

Goodness-of-fit measures of evenness: a new tool for exploring changes in community structure

A. C. STUDENY,^{1,†} S. T. BUCKLAND,¹ J. B. ILLIAN,¹ A. JOHNSTON,² AND A. E. MAGURRAN³

¹Centre for Research into Ecological and Environmental Modelling, University of St Andrews, The Observatory, Buchanan Gardens, St Andrews, Fife KY16 9LZ United Kingdom

²British Trust for Ornithology, The Nunnery, Thetford IP24 2PU United Kingdom

³Scottish Oceans Institute, School of Biology, University of St Andrews, St Andrews, Fife KY16 8LB United Kingdom

Abstract. Growing concern about the fate of biodiversity, highlighted by the Convention on Biological Diversity's 2010 and 2020 targets for stemming biodiversity loss, has intensified interest in methods of assessing change in ecological communities through time. Biodiversity is a multivariate concept, which cannot be well-represented by a single measure. However, diversity profiles summarize the multivariate nature of multi-species datasets, and allow a more nuanced interpretation of biodiversity trends than unitary metrics. Here we introduce a new approach to diversity profiling. Our method is based on the knowledge that an ecological community is never completely even and uses this departure from perfect evenness as a novel and insightful way of measuring diversity. We plot our measure of departure as a function of a free parameter, to generate "evenness profiles". These profiles allow us to separate changes due to dominant species from those due to rare species, and relate these patterns to shifts in overall diversity. This separation of the influence of dominance and rarity on overall diversity enables the user to uncover changes in diversity that would be masked in other methods. We discuss profiling techniques based on this parametric family, and explore its connections with existing diversity indices. Next, we evaluate our approach in terms of predicted community structure (following Tokeshi's niche models) and present an example assessing temporal trends in diversity of British farmland birds. We conclude that this method is an informative and tractable parametric approach for quantifying evenness. It provides novel insights into community structure, revealing the contributions of both rare and common species to biodiversity trends.

Key words: biodiversity; divergence measures; diversity indices; diversity ordering; diversity profiles; evenness; goodness-of-fit statistics; Hill's diversity numbers; rare species; species abundance distribution.

Received 7 September 2010; revised 15 November 2010; accepted 18 November 2010; **published** 9 February 2011. Corresponding Editor: N. Gotelli.

Citation: Studeny, A. C., S. T. Buckland, J. B. Illian, A. Johnston, and A. E. Magurran. 2011. Goodness-of-fit measures of evenness: a new tool for exploring changes in community structure. *Ecosphere* 2(2):art15. doi:10.1890/ES10-00074.1

Copyright: © 2011 Studeny et al. This is an open-access article distributed under the terms of the Creative Commons Attribution License, which permits restricted use, distribution, and reproduction in any medium, provided the original author and sources are credited.

† **E-mail:** angelika@mcs.st-and.ac.uk

INTRODUCTION

Species abundance distributions describe the abundance of species in an ecological community and are used to assess changes along spatial and temporal gradients or as a result of anthropogenic impacts (McGill et al. 2007, Dornelas et al. 2010). A large number of measures have been

proposed as indices of the diversity represented by a species abundance distribution (Magurran 2004), the most prominent being Shannon's entropy and Simpson's index (Shannon 1948, Simpson 1949, Buckland et al. 2005). These measures typically take account of both richness, that is the number of species present, and evenness, that is how equally abundant these

species are. In addition there are classes of measures that focus on either the richness or evenness component of diversity. Natural communities are never perfectly even (McGill et al. 2007). However the degree of unevenness in abundance provides important insights into the mechanisms that structure a community (Tokeshi 1996) and the extent to which it is disturbed (Drobnier et al. 1998).

One drawback of diversity statistics such as the Shannon and Simpson indices is that they offer only a single insight into community structure, and thus may be relatively uninformative and not easily compared. An alternative approach is diversity ordering (e.g., Hill 1973, Patil and Taillie 1982), which embeds indices in parametric families and thus provides a much more complete picture of the diversity of an assemblage. For example Hill's diversity family links species richness, the Shannon and Simpson indices, along with the Berger-Parker dominance measure, via a free parameter and allows investigators to construct a diversity profile of a community that can be used to make comparisons across space and/or time.

