JOURNAL OF Evolutionary Biology

doi: 10.1111/jeb.12986

REVIEW

Ordinary least squares regression is indicated for studies of allometry

J. T. KILMER & R. L. RODRÍGUEZ

Behavioral and Molecular Ecology Group, Department of Biological Sciences, University of Wisconsin–Milwaukee, Milwaukee, WI, USA

Keywords:

linear regression;
measurement error;
scaling relationships;
slope underestimation.

Abstract

When it comes to fitting simple allometric slopes through measurement data, evolutionary biologists have been torn between regression methods. On the one hand, there is the ordinary least squares (OLS) regression, which is commonly used across many disciplines of biology to fit lines through data, but which has a reputation for underestimating slopes when measurement error is present. On the other hand, there is the reduced major axis (RMA) regression, which is often recommended as a substitute for OLS regression in studies of allometry, but which has several weaknesses of its own. Here, we review statistical theory as it applies to evolutionary biology and studies of allometry. We point out that the concerns that arise from measurement error for OLS regression are small and straightforward to deal with, whereas RMA has several key properties that make it unfit for use in the field of allometry. The recommended approach for researchers interested in allometry is to use OLS regression on measurements taken with low (but realistically achievable) measurement error. If measurement error is unavoidable and relatively large, it is preferable to correct for slope attenuation rather than to turn to RMA regression, or to take the expected amount of attenuation into account when interpreting the data.

Introduction

Biologists must often decide how best to fit a line through data. Line-fitting methods are used in many biological disciplines, and are particularly important for fields like the study of allometry, in which the slopes of fitted lines are themselves data of interest. In the study of allometry, slopes are used to describe how a trait (often a morphological structure) scales with overall body size (Huxley, 1932). Researchers take measurements of trait size and body size across a sample of individuals, and then analyse scaling relationships on log–log scatter plots with body size on the *x*-axis and trait size on the *y*-axis. The slope (b) of the line fit through the points describes the allometric scaling of the structure. Structures may scale in direct proportion to body size ('isometry'), indicated by $b = 1$, or they may scale more steeply ('positive allometry', $b > 1$) or more shallowly ('negative allometry', $b < 1$). Allometric slopes are of interest for evolutionary biologists, because they are key for addressing hypotheses about the action of natural selection, sexual selection, and of factors such as size constraints and condition dependence (Eberhard *et al.*, 1998, 2009; Bonduriansky & Day, 2003; Egset *et al.*, 2012; Emlen *et al.*, 2012; Fromhage & Kokko, 2014; Rodríguez *et al.*, 2015).

Defining allometric relationships as the slope of the line fit through size measurement data sounds simple, but it is complicated by the availability of several different line-fitting techniques that make different assumptions and return different slopes for the same data set. This has generated considerable discussion in the evolutionary literature, with no apparent consensus regarding which method is best suited for allometric studies (Madansky, 1959; Kuhry & Marcus, 1977; McArdle, 1988; Eberhard *et al.*, 1999; Green, 1999; Smith, 2009).

Correspondence: Joseph T. Kilmer, Behavioral and Molecular Ecology Group, Department of Biological Sciences, University of Wisconsin–Milwaukee, Lapham Hall, 3209 N. Maryland Ave. Milwaukee, WI 53201, USA.
Tel.: +1 (414) 229 4214; fax: +1 414 229 3926;
e-mail: jtkilmer@uwm.edu

Here, we point out that this uncertainty arises from a misapprehension by biologists about established statistical theory, which states that there is a simple line-fitting method that is best suited for research on allometry.

Why there is confusion about methods for allometry: the problem of measurement error

The discussion of which regression method to use largely revolves around how to deal with the problem of measurement error. Measurement error is imprecision in how measurements are taken, and it results in data that do not perfectly reflect the true size of the measured structures. It can come from using imprecise equipment or from inexperience in the person taking the measurements (e.g. lack of skill in aligning and focusing on measurement landmarks). Thus, the values used to calculate slopes are not the true sizes of the structures, but rather the true size plus or minus a little bit, which comes from classical measurement error (ε). We can model measurement error as following a normal distribution with a mean of zero and a standard deviation that varies based on the precision of the measurements taken. The variables that are plugged into the slope formulae are not the true values, x and y, but rather the observed values, each with their own measurement error: X and Y, where $X = x + \varepsilon_x$ and $Y = y + \varepsilon_y$, with ε_x assumed independent of x and ε_y independent of y (Warton *et al.*, 2006). When measurement error is present in a variable, that variable's variance differs from what it would have been without any measurement error – that is $\mathrm{var}(X) > \mathrm{var}(x)$, and $\mathrm{var}(Y) > \mathrm{var}(y)$.

The two most commonly used line-fitting methods in studies of allometry make different assumptions about measurement error, and their slope estimates are influenced by measurement error in different ways (McArdle, 1988; Smith, 2009; Sokal & Rohlf, 2012). These methods are ordinary least squares (OLS) regression and reduced major axis (RMA) regression (also known as the standard major axis) (McArdle, 1988; Bonduriansky, 2007; Smith, 2009).

Ordinary least squares regression fits a line to bivariate data such that the (squared) vertical distance from each data point to the line is minimized across all data points (Fig. 1a) (Sokal & Rohlf, 2012). The slope of this line is described by the equation $b_{\mathrm{OLS}} = \mathrm{cov}(x,y)/\mathrm{var}(x)$ (Sokal & Rohlf, 2012). Therefore, OLS slopes change if there is either a change in how x and y covary or a change in the variance of the x-axis variable.

