Evolutionary Ecology Research, 2014, **16**: 689–704

Choice of regression model for isodar analysis

Jack W. Bradbury and Sandra L. Vehrencamp

Department of Neurobiology and Behavior, Cornell University, Ithaca, New York, USA

ABSTRACT

Background: Isodar analysis allows us to assess the multiple influences of density on relative fitnesses in alternative habitats. It relies on linear regressions. But isodar analyses have a non-classical error structure that complicates the choice of regression method. Prior comparisons of alternative regression methods using simulated data failed to emulate the type of error structure encountered in isodar studies.

Question: Which linear regression method gives the most accurate slope and intercept estimates in isodar analyses?

Method: Simulate datasets that mimic typical isodar studies but have known functional slopes and intercepts, known levels of measurement error, and known equation error. Generate 50 randomly drawn replicates for each combination of input parameters and compare the true slopes to those estimated by each of the four most widely used linear regression methods (ordinary least squares, major axis, geometric mean, and bisected angle protocols). Identify the most reliable method for isodar analyses.

Range of key variables: In total, 14,256 combinations of parameters were examined. Parameter ranges were: mean number of XY pairs/dataset (20–640), coefficient of variation around mean numbers (5–40%), true slopes (0.5–2.0), ratio of ideal free distribution Y mean to X mean (1.2, 1.4), equation error (1–20%), and measurement error (0.1–10%).

Results: All four methods exhibited biases. Bias declined as the Pearson correlation between X and Y increased. Ordinary least squares always underestimated true slopes, often severely. The geometric mean and bisected angle protocols underestimated true slopes greater than one, and overestimated those less than one. Major axis regressions showed the reverse pattern of biases and could behave erratically. All four methods gave unreliable slope estimates for Pearson correlation coefficients less than 0.4. For Pearson correlations above that threshold, the geometric mean method had the least bias and remains the most conservative choice for isodar studies.

Keywords: equation error, isodars, measurement error, Model II regression.

INTRODUCTION

In a recently published study (Bradbury *et al.*, 2015), we mapped sward biomass densities and grazing Thomson's gazelles (*Eudorcas thomsonii*) in Kenya to determine whether the animals

Correspondence: J.W. Bradbury, Department of Neurobiology and Behavior, Cornell University, Ithaca, NY 14853, USA. e-mail: jwb25@cornell.edu

were distributed over the resource in an ideal free manner (*sensu* Fretwell and Lucas, 1969). Previous attempts to verify ideal free distribution theory have met with mixed success (Emlen *et al.*, 2003). In part, this is due to the widespread belief that the theory predicts proportional settlement where the fraction of animals in each alternative habitat at equilibrium equals the fraction of resource there (Parker, 1978; Morris, 1994). Proportional settlement is a prediction only if all the resource in each habitat is used. Each of the N animals in a given habitat must then consume $1/N$ of the total resource there. Our gazelles were clearly not consuming all of the grass available, so we looked for another approach that did not make this assumption. We chose isodar analysis, which assumes that each additional forager in a habitat depresses fitness there by the same amount. This makes density dependence a linear instead of hyperbolic function.

Morris (1987) created isodar analysis as a more general extension of ideal free distribution theory. When applied to habitat selection, the method examines a focal species' dispersion over pairs of alternative habitats and requires multiple samples of the concurrent densities of settlers in each. The goal is to identify both quantitative and qualitative differences between the habitats that modulate settlement. A regression of the densities in the presumed more suitable habitat on densities in the less suitable one generates a slope and intercept used to infer the type of density-dependent 'rules' used by settling animals within this set of habitats (Morris, 1988). When isodar intercepts are significantly positive, the habitat assigned to the ordinate is inferred to be quantitatively superior to that assigned to the abscissa. Qualitative differences between the habitats, such as different foraging efficiencies, can either enhance or undermine the quantitative ones as overall ambient densities vary. If the two factors affect fitness similarly, the slope of the isodar will be greater than one. If they work against each other, it will have a value between zero and one, and it is possible that the habitat that is most densely settled will shift (crossover) as density varies. If the habitats are the same qualitatively, then the slope will be one. If they are also similar quantitatively, then settlers will be indifferent to habitat (intercept = 0, slope = 1), making it effectively a single habitat. If there is no density dependence, settlers will crowd into the quantitatively more suitable habitat regardless of overall density. The associated isodar will have an indeterminable slope. The critical questions thus boil down to whether the isodar intercept is positive, whether the slope is significantly greater than zero, and if so, is it significantly different from one.

Having obtained a suitable gazelle dataset for isodar analysis, we faced the problem of which type of linear regression protocol to use. A number of published isodar studies have used ordinary least square regressions (hereafter OLS). This method, also called a Model I regression, assumes that errors in the abscissa variable are sufficiently small that they can be ignored. It then extracts the regression line that minimizes deviations along the ordinate axis (Sokal and Rohlf, 1995). As Morris (1987) pointed out, this assumption is routinely violated with isodar studies: any animal mis-assigned to a habitat counts as one more in that habitat but also as one less in the other [an example of 'measurement error' (*sensu* Fuller, 1987)]. Similarly, animals that ignored current settlement rules during mapping, perhaps because of preoccupation with activities unrelated to quantitative or qualitative habitat differences, would also increase the deviations around the regression line along both axes ['equation error' (*sensu* Fuller, 1987)]. The recommended solution to this problem is Model II regression (McArdle, 1988; Sokal and Rohlf, 1995; Legendre and Legendre, 1998). There are over 20 alternative Model II protocols to choose from, each with its own assumptions and biases (McArdle, 1988, 2003; Isobe *et al.*, 1990; Jolicoeur, 1990; Babu and Feigelson, 1992; Kimura, 1992; Angleton and Bonham, 1995; Carroll and Ruppert, 1996;

Schmid *et al.*, 2000; Leng *et al.*, 2007; Smith, 2009; Ludbrook, 2010; Xu, 2014). All of these alternatives derive from the same maximum likelihood estimator of slopes called the 'structural model' (Kendall and Stuart, 1973; Fuller, 1987; McArdle, 1988). The equation for this model contains one critical term, the ratio of the ordinate error to that of the abscissa. The different Model II protocols make different assumptions about the value of this unknown ratio (Jolicoeur, 1990). The three most widely used, and frequently recommended, Model II methods are the major axis (MA), geometric mean (GM), and bisected angle (BS) protocols.

