Body size and the relative abundance of species

Werner Ulrich

Department of Animal Ecology, Institute of Ecology and Protection of Environment, Nicolaus Copernicus University, Gagarina 9, 87 – 100 Toruń, Poland
e-mail: ulrichw@uni.torun.pl

Abstract. Existing models of species abundance distributions (SADs) can be divided into those that are based on concepts of common limited niche space (niche apportionment models, neutral models) and those that invoke standard statistical distributions (e.g. log-series, lognormal). While the first type of models assumes that competitive interactions lead to observed SADs, the models of the second type appear to be mainly statistical descriptors of SADs without deeper biological meaning. None of the models explicitly includes species body size as a factor influencing species abundances. Further, with the exception of recent neutral models they are not embedded into basic ecological and evolutionary models to explain local diversity and ecosystem functioning. Here I present a new random walk model of species abundances that is based on two well known ecological distributions, the abundance – body weight distribution and the species – body weight distribution to define long-term upper abundance boundaries (carrying capacities). I show that a simple random walk of species abundances around the carrying capacities not only generates observed SADs but is also able to explain other patterns of community structure like core – satellite distributions, temporal patterns of species turnover, variance – mean ratios, and biomass distributions.

Key words: species – abundance distribution, dominance – rank orders, lognormal distribution, random walk, core – satellite species, metabolic theory, temporal variability.

1. Introduction

The study of species – abundance distributions (SADs) has a long tradition in ecology (for reviews see Tokeshi 1999; Magurran 2003, 2005), starting from early distribution orientated models of Motomura (1932) and Fisher et al. (1943) to recent neutral (Bell 2001; Hubbell 2001; Ulrich 2007; Ulrich & Zalewski 2007) and fractal (Mouillot 2000) approaches. The study of SADs is closely related to the question about the causes of commonness and rarity (Gaston 1994; Kunin & Gaston 1997; Magurran 2007; McGill et al. 2007).

Much attention gained SAD models that involve the concept of ecological niches. They assume a common limiting resource that is divided among the species of a local community. Most prominent among these models are the broken stick model of MacArthur (1957), the sequential breakage models of Sugihara (1980), and the niche apportionment models of Tokeshi (1990, 1996). According to these models a common limiting resource (or the total niche space) is divided among the species by a two step process (Tokeshi 1996, 1999). First the niche space is divided into two parts. The second step implies the choice of one of these parts for further division. Different probability distributions for both steps generate now a whole series of models.

However, despite of the seeming simplicity of the algorithms all these models rely on several hidden assumptions that make them difficult to apply and to test. First, the concept of a common limiting resource that is divided among the species (MacArthur 1957; Tokeshi 1990) immediately implies severe interspecific competition. Further, niche division is assumed to be the dominating process in assigning abundances. Any abundance influencing mechanism other than resource division would bias the apportionment algorithms and produce deviating SADs. Second, niche based
models do not consider fluctuations in abundance while either dealing only with point estimates of abundance (for instance the yearly abundance) or while averaging over longer times. Third, the models do not include temporal and spatial species turnover (Ulrich 2001, 2007). Hence these models either assume communities to be stable with respect to its species composition or silently assume that immigrating species occupy preexisting empty niches after local species extinction. Fourth, niche apportionment models necessarily apply only to the local, the ecosystem scale. At the regional scale common limiting resources for which all species compete are hard to envision.

A second group of models assumes that SADs follow certain statistical distributions. Most prominent among these are the log-series (Fisher et al. 1943) expected from random samples from a spatially or temporally Poisson distributed community (Tokeshi 1999; Magurran & Henderson 2003) and the lognormal (Preston 1962) frequently associated with the central limit theorem of statistics (May 1975). While these models are less forced by the problems of the niche apportionment models they appear to be mainly statistical descriptors of SADs without deeper biological meaning (Tokeshi 1999). Their parameters are difficult to interpret within the framework of existing ecological models on abundance and diversity. Nevertheless they are used as null models (McGill 2003; McGill et al. 2007) assuming that any deviation for instance from the lognormal should be caused by species specific and environmental factors (Ugland & Gray 1982; Nummelin 1998; but see Williamson & Gaston 2005).