Because OLS regression fits a line using vertical residuals, it assumes that the values on the horizontal axis were measured perfectly and that any deviations of data points from the regression line are due to the variable plotted on the vertical axis (Sokal & Rohlf, 2012). There

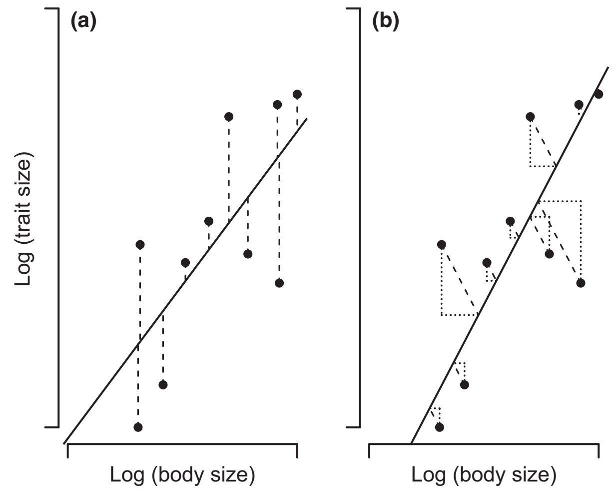

Fig. 1 Fitting a line to bivariate data using ordinary least squares (OLS) regression (a) and reduced major axis (RMA) regression (b). Both panels display the same hypothetical data set with different regression slopes (solid lines) fit through the data. Both OLS and RMA regressions fit slopes by minimizing the sum of the residuals (dashed lines), but they differ in their treatment of residuals. OLS regression (a) uses vertical residuals, whereas RMA regression (b) uses diagonal residuals that have slopes opposite to the slope of the regression line. RMA residuals can be thought of as hypotenuses of right triangles formed between each data point and the regression line such that the two shorter sides of the triangles are parallel to the axes and are proportional in length to the standard deviation of the variables plotted on their respective axes (as shown with dotted lines). The values shown here were generated in R version 3.1.3 (R Core Team, 2015). The x-axis values were sampled randomly from a specified range, and the y-axis values were generated by multiplying x-values by a specified slope and adding some scatter (representing natural variation among individuals). The scatter was randomly sampled for each point from a normal distribution with a mean of 0 and a standard deviation that was approximately 90% of the standard deviation of the y-axis variable.

are many types of studies in which the predictor variable is measured with absolute accuracy, but in allometry, the x-axis variable is a series of measurements that are naturally susceptible to some amount of error, and so this assumption of OLS regression is not met.

When measurement error (ε) is present, the formula for the OLS slope becomes: $b'_{\mathrm{OLS}} = \mathrm{cov}([x + \varepsilon_x], [y + \varepsilon_y])/\mathrm{var}(x + \varepsilon_x) = \mathrm{cov}(X, Y)/\mathrm{var}(X)$, where b' denotes a slope calculated with measurement error. The expected effect of measurement error in the numerator is neutral on average; measurement error does not consistently bias covariance in one direction or another. By contrast, measurement error will increase the denominator, var (X), leading to smaller estimates of OLS slope (Carroll *et al.*, 2006; Fuller, 2006; Sokal & Rohlf, 2012). In other words, the observed slope, b'_{OLS}, is usually lower than the slope without measurement error, b_{OLS}. If the difference between these two is large, then researchers can run into trouble if they see b'_{OLS} and assume that it equals b_{OLS} – doing so would lead them to believe that a particular scaling relationship is shallower than it truly is. The attenuating effect of measurement error on OLS slopes is widely discussed in the evolutionary literature (McArdle, 1988; Warton *et al.*, 2006; Bonduriansky, 2007; Smith, 2009). Because of it, some authors recommend that OLS regression be avoided altogether in studies of allometry (Ricker, 1973; Green, 1999) or that it be used only under a restricted range of conditions (McArdle, 1988; Bonduriansky, 2007; Forstmeier, 2011; Legendre & Legendre, 2012).

Reduced major axis regression also minimizes collective distance between data points and the line, but rather than being vertical, these distances are diagonal (Warton *et al.*, 2006); they can be thought of as hypotenuses of right triangles that have two sides parallel to the axes, each side being proportional in length to the standard deviation of the variable plotted on that axis (Fig. 1b). RMA slopes are calculated as $b_{RMA} = \mathrm{sd}(y)/\mathrm{sd}(x)$, and they can be positive or negative depending on the sign of the correlation (Sokal & Rohlf, 2012). Therefore, RMA slopes change if the variance of x or y changes.

To avoid OLS attenuation, some researchers prefer RMA regression, because it allows for measurement error in both axes (Green, 1999; Laws, 2003). Specifically, it assumes that each variable has error proportional to its standard deviation (McArdle, 1988). In the presence of measurement error, the formula for RMA slopes becomes: $b'_{RMA} = \mathrm{sd}(y + \varepsilon_y)/\mathrm{sd}(x + \varepsilon_x) = \mathrm{sd}(Y)/\mathrm{sd}(X)$. Thus, when measurement error affects each axis proportionately, an RMA slope with measurement error will be very close to what it would have been without measurement error (i.e. $b'_{RMA} \approx b_{RMA}$). This is why some authors believe that RMA slopes are more robust to measurement error than OLS slopes.

