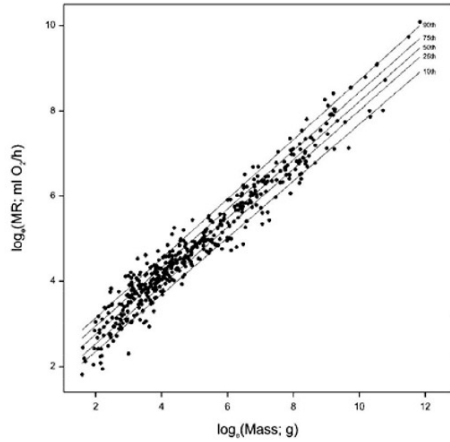


Fig. 3. Fitted linear quantile regression lines for White *et al.* (2006) mammal data at 10th, 25th, 50th, 75th and 90th percentiles

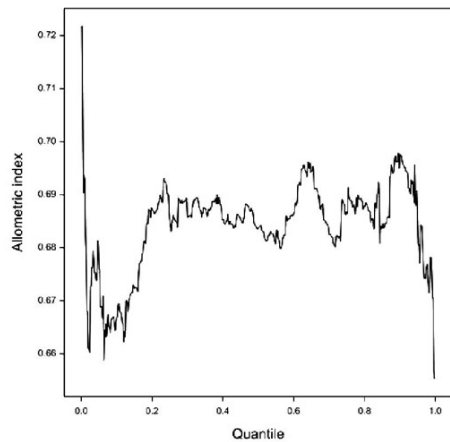


Fig. 4. Variability of the linear quantile slopes, *i.e.* allometric index, from $\log_e(\text{MR})$ versus $\log_e(\text{Mass})$ over the entire quantile range

Table 4 shows allometry index estimates using nonlinear, type I, type II and quantile regression models for each of the five classes of animals in the White *et al.* (2006) data set. In these data, there are some outliers which will undermine the weighted regression model (possibly an indirect way to detect outlier problems).

Bias correction in the log-scale allometric equation

The logarithm of the linear scale mean is not equal to the mean of logged values. This difference is equal to the gap between the geometric mean (GM) and arithmetic mean (AM) in the original observations. Anti-log of log-scale mean equals GM, whilst for AM a correction derived by Finney (1941) and illustrated by Sprugel (1983) needs to be applied. Finney (1941) showed that:

$$AM = e^{\hat{\mu} + \hat{\sigma}^2/2} = e^{\hat{\mu}} e^{\hat{\sigma}^2/2}$$

where $\hat{\mu}$ and $\hat{\sigma}^2$ are estimates of mean and variance on the log-scale respectively. Sprugel (1983) called this variance estimate SEE which is also known as mean square prediction error (MSPE), *i.e.*

$$SEE = \sqrt{\sum \{ \log(y_i) - \log(y_{i \text{ fitted}}) \}^2 / (N - 2)}$$

where $\log(y_{i \text{ fitted}})$ are the log-scale fitted values (we use \log_e here but this can be converted to any other base) and N is the sample size. The correction factor $e^{\hat{\sigma}^2/2}$ must be applied if log-scale quantities are to be transformed back to the original scale.

Allometry data compression

In the Deming, MA and RMA models, it is implicitly assumed that measurement errors are equal across the whole data range. In order to illustrate the relevant methodology for handling individual observation variances, we arbitrarily compress the example data into 40 groups on the basis of MR, temperature and mass data size using non-hierarchical cluster analysis. The mean and variance of these groups were used to calculate linear functional relationship estimates using the FREML software of Ripley and Thompson (1987) as implemented in an Excel Add-In (Royal Society of Chemistry, 2002). Thus, using means and variances of these 40 groups the fitted linear equation was:

$$\log_e(\text{MR}) = 1.4716 (\pm 0.10175) + 0.6788 (\pm 0.01828) \log_e(\text{Mass})$$

The fitted line with both y-axis and x-axis error bars is shown in Fig. 5.

