Ecology Letters, (2008) **11**: 1216–1228 doi: 10.1111/j.1461-0248.2008.01236.x

LETTER

Three allometric relations of population density to body mass: theoretical integration and empirical tests in 149 food webs

Daniel C. Reuman,[1*†] Christian Mulder,[2] Dave Raffaelli[3] and Joel E. Cohen[1]

[1]*Laboratory of Populations, Rockefeller University, 1230 York Avenue, Box 20, New York, NY 10065, USA*

[2]*Department of Ecology, RIVM, 9 Antonie van Leeuwenhoeklaan, Box 1, Bilthoven 3720 BA, The Netherlands*

[3]*Environment Department, University of York, York YO10 5DD, UK*

**Correspondence: E-mail: d.reuman@imperial.ac.uk*

[†]*From September 2007: Imperial College London, Silwood Park Campus, Buckhurst Road, Ascot, Berkshire SL5 7PY, UK.*

Abstract

Predicting species population density–body mass scaling in community food webs (henceforth *webs*) is important for conservation and to understand community structure. Very different types of scaling have been studied, based on either individuals or species. The *individual size distribution* (ISD) describes the distribution of individual-organism body masses regardless of taxonomy, and contains the same information as the abundance spectrum. Focusing instead on species, the *local size–density relationship* (LSDR) plots population densities vs. mean body masses of species. The distribution of species mean body masses (the *species-mean-size distribution*, SMSD) is also important but previously little studied in webs. We here combine and formalize theory of several authors to predict: how these three descriptions are related; the forms of the LSDR and ISD; and variation in scaling among webs. We describe empirically the SMSDs of two pelagic, one estuarine, and 146 soil webs by power laws and generalizations. We test theory and find it broadly validated.

Keywords

Allometry, biodiversity, biomass spectrum, body mass, metabolism, numerical abundance, population density, power law, size spectrum, truncated Pareto distribution.

Ecology Letters (2008) 11: 1216–1228

INTRODUCTION

In macroecology, species population density and body-mass-specific taxonomic diversity vary with body mass according to power laws. The scaling of population densities (N) against average body masses ($\bar{M}$) of species from a single broad clade like birds or mammals often follows a power law, $N = a\bar{M}^b$, and this fact has played a central role in macroecology (Marquet *et al.* 2005; White *et al.* 2007). The approximate scaling exponent $b = -3/4$ has been supported theoretically and empirically (Damuth 1981; Peters 1983; Nee *et al.* 1991; West *et al.* 1997; Brown *et al.* 2004). Body-mass-specific taxonomic diversity is also a power law of body mass: if species data from a clade, gathered globally or regionally, are grouped by average body mass in bins of equal width on a log M axis, then the number of species, D, in each bin depends on the central bin M values by a power law (Marquet *et al.* 2005; see also Niklas *et al.* 2003; who use linear-scale uniform bins). The exponent $b = -3/4$ has been theoretically predicted for this power law as well (Marquet *et al.* 2005), but older datasets and different theory support $b = -2/3$ for large-mass categories and non-power-law behaviour for small-mass categories.

By contrast with studies of a single broad clade, whether at local, regional or global spatial scales, local community food webs (henceforth *webs*) comprise all organisms occurring in a location. In webs, power law relationships among population density, taxonomic diversity and body mass are connected to each other and to ecological processes differently from the ways they are connected in macroecology. These connections, while important to the structure of webs, are incompletely understood. Therefore we here analyse scaling in webs.

Population density–body mass scaling has been studied intensively for decades in webs using plots of log total density in log(M) bins (often called abundance spectra; Kerr & Dickie 2001). Such plots approximate log frequency distributions (log probability density functions) of

© 2008 Blackwell Publishing Ltd/CNRS

individual-organism log body masses, ignoring taxonomy. The notation M denotes individual body mass or body mass as a general variable, whereas $\bar{M}$ will be used for the mean body mass of a taxon. The *individual size distribution* (ISD) is the probability density function (pdf) of individual-organism body masses (not log body masses), again ignoring taxonomy. The ISD is a transformed abundance spectrum and captures the same information in different form. It is a power law with exponent λ if and only if the abundance spectrum is linear with slope $\lambda + 1$ (Theory; Methods; Andersen & Beyer 2006; White *et al.* 2008).

A very different type of web population density–body mass scaling results when the unit is a taxon, rather than an individual organism: scatter plots of N-vs.-$\bar{M}$ or log(N)-vs.-log($\bar{M}$) for web taxa (here called *local size-density relationships* (White *et al.* 2007), or LSDRs) have recently been studied, sometimes with axes exchanged (Marquet *et al.* 1990; Cyr *et al.* 1997; Leaper & Raffaelli 1999; Cyr 2000; Schmid *et al.* 2000; Cohen *et al.* 2003; Reuman & Cohen 2004; Jonsson *et al.* 2005; Mulder *et al.* 2005a; Reuman & Cohen 2005; Woodward *et al.* 2005; Long *et al.* 2006; Reuman *et al.* 2008). The conceptual and empirical differences between the LSDR and the ISD were recognized by Damuth (1994), Jonsson *et al.* (2005), Jennings *et al.* (2007), White *et al.* (2007) and others. Both relationships were often but not always approximate power laws for local community webs, with exponents that differed from each other and that varied widely by web (Cyr *et al.* 1997; Cyr 2000; Jonsson *et al.* 2005; Jennings *et al.* 2007; Reuman *et al.* 2008). Power-law exponents of the LSDR and ISD are here called the LSDRE and ISDE.

The frequency distribution of mean taxon body masses in a web, here called the *species-mean-size distribution*, or SMSD, differs from the LSDR and the ISD. Little work has tested the power law hypothesis for web SMSDs, examined SMSD exponents (here called SMSDEs), or made clear the relationships among the SMSD, the ISD, and the LSDR in webs (but see Jonsson *et al.* 2005; Rossberg *et al.* 2008).

LSDRs and SMSDs in webs have been examined empirically less often than ISDs, but are just as important. Conservation of species requires understanding relative species population densities, which are revealed by the LSDR, rather than relative population densities of size categories, which are revealed by the ISD. Theory predicts a power law for web ISDs, and also for N-vs.-$\bar{M}$ relationships for webs aggregated to trophic levels, and gives a formula for the ISDE (Cyr 2000; Brown & Gillooly 2003; Brown *et al.* 2004). Predictions of the theory were verified with terrestrial arthropods and a marine pelagic system (Jennings & Mackinson 2003; Meehan 2006a,b; Meehan *et al.* 2006). The theory can be adapted to explain the LSDR if the SMSD is a power law and the SMSDE is known (Damuth 1994; Cyr 2000; Jonsson *et al.* 2005). Unsurprisingly, theory for the ISD (Brown & Gillooly 2003; Brown *et al.* 2004) does not apply directly to the LSDR. Empirical LSDREs vary over a wider range than that predicted for ISDEs (Reuman *et al.* 2008).

A broad understanding of body-mass allometry requires theory to describe and data to test the relationships among the three types of allometry defined above (White *et al.* 2007). We here unify, formalize, and test theory to predict the form of the LSDR and the ISD and the relationship among them and the SMSD.

To help the reader, we give idealized examples. Consider webs with a set B of basal species, a set I of intermediate species, and a set T of top species. Each basal species has average body mass $\bar{M} = 1$ mg, each intermediate species has $\bar{M} = 1 \times 10^4$ mg = 10 g, and each top species has $\bar{M} = 1 \times 10^8$ mg = 100 kg. Assume that population production is $P_j \propto N_j \bar{M}_j^{3/4}$ for species j (Peters 1983; Brown *et al.* 2004); that all species in each trophic level have about the same log(N) (i.e. individuals are equally numerous in each species at a given level); that body mass variation within species is much smaller than variation among species; and that each trophic level except basal species consumes all the production of the level immediately below it and converts 10% into production available to the next trophic level, if any. Setting quotients of total productions of successive trophic levels equal to 10% implies

$$\frac{\bar{M}_{B_l}^{3/4} \sum_l N_{B_l}}{\bar{M}_{I_m}^{3/4} \sum_m N_{I_m}} = \frac{\bar{M}_{I_m}^{3/4} \sum_m N_{I_m}}{\bar{M}_{T_n}^{3/4} \sum_n N_{T_n}} = 10. \quad (1)$$

The quotients of masses are equal to 1/10 000; replacing these and computing logs gives

$$\begin{aligned} &\log\Big(\sum_l N_{B_l}\Big) - \log\Big(\sum_m N_{I_m}\Big) \\ &= \log\Big(\sum_m N_{I_m}\Big) - \log\Big(\sum_n N_{T_n}\Big) = 4. \end{aligned} \quad (2)$$

If the number D of species in each level is the same (e.g. one species per level, Fig. 1a), then eqn 2 implies that the regression slope for the log(N)-vs.-log($\bar{M}$) scatter plot is −1; this is the LSDRE. Separating species into three equal bins by log($\bar{M}$), computing log total population density for each bin, and computing the regression slope of these quantities against the log(M) values at the bin centres also gives −1, an estimate of the abundance spectrum and therefore of ISDE + 1. The log(D)-vs.-log(M) scatter plot (using three bins) is flat (regression slope 0), so SMSDE + 1 = 0 (Theory; Methods) and LSDRE = (ISDE + 1)− (SMSDE + 1) holds.

If the number of species decreases (respectively, increases) by an order of magnitude for each trophic level going up (e.g. 100 basal, 10 intermediate, and one top

© 2008 Blackwell Publishing Ltd/CNRS

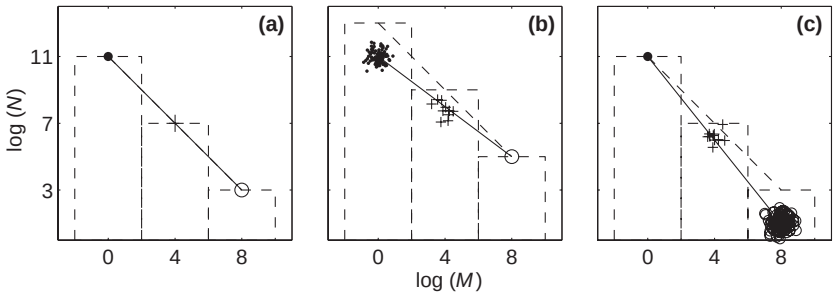


Figure 1 Illustrative relationships among the local size-density relationship (LSDR; solid lines), the abundance spectrum (dashed lines; the abundance spectrum slope is the individual size distribution exponent plus 1, or ISDE + 1), and the species-mean-size distribution (SMSD) in three hypothetical food webs. Dots are basal species (1 in a, c; 100 in b); plus signs are intermediate species (1 in a; 10 in b, c); open circles are top species (1 in a, b; 100 in c). LSDR exponents (LSDREs) are −1 (a), −3/4 (b), −5/4 (c), whereas all abundance spectrum slopes are −1 (so ISDEs are −2). SMSD exponents (SMSDEs; not shown) are 0 (a), −1/4 (b), and 1/4 (c). In all panels, LSDRE = (ISDE + 1)−(SMSDE + 1) = ISDE−SMSDE.

species, Fig. 1b; respectively, one basal, 10 intermediate, and 100 top, Fig. 1c), then eqn 2 implies that the log(N)-vs.-log($\bar{M}$) regression slope, equal to the LSDRE, is −3/4 (respectively, −5/4). The bin-specific log-density-vs.-log(M) slope, equal to ISDE + 1, is always −1 and the bin-specific log(D)-vs.-log(M) slope, equal to SMSDE + 1, is −1/4 (respectively, +1/4). In all cases considered, LSDRE = (ISDE + 1)−(SMSDE + 1) = ISDE−SMSDE. In these examples and in general, LSDRE and the relative population densities of species of different average body sizes can be predicted from the other exponents.

We here formalize the theory illustrated above, synthesizing, elaborating and clarifying ideas of Damuth (1994), Cyr (2000), Brown & Gillooly (2003), Brown *et al.* (2004), Jonsson *et al.* (2005) and others. In addition, for the first time, we empirically characterize SMSD from a diverse collection of food webs as power laws. We test our unified theory systematically against 146 soil webs from Dutch agroecosystems, two pelagic webs from Tuesday Lake, Michigan and the Ythan Estuary web, Scotland, empirically unifying three types of allometry for webs under one framework.

THEORY

Theory for individual size distributions

We explain and elaborate the theory of Brown *et al.* (2004). See also Cyr (2000), Brown & Gillooly (2003), Jennings & Mackinson (2003), Jonsson *et al.* (2005). Denote the ISD by $f_I(M)$. We do not assume a particular form for $f_I(M)$. Ecosystem boundaries and physiological limitations ensure that $f_I(M) = 0$ for M outside some interval $[a_I, b_I]$ with $0 < a_I < b_I$. Denote the trophic transfer efficiency or Lindeman efficiency of the web by α, and denote the average consumer-to-resource body-mass ratio by β; both are assumed constant or not systematically varying with M. We assume $\beta > 1$, i.e. that bigger organisms eat smaller organisms. Population production of organisms in the mass range $R = [M, M + \mathrm{d}M]$, defined as the total growth biomass plus the total reproduction biomass of these organisms per unit time, is approximately $\mathrm{Prod}(M)\mathrm{d}M$, where $\mathrm{Prod}(M) \propto f_I(M)M^{3/4}$ (Peters 1983; West *et al.* 1997; Brown *et al.* 2004). This production is distributed across the range of body masses in a way that depends on the life histories of individuals of mass M. Production in the form of individual growth will occur within or slightly above the mass range R. Production in the form of reproduction will occur at or near species minimum masses for the species with reproductive individuals in R. We assume that production $\mathrm{Prod}(M)\mathrm{d}M$ occurs mainly within the mass range R, or close to it relative to the consumer–resource mass ratio β. This assumption is more likely to be valid if most production takes the form of growth, or if variation in individual body masses within species is small, so that the minimum sizes of most species are relatively close to their mean body sizes (Discussion).

Denote by $g_I(u)$ the pdf of $u = \log(M)$, so that $\log(g_I(u))$ is the abundance spectrum. Then

$$g_I(u) = \ln(10) \times f_I(M) \times M \qquad (3)$$

(Appendix S1; Andersen & Beyer 2006; White *et al.* 2008). Log denotes logarithm base 10 and ln denotes natural logarithm.

Because β does not vary systematically with M, individuals in the body mass range R eat mainly individuals in the range $R/\beta = [M/\beta, M/\beta + \mathrm{d}M/\beta]$. Because population production $\mathrm{Prod}(M/\beta)\mathrm{d}M/\beta$ of the mass category R/β occurs mainly within that mass category or close to it, by making the further assumption that all available production is consumed we set the production of R equal to α times the production of R/β:

© 2008 Blackwell Publishing Ltd/CNRS

$$M^{3/4} f_I(M)\mathrm{d}M = \alpha\left(\frac{M}{\beta}\right)^{3/4} f_I\left(\frac{M}{\beta}\right)\frac{\mathrm{d}M}{\beta}. \quad (4)$$

Metabolism and production also depend on temperature, but for the webs of this study, all organisms had body temperature equal to ambient temperature (although ambient temperature differed by web). If temperature factors were included in eqn 4 they would cancel under plausible assumptions. Eqn 4 reduces (Appendix S1) to

$$g_I(u) \propto 10^{\left(u\left(\frac{\log(\alpha)}{\log(\beta)}-\frac{3}{4}\right)\right)} = M^{\left(\frac{\log(\alpha)}{\log(\beta)}-\frac{3}{4}\right)}, \quad (5)$$

or equivalently (see eqn 3),

$$f_I(M) \propto M^{\frac{\ln(\alpha)}{\ln(\beta)}-\frac{7}{4}}. \quad (6)$$

Constants of proportionality are determined by the requirement that pdfs integrate to 1. The function that is a power law distribution $f_{\mathrm{tp}} \propto M^{\lambda_{\mathrm{tp}}}$ on $[a_{\mathrm{tp}}, b_{\mathrm{tp}}]$ and is $f_{\mathrm{tp}} = 0$ outside $[a_{\mathrm{tp}}, b_{\mathrm{tp}}]$ (where λ_{tp} can be any real number and $0 < a_{\mathrm{tp}} < b_{\mathrm{tp}}$) is called a truncated Pareto distribution, and is the predicted form of the ISD. The predicted ISDE is therefore

$$\lambda_I = \frac{\log(\alpha)}{\log(\beta)} - \frac{7}{4}. \quad (7)$$

Eqns 5–7 were empirically supported for systems aggregated to trophic levels (Meehan 2006a,b; Meehan *et al.* 2006) and for a marine fishery (Jennings & Mackinson 2003).

Theory for the local size-density relationship

Denote by $f_S(\bar{M})$ the pdf of species average body masses, which must be zero outside some interval $[a_S, b_S]$. We will demonstrate empirically that

$$f_S(\bar{M}) \propto \bar{M}^{\lambda_S} \quad (8)$$

in $[a_S, b_S]$ and 0 elsewhere is often an excellent approximation for web data, although we present no theoretical explanation for this fact.

The population density or number of organisms in the mass range R is proportional to $f_I(M) \times \mathrm{d}M$ and those organisms are partitioned into a number of species proportional to $f_S(M) \times \mathrm{d}M$. If eqn 8 holds, then the expected density of a species of average mass $\bar{M}$ is predicted to scale approximately as

$$N \propto \frac{f_I}{f_S} \propto \bar{M}^{\lambda_I - \lambda_S} \quad (9)$$

if the range of individual body masses within species is not too large relative to the range of body masses in the web (Damuth 1994; Jonsson *et al.* 2005). Thus, the LSDR is predicted to follow a power law with LSDRE

$$\begin{aligned}\lambda_L &= \lambda_I - \lambda_S = \frac{\log(\alpha)}{\log(\beta)} - \frac{7}{4} - \lambda_S \\ &= \left(\frac{\log(\alpha)}{\log(\beta)} - \frac{3}{4}\right) - (\lambda_S + 1) \\ &= (\lambda_I + 1) - (\lambda_S + 1)\end{aligned} \quad (10)$$

whenever f_S follows a truncated Pareto distribution. Derivations do not depend on any theory that claims to explain why metabolic rates scale as $M^{3/4}$, but only on the phenomenology that metabolic rates scale as $M^{3/4}$.