Major axis (also called orthogonal) regressions assume that the unknown error ratio is unity and errors in X and Y are fully symmetrical. The Pearson correlation between Y and X appears explicitly in the relevant equations (Babu and Feigelson, 1992). The method finds the regression slope that minimizes the sum of the squared perpendicular distances between each XY point and the line. This is equivalent to finding the longest axis in the 95% confidence ellipse around the XY points. Leng *et al.* (2007) reviewed the algebraic differences between methods, and compared OLS, GM, and MA regressions using simulated data. They concluded that MA gave the most reliable slope values when there is significant abscissa error.

A second Model II method, bisected angle, also assumes symmetrical errors in Y and X and incorporates the correlation between the variables in its formulation (Babu and Feigelson, 1992). It finds the line that bisects the angle between the lines formed by plotting the OLS regressions of Y on X and X on Y on the same graph. Babu and Feigelson (1992) used simulated data to compare OLS, MA, GM, and BS methods and concluded that the BS method was most accurate when there is abscissa error, particularly for smaller sample sizes.

The geometric mean (also called reduced major axis) method assumes that the unknown error variance ratio is proportional to the ratio of the Y and X sample variances (e.g. the errors need not be symmetrical). The slope estimate is computed as the geometric mean of the slopes of the OLS regression of Y on X and the reciprocal of the OLS regression of X on Y. In practice, the GM slope is approximately equal to the ratio of the standard deviations of Y and X. Unlike the prior two methods, GM regressions do not explicitly incorporate the correlation between Y and X: even if the two variables are uncorrelated, a GM regression would dutifully crank out a finite slope value. Given this problem, Jolicoeur (1990) recommended that GM only be used when sample sizes are greater than 20 and correlation coefficients between variables are greater than 0.6. McArdle (1988) compared OLS, MA, and GM in simulated datasets based on the full structural equation with specified ratios of error variances. He found that GM generally did best, although the best method could shift as the ratio of error variances became more extreme.

This disagreement about which regression was most reliable given abscissa error has persisted. After reviewing the many conflicting claims, Ludbrook (2010) concluded that the outcomes of the various simulation studies were sufficiently diverse as to be 'almost uninterpretable'. Other authors have dodged the choice between methods by trying to correct for abscissa error: if the errors in the two variables are uncorrelated, bootstrap sampling and invocation of relevant equations can estimate the abscissa error and correct for its bias (Fuller, 1987; Carroll *et al.*, 2006; Buonaccorsi, 2010). However, this and similar approaches do not help solve the problem for isodar analyses. Errors in isodar regressions are invariably correlated, generating 'non-classical error structure'. Non-classical errors in regression are not easily quantified or corrected analytically (Reilman *et al.*, 1985; Erickson, 1993; Kelly, 2007; Chen *et al.*, 2011; O'Neill and Sweetman, 2013; Muff *et al.*, 2015). None of the previous simulations seeking the 'best method' examined datasets with non-classical error structures.

Faced with this confusion, we performed both OLS and GM regressions on our gazelle data. Ordinary least squares yielded shallow slopes much less than one and large positive intercepts, implying that gazelles likely change their habitat preference with increasing density. The geometric mean method yielded steeper slopes just slightly (and not always statistically) less than one with a positive intercept marginally larger than zero, implying that a crossover in habitat preference was less likely. The differences were important because crossovers are incompatible with constant proportional settlement. We argued in the paper that the true values were likely between these two estimates and that crossovers were likely and proportional settlement unlikely. However, we could not know for sure what the true slopes were from the regressions alone. The simulations in this paper attempt to resolve that uncertainty in our work while providing advice for future applications of isodar theory.

We constructed datasets with typical isodar error structures and a wide range of sample sizes and levels of measurement and equation error. We then compared the slopes estimates from OLS, MA, GM, and BS regressions to the known true values used to create the datasets. We focus on slopes because bias in our slope estimates is negatively correlated with bias in intercepts. This arises because all four methods force the regression line through the means of the two variables. The joint means act like a fulcrum, so that a steeper slope means a lower intercept and vice versa. We also pay close attention to the Pearson correlation between the two variables, as this can be measured in any isodar study and, as we show, provides important insights into the interpretation of the regression results. In brief, we find that slope bias exists in all four methods, bias magnitudes are quite sensitive to the correlation between the variables, and surprisingly, the direction of Model II bias can shift depending on whether the true slope is greater or less than one.

METHODS

The gazelle system

We used our gazelle study as a model for our simulations. Densities of grazing gazelles in the 'rich' half of our study area (grass biomass densities higher than the median) during any given sample constituted our Y regression variable (ordinate). Densities in the remaining 'poor' half of the study area during that same sample were assigned to the X variable (abscissa). Overall densities varied widely over our one-year study, and this permitted the accumulation of multiple XY pairs over a broad range of abscissa values. If every gazelle followed the same settlement rule and we made no mapping errors, the correlation between Y and X would be perfect and all regression methods would yield the same slope and intercept values (Sokal and Rohlf, 1995; Legendre and Legendre, 1998). In practice, not every mapped gazelle was likely to attend to the settlement rule (equation error), and we surely mis-assigned some animals close to habitat boundaries (measurement error). This is why our plots of Y versus X did not show straight lines but instead clouds of points around a general trend. There were thus significant errors in both axes. Similar issues arise in many published papers using isodar analysis. Any comparison of regression methods for isodar analysis needs to replicate datasets with these kinds of correlated errors.