Table 4. Allometry index estimates using nonlinear, type I, type II and quantile regression models from unadjusted White *et al.* (2006) data set.

Method	Mammals	Birds	Reptiles ^{@1}	Fish	Amphibians ^{@2}
Nonlinear ($y = ax^b$)	0.879 (± 0.0136)	0.724 (± 0.0248)	0.416 (± 0.1170)	0.786 (± 0.0156)	0.529 (± 0.0447) ^{@2}
Nonlinear bootstrap	0.843 (± 0.0833)	0.721 (± 0.0358)	0.518 (± 0.0525)	0.784 (± 0.0303)	0.598 (± 0.0477) ^{@2}
<i>With data transformed to natural log-scale</i>					
OLS	0.684 (± 0.0076)	0.645 (± 0.0181)	0.752 (± 0.0189)	0.826 (± 0.0094)	0.666 (± 0.0139)
OLS bootstrap	0.684 (± 0.0090)	0.644 (± 0.0159)	0.751 (± 0.0978)	0.826 (± 0.0110)	0.667 (± 0.0155)
OLS weighted	0.786 (± 0.0085)	0.658 (± 0.0163)	0.760 (± 0.0259)	0.931 (± 0.0778)	0.780 (± 0.0114) ^{@2}
RMA	0.703 (± 0.0084)	0.665 (± 0.0149)	0.858 (± 0.0191)	0.883 (± 0.0114)	0.758 (± 0.0153) ^{@2}
RMA bootstrap	0.703 (± 0.0080)	0.666 (± 0.0132)	0.857 (± 0.0173)	0.883 (± 0.0109)	0.758 (± 0.0153) ^{@2}
MA	0.697 (± 0.0088)	0.657 (± 0.0155)	0.840 (± 0.0211)	0.875 (± 0.0120)	0.730 (± 0.0170) ^{@2}
MA bootstrap	0.696 (± 0.0093)	0.654 (± 0.0169)	0.840 (± 0.0218)	0.876 (± 0.0113)	0.733 (± 0.0180) ^{@2}
Quantile linear at median	0.684 (± 0.0112)	0.641 (± 0.0211)	0.750 (± 0.0157)	0.788 (± 0.0149)	0.688 (± 0.0170) ^{@2}
Quantile linear at median weighted	0.781 (± 0.0355)	0.649 (± 0.0724)	No fit due to outliers	0.687 (± 0.0420)	No fit due to outliers
Dataset size	469	82	483	1107	681

^{@1}Some very large size outliers cause fitting problems.

^{@2}Unreliable estimates (dataset dominated by majority of small amphibians).

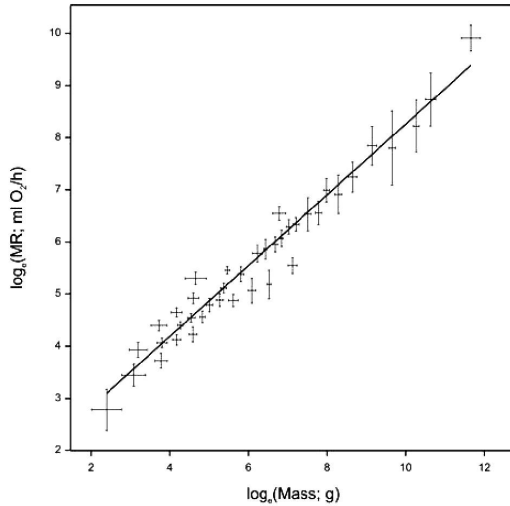


Fig. 5. Allometric relationship (solid line) between $\log_e(\text{MR})$ and $\log_e(\text{Mass})$ for White *et al.* (2006) mammal data as a ML regression with variable λ estimated using the FREML software of Ripley and Thompson (1987). Error bars represent ± 1 SD about the means of each of the 40 groups.

