DYNAMIC FOOD WEBS:
MULTISPECIES
ASSEMBLAGES, ECOSYSTEM
DEVELOPMENT, AND
ENVIRONMENTAL CHANGE

Peter de Ruiter
Utrecht University
Utrecht, The Netherlands

Volkmar Wolters
Justus-Liebig University
Giessen, Germany

John C. Moore
University of Northern Colorado
Greeley, Colorado, USA

Managing Editor
Kimberly Melville-Smith
University of Northern Colorado
Greeley, Colorado, USA

4.1 | SPECIES' AVERAGE BODY MASS AND NUMERICAL ABUNDANCE IN A COMMUNITY FOOD WEB: STATISTICAL QUESTIONS IN ESTIMATING THE RELATIONSHIP

Joel E. Cohen and Stephen R. Carpenter

The quantitative patterns and mechanisms of the relationship between body mass and abundance have been examined in many ecological settings (Colinvaux, 1978; Damuth, 1981; Peters, 1983; Griffiths, 1992; Brown, 1995a, p. 94; Griffiths, 1998; Blackburn and Gaston, 1999; Leaper and Raffaelli, 1999; Kerr and Dickie, 2001; Russo et al., 2003).

The purpose of this chapter is to examine in detail some of the statistical foundations of estimating the quantitative relationship between average body mass and abundance. The empirical example analyzed here is the pelagic food web (hereafter, simply "web") of Tuesday Lake, Michigan (Carpenter and Kitchell, 1993a). The data, provided by Stephen R. Carpenter, are given in full by Jonsson et al. (2005).

Blackburn and Gaston (1999, p. 182) referred to the study of Damuth (1981) as "the single most influential study of the interspecific abundance-body size relationship." Blackburn and Gaston plotted Damuth's 467 data points on 307 species of mammalian terrestrial primary consumers in a graph in which the ordinate (vertical axis) was log N ($\log_{10}$ number of individuals per km^2) and the abscissa (horizontal axis) was log M ($\log_{10}$ body mass, not further specified as to average, adult, or maximal). According to Blackburn and Gaston, the ordinary least squares regression of the linear model

$$\log(N) = \beta_1 \log(M) + \gamma_1 \qquad (1)$$

gave an estimated slope of $\beta_1 = -0.75$. Other data yielded other estimates of the slope (Peters, 1983; Peters and Wassenberg, 1983; Peters and Raelson, 1984; Griffiths, 1992; Cyr et al., 1997a; Cyr et al., 1997b; Griffiths, 1998; Leaper and Raffaelli, 1999). (Throughout the appendices of Peters (1983), the columns headed "Independent variable" should be headed "Dependent variable.")

Enquist et al. (1998, p. 164), Cohen et al. (2003, p. 1784), and most plant ecologists plotted log M on the ordinate and log N on the abscissa. Many approximated the data by ordinary least squares regression of the linear model

$$\log(M) = \beta_2 \log(N) + \gamma_2 \qquad (2)$$

If all data fell exactly on a straight line, then both equations would be exact descriptions, the slope coefficients would be related by $\beta_1 = 1/\beta_2$, and the choice between Eqn. 1 and Eqn. 2 would not matter. But $\beta_1 = 1/\beta_2$ need not hold when there is random variation in the relation between body mass and abundance. It is not clear initially whether Eqn. 1 or Eqn. 2 is a more useful description of patterns in a community of plant and animal species with different modes of growth, some determinate and some indeterminate.

The results show that the two regressions may appear contradictory unless careful attention is paid to how well the data satisfy the assumptions of each linear regression model. Most of the assumptions underlying ordinary least squares regression of the linear model Eqn. 1 cannot be rejected by the Tuesday Lake data. An important limitation of this conclusion is that the assumption of negligible error variance in the measurement of log M cannot be tested with the available data. Most of

the assumptions underlying Eqn. 2 can be rejected by the Tuesday Lake data. This information is sufficient to guide the choice of a linear regression model. However, the direction of causation cannot be determined on statistical grounds from static observational data.

MATERIALS AND METHODS

Data

Tuesday Lake is a mildly acidic lake in Michigan (89°32' W, 46°13' N) (Carpenter and Kitchell, 1993a) with surface area 0.9 ha. The fish populations were not exploited and the drainage basin was not developed. Data collected in 1984 and again in 1986 consisted of a list of species, and for each species, its predator species and its prey species (for the body sizes and life stages present in the lake in each year), its average body mass M (kg fresh weight per individual), and its numerical abundance N (individuals/m^3 in the non-littoral epilimnion, where the trophic interactions take place). In 1985, the three planktivorous fish species were removed and replaced by a single piscivorous fish species. In 1984, the 56 biological species consisted of 31 phytoplankton species, 22 zooplankton species, and 3 fish species. In 1986, the 57 species consisted of 35 phytoplankton species, 21 zooplankton species, and one fish species. Jonsson et al. (2005) gave details of sampling methods, the raw data, and significant limitations of the data, including the delicate estimation of the density of zooplankton species. Our analyses used biological species rather than trophic species. Two or more biological species are "lumped" into a single trophic species if the two or more biological species have an identical set of consumer species and an identical set of resource species. There were 27 trophic species in 1984 and 26 trophic species in 1986 (Jonsson et al., 2005). Cohen et al. (2003) and Reuman and Cohen (2004a, 2004b) analyzed other aspects of the data.

Definitions

The biomass B (kg/m^3) of a species is its average body mass M times its numerical abundance N. Log throughout means $\log_{10}$.

The rank(M) of a species equals 1 if that species has the largest value of M, equals 2 if that species has the second largest value of M, and so on; the rank is the order from largest to smallest average body mass. In 1984, the 56 body masses were described by 55 unique values; in 1986, the

57 body masses were described by 55 unique values. The species with tied body masses retained the ordering of the listing of species in the data appendices of Jonsson et al. (2005). Rank(N) is similarly defined by rank ordering the species from largest to smallest values of N and by retaining the ordering of Jonsson et al. (2005) in case of ties. The rank (with values 1, 2, 3, ...) increases as the numerical value of the variable decreases.