Setup of simulation datasets

We used MATLAB (v.R2014b, The Mathworks, Inc.) to create the required XY datasets over a wide range of numbers of XY points in each dataset, mean animals per XY point, and levels of both equation and measurement errors (code available at evolutionary-ecology.com/data/2960Appendix.pdf). In total, we examined 14,256 combinations of the following parameters:

- *Sample size*: Each dataset contained 20, 40, 80, 160, 320, or 640 XY points.
- *Initial numbers of animals assigned to each X in a dataset*: Each X value was drawn from a normal random distribution given an assigned mean of 25, 50, or 100 gazelles and an assigned coefficient of variation of 5%, 10%, 20%, or 40%. These values mimicked the numbers of foraging gazelles per 25 hectares in our field study. These X animals were initially assigned to the abscissa habitat.
- *True slopes*: The settlement rule was assumed to be linear and thus of the form $Y = mX + b$. We examined true slopes, m, of 0.5, 0.6, 0.7, 0.8, 0.9, 1.0, 1.2, 1.4, 1.6, 1.8, and 2.0.
- *Intercepts*: Intercepts (b) of all regression methods were computed using the slope estimate and the means of the two variables ($\bar{X}$, $\bar{Y}$). We thus defined a parameter $R = \bar{Y}/\bar{X}$ for which we supplied a set of possible values. The resulting assigned intercept for the settlement rule was then $b = \bar{X}(R - m)$. The value of $\bar{X}$ was the actual mean of all the X values in each dataset. We used two values of R based on our gazelle field data: 1.2 and 1.4. Both implied that the ordinate habitat was quantitatively superior to the abscissa habitat (Morris, 1987).
- *Initial numbers of Y animals matched with each X in a dataset*: We then applied the linear rule $Y = mX + b$ for each X value in the current sample to determine the corresponding numbers of animals, Y, assigned to the ordinate habitat. The resulting XY pairs in a given dataset thus constituted a sample that followed the linear settlement rule perfectly.
- *Equation error addition*: Distracted gazelles initially assigned to either habitat by the rule might subsequently wander into the other habitat before mapping. This creates equation error for both Y and X values. We incorporated this effect in the simulations by selecting a fraction NE of all ($X + Y$) animals in each XY sample at random, and for each animal, drawing a number between zero and one from a uniform random distribution. If the number was ≥0.5, the animal was moved from its rule-assigned habitat to the alternative; if the random number was <0.5, habitat assignment was unchanged. Because the NE animals were chosen from the overall pool, some formerly X animals were added to Y and subtracted from the prior X, and some formerly Y animals were added to the X and subtracted from the prior Y. Because of this random selection, the numbers of X and Y animals moved were not necessarily equal. The NE values used were 0.01, 0.10, and 0.20.
- *Measurement error addition*: After reassignments based on equation errors, we added measurement error using the same method but with a different fraction ME. Note that, by chance, some animals reassigned to a different habitat by equation error might be reassigned back to their original rule-assignment by measurement error. The latter was presumed to be less than equation error, so random fractions of the total pool selected for consideration were ME values of 0.001, 0.01, and 0.10.
- *Relative abscissa error variance*: We then computed the abscissa variances before (V_1) and after (V_2) adding the two error modifications. The relative contribution of the two kinds of error to the overall abscissa variance was then calculated as $(V_2 - V_1)/V_2$.

Regression analyses

We created 50 different XY datasets for each of the 14,256 combinations of parameter values (total $N = 712{,}800$). We then applied the OLS, MA, GM, and BS regression routines written by Edward Peltzer (2009) for MATLAB to each. We calculated the mean estimates of slope, intercept, slope standard error, intercept standard error, Pearson correlation between X and Y, and bias for each method (defined as the difference between the slope estimated by that method and the true value) across the 50 datasets for each parameter combination. We assembled these means into an overall data table, looked for trends, and related any patterns seen to input parameters with multiple regression. We transformed distributions as necessary to ensure normal residuals. The best regression models were identified using corrected Akaike Information Criterion (AICc) values. All statistical tests were performed in JMP Pro (v.12.0.0, SAS Institute, Inc., Cary, NC).

RESULTS

Variation and correlation

Values of the Pearson correlation coefficient, ρ, for the XY datasets ranged from −0.853 to 0.998, with a strong skew towards higher values. Relative abscissa error ranged from 0.01 to 0.89 with a strong skew towards lower values. Pooling all datasets, the values of ρ were strongly and negatively correlated with the corresponding relative abscissa variances ($r^2 = 0.934$, $N = 14{,}256$, $P < 0.0001$). Dividing the datasets based on the 11 true (assigned) slope values, Pearson ρ values were even more tightly linked to the underlying abscissa error variance (quadratic fits, all $r^2 \geq 0.997$, each $N = 1296$, each $P < 0.0001$). Focusing on the 11,659 samples with Pearson $\rho \geq 0$, multiple regressions showed that all input parameters contributed significantly to abscissa error, and all but the numbers of XY points in a dataset contributed to the Pearson ρ values (Table 1). The directions of effects are instructive. Increases in the numbers of animals associated with each XY point, and higher coefficients of variation in those numbers, decreased abscissa error and increased Pearson ρ. Increased levels of equation error, measurement error, and the ratio $R = \bar{Y}/\bar{X}$ all had the opposite effects. We used Pearson's ρ as the main independent variable in subsequent contrasts because it is closely linked to the (usually unknown) abscissa error variance, and can be measured in any field study. We limited our attention to non-negative ρ values because all of our assigned simulation slopes were positive.

Absolute bias of the four regression methods

Figures 1 and 2 show how bias in each of the four methods changed as the overall mean Pearson ρ varied from 0 to 1.0. The bias for the OLS regressions was consistently negative, thus underestimating the true slope. Bias decreased (closer to 0) as Pearson ρ increased (reaching 0 when $\rho = 1.0$). Note that despite a wide range of sample sizes, coefficients of variation in the number of animals in a sample, and error rates, the points are quite tightly clustered around the OLS line. This results in small confidence intervals for the slope estimate at any given Pearson ρ even when there is considerable bias. Note also that the spread of points never included the true value of the slope until ρ was quite high.

All three Model II regressions were noisier than the OLS ones, and all three became highly unstable at Pearson ρ values below 0.4. Whereas bias for GM and BS grew

Table 1. Results of multiple regressions predicting relative abscissa error and Pearson ρ from input parameters used to generate isodar-like datasets

	Dependent variables			
	Relative abscissa error		Pearson ρ	
Input parameters	Estimate	*t*-test ratio	Estimate	*t*-test ratio
Sample size coefficient of variation	−1.15	−206.6*	2.17	186.9*
Level of equation error	1.11	117.4*	−1.92	−100.2*
Mean animals per sample	−0.002	−78.4*	0.003	69.2*
Level of measurement error	1.14	69.1*	−1.96	−58.5*
$R = \bar{Y}/\bar{X}$	0.04	29.3*	−0.06	−25.9*
True slope value	0.02	10.9*	0.15	49.5*
Number of XY points in dataset	1.73	5.2*	4.29	0.6
Overall r^2	0.830*		0.794*	

Note: Sign of estimate indicates direction of effect, and *t*-test ratios indicate overall contribution to regression.
* Significant variable (all significant effects with $P < 0.0001$).

increasingly negative as ρ approached 0, bias in MA showed increasingly extreme bias values in both positive and negative directions. Since each method was based on the same data, MA was clearly responding differently to the underlying variation in parameters.