Ecological interpretations of theoretical predictions

We consider limiting cases of eqns 5–7, 9 and 10 to link theoretical predictions to ecological intuition. The term −3/4 in eqn 10 corresponds to the benchmark of energetic equivalence of species. The energetic equivalence hypothesis (Damuth 1981) assumes that each species extracts about the same amount of energy from the environment, and predicts from this assumption that species N will depend on $\bar{M}$ by a power law with exponent −3/4. Energetic equivalence holds for our model if $\alpha = 1$, corresponding to efficient trophic transfer of energy, and if about the same number of species occur in each pair of body mass ranges $\bar{R} = [\bar{M}_1, \bar{M}_2]$ and $\bar{R}/\beta = [\bar{M}_1/\beta, \bar{M}_2/\beta]$. Because most individuals of species with $\bar{M}$ in $\bar{R}$ feed primarily on individuals of species with $\bar{M}$ in $\bar{R}/\beta$, fewer (respectively, more) species in $\bar{R}$ than in $\bar{R}/\beta$ will mean more (respectively, less) energy available per species in $\bar{R}$ than in $\bar{R}/\beta$, given efficient trophic transfer. These assumptions correspond, respectively, to $\log(\alpha) = 0$ and $\lambda_S = -1$, so that eqn 10 predicts $\lambda_L = -3/4$ as does the energetic equivalence hypothesis.

Corrections to this baseline case correspond to other terms in eqn 10. The term $\log(\alpha)/\log(\beta)$, a correction for the inefficiency of trophic transfer of energy, is negative: less efficient trophic transfer (smaller α) contributes to a more negative LSDRE (a steeper LSDR and less abundant large-$\bar{M}$ species relative to small-$\bar{M}$ species; Cyr 2000). The effects of inefficient trophic transfer on LSDR are magnified for webs with low β. Such webs have many inefficient trophic links over a fixed range of $\bar{M}$.

The term $\lambda_I + 1$ of eqn 10 is a correction for power-law increases or decreases in species diversity with $\log(M)$. If diversity decreases with increasing $\log(M)$, this term may cancel $\log(\alpha)/\log(\beta)$ so that energetic equivalence holds and LSDRE = −3/4 even if trophic transfer is inefficient. In that case, the reduced energy available to large-$\bar{M}$ top predators is partitioned among fewer species, so that the total energy available per species is the same as for basal species. LSDRE is determined by the competing influences of trophic transfer losses and changes in species diversity

© 2008 Blackwell Publishing Ltd/CNRS

with $\bar{M}$. LSDRE can be greater than, less than, or equal to −3/4, whereas ISDE + 1 is not greater than −3/4. The unit of analysis for the LSDR and SMSD theory presented here has been species. The analysis applies equally to species and to higher taxonomic levels.

METHODS

Testing theory

Theory predicts that: (1) ISD is a power law; (2) LSDR is a power law whenever the SMSD is a power law; (3) ISDE + 1 = $\log(\alpha)/\log(\beta)-3/4$ (eqn 7); (4) LSDRE + SMSDE + 1 = $\log(\alpha)/\log(\beta)-3/4$ (eqn 10). We tested (1) and (2) by comparing the power law hypotheses against more general alternative models.

To test (3) and (4), ISDE + 1 and LSDRE + SMSDE + 1 were estimated, with confidence intervals, and results were compared with $\log(\alpha)/\log(\beta)-3/4$. For the Tuesday Lake and Ythan Estuary webs, which have highly resolved trophic data, $\log(\beta)$ was the mean of $\log(\bar{M}_c/\bar{M}_r)$ over all trophic links, where $\bar{M}_c$ and $\bar{M}_r$ denote the mean body masses of the consumer and resource taxa (Appendix S6). If the confidence intervals of ISDE + 1 (respectively, LSDRE + SMSDE + 1) contained $\log(\alpha)/\log(\beta)-3/4$ for α in a reasonable range (10–30%), prediction (3) [respectively, (4)] was not rejected. Highly resolved trophic data were unavailable for the soil webs, so $\log(\alpha)/\log(\beta)$ was unknown, but was very likely to be negative. Prediction (3) [respectively, (4)] was rejected for a soil web only if the confidence intervals of ISDE + 1 (respectively, LSDRE + SMSDE + 1) lay entirely above −3/4.

Statistical estimators of exponents and hypothesis tests

Data for each system included a list of taxa and the mean body mass ($\bar{M}$) and population density (N) of each taxon. For each web separately, we tested the hypothesis that the LSDR was a power law and estimated the exponent by fitting the models $\log(N) = b \times \log(\bar{M}) + a$ and $\log(N) = e \times \log(\bar{M})^2 + d \times \log(\bar{M}) + c$ using ordinary least squares regression. The models were compared with an F-test. If the latter model explained significantly more variation in $\log(N)$ (1% level), the power law form of the LSDR was rejected. The slope b of the linear model was the LSDRE.

If $f_S(\bar{M}) \propto \bar{M}^{\lambda_S}$ on $[a_S, b_S]$ and 0 elsewhere, then the pdf of $\bar{u} = \log(\bar{M})$ is

$$g_S(\bar{u}) \propto 10^{\bar{u}(\lambda_S+1)} \tag{11}$$

on $[\log(a_S), \log(b_S)]$ and 0 elsewhere (Appendix S1). The log of this expression can be called the *diversity spectrum*. The exponent in eqn 11 is linear in $\bar{u}$. A quadratic generalization of the truncated Pareto distribution has pdf

$$G_S(\bar{u}) \propto 10^{\gamma_S \times \bar{u} + \eta_S \times \bar{u}^2} \tag{12}$$

on $[\log(a_S), \log(b_S)]$ and 0 elsewhere. For each web separately, we fitted the truncated Pareto distribution and this generalization using maximum likelihood, the method recommended by White *et al.* (2008) (see also Aban *et al.* 2006). Simple numeric optimization was required (Appendices S2 and S3). The two models were compared by the likelihood ratio test. If eqn 12 was a significantly better fit (1% level), the power law form of the SMSD was rejected.

If $\eta_S < 0$, G_S is a normal distribution in $\bar{u}$ with standard deviation $\sigma^2 = -1/(2 \times \ln(10) \times \eta_S)$ and mean $\mu = \sigma^2 \times \ln(10) \times \gamma_S$, truncated at $\log(a_S)$ and $\log(b_S)$. We therefore compared the truncated Pareto distribution to an alternative that includes a (truncated) log-normal distribution, often used to characterize body mass distributions.

The maximum likelihood estimator of λ_S was biased for the sample sizes of this study (n = 30 to 96 taxa), so we developed a bias-corrected estimator of λ_S for comparisons with theory. Simulations indicated that the new estimator was essentially unbiased for $n \geq 30$ and for λ_S in the range of our webs (Appendix S2).

Confidence intervals for LSDRE, SMSDE, and LSDRE + SMSDE + 1 were estimated by resampling. For each web, 5000 resamplings with replacement were taken from the empirical joint (N, $\bar{M}$) distribution of the web. For each resampling, as many points were selected as taxa in the web. LSDRE, SMSDE, and LSDRE + SMSDE + 1 were calculated for each resampling. These confidence intervals of LSDRE + SMSDE + 1 captured any covariance of LSDRE and SMSDE. Confidence intervals of LSDRE and SMSDE could have been calculated without resampling, but resampling was the simplest way to get confidence intervals for the sum.

Individual-organism M data were unavailable for the webs of this study, so the ISDE could not be estimated by a method with a strong probabilistic foundation without making precise assumptions about how individual-organism M was distributed within taxa. Instead, we used our assumption, revisited in the Discussion, that individual-organism M did not vary too much within taxa relative to the range of M values in a whole web. ISDE + 1 was then estimated by a commonly used method based on separating data into bins of uniform width on a log scale. Each web's range of $\log(\bar{M})$ values was divided evenly into 10 bins, and the log total population density of all organisms with $\log(\bar{M})$ in each bin was regressed against log-scale bin centers. ISDE + 1 was the slope of this regression. Confidence intervals of ISDE + 1 were derived from this regression. The power-law form of the ISD was rejected only if linearity of this regression was rejected. This estimator of ISDE + 1 may have been biased toward 0 (Appendix S5).

© 2008 Blackwell Publishing Ltd/CNRS

For comparison with our bin-based estimator, we adapted to the context of the available data a method recommended by White *et al.* (2008) based on fitting the theoretical cumulative distribution function (cdf) of a truncated Pareto distribution to an empirically estimated cdf (see also Johnson *et al.* 1994, p. 580). Results using the cdf-based method were qualitatively very similar to those from the binning method (Appendix S5). We focused primarily on bin-based results in the main text because those methods were more intuitive, and because results were similar for both methods. ISD results were approximate but sufficient for testing the theory, also approximate.

Data

All webs and methods by which data were gathered are described in Appendix S6 and (Hall & Raffaelli 1991; Carpenter & Kitchell 1993; Cohen *et al.* 2003; Mulder *et al.* 2003; Jonsson *et al.* 2005; Mulder *et al.* 2005a,b). The 146 soil webs were sampled from five types of active-management farms (organic, conventional, intensive and super-intensive farms, and pastures), as well as from winter farms (arable fields untreated at the time of sampling) and unmanaged Scots pine plantations (here called forests) in the Netherlands.

General methods

Computations used Matlab version 7.4.0.287 (R2007a) and the Matlab optimization toolbox (The Mathworks, Inc., Natick, MA, USA). We set 1% significance to detect flagrant violations of theory.

RESULTS

Individual size distributions

Theory predicted a truncated Pareto ISD. This hypothesis could not be rejected for our webs, with a few exceptions. A truncated Pareto distribution was statistically rejected using bin-based methods only for the Ythan Estuary, and using cdf-fitting methods (Appendix S5) only for the Ythan Estuary and two forests (Table S1). For the Ythan Estuary, however, the best-fitting generalized distribution was almost indistinguishable from a truncated Pareto (Fig. 2c); the difference was considered unimportant and was ignored. A truncated Pareto distribution was visually a poor description of the ISDs of organic farms ID 222–232 (e.g. Fig. 2e), although it could not be rejected statistically, probably only because power was low.

Excluding organic farms ID 222–232, which probably had non-truncated-Pareto ISDs, only forests had ISDEs that falsified the quantitative predictions of theory (eqn 7). For the Tuesday Lake and Ythan Estuary webs, for reasonable values of α, both bin-based and cdf-fitting confidence intervals of ISDE + 1 contained $\log(\alpha)/\log(\beta)$ $-3/4$ (Fig. 3, Fig. S6). For all soil webs except five forests, bin-based confidence intervals of ISDE + 1 contained values less than $-3/4$. For all soil webs except 13 forests and super-intensive farm ID 157, cdf-fitting confidence intervals of ISDE + 1 contained values less than $-3/4$ (Table S1). Mean ISDE + 1 values by ecosystem type were less than or close to $-3/4$, except for forests (Fig. 4), using either bin-based or cdf-fitting estimates. We argue in Appendix S5 that our bin-based estimator of ISDE + 1 may have been biased toward 0, so true values of ISDE + 1 may have been smaller than those reported in Fig. 4.

Species-mean-size distributions

All non-forest webs and one forest (ID 169) had truncated Pareto SMSD. For ID 169 and all non-forests except super-intensive farm ID 158 and pasture ID 192, the null-hypothesis that the SMSD was truncated Pareto could not be rejected in favour of the generalized distribution of eqn 12 (Figs 2 and 4; Table S2). The best-fitting generalized distribution for IDs 158 and 192 did not differ much from a truncated Pareto distribution; the difference was considered unimportant and was ignored. Truncated Pareto distributions were excellent descriptions of SMSDs for non-forests and ID 169 (e.g. Fig. 2).

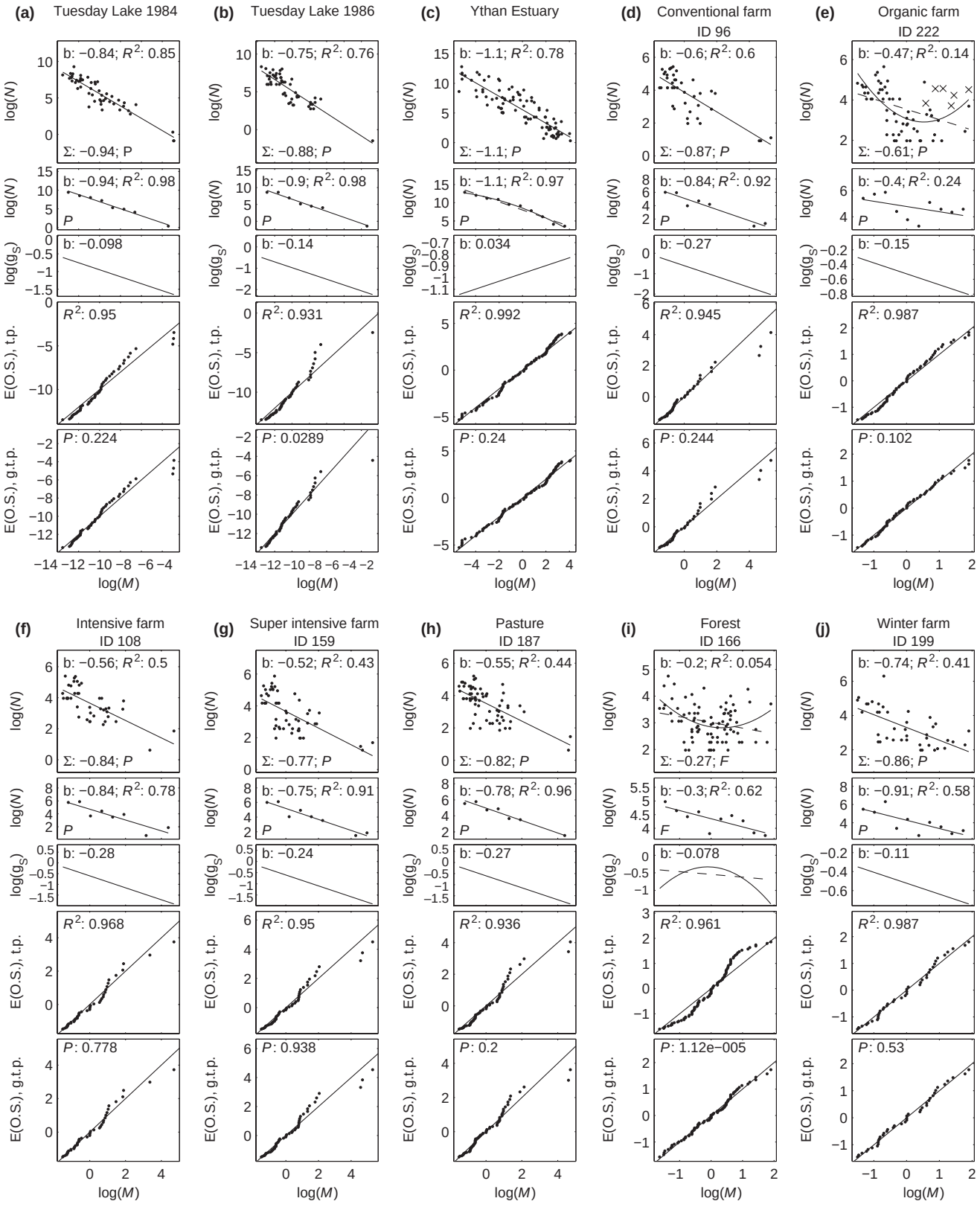
In contrast, all forests except ID 169 had truncated log normal SMSD. The null-hypothesis of a truncated Pareto SMSD was rejected for those webs (Table S2). The best-fitting generalized distributions differed substantially from a truncated Pareto (Fig. 2), and had negative quadratic coefficient, implying a truncated normal distribution. The generalized distribution was an excellent description of all SMSDs of this study (e.g. Fig. 2). Quadratic coefficients of the generalized distribution were never significantly positive for any web.

All truncated Pareto SMSDs had SMSDE < -1 (corresponding to decreasing taxonomic diversity in log-scale-uniform bins of increasing $\log(\bar{M})$), except the Ythan Estuary and forest ID 169. Truncated Pareto SMSDs had SMSDEs that differed by ecosystem type (Figs 2, 4; Table S2).

Local size–density relationships

Power-law LSDR could not be statistically rejected for our webs, except for organic farms ID 222–232, four forests, three winter farms, and super-intensive farm ID 158 (Figs 2 and 4; Table S2; Reuman *et al.* 2008). The departure from power law behaviour of farm ID 158 was very minor and was ignored. Three of four forests with non-power-law LSDR had non-truncated-Pareto SMSD (Table S2); these webs do not falsify the prediction of theory that the LSDR

© 2008 Blackwell Publishing Ltd/CNRS



© 2008 Blackwell Publishing Ltd/CNRS

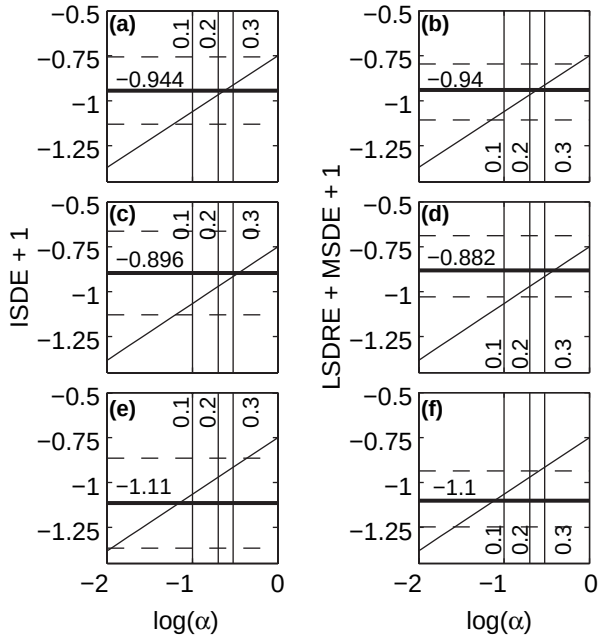


Figure 3 Tests of eqn 7 (a, c, e) and eqn 10 (b, d, f) using data of Tuesday Lake 1984 (a, b) and 1986 (c, d), and the Ythan Estuary (e, f). Heavy solid horizontal lines are point estimates of ISDE + 1 (a, c, e) or LSDRE + SMSDE + 1 (b, d, f). Dashed horizontal lines are 99% confidence intervals of these values. Diagonal lines are $\log(\alpha)/\log(\beta) - 3/4$, predicted by eqns 7 and 10 to equal *y*-axis values for the true value of $\log(\alpha)$. Predicted values fell within confidence intervals and close to point estimates for reasonable values of α. Vertical lines are labelled with benchmark α-values.

will be a power law because that prediction presupposes that the SMSD is a power law. Organic farms ID 222–232 had non-power law LSDR because enchytraeids (potworms) were more abundant than would have been expected from a power law (Fig. 2). Enchytraeid populations in those farms were promoted by organic fertilizer inputs (Mulder 2006; Mulder *et al.* 2006; Reuman *et al.* 2008). We cannot explain why one forest (ID 169) and three winter farms (IDs 201, 213, 217) had non-power-law LSDR.