Theory

In the absence of measurement error or random fluctuation, numerical abundance N depends allometrically on average body mass M if and only if biomass B depends allometrically on M. For if $b > 0$ is the exponent in an allometric relation $N \propto M^{-b}$, then $B = MN \propto M^{1-b}$ and conversely. Thus the study of the relation between the numerical abundance and the body mass of biological species is related to the study of the biomass spectrum, as has been widely recognized (Platt and Denman, 1977, 1978), but if the biomass spectrum uses size classes rather than biological species as the unit of analysis, then the relation depends on the number of biological species within each size class. Likewise, N depends allometrically on M if and only if M depends allometrically on N, for $N \propto M^{-b}$ if and only if $M \propto N^{-1/b}$, so biomass B also depends allometrically on numerical abundance N according to $B = MN \propto N^{1-1/b}$.

The two cases of special interest are b = 3/4 and b = 1. If b = 3/4 (as Damuth, 1981 suggested), then $N \propto M^{-3/4}$ so $B = MN \propto M^{1/4}$, and the biomass of a species increases with average body mass. If b = 1, then $N \propto M^{-1}$ and B is the same for every species, regardless of M or of N or of rank(M) or of rank(N). In this case, $c = MN$ implies that $\log N = \log c - \log M$ and $\log M = \log c - \log N$, that is, all species should fall on a line with slope –1 in the plane with coordinates $\log M$ and $\log N$, regardless of which coordinate is the abscissa and which is the ordinate.

Methods

Computations were done using the statistics toolbox and other functions of Matlab, version 6.5.0.180913a (http://www.mathworks.com/, Release 13). Linear regressions were done using 'regress.' All regressions used all species, whether they were connected to the main web or isolated (i.e., not connected to most other species).

Five principal assumptions must be satisfied to justify the probability values and confidence intervals generated by ordinary least squares lin-

ear regression analysis of the model y=a+bx+ε. In a sample of size n, for any data point (x_i, y_i), with i = 1, . . ., n, the residual r_i is defined as the difference in the vertical direction $r_i = y_i - (a+bx_i)$ between the observed ordinate y_i and the predicted ordinate $y_{pred,i} = a+bx_i$ (pred = predicted) given by the linear model, where a and b are the least-squares estimates of the regression coefficients. The five assumptions are *linearity* of the average (conditional expectation) of y as a function of x, *normality* of the residuals ε (with unknown variance), *homoscedasticity* (i.e., the variance of the residuals is independent of x), serial *independence* of the residuals with increasing x, and *no error in the value of x*. A diagnostic Matlab function 'regressiontest' was written to assess the validity of the first four of these assumptions for any pair x, y of data vectors of equal length.

This function evaluated *linearity* in two ways, by classical hypothesis testing and by using the Akaike information criterion. The classic test of nonlinearity was performed by fitting a quadratic equation y=a+bz+cz²+ε, where z=x−mean(x) was a centered translation of x used to reduce the collinearity between x and x². The null hypothesis of linearity was rejected if the confidence interval for the parameter c did not include 0.

To use the Akaike information criterion, the corrected Akaike information criterion AIC_c (Burnham and Anderson, 2002, pp. 63, 66) was computed once for the residuals from the linear model and then again separately for the residuals from the quadratic model, using in each case the formula:

$$AIC_c = n \log \left(\sum_{i=1}^{n} r_i^2 / n \right) + 2Kn / (n - K - 1). \qquad (3)$$

For the linear model, K=3. For the quadratic model, K=4. Intuitively speaking, Eqn. 3 says that, in its application to ordinary least squares regression, AIC_c is a rescaling of the mean squared residual, and the smaller AIC_c the better the model fits. Then

$$\Delta = | AIC_c \text{ (linear model)} - AIC_c \text{ (quadratic model)} | \qquad (4)$$

was computed to compare the goodness of fit of the linear and quadratic models for the same set of (x, y) data. Although statistical significance could not be assigned to any value of Δ, a value of 10 or more is interpreted to mean that the model with the higher AIC_c is essentially without empirical support (Burnham and Anderson, 2002, e.g., p. 226).

Whether the residuals from the regression line were *approximately normal* with unknown variance was examined in three ways. First, for visual inspection, the Matlab function 'qqplot(x)' plotted the quantiles of the residuals r_i as a function of the quantiles of the normal distribution. The more nearly the residuals were normally distributed, the more nearly the quantile-quantile plot approximated a straight line. Second and third, the Matlab functions 'jbtest' performed the Jarque-Bera test and 'lillietest' performed the Lilliefors test of normality with unknown mean and variance. Neither test was impressively sensitive to deviations from normality. For example, at the 0.01 level of significance, 'lillietest' failed to reject normality as a model for the first 358 natural numbers 1, 2, . . . , 358, but did reject normality as a model for the first 359 natural numbers. Similarly, at the 0.01 level of significance, 'jbtest' failed to reject normality as a model for the first 147 natural numbers but did reject normality as a model for the first 148 natural numbers. For samples as small as the 56 species in 1984 or the 57 species in 1986, these tests would not easily detect minor deviations from normality.

If the residuals were *homoscedastic*, there should be no trend in the residuals, or in the absolute value of the residuals, as a function of either the independent variable x or of the predicted value of the dependent variable y_{pred}. Homoscedasticity was tested by fitting linear and quadratic regressions of r_i against $y_{pred,i}$. Homoscedasticity was further tested by fitting linear and quadratic regressions of the absolute value of the residuals $|r_i|$ against $y_{pred,i}$. The null hypothesis of homoscedasticity was rejected if any of the confidence intervals (from either the linear or the quadratic models) of the coefficients of the linear or quadratic terms did not include 0.

The *independence* of successive residuals was tested by sorting the residuals in the order determined by increasing y_{pred} and then comparing computed values of the Durbin-Watson statistic for the residuals against tabulated critical values (Stuart and Ord, 1991, p. 1077 for the formula, pp. 1245–1246 for critical values). The computation of the Durbin-Watson statistic was programmed from scratch. Each numerical result was identical to that obtained from the publicly available Matlab function dwatson.m, accessed from http://econpapers.hhs.se/software/bocbocode/t850802.htm on January 15, 2004.

Testing the assumption of no error in x would require replicate measurements of *M* and *N* for all species. Unfortunately such data are not available for Tuesday Lake. See the discussion and conclusion subsection on *Errors in measurement*.