Table 2 summarizes the effects of the various input parameters on bias in each of the three Model II methods. All of the parameters made significant contributions to GM and BS biases. Bias in MA depended much less on input parameters, only responding to a subset of them, and often responding in the opposite direction to the responses of GM and BS. Comparison of Figs. 1 and 2 shows that the direction of bias for all three Model II protocols when Pearson $\rho \geq 0.4$ depended on whether the true slope was less than or greater than one. The MA protocol underestimated true slopes less than one, but overestimated true slopes greater than one; GM and BS showed the opposite behaviours. As with OLS, slope estimates for all Model II methods gradually converged on the true value as the Pearson ρ increased. While GM and BS were nearly linear in their convergence on the true value, MA converged along an asymptotic trajectory. This resulted in the absolute value of MA bias being greater than that for GM at low ρ values, but less than that for GM at high ρ values. For any given Pearson ρ, the average bias for BS was consistently greater than that for GM; otherwise, these two methods showed similar behaviours. The fact that the three Model II estimates reversed the direction of bias depending on whether the true slope was greater or less than one suggested that they should have nearly zero bias when the true slope was 1.0. This was in fact the case: the median bias (and 95% confidence intervals) for the Model II methods when the true slope equalled 1.0 and the Pearson ρ was 0.4 or higher were: MA = 0.0012 (−0.0184, 0.0779), GM = 0.0008 (−0.0184, 0.0305), and BS = 0.0007 (−0.0181, 0.0271).

Relative magnitudes of OLS and GM biases

The two most widely used regression methods in isodar studies are OLS and GM. Whether they respond similarly to varying Pearson ρ and true slope values is thus important in method selection. Figures 1 and 2 indicate that the absolute magnitude of OLS bias at any

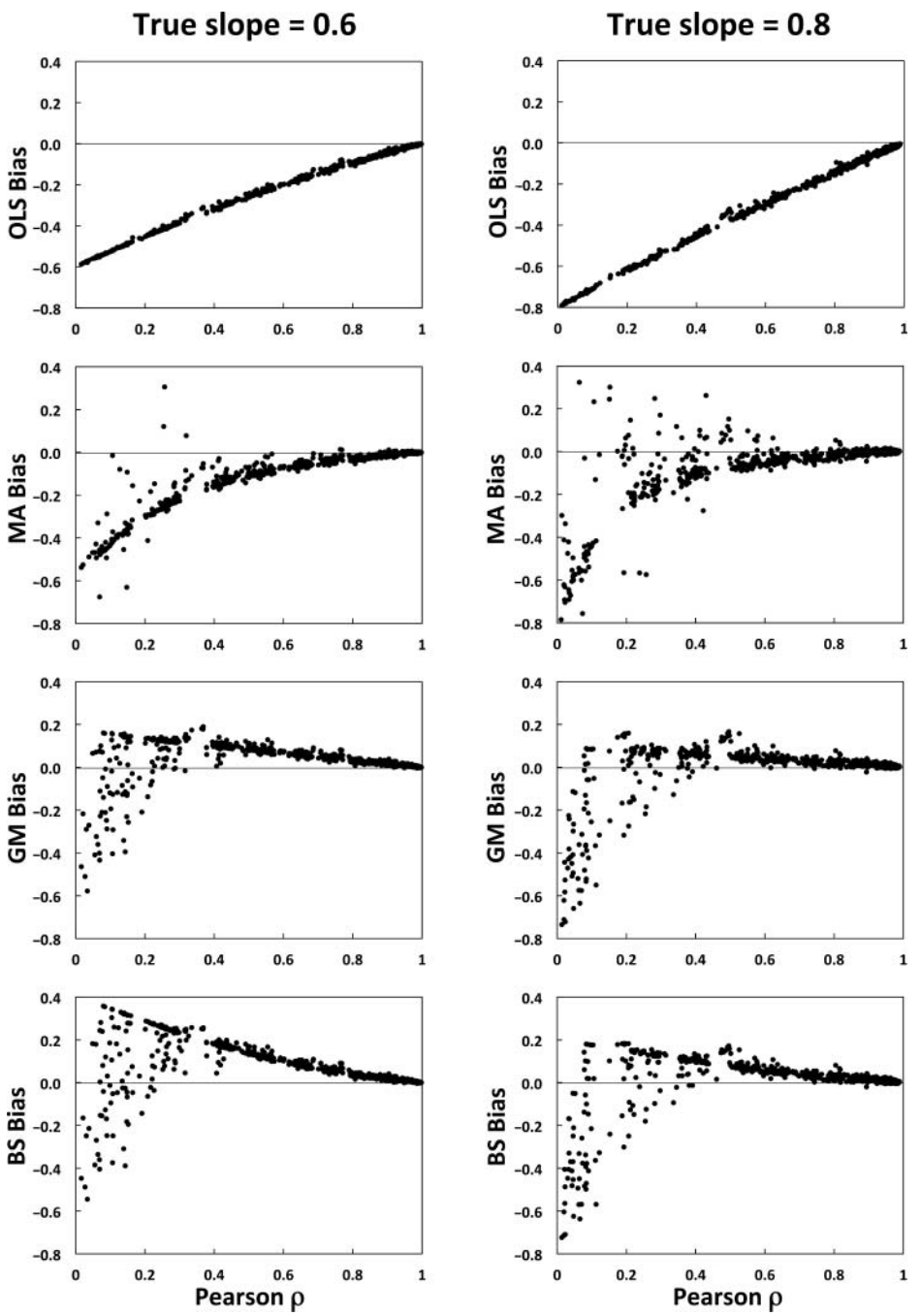

Fig. 1. Plots of bias in slope estimates versus the Pearson correlation coefficients calculated between regression variables for four alternative linear regression methods. Columns of plots show results for two true (initially assigned) slopes less than unity. Each point is the mean of 50 random samples with the same combination of input parameters. Horizontal lines demarcate absence of bias. Number of points (N) varies due to stochasticity in fraction of samples with Pearson $\rho \geq 0$. $N = 961$ (0.6 slope) and 1024 (0.8 slope).