All webs that had power-law LSDR and SMSD had confidence intervals of LSDRE + SMSDE + 1 containing theoretically predicted or possible values, except one. Tuesday Lake and the Ythan Estuary had confidence intervals containing $\log(\alpha)/\log(\beta) - 3/4$ for reasonable values of α (Fig. 3). Of soil webs with power-law SMSD and LSDR, only super-intensive farm ID 157 had confidence intervals entirely above $-3/4$ (Table S2). Means by ecosystem type of LSDRE + SMSDE + 1 were less than or very close to $-3/4$ (Fig. 4).

LSDRE < ISDE + 1 was predicted when SMSDE + 1 > 0, and LSDRE > ISDE + 1 was predicted when SMSDE + 1 < 0 (eqn 10). SMSDE + 1 was positive only for the Ythan Estuary and forest ID 169, for which LSDRE < ISDE + 1 held. For 139 of the 147 webs with SMSDE + 1 < 0, LSDRE > ISDE + 1 held. Of the 8 remaining webs, six were organic farms (ID 222–225, 228, 232) and two were winter farms (ID 212 and 236). The correction SMSDE + 1 for systematic changes in diversity with M clearly contributed to variation in LSDRE in the webs of this study.

We tested to what degree the model LSDRE = ISDE-SMSDE explained variation in data, compared with the null models LSDRE = k [where k was estimated from data as k = mean(LSDRE)] and LSDRE = ISDE + 1 (which corresponds to omitting the correction for systematic changes in diversity). Using all 149 sites, the sum of squares of LSDRE − mean(LSDRE), here called SS, was 3.10, the sum of squares of LSDRE − (ISDE + 1) was 7.57, and the sum of squares of LSDRE − ISDE + SMSDE, here called SSE, was 1.25. The model LSDRE = ISDE − SMSDE therefore explained 1−SSE/SS = 59.6% of the variation in LSDRE, although it had 0 degrees of freedom, less than the model LSDRE = k (which had one degree of freedom).

Taxonomic resolution

The SMSD and LSDR are supposed to describe allometric relationships for species, but only the Tuesday Lake data

Figure 2 One site from each of the ecosystem categories (see titles). Straight lines in top panels are ordinary least squares (OLS) regressions through taxa (the LSDR). Second panels show log total population density in $\log(M)$ bins and abundance spectra (slope = ISDE+1) as estimated by OLS regression. Third panels show the estimated log probability density function (pdf) of log taxon mean body masses (the taxonomic diversity spectrum; eqn 11), which has slope s = SMSDE+1; the estimated SMSD is proportional to $\bar{M}^{s-1}$. Straight lines in the top three panels are dashed and accompanied by the appropriate best-fitting quadratic model if that model was a significantly better fit (1% level; Methods). Σ is LSDRE + SMSDE + 1, predicted by theory to equal $\log(\alpha)/\log(\beta) - 3/4$ (eqn 10). The slope of the line in each panel is b; R^2 is the squared correlation coefficient. Fourth and fifth panels are probability plots comparing data to, respectively, the best-fitting truncated-Pareto (tp) SMSD and the best-fitting quadratic generalized truncated-Pareto (gtp; Methods). They show the expected values of the order statistics (O.S.) of $\log(\bar{M})$ under the best-fitting distribution (tp and gtp, respectively) vs. ranked taxon log mean body masses. *P*-values on bottom panels are from a likelihood ratio test comparison of the tp and gtp SMSDs. An *F* or *P* appears in the lower left of the top panel (respectively, second panel) according to whether eqn 7 (respectively, eqn 10) was falsified by each web or not (see Methods, Testing theory). In the organic farm (E), enchytraeids were plotted with x symbols.

© 2008 Blackwell Publishing Ltd/CNRS

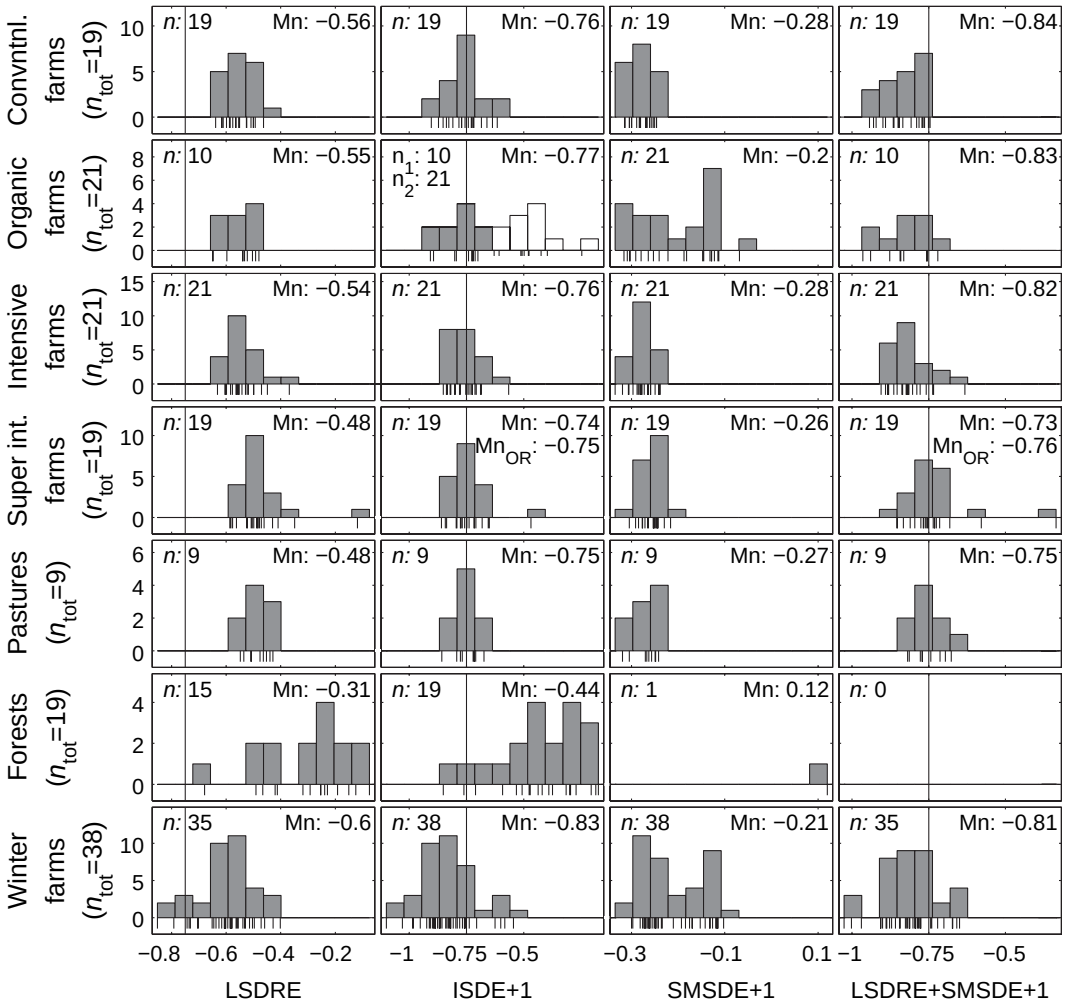


Figure 4 Histograms of scaling exponents for soil webs. Tick marks below the horizontal axis are individual web exponents. For the LSDRE, all webs were included for which a power law form for the LSDR was not statistically rejected (Methods). The super-intensive farm ID 158 was also included because its departure from power-law behaviour was minor. ISDs could not be statistically distinguished from power laws for soil webs using bin-based methods (Methods), but for some organic farms (ID 222–232) a power law was visually a poor description (e.g. Fig. 2e). These farms also had non-power law LSDRs. They are distinguished with open histogram bars and smaller marks. The mean (Mn) of ISDE + 1 does not include them. ISDE histograms use bin-based estimates (but see Fig. S6). SMSDEs were included whenever a power law SMSD could not be statistically rejected in favour of a generalized distribution (Methods). The super-intensive farm ID 158 and the pasture ID 192 were also included because the best-fitting generalized distribution, while a better fit, differed little from a power law. Histograms of LSDRE + SMSDE + 1 included sites for which power laws described LSDR and SMSD, as well as farms ID 158 and ID 192. Mn_{OR} denotes the mean with the right-most outlier point removed. Vertical lines demarcate −3/4. The number *n* in each panel is the number of sites included. The total number of sites of each type is in parentheses at left.

were uniformly resolved to species. To investigate whether imperfect taxonomic resolution could have caused inaccuracies in results based on the other webs of this study, we artificially lumped Tuesday Lake data to several levels of resolution, recomputed exponents and retested theory with lumped webs (unpublished results, Reuman, D.C., and Schittler, D.N.). For Tuesday Lake, all exponents and theories were insensitive to various kinds of lumping (including even and uneven lumping across the body mass range of the system) down to a surprisingly low level of resolution, much lower than the resolution of the other webs of this study.

DISCUSSION

We synthesized, elaborated, and formalized theoretical ideas of Damuth (1994), Cyr (2000), Brown & Gillooly (2003), Brown *et al.* (2004), Jonsson *et al.* (2005) and others on body

© 2008 Blackwell Publishing Ltd/CNRS

mass allometry in local community food webs and systematically tested theory for a large and diverse collection of webs. Theory was broadly validated. Exceptions provided additional ecological insight. Although many theories explained relative population densities or biomasses of trophic levels or size categories in assemblages or webs (Kerr 1974; Silvert & Platt 1980; Lurie *et al.* 1983; Borgmann 1987; Han & Straskraba 2001; Andersen & Beyer 2006; Pope *et al.* 2006), the integrated theory tested here predicted relative population densities of both mass categories and highly resolved taxa and explained web-to-web variation in relative densities. The theory required SMSD as input. We provided the first power-law description of SMSDs for a large collection of webs. Our results integrated three types of allometry theoretically and empirically.

Theory revisited

The theory of Brown *et al.* (2004) assumed that species could be aggregated into distinct trophic levels not overlapping in body mass; then the population production of each level was distributed exclusively over a range of body masses vulnerable to predation from the next level, making it natural to set α times the production of one trophic level equal to the production of the next higher trophic level. To avoid assuming trophic levels, we assumed that production of a mass category $R = [M, M + \mathrm{d}M]$ was distributed mainly over nearby (relative to the average consumer-resource mass ratio β) parts of the web's range of body masses; we then related the production of consumers and resources by $\mathrm{Prod}(R) = \alpha\mathrm{Prod}(R/\beta)$ (eqn 4). If only a fraction, ρ, of $\mathrm{Prod}(R/\beta)$ (e.g. the growth component) had body mass vulnerable to consumption by individuals in R, one could set $\mathrm{Prod}(R) = \alpha \times \rho \times \mathrm{Prod}(R/\beta)$, but further modification would be needed to reflect production from other mass categories occurring in R/β, for instance if the size at birth or hatching of a larger species fell in R/β.

An elaborated theory of ISD could use mass-dependent production kernels and feeding kernels, so that both production and feeding of organisms in R would be distributed over a range. Existing diet and life-history data and theory may suffice to construct such a model for a marine system (e.g. Charnov & Gillooly 2004; Andersen & Beyer 2006). The assumptions of the ISD theory (eqn 6) and Brown *et al.* (2004) probably do not hold for marine and other systems dominated by indeterminate growers. Indeterminate growers (e.g. fish) can pass though multiple trophic levels as they grow, and the investment of large individuals in reproduction (which represents production in small mass categories) can be larger than their investment in growth. Describing production and feeding kernels for organisms in R in a way that depends only on M and not on the species composition of R may also be more tractable for marine systems, which are heavily size-structured. The theory of Brown *et al.* (2004) has been supported by data from a marine system (Jennings & Mackinson 2003), though assumptions of the theory were violated. Other mechanisms make the same predictions for marine systems (e.g. Andersen & Beyer 2006).

For estimates of the LSDR, we assumed that the variation in individual-organism M within taxa was much smaller than among taxa. That assumption is far from true for systems dominated by indeterminate growers, so the LSDR theory is also not expected to apply to those systems. For marine and other systems, models should be developed and compared with data (currently rare) that explicitly relate M distributions within taxa to M distributions in whole webs and distributions of characteristic taxon body masses. Andersen & Beyer (2006) take important steps in this direction starting from different hypothesized mechanisms.

Our theory was useful when it predicted accurately because it illuminated potential mechanisms and when it failed because it provided a benchmark from which to measure deviations. Theoretical predictions about the ISD were within the likely margin of error of available ISD estimators, except for forests and some organic farms. Violations of theory by some organic farms were explained by subsidies, but forests had higher ISDEs than predicted, for unknown reasons. How were observed population densities of large-M organisms in soils under pine-plantation supported? Some form of external subsidy may have played a role, but fertilizers were not added. Other possible explanations included increasing trophic transfer efficiency or increasing consumer–resource body-mass ratios with increasing M (Brose *et al.* 2006) or a web not primarily body-mass structured. All of these possibilities violated model assumptions, and may need to be included in future modelling. In addition, the theory of this study predicted only the expected dependence of taxon densities N on mean body mass $\bar{M}$. Understanding the variation of individual taxa from the trend remains a challenge.

Species-mean-size distributions

In our data, SMSDs were often truncated Pareto distributions with SMSDE < -1, corresponding to a linear diversity spectrum of slope SMSDE $+ 1 < 0$. In related but distinct macroecological research, Hutchinson & MacArthur (1959) and May (1978) theorized that the global or regional diversity spectrum of a clade should have an approximately linear right tail with slope $-2/3$. Blackburn & Gaston (1994) and Loder *et al.* (1997) supported rough linearity for the right tail with slopes varying across clades and scales. Marquet *et al.* (2005) found linearity across all size bins of a diversity spectrum for South American mammals, with slope about $-3/4$. Histograms of $\log(\bar{M})$ of mammal species from

© 2008 Blackwell Publishing Ltd/CNRS

local communities were approximately uniform, corresponding to flat diversity spectra (Brown & Nicoletto 1991). Our study differed from the above studies and provided novel results by examining all or most taxa from local communities, rather than clade-specific datasets. Theoretical explanations for web diversity spectra and variation in web SMSDEs should be developed, perhaps through models that draw species from clade-specific regional pools described by macroecological diversity spectra.

Some conclusions of Jonsson *et al.* (2005) about the SMSD of Tuesday Lake agreed with our study; others differed superficially because some methods of Jonsson *et al.* (2005) were less probabilistically refined. Jonsson *et al.* (2005) confirmed for Tuesday Lake the prediction of Cohen (1991) that if $\bar{M}_i$ is the mean mass of the ith-heaviest taxon, then $\bar{M}_i \propto i^{\beta}$. For Pareto-distributed $\bar{M}_i$, this is expected, so their results support ours. A falloff from the trend $\bar{M}_i \propto i^{\beta}$ at large $\bar{M}_i$ is expected for truncated-Pareto $\bar{M}_i$; Jonsson *et al.* (2005) observed a falloff. By visual inspection of a histogram of $\log(\bar{M})$ values, Jonsson *et al.* (2005) suggested that a log-hyperbolic distribution of $\bar{M}$ may better account for large values than a log-normal distribution, but here we showed that a truncated Pareto distribution can do the same.

ACKNOWLEDGEMENTS

The authors thank Simon Jennings, Ethan White, Russ Lande, and an anonymous referee for helpful criticisms. D.C.R. and J.E.C. thank Priscilla K. Rogerson for assistance, and Daniella Schittler for help with the lumping results summarized here. J.E.C. thanks Mr and Mrs W.T. Golden and family for hospitality. D.C.R. and J.E.C. were partially supported by United States National Science Foundation grant DMS 0443803. C.M. was supported by the Netherlands Ministry of Housing, Spatial Planning, and Environment (VROM).

REFERENCES

Aban, I.B., Meerschaert, M.M. & Panorska, A.K. (2006). Parameter estimation for the truncated Pareto distribution. *J. Am. Stat. Assoc.*, 101, 270–277.

Andersen, K.H. & Beyer, J.E. (2006). Asymptotic size determines species abundance in the marine size spectrum. *Am. Nat.*, 168, 54–61.

Blackburn, T.M. & Gaston, K.J. (1994). Animal body size distributions: patterns, mechanisms and implications. *Trends Ecol. Evol.*, 9, 471–474.

Borgmann, U. (1987). Models of the slope of, and biomass flow up, the biomass size spectrum. *Can. J. Fish. Aquat. Sci.*, 44, 136–140.

Brose, U., Jonsson, T., Berlow, E.L., Warren, P., Banašek-Richter, C., Bersier, L.F. *et al.* (2006). Consumer–resource body-size relationships in natural food webs. *Ecology*, 87, 2411–2417.

Brown, J.H. & Gillooly, J.F. (2003). Ecological food webs: high-quality data facilitate theoretical unification. *Proc. Natl Acad. Sci. USA*, 100, 1467–1468.

Brown, J.H. & Nicoletto, P.F. (1991). Spatial scaling of species composition: body masses of North American land mammals. *Am. Nat.*, 138, 1478–1512.

Brown, J.H., Gillooly, J.F., Allen, A.P., Savage, V.M. & West, G.B. (2004). Toward a metabolic theory of ecology. *Ecology*, 85, 1771–1789.

Carpenter, S.R. & Kitchell, J.F. (1993). *The Trophic Cascade in Lakes*. Cambridge University Press, Cambridge.

Charnov, E.L. & Gillooly, J.F. (2004). Size and temperature in the evolution of fish life histories. *Integr. Comp. Biol.*, 44, 494–497.

Cohen, J.E. (1991). Food webs as a focus for unifying ecological theory. *Ecol. Int. Bull.*, 19, 1–13.

Cohen, J.E., Jonsson, T. & Carpenter, S.R. (2003). Ecological community description using the food web, species abundance, and body size. *Proc. Natl Acad. Sci. USA*, 100, 1781–1786.

Cyr, H. (2000). The allometry of population density and interannual variability. In: *Scaling in Biology* (eds Brown, J.H. & West, G.B.). Oxford University Press, Oxford, pp. 267–295.

Cyr, H., Downing, J.A. & Peters, R.H. (1997). Density body size relationships in local aquatic communities. *Oikos*, 79, 333–346.