Recently, neutral model approaches (Solé & Alonso 1998; Hubbell 2001; Bell 2001) proved to produce SADs similar to those found in nature (Chave 2004; Ulrich 2004; McGill et al. 2006). According to the most elaborate of these approaches, the ecological drift model (EDM) of Hubbell (2001), local communities are assembled (irrespective of species membership) by simple birth and death processes, by dispersion, and by speciation events. Nevertheless, to produce realistic results EDM has to assume that the total number of individuals (the carrying capacity) of the focal community is fixed and limited by the available amount of resources. Hence, EDM still relies on competition (between individuals) for a limited total niche space (Hubbell 2003; Hubbell & Lake 2003) as a main driver of community assembly. However, it switches the focus from species as the major players in the competition game to individuals abolishing thus the distinction between inter- and intraspecific competition. As the niche apportionment models it is a theory for ecologically equivalent species.

None of the above models considers body size as a factor that affects abundances. The neutral and niche orientated models rather explicitly (Hubbell 2001) or implicitly assume that all species are of similar size because different sized species would also be limited by different resources contrary to the basic assumptions of these models. However, even a short survey of the literature (Ulrich & Ollik 2003) reveals many examples where niche models were applied to communities spanning over several magnitudes of body size. On the other hand, there is an extensive literature on the shapes of abundance – body weight distributions (AWDs; cf. Blackburn & Gaston 1997 and Gaston & Blackburn 2000 for reviews) and the recent development of the metabolic theory of ecology (Brown et al. 2004; Savage et al. 2004) highlights the need to incorporate body size into models of relative abundance.

2. The model

The present model uses two basic ecological distributions. First, most communities have more medium sized species than very large or very small ones. This well know species – body weight distribution (SWD) is for many vertebrate communities right skewed lognormal (Kozlowski & Gavelczyk 2002; Smith et al. 2004), whereas for invertebrate communities the available evidence rather points to approximately symmetrical lognormal distributions (Chislenko 1981; Novotny & Kindlmann 1996; Ulrich 2006). The present model uses the well-known Weibull distribution to generate symmetric and skewed lognormal distributions of species numbers S along the body weight axis

$$S = S_0 \left( \frac{w}{\eta} \right)^{\beta - 1} e^{-\left( \frac{w}{\eta} \right)^{\beta}}$$

(1)

where $\beta$ and $\eta$ are the shape parameters and $w$ denote log-transformed body weight classes.

Second, according to the metabolic theory of ecology (Brown et al. 2004) long-term carrying capacities $K$ of species (equilibrium abundances) scale allometrically to body weight $W$ (the carrying capacity – body weight distribution KWD). Despite some discussion around the slope value (White & Seymour 2003; Kozlowski & Konarzewski 2004; Brown et al. 2005; Ulrich et al. 2005; Farrell-Gray & Gotelli 2005; Reich et al. 2006) there is ample evidence that at least animal KWDs scatter around a slope of – 0.75 (Brown et al. 2004, 2005; Savage et al. 2004). Hence

$$K = K_0 W^{-0.75}$$

(2)

with $K_0$ being a normalizing constant. The combination of the SWD (eq. 1) and the KWD (eq. 2) provides now the distribution of carrying capacities for a community of $S$ species spanning over a range of body sizes $W$ (Fig. 1). This distribution is predicted to be left skewed lognormal in the case of a right skewed SWD (Fig. 1B) and lognormal in the case of a symmetrical SWD (not shown).
However, species abundances are not fixed at the carrying capacities. The present model further assumes that species abundances \( N \) fluctuate randomly in time around the respective carrying capacities \( K \). Such temporal abundance fluctuations can be modelled by a simple random walk around \( K \) (the drunkeners walk).

\[
N_{i+1} = N_i \text{norm}(0, \sigma) + \phi(K, N_i)
\]

In this model \( \text{norm}(0, \sigma) \) is a normally distributed random variate with mean zero and standard deviation \( \sigma \). The additive term \( \phi(K, N) \) shifts the mean of \( \text{norm} \) to a positive value (higher probability of population increase) for abundances below \( K \) and to a negative value (higher probability of population decrease) for abundances above \( K \). Most simply one can assume that the strength of the attractor \( K \) is similar for all species. To make the value of \( \phi \) dependent on the distance of \( N_i \) from \( K \) we can define

\[
\phi(K, N_i) = \frac{\ln(K) - \ln(N_i)}{2\ln(K)}
\]