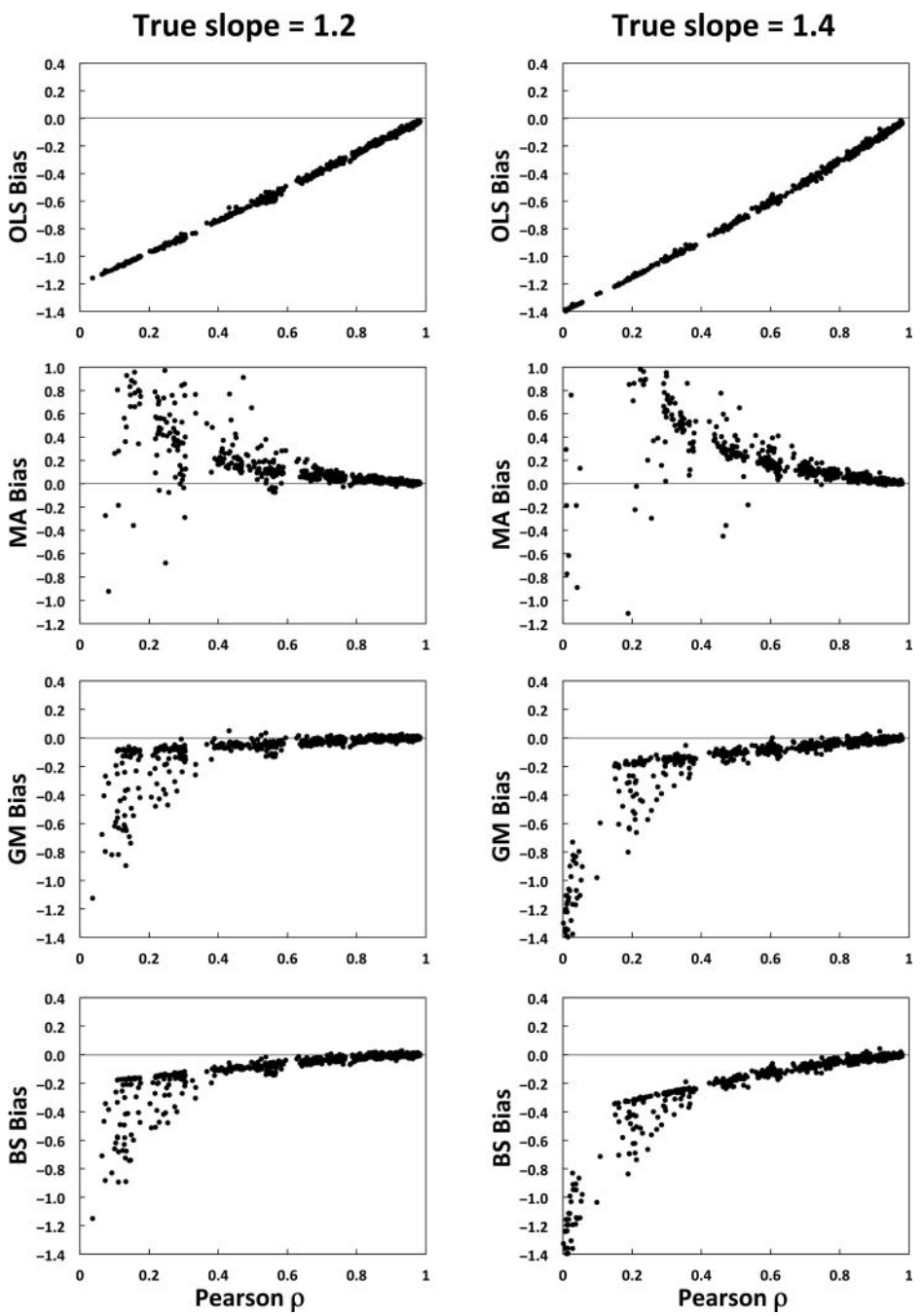

Fig. 2. Bias plots similar to those in Fig. 1 but for true slopes greater than unity. Note the different scale on the MA ordinate. $N = 1087$ (1.2 slope) and 1126 (1.4 slope).

Table 2. Results of multiple regressions of predicting bias among Model II protocols from input parameters for datasets with Pearson $\rho \geq 0$

	MA		GM		BS	
Input parameters	*t*-test ratio	*P*	*t*-test ratio	*P*	*t*-test ratio	*P*
True slope value	5.19	<0.0001	−0.13	<0.0001	−0.19	<0.0001
Sample size coefficient of variation	−3.95	<0.0001	0.46	<0.0001	0.55	<0.0001
Level of equation error	0.69	0.486	−0.41	<0.0001	−0.49	<0.0001
Level of measurement error	2.21	0.027	−0.41	<0.0001	−0.47	<0.0001
Mean animals per sample	−3.01	0.003	0.0005	<0.0001	0.0006	<0.0001
Number of points in dataset	1.61	0.106	5.60	<0.0001	5.81	<0.0001
$R = \bar{Y}/\bar{X}$	−0.92	0.358	−0.01	<0.0001	−0.01	<0.0001
Overall r^2 and *P*	0.005	<0.0001	0.250	<0.0001	0.308	<0.0001

Table 3. Relative bias of OLS and GM regressions associated with true slope values

	% Deviation for OLS regressions		% Deviation for GM regressions	
True slope	Pearson $\rho = 0.6$	Pearson $\rho = 0.8$	Pearson $\rho = 0.6$	Pearson $\rho = 0.8$
0.5	−32.1	−14.8	14.8	7.8
0.6	−33.0	−15.8	11.7	5.6
0.7	−34.6	−16.9	7.7	3.6
0.8	−36.7	−17.8	5.6	2.8
0.9	−37.5	−19.2	2.9	1.5
1.0	−39.9	−19.8	0.2	0.1
1.2	−41.0	−22.1	−2.9	−1.6
1.4	−43.6	−22.7	−6.1	−2.9
1.6	−45.0	−23.1	−8.1	−4.2
1.8	−45.8	−24.2	−9.8	−5.3
2.0	−46.5	−24.7	−12.4	−6.1