Damuth, J. (1981). Population density and body size in mammals. *Nature*, 290, 699–700.

Damuth, J. (1994). No conflict among abundance rules. *Trends Ecol. Evol.*, 9, 487–487.

Hall, S.J. & Raffaelli, D. (1991). Food web patterns: lessons from a species rich web. *J. Anim. Ecol.*, 60, 823–842.

Han, B.-P. & Straskraba, M. (2001). Size dependence of biomass spectra and abundance spectra: the optimal distributions. *Ecol. Modell.*, 145, 175–187.

Hutchinson, G.E. & MacArthur, R.H. (1959). A theoretical ecological model of size distributions among species of animals. *Am. Nat.*, 93, 117–125.

Jennings, S. & Mackinson, S. (2003). Abundance–body mass relationships in size-structured food webs. *Ecol. Lett.*, 6, 971–974.

Jennings, S., d'Oliveira, J.A.A. & Warr, K.J. (2007). Measurement of body size and abundance in tests of macroecological and food web theory. *J. Anim. Ecol.*, 76, 72–82.

Johnson, N.L., Kotz, S. & Balakrishnan, N. (1994). *Continuous Univariate Distributions*, 2nd edn. John Wiley & Sons Inc., New York.

Jonsson, T., Cohen, J.E. & Carpenter, S.R. (2005). Food webs, body size, and species abundance in ecological community description. *Adv. Ecol. Res.*, 36, 1–84.

Kerr, S.R. (1974). Theory of size distributions in ecological communities. *J. Fish. Res. Board Can.*, 31, 1859–1862.

Kerr, S.R. & Dickie, L.M. (2001). *The Biomass Spectrum: A Predator–Prey Theory of Aquatic Production*. Columbia University Press, New York.

Leaper, R. & Raffaelli, D. (1999). Defining the abundance body-size constraint space: data from a real food web. *Ecol. Lett.*, 2, 191–199.

Loder, N., Blackburn, T.M. & Gaston, K.J. (1997). The slippery slope: towards an understanding of the body size frequency distribution. *Oikos*, 78, 195–201.

Long, Z.T., Steiner, C.F., Krumins, J.A. & Morin, P.J. (2006). Species richness and allometric scaling jointly determine biomass in model aquatic food webs. *J. Anim. Ecol.*, 75, 1014–1023.

© 2008 Blackwell Publishing Ltd/CNRS

Lurie, D., Valls, J. & Wagensberg, J. (1983). Thermodynamic approach to biomass distribution in ecological systems. *Bull. Math. Biol.*, 45, 869–872.

Marquet, P.A., Navarrete, S.A. & Castilla, J.C. (1990). Scaling population-density to body size in rocky intertidal communities. *Science*, 250, 1125–1127.

Marquet, P.A., Quinones, R.A., Abades, S., Labra, F., Tognelli, M., Arim, M. *et al.* (2005). Scaling and power-laws in ecological systems. *J. Exp. Biol.*, 208, 1749–1769.

May, R.M. (1978). The dynamics and diversity of insect faunas. In: *Diversity of Insect Faunas* (eds Mound, L.A. & Waloff, N.). Blackwell, Oxford, pp. 188–204.

Meehan, T.D. (2006a). Energy use and animal abundance in litter and soil communities. *Ecology*, 87, 1650–1658.

Meehan, T.D. (2006b). Mass and temperature dependence of metabolic rate in litter and soil invertebrates. *Physiol. Biochem. Zool.*, 79, 878–884.

Meehan, T.D., Drumm, P.K., Farrar, R.S., Oral, K., Lanier, K.E., Pennington, E.A. *et al.* (2006). Energetic equivalence in a soil arthropod community from an aspen-conifer forest. *Pedobiologia*, 50, 307–312.

Mulder, C. (2006). Driving forces from soil invertebrates to ecosystem functioning: the allometric perspective. *Naturwissenschaften*, 93, 467–479.

Mulder, C., De Zwart, D., Van Wijnen, H.J., Schouten, A.J. & Breure, A.M. (2003). Observational and simulated evidence of ecological shifts within the soil nematode community of agroecosystems under conventional and organic farming. *Funct. Ecol.*, 17, 516–525.

Mulder, C., Cohen, J.E., Setälä, H., Bloem, J. & Breure, A.M. (2005a). Bacterial traits, organism mass, and numerical abundance in the detrital soil food web of Dutch agricultural grasslands. *Ecol. Lett.*, 8, 80–90.

Mulder, C., Van Wijnen, H.J. & Van Wezel, A.P. (2005b). Numerical abundance and biodiversity of below-ground taxocenes along a pH gradient across the Netherlands. *J. Biogeogr.*, 32, 1775–1790.

Mulder, C., Den Hollander, H., Schouten, T. & Rutgers, M. (2006). Allometry, biocomplexity, and web topology of hundred agro-environments in The Netherlands. *Ecol. Complex.*, 3, 219–230.

Nee, S., Read, A.F., Greenwood, J.J.D. & Harvey, P.H. (1991). The relationship between abundance and body size in British birds. *Nature*, 351, 312–313.

Niklas, K.J., Midgley, J.J. & Rand, R.H. (2003). Size-dependent species richness: trends within plant communities and across latitude. *Ecol. Lett.*, 6, 631–636.

Peters, R.H. (1983). *The Ecological Implications of Body Size*. Cambridge University Press, Cambridge.

Pope, J.G., Rice, J.C., Dann, N., Jennings, S. & Gislason, H. (2006). Modelling an exploited marine fish community with 15 parameters – results from a simple size-based model. *ICES J. Mar. Sci.*, 63, 1029–1044.

Reuman, D.C. & Cohen, J.E. (2004). Trophic links' length and slope in the Tuesday Lake food web with species' body mass and numerical abundance. *J. Anim. Ecol.*, 73, 852–866.

Reuman, D.C. & Cohen, J.E. (2005). Estimating relative energy fluxes using the food web, species abundance, and body size. *Adv. Ecol. Res.*, 36, 137–182.

Reuman, D.C., Mulder, C., Banašek-Richter, C., Cattin Blandenier, M.-F., Breure, A.M., Den Hollander, H. *et al.* (2008). Allometry of body size and abundance in 166 food webs. *Adv. Ecol. Res.*

Rossberg, A.G., Ishii, R., Amemiya, T. & Itoh, K. (2008). The top down mechanism for body-mass abundance scaling. *Ecology*, 89, 567–580.

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

Silvert, W. & Platt, T. (1980). Dynamic energy-flow model of the particle-size distribution in pelagic ecosystems. In: *Evolution and Ecology of Zooplankton Communities* (ed. Kerfoot, W.C.). The University Press of New England, Hanover, N.H., pp. 754–763.

West, G.B., Brown, J.H. & Enquist, B.J. (1997). A general model for the origin of allometric scaling laws in biology. *Science*, 276, 122–126.

White, E.P., Ernest, S.K.M., Kerkhoff, A.J. & Enquist, B.J. (2007). Relationships between body size and abundance in ecology. *Trends Ecol. Evol.*, 22, 323–330.

White, E.P., Enquist, B.J. & Green, J.L. (2008). On estimating the exponent of power-law frequency distributions. *Ecology*, 89, 905–912.

Woodward, G., Speirs, D.C. & Hildrew, A.G. (2005). Quantification and resolution of a complex, size-structured food web. *Adv. Ecol. Res.*, 36, 85–135.

SUPPORTING INFORMATION

Additional supporting information may be found in the online version of this article:

Figure S1 Bias and variability of parameter estimates from fitting Pareto and power function distributions to data from a truncated Pareto distribution.

Figure S2 Bias, standard deviation, and mean squared error of the MLE of the exponent λ_{tp} of the truncated Pareto distribution and the bias-corrected MLE of the same exponent.

Figure S3 Bias, standard deviation, and mean squared error of an estimate of the exponent λ_{tp} of the truncated Pareto distribution, using 10 bins uniform on a logarithmic scale.

Figure S4 A straightforward adaptation of cdf fitting to the case where only taxon means $\bar{M}$ and population densities N are available, compared to a generalized method.

Figure S5 ISDE+1 as estimated by the generalized-cdf fitting method of Appendix S5 versus as estimated by the bin-based method of Methods.

Figure S6 ISDE results using cdf-fitting methods.

Table S1 ISDE estimates for the 149 webs of this study.

Table S2 SMSDE and LSDRE estimates for the 149 webs of this study.

Appendix S1 Details of the derivation of the theory.

Appendix S2 Fitting the truncated Pareto distribution.

Appendix S3 Fitting a generalized truncated Pareto distribution.

© 2008 Blackwell Publishing Ltd/CNRS

Appendix S4 Pareto vs. truncated Pareto distributions.
Appendix S5 Estimating the individual size distribution exponent.
Appendix S6 Data.
Appendix S7 Other theories and mechanisms.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.

Editor, Owen Petchey
Manuscript received 22 January 2008
First decision made 25 February 2008
Second decision made 3 June 2008
Third decision made 16 July 2008
Manuscript accepted 30 July 2008

© 2008 Blackwell Publishing Ltd/CNRS

Online Appendices

Three allometric relations of population density to body mass: Theoretical integration and empirical tests in 149 food webs

Daniel C. Reuman[1], Christian Mulder[2], Dave Raffaelli[3], Joel E. Cohen[4]

[1]LABORATORY OF POPULATIONS, ROCKEFELLER UNIVERSITY, 1230 YORK AVENUE, BOX 20, NEW YORK, NY 10065, U.S.A. FROM SEPTEMBER 2007: IMPERIAL COLLEGE LONDON, SILWOOD PARK CAMPUS, BUCKHURST ROAD, ASCOT, BERKSHIRE, SL5 7PY, UNITED KINGDOM. E-MAIL: D.REUMAN@IMPERIAL.AC.UK (CORRESPONDING AUTHOR).

[2]DEPARTMENT OF ECOLOGY, RIVM, 9 ANTONIE VAN LEEUWENHOEKLAAN, BOX 1, BILTHOVEN, 3720 BA, THE NETHERLANDS. E-MAIL: CHRISTIAN.MULDER@RIVM.NL

[3]ENVIRONMENT DEPARTMENT, UNIVERSITY OF YORK, YORK, YO10 5DD, U.K. E-MAIL: DR3@YORK.AC.UK

[4]LABORATORY OF POPULATIONS, ROCKEFELLER AND COLUMBIA UNIVERSITIES, BOX 20, 1230 YORK AVENUE, NEW YORK, NY 10065, U.S.A. E-MAIL: COHEN@ROCKEFELLER.EDU

Appendix S1: Details of the derivation of the theory

Eq. 4 reduces immediately to

Eq. S1 $$f_I(M) = f_I\left(\frac{M}{\beta}\right)\frac{\alpha}{\beta^{7/4}}.$$

Multiplying both sides by ln(10)*M and taking logs gives

Eq. S2 $$\log(\ln(10)*M*f_I(M)) = \log(\ln(10)*\frac{M}{\beta}*f_I\left(\frac{M}{\beta}\right)\frac{\alpha}{\beta^{3/4}}).$$

For any s and t,

Eq. S3 $$\int_s^t f_I(M)dM = \text{Prob}(s \le M \le t) = \text{Prob}(\log(s) \le u \le \log(t))$$

$$= \int_{\log(s)}^{\log(t)} g_I(u)du = \int_s^t g_I(\log(M))\frac{dM}{M\ln(10)},$$

where the last equality follows by making the substitutions u=log(M) and du=dM/(M*ln(10)). Since this holds for any s and t, we know the first and last integrands are equal, so

Eq. S4 $$\ln(10)*M*f_I(M) = g_I(u).$$

An analogue of Eq. S4 holds for f_S and g_S, too, by the same reasoning. Eq. S2 becomes

Eq. S5 $$\log(g_I(u)) = \log\left(\frac{\alpha}{\beta^{3/4}}\right) + \log(g_I(u - \log(\beta))).$$

Eq. S5 means the value of the abundance spectrum, $\log(g_I(u))$, at u is $\log\left(\frac{\alpha}{\beta^{3/4}}\right)$ plus its value at u-log(β). One functional form satisfying this requirement is $\log(g_I(u))$ linear in u with slope $\log(\alpha/\beta^{3/4})/\log(\beta) = \log(\alpha)/\log(\beta) - 3/4$, from which Eqs. 5-7 follow.

Eq. 4 and Eq. S5 are recursive relationships that determines the structure of f_I and g_I over the interval $[\beta^n*a_I, \beta^{n+1}*a_I]$ in terms of its structure over the interval $[\beta^{n-1}*a_I, \beta^n*a_I]$. Irregularities in f_I and g_I over $[a_I, \beta*a_I]$ are predicted to propagate through the ISD, superimposed on the general trend of Eqs. 6 and S5. In practice, random variation will mitigate this effect. The effect still may be visible, when relatively well-defined trophic levels exist, in the form of within-trophic-level relationships between density and body mass that differ from overall patterns and are superimposed on the general trend (e.g., Brown & Gillooly 2003; Jonsson *et al.* 2005). In summary, Eq. S5 implies that the overall pattern of the abundance spectrum will be a line, but periodic departures from this trend may be observed.

Appendix S2: Fitting the truncated Pareto distribution

Aban *et al.* (2006) gave the maximum likelihood estimator (MLE) of the exponent λ_{tp} of a truncated Pareto distribution with unknown truncation parameters a_{tp} and b_{tp} as the solution of an equation that can be easily solved numerically. Although the equation need not always have a solution, Aban *et al.* (2006) proved that it has a solution with probability tending to 1 as sample size increases, and they proved that the solution is unique when it exists. A solution existed for each of the mean body-mass distributions of the webs of this study. The standard theorems about asymptotic properties of MLEs do not apply to the MLE of λ_{tp} because regularity conditions of the theorems are not satisfied. But Aban *et al.* (2006) proved that the MLE of λ_{tp} is a consistent estimator of λ_{tp}, and that it is asymptotically normal and unbiased.

For comparisons with theory, it is important to use an unbiased estimator of λ_{tp}. But the MLE of λ_{tp} is biased for small sample sizes (Fig. S2). We corrected for this bias. For $a_{tp} = 1$, $\log(b_{tp})$ ranging from 2 to 9.5 in increments of 0.5, n (the sample size) ranging from 10 to 100 in increments of 5, and λ_{tp} ranging from -6 to 6 in increments of 0.1, we generated points from the truncated Pareto distribution and computed the MLE of λ_{tp}. This was repeated 500 times for each combination of b_{tp}, n and λ_{tp}. Bias of the MLE was estimated for each parameter combination. The function $-1.3814*(\lambda_{tp}+1)/n^{0.83428}$ explained 98.5% of the variation in the estimated bias. The parameters of this function were determined by nonlinear least-squares regression. This suggests the bias-corrected estimator $\hat{\lambda}_{tp} + 1.3814*(\hat{\lambda}_{tp}+1)/n^{0.83428}$, where $\hat{\lambda}_{tp}$ is the MLE. We reran the same simulations with the bias-corrected estimator in place of the MLE and found (e.g., Fig. S2) that the bias-corrected estimator had very low bias for $n \geq 30$ and λ_{tp} between -6 and 6 (which amply covered the range of values in this study). The variance of the bias-corrected estimator is a factor of $1+1.3814/n^{0.83428}$ greater than the variance of the MLE, a modest increase for $n \geq 30$. The mean squared error of the bias-corrected and MLE estimators are about the same (Fig. S2).

The uniformly minimum variance unbiased estimator (UMVUE) of λ_{tp}, derived by Beg (1983), equals $-\theta-1$ where

Eq. S6
$$\theta = (n-3)\frac{\sum_j (-1)^j {}_{n-2}C_j \delta_{j+1}(y,z)^{n-4}}{\sum_j (-1)^j {}_{n-2}C_j \delta_{j+1}(y,z)^{n-3}},$$

Eq. S7
$$\delta_j(y,z) = t - ny - j(z-y),$$

y is the minimal data point, z is the maximal one, t is the sum of the data, n is the number of data points, ${}_xC_y$ is the binomial coefficient, and the sums are from j = 0 to $j = \lfloor (t - ny)/(z - y) \rfloor$ where $\lfloor \ \rfloor$ represents rounding toward $-\infty$. For large enough n and λ_{tp} bigger than about -1.5 the UMVUE cannot be computed using most commonly available computing packages (including Matlab and R). The alternating summands in the numerator and denominator of θ can be larger than 10^{300}, but the sums themselves can be many orders of magnitude smaller. Since most computational software only stores the most significant 15-20 decimal digits of each variable in computer memory, cancellation of very large summands leads to erroneous results. The UMVUE could be computed efficiently using an arbitrary-precision arithmetic package such as the GMP package available for the C programming language (gmplib.org). Overcoming the numeric hurdles intrinsic in the UMVUE formula was beyond the scope of this study.

For large sample sizes, the MLE of λ_{tp} is essentially unbiased and is recommended. For small sample sizes, the MLE, if substantially biased, can be corrected or the UMVUE can be used if some numeric hurdles are overcome. The bias-corrected estimate and the UMVUE are more useful for studies of taxon mean trait distributions than for studies of individual-organism trait distributions, since the latter will typically involve larger sample sizes.

An untruncated Pareto distribution has pdf given by a power law on $[a_p, \infty)$ and equal to 0 elsewhere. Fitting a Pareto distribution is simpler than fitting a truncated Pareto distribution (Johnson *et al.* 1994; Aban *et al.* 2006; White *et al.* 2008). The exponent of a truncated Pareto distribution is sometimes estimated by ignoring the upper truncation: a Pareto distribution is fitted and the exponent is taken as an estimate of the exponent of the best-fitting truncated Pareto distribution (Page 1968). For some parameters this approximation is reasonable. But tests showed that the approximation introduces error for the data of this study (Appendix S4), so we did not use it.

Appendix S3: Fitting a generalized truncated Pareto distribution

We denoted the pdf of our generalized truncated Pareto distribution by $F_{gtp}(x)$, which is equal to 0 outside $[a_{gtp}, b_{gtp}]$. The truncation points $0 < a_{gtp} < b_{gtp}$ were taken to be unknown parameters. The distribution was described (Methods) by giving the pdf of w=log(x), which is $G_{gtp}(w) \propto 10^{\gamma_{gtp} * w + \eta_{gtp} * w^2}$ on $[\log(a_{gtp}), \log(b_{gtp})]$ and 0 elsewhere.