Such a random walk has a fixed lower boundary (zero) where a species goes locally extinct. However, there is no sharp upper boundary. Species might achieve abundances well above \( K \). This is not uncommon and for instance arthropods frequently reach abundance peaks that are several hundred times higher than normally. Below I will show that this inherent asymmetry in abundance boundaries predicts not only observed SAD shapes but also a series of other patterns frequently observed in real communities.

### 3. Methods

For the simulations below I used either the right skewed lognormal SWD model \( (\beta = 2.5; \text{Fig. 1}) \) or a symmetrical lognormal SWD with 200 species and \( \eta = 0.5 \) (one \( \log_2 \) weight class; a community of similar sized species) and \( \eta = 5 \) (ten \( \log_2 \) weight classes; a community spanning over three orders of magnitude in body weight, shown in Fig. 1). \( K_0 \) was arbitrarily set to 10,000. These parameter settings assign the least abundant species carrying capacities around 150, values well above the extinction threshold of one.

For each random walk I used 10,000 steps. This number proved to be enough to get a stable pattern. If a species went extinct \( (N_i < 1) \) it was replaced by a new species of the same body weight, which started at low abundance \( (N_i = 2) \). I used the last 50 time steps to study extinction probabilities and temporal patterns of species occurrences. All computations were done using the program RAD (Ulrich 2002, available at http://www.uni.torun.pl/~ulrichw).

I used the mean – variance ratio in form of the index of Lloyd (1967) \( J = \sigma^2 / \mu^2 - 1/\mu + 1 \) to assess the degree of temporal variability. Regression slopes refer always to model I least square regressions. Basic statistics were done with Statistica 7.1 (StatSoft 2005). Errors refer always to standard errors.
4. Model predictions

4.1. SAD shapes

In the case of low temporal variability ($\sigma = 0.01$) and a right skewed SWD ($\beta = 2.5$) the present models predicts asymmetrical sigmoid SADs (left skewed in Preston plots) irrespective of whether the community is made of similar sized species (Fig. 2A) or spans over a wide range of body sizes (Fig. 2B). Higher temporal variability ($\sigma = 2$), shifts the SADs towards shapes that lack their lower curvature typical for the sigmoid class of models (Fig. 2C, D). The distributions become more and more right skewed (Fig. 2C, D).

In the case of low temporal variability ($\sigma = 0.01$) asymmetrical right skewed SWDs produce left skewed S-shaped SADs at wider ranges of body sizes. However, for communities of similar sized species (one to two log$_2$ body size classes) the model predicts symmetrical SADs (Fig. 3A). A symmetrical SWD produced symmetrical S-shaped SADs irrespective of the range of body sizes (Fig. 3B). A higher temporal variability ($\sigma = 2$) changes this picture. Irrespective of the SWD shape a trend towards right skewed SADs at higher ranges of body sizes is predicted (Fig. 3C, D). The SADs become less sigmoid shaped while progressively lacking their lower curvature (Fig. 2C, D). However, communities of similar sized species are again predicted to have symmetrical sigmoid shaped SADs (Fig. 3C, D).

4.2. Temporal patterns

At low temporal variability (low values of the parameter $\sigma$) species remain near their carrying capacities during the random walk and extinction events ($N_t < 1$) are rare. Temporal species turnover is predicted to increase with higher temporal variability (higher values of $\sigma$) due to the accumulation of rare species that enter the assemblage infrequently. The result is a significant bimodal distribution of persistence with a group of core species that persisted over time and a group of satellite species that occurred infrequently (Fig. 4).

Intuitively, the model parameter $\sigma$ that influences the degree of variability around the carrying capacity should directly influence the temporal variability of species. In-
Body size and the relative abundance of species

Indeed, a higher model parameter $\sigma$ resulted in a higher temporal variability (estimated by the Lloyd index $J$) than at lower values of the parameter $\sigma$. In the latter case, $J$ has a value around unity independent of the mean abundance. At higher temporal variability, $J$ increases with mean abundance (Fig. 5A). However, irrespectively of a low or high temporal variability parameter $\sigma$ and the span of body weights, the present model predicts that temporal variability is related to mean abundance according to a proportional rescaling process in which the variance in abundance scales to the square of mean abundance (Fig. 5B).

