

## LETTER

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

Daniel C. Reuman,<sup>1\*†</sup> Christian Mulder,<sup>2</sup> Dave Raffaelli<sup>3</sup> and Joel E. Cohen<sup>1</sup>

<sup>1</sup>Laboratory of Populations, Rockefeller University, 1230 York Avenue, Box 20, New York, NY 10065, USA

<sup>2</sup>Department of Ecology, RIVM, 9 Antonie van Leeuwenhoeklaan, Box 1, Bilthoven 3720 BA, The Netherlands

<sup>3</sup>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

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  $\lambda$  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

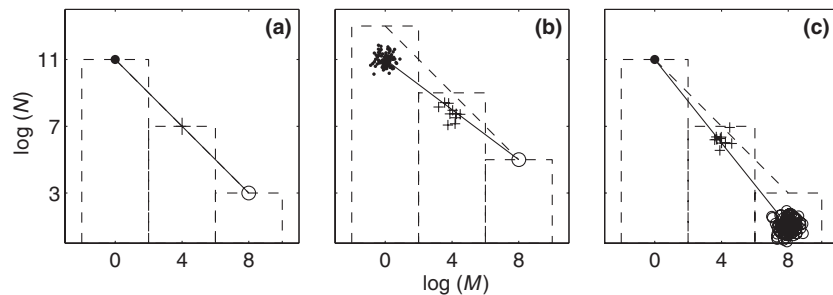
$$\frac{\bar{M}_{B_l}^{3/4} \sum_l N_{B_l}}{\bar{M}_{I_m}^{3/4} \sum_m N_{I_m}} = \frac{\bar{M}_{I_m}^{3/4} \sum_m N_{I_m}}{\bar{M}_{T_n}^{3/4} \sum_n N_{T_n}} = 10. \quad (1)$$

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

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

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  $\log(\bar{M})$  values at the bin centres also gives  $-1$ , an estimate of the abundance spectrum and therefore of ISDE + 1. The  $\log(D)$ -vs.- $\log(\bar{M})$  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



**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,  $\text{LSDRE} = (\text{ISDE} + 1) - (\text{SMSDE} + 1) = \text{ISDE} - \text{SMSDE}$ .

species, Fig. 1b; respectively, one basal, 10 intermediate, and 100 top, Fig. 1c), then eqn 2 implies that the  $\log(N)$ -vs.- $\log(M)$  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,  $\text{LSDRE} = (\text{ISDE} + 1) - (\text{SMSDE} + 1) = \text{ISDE} - \text{SMSDE}$ . 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  $\alpha$ , and denote the average consumer-to-resource body-mass ratio by  $\beta$ ; 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  $R = [M, M + dM]$ , defined as the total growth biomass plus the total reproduction biomass of these organisms per unit time, is approximately  $\text{Prod}(M)dM$ , where  $\text{Prod}(M) \propto f_I(M)M^{3/4}$  (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  $\text{Prod}(M)dM$  occurs mainly within the mass range  $R$ , or close to it relative to the consumer–resource mass ratio  $\beta$ . 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

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

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

Because  $\beta$  does not vary systematically with  $M$ , individuals in the body mass range  $R$  eat mainly individuals in the range  $R/\beta = [M/\beta, M/\beta + dM/\beta]$ . Because population production  $\text{Prod}(M/\beta)dM/\beta$  of the mass category  $R/\beta$  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  $\alpha$  times the production of  $R/\beta$ :

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

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  $f_{tp} \propto M^{\lambda_{tp}}$  on  $[a_{tp}, b_{tp}]$  and is  $f_{tp} = 0$  outside  $[a_{tp}, b_{tp}]$  (where  $\lambda_{tp}$  can be any real number and  $0 < a_{tp} < b_{tp}$ ) 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  $f_I(M) \times dM$  and those organisms are partitioned into a number of species proportional to  $f_S(M) \times dM$ . 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  $\alpha$ ) 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  $\beta$ . 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

with  $\bar{M}$ . LSDRE can be greater than, less than, or equal to  $-3/4$ , whereas  $\text{ISDE} + 1$  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)  $\text{ISDE} + 1 = \log(\alpha)/\log(\beta) - 3/4$  (eqn 7); (4)  $\text{LSDRE} + \text{SMSDE} + 1 = \log(\alpha)/\log(\beta) - 3/4$  (eqn 10). We tested (1) and (2) by comparing the power law hypotheses against more general alternative models.