We used an MLE of the parameters of the generalized truncated Pareto distribution. It is easy to see that the MLEs of a_{gtb} and b_{gtb} are the minimum and maximum data points, respectively. We optimized the log likelihood function over the other two parameters, starting separately from each of five different initial parameter pairs using the Nelder-Mead simplex algorithm (Nelder & Mead 1965). One of the initial parameter pairs set γ_{gtp} equal to the bias-corrected MLE of the parameter λ_{tp} of the truncated Pareto distribution and set η_{gtp} equal to 0.

Appendix S4: Pareto versus truncated Pareto distributions

Three probability distributions are considered in this appendix. A *Pareto distribution* has probability density function (pdf) $f_p(x) \propto x^{\lambda_p}$ on $[a_p, \infty)$ and 0 elsewhere, where the constant of proportionality is determined by the requirement that a pdf integrate to 1. Here, $\lambda_p < -1$ is required so that x^{λ_p} can be rescaled to integrate to 1, and $a_p > 0$ is assumed unknown. A *truncated Pareto distribution*, introduced in the main text, has pdf $f_{tp}(x) \propto x^{\lambda_{tp}}$ on $[a_{tp}, b_{tp}]$ and 0 elsewhere, where λ_{tp} is an arbitrary real number and $0 < a_{tp} < b_{tp}$ are assumed unknown. A *power function distribution* has pdf $f_{pf}(x) \propto x^{\lambda_{pf}}$ on $(0, b_{pf}]$ and 0 elsewhere, where $\lambda_{pf} > -1$ is required so that $x^{\lambda_{pf}}$ can be rescaled to integrate to 1, and $b_{pf} > 0$ is unknown.

Maximum likelihood estimators (MLE) and uniformly minimum variance unbiased estimators (UMVUE) of λ_p, λ_{tp}, and λ_{pf} are available, but not all of these estimators are equally easy to compute. Simple formulas for the MLE and the UMVUE of λ_p were given by Johnson *et al.* (1994, p. 582-583). Since a power function distribution becomes a Pareto distribution under the transformation $w = 1/x$, these estimators can easily be transformed to obtain an MLE and a UMVUE of the parameter λ_{pf}. An MLE and a UMVUE of the truncated Pareto parameter λ_{tp} are also available, but are much more difficult to compute (Beg 1983; Aban *et al.* 2006; see also Appendix S3).

In the main text, we argued that the truncated Pareto distribution is more appropriate for describing body mass data because of ecological and physiological limitations on the range of possible masses. The truncated Pareto distribution is also more appropriate if sampling methods precluded detection of masses outside of some range. Perhaps because of the difficulty of estimating λ_{tp}, the Pareto and power function distributions are sometimes used instead of the truncated Pareto distribution in cases where data were most likely bounded above (e.g., earthquake intensities, body sizes). Page (1968) suggested this approximation in cases where the largest data point is more than 100 times the

smallest data point. We explored in several ways the possibility that error can be introduced by fitting Pareto and power function distributions to data from finite systems or from the truncated Pareto distribution (i.e., we explored when using an estimate of λ_p as an estimate of λ_{tp} can cause error).

First, for each of several values of λ_{tp} spanning the range -7 to 5, we independently generated 100 data points from a truncated Pareto distribution with $a_{tp} = 1$ and $b_{tp} = 1000$, and estimated λ_{tp} with the UMVUE of λ_p for the Pareto distribution. We denote this estimator λ_1. We also estimated λ_{tp} by fitting a power function distribution, again using a UMVUE. We denote this estimator λ_2. We also used a third estimator, here denoted λ_3, of the truncated-Pareto parameter λ_{tp}, equal to λ_1 or λ_2 according to whether the maximum likelihood of the data was greater for a Pareto distribution or a power function distribution, respectively. The experiment was repeated 5000 times for each value of λ_{tp} and the bias, variance, and mean squared error of each estimator was plotted against λ_{tp} (Fig. S1A). All three estimators were substantially biased for λ_{tp} close to -1. Since λ_1 was always less than -1, it was heavily biased for $\lambda_{tp} > -1$. Since λ_2 was always greater than -1, it was heavily biased for $\lambda_{tp} < -1$. λ_3 was bimodal for λ_{tp} close to -1 (Fig. S1B). λ_1 is an unbiased estimator of λ_p, and λ_2 is an unbiased estimator of λ_{pf}, but neither is an unbiased estimator of λ_{tp}, nor is λ_3.

A substantial fraction of datasets of size 100 with points drawn independently from a truncated Pareto distribution with $a_{tp} = 1$, $b_{tp} = 1000$ and λ_{tp} not close to 0 will not span 2 orders of magnitude. Since Page (1968) suggested that the estimator λ_1 could be used as an approximation of λ_{tp} when the largest data point is more than 100 times the smallest data point, to further test the suggestion, we repeatedly generated 100 independent points from a truncated Pareto distribution with $a_{tp} = 1$, $b_{tp} = 1000$, $\lambda_{tp} = -1.1$ until we had 5000 datasets each spanning at least two orders of magnitude. The mean value of λ_1 for these datasets was -1.3268, showing substantial bias for λ_{tp} close to 1 even for datasets covering a wide range.

Finally, we compared the values of λ_1, λ_2, and λ_3 estimated for the SMSDs of the webs of this study to those of the bias-corrected MLE described in Methods and Appendix S3, here called λ_4. The median, 95th and 99th percentiles, and the maximum of $|\lambda_1 - \lambda_4|$ over our 149 webs were 0.0105, 0.1915, 0.2283, and 0.3955. The same percentiles for $|\lambda_2 - \lambda_4|$ were 0.3453, 0.4050, 0.4645, and 0.4783. The same percentiles for $|\lambda_3 - \lambda_4|$ were 0.0105, 0.1915, 0.2266, 0.2312. The Akaike Information Criterion, corrected for small sample size (AIC_c; Burnham & Anderson 2002, p. 66) was computed for each type of distribution and for each of our 149 webs. The AIC_c for the truncated Pareto minus the AIC_c for the Pareto was never greater than 2, but was less than -2 in 54 of 149 webs, supporting the choice of the

truncated Pareto as a generally better description of the data of this study. The AIC_c for the truncated Pareto minus the AIC_c for the power function distribution was less than -2 for all 149 webs.

In summary: λ_1 is a biased estimator of λ_{tp} when λ_{tp} is not substantially less than -1; λ_2 is a biased estimator of λ_{tp} when λ_{tp} is not substantially greater than -1; λ_3 is a biased estimator of λ_{tp} when λ_{tp} is close to -1; and λ_1, λ_2, and λ_3 can be poor approximations of λ_4, which is an essentially unbiased estimator of λ_{tp} for a wide range of biologically reasonable λ_{tp} values (Appendix S3). Our estimates λ_4 took values less than, greater than, and often close to -1, so the tests of this appendix justify the computational effort required to compute λ_4. The estimators λ_1, λ_2, and λ_3 of λ_{tp} should be used, as suggested by Page (1968), only when one can be confident that λ_{tp} is substantially less than -1 (respectively, greater than -1, far from -1). For data that may have been bounded above, one should consider carefully the decision to use a Pareto distribution in place of a truncated Pareto distribution; comparison of the two distributions using the AIC_c can be helpful in making this decision. The data of this study are at least as well described by the truncated Pareto distribution as by either the Pareto or power function distributions, according to the AIC_c.

Appendix S5: Estimating the individual size distribution exponent

Bias of common bin-based estimators

We plotted the bias, variance, and mean squared error of a commonly used method of estimating λ_{tp}, based on separating data into log-scale-uniform bins. For each λ_{tp} from -7 to 5 in increments of 0.1, 5000 datasets, d, of size 100 were drawn (each point independently) from a truncated Pareto distribution with $a_{tp} = 1$ and $b_{tp} = 1000$. For each dataset, λ_{tp} was estimated by dividing the interval [log(min(d)), log(max(d))] into 10 equal bins, computing the log number of points in each bin, regressing against log-scale bin centres (excluding empty bins), and subtracting 1 from the regression slope. Bias, variance, and mean squared error were plotted against λ_{tp} (Fig. S3). The estimator was biased toward -1. The process was repeated using 500 datasets of size 1000 for each λ_{tp}. Bias patterns were similar but reduced. The method used in the main text to estimate ISDE from taxon $\overline{M}$ and N data (Methods) is an intuitive adaptation of the above method. Although it is not possible to quantitatively estimate the bias, variance, or mean squared error of the adapted estimator without making assumptions about distributions of individual-organism M within taxa, the adapted method may

have patterns of bias similar to the method it is based upon: the ISDEs we report may be closer to -1 than true ISDEs.

Cdf-fitting method

Bin-based methods have been criticized for their bias and variance and for the arbitrary choices they involve, such as the number of bins used White *et al.* (2008). Preferred methods such as the UMVUE and the MLE (Appendix S3) cannot be adapted to estimate the ISDE from taxon mean body masses $\overline{M}$ and population densities N without making assumptions about distributions of individual-organism M within taxa. UMVUE and MLE formulas refer to minimum and maximum observed individual body masses (Appendix S3; Aban *et al.* 2006), unknown if only the taxon means $\overline{M}$ were recorded. We adapted the cdf-fitting method recommended by White *et al.* (2008) (see also Johnson *et al.* 1994, p. 580) to the context of the available data in order to compare cdf-based estimates to bin-based estimates of the ISDE.

If mean body masses of taxa in a web are sorted in ascending order, $\overline{M}_1$, …, $\overline{M}_n$, then

Eq. S8 $$\varphi_{emp}(\overline{M}_i) = \frac{wN_i + \sum_{j=1}^{i-1} N_j}{\sum_{j=1}^{n} N_j}$$

is an empirical approximation of the cdf of the ISD at $\overline{M}_i$, where w is taken to be 0.5. A straightforward adaptation of the cdf-fitting method would numerically minimize the summed squared differences, at the $\overline{M}_i$, between this empirical cdf and the theoretical cdf of the ISD, under the assumptions that the ISD is a truncated Pareto distribution with parameters λ_{tp}, a_{tp}, and b_{tp}. The theoretical cdf is

Eq. S9 $$\varphi_{theor}(M) = \frac{M^{\lambda_{tp}+1} - a_{tp}^{\lambda_{tp}+1}}{b_{tp}^{\lambda_{tp}+1} - a_{tp}^{\lambda_{tp}+1}}$$

for $\lambda_{tp} \neq -1$ and

Eq. S10 $$\varphi_{theor}(M) = \frac{\ln(M) - \ln(a)}{\ln(b) - \ln(a)}$$

for $\lambda_{tp} = -1$. However, the empirical cdf of Eq. S8 is often a statistically poor target to which to match a theoretical cdf. If the population density of one or a few taxa is a substantial portion of the total

population density of all taxa in a system (a common occurrence in webs), then the values of Eq. S8 will cluster close to 0 and 1 (Fig. S4). In that case, the best fit of Eq. S9 with Eq. S8 is determined primarily by a few taxa. It may be possible to transform the empirical and theoretical cdfs to mitigate this problem, but any of several transformations could be used, arbitrarily, with different results. For these reasons, we did not use this straightforward adaptation of the cdf-fitting method.

We developed a natural generalization of the cdf-fitting method that circumvents the problems stated in the previous paragraph without requiring arbitrary choices. "Cdf-fitting" in the main text refers to the improved method we now describe. The theoretical cdf of a truncated Pareto ISD with parameters λ_{tp}, a_{tp}, and b_{tp} is

Eq. S11 $$\varphi_{theor}(M) = \frac{\int_{a_{tp}}^{M} t^{\lambda_{tp}} dt}{\int_{a_{tp}}^{b_{tp}} t^{\lambda_{tp}} dt}.$$

We defined a generalized theoretical cdf

Eq. S12 $$\varphi_{theor}^{gen}(M) = \frac{\int_{a_{tp}}^{M} h(t) t^{\lambda_{tp}} dt}{\int_{a_{tp}}^{b_{tp}} h(t) t^{\lambda_{tp}} dt},$$

where $h(t) = t^s$. This can be reduced to a simple formula. If $s = 1$, Eq. S12 is the proportion of biomass in the system with body mass less than M. An empirical approximation of the generalized cdf at $\overline{M}_i$ is given by

Eq. S13 $$\varphi_{emp}^{gen}(\overline{M}_i) = \frac{wN_i\overline{M}_i^{\,s} + \sum_{j=1}^{i-1} N_j \overline{M}_j^s}{\sum_{j=1}^{n} N_j \overline{M}_j^s},$$

where $w = 0.5$. For any fixed s, it is straightforward to minimize numerically the summed squared differences, at the $\overline{M}_i$, between the empirical and theoretical generalized cdfs.

We selected s to minimize variation in the quantities $N_i\overline{M}_i^{\,s}$, thereby optimizing the evenness of $\varphi_{emp}^{gen}(\overline{M}_i)$ over the range [0, 1] and making the empirical generalized cdf a statistically reasonable target to compare to the theoretical generalized cdf. A natural choice for s was -b, where

$\log(N) = b * \log(\overline{M}) + a + \varepsilon$ is the ordinary least squares best fit. This choice minimizes var(ε), equal to $\mathrm{var}(\log(N\overline{M}^{-b}))$ (Fig. S4). We estimated IDSEs for all webs of this study by generalized-cdf fitting using this procedure.

For each web of this study, confidence intervals of the generalized-cdf estimate of the ISDE were computed by the resampling scheme of Methods, using 500 resamplings instead of 5000. A quadratic generalization of the truncated Pareto distribution (like Eq. 12, but using individual body masses instead of taxon means) was also fitted by the generalized-cdf method, and the resampling scheme (here using 100 resamplings instead of 5000) was used to determine approximate confidence intervals of the quadratic coefficient. The hypothesis that the ISD was a truncated Pareto distribution was rejected if these confidence intervals did not contain 0. All ISDE results obtained by generalized-cdf fitting are in Table S1, alongside the same results obtained by binning. R^2 (over all 149 webs) between bin-based and generalized-cdf estimates of ISDEs was 83.4% (Fig. S5). The main results of this study were qualitatively similar regardless of whether ISDE was estimated by binning or by generalized-cdf fitting (Fig. S6).

Probabilistically grounded estimators

Individual-organism M data, though presently rare for webs, could enable probabilistically grounded estimates of the ISDE and more exacting tests of the ISD theory of this study. Different sampling methods are often used for different species or size ranges within a web, so even when individual M data are available, it may be impossible to view measured values as random samples from the ISD, and therefore inappropriate to use the MLE or UMVUE. A probabilistically grounded estimator of the ISDE based on realistically obtainable web data is a useful topic for future research. It may be necessary to model the distribution of organism M values within taxa, an important challenge in itself.

Appendix S6: Data

The two pelagic webs were from Tuesday Lake, a mildly acidic lake in Michigan, U.S.A. (Carpenter & Kitchell 1993; Cohen *et al.* 2003; Jonsson *et al.* 2005, which contains the full data). In 1984 and separately in 1986, all taxa found in the non-littoral epilimnion of Tuesday Lake (except parasites and bacteria) were resolved to species. Three fish species were removed and another was introduced in 1985. About 50% species turnover occurred between 1984 and 1986. Only taxa connected to the webs were included for this study (50 species in 1984, 51 in 1986). Six unconnected species were found in

each year. $\overline{M}$ was measured in kg fresh mass. N was in individuals per m^3 of the non-littoral epilimnion, where all trophic interactions occurred.

The estuarine web was in the Ythan Estuary, Scotland (Hall & Raffaelli 1991). Of the 91 taxa observed, 73 were species and half of the rest were genera. Particulate organic matter, a resource for some taxa in the Ythan Estuary, was omitted from this study because $\overline{M}$ and N estimates were difficult to obtain. $\overline{M}$ was in g fresh mass per individual; population density was absolute numbers of individuals in the whole estuary (the unit were numbers per area A, where A was the area of the estuary). The full data will be published elsewhere (Schittler, Reuman, Raffaelli & Cohen, in preparation).

All taxa observed in the rhizosphere of the 146 Dutch agroecosystems were measured for $\overline{M}$ and N except fungi and protists. Of 169 nematode, 186 microarthropod, and 17 oligochaete taxa identified in any of the 146 soil webs, approximately 78% of the nematodes, 83% of the microarthropods, and 100% of the oligochaetes were genera; the rest were families. Bacterial $\overline{M}$ and N considered all bacteria as one taxon (Mulder *et al.* 2005a; Mulder *et al.* 2005b); these data were omitted because this level of resolution was very different from that used for other taxa. Units of $\overline{M}$ were μg dry mass and units of N were individuals per m^2. $\overline{M}$ and N were from direct measurement, but the $\overline{M}$ values for each taxon were assumed to be the same at all sites (Mulder *et al.* 2005a; Mulder *et al.* 2005b). Samples were from the top 10 cm of soil. For possible access to the soil web data, contact C.M., who intends to publish the webs at a future date.

The theory of this study used a fixed value of β, but $\overline{M}_c/\overline{M}_r$ varied within each of the Tuesday Lake and Ythan Estuary webs. To test theory, a representative value of β was needed for each web. We chose the central value β=mean(log($\overline{M}_c/\overline{M}_r$)) of each web's distribution of $\overline{M}_c/\overline{M}_r$ because the $\overline{M}_c/\overline{M}_r$ were closer to a log-normal distribution than to a normal distribution for all three webs. Mean(log($\overline{M}_c/\overline{M}_r$)) was more representative of a typical trophic link in each web than log(mean($\overline{M}_c/\overline{M}_r$)). The latter was unduly influenced by a few links with large $\overline{M}_c/\overline{M}_r$, which represented feeding relationships that may have been uncommon. There is no reason to believe that links with large $\overline{M}_c/\overline{M}_r$ have influence on patterns of relative population density in webs in proportion

to their large influence on log(mean($\overline{M}_c/\overline{M}_r$)). Mean(log($\overline{M}_c/\overline{M}_r$)) and $10^{\text{mean}(\log(\overline{M}_c/\overline{M}_r))}$ were respectively closer to the medians of the distributions of log($\overline{M}_c/\overline{M}_r$) and $\overline{M}_c/\overline{M}_r$ than were log(mean($\overline{M}_c/\overline{M}_r$)) and mean($\overline{M}_c/\overline{M}_r$).