Studies on allometry vary considerably in their choice of method. For instance, we surveyed three journals that often feature allometric studies (*Biological Journal of the Linnean Society*, *Evolution*, and *Journal of Evolutionary Biology*) for the years 2010–2015. We searched for 'allometry' on *Web of Science* targeting these journals. We found 33 papers that dealt with static allometry. Of these, 42% used RMA regression either exclusively (30%) or together with OLS regression (12%), and 58% used OLS regression exclusively. If one of these methods is well suited for allometric research and the other is not (and the estimated slopes are different), this represents a problem for the field.

Why measurement error need not be a big problem for OLS slopes

Although measurement error in the x-axis attenuates OLS slopes, to conclude that OLS slopes are always unreliable descriptors of allometric relationships (as some evolutionary biologists do) is to ignore an important piece of statistical theory: the degree of OLS slope attenuation is a function of the magnitude of measurement error in X (McArdle, 2003; Carroll *et al.*, 2006; Fuller, 2006; Smith, 2009; Egset *et al.*, 2012); that is, as measurement error in X increases, attenuation gets worse, and OLS slopes are more strongly biased towards zero. Conversely, when measurement error in X is close to zero, slope attenuation is close to zero, and b'_{OLS} is close to b_{OLS}.

The relationship between measurement error and slope attenuation can be deduced from the formula for OLS slope that includes measurement error. The cause of OLS attenuation is the inflation of the denominator in the equation for b'_{OLS}, $\mathrm{var}(x + \varepsilon_x)$. If we note that $\mathrm{var}(x + \varepsilon_x) = \mathrm{var}(x) + \mathrm{var}(\varepsilon_x)$, then we can see that when $\mathrm{var}(\varepsilon_x)$ is large relative to $\mathrm{var}(x)$, the denominator is substantially larger, and b'_{OLS} is strongly biased towards zero. When, on the other hand, $\mathrm{var}(\varepsilon_x)$ is small compared to $\mathrm{var}(x)$, then $\mathrm{var}(X)$ is close to $\mathrm{var}(x)$, and an OLS slope calculated with measurement error in x will be close to what it would have been without measurement error (i.e. $b'_{OLS} \approx b_{OLS}$). This means that OLS slopes can serve as useful indicators of scaling patterns even in the presence of measurement error, as long as that error is relatively small (Fig. 2). The effect of measurement error on an OLS slope can be summarized in terms of an attenuation factor (λ), as follows: $\lambda = \mathrm{var}(x)/[\mathrm{var}(x) + \mathrm{var}(\varepsilon_x)]$, and $b'_{OLS} = \lambda * b_{OLS}$ (Carroll *et al.*, 2006). Note again that the smaller that $\mathrm{var}(\varepsilon_x)$ is, the closer b'_{OLS} will be to b_{OLS}.

Knowing that OLS slopes are robust to small levels of measurement error is helpful only if those levels are reasonably achievable in the real world. Theoretical studies often use dramatic levels of measurement error to illustrate the attenuation point (e.g. as we do in Fig. 2a). To get a sense for the levels of measurement error that are typical of empirical studies, we searched for 'static allometry' and 'measurement error' on Web of Science. We found five papers on allometry that report measurement error (Table 1). When looking at the relative magnitude of measurement error variance as a proportion of total variance in X – i.e. $\mathrm{var}(\varepsilon_x)/\mathrm{var}(X)$ – we found that very few traits exceeded 0.1, and the majority of them were below 0.05. This means that for most traits in these studies, measurement error accounted for < 5% of the total variance. At 5% measurement error, we can expect OLS slopes to be attenuated by about 5%, so a structure that scales isometrically ($b = 1$) would be described by a slope of $b' \approx 0.95$ (see Fig. 2b). Although, strictly speaking, the slope is attenuated, the bias is small enough that it is not likely to cause confusion in analysis, whether the goal is (i) assigning slopes to broad categories (e.g. isometry, negative allometry, positive allometry); (ii) statistically testing whether slopes are different from reference values (e.g. 1 or 0); or (iii) testing whether

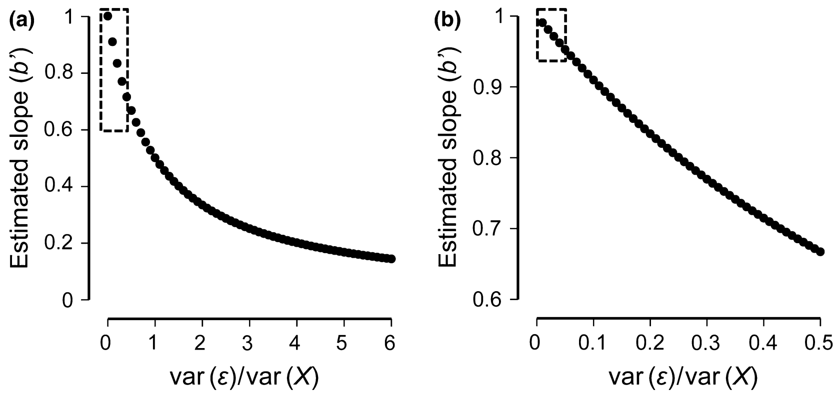

Fig. 2 Effect of measurement error in X on slope estimation by ordinary least squares (OLS) regression. We show the estimated slope (b') (given a true slope of $b = 1$) as a function of the relative magnitude of the error variance and total variance in X – i.e. var(ε_x)/var(X) (see text). Attenuation of b' calculated as $b' = b * 1/[1 + \text{var}(\varepsilon_x)/\text{var}(X)]$. With this formulation, if var(ε_x) = 0 then $b' = b = 1$. This is equivalent to the formulation of var(x)/[var(x) + var(ε_x)] in the text but makes the x-axis for this figure more straightforward. (a) As the relative magnitude of measurement error increases, b' is attenuated. Note, however, that attenuation is severe only when measurement error is very large. (b) This panel shows the area highlighted in the dashed rectangle in (a). Here, attenuation of b' is roughly in a one-to-one relationship with the relative magnitude of measurement error, and quite minor when measurement error is within reasonably achievable levels (dashed rectangle; see text).