To test (3) and (4),  $\text{ISDE} + 1$  and  $\text{LSDRE} + \text{SMSDE} + 1$  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  $\text{ISDE} + 1$  (respectively,  $\text{LSDRE} + \text{SMSDE} + 1$ ) contained  $\log(\alpha)/\log(\beta) - 3/4$  for  $\alpha$  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  $\text{ISDE} + 1$  (respectively,  $\text{LSDRE} + \text{SMSDE} + 1$ ) 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  $\log(N) = c \times \log(\bar{M})^2 + d \times \log(\bar{M}) + e$  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

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

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

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

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

Confidence intervals for LSDRE, SMSDE, and  $\text{LSDRE} + \text{SMSDE} + 1$  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  $\text{LSDRE} + \text{SMSDE} + 1$  were calculated for each resampling. These confidence intervals of  $\text{LSDRE} + \text{SMSDE} + 1$  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.  $\text{ISDE} + 1$  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.  $\text{ISDE} + 1$  was the slope of this regression. Confidence intervals of  $\text{ISDE} + 1$  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  $\text{ISDE} + 1$  may have been biased toward 0 (Appendix S5).

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  $\alpha$ , 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

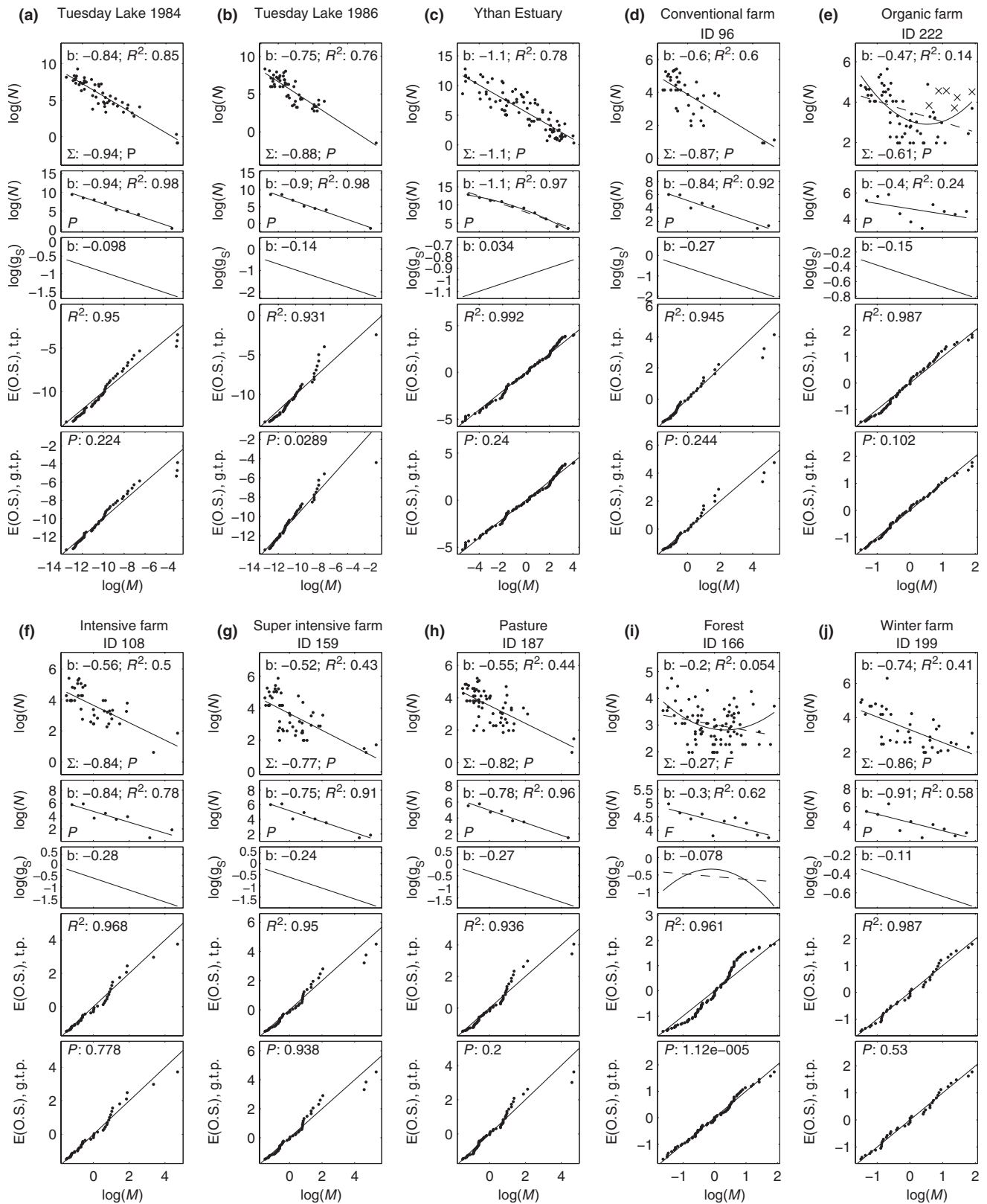
All non-forest webs and one forest (ID 169) had truncated Pareto SMSD. For ID 169 and all non-forests except super-intensive farm ID 158 and pasture ID 192, the null-hypothesis that the SMSD was truncated Pareto could not be rejected in favour of the generalized distribution of eqn 12 (Figs 2 and 4; 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).