Although several families of diversity indices have been proposed (for an overview see Tothmeresz 1995), families of evenness measures have rarely been considered and none has found wide application in ecological studies (Ricotta 2003). Among existing parametric diversity indices, Hill's diversity numbers are widely used (Hill 1973, Hill 1997). This family tends to perform better than other measures (Tothmeresz 1995) and has lately attracted renewed interest (Jost 2006, Chao et al. 2010, Jost 2010). Quotients of Hill's diversity numbers can be used to quantify evenness, but have been criticized because they do not carry direct ecological meaning (Peet 1974).

In this paper, we examine a family of divergence measures introduced by Read and Cressie (1988) in the context of a comprehensive study of goodness-of-fit statistics. It allows us to construct diversity profiles which capture both the dominance and the rarity characteristics of a community. Existing diversity profiling techniques shift focus from the number of species (as seen in a sample) towards the dominance of one species, thus combining different aspects of the species abundance distribution. For a perfectly even

community with no dominant species, such a profile is flat. Real communities are never perfectly even of course, but unevenness can have different causes, one of which is strong dominance by one or two species. Our method quantifies the relative contributions of dominant and rare species to unevenness. Thus it can distinguish between communities with a few highly dominant species and those where common species are fairly balanced in abundance, with the main contribution to unevenness coming from the rare species. We evaluate the performance of our measure of evenness, and its ability to discriminate between species abundance distributions.

All information on species abundance distributions is based on finite samples from communities (Smith and Wilson 1996, Hill 1997, Jost 2010). The family of evenness measures considered here has a direct sample-based analogue in Read and Cressie's family of goodness-of-fit statistics. In statistical terms, all members of this family are goodness-of-fit statistics commonly used to infer information on a population from a sample. They provide a natural connection between the divergence of a species abundance distribution from evenness and finite samples from this distribution. In other words, instead of looking at the species abundance distribution itself, they measure the plausibility that a finite sample was taken from a community with a completely even distribution.

Similar to diversity profiling, this family is based on a free parameter λ and we can construct "evenness profiles" by plotting the divergence measure as a function of this parameter. In contrast with traditional profiles, the parameter range is not restricted to positive values.

To evaluate this divergence/goodness-of-fit approach to evenness, we consider abundance distributions simulated from Tokeshi's (1990, 1996) niche models. These generate communities with varying degrees of evenness, and so provide an ideal context for testing the properties of the evenness profiles based on the family of power divergence measures, including sampling effects.

We first explain the general idea of a goodness-of-fit approach to diversity assessment and show that common goodness-of-fit statistics are transformations of the classical diversity indices, Shannon's and Simpson's (see *General idea*). A

generalization in terms of a family of measures is then given which leads to a parametric evenness index (see *A parametric evenness index...: The family of power divergence statistics*). We plot this index as a function of the parameter. The resulting evenness profiles and their properties are discussed and compared to alternative diversity profiling techniques (see *A parametric evenness index...: Evenness profiles...*). We explore the links with existing single parameter families of diversity indices (see *A parametric evenness index...: Connection to other parametric...*). Finally we illustrate and evaluate the evenness profiles, first using simulated species abundance distributions from the Tokeshi models (see *Applications: Evenness profiles...*), and then by quantifying trends in British farmland birds (see *Applications: Assessing trends in evenness...*). We conclude that evenness profiles based on the family of divergence measures provide more information than existing parametric index families, in particular with respect to rare species.

GENERAL IDEA

Terminology and notation

We assume that the total number of species $S > 1$ present in a region of interest is fixed over time. (We return to this assumption later.) The individuals belonging to these species form our community. The species abundance distribution of this community is given by the vector

$$\mathbf{N} = (N_1, \dots, N_S), \quad (1)$$

where N_i is the number of individuals of species i in the region of interest, $i = 1, \dots, S$. Typically, diversity indices are defined based on the species proportions given by

$$\mathbf{p} = (p_1, \dots, p_S) = \left(\frac{N_1}{N}, \dots, \frac{N_S}{N} \right) \quad (2)$$

where $N = \sum_{i=1}^S N_i$ is the total abundance.