Note: Cell values are the mean bias for each parameter combination divided by the true slope value × 100.

given Pearson ρ was larger than that for GM. Table 3 confirms this point. The bias of the OLS estimates continued to become larger as the true slope increased. The GM biases were nearly zero at a true slope of one but increased in absolute value as the true slope became increasingly different from one. However, over the range of true slopes examined, the absolute value of the GM bias remained less than half of that for OLS. The increasing absolute magnitude of the OLS bias as true slopes increased may seem surprising. Examination of Figs. 1 and 2 reveals that the plot of OLS slope estimates versus Pearson ρ is actually steeper for higher true value slopes. Hence, at any given Pearson ρ, the OLS slope estimate would deviate even more at higher true slope values.

Relationships between OLS and GM regression slopes

Theoretically, the slope of an OLS regression should equal the product of the relevant Pearson ρ and the slope of the corresponding GM regression (Babu and Feigelson, 1992; Legendre and

Legendre, 1998; Smith, 2009). Our simulated samples verified this expectation: a linear (OLS) regression of OLS slope estimates on the corresponding ρ*GM slope estimates using all 11,659 parameter combinations with a Pearson $\rho \geq 0$ had a slope of 0.991 and an intercept of 0.011 ($r^2 = 0.9994$, $P < 0.0001$). The regression was even tighter if restricted to datasets with a Pearson $\rho \geq 0.4$: slope = 0.9995, intercept = 0.0007, and $r^2 = 0.99995$. One consequence of this tight relationship is that no combination of OLS and GM slope estimates would provide any more information about the true slope values than is available from the GM slope and Pearson ρ values alone. Since GM regression appears to provide the least bias in this system, at least for Pearson ρ values ≥ 0.4, the next obvious question is whether one can correct for bias in GM regressions without knowledge of the underlying ratio of error variances.

Patterns of bias in GM regressions

For datasets with Pearson ρ values ≥ 0.4, plots of GM slope estimates versus Pearson ρ for any given true slope showed quite linear relationships. Bias decreased at higher ρ values. The 'bias slopes' for these linear relationships were negative for true XY slopes < 1.0, nearly zero for a true slope of 1.0, and positive for true slopes > 1.0. Bias slopes thus appeared to increase monotonically with increasing true slopes. We modelled this relationship as follows: GM slope estimates = $A + a(T)^*\rho$ where $a(T)$ is the bias slope for a given true (assigned) dataset slope T. Note that GM bias is simply the difference between the GM slope estimate and the true value. Thus $a(T)$ is also the slope of the bias relationships shown in Figs. 1 and 2. Since the GM slope estimate equals T when $\rho = 1.0$, then $A = T - a(T)$. Suppose $a(T)$ can also be modelled as a linear relationship, e.g. $a(T) \approx Q - w^*T$. Substituting this into the equation for GM slope and solving for T, we get:

$$T = \frac{GM\ slope - Q\,(1-\rho)}{1 + w\,(1-\rho)} \tag{1}$$

A regression of GM slope estimates on Pearson ρ separately for each T indicated that $a(T)$ was indeed a monotonically increasing function of T. A linear fit for $\rho \geq 0.4$ ($r^2 = 0.986$, $P < 0.0001$) yielded estimates of $Q = 0.507$ and $w = 0.528$. Resulting distributions of corrected T estimates were normal for true slopes less than one, providing accurate confidence intervals, but were increasingly skewed to lower values for true slopes larger than one (Table 4). Still, this formula provided reasonable estimates of the true slope values using only the GM estimate and the concurrent Pearson ρ. Our fitted coefficients are unlikely to be general, but this exercise points out that unlike OLS, the Pearson ρ provides additional information about true slopes beyond a GM analysis alone.

Since GM slope estimates $\approx T - a(T) + a(T)^*\rho$, and OLS slope estimates = ρ*GM slope estimates, OLS slopes $\approx \rho(T - a(T)) + a(T)^*\rho^2$. This implies that the true relationship between OLS slope estimates and the corresponding Pearson ρ values should be quadratic with an intercept of zero. In fact, this is the case. For each of the true slopes used in our simulations, a quadratic fit gave a lower AICc value than did a linear one. Both the linear and quadratic coefficients were significantly greater than zero, and the intercepts were all very close to zero.

Table 4. Corrected slope estimates and confidence limits (CL) calculated from the GM slope estimate, estimated GM bias slope, and Pearson ρ

True slope	GM bias slope	Corrected estimate	Lower 95% CL	Upper 95% CL
0.5	–0.192	0.498	0.485	0.516
0.6	–0.171	0.600	0.587	0.623
0.7	–0.138	0.700	0.684	0.724
0.8	–0.115	0.800	0.777	0.830
0.9	–0.069	0.898	0.864	0.950
1.0	–0.006	0.991	0.956	1.019
1.2	0.101	1.191	1.117	1.212
1.4	0.215	1.385	1.289	1.409
1.6	0.325	1.581	1.457	1.607
1.8	0.459	1.775	1.584	1.807
2.0	0.615	1.967	1.732	2.005

Note: See text for relevant formulae.

Point spread, standard deviations, and confidence intervals

The points in the OLS plots of Figs. 1 and 2 appear more tightly clustered than those for the corresponding GM plots. Given the algebraic relationship between OLS and GM slopes, we might expect the range of OLS slope estimates for any given Pearson ρ value to equal the concurrent range of GM estimates multiplied by ρ. The tighter clustering of OLS points in Figs. 1 and 2 suggests this is the case. We verified this quantitatively by performing separate linear regressions of each of the two methods' slope estimates versus Pearson ρ for each of our true slope values. We then saved the ordinate residuals and computed the ranges required to encompass 95% of the points. We also identified the minimal Pearson ρ above which the true slope was included within the regressions' 95% confidence limits.