RESULTS

Regression Coefficients of Allometric Relations

Scatter plots of log *N* as a function of log *M* in 1984 (Figure 1A) and 1986 (Figure 1B) suggested that a linear relation is plausible. When log *N* was regressed as a linear function of log *M* (following Damuth, 1981) separately in 1984 and 1986, the point estimates of the slope coefficient were −0.83 and −0.74 (Table 1). The 99% confidence intervals for the slopes included −0.75 and excluded −1. Scatter plots of log *B* as a function of log *M* in 1984 (Figure 2A) and 1986 (Figure 2B) suggested no very clear, but perhaps an increasing, relation. When log *B* was regressed as a linear function of log *M* separately in 1984 and 1986, the point estimates of the slope coefficient were 0.17 and 0.26. Since log B = log M + log N, it follows mathematically, and is observed numerically, that $\beta_1 + 1 = \mu_1$ and that $\beta_2 + 1 = \nu_1$, using the notation of Table 1. Both 99% confidence intervals for the slopes included 0.25 and excluded 0. The regression coefficients of log *B* as a linear function of log rank(*M*) were significantly negative
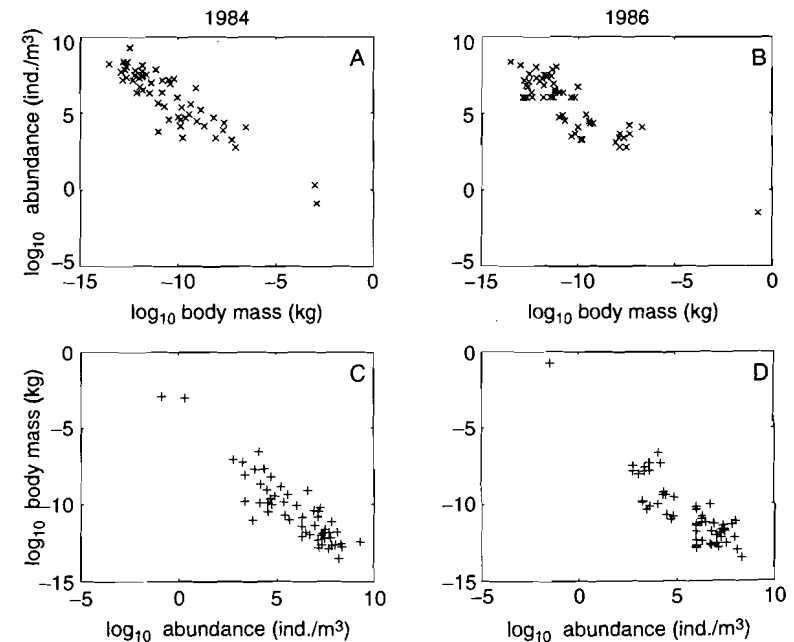

FIGURE 1 | Numerical abundance of species as a function of average body mass in Tuesday Lake, Michigan, in 1984 (**A**) and 1986 (**B**) and average body mass as a function of numerical abundance in 1984 (**C**) and 1986 (**D**).

Table 1: Slope coefficients of linear regressions between log numerical abundance log N and log average body mass log M of all species in Tuesday Lake in 1984 (56 species) and 1986 (57 species), and regressions of log biomass log B as a function of log M, log N, log rank(M), and log rank(N). The 99% confidence intervals for the slopes are in parentheses. The putative p value is the nominal statistical significance with which the data reject the null hypothesis that the slope is zero.

x	y	year	slope param- eter	point estimate (99% confidence interval)	r^2	putative p
log M	log N	1984	β_1	−0.8271 (−0.96, −0.70)	0.84	<0.001
log N	log M	1984	β_2	−1.0178 (−1.18, −0.86)	0.84	<0.001
log M	log N	1986	β_1	−0.7397 (−0.89, −0.59)	0.75	<0.001
log N	log M	1986	β_2	−1.0149 (−1.23, −0.80)	0.75	<0.001
log M	log B	1984	μ_1	0.1729 (0.04, 0.30)	0.19	<0.001
log N	log B	1984	ν_1	−0.0178 (−0.18, 0.14)	0.002	0.77
log M	log B	1986	μ_1	0.2603 (0.11, 0.41)	0.27	<0.001
log N	log B	1986	ν_1	−0.0149 (−0.23, 0.20)	0.001	0.85
log rank (M)	log B	1984	μ_2	−1.0762 (−1.89, −0.26)	0.19	<0.001
log rank (N)	log B	1984	ν_2	−0.2750 (−1.17, 0.62)	0.01	0.42
log rank (M)	log B	1986	μ_2	−1.4506 (−2.31, −0.59)	0.27	<0.001
log rank (N)	log B	1986	ν_2	−0.5807 (−1.57, 0.40)	0.04	0.12

(p < 0.001). These statistical results are all consistent with $N \propto M^{-3/4}$ and inconsistent with $N \propto M^{-1}$.

When the independent variable is changed from average body mass to numerical abundance, scatter plots of log M as a function of log N in 1984 (Figure 1C) and 1986 (Figure 1D) likewise suggested that a linear relation is plausible. However, when log M was regressed as a linear function of log N separately in 1984 and 1986, the point estimates of the slope coefficient were −1.02 and −1.01 (Table 1). The 99% confidence intervals for the slopes excluded −1.33 = 1/(−0.75) and included −1. Scatter plots of log B as a function of log N in 1984 (Figure 2C) and 1986 (Figure 2D) suggested no clear, but perhaps a decreasing, relation. When log B was regressed as a linear function of log N separately in 1984 and 1986, the point estimates of the slope coefficient were −0.02 and −0.01. Both 99% confidence intervals for the slopes included 0 and
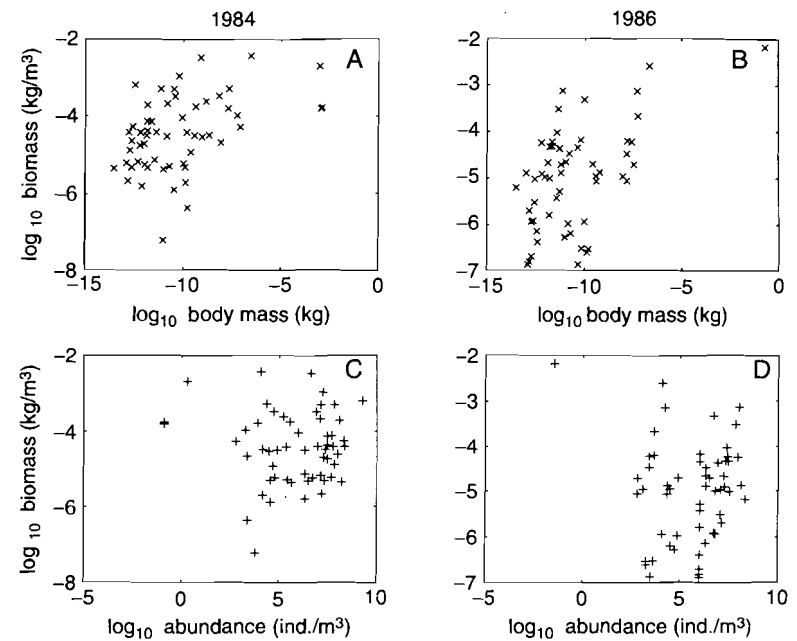
FIGURE 2 | Biomass of species as a function of average body mass in Tuesday Lake, Michigan, in 1984 (A) and 1986 (B) and biomass as a function of numerical abundance in 1984 (C) and 1986 (D).