Table 1 Empirical measurement error values reported in papers on allometry. Values are categorized as x-axis or y-axis traits, depending on how they were used in the study (x-axis for body size indicators, y-axis for traits of allometric interest). Measurement error is calculated as 1 – repeatability, which is equivalent to var(ε_x)/var(X). We show the mean relative size of measurement error among traits in the data set ± SD, followed by the number of traits.

Source	*x*-axis traits	*y*-axis traits
Al-Wathiqui & Rodríguez (2011)	0.0002 (1)	0.011 ± 0.008 (7)
Nava-Bolaños *et al.* (2012) (first data set)	0.020 ± 0.010 (8)	0.018 ± 0.015 (12)
Nava-Bolaños *et al.* (2012) (second data set)	0.039 ± 0.022 (8)	0.056 ± 0.028 (8)
Egset *et al.* (2012)	0.01 (1)	0.11 (1)
Nava-Bolaños *et al.* (2014)	0.042 ± 0.031 (6)	0.052 ± 0.016 (12)
Kilmer & Rodríguez (2015)	0.049 ± 0.078 (7)	0.050 ± 0.057 (19)

the slopes of two traits are different from each other. If, on the other hand, the purpose is precise slope estimation, it is possible to use the attenuation factor to correct for slope attenuation.

We recognize that the above is a small sample of studies and that there may be under-reporting in studies with larger relative errors. However, our sample does indicate that low relative errors that present no serious problem for OLS slope estimation are realistically achievable in empirical studies.

Checking for measurement error (especially in the x-axis measure of body size) is straightforward. This can be done by taking multiple measurements of each individual and calculating measurement repeatability, which is the proportion of total variance explained by individual identity (bound between 0 and 1). Measurement error can then be calculated as: *measurement error* = 1 – *measurement repeatability*, or the proportion of residual variance. Once measurement error in X is known, it becomes clear whether the OLS slope estimate is strongly attenuated or not. When measurement error is relatively large and unavoidable, there are methods available to correct for it (e.g. Kuhry & Marcus, 1977; McArdle, 1988; Carroll & Ruppert, 1996; Carroll *et al.*, 2006). We would, however, recommend improving the technique or the instrumentation whenever possible; otherwise, the measurements themselves cease being good descriptors of structure size, and when the error in slope estimation is very large, correction methods cease to be effective (Hansen & Bartoszek, 2012). A complementary approach may be to take the expected magnitude of attenuation into account when interpreting the results. This may be particularly useful when using complex statistical models built on OLS regression.

It is important to point out that calculating measurement error requires taking at least twice as many measurements as are needed for slope estimation. This may make error estimation prohibitive in terms of time and costs for large data sets. In some situations, it may be more practical for the researcher to first hone their measurement technique by running through practice rounds, taking multiple measurements per individual and calculating measurement error. Once measurement error is sufficiently low in the training rounds, the researcher can move on to collecting the full data set, measuring each structure only once per individual.

Why RMA regression does not fix the problem of measurement error

Besides the above argument for why OLS regression is adequate for allometric studies, there are several reasons why RMA regression is not a useful alternative. As we have seen, RMA regression does not take the covariance between traits into account for estimating the slope, so it does not describe functional scaling relationships. Some authors recommend RMA regression in cases in which there is 'codependency' between the traits involved – that is when x and y are exchangeable (e.g. Smith, 2009; Forstmeier, 2011; Legendre & Legendre, 2012). This will rarely be the case in studies of allometry, where the goal is to determine how traits scale relative to body size – it makes more sense to say that a beetle's horn is large for its body than to say that its body is small for its horn. With body size plotted on the x-axis and trait size plotted on the y-axis, variation in trait size relative to body size is best described as vertical deviations from the underlying scaling function. These vertical deviations are precisely what OLS regression models using vertical (rather than horizontal or diagonal) residuals (Fig. 1).

An additional concern is that RMA regression requires that data scatter be low and come only from measurement error that was proportional in both axes. For empirical measurement data, this is often not the case. In fact for many studies, measurement error may contribute only a small proportion of the scatter in a data set (e.g. Table 1). The rest comes from natural variation – that is, differences among individuals in precisely how big a structure is relative to body size. Individuals of the same body size are likely to have different sizes of a given trait, due to differences in environmental or genetic inputs. Consequently, even with no measurement error, an allometric data set will have some scatter from natural variation, meaning that the correlation (r) between trait size and body size will be less than perfect (i.e. $|r| < 1$). If we note that $b_{RMA} = b_{OLS}/|r|$ (Sokal & Rohlf, 2012), then it follows that for empirical measurement data sets, b_{RMA} can nearly always be expected to be higher than b_{OLS}. Some authors have interpreted such differences in slope to be the dreaded OLS attenuation (e.g. Hayes & Shonkwiler, 2001), but OLS attenuation, as we have seen, *only* comes from measurement error in the x-axis variable. Without measurement error, OLS slopes would be perfectly reliable but still shallower than RMA slopes. Therefore, these differences in slope do not reflect a weakness of OLS regression, but rather the tendency of RMA slopes to confound steepness of a relationship with dispersion of data. Studies that use RMA regression to fit slopes to data with moderate-to-high natural variation run the risk of concluding that relationships are steeper than the true underlying scaling pattern (Hansen & Bartoszek, 2012; Pélabon *et al.*, 2014). For example, negative allometry could be mistaken for isometry or even positive allometry.