As expected, the range of variation around the slope estimates for Pearson ρ values above 0.4 was higher for GM regressions than OLS ones (Table 5). Both methods showed a slight drop in point spread as the true slope approached 1.0, and increased spreads for true slopes further from 1.0. Although the spreads for both methods increased with the deviation of true slope value from 1.0, so did the concurrent biases. The result, for both methods, was that the minimum Pearson ρ at which the point spread included the true value was higher when the deviation of the true slope from 1.0 was also higher. At any given true slope value, the minimum ρ required to include the true value in the 95% spread was lower for GM than for OLS. This means that the confidence intervals for an OLS slope estimate might fail to include the true slope value even when those for the concurrent GM regression would do so. At the same time, the noisier GM protocol means that any given estimate might or might not be close to the true value.

DISCUSSION

Our simulations provide some likely explanations for the conflicting recommendations in the literature about which Model II method is superior. As predicted in the textbooks (Sokal and Rohlf, 1995; Legendre and Legendre, 1998), all of the regression methods examined converged on

Table 5. Variation around OLS and GM regression lines and minimum Pearson ρ required to include true values within regression confidence interval (CI)

	95% Range of plot residuals		Minimum ρ to include true value in 95% CI	
True slope	OLS	GM	OLS	GM
0.5	0.026	0.028	0.951	0.914
0.6	0.026	0.032	0.964	0.891
0.7	0.031	0.037	0.969	0.857
0.8	0.031	0.047	0.972	0.768
0.9	0.049	0.073	0.973	0.569
1.0	0.035	0.049	0.983	0.400
1.2	0.040	0.053	0.987	0.788
1.4	0.047	0.064	0.989	0.868
1.6	0.065	0.078	0.990	0.900
1.8	0.080	0.096	0.990	0.900
2.0	0.078	0.078	0.991	0.940

Note: Variation was estimated by saving the residuals from regressions of estimated slopes on the Pearson ρ values for each true slope. These residuals quantify the spread of points around these regressions. The average range of this spread along the ordinate axis that includes 95% of the estimates is given for each method in columns 2 and 3.

the same true slope values as the correlation between the ordinate and abscissa variables approached one. Because GM and BS did so linearly whereas MA did so asymptotically, MA gave more accurate slope estimates once the Pearson ρ was sufficiently large. This might explain Leng and colleagues' (2007) claim that MA is the more reliable method. However, this was only true for the highest Pearson ρ values, at which point all the methods gave fairly accurate slope estimates.

Babu and Feigelson (1992) have argued that BS is generally the superior method. In our simulations, GM and BS yielded very similar slope estimates, responded to input variation in similar ways, and showed the same reversals in bias depending on the value of the true slope compared to one. However, GM was slightly but consistently more accurate than BS for all Pearson ρ values ≥ 0.4. The difference between our results and those of Babu and Feigelson may be due to the different error structures examined, or even to stochastic variations, since the two methods often give quite similar estimates.

Finally, our results confirm the conclusion of McArdle (1988) that GM is generally the most reliable method when both variables have errors. However, he worried that none of the methods would give trustworthy estimates when the errors in the two variables were correlated (McArdle, 2003). Our simulations hopefully allay some of those concerns, at least for isodar-like datasets and Pearson ρ values ≥ 0.4.

Although GM usually had the least absolute bias among these four regression methods, this does not automatically make it the method of universal choice. The four methods differ in their sensitivities to the relevant Pearson ρ, and to the direction of their biases. One thus needs to match the acceptable biases to the question being asked. For isodar studies, the critical question is usually whether the regression slope is significantly less than, equal to, or greater than one. One should first check that the Pearson ρ value for the XY correlation is at least 0.4; all the methods produce unreliable slope estimates when this condition is not met. Given that it is, one should then consider that OLS always underestimates the true

slopes, GM and BS overestimate slopes less than one and underestimate those greater than one, and MA overestimates slopes greater than one and underestimates those less than one. A regression yielding a slope estimate significantly less than one should thus be accepted if the method was GM or BS; a result implying a true slope significantly greater than one should be accepted if the method was OLS, GM, or BS. Confidence intervals for these methods that cannot exclude a true slope of one should be coupled with an examination of the Pearson ρ value for the dataset: the higher this correlation, the more likely the true slope is indeed one.

Our simulations indicate that MA should be used with caution. We confirmed McArdle's (1988) observation of highly erratic slope estimates by MA. He only examined two Pearson ρ values, both greater than 0.4, but even at $\rho = 0.5$ he observed wildly deviant slope values. He attributed these jumps to variations in the ratio of the X and Y error variances. Equation error had no effect on MA bias in our regressions, and measurement error had only a marginal effect. The major factors modulating MA bias were those associated with the number of animals in a dataset. These results contrast strongly with our regressions for GM and BS, which both showed significant links between bias levels and all of the input parameters. Because of the way MA fits a line to data, it is potentially more sensitive to outliers than either GM or BS, and this was indeed what Ludbrook (2010) found. Whatever the causes of MA's erratic behaviour, the general import is that it is a generally less reliable method than GM or BS for isodar studies.

These simulations essentially confirm our tentative conclusion in our gazelle study: the true slopes of our regressions were most likely between the OLS and GM estimates. Both OLS and GM yielded slopes less than one, but the confidence limits for GM could not exclude a slope of one. Given the Pearson ρ values in our gazelle datasets, and the relative biases listed in Table 3, the slope estimates in our GM regressions were likely about 8% too high. Incorporating this correction brings the GM slope estimates below one. Isodar slopes less than one imply the potential for crossover shifts in habitat preference as overall densities vary (Morris, 1988). We observed several such crossover events in the gazelle study, both as overall densities increased and then later as they decreased again.

Overall, our simulations reinforce the claim that GM regression is usually the better method to test isodar hypotheses. Geometric mean contrasts are conservative, in that a significant deviation from a slope of one is likely reliable despite any bias. However, if the suitabilities of the two habitats being contrasted are not strongly different, the associated isodar slope may truly but only slightly differ from one, and no Model II regression protocol is likely to exclude a null model of one with confidence. But then if the two habitats are that similar from the animal's point of view, this is the right answer!