excluded −1/4 and −1/3. The regression coefficients of log B as a linear function of log rank(N) were not significantly different from 0 (p > 0.12). These statistical results are all consistent with $N \propto M^{-1}$ and inconsistent with $N \propto M^{-3/4}$

The 99% confidence interval (−1.18, −0.86) for β_2, the 1984 slope of Eqn. 2, implies a confidence interval (1/−0.86, 1/−1.18) = (−1.16, −0.85) for the slope coefficient in a linear equation of the form of Eqn. 1 (though not for the least squares fit of Eqn. 1), and this interval intersects the 99% confidence interval (−0.96, −0.70) for β_1, the slope coefficient in Eqn. 1, in the range (−0.96, −0.85), which is incompatible with Damuth's estimate of slope of −3/4 and with constant biomass across species (slope −1). Similarly, for 1986, the 99% confidence interval (−1.23, −0.80) for the 1986 slope of Eqn. 2 implies a confidence interval (1/−0.80, 1/−1.23) = (−1.25, −0.81) which intersects the 99% confidence interval (−0.89, −0.59) for the slope coefficient in Eqn. 1 in the range (−0.89, −0.81), which is incompatible with Damuth's estimate of slope of −3/4 and with constant biomass across species (slope −1). If one hopes that the slope coefficients are

the same in 1984 and 1986, then the intersecting confidence intervals for the slope coefficient in Eqn. 1 in 1984 of (–0.96, –0.85) and in 1986 of (–0.89, –0.81) themselves intersect only in the range (–0.89, –0.85).

To resolve the contradiction between the apparent conclusions in the first two paragraphs of this section of Results regarding the slopes –3/4 and –1, we investigated whether the data conformed to the assumptions that justified the computed p values and confidence intervals.

Conformity of Data to Regression Assumptions

Quantile-quantile plots of the residuals vs. a standard normal distribution were examined visually. Discrepancies from a straight line were apparent in some cases. However, the Jarque-Bera and Lilliefors tests failed to reject the null hypothesis of composite normality for the distribution of the residuals even at the 0.05 level of significance in any of the regressions. In none of the regressions was there significant evidence even at the 0.05 level of significance for a linear trend or for serial correlation in the residuals. In these respects, the data did not deviate detectably from the assumptions underlying the linear regression model.

In other respects, the data rejected some assumptions underlying the linear regression model (see Table 2). Quadratic regressions showed that log M and log B were significantly nonlinear functions of log N: the coefficient of the quadratic term differed from 0 according to the classical test of statistical significance with $0.01 < p < 0.05$ in 1984 and $p < 0.01$ in 1986 (Table 2, column 4). The strong rejection of linearity for both models in 1986 was exactly confirmed by the Akaike information criterion: in these two regressions, and in no others in Table 2 (column 5), AIC_c for the linear model exceeded AIC_c for the quadratic model by more than 10. In short, classical hypothesis testing (with $p < 0.01$) and the Akaike information criterion (with $\Delta > 10$) agreed perfectly in identifying marked nonlinearities. When the evidence against linearity was weaker, as in 1984, the classic test and the Akaike information criterion agreed well though imperfectly. Excepting only the value of Δ in 1986 for the regression of log B as a function of log rank(N), shown in the last line of Table 2, all of the values of Δ between 2.27 and 10 identified models that the classic test rejected as nonlinear with $0.01 < p < 0.05$, and vice versa.

As the models in which log M and log B were quadratic functions of log N were markedly better in 1986 than the models in which log M and log B were linear functions of log N, it is of biological interest to report the estimated 99% confidence intervals of the coefficients of the quadratic models for 1986. Let $z = \log N - \text{mean}(\log N)$ be the centered log numerical abundance of each species. Then in 1986:

Table 2: Quantitative tests of assumptions underlying putative p values in Table 1. Cells left blank represent no statistically significant violation of the assumptions of linear regression due to the property in the column heading. Cells with + represent a statistically significant violation with probability between 0.01 and 0.05. Cells with ++ represent a statistically significant violation with probability < 0.01. Columns are omitted for the Jarque-Bera and Lilliefors tests of composite normality, for the presence of linear trend in the residuals, and for the presence of serial correlation in the residuals because none of these tests was significant at the 0.05 level. In columns 6 and 7, AIC_c is defined in Eqn. 3. In column 5, Δ is defined in Eqn. 4.

column 1	2	3	4	5	6	7	8	9	10
			classic non-linearity	Δ	AIC_c linear model	AIC_c quadratic model	residuals quadratic	absolute residuals linear	absolute residuals quadratic
x	y	year							
log M	log N	1984		2.26	–9.96	–7.71			
log N	log M	1984	+	2.28	1.65	–0.63	+		
log M	log N	1986		2.02	–2.77	–0.75			
log N	log M	1986	++	10.38	15.26	4.88	++	+	
log B	log M	1984		2.26	–9.96	–7.71			
log B	log N	1984	+	2.28	1.65	–0.63	+		
log M	log B	1986		2.02	–2.77	–0.75			
log N	log B	1986	++	10.38	15.26	4.88	++	+	
log rank(M)	log B	1984		2.09	–9.89	–7.80		+	
log rank(N)	log B	1984		1.65	1.05	2.70		+	
log rank(M)	log B	1986		0.42	–2.57	–2.99			
log rank(N)	log B	1986		2.32	12.78	15.10		++	+