DISCUSSION

Energy balance analysis

Because of measurement errors, slope attenuation is largest for OLS and mitigated to different degrees by the three type II models discussed here. Scale and variance shrinkage of LW^b and energy balance components have consequences when different values of the metabolic scaling index are used. Estimates of linear regression parameters and their meaningful functions are also affected both by LW scaling and type I and type II regression models (Fig. 6). It is therefore necessary to account for these effects with an appropriate regression model when conducting literature reviews and further data analyses such as modelling and meta-analysis. Smaller differences on the analysis scale may translate to larger differences on the original scale.

Allometry analysis on the log-scale

In order to cover the whole domain of a species, a large data set is needed which can have implicit shortcomings due to data coming from

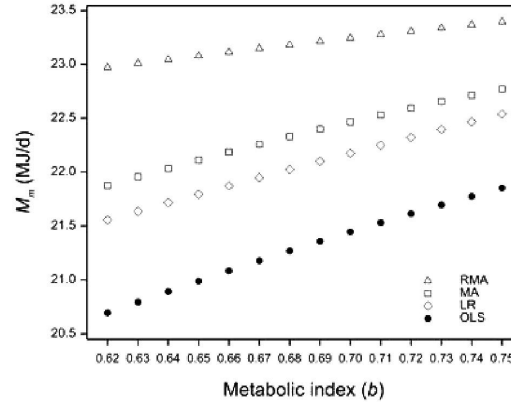


Fig. 6. Influence of metabolic scaling index (b) and regression model on estimates of metabolisable energy for maintenance of a steer with live weight of 142.4 kg.

different sources and also not covering the whole range adequately. That results in distribution fragmentation as illustrated above for our example data. We have used bootstrap (Efron and Tibshirani, 1993) to verify regression parameter estimates and their standard errors. Allometry analysis procedure is quite well established but one needs to be aware of some of the log-transformation effects such as unequal data size shrinkage and need for an appropriate regression model to deal with changing variances. Inevitably, a type II regression model will be needed to deal with consequences of log-transformation in relation to scale and variances.

CONCLUSION

In this paper we have explained and illustrated problems due to the metabolic scaling index used to calculate MW from an animal's LW which is needed to scale energy balance components, *e.g.* ER and MEI. Regression slope of ER versus MEI is used as an estimate of efficiency of energy retention for growth k_g , and the x -axis intercept as an estimate of maintenance energy M_m . A suitable regression model is needed to cope with changes in variances caused by scaling with metabolic weight at a chosen value of the metabolic index b . Similarly log-transformation to linearize the allometry equation for estimation of the metabolic index has associated problems. As illustrated above, this transformation has a disproportional effect on both data scale and variances. As in the case of energy balance data

analysis, an appropriate regression model is necessary. Furthermore, in taking regression quantities back to the original scale, bias corrections need to be carried out.

ACKNOWLEDGEMENT

Rothamsted Research is sponsored by the Biotechnology and Biological Sciences Research Council, UK. The Canada Research Chairs Program (National Science and Engineering Council, Ottawa) is thanked for part funding.