The model further relates extinction probabilities to the degree of temporal variability $\sigma$ and to the carrying capacity $K$. The higher $\sigma$ and the lower $K$ were the higher was the probability of extinction. Extinction probability is predicted to be inversely correlated to $K$ (Fig. 6A) and to the mean abundance of a species (Fig. 6B). Mean abundance appeared to be a better predictor of extinction probability than $K$ (Fig. 6).

### 4.3. Biomasses

At high temporal variability ($\sigma = 2$), the present model generated typical AWDs that were in most model runs triangular (Fig. 6A). Using arithmetic mean abundances per log$_2$ weight class AWDs were only approximately allometric and had always slopes less then the value of $-\frac{3}{4}$ predicted for the KWD slope (Fig. 7A). Further, the resulting biomass – body weight distribution (BWD) is characterized by a high scatter of data points. Again, using mean

---

**Figure 3.** The dependence of the skewness of the SAD on the range of body sizes. A: symmetrical SWD, low variability ($\sigma = 0.01$); B: right skewed SWD, low variability ($\sigma = 0.01$); C; symmetrical SWD, high variability ($\sigma = 2.00$); D; right skewed SWD, high variability ($\sigma = 0.01$)

**Figure 4.** High temporal variability (same assemblage as in Fig. 2A) produces a growing number of infrequent species resulting in a core – satellite species distribution of temporal occurrence. Given is the frequency of species that occurred in the last 50 steps of the random walk
abundances per weight class the model predicts an initial increase in biomass and a rather constant biomass at medium to high weight classes (Fig. 7B).

At low temporal variability (σ = 0.01) AWDs are predicted to follow power functions with slopes similar to the underlying KWD (Fig. 7C). Biomass increased with body weight (Fig. 7D) and the slopes of the BWD scattered around 1/4 for an initial KWD slope of –0.75 (not shown).

5. Discussion

The present model is based on three major assumptions, an allometric KWD with a negative slope, a lognormal SWD, and a random walk of species abundances without sharp upper abundance boundary. In this way the model reduces the task to explain observed SADs to the explanation of two basic ecological distributions. The allometric KWD is well established (Cyr et al. 1997; Brown et al. 2004) and recent metabolic theory (Savage et al. 2005; Brown et al. 2005) seems to come close to a general derivation of this distribution from first principles. The lognormal SWD is also one of the best documented macroecological patterns (Loder 1997; Gaston & Blackburn 2000; Kozłowski & Gawełczyk 2002; Ulrich 2006). However, there are still several different models that aim at explaining the pattern (McKinney 1990; Maurer et al. 1992; Brown et al. 1993; Kozłowski & Weiner 1997; Kindlmann et al. 1999; Ulrich 2006).

Random walks are a common tool for modelling temporal variability of species abundances, dispersal, species
turnover, and extinction probabilities (Foley 1994; Byers 2001). At first sight it seems therefore astonishing that there are no elaborated random walk models of relative abundance although there are some scattered attempts in the statistical literature (Engen & Lande 1996; Bramson et al. 1999; Ulrich 2001). The dynamics model of Hughes (1984) can also be viewed as a random walk model although it has never been tested seriously due to its complicated structure. One reason for this lack of true random walk SAD models is surely that it is difficult to invoke upper and lower abundance boundaries that have reasonable ecological interpretations (Pollard et al. 1987; Crowley 1992). A simple random walk with reflecting boundaries (the Ornstein Uhlenbeck model) gives always symmetrical lognormal SADs. Further, fixed boundaries have an undesired effect on evenness. Because at larger community size more species are placed within the abundances boundaries evenness will be closely correlated to species richness. These are not realistic features as has been demonstrated by Williamson & Gaston (2005). In turn, random walks without boundaries result immediately in unrealistic high abundances of some species and the span of abundances is closely correlated to the number of model steps.