$$\log M = (-11.29, -10.46) + (-1.07, -0.62)z + (+0.02, +0.15)z^2, \qquad (5)$$

$$\log B = (-5.70, -4.87) + (-0.07, +0.38)z + (+0.02, +0.15)z^2. \qquad (6)$$

Because $\log B = \log M + \log N$, it is expected mathematically (and observed numerically here) that the confidence intervals (and point parameter estimates) of the coefficients of the quadratic terms in Eqn. 5 and Eqn. 6 are identical, and that the confidence intervals (and point parameter estimates) of the coefficients of the linear terms in Eqn. 5 and Eqn. 6 differ by exactly 1. It is of more biological interest to compare the interval $(-1.07, -0.62)$ of the linear coefficient in Eqn. 5 with the interval $(-1.23, -0.80)$ of the linear coefficient of the corresponding *linear* model in Table 1, line 4: while both intervals include a slope estimate of -1, the latter interval excludes both $-2/3$ and $-3/4$ while the former interval (which incorporates the influence of the quadratic term) includes both $-2/3$ and $-3/4$. Similarly, the interval $(-0.07, +0.38)$ of the linear coefficient in Eqn. 6 includes both $1/3$ and $1/4$, while the interval $(-0.23, -0.20)$ of the linear coefficient of the corresponding *linear* model in Table 1, line 8, excludes both $1/3$ and $1/4$.

The residuals of the regressions of $\log M$ and $\log B$ as linear functions of $\log N$ had significant quadratic dependence on the predicted value of the linear regression, and therefore on $\log N$ (Table 2, column 8). In 1984 and 1986, linear regressions of $\log B$ on $\log \text{rank}(N)$ had absolute residuals that had a statistically significant linear dependence (column 9), and in 1986 a significant quadratic dependence (column 10), on the predicted value of the linear regression, and therefore on $\log \text{rank}(N)$. All 6 regressions in which the independent variable (x) was $\log N$ or $\log \text{rank}(N)$ displayed statistically significant deviations from the assumptions of linear regression, whereas none of the 6 regressions in which the independent variable (x) was $\log M$ or $\log \text{rank}(M)$ displayed any statistically significant deviation.

Artificial Example

The following artificial example emphasizes that the linear regression of y on x may satisfy the underlying assumptions while the linear regression of x on y may not. Suppose the random variable X took the values 1, 2, 3, or 4 each with probability $1/4$, and suppose the random variable Y took the values $1 + 10X + \varepsilon$, where each realization of ε was an independent normal random variable with mean 0 and variance 1. By construction, (X, Y) satisfied perfectly the assumptions of the linear

regression model: normality, linearity, homoscedasticity, independence, and exact knowledge of X. Figure 3A illustrates the distribution of (X, Y) with 100 values of ε, and therefore 100 values of Y, for each value of X. The quantile-quantile plot (see Figure 3B), the distribution of the (signed) residuals (see Figure 3C), and the distribution of the absolute residuals (see Figure 3D), as well as all of the quantitative statistics, were compatible with the assumptions of linear regression.

When X and Y were exchanged, the graph of the function (Figure 4A) approximated a step function. While the marginal distribution of the residuals was very nearly normal (see Figure 4B), neither the residuals nor the absolute residuals were serially independent for increasing values of the linear prediction (see Figure 4 C,D). The Durbin-Watson statistic easily detected these departures from serial independence.
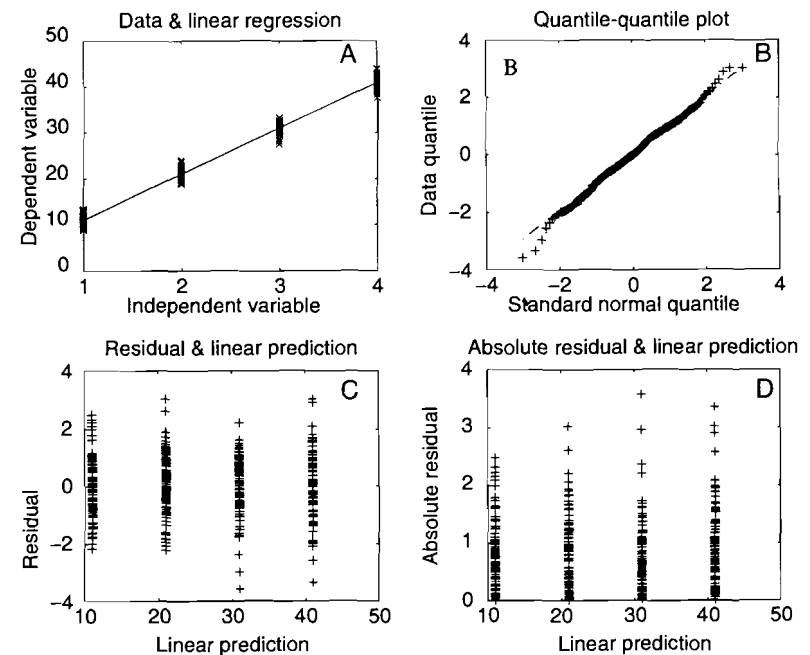

FIGURE 3 | A, Artificial data (X, Y) that satisfy the assumptions of the linear regression model regarding normality, linearity, homoscedasticity, independence and no error variance in x (see text for details). B, Quantile-quantile plot. C, Residuals as a function of the linear prediction. D, Absolute residuals as a function of the linear prediction.
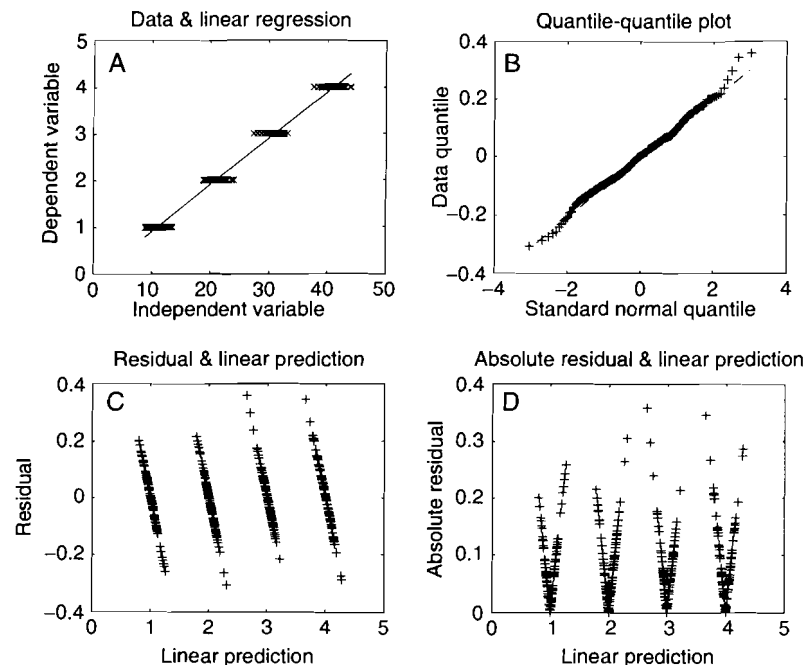
FIGURE 4 | A, The same artificial data as in Figure 3 but with axes exchanged: (Y, X) data do not satisfy linearity, homoscedasticity or independence. **B,** Quantile-quantile plot. **C,** Residuals as a function of the linear prediction. **D,** Absolute residuals as a function of the linear prediction.