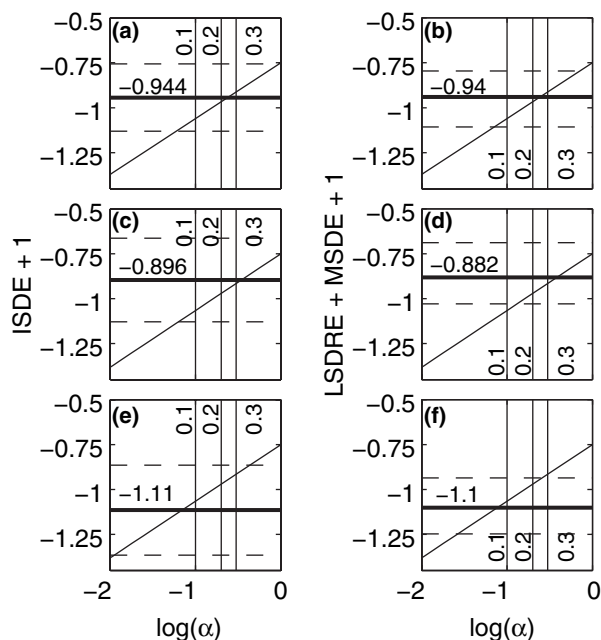
In contrast, all forests except ID 169 had truncated log normal SMSD. The null-hypothesis of a truncated Pareto SMSD was rejected for those webs (Table S2). The best-fitting generalized distributions differed substantially from a truncated Pareto (Fig. 2), and had negative quadratic coefficient, implying 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





**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  $\alpha$ . Vertical lines are labelled with benchmark  $\alpha$ -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  $\alpha$  (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).