The present model overcomes these problems by referring to species specific mean abundances (long-term carrying capacities) around which abundances fluctuate. This seems a realistic model feature. Many long-term studies of abundance fluctuations identified a pattern where trend corrected medium abundances were more frequent than very high or very low ones (Taylor et al. 1980). Further, in natural populations the variability – mean relationship is frequently of the allometric rescaling type ($\sigma^2 = \mu^z$) with $z$ being around 2 (Taylor’s power law, Taylor 1961; Taylor et al. 1982). The present model predicts the same pattern (Fig. 5B).

The present model is still species centred like the niche and the distribution oriented classes of models. However, it differs from the niche models in not referring to any common resources or niche space. It thus overcomes the problems these models are confronted with. The model can also be applied to different ecological scales from local to continental, while all niche orientated models strictly apply only to a local scale where it can reasonable be assumed that competitive effects influence community structure. Neutral models of relative abundance, on the other hand, are basically individual orientated and do not differentiate

Figure 7. A: At high temporal variability ($\sigma = 2$), the random walk model produces AWDs typical for many animal assemblages. In the majority of runs a marked triangular shape appeared. B: the resulting biomass – body weight relationship. At low temporal variability ($\sigma = 0.01$) AWDs and BWDs are predicted to follow power functions. Open circles: raw values, black squares: arithmetic mean values per log body weight class. Regression slopes for the mean values: A: $z = -0.63$; B: $z = 0.33$; C: $z = -0.79$; D: $z = 0.21$. Data from the same assemblages as in Figure 2.
between intra- and interspecific processes (Bell 2001; Hubbell 2001; Chave 2004; Ulrich 2004). They refer to ecologically equivalent individuals of the same trophic level and of similar body size. Hence, they apply to local and regional spatial scales and also predict patterns of temporal variability. Neutral models predict skewed sigmoid SAD shapes at the local and log-series SADs at the regional, the metacommunity scale (at least under point mutation specification, Hubbell 2001). The latter prediction differs from that of the present model which predicts the same SAD shapes at different spatial scales. Regional or continental SAD shapes have been studied only for a small number of species rich taxa Ulrich and Ollik (2003), Nee et al. (1992), Cantos and Gomez-Mazaneque (1998), Blackburn and Gaston (1997) reported asymmetric sigmoid SADs and an excess of rare species for birds. A similar pattern was found in tropical trees (Hubbell 2001). The data of Purtauf et al. (2005) however imply a log-series SAD for the regional abundance of ground-beetles.

The distribution oriented models like the lognormal, the geometric, and the log-series were always seen as mere descriptors of community structure without deeper ecological interpretation of their parameters (May 1975). The same critiques holds for the niche orientated models. The niche division probabilities they use were never critically examined and interpreted. The three parameters of the present model, however, have clear ecological interpretations. They refer to the KWD slope that follows from metabolic theory, the skewness of the SWD and the degree of temporal variability \( \sigma \). All three parameters can in principle be measured and be used to derive the associated SAD. Comparing this SAD with the observed one allows therefore for a straightforward testing of the model. This feature differs from the niche based and the neutral models which are notoriously hard to falsify.

The present model makes a series of predictions about the shape of SADs that can be compared to those of the above two classes of models. First, it predicts that the basic SAD shape is sigmoid. This contrasts to all models that predict geometric SAD shapes like the Motomura model (Motomura 1932), the random assortment models of Tokeshi (1990) or power function shapes like the fractal model of Mouillot et al. (2000). However, a clear sigmoid shape should only appear at comparably low temporal variability. At higher variability and therefore a higher degree of species turnover (Fig. 4), the SAD is predicted to lack its lower curvature and to have in its lower part a shape similar to a log-series. Indeed, Fisher et al. (1943) derived the log-series under the assumption that species occurrence (for instance due to an extinction/immigration equilibrium) is a random Poisson process (Magurran & Henderson 2003; Chave 2004).