Appendix S7: Other theories and mechanisms

It is possible to interpret theory in Brown *et al.* (2004; pp. 1785-1786, their Eq. 13) and Jennings & Mackinson (2003) so that predictions do not agree with our theoretical predictions or data (comments of an anonymous referee). However, if so interpreted, predictions of those theories would also disagree with the data presented in the same studies. Eq. 13 of Brown *et al.* (2004) read, in part:

Eq. 13 $$N = N_0\left(\frac{M}{M_0}\right)^{[\log(\alpha)/\log(\beta)]-3/4},$$

where N was called the "total number of organisms of a given size" (Brown *et al.* 2004, p. 1785), and N_0 and M_0 were constants for a web. If N in Eq. 13 is interpreted as the total number of organisms in the interval [M, M+dM], then Eq. 13 constitutes a prediction that $\lambda_I = \frac{\log(\alpha)}{\log(\beta)} - \frac{3}{4}$, conflicting with Eq. 7. However, if N in Eq. 13 is interpreted as the total number of organisms in the interval [u, u+du], where u=log(M), then Eq. 13 predicts that the abundance spectrum slope is $\frac{\log(\alpha)}{\log(\beta)} - \frac{3}{4}$, and that therefore $\lambda_I = \frac{\log(\alpha)}{\log(\beta)} - \frac{7}{4}$, in agreement with Eq. 7. We assumed the latter interpretation, because Brown *et al.* (2004) compared their prediction $\frac{\log(\alpha)}{\log(\beta)} - \frac{3}{4}$ to the classical result that the marine abundance spectrum slope is -1 (while α is often taken to be 10% and a typical β in marine systems is about 10000). Jennings & Mackinson (2003) also compared $\frac{\log(\alpha)}{\log(\beta)} - \frac{3}{4}$ to the slope of an abundance spectrum. Confusion may result when it is not clearly indicated whether theory is intended to predict the abundance spectrum slope or the ISDE. The probabilistic approach and language of White *et al.* (2008) and this study can help clarify theoretical statements and their comparison with data.

Another interpretation of the theory of Brown *et al.* (2004; pp. 1785-1786, their Eq. 13) is also possible, but does not agree with our data. In previous work, we compared the LSDREs of 166 webs to

the predictions of two null models (Reuman *et al.* In press). One of the null models was the prediction that the LSDRE (rather than the abundance spectrum slope, as predicted in Appendix S1) should equal

Eq. 14 $$\lambda_L = \frac{\log(\alpha)}{\log(\beta)} - \frac{3}{4}.$$

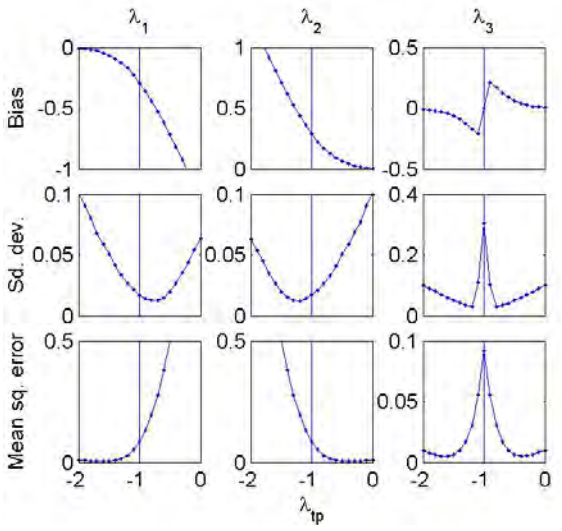
Eq. 14 is a possible interpretation of the way theory is presented in Brown *et al.* (2004) and especially in Brown & Gillooly (2003). Eq. 14 is empirically unsupported, as demonstrated here (last paragraph of *Local size-density relationships* in Results) and in Reuman *et al.* (In press). A superficial reading of Reuman *et al.* (In press) and the current study may suggest that the theories of Brown *et al.* (2004) were falsified by the former study and validated by the latter, but we emphasize that in Reuman *et al.* (In press), we compared the interpretation Eq. 14 with data, and here we compared a different interpretation, augmented by other theory (Appendix S1). We recognized in Reuman *et al.* (2004) that Eq. 13 of Brown *et al.* (2004) was intended to predict the abundance spectrum and not the LSDRE; testing Eq. 14 was for comparative purposes and because a precise and integrated theory of the LSDR was not available at that time. Nevertheless, confusion may result when it is not explicitly indicated whether theory is intended to predict the LSDR or the ISD and abundance spectrum.

Formulating theory intended to predict ISDs and abundance spectra using the term "trophic level," as done by Brown & Gillooly (2003) and Brown *et al.* (2004) may also create confusion. In much historical literature, the concept of the trophic level has often been applied to species or higher taxa, ignoring ontogenetic changes in individual feeding. Arriving at predictions for the ISD, which ignores taxonomy, using concepts such as trophic level could create confusion, unless terms are precisely defined or redefined.

This study showed that a mechanistic model explained variation in observed allometric patterns among webs. Our results supported the causal hypothesis in the model but did not confirm causation nor exclude other important mechanisms. The model of Silvert & Platt (1980) included mass-specific mortality and transfer of biomass (by a combined growth and consumption term) from small-mass categories to larger-mass categories. Andersen & Beyer (2006) modelled consumer foraging velocities and search volumes. Rossberg *et al.* (2008) modelled several mechanisms combined, including evolution, migration from and to other communities, and explicit top-down effects. A topic for future research is a detailed comparison of the ability of models to explain variation among webs in observed allometric patterns. Such a comparison could illuminate the relative importance of the proposed mechanisms for structuring communities.

Figure S1: Bias and variability of parameter estimates from fitting Pareto and power function distributions to data from a truncated Pareto distribution. (A) For each value of λ_{tp}, 100 values were independently chosen from the truncated Pareto distribution with $a_{tp} = 1$ and $b_{tp} = 1000$; λ_1, λ_2, and λ_3 were computed. This experiment was repeated 5000 times for each λ_{tp}. The bias, standard deviation, and mean squared error of each estimator were plotted. Increments of 0.1 in λ_{tp} from -7 to 5 were used. (B) For each of the λ_{tp} values -1.05, -1, and -0.95, 5000 datasets of 100 points each were generated from a truncated Pareto distribution with $a_{tp} = 1$ and $b_{tp} = 1000$ and histograms of λ_3 estimates were plotted.

A



B

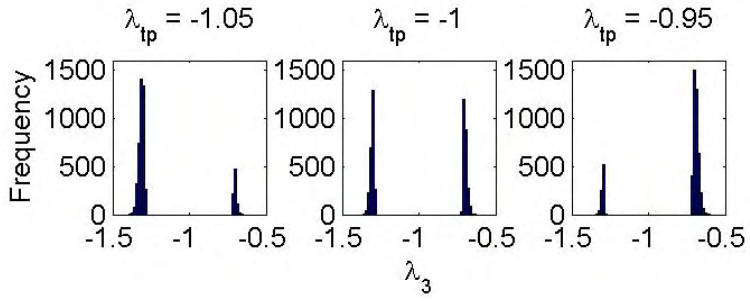


Figure S2: Bias, standard deviation, and mean squared error of the MLE of the exponent λ_{tp} of the truncated Pareto distribution (left) and the bias-corrected MLE of the same exponent (right). For these plots, $a_{tp} = 1$ and $b_{tp} = 1000$, and 100 data points were independently generated 5000 times for each value of λ_{tp}, so 5000 estimates were generated for each λ_{tp}.

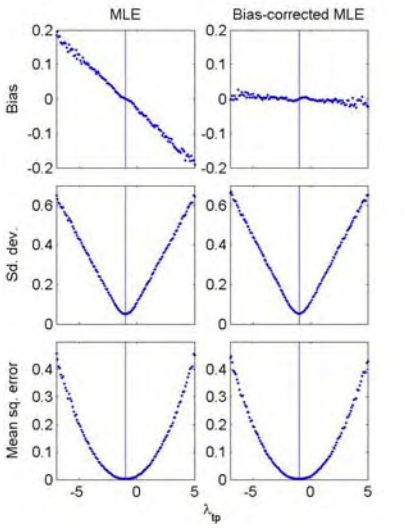


Figure S3: Bias, standard deviation, and mean squared error of an estimate of the exponent λ_{tp} of the truncated Pareto distribution, using 10 bins uniform on a logarithmic scale. For these plots, $a_{tp} = 1$ and $b_{tp} = 1000$ and 100 data points (left) and 1000 data points (right) were independently generated 5000 times (left) and 1000 times (right) for each value of λ_{tp}, so 5000 estimates (left) and 1000 estimates (right) were generated for each λ_{tp}.

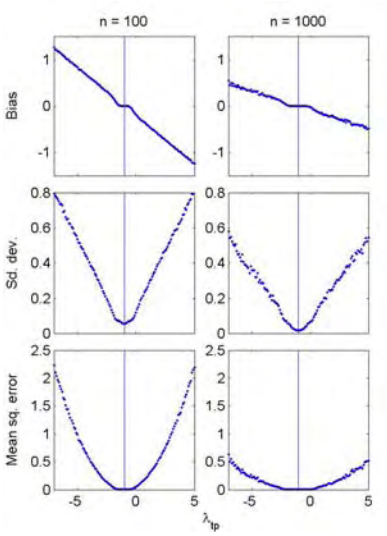


Figure S4: A straightforward adaptation of cdf fitting to the case where only taxon means $\overline{M}$ and population densities N are available (left), compared to a generalized method (right). A few species comprised most of the total population density in Tuesday Lake, 1984, so the empirical cdf of Eq. S8 had most points close to 1, and was a poor target to compare with a theoretical cdf. A generalized cdf (right) was a statistically better target for comparison with a generalized theoretical cdf (Appendix S5). Results were often unrealistic for the cdf fitting method but more realistic for the generalized-cdf fitting method. For instance, cdf fitting gave ISDE+1 = -0.37 for Tuesday Lake 1984, an unrealistic value based primarily on the fit of the cdf to the smallest 10-15 species. Generalized-cdf fitting and the bin-based method of Methods both gave -0.94; these estimates were not disproportionately influenced by any subset of species.

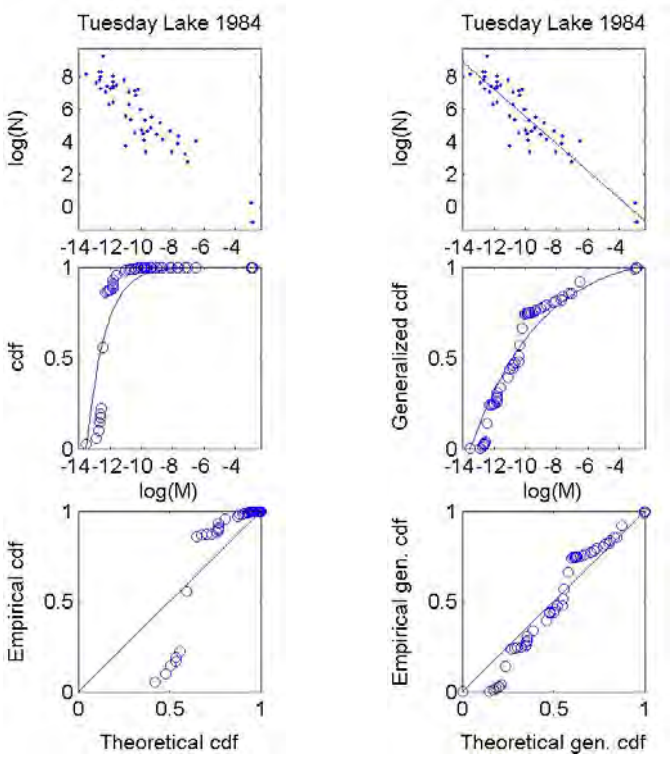


Figure S5: ISDE+1 as estimated by the generalized-cdf fitting method of Appendix S5 versus as estimated by the bin-based method of Methods. The line is y = x.

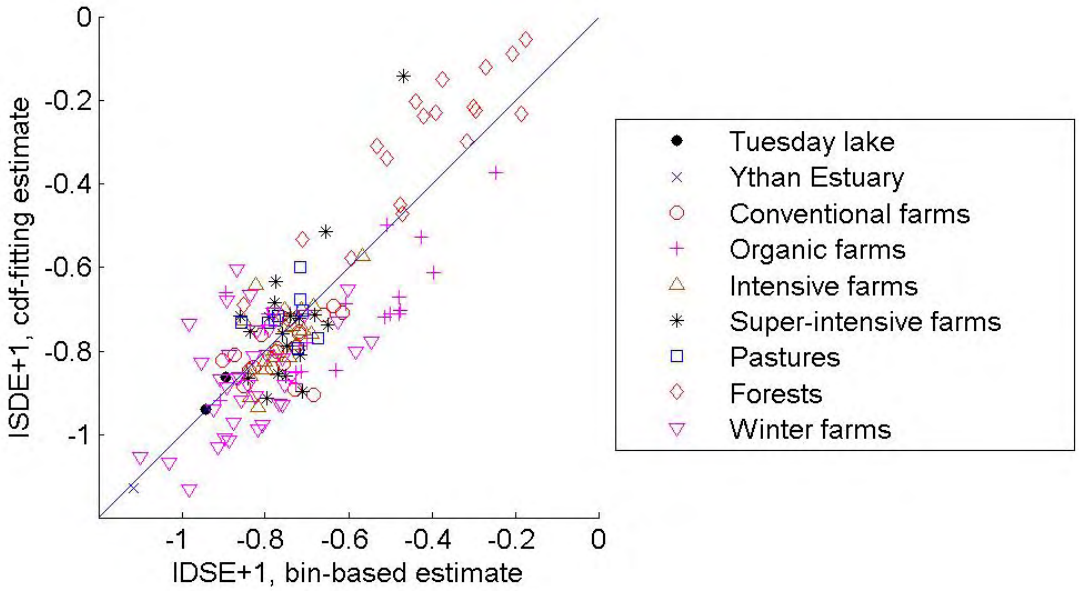


Figure S6: (Left) Same as Fig. 3A, C, E but using the generalized-cdf estimator of ISDE. (Right) Same as the second column of panels of Fig. 4, but using the generalized-cdf estimator of ISDE. Webs were excluded from these histograms only if they had non-power-law ISD, determined using generalized-cdf methods as described in Appendix S5. Only two of 146 soil webs were excluded, both forests (Table S1). The mean ISDE in the panel for super-intensive farms was computed with and without the largest outlier value included. Mn_{OR} denotes the mean with the right-most outlier point removed.

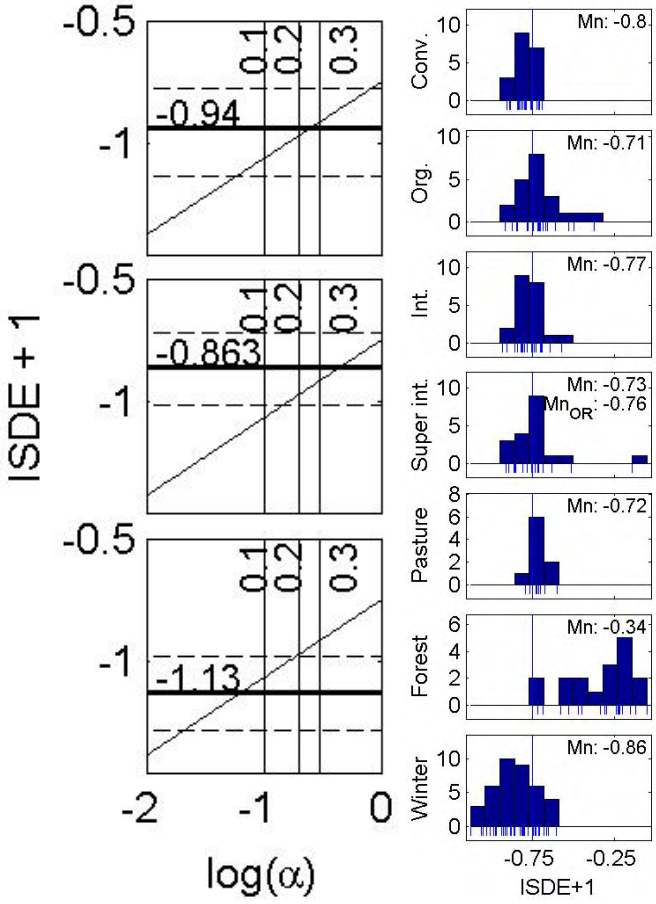


Table S1: Individual size distribution exponent (ISDE) estimates for the 149 webs of this study. The entry under "Web name" for soil webs is the web ID number, also used in Fig. 2. Web types are: P = pelagic, E = estuarine, 1 = conventional farm, 2 = organic farm, 3 = intensive farm, 4 = super-intensive farm, 5 = pasture, 6 = forest, 7 = winter farm. Column 3 contains estimates of ISDE+1 using the bin-based method (Methods); columns 4-5 contain 99% confidence intervals. Column 6 contains "yes" if bin-based methods could not reject the hypothesis that the ISD was a truncated Pareto distribution (1% level). Column 7 contains estimates of ISDE+1 using the generalized-cdf fitting method of Appendix S5; columns 8-9 contain 99% confidence intervals. Column 8 contains "yes" if generalized cdf-based methods could not reject the hypothesis that the ISD was a truncated Pareto (1% level). Confidence intervals entirely above -3/4 are highlighted in red.