It is also important to note that RMA regression, too, can misestimate slopes when its assumptions about measurement error are not met. Given that $b'_{RMA} = \text{sd}(y + \varepsilon_y)/\text{sd}(x + \varepsilon_x) = \text{sd}(Y)/\text{sd}(X)$, we can see that b'_{RMA} will only be a good estimate of b_{RMA} when the effects of measurement error are proportional between the two variables (McArdle, 1988; Smith, 2009). If the y-axis variable were measured with proportionately more error than the x-axis variable, then sd(Y) would tend to increase relative to sd(X), and b_{RMA} would be overestimated (Fig. 3). Conversely, if X were measured

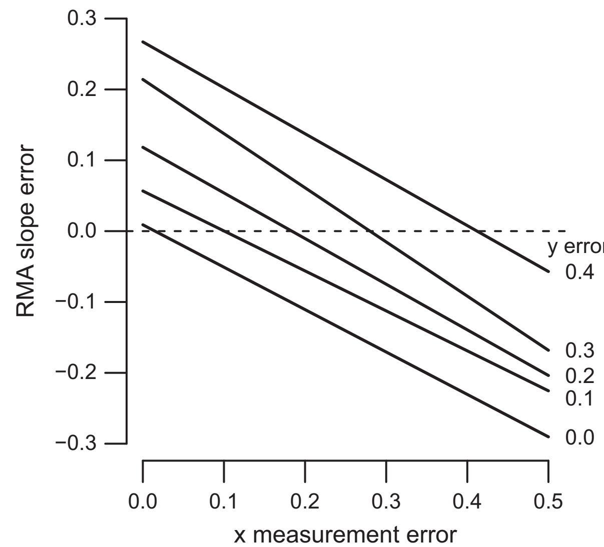

Fig. 3 Effects of measurement error on reduced major axis (RMA) slope estimation for different combinations of measurement error in X and Y. Data are the result of simulations that we ran in R (version 3.1.3; R Core Team, 2015). The simulations generated random bivariate allometric data sets and added measurement error to the variables. The exact amount of measurement error was obtained by making repeated measurements and calculating the proportion of total variance that was attributed to error, that is $\text{var}(\varepsilon_x)/\text{var}(X)$ (see text). The simulations then calculated ordinary least squares and RMA slopes for the data without measurement error (b) and the data with measurement error (b'), and then calculated proportional slope error as $(b' - b)/b$. Negative values of slope error represent underestimation (or attenuation), and positive values represent overestimation of slope. Dashed line indicates where slope error equals 0. Solid lines show the overall relationship between slope error and measurement error from a series of simulations (individual data points not shown). These lines differ in their levels of Y measurement error, as indicated by the values on the right. Because RMA slopes are calculated as sd $(y)/\text{sd}(x)$, RMA slopes are estimated with the least error when measurement error in X equals measurement error in Y. When measurement error in Y is greater than error in X, RMA slopes tend to be overestimated (positive slope error values), and when X measurement error is greater, RMA slopes tend to be underestimated (negative slope error values).

with proportionately more error than Y, then RMA slopes would be biased towards zero (Fig. 3) – ironically, this is the exact problem that researchers try to avoid using RMA regression in the first place. Thus, the use of RMA slopes does not relieve the concerns that arise from measurement error.

Another way of viewing the problem is to note that any change in the variance of either x or y will change the estimated RMA slope, regardless of whether this arises from a change in the steepness of the relationship or from a change in the scatter of the data (Al-Wathiqui & Rodríguez, 2011; Egset *et al.*, 2012; Hansen & Bartoszek, 2012; Voje & Hansen, 2012) (Fig. 4). OLS slopes, on the other hand, distinguish between the steepness of the relationship and the scatter of the data because OLS regression takes the covariance between the two variables into account (Eberhard *et al.*, 1998) (Fig. 4).

The one advantage of using RMA slopes is that as long as measurement error affects both variables proportionately, then b'_{RMA} will be close to b_{RMA} even when the overall magnitude of measurement error is high. But this does not solve the larger problem of RMA slopes overestimating relationships in the presence of natural variation.

On the (un)importance of the relative magnitude of measurement error in *X* and *Y*

Some authors recommend that the decision to use one regression method over the other be based on the relative magnitude of measurement error in X and Y. One commonly recommended rule of thumb is that the OLS method should only be used when error in Y is more than three times larger than the error in X, and that otherwise RMA should be used (McArdle, 1988; Bonduriansky, 2007; Legendre & Legendre, 2012). An alternative recommendation is to use the OLS method when the sampling variance (natural vertical deviations from the regression line) plus the measurement error variance in Y together is more than three times larger than the measurement error variance in X (McArdle, 1988; Smith, 2009). Note that the latter better fits the formulation of McArdle (1988) and that it will almost always favour OLS regression (unless measurement error in X is very large).