Second, the present model predicts that the skewness of the SAD depends on the range of body sizes within the community. A marked skewness is only predicted for communities of different sized species. Left skewed SADs that captured so much recent attention (Tokeshi 1999; Gaston & Blackburn 2000) are only predicted for communities with also skewed underlying SWDs and comparably low degrees of temporal variability (Fig. 3). Indeed the best examples for left skewed SADs with an excess of rare species stem mainly from birds (Nee et al. 1991; Gaston & Blackburn 2000) and trees (Hubbell 2001) that have comparably stable population sizes. Most reported invertebrate SADs, in turn, are right skewed (Hughes 1984; Morse et al. 1988; Novotny & Basset 2000; Ulrich 2005). Nevertheless, the relationship between SAD skew and body size has apparently not been studied systematically.

Third, a series of recent papers (Magurran & Henderson 2003; Ulrich & Ollik 2004; Connolly et al. 2005; Ulrich & Zalewski 2006, 2007) reported different SAD shapes of spatially or temporally defined core and satellite species. Core species were found to follow lognormal distributions. Satellite species, on the other hand, appeared to follow either a log-series (Magurran & Henderson 2003) or power functions (Ulrich & Ollik 2004; Ulrich & Zalewski 2006) as expected if spatial or temporal occurrence patterns were self-similar. The present model predicts exactly such a distinction (Fig. 8) with core species having a symmetrical sigmoid SAD and satellite species following a power function SAD. This result implies that it might not be necessary to invoke ecological differences between core and satellite species to explain observed SAD shapes (Magurran & Henderson 2003). It might most parsimonious be explained as being a side effect of the temporal processes that generate observed SADs.

Fourth, the present model predicts shapes of body size dependent distributions. In particular it explains the often observed triangular shape of the ABW as a statistical artefact due to uneven species numbers per size class (Fig. 7A, C). A similar explanation has already been favoured by Ulrich (1999) and Gaston & Blackburn (2000). Further, metabolic theory predicts the total biomass of a community to rise allometrically to body weight with a slope of \( \frac{1}{4} \) (Brown et al. 2004). The present model predicts a very similar pattern (Fig. 7D). However, at higher temporal variability the scatter of data points causes that community biomass rises significantly only at lower body weights. For higher body weights statistical averaging on a log scale causes biomasses per weight class to be rather constant (Fig. 7B). Very similar patterns have been reported for tropical arthropods (Stork & Blackburn 1993) and temperate Hymenoptera (Ulrich 1999, 2005).

In summary, the present model predicts SAD shapes and associated ecological patterns like core – satellite distributions, extinction probabilities, and body weight dependent distributions to be mainly dependent on the degree of temporal variability. It is tempting to relate this
Body size and the relative abundance of species

prediction to the traditional differentiation between K- and r-selected species (Pianka 1970). r-selected species are particular common in arthropod communities where many species have pronounced abundance fluctuations. K-species, on the other hand, are often vertebrates with comparably moderate abundance fluctuations. Hence, the present model predicts vertebrates to have predominantly sigmoid SAD shapes, low local extinction probabilities, and body weight dependent distributions in accordance to metabolic theory. Arthropod communities should often have SAD shapes without lower curvature (right skewed in Preston plots), high extinction probabilities, and widely varying body weight dependent distributions that only moderately follow the predictions of metabolic theory.

Acknowledgments

I thank Marcin Ollik and Simone Fattorini for critical and valuable suggestions on the manuscript. Miss Hazel Pearson kindly improved my English. This work was in part supported by a grant from the State Committee for Scientific Research No. 3 F04F 034 22.

References

Chislenko L. L., 1981, Structure of fauna and flora as dependent on organismal body size, Moscow Univ. Press, in Russian.

Foley P., 1994, Predicting extinction times from environmental stochasticity and carrying capacity, Conservation Biology 8: 124 – 137.


Novotny V. & Basset Y., 2000, Rare species in communities of tropical insect herbivores: pondering the mystery of singletons, Oikos 89: 564 – 572.

Novotny V. & Kindlmann P., 1996, Distribution of body sizes in arthropod taxa and communities, Oikos 75: 75 – 82.


Smith F. A. et al., 2004, Similarity of Mammalian body size across the taxonomic hierarchy and across space and time, American Naturalist 163: 672-691.


White C. R. & Seymour R. S., 2003, Mammalian basal metabolic rate is proportional to body mass (2/3), Proceedings of the National Academy of Sciences USA 100: 4046-4049.