Web name	Web type	Bin ISDE, +1	Lower 99% conf.	Upper 99% conf.	Trunc. Pareto ?	cdf ISDE, +1	Lower 99% conf.	Upper 99% conf.	Trunc. Pareto ?
TL 84	P	-0.9438	-1.1309	-0.7566	yes	-0.9398	-1.1311	-0.77262	yes
TL 86	P	-0.896	-1.1283	-0.6637	yes	-0.8628	-1.0128	-0.71939	yes
Ythan	E	-1.1147	-1.3652	-0.8642	no	-1.1285	-1.2768	-0.97934	no
ID 95	1	-0.6158	-1.1801	-0.0515	yes	-0.7093	-1.1237	-0.43502	yes
ID 96	1	-0.8375	-1.2694	-0.4056	yes	-0.8455	-1.1356	-0.64081	yes
ID 97	1	-0.8095	-1.3532	-0.2658	yes	-0.7612	-1.091	-0.48895	yes
ID 98	1	-0.9039	-1.1632	-0.6446	yes	-0.8226	-1.2855	-0.34274	yes
ID 99	2	-0.7944	-1.163	-0.4257	yes	-0.7108	-1.2139	-0.45558	yes
ID 100	1	-0.7708	-0.985	-0.5565	yes	-0.8001	-1.1184	-0.50622	yes
ID 101	1	-0.7723	-1.4121	-0.1325	yes	-0.7939	-1.2199	-0.52647	yes
ID 102	1	-0.6849	-1.4873	0.1175	yes	-0.9065	-1.5917	-0.63559	yes
ID 103	1	-0.7289	-1.2378	-0.22	yes	-0.8925	-1.1496	-0.71894	yes
ID 104	1	-0.7255	-0.9676	-0.4833	yes	-0.7901	-0.99514	-0.59208	yes
ID 105	1	-0.8276	-1.2504	-0.4049	yes	-0.8396	-1.1949	-0.5823	yes
ID 106	3	-0.5659	-0.9929	-0.1389	yes	-0.5728	-0.9072	-0.17129	yes
ID 107	3	-0.7077	-1.367	-0.0483	yes	-0.7503	-1.305	-0.12272	yes
ID 108	3	-0.8381	-1.5094	-0.1668	yes	-0.8434	-1.1424	-0.55341	yes
ID 109	2	-0.741	-1.1894	-0.2926	yes	-0.7096	-1.1777	-0.42119	yes
ID 110	3	-0.8356	-1.3277	-0.3435	yes	-0.8593	-1.211	-0.40697	yes
ID 111	1	-0.6366	-1.6131	0.3398	yes	-0.691	-1.167	-0.37895	yes
ID 112	1	-0.7421	-1.3073	-0.1768	yes	-0.7243	-1.1665	-0.27452	yes
ID 113	1	-0.8724	-1.2816	-0.4633	yes	-0.8097	-1.1627	-0.54856	yes
ID 114	2	-0.7004	-1.1157	-0.285	yes	-0.7787	-1.0031	-0.45425	yes
ID 115	3	-0.7495	-1.3996	-0.0994	yes	-0.7429	-1.2336	-0.44963	yes
ID 116	3	-0.7349	-1.2237	-0.2461	yes	-0.7388	-1.0605	-0.4483	yes
ID 117	1	-0.7562	-1.0999	-0.4125	yes	-0.8315	-1.1878	-0.57098	yes
ID 118	1	-0.7178	-1.1277	-0.3079	yes	-0.7546	-1.0356	-0.55186	yes
ID 119	3	-0.7535	-1.1091	-0.3978	yes	-0.6987	-1.1112	-0.24855	yes
ID 120	3	-0.8182	-1.3238	-0.3126	yes	-0.9352	-1.2026	-0.68474	yes
ID 121	3	-0.7765	-1.2285	-0.3245	yes	-0.7987	-1.0631	-0.54718	yes

ID 122	1	-0.7833	-1.3331	-0.2335	yes	-0.8427	-1.1484	-0.66299	yes
ID 123	3	-0.689	-1.0871	-0.2909	yes	-0.7584	-1.0395	-0.4463	yes
ID 124	3	-0.8217	-1.3063	-0.3371	yes	-0.6418	-0.96202	-0.35439	yes
ID 125	3	-0.8518	-1.252	-0.4515	yes	-0.7276	-1.0446	-0.50988	yes
ID 126	3	-0.7246	-1.1166	-0.3327	yes	-0.7848	-1.0881	-0.43364	yes
ID 127	1	-0.6605	-1.0938	-0.2272	yes	-0.7133	-0.97787	-0.42487	yes
ID 128	3	-0.8355	-1.3941	-0.2769	yes	-0.9105	-1.3096	-0.56621	yes
ID 129	3	-0.6853	-1.0895	-0.2811	yes	-0.6909	-1.0115	-0.19204	yes
ID 130	3	-0.8025	-1.5634	-0.0415	yes	-0.8243	-1.2879	-0.4312	yes
ID 131	3	-0.7129	-1.1569	-0.2689	yes	-0.701	-1.0743	-0.32808	yes
ID 132	1	-0.8522	-1.4652	-0.2391	yes	-0.8844	-1.2082	-0.54569	yes
ID 133	3	-0.7269	-1.0586	-0.3951	yes	-0.7855	-1.1079	-0.5606	yes
ID 134	3	-0.7808	-1.1273	-0.4343	yes	-0.8137	-1.1683	-0.53436	yes
ID 135	3	-0.7984	-1.3413	-0.2555	yes	-0.8453	-1.1888	-0.60943	yes
ID 136	1	-0.7193	-1.2175	-0.2212	yes	-0.7594	-1.0611	-0.3329	yes
ID 137	2	-0.7125	-1.1803	-0.2447	yes	-0.8504	-1.1157	-0.60169	yes
ID 138	2	-0.802	-1.2186	-0.3855	yes	-0.7457	-1.0695	-0.50848	yes
ID 139	2	-0.894	-1.4769	-0.3111	yes	-0.6597	-1.1502	-0.17593	yes
ID 140	2	-0.7263	-1.1827	-0.2699	yes	-0.8509	-1.2435	-0.62629	yes
ID 141	2	-0.9075	-1.4493	-0.3658	yes	-0.9199	-1.3173	-0.59697	yes
ID 142	2	-0.7199	-1.1274	-0.3125	yes	-0.7867	-1.0791	-0.54854	yes
ID 143	2	-0.7256	-1.1576	-0.2935	yes	-0.878	-1.1181	-0.61712	yes
ID 144	3	-0.7429	-1.0692	-0.4165	yes	-0.8146	-1.0723	-0.55167	yes
ID 145	4	-0.6549	-1.1033	-0.2066	yes	-0.5127	-0.75996	-0.30693	yes
ID 146	4	-0.7174	-1.1598	-0.275	yes	-0.7213	-1.1664	-0.36068	yes
ID 147	4	-0.7156	-1.2407	-0.1905	yes	-0.8093	-1.0764	-0.56164	yes
ID 148	4	-0.8369	-1.3143	-0.3595	yes	-0.7528	-1.095	-0.5201	yes
ID 149	4	-0.7389	-1.1306	-0.3472	yes	-0.7164	-1.0166	-0.28551	yes
ID 150	4	-0.8424	-1.3127	-0.3721	yes	-0.8662	-1.2874	-0.55649	yes
ID 151	4	-0.7484	-1.0831	-0.4137	yes	-0.7868	-1.098	-0.51626	yes
ID 152	4	-0.769	-1.1005	-0.4374	yes	-0.8542	-1.1481	-0.56951	yes
ID 153	4	-0.7095	-1.1535	-0.2655	yes	-0.897	-1.2561	-0.61942	yes
ID 154	4	-0.7771	-1.085	-0.4692	yes	-0.6847	-0.99892	-0.39403	yes
ID 155	4	-0.7918	-1.4974	-0.0862	yes	-0.7177	-0.94928	-0.47265	yes
ID 156	4	-0.8598	-1.3397	-0.38	yes	-0.716	-1.0725	-0.30304	yes
ID 157	4	-0.469	-1.063	0.125	yes	-0.1409	-0.2928	0.56387	yes
ID 158	4	-0.795	-1.1994	-0.3907	yes	-0.9123	-1.2308	-0.6657	yes
ID 159	4	-0.7501	-1.1086	-0.3915	yes	-0.8598	-1.2	-0.64875	yes
ID 160	4	-0.7578	-1.1398	-0.3759	yes	-0.7588	-1.0452	-0.26966	yes
ID 161	4	-0.6821	-1.0793	-0.285	yes	-0.7141	-0.96124	-0.44405	yes
ID 162	4	-0.7742	-1.223	-0.3255	yes	-0.6331	-0.98204	-0.42674	yes
ID 163	4	-0.6505	-0.9935	-0.3076	yes	-0.7376	-1.0439	-0.41681	yes
ID 164	6	-0.7611	-1.1632	-0.359	yes	-0.7171	-1.0174	-0.22375	yes
ID 165	6	-0.5926	-0.9833	-0.2019	yes	-0.5764	-0.85485	-0.27401	yes
ID 166	6	-0.2994	-0.5802	-0.0187	yes	-0.2155	-0.57711	0.13274	yes

ID 167	6	-0.4376	-1.0002	0.1249	yes	-0.202	-0.62648	0.19595	yes
ID 168	6	-0.2058	-0.6769	0.2653	yes	-0.0869	-0.31126	0.19409	no
ID 169	6	-0.1745	-0.8065	0.4575	yes	-0.0509	-0.62465	0.52657	yes
ID 170	6	-0.3905	-0.8132	0.0321	yes	-0.2288	-0.74322	0.15064	yes
ID 171	6	-0.3745	-0.9611	0.2122	yes	-0.1485	-0.52035	0.2392	yes
ID 172	6	-0.2716	-0.7826	0.2393	yes	-0.1193	-0.56687	0.38141	yes
ID 173	6	-0.4196	-0.947	0.1079	yes	-0.2366	-0.66621	0.14258	yes
ID 175	6	-0.5076	-0.8937	-0.1215	yes	-0.3369	-0.6489	0.03426	yes
ID 176	6	-0.7108	-1.3862	-0.0354	yes	-0.5314	-1.0273	-0.05676	yes
ID 177	6	-0.5323	-1.062	-0.0025	yes	-0.3082	-0.60932	0.12758	yes
ID 178	6	-0.4716	-0.8135	-0.1297	yes	-0.4703	-0.79129	-0.00192	yes
ID 179	6	-0.2954	-0.7761	0.1853	yes	-0.2215	-0.52223	0.12869	yes
ID 180	6	-0.3149	-0.5281	-0.1017	yes	-0.2984	-0.62866	-0.06245	yes
ID 181	6	-0.1854	-0.6144	0.2436	yes	-0.2313	-0.43512	0.01904	no
ID 182	6	-0.4765	-0.7456	-0.2074	yes	-0.4498	-0.78458	-0.07077	yes
ID 183	6	-0.8516	-1.2231	-0.4801	yes	-0.6892	-1.0174	-0.24504	yes
ID 184	5	-0.7927	-1.2296	-0.3557	yes	-0.7322	-1.0099	-0.49942	yes
ID 185	5	-0.7216	-1.2592	-0.184	yes	-0.7935	-1.1246	-0.58133	yes
ID 186	5	-0.6739	-1.2255	-0.1224	yes	-0.7701	-1.1476	-0.386	yes
ID 187	5	-0.7768	-1.0715	-0.482	yes	-0.7273	-1.0085	-0.23995	yes
ID 188	5	-0.7168	-1.1243	-0.3093	yes	-0.5981	-0.92362	-0.31727	yes
ID 189	5	-0.7156	-1.2611	-0.1701	yes	-0.677	-1.1355	-0.21925	yes
ID 190	5	-0.7685	-1.2104	-0.3266	yes	-0.7156	-1.0551	-0.38267	yes
ID 191	5	-0.7097	-0.9144	-0.5051	yes	-0.7022	-0.94804	-0.42077	yes
ID 192	5	-0.8581	-1.1105	-0.6058	yes	-0.7322	-1.0203	-0.3695	yes
ID 193	7	-0.759	-1.2674	-0.2505	yes	-0.9296	-1.2634	-0.7103	yes
ID 194	7	-0.7969	-1.7326	0.1387	yes	-0.8085	-1.5156	-0.07994	yes
ID 196	7	-0.8186	-1.5337	-0.1035	yes	-0.989	-1.466	-0.63793	yes
ID 197	7	-0.857	-1.6292	-0.0849	yes	-0.918	-1.2868	-0.4044	yes
ID 198	7	-0.5454	-1.9195	0.8286	yes	-0.7776	-1.1889	-0.13819	yes
ID 199	7	-0.913	-1.8393	0.0132	yes	-1.0315	-1.5942	-0.45678	yes
ID 200	7	-0.8236	-1.2227	-0.4246	yes	-0.9069	-1.2818	-0.54884	yes
ID 201	7	-0.984	-1.8159	-0.152	yes	-1.1333	-1.5407	-0.85064	yes
ID 202	7	-0.6255	-1.1748	-0.0762	yes	-0.7289	-1.109	-0.42212	yes
ID 203	7	-1.0297	-1.4722	-0.5872	yes	-1.0675	-1.4372	-0.73872	yes
ID 204	7	-0.8861	-1.9709	0.1986	yes	-1.0133	-1.3577	-0.72591	yes
ID 205	7	-0.9832	-2.1556	0.1892	yes	-0.7352	-2.1029	0.01887	yes
ID 206	7	-0.8986	-1.3677	-0.4295	yes	-1.0096	-1.2409	-0.59919	yes
ID 207	7	-0.924	-1.5283	-0.3197	yes	-0.9407	-1.265	-0.5046	yes
ID 208	7	-0.8919	-1.831	0.0473	yes	-0.8863	-1.4517	0.00454	yes
ID 209	7	-0.7593	-1.6943	0.1757	yes	-0.8184	-1.152	-0.28743	yes
ID 210	7	-0.8351	-1.3952	-0.275	yes	-0.6644	-0.96101	-0.22182	yes
ID 211	7	-0.864	-1.4934	-0.2346	yes	-0.8624	-1.1678	-0.5227	yes
ID 212	7	-0.5833	-1.756	0.5895	yes	-0.8023	-1.2084	-0.32336	yes
ID 213	7	-0.7678	-1.589	0.0533	yes	-0.9279	-1.3666	-0.27257	yes

ID 214	7	-0.8765	-1.3271	-0.4259	yes	-0.9727	-1.2263	-0.57924	yes
ID 215	7	-0.7061	-1.6855	0.2733	yes	-0.7134	-1.1951	-0.13971	yes
ID 216	7	-1.0998	-1.5782	-0.6215	yes	-1.0549	-1.3493	-0.74651	yes
ID 217	7	-0.8079	-1.4543	-0.1615	yes	-0.9767	-1.2746	-0.6666	yes
ID 218	7	-0.7543	-1.6176	0.1091	yes	-0.8826	-1.2914	-0.57106	yes
ID 219	7	-0.7908	-1.4203	-0.1612	yes	-0.7479	-1.042	-0.39079	yes
ID 220	7	-0.8409	-1.3857	-0.2961	yes	-0.8802	-1.235	-0.56643	yes
ID 221	7	-0.6004	-1.0615	-0.1394	yes	-0.6524	-0.86088	-0.29877	yes
ID 222	2	-0.3972	-1.2381	0.4437	yes	-0.6109	-0.97529	-0.18218	yes
ID 223	2	-0.4797	-1.1415	0.1822	yes	-0.6715	-1.0525	-0.2096	yes
ID 224	2	-0.4794	-1.1463	0.1875	yes	-0.7115	-1.182	-0.12806	yes
ID 225	2	-0.4246	-1.1536	0.3045	yes	-0.5278	-1.12	0.18017	yes
ID 226	2	-0.4988	-1.2179	0.2203	yes	-0.7094	-1.1991	0.02064	yes
ID 227	2	-0.5072	-1.6433	0.629	yes	-0.4986	-1.239	0.03248	yes
ID 228	2	-0.2466	-1.5268	1.0336	yes	-0.3728	-1.1316	0.08513	yes
ID 229	2	-0.5128	-1.5175	0.4919	yes	-0.7181	-1.2182	-0.08554	yes
ID 230	2	-0.6075	-1.4176	0.2025	yes	-0.6874	-1.2684	-0.1855	yes
ID 231	2	-0.6299	-1.8336	0.5737	yes	-0.8454	-1.2714	-0.13584	yes
ID 232	2	-0.4773	-1.2294	0.2748	yes	-0.703	-1.0718	-0.10091	yes
ID 233	7	-0.8675	-1.3114	-0.4237	yes	-0.8666	-1.1117	-0.64874	yes
ID 234	7	-0.8872	-1.4026	-0.3718	yes	-0.8095	-1.0528	-0.47535	yes
ID 235	7	-0.8292	-1.1561	-0.5023	yes	-0.8123	-1.0575	-0.56161	yes
ID 236	7	-0.7378	-1.3571	-0.1184	yes	-0.8604	-1.3977	-0.38266	yes
ID 237	7	-0.8213	-1.5578	-0.0847	yes	-0.7506	-1.1324	-0.3437	yes
ID 238	7	-0.908	-1.3444	-0.4716	yes	-0.868	-1.2241	-0.44682	yes
ID 239	7	-0.9521	-1.3845	-0.5198	yes	-0.8292	-1.0727	-0.5871	yes
ID 240	7	-0.8689	-1.4408	-0.297	yes	-0.6031	-0.91128	-0.13982	yes
ID 241	7	-0.893	-1.3604	-0.4256	yes	-0.679	-1.0393	-0.39168	yes
ID 242	7	-0.7797	-1.363	-0.1965	yes	-0.7073	-1.0469	-0.21169	yes

414

Table S2: Species-mean-size distribution exponent (SMSDE) and local size-density relationship exponent (LSDRE) estimates for the 149 webs of this study. The entry under "Web name" for soil webs is the web ID number, also used in Fig. 2. Web types are: P = pelagic, E = estuarine, 1 = conventional farm, 2 = organic farm, 3 = intensive farm, 4 = super-intensive farm, 5 = pasture, 6 = forest, 7 = winter farm. Column 3 contains estimates of SMSDE using the bias-corrected maximum likelihood method. Column 4 contains "yes" if the hypothesis that the SMSD was truncated Pareto could not be rejected (1% level) in favour of the generalization mentioned in Methods. Column 5 contains estimates of the LSDRE; column 6 contains "yes" if the hypothesis that the LSDRE was a power law could not be rejected. If the 99% confidence intervals of SMSDE+LSDRE+1 that occur in the last two columns are entirely above -3/4, then they are highlighted in red.