DISCUSSION AND CONCLUSIONS

Principle Findings

As many have pointed out before, studies of the connection between body size and numerical abundance should carefully investigate whether the probability statements associated with linear regression are justified by the properties of the data being analyzed (Strathmann, 1990; Blackburn and Gaston, 1997, 1998, 1999; Russo et al., 2003; and references they cite). The present study may be the first to examine carefully both choices of independent variable, body size and numerical abundance, for a single set of data. These data include essentially all plant and animal species in a single pelagic community with a known food web. In linear regressions of the Tuesday Lake data, it is more defensible to use $\log M$ or $\log \text{rank}(M)$ as an independent variable and $\log N$ or $\log B$ as the dependent variable than it is to use $\log N$ or $\log \text{rank}(N)$ as an independent variable and $\log M$ or $\log B$ as the dependent variable. On the other

hand, using $\log N$ as the independent variable seems to be more sensitive at picking up a nonlinear relationship, if that nonlinearity is real. The generality of these conclusions for other webs remains to be determined.

Ecological Significance

Choosing an appropriate statistical model matters ecologically. For example, in Table 1, in the models where the underlying statistical assumptions were not rejected by the data, the 99% confidence intervals for the allometric exponent β_1 did not include -1, but rather lay above -1. Equivalently, the 99% confidence intervals for the allometric exponent μ_1 lay above 0. That is, in Tuesday Lake in 1984 and 1986, the bigger the average individual body mass of a biological species, the larger the biomass of that species, on the average.

The important conclusion that the biomass of biological species in Tuesday Lake increased with the average individual body mass of biological species was not evident in Table 1 when the independent variable was log N. However, in these cases the confidence intervals of the slopes were statistically unjustified. If unjustified models were to be used, an important conclusion about the upward trend of biomass with increasing average individual body mass of biological species in the Tuesday Lake pelagic food web would be lost.

Variation Among Studies

Among the many studies of the relation between numerical abundance and body size, Leaper and Raffaelli (1999, their Table 2) tabulated 31 estimates of the exponent b in the relation $N \propto M^{-b}$, in addition to two estimates of their own for the Ythan estuary. Kerr and Dickie (2001) reviewed many studies of biomass spectra, which are related through the frequency distribution of species according to body size. These studies differ in numerous significant respects, including: the units of observation, the universe of units of observation, the measurement of body size, and the choice of independent variable (body size or abundance). We consider each of these four differences in turn.

First, the units of observation have been diverse. Examples include: functional groups, such as detritus, phytoplankton, and macroalgae (Leaper and Raffaelli, 1999, p.192); taxonomic groups resolved as nearly as possible to biological species (Leaper and Raffaelli, 1999, p. 192, and the present study); size classes without regard to taxonomic identification (Jennings and Mackinson, 2003); and narrowly defined size classes within a mixture of species and broad taxonomic units, for example,

heterotrophic bacteria, *Prochlorococcus, Synechococcus*, and ultra- and nano-plankton (Rinaldo et al., 2002). Regressions of numerical abundance against body size using taxonomic units of observation are necessarily affected by the distribution of body size within the taxonomic units and may be biased by failure to sample some life history stages adequately.

Such regressions may differ from regressions of numerical abundance against body size using size classes as the units of observation. Leaper and Raffaelli (1999) showed that improving the taxonomic resolution among organisms of small body size substantially decreased the estimate of b in the Ythan estuary from 1.03 to 0.63, and observed that it was difficult to compare studies with differing degrees of taxonomic resolution. Russo et al. (2003) compared estimates of b obtained by using individual biological species, ecological guilds (frugivore-omnivore, granivore, insectivore, raptor), and groupings based on phylogenetic relatedness.

Second, the units of observation have been selected from different universes. Leaper and Raffaelli (1999, p. 192) usefully classify studies as direct observations of a local community; compilations of community data; global compilations; and compilations of data on one broad taxon, such as birds or terrestrial herbivorous mammals. Russo et al. (2003) analyzed direct observations of sympatric assemblages of tropical birds in three local rain forest habitats. The Tuesday Lake data were direct observations of phytoplankton, zooplankton, and fishes, excluding microbes and benthic and littoral organisms (Jonsson et al., 2005).

Third, how body size is measured has two aspects: how is the size of a single individual measured, and how is a composite indicator of size for the units of observation derived from the measurements for individuals? The size of a single individual may be measured by wet weight (this study), volume (Rinaldo et al., 2002), dry weight (Leaper and Raffaelli, 1999), length (converted to mass by an allometric relation between length and mass, a conversion that introduces additional error), or as a vector of weights of each component chemical element (e.g., by carbon and phosphorus content, Sterner and Elser, 2002). The composite indicator of size may be the average (this study, Enquist et al., 1998, p. 164), the "mean adult body-size" (Leaper and Raffaelli, 1999), or the maximum body size. Only the average body size has the convenient property that the species' biomass is the product of the numerical abundance times the average body mass. On the other hand, the average body size depends on the body size distribution within the unit of observation.

The other indicators of population body size, such as adult or maximum size, may not be well defined for species of indeterminate growth.