REFERENCES

- Atkinson, A.C. 2003. Horvitz' rule, transforming both sides and the design of experiments for mechanistic models. *Journal of the Royal Statistical Society, Series C (Applied Statistics)*. **52**: 261-278.
- Box, G.E.P. and Cox, D.R. 1964. An analysis of transformations (with discussion). *Journal of the Royal Statistical Society, Series B*. **26**: 211-246.
- Deming, W.E. 1943. *Statistical adjustment of data*. John Wiley and Sons, New York, 184 pp.
- Dhanoa, M.S. and Sanderson, R. 2010. Comment on 'The structural relationship: regression in biology'. *Canadian Journal of Zoology*. **88**: 821-823.
- Dhanoa, M.S., Sanderson, R. and France, J. 2001. Dependence of k_f and maintenance estimates on the choice of regression model: Model II regression. Energy metabolism of farm animals, EAAP Publication No.103 (Chwalibog, A. and Jakobsen, K. eds.), 43-46. Wageningen Pers, Wageningen.
- Dhanoa, M.S., Sanderson, R., Lopez, S., Dijkstra, J., Kebreab, E. and France, J. 2007. Alternative regression approaches when modelling energy components. Energy and protein metabolism and nutrition, EAAP publication number 124 (Ortigue-Marty, I., Miraux, N. and Brand-Williams, W. eds.), 593-594. Wageningen Academic Press, Wageningen, the Netherlands.
- Dhanoa, M.S., Sanderson, R., Lopez, S., Dijkstra, J., Kebreab, E. and France, J. 2010a. Regression procedures for relationships between random variables. Modelling nutrient digestion and utilization in farm animals (Sauvant, D., Van Milgen, J., Faverdin, P. and Friggens, N. eds.), 31-39. Wageningen Academic Publishers, Wageningen, The Netherlands.
- Dhanoa, M.S., Sanderson, R., Lopez, S., Dijkstra, J., Kebreab, E. and France, J. 2010b. Estimating maintenance energy using type I and type II regression models. 3rd EAAP International symposium on energy and protein metabolism and nutrition, EAAP publication no. 127 (Crovetto, G.M. ed.), 559-560. Wageningen Academic Publishers, Wageningen, the Netherlands.
- Dhanoa, M.S., Sanderson, R., Lopez, S., Kebreab, E. and France, J. 2015. Nonlinear effects of metabolic scaling on energy balance parameters from type I and type II linear regression models. *Canadian Journal of Animal Science*. **95**: 642-643.
- Efron, B. and Tibshirani, R.J. 1993. *An introduction to the bootstrap*. Chapman and Hall, London.
- Finney, D.J. 1941. On the distribution of a variate whose logarithm is normally distributed. Supplement to the *Journal of the Royal Statistical Society*. **7**: 155-161.
- Guest, P.G. 1961. *Numerical methods of curve fitting*. Cambridge University Press, Cambridge.
- Huxley, J.S. 1924. Constant differential growth ratios. *Nature, London*. **114**: 895-896.
- Kendall, M.G. and Stuart, A., 1966. *The advanced theory of statistics (Volume 3)*. Charles Griffin and Co. Ltd., London.
- Koenker, R. 2005. *Quantile regression*. Cambridge University Press, New York.
- Ku, H., 1966. Notes on the use of propagation error formulas. *Journal of Research of the National Bureau of Standards - C. Engineering and Instrumentation*. **70 C**: 263-273.
- Madansky, A., 1959. The fitting of straight lines when both variables are subject to error. *Journal of the American Statistical Association*. **54**: 173-205.
- Mandel, J., 1964. *The statistical analysis of experimental data*. John Wiley and Sons, New York.
- McArdle, B.H. 1988. The structural relationship: regression in biology. *Canadian Journal of Zoology* **66**: 2329-2339.
- Ripley, B.D. and Thompson, M. 1987. Regression techniques for the detection of analytical bias. *Analyst*. **112**: 377-383.
- Royal Society of Chemistry (Analytical Methods Committee). 2002. Fitting a linear functional relationship to data with error on both variables. Technical Brief Number 10.
- Sanderson, R., Dhanoa, M.S., Thomas, C. and Beever, D.E. 1995. Comparisons of calorimetry and comparative slaughter for estimating energy retention and k_f by young steers. *Animal Science*. **60**: 550.
- Snedecor, G.W. and Cochran, W.G. 1980. *Statistical methods (2nd Edition)*. Iowa State University Press, Ames, IA, USA.
- Sprugel, D.G. 1983. Correcting for bias in log-transformed allometric equations. *Ecology*. **64**: 209-210.
- White, C.R., Phillips, N.F. and Seymour, R.S. 2006. The scaling and temperature dependence of vertebrate metabolism. *Biology Letters* **2**: 125-127.
- Wikipedia. 2016. *Propagation of uncertainty*. Available from: http://en.wikipedia.org/wiki/Propagation_of_uncertainty. [26 January 2016].