Web name	Web type	SMSDE	Trunc. Pareto?	LSDRE	Power law?	SMSDE+ LSDRE+1	Lower 99% conf.	Upper 99% conf.
TL 84	P	-1.0984	Yes	-0.8413	Yes	-0.9397	-1.1074	-0.7958
TL 86	P	-1.1355	Yes	-0.7461	Yes	-0.8816	-1.0293	-0.6892
Ythan	E	-0.9664	Yes	-1.1347	Yes	-1.1011	-1.2467	-0.9349
ID 95	1	-1.2831	Yes	-0.4627	Yes	-0.7458	-1.0185	-0.4268
ID 96	1	-1.2664	Yes	-0.599	Yes	-0.8654	-1.1562	-0.6651
ID 97	1	-1.2818	Yes	-0.5658	Yes	-0.8477	-1.1241	-0.5129
ID 98	1	-1.3042	Yes	-0.6387	Yes	-0.9429	-1.2585	-0.5243
ID 99	2	-1.264	Yes	-0.4921	Yes	-0.7561	-1.01	-0.5567
ID 100	1	-1.2997	Yes	-0.5634	Yes	-0.863	-1.163	-0.5596
ID 101	1	-1.3047	Yes	-0.6166	Yes	-0.9212	-1.3512	-0.6108
ID 102	1	-1.2889	Yes	-0.6106	Yes	-0.8995	-1.3785	-0.6715
ID 103	1	-1.2803	Yes	-0.5089	Yes	-0.7892	-1.1435	-0.597
ID 104	1	-1.2575	Yes	-0.5242	Yes	-0.7817	-1.048	-0.6005
ID 105	1	-1.2992	Yes	-0.551	Yes	-0.8502	-1.266	-0.6074
ID 106	3	-1.2638	Yes	-0.3684	Yes	-0.6321	-0.8493	-0.2791
ID 107	3	-1.2384	Yes	-0.499	Yes	-0.7374	-1.2939	-0.2889
ID 108	3	-1.2776	Yes	-0.5598	Yes	-0.8374	-1.177	-0.5074
ID 109	2	-1.2405	Yes	-0.4797	Yes	-0.7202	-1.0379	-0.4562
ID 110	3	-1.2967	Yes	-0.5842	Yes	-0.8809	-1.2086	-0.4805
ID 111	1	-1.2463	Yes	-0.5259	Yes	-0.7722	-1.3232	-0.3699
ID 112	1	-1.2537	Yes	-0.5539	Yes	-0.8076	-1.0805	-0.5
ID 113	1	-1.2466	Yes	-0.5879	Yes	-0.8345	-1.1179	-0.6034
ID 114	2	-1.2529	Yes	-0.5036	Yes	-0.7564	-1.1169	-0.5688
ID 115	3	-1.3348	Yes	-0.5289	Yes	-0.8637	-1.2577	-0.4522
ID 116	3	-1.3053	Yes	-0.5185	Yes	-0.8238	-1.1003	-0.4758
ID 117	1	-1.3135	Yes	-0.6165	Yes	-0.93	-1.2707	-0.7003
ID 118	1	-1.2508	Yes	-0.5004	Yes	-0.7512	-1.0799	-0.5604
ID 119	3	-1.2878	Yes	-0.4702	Yes	-0.758	-1.0525	-0.4347
ID 120	3	-1.2589	Yes	-0.6321	Yes	-0.891	-1.2013	-0.713
ID 121	3	-1.2649	Yes	-0.5543	Yes	-0.8192	-1.116	-0.6105
ID 122	1	-1.2612	Yes	-0.5854	Yes	-0.8467	-1.1543	-0.6559

ID 123	3	-1.2491	Yes	-0.604	Yes	-0.8531	-1.1413	-0.5831
ID 124	3	-1.2592	Yes	-0.5625	Yes	-0.8217	-1.0324	-0.4389
ID 125	3	-1.2403	Yes	-0.577	Yes	-0.8173	-1.08	-0.636
ID 126	3	-1.2724	Yes	-0.5421	Yes	-0.8145	-1.0824	-0.4675
ID 127	1	-1.2683	Yes	-0.5003	Yes	-0.7686	-1.0589	-0.4842
ID 128	3	-1.3066	Yes	-0.5989	Yes	-0.9055	-1.2786	-0.45
ID 129	3	-1.2781	Yes	-0.4996	Yes	-0.7778	-1.0436	-0.3402
ID 130	3	-1.2805	Yes	-0.6031	Yes	-0.8836	-1.3223	-0.5238
ID 131	3	-1.2418	Yes	-0.5607	Yes	-0.8025	-1.1433	-0.4478
ID 132	1	-1.3147	Yes	-0.5761	Yes	-0.8908	-1.2088	-0.4785
ID 133	3	-1.2852	Yes	-0.449	Yes	-0.7342	-0.9665	-0.5323
ID 134	3	-1.2819	Yes	-0.5529	Yes	-0.8348	-1.0928	-0.4606
ID 135	3	-1.3192	Yes	-0.5637	Yes	-0.8829	-1.2186	-0.5569
ID 136	1	-1.2699	Yes	-0.4935	Yes	-0.7634	-1.0155	-0.4075
ID 137	2	-1.3029	Yes	-0.5381	Yes	-0.841	-1.1722	-0.6141
ID 138	2	-1.2216	Yes	-0.5327	Yes	-0.7544	-1.0163	-0.5794
ID 139	2	-1.2908	Yes	-0.6489	Yes	-0.9397	-1.3497	-0.512
ID 140	2	-1.2795	Yes	-0.5974	Yes	-0.8769	-1.2773	-0.6646
ID 141	2	-1.3159	Yes	-0.6488	Yes	-0.9648	-1.3163	-0.5496
ID 142	2	-1.3088	Yes	-0.5224	Yes	-0.8312	-1.1121	-0.5246
ID 143	2	-1.3038	Yes	-0.5397	Yes	-0.8435	-1.2534	-0.6512
ID 144	3	-1.2657	Yes	-0.5231	Yes	-0.7888	-1.0882	-0.6016
ID 145	4	-1.2295	Yes	-0.3491	Yes	-0.5786	-0.7963	-0.3014
ID 146	4	-1.2913	Yes	-0.5041	Yes	-0.7953	-1.1394	-0.4986
ID 147	4	-1.277	Yes	-0.4987	Yes	-0.7757	-1.0078	-0.5628
ID 148	4	-1.2763	Yes	-0.5762	Yes	-0.8524	-1.1566	-0.4661
ID 149	4	-1.2435	Yes	-0.5243	Yes	-0.7678	-1.0285	-0.3753
ID 150	4	-1.248	Yes	-0.5622	Yes	-0.8102	-1.151	-0.6131
ID 151	4	-1.2445	Yes	-0.4861	Yes	-0.7305	-0.9795	-0.496
ID 152	4	-1.2531	Yes	-0.4824	Yes	-0.7356	-0.9872	-0.4894
ID 153	4	-1.2516	Yes	-0.5081	Yes	-0.7597	-1.1061	-0.5356
ID 154	4	-1.2509	Yes	-0.4302	Yes	-0.6812	-0.935	-0.4241
ID 155	4	-1.2529	Yes	-0.478	Yes	-0.7309	-1.0233	-0.4845
ID 156	4	-1.271	Yes	-0.5824	Yes	-0.8534	-1.0966	-0.4859
ID 157	4	-1.2161	Yes	-0.1188	Yes	-0.3349	-0.4916	0.3008
ID 158	4	-1.246	No	-0.5866	No	-0.8326	-1.1434	-0.5676
ID 159	4	-1.2435	Yes	-0.5226	Yes	-0.7661	-1.0708	-0.5661
ID 160	4	-1.285	Yes	-0.4709	Yes	-0.7559	-1.0172	-0.4941
ID 161	4	-1.2645	Yes	-0.4881	Yes	-0.7526	-1.0343	-0.5455
ID 162	4	-1.3047	Yes	-0.4099	Yes	-0.7146	-0.9493	-0.4062
ID 163	4	-1.2642	Yes	-0.4608	Yes	-0.725	-0.9739	-0.4575
ID 164	6	-1.2325	No	-0.4656	Yes	-0.6981	-0.8983	-0.2444
ID 165	6	-1.1156	No	-0.4124	Yes	-0.5281	-0.7917	-0.1921
ID 166	6	-1.0779	No	-0.195	No	-0.2729	-0.5653	0.0491
ID 167	6	-1.1145	No	-0.2533	Yes	-0.3678	-0.667	0.1129

ID 168	6	-1.0684	No	-0.075	Yes	-0.1433	-0.3927	0.1835
ID 169	6	-0.8806	Yes	-0.368	No	-0.2486	-0.6254	0.1478
ID 170	6	-1.0991	No	-0.3438	No	-0.4429	-0.7663	-0.0456
ID 171	6	-1.0837	No	-0.2286	Yes	-0.3123	-0.603	0.0954
ID 172	6	-1.1327	No	-0.1245	Yes	-0.2572	-0.5699	0.2005
ID 173	6	-1.0899	No	-0.3184	Yes	-0.4084	-0.6789	-0.0387
ID 175	6	-1.1202	No	-0.2535	Yes	-0.3737	-0.6166	0.0094
ID 176	6	-1.0821	No	-0.679	Yes	-0.7611	-1.0577	-0.3699
ID 177	6	-1.0915	No	-0.2919	Yes	-0.3834	-0.6229	-0.0462
ID 178	6	-1.0692	No	-0.4203	Yes	-0.4895	-0.7609	-0.1668
ID 179	6	-1.0889	No	-0.1913	Yes	-0.2802	-0.5489	0.1003
ID 180	6	-1.1119	No	-0.184	No	-0.2959	-0.5369	-0.0286
ID 181	6	-1.0189	No	-0.1507	Yes	-0.1696	-0.4232	0.1375
ID 182	6	-1.1041	No	-0.2391	Yes	-0.3432	-0.6128	-0.0643
ID 183	6	-1.2227	No	-0.4909	Yes	-0.7136	-0.9273	-0.2805
ID 184	5	-1.2568	Yes	-0.4391	Yes	-0.6959	-0.9325	-0.4062
ID 185	5	-1.2419	Yes	-0.5353	Yes	-0.7772	-1.1082	-0.5614
ID 186	5	-1.2673	Yes	-0.4757	Yes	-0.7429	-1.1149	-0.5253
ID 187	5	-1.2703	Yes	-0.5483	Yes	-0.8187	-1.0507	-0.523
ID 188	5	-1.248	Yes	-0.4286	Yes	-0.6766	-0.9226	-0.4331
ID 189	5	-1.3045	Yes	-0.5089	Yes	-0.8134	-1.1495	-0.4482
ID 190	5	-1.319	Yes	-0.4524	Yes	-0.7714	-1.0747	-0.3387
ID 191	5	-1.2492	Yes	-0.464	Yes	-0.7132	-0.9703	-0.3712
ID 192	5	-1.2628	No	-0.5079	Yes	-0.7708	-1.0173	-0.4114
ID 193	7	-1.2514	Yes	-0.594	Yes	-0.8455	-1.1773	-0.5826
ID 194	7	-1.1027	Yes	-0.5555	Yes	-0.6582	-1.1319	-0.0604
ID 196	7	-1.2723	Yes	-0.6055	Yes	-0.8778	-1.3157	-0.3866
ID 197	7	-1.1321	Yes	-0.6514	Yes	-0.7834	-1.1026	-0.466
ID 198	7	-1.1528	Yes	-0.4967	Yes	-0.6495	-1.1273	-0.1667
ID 199	7	-1.1149	Yes	-0.7434	Yes	-0.8584	-1.2775	-0.4446
ID 200	7	-1.2482	Yes	-0.5635	Yes	-0.8117	-1.1323	-0.5856
ID 201	7	-1.1152	Yes	-0.8369	No	-0.9521	-1.3402	-0.5241
ID 202	7	-1.2489	Yes	-0.4721	Yes	-0.721	-1.0311	-0.4209
ID 203	7	-1.2972	Yes	-0.7047	Yes	-1.0019	-1.3566	-0.6491
ID 204	7	-1.1698	Yes	-0.7309	Yes	-0.9006	-1.3117	-0.5075
ID 205	7	-1.1316	Yes	-0.645	Yes	-0.7766	-1.2325	-0.3042
ID 206	7	-1.2601	Yes	-0.6266	Yes	-0.8867	-1.0802	-0.494
ID 207	7	-1.2692	Yes	-0.6017	Yes	-0.8708	-1.1568	-0.4106
ID 208	7	-1.1701	Yes	-0.7356	Yes	-0.9058	-1.2986	-0.45
ID 209	7	-1.1841	Yes	-0.5596	Yes	-0.7437	-1.1645	-0.3187
ID 210	7	-1.2615	Yes	-0.5315	Yes	-0.7929	-1.0313	-0.339
ID 211	7	-1.2421	Yes	-0.5837	Yes	-0.8258	-1.0654	-0.6147
ID 212	7	-1.1756	Yes	-0.637	Yes	-0.8126	-1.2946	-0.4136
ID 213	7	-1.1261	Yes	-0.7596	No	-0.8857	-1.2793	-0.4687
ID 214	7	-1.273	Yes	-0.6161	Yes	-0.8891	-1.0997	-0.5045

ID 215	7	-1.1193	Yes	-0.5813	Yes	-0.7006	-1.092	-0.2076
ID 216	7	-1.2338	Yes	-0.7921	Yes	-1.0259	-1.3692	-0.6675
ID 217	7	-1.142	Yes	-0.7313	No	-0.8733	-1.2276	-0.4563
ID 218	7	-1.1163	Yes	-0.7022	Yes	-0.8185	-1.2414	-0.4411
ID 219	7	-1.2696	Yes	-0.4019	Yes	-0.6715	-0.9475	-0.0993
ID 220	7	-1.2998	Yes	-0.6071	Yes	-0.9069	-1.1741	-0.5079
ID 221	7	-1.2455	Yes	-0.4276	Yes	-0.673	-0.8679	-0.3253
ID 222	2	-1.1476	Yes	-0.4656	No	-0.6133	-1.1516	-0.1893
ID 223	2	-1.128	Yes	-0.5126	No	-0.6406	-1.1419	-0.2049
ID 224	2	-1.1288	Yes	-0.5005	No	-0.6293	-1.0574	-0.2437
ID 225	2	-1.1222	Yes	-0.4252	No	-0.5475	-1.0518	-0.1673
ID 226	2	-1.0686	Yes	-0.4496	No	-0.5182	-1.0539	-0.0657
ID 227	2	-1.1874	Yes	-0.4597	No	-0.6471	-1.1738	-0.2366
ID 228	2	-1.145	Yes	-0.4847	No	-0.6298	-1.1725	-0.1325
ID 229	2	-1.1317	Yes	-0.4507	No	-0.5823	-1.0763	-0.2076
ID 230	2	-1.182	Yes	-0.4726	No	-0.6546	-1.1646	-0.264
ID 231	2	-1.1148	Yes	-0.6229	No	-0.7377	-1.3119	-0.2581
ID 232	2	-1.113	Yes	-0.5201	No	-0.6331	-1.0564	-0.277
ID 233	7	-1.2358	Yes	-0.5859	Yes	-0.8217	-1.008	-0.4456
ID 234	7	-1.2102	Yes	-0.5821	Yes	-0.7923	-0.9883	-0.3421
ID 235	7	-1.263	Yes	-0.5384	Yes	-0.8014	-1.0062	-0.4733
ID 236	7	-1.1177	Yes	-0.8518	Yes	-0.9695	-1.3664	-0.5881
ID 237	7	-1.1928	Yes	-0.6035	Yes	-0.7962	-1.1101	-0.4596
ID 238	7	-1.265	Yes	-0.5775	Yes	-0.8426	-1.1308	-0.421
ID 239	7	-1.2613	Yes	-0.5146	Yes	-0.7758	-0.9758	-0.4163
ID 240	7	-1.2811	Yes	-0.4593	Yes	-0.7404	-1.021	-0.2475
ID 241	7	-1.2557	Yes	-0.5305	Yes	-0.7862	-0.9884	-0.3141
ID 242	7	-1.2762	Yes	-0.5033	Yes	-0.7794	-0.9762	-0.4019

426

ADDITIONAL REFERENCES

Aban I.B., Meerschaert M.M. & Panorska A.K. (2006). Parameter estimation for the truncated Pareto distribution. *J. Amer. Stat. Assoc.*, 101, 270-277.

Andersen K.H. & Beyer J.E. (2006). Asymptotic size determines species abundance in the marine size spectrum. *Am. Nat.*, 168, 54-61.

Beg M.A. (1983). Unbiased estimators and tests for truncation and scale parameters. *Amer. J. Math. Management Sci.*, 3, 251-274.

Brown J.H. & Gillooly J.F. (2003). Ecological food webs: High-quality data facilitate theoretical unification. *Proc. Natl. Acad. Sci. USA*, 100, 1467-1468.

Brown J.H., Gillooly J.F., Allen A.P., Savage V.M. & West G.B. (2004). Toward a metabolic theory of ecology. *Ecology*, 85, 1771-1789.

Burnham K.P. & Anderson D.R. (2002). *Model Selection and Multimodel Inference: A Practical Information Theoretic Approach*. 2 edn. Springer, New York.

Carpenter S.R. & Kitchell J.F. (1993). *The Trophic Cascade in Lakes*. Cambridge University Press, Cambridge.

Cohen J.E., Jonsson T. & Carpenter S.R. (2003). Ecological community description using the food web, species abundance, and body size. *Proc. Natl. Acad. Sci. USA*, 100, 1781-1786.

Hall S.J. & Raffaelli D. (1991). Food web patterns: lessons from a species rich web. *J. Anim. Ecol.*, 60, 823-842.

Jennings S. & Mackinson S. (2003). Abundance-body mass relationships in size-structured food webs. *Ecol. Lett.*, 6, 971-974.

Johnson N.L., Kotz S. & Balakrishnan N. (1994). *Continuous Univariate Distributions*. 2 edn. John Wiley & Sons, Inc., New York.

Jonsson T., Cohen J.E. & Carpenter S.R. (2005). Food webs, body size, and species abundance in ecological community description. *Adv. Ecol. Res.*, 36, 1-84.

Mulder C., Cohen J.E., Setälä H., Bloem J. & Breure A.M. (2005a). Bacterial traits, organism mass, and numerical abundance in the detrital soil food web of Dutch agricultural grasslands. *Ecol. Lett.*, 8, 80-90.

Mulder C., Van Wijnen H.J. & Van Wezel A.P. (2005b). Numerical abundance and biodiversity of below-ground taxocenes along a pH gradient across the Netherlands. *J. Biogeogr.*, 32, 1775-1790.

Nelder J.A. & Mead R. (1965). A simplex method for function minimization. *Computer Journal*, 7, 308-313.

Page R. (1968). Aftershocks and microaftershocks of the great Alaska earthquake of 1964. *Bull. Seismol. Soc. Am.*, 58, 1131-1168.

Reuman D.C., Mulder C., Banasek-Richter C., Cattin Blandenier M.-F., Breure A.M., Den Hollander H., Kneitel J.M., Raffaelli D., Woodward G., Cohen J.E. (In press) Allometry of body size and abundance in 166 food webs. *Adv. Ecol. Res.*

Rossberg A.G., Ishii R., Amemiya T. & Itoh K. (2008). The top down mechanism for body-mass abundance scaling. *Ecology*, 89, 567-580.

Silvert W. & Platt T. (1980). Dynamic energy-flow model of the particle-size distribution in pelagic ecosystems In: *Evolution and Ecology of Zooplankton Communities* (ed. Kerfoot WC). The University Press of New England Hanover, N.H., pp. 754-763.

White E.P., Enquist B.J. & Green J.L. (2008). On estimating the exponent of power-law frequency distributions. *Ecology*, 89, 905-912.