Fourth, which variable (body size or abundance) is viewed as independent may depend on the observer and the unit of observation. As previously noted, most plant ecologists have plotted log M on the ordinate and log N on the abscissa. Damuth (1981), Blackburn and Gaston (1999), Leaper and Raffaelli (1999) and most animal ecologists have reversed the assignment of variables to axes. If there were no errors of measurement in log M and log N and no stochastic variations from an exact allometric relation between M and N, the choice would be immaterial. In the presence of deviations from an allometric relation and error variance in measurements, the choice could substantially affect the estimated coefficients of the allometric relation.

Errors in Measurement

Given replicate measurements of M and N for all species, an explicit allometric model that allows for measurement errors in both variables M and N is:

$$\log M = \mu + \upsilon,$$
$$\log N = \eta + \omega,$$
$$0 = \beta_0 + \beta_1\mu + \beta_2\eta + \varepsilon,$$

where υ, ω, and ε are independent, normally distributed errors with mean zero, μ is the true but unobservable log M and η is the true but unobservable log N. One could estimate the regression coefficients β and the variances of υ, ω, and ε using Error-in-Variables regression (Clutton-Brock, 1967; Reilly and Patino-Leal, 1981). The estimates of the β_is will be sensitive to the variances var(υ) and var(ω). Alternative statistical models such as the general structural relation, reduced major axis regression and ordinary least squares, and further references, are discussed by Griffiths (1992, 1998) and Russo et al. (2003, p. 272), among others.

For pelagic ecosystems, measurements of M tend to have lower observational error than measurements of N. For subsets of communities similar to Tuesday Lake, S.R.C. has estimated the measurement error in M and N using replicates and found the error variance in N to be larger than the error variance in M, frequently by as much as two orders of magnitude. This difference in error variance results in part from the sampling design, which could allocate replicates to increase the precision of either M or N. For patchy or mobile species, it can be laborious to

increase the precision of N by sampling more intensively. Estimating N requires handling many individual organisms. With only the limited extra effort of measuring each individual's M, it is possible to build up a large number of replicates to increase the precision of measurement of M. Thus attempts to improve the precision of measurement of N may yield even further improvements in the precision of measurement of M.

It is tempting to speculate that the difference between animal and plant ecologists in which variable they choose as independent and plot on the abscissa may result from differences in the difficulty of measuring M versus N for different types of organisms. Trees (and most terrestrial plants) stand still to be counted but it is often not easy to estimate their mass. For example, considerable error may be involved in estimating tree mass from trunk diameter. On the other hand, the numerical abundance of animals is often difficult to measure precisely but measurements of body mass may be relatively easy to replicate, especially for species with determinate growth. Where the direction of causality is not obvious, perhaps animal and plant ecologists are both inclined to put the more precise measurement on the abscissa.

Causality

A linear regression model is not informative about causality because the direction and quantitative form of causation cannot be determined on statistical grounds from static observational data. The statistically defensible regression model in Eqn. 1 suggests but does not prove that the average individual body mass of a species determined the species' numerical abundance in Tuesday Lake in 1984 and 1986.

Damuth (1981) suggested that M determines N for energetic reasons. If M is fixed for each unit of observation (e.g., species, other taxonomic unit, or size class); and if the energy required to support each individual is aM^b; and if the energy available to support each unit of observation is a constant E; then the number N of individuals that can be supported by the available energy in each unit of observation is $E/(aM^b)$ (i.e., $N \propto M^{-b}$). The fixed amount of energy and the scaling of metabolic rate with body mass establish a relationship between M and N and could equally well argue for M as a function of N. Damuth's crucial assumption, for the present discussion, is to regard M as fixed and N as variable. This and other mechanisms are reviewed in greater detail by Blackburn and Gaston (1999).

An alternative causal scheme is that the abundance of a species is fixed by underlying birth and death rates, and that the average body size adjusts up or down according to the number of individuals who make a living

from a fixed supply of food or light or other essential input. Causation could run from numerical abundance to body mass in a nonlinear relation with a nonstandard error structure. Such an alternative causal model is not the simplest model, given our data, but it is also not logically impossible when species interact in a community web, as in Tuesday Lake.

The assumption that M is fixed for each unit of observation can be attacked at the level of the individual, at the level of the population, and at the level of the community.

At the level of the individual, density-dependent body growth is well known in plants, animals and other organisms. A large number of seedlings or young-of-year may, when resources are limited, result in reduced growth for most or all individuals. The mechanistic basis of the determination of individual body size is little understood (Hafen and Stocker, 2003).

At the level of the population (excluding those units of observation defined by size alone, such as a size class, for which size is necessarily fixed or bounded), a familiar feature of the demography of age-structured populations is that rapidly growing populations have a higher proportion of individuals in young age classes and a lower proportion of individuals in old age classes than stationary or declining age-structured populations with exactly the same life table (that is, holding the life table constant and varying only the intrinsic rate of natural increase or Malthusian parameter usually denoted by r, not to be confused with the residual r). In such a case, if body mass increases with age, then a more rapidly growing population (with a younger age structure) will have smaller average body mass than a stationary or declining population, even if the life tables, fertility schedules, and age-specific growth schedules are identical in the populations being compared. Roff (1986, p. 317) argued that "although an increase in body size increases fecundity and tends to increase [the Malthusian parameter] r, the concomitant increase in development time decreases r"; the observed distribution of body sizes in a population results from the balance between these countervailing selective forces and has, in some instances, been demonstrated to be under genetic control.

Changes in the size distribution within a population influence, not only the average body size, but also the average energy consumption per unit of mass if younger, faster-growing individuals have higher respiration rates per unit of mass than older, slower-growing individuals, as indicated by Riisgård (1998). If the average energy consumption per unit of mass influences a species' numerical abundance (as in the energetic model of Damuth), then population age-structure or size-structure influences both M and N and may affect the estimated allometric exponent b.

At the level of the community, trophic, competitive and mutualistic interactions may affect the average body size of a population. For example, predators (including human harvesters) may preferentially remove vulnerable small individuals or large trophy individuals. In Tuesday Lake, all but a handful of species were linked by the food web and were therefore subject to influence by resource species or consumer species. Even the half dozen isolated species were subject to competition (for light or nutrients, for example) that could affect average body size.

Some ecologists view a species' average body mass as determined on an evolutionary time scale and its numerical abundance as responding on an ecological time scale. However, a fuller view suggests that a species' average body mass is determined on at least three time scales: ontogenetic (by the way an individual's genes guide its development in interaction with its proximate environment), ecological (by how individuals interact with other individuals of the same and different species and with the abiotic environment), and evolutionary (by how individuals' selective advantages, resulting from their ontogeny and ecological interactions, contribute to the heritable characteristics of the next generation). The balance among ontogenetic, ecological, and evolutionary influences on a species' average body mass varies among species and, for a given species, according to the biotic and abiotic surroundings of the species.