There are problems with recommendations that focus on relative error magnitudes in X and Y. First, they imply that OLS and RMA slopes are interchangeable –

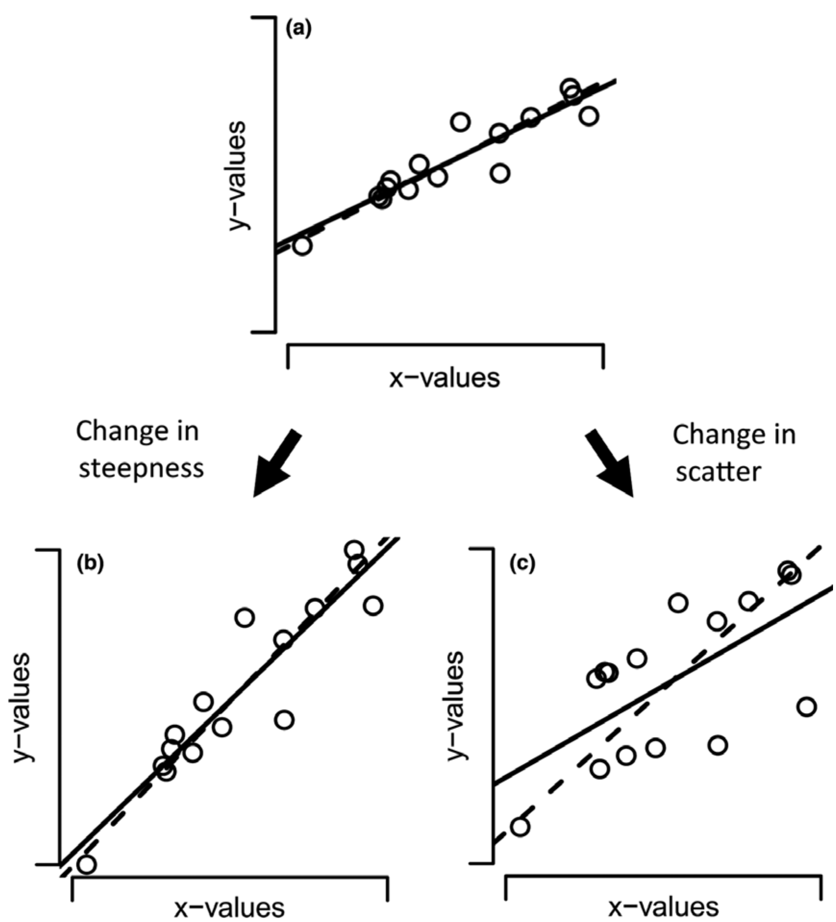

Fig. 4 How slopes estimated by ordinary least squares (OLS) and RMA regression are influenced by the steepness of the $y \sim x$ relationship and dispersion of the data. (a) OLS slope (solid line) and RMA slope (dashed line) fit through sample data with moderate slope and scatter. (b) OLS and RMA slopes fit through data with a steeper $y \sim x$ relationship, but with the same degree of scatter as (a). (c) OLS and RMA slopes fit through data with the same $y \sim x$ relationship as (a), but with greater scatter. In both (b) and (c), the RMA slope is steeper than in (a), because RMA slopes are calculated as sd(y)/sd(x), and the standard deviation of the y-values has increased in both of these cases. Thus, RMA slopes may confound scatter in a data set with the steepness of the relationship between two variables. By comparison, OLS slopes are sensitive to changes in the steepness of relationships (b), but not to changes in scatter (c), making them better suited for describing bivariate relationships.

that one should be used in one case, and the other should be used in another case for the same purpose. This is incorrect because, as explained above, the two methods describe different aspects of a data set, and cannot be compared to one another. Whereas one method describes a relationship between two variables, taking covariance into account, the other method simply describes the relative variance of the two variables.

The second problem is a focus on the wrong factor. The recommendation comes from simulations that looked at slope error as a function of the ratio of natural variation plus measurement error in *Y* and measurement error in *X* (McArdle, 1988). OLS slopes had lower slope error than RMA slopes only when the above variation in *Y* was over three times greater than in *X*. This is correct, but it overcomplicates the issue. As we have seen, the sole cause of OLS attenuation is measurement error in *X*. Measurement error in *Y* may increase the range of potential slopes that OLS calculates by altering the covariance, but it will not consistently bias slopes in a particular direction (compare panels a and b in Fig. 5). If *X* has 5% measurement error, then on average, OLS slopes will be attenuated by approximately 5%, regardless of whether measurement error in *Y* is 1%, 5% or 15%. This contrasts with the behaviour of RMA regression, where proportionately more error in *Y* overestimates the slope and proportionately more error in *X* attenuates the slope (Figs 3 and 5c,d).

Third, although the effect of ε in the numerator is neutral on average for OLS slopes, meaning that it introduces no net bias, it adds noise to the slope estimate (Fig. 5b). This highlights the advantage of keeping measurement error low across the board.