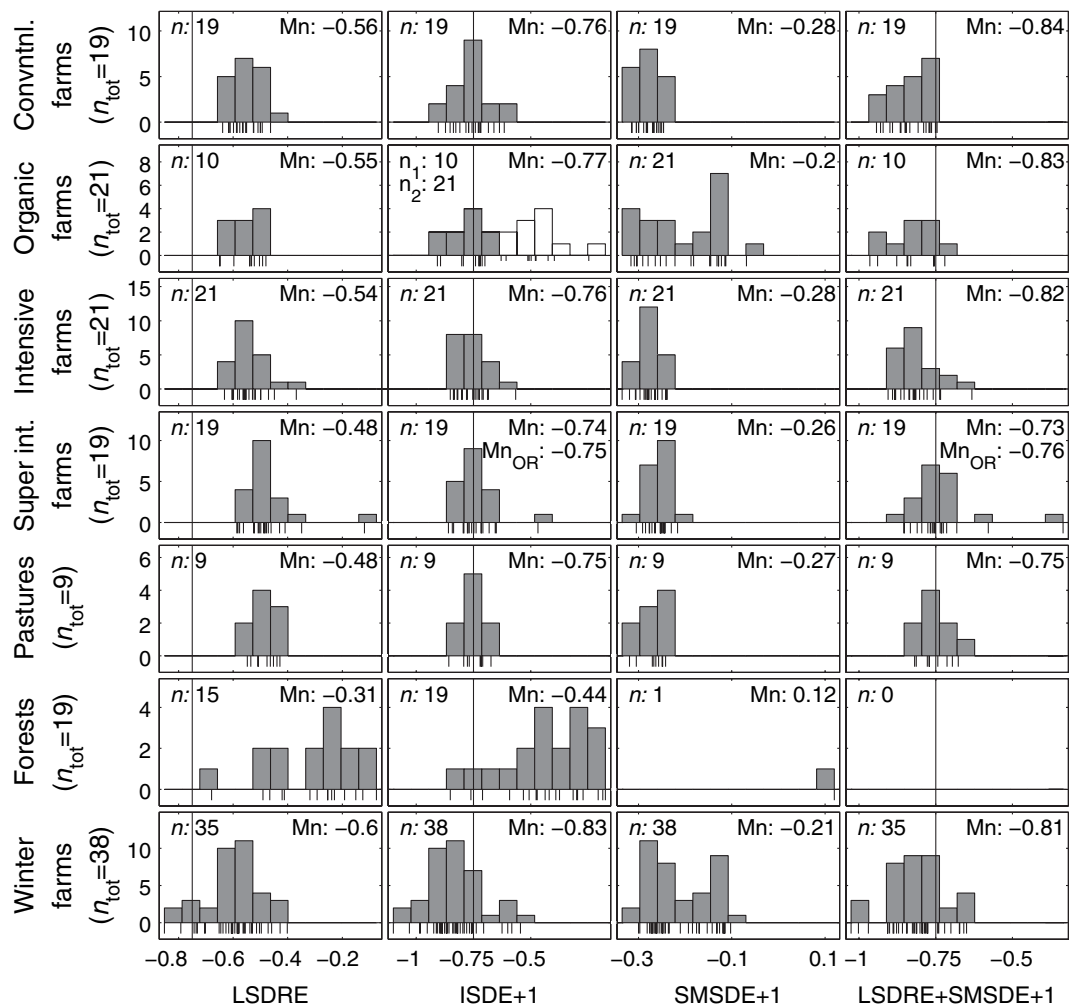
$LSDRE < ISDE + 1$  was predicted when  $SMSDE + 1 > 0$ , and  $LSDRE > ISDE + 1$  was predicted when  $SMSDE + 1 < 0$  (eqn 10).  $SMSDE + 1$  was positive only for the Ythan Estuary and forest ID 169, for which  $LSDRE < ISDE + 1$  held. For 139 of the 147 webs with  $SMSDE + 1 < 0$ ,  $LSDRE > ISDE + 1$  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  $k = \text{mean}(LSDRE)$ ] and  $LSDRE = ISDE + 1$  (which corresponds to omitting the correction for systematic changes in diversity). Using all 149 sites, the sum of squares of  $LSDRE - \text{mean}(LSDRE)$ , 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).  $\Sigma$  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.



**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<sub>OR</sub> 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

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  $\alpha$  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  $R = [M, M + dM]$  was distributed mainly over nearby (relative to the average consumer–resource mass ratio  $\beta$ ) parts of the web's range of body masses; we then related the production of consumers and resources by  $\text{Prod}(R) = \alpha \text{Prod}(R/\beta)$  (eqn 4). If only a fraction,  $\rho$ , of  $\text{Prod}(R/\beta)$  (e.g. the growth component) had body mass vulnerable to consumption by individuals in  $R$ , one could set  $\text{Prod}(R) = \alpha \times \rho \times \text{Prod}(R/\beta)$ , but further modification would be needed to reflect production from other mass categories occurring in  $R/\beta$ , for instance if the size at birth or hatching of a larger species fell in  $R/\beta$ .

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 through 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  $\text{SMSDE} < -1$ , corresponding to a linear diversity spectrum of slope  $\text{SMSDE} + 1 < 0$ . 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

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  $i$ th-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 inter-annual 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., Krumsins, 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.

- 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 taxocenoses 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 Blandinier, 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  $\lambda_{tp}$  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  $\lambda_{tp}$  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.

**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