Body size (individually and on the average) and numerical abundance interact dynamically, on multiple time scales, within and between species or other units of observation, and through biotic interaction with the abiotic environment. The Tuesday Lake data demonstrate, once again, that statistical care is required in assessing even the simplest pattern generated by this dynamic process.

ACKNOWLEDGMENTS

Species' average body mass and numerical abundance in a community food web:

statistical questions in estimating the relationship

JOEL E. COHEN[1] and STEPHEN R. CARPENTER[2]

[1]Laboratory of Populations, Rockefeller & Columbia Universities, 1230 York Avenue, Box 20, New York, NY 10021, U.S.A., e-mail: cohen@rockefeller.edu
[2]Center for Limnology, University of Wisconsin, 680 North Park Street, Madison, Wisconsin 53706, srcarpen@facstaff.wisc.edu

In: *Dynamic Food Webs: Multispecies Assemblages, Ecosystem Development, and Environmental Change*, edited by Peter C. de Ruiter, Volkmar Wolters, John C. Moore. Academic Press, 2005. (Proceedings of the Food Web 2003 Giessen conference)

References

Blackburn, T.M. and Gaston, K.J., 1997. A critical assessment of the form of the interspecific relationship between abundance and body size in animals. Journal of Animal Ecology 66, 233-249.

Blackburn, T.M. and Gaston, K.J., 1998. Some methodological issues in macroecology. American Naturalist 151, 68-83.

Blackburn, T.M. and Gaston, K.J., 1999. The relationship between animal abundance and body size: a review of the mechanisms. Advances in Ecological Research 28, 181-210.

Brown, James H., 1995. Macroecology. University of Chicago Press, Chicago and London.

Carpenter, S.R. and Kitchell, J.F., 1993. The trophic cascade in lakes. Cambridge University Press, Cambridge, UK.

Clutton-Brock, M., 1967. Likelihood distributions for estimating functions when both variables are subject to error. Technometrics 9, 261-269.

Cohen, J.E., Jonsson, T. and Carpenter, S., 2003. Ecological community description using the food web, species abundance, and body size. PNAS 100, 1781-1786.

Colinvaux, P., 1978. Why Big Fierce Animals Are Rare: An Ecologist's Perspective. Princeton University Press, Princeton.

Cyr, H., Downing, J.A. and Peters, R.H., 1997a. Density-body size relationships in local aquatic communities. Oikos 79, 333-346.

Cyr, H., Peters, R.H. and Downing, J.A., 1997b. Population density and community size structure: comparison of aquatic and terrestrial systems. Oikos 80, 139-149.

Damuth, J., 1981. Population density and body size in mammals. Nature 290, 699-700.

Enquist, B.J., Brown, J.H. and West, G.B., 1998. Allometric scaling of plant energetics and population density. Nature 395, 163-165.

Griffiths, D. 1992. Size, abundance, and energy use in communities. Journal of Animal Ecology 61, 307-316.

Griffiths, D. 1998. Sampling effort, regression method, and the shape and slope of size-abundance relations. Journal of Animal Ecology 67, 795-804.

Hafen, E. and Stocker, H., 2003. How are the sizes of cells, organ, and bodies controlled? PLoS Biology 1, 319-323.

Jennings, S. and Mackinson, S., 2003. Abundance-body mass relationships in size-structured food webs. Ecology Letters 6, 971-974.

Jonsson, T., Cohen, J.E. and Carpenter, S.R., In press. Food webs, body size and species abundance in ecological community description. Advances in Ecological Research.

Kerr, S.R. and Dickie, L.M., 2001. The Biomass Spectrum: A Predator-Prey Theory of Aquatic Production. Columbia University Press, New York.

Leaper, R. and Raffaelli, D., 1999. Defining the abundance body-size constraint space: data from a real food web. Ecology Letters 2, 191-199.

Peters, R.H., 1983. The Ecological Implications of Body Size. Cambridge University Press, New York.

Peters, R.H. and Raelson, J.V., 1984. Relations between individual size and mammalian population density. American Naturalist 124, 498-517.

Peters, R.H. and Wassenberg, K., 1983. The effect of body size on animal abundance. Oecologia 60, 89-96.

Platt, T. and Denman, K., 1977. Organisation in the pelagic ecosystem. Helgoländer Meeresunters 30, 575-581.

Platt, T. and Denman, K., 1978. The structure of pelagic marine ecosystems. Rapports et Procés-Verbaux des Réunions, Conseil International pour l'Exploration de la Mer 173, 60-65.

Reilly, P.M. and Patino-Leal, H., 1981. A Bayesian study of the error-in-variables model. Technometrics 23, 221-231.

Reuman, D.C. and Cohen, J.E., 2004a. Trophic links' length and slope in the Tuesday Lake food web with species' body mass and numerical abundance. Journal of Animal Ecology 73, 852-866.

Reuman, D.C. and Cohen, J.E., 2004b. Estimating relative energy fluxes using the food web, species abundance, and body size. Advances in Ecological Research.

Riisgård, H.U., 1998. No foundation of a "3/4 power scaling law" for respiration in biology. Ecology Letters 1, 71-73.

Rinaldo, A., Maritan, A., Cavender-Bares, K.K. and Chisholm, S.W., 2002. Cross-scale ecological dynamics and microbial size spectra in marine ecosystems. Proceedings of the Royal Society London Series B 269, 2051-2059.

Roff, D.A., 1986. Predicting body size with life history models. BioScience 36, 316-323.

Russo, S.E., Robinson, S.K. and Terborgh, J., 2003. Size-abundance relationships in an Amazonian bird community: implications for the energetic equivalence rule. American Naturalist 161, 267-283.

Sterner, R.W. and Elser, J.J., 2002. Ecological Stoichiometry: The Biology of Elements from Molecules to the Biosphere. Princeton University Press, Princeton.

Strathmann, R.R., 1990. Testing size-abundance rules in a human exclusion experiment. Science 250, 1091.

Stuart, A. and Ord, J.K., 1991. Kendall's Advanced Theory of Statistics Vol. 1 and 2. 5th Ed. Oxford University Press.