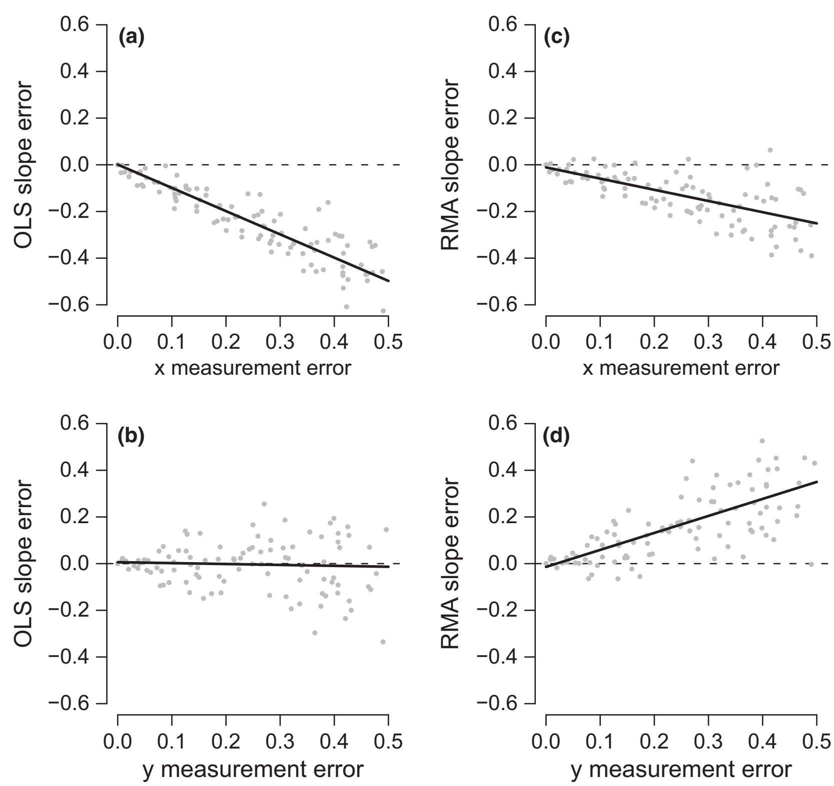

Fig. 5 Effects of measurement error on ordinary least squares (OLS) (a–b) and RMA (c–d) slope estimation. Each grey point is the result of a single simulation that we ran in R version 3.1.3 (R Core Team, 2015), following the framework of Fig. 3. Measurement error expressed as a proportion of total variance, as $\mathrm{var}(\varepsilon_x)/\mathrm{var}(X)$. Slope error calculated as $(b' - b)/b$, where b is the slope of a line fit through data without measurement error, and b' is the slope of a line fit through the same data with measurement error. Negative values of slope error represent underestimation (or attenuation), and positive values represent overestimation of slope. Dashed lines indicate where slope error equals 0. (a) Error in *Y* remains at 0, whereas error in *X* varies from 0 to 0.5. As measurement error in *X* increases, OLS slopes are increasingly biased downwards. In other words, attenuation increases in nearly direct proportion with *X* measurement error (and see Fig. 2). (b) Error in *X* remains at 0, whereas error in *Y* varies from 0 to 0.5. As measurement error in *Y* increases, the range of OLS slope error increases, but there is no net bias in either direction, as shown by the regression line. (c–d) Same as a–b, except for RMA slopes. As *X*-error increases relative to *Y*-error, RMA slopes are underestimated (c), and as *Y*-error increases relative to *X*-error, they are overestimated.

Line-fitting for different types of allometry

Our comments are broadly applicable to studies of static and developmental allometry, in which each data point is a single precise measure from one individual, and the data set is composed of measurements from a sample of adults (static allometry) or individuals spanning a range of ages (developmental allometry) from a single species. An additional problem arises in studies of evolutionary allometry, which looks at scaling relationships across different species, with each data point representing the species mean for body size and structure size. As means, these data points naturally contain uncertainty in the form of standard error – the sample mean may differ from the population mean. This discrepancy between observed values and true values is analogous to the measurement error in static or developmental allometry. In these cases, if the standard errors are not small relative the total variance across the data points, it may be necessary to correct for bias in the slope estimates (Hansen & Bartoszek, 2012) (but recall that when error is large, correction becomes ineffective; Hansen & Bartoszek, 2012). A complementary approach may be to conduct a comparative study of within-species scaling (e.g. Emlen *et al.*, 2007). One advantage of this approach is that it yields a more detailed view of the evolution of static allometry across species (for example, a line fit across species means may show isometry, but patterns within species may range from negative allometry to positive allometry).

Conclusion

Evolutionary biologists should note that statistical theory recommends the use of OLS regression over RMA regression for studies of allometry. This is because: (i) the problem of slope attenuation by OLS regression can realistically be mitigated by ensuring that measurement error is low, or by correcting for it; and (ii) OLS regression actually describes allometric scaling, whereas RMA regression does so only in a narrow range of conditions. When it comes to describing allometric relationships, RMA regression is the wrong fix to a small problem.

Acknowledgments

We thank Gerlinde Höbel, Bretta Speck, Michael Morrissey and an anonymous reviewer for their insightful comments and suggestions. Funding was provided by National Science Foundation grant IOS-1120790 to R.L.R. and K.D. Fowler-Finn.

Conflict of interest

The authors declare no conflict of interests.

References

Al-Wathiqui, N. & Rodríguez, R.L. 2011. Allometric slopes not underestimated by ordinary least squares regression: a case study with *Enchenopa* treehoppers (Hemiptera: Membracidae). *Ann. Entomol. Soc. Am.* **104**: 562–566.

Bonduriansky, R. 2007. Sexual selection and allometry: a critical reappraisal of the evidence and ideas. *Evolution* **61**: 838–849.

Bonduriansky, R. & Day, T. 2003. The evolution of static allometry in sexually selected traits. *Evolution* **57**: 2450–2458.

Carroll, R.J. & Ruppert, D. 1996. The use and misuse of orthogonal regression in linear errors-in-variables models. *Am. Stat.* **50**: 1–6.

Carroll, R.J., Ruppert, D., Stefanski, L.A. & Crainiceanu, C.M. 2006. *Measurement Error in Nonlinear Models: A Modern Perspective*, 2nd edn. Chapman & Hall/CRC, Boca Raton, FL.