McLoughlin *et al.* (2010) argued that isodar analysis is best suited to settlement across different habitats whereas other approaches, such as resource utilization functions, should be used for within-habitat choices. This recommendation depends critically on the scale and definition of habitat. As long as the scale is large enough to enable estimates of density, then differences in densities (and thus isodars) define the relevant habitats (Morris, 2003). This is what we did with the gazelle swards, dividing them into contiguous zones where local grass biomass densities were either greater than or less than the median. The gazelles moved into and out of these zones as if they were alternative habitats, and the resulting isodar analyses provided major insights into the determinants of foraging dispersions at this fine scale. The isodar method just requires careful attention to the type of regression protocol, the Pearson ρ correlation between ordinate and abscissa variables, and whether the estimated slopes are

greater or less than unity. Having done so, one can then get on with the important business of using isodars to infer the evolutionary and ecological significance of habitat selection.

ACKNOWLEDGEMENTS

We thank Dr. Jacob Bien, Department of Biological Statistics and Computational Biology, Cornell University for helpful discussions about this study. EER Editor Douglas Morris and two anonymous reviewers also provided very helpful suggestions that improved the manuscript. Cornell University's Emeritus Faculty Program provided funding of licenses for relevant software.

REFERENCES

Angleton, G.M. and Bonham, C.D. 1995. Least-squares regression vs. geometric mean regression for ecotoxicology studies. *Appl. Math. Comput.*, **72**: 21–32.

Babu, G.J. and Feigelson, E.D. 1992. Analytical and Monte-Carlo comparisons of six different linear least squares fits. *Communications in Statistics: Simulation and Computation*, **21**: 533–549.

Bradbury, J.W., Vehrencamp, S.L. and Clifton, K.E 2015. The ideal free antelope: foraging dispersions. *Behav. Ecol.*, **26**: 1303–1313.

Buonaccorsi, J.P. 2010. *Measurement Error: Models, Methods, and Applications.* New York: Chapman & Hall/CRC.

Carroll, R.J. and 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., Stefansky, L.A. and Crainiceanu, C.M. 2006. *Measurement Error in Nonlinear Models.* New York: Chapman & Hall/CRC.

Chen, X.H., Hong, H. and Nekipelov, D. 2011. Nonlinear models of measurement errors. *J. Econ. Lit.*, **49**: 901–937.

Emlen, J.M., Freeman, D.C., Kirshhoff, M.D., Alados, C.L., Escos, J. and Duda, J.J. 2003. Fitting population models from field data. *Ecol. Model.*, **162**: 119–143.

Erickson, T. 1993. Restricting regression slopes in the errors-in-variables model by bounding the error correlation. *Econometrica*, **61**: 959–969.

Fretwell, S.D. and Lucas, H.L.J. 1969. On territorial behavior and other factors influencing habitat distribution in birds. *Acta Biotheor.*, **19**: 16–36.

Fuller, W.A. 1987. *Measurement Error Models.* New York: Wiley.

Isobe, T., Feigelson, E.D., Akritas, M.G. and Babu, G.J. 1990. Linear-regression in astronomy. 1. *Astrophys. J.*, **364**: 104–113.

Jolicoeur, P. 1990. Bivariate allometry: interval estimation of the slopes of the ordinary and standardized normal major axes and structural relationship. *J. Theor. Biol.*, **144**: 275–285.

Kelly, B.C. 2007. Some aspects of measurement error in linear regression of astronomical data. *Astrophys. J.*, **665**: 1489–1506.

Kendall, M.G. and Stuart, A. 1973. *The Advanced Theory of Statistics*, 3rd edn. New York: Hafner.

Kimura, D.K. 1992. Symmetry and scale dependence in functional-relationship regression. *Syst. Biol.*, **41**: 233–241.

Legendre, P. and Legendre, L. 1998. *Numerical Ecology*, 2nd English edn. New York: Elsevier.

Leng, L., Zhang, T., Kleinman, L. and Zhu, W. 2007. Ordinary least square regression, orthogonal regression, geometric mean regression, and their applications in aerosol science. *J. Phys. Conf. Ser.*, **78**: 1–5.

Ludbrook, J. 2010. Linear regression analysis for comparing two measurers or methods of measurement: but which regression? *Clin. Exp. Pharmacol. Physiol.*, **37**: 692–699.

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.

McLoughlin, P.D., Morris, D.W., Fortin, D., Vander Wall, E. and Contasti, A.L. 2010. Considering ecological dynamics in resource selection functions. *J. Anim. Ecol.*, **79**: 4–12.

Morris, D.W. 1987. Tests of density-dependent habitat selection in a patchy environment. *Ecol. Monogr.*, **57**: 269–281.

Morris, D.W. 1988. Habitat-dependent population regulation and community structure. *Evol. Ecol.*, **2**: 353–269.

Morris, D.W. 1994. Habitat matching: alternatives and implications to popultions and communities. *Evol. Ecol.*, **8**: 387–406.

Morris, D.W. 2003. Toward an ecological synthesis: a case for habitat selection. *Oecologia*, **136**: 1–13.

Muff, S., Riebler, A., Held, L., Rue, H. and Saner, P 2015. Bayesian analysis of measurement error models using integrated nested Laplace approximations. *J. R. Stat. Soc. C: Appl. Stat.*, **64**: 231–252.

O'Neill, D. and Sweetman, O 2013. The consequences of measurement error when estimating the impact of BMI on labor market outcomes. *IZA J. Labor Econ.*, **2**: 3.

Parker, G.A. 1978. Searching for mates. In *Behavioural Ecology*, 1st edn. (J.R. Krebs and N.B. Davies, eds.), pp. 214–244. Oxford: Blackwell Science.

Peltzer, E.T. 2009. *Model I and model II regressions*. Monterey Bay Aquarium Research Institute [http://www.mbari.org/staff/etp3/regress/index.htm].

Reilman, M.A., Gunst, R.F. and Lakshminarayanan, M.Y. 1985. Structural model estimation with correlated measurement errors. *Biometrika*, **72**: 669–672.

Schmid, P.E., Tokeshi, M. and Schmid-Araya, J.M. 2000. Relation between population density and body size in stream communities. *Science*, **289**: 1557–1560.

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. and Rohlf, F.J. 1995. *Biometry: The Principles and Practice of Statistics in Biological Research*, 3rd edn. New York: W.H. Freeman.

Xu, S.J. 2014. A property of geometric mean regression. *Am. Stat.*, **68**: 277–281.