Eberhard, W.G., Huber, B.A., Rodríguez, R.L., Briceño, R.D., Salas, I. & Rodriguez, V. 1998. One size fits all? Relationships between the size and degree of variation in genitalia and other body parts in twenty species of insects and spiders. *Evolution* **52**: 415–431.

Eberhard, W.G., Huber, B.A. & Rodríguez, R.L. 1999. Don't forget the biology: a reply to Green. *Evolution* **53**: 1624–1627.

Eberhard, W.G., Rodríguez, R.L. & Polihronakis, M. 2009. Pitfalls in understanding the functional significance of genital allometry. *J. Evol. Biol.* **22**: 435–445.

Egset, C.K., Hansen, T.F., Le Rouzic, A., Bolstad, G.H., Rosenqvist, G. & Pélabon, C. 2012. Artificial selection on allometry: change in elevation but not slope. *J. Evol. Biol.* **25**: 938–948.

Emlen, D.J., Lavine, C.L. & Ewen-Campen, B. 2007. On the origin and evolutionary diversification of beetle horns. *Proc. Natl. Acad. Sci. USA* **104**: 8661–8668.

Emlen, D.J., Warren, I.A., Johns, A., Dworkin, I. & Lavine, L.C. 2012. A mechanism of extreme growth and reliable signaling in sexually selected ornaments and weapons. *Science* **337**: 860–864.

Forstmeier, W. 2011. Women have relatively larger brains than men: a comment on the misuse of general linear models in the study of sexual dimorphism. *Anat. Rec.* **294**: 1856–1863.

Fromhage, L. & Kokko, H. 2014. Sexually selected traits evolve positive allometry when some matings occur irrespective of the trait. *Evolution* **68**: 1332–1338.

Fuller, W.A. 2006. *Measurement Error Models*. John Wiley & Sons Inc., Hoboken, NJ.

Green, A.J. 1999. Allometry of genitalia in insects and spiders: one size does not fit all. *Evolution* **53**: 1621–1624.

Hansen, T.F. & Bartoszek, K. 2012. Interpreting the evolutionary regression: the interplay between observational and biological errors in phylogenetic comparative studies. *Syst. Biol.* **61**: 413–425.

Hayes, J.P. & Shonkwiler, J.S. 2001. Morphological indicators of body condition: useful or wishful thinking? In: *Body Composition Analysis of Animals: A Handbook of Non-Destructive Methods* (J.R. Speakman, ed.), pp. 8–38. Cambridge University Press, Cambridge.

Huxley, J. 1932. *Problems of Relative Growth*. Methuen & Co., Ltd., London.

Kilmer, J.T. & Rodríguez, R.L. 2015. Do structures with sexual contact functions evolve negative static allometries? A case study with the harvestman *Leiobunum vittatum* (Opiliones Sclerosomatidae). *Ethol. Ecol. Evol*. doi:10.1080/03949370.2015.1087432.
Kuhry, B. & Marcus, L.F. 1977. Bivariate linear models in biometry. *Syst. Zool.* **26**: 201–209.
Laws, E.A. 2003. Mesozooplankton grazing and primary production: an alternative assessment. *Limnol. Oceanogr.* **48**: 1357–1359.
Legendre, P. & Legendre, L. 2012. *Numerical Ecology*, 3rd English edn. Elsevier, Oxford.
Madansky, A. 1959. The fitting of straight lines when both variables are subject to error. *J. Am. Stat. Assoc.* **54**: 173–205.
McArdle, B.H. 1988. The structural relationship: regression in biology. *Can. J. Zool.* **66**: 2329–2339.
McArdle, B.H. 2003. Lines, models, and errors: regression in the field. *Limnol. Oceanogr.* **48**: 1363–1366.
Nava-Bolaños, A., Córdoba-Aguilar, A. & Munguía-Steyer, R. 2012. A test of genital allometry using two damselfly species does not produce hypoallometric patterns. *Ethology* **118**: 203–213.
Nava-Bolaños, A., Sánchez-Guillén, R.A., Munguía-Steyer, R. & Córdoba-Aguilar, A. 2014. Isometric patterns for male genital allometry in four damselfly species. *Acta Ethol.* **17**: 47–52.
Pélabon, C., Firmat, C., Bolstad, G.H., Voje, K.L., Houle, D., Cassara, J. *et al.* 2014. Evolution of morphological allometry. *Ann. N. Y. Acad. Sci.* **1320**: 58–75.
R Core Team. 2015. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna. http://www.R-project.org/.
Ricker, W.E. 1973. Linear regressions in fishery research. *J. Fish. Res. Board Can.* **30**: 409–434.
Rodríguez, R.L., Araya-Salas, M., Gray, D.A., Reichert, M.S., Symes, L.B., Wilkins, M.R. *et al.* 2015. How acoustic signals scale with individual body size: common trends across diverse taxa. *Behav. Ecol.* **26**: 168–177.
Smith, R.J. 2009. Use and misuse of the reduced major axis for line-fitting. *Am. J. Phys. Anthropol.* **140**: 476–486.
Sokal, R.R. & Rohlf, F.J. 2012. *Biometry*, 4th edn. W.H. Freeman and Co., New York.
Voje, K.L. & Hansen, T.F. 2012. Evolution of static allometries: adaptive change in allometric slopes of eye span in stalk-eyed flies. *Evolution* **67**: 453–467.
Warton, D.I., Wright, I.J., Falster, D.S. & Westoby, M. 2006. Bivariate line-fitting for allometry. *Biol. Rev.* **81**: 259–291.

Received 2 November 2015; revised 27 September 2016; accepted 3 October 2016