For a perfectly even community, $\mathbf{p}^* = (p_1^*, \dots, p_S^*) = (1/S, \dots, 1/S)$.

Usually quantities in Eqs. 1 and 2 are not known and summary statistics of the species abundance distribution must be estimated using information from samples taken according to a specified survey design. A sample will be denoted by $\mathbf{n} = (n_1, \dots, n_S)$ where $n = \sum_{i=1}^S n_i$ is the sample size. Note that zeros are allowed: n_i is zero if species i is not recorded in the sample. A

zero count might arise in various ways: the species may have disappeared from the species pool; it may be temporarily absent from the region at the time of the survey (e.g., migrants); it may be present but absent from the sampled plots (more likely for rare species); it may be present on at least one sample plot, but evade detection (more likely for rare and elusive species).

The species' sample proportions are given by $\mathbf{q} = (q_1, \dots, q_S) = (n_1/n, \dots, n_S/n)$. For the following theory to hold, samples are assumed to be taken according to a multinomial distribution.

Perfect evenness as the yardstick

We quantify the evenness of the species abundance distribution \mathbf{p} by comparing it to the perfectly even abundance distribution given by \mathbf{p}^* (serving as a "null model", Gotelli and Graves 1996).

To do so, we use the connection between standard goodness-of-fit statistics and a specific group of divergence measures. Keeping in mind that the individual abundances in Eq. 1 and species proportions in Eq. 2 are usually unknown, we assess the departure from the null model of perfect evenness based on the observed abundances \mathbf{q} . Goodness-of-fit statistics offer a natural method to do so. Taking the two most commonly used goodness-of-fit statistics, the likelihood ratio G and Pearson's X^2 , we have

$$G = 2 \sum_{i=1}^S n_i \log \left(\frac{n_i S}{n} \right), \quad (3)$$

$$X^2 = \sum_{i=1}^S \frac{(n_i - n/S)^2}{n/S}. \quad (4)$$

Instead of comparing these statistics to a χ^2 -distribution, we use their values as measures of the degree of departure of the true species abundance distribution from perfect evenness. More precisely, if we assess the evenness contained in \mathbf{p} by the divergence measures $\sum_{i=1}^S p_i \log(p/p_i^*)$ (the well-known Kullback-Leibler divergence) and $\sum_{i=1}^S p_i [p_i/p_i^* - 1]$, then $G/2n$ and $X^2/2n$ provide sample estimates of these divergences. In fact, they are asymptotically unbiased maximum likelihood estimators. They are also simple transformations of the most prominent classical diversity measures, Shannon's and

Simpson's indices (see Appendix A).

A PARAMETRIC EVENNESS INDEX BASED ON GOODNESS-OF-FIT STATISTICS

The family of power divergence statistics

Read and Cressie introduced a family of goodness-of-fit statistics which includes both G and X^2 (Cressie and Read 1984, Read and Cressie 1988). This family provides a generalization of the idea to use goodness-of-fit statistics as estimators of evenness. Introducing a free parameter λ , Read and Cressie derived the following parametric form for a generalized goodness-of-fit statistic. By changing values for λ , we can switch between different goodness-of-fit statistics (where continuous limits are taken for the initially undefined parameter values $\lambda = -1$ and $\lambda = 0$)

$$I_n(\lambda) = \frac{2}{\lambda(\lambda + 1)} \sum_{i=1}^S n_i \left[\left(\frac{n_i}{n/S} \right)^\lambda - 1 \right]. \quad (5)$$

The statistics G and X^2 are special cases, corresponding to $\lambda = 0$ and $\lambda = 1$, respectively (for further details see Appendix B). Analogously to G and Pearson's X^2 , when divided by $2n$ this family of goodness-of-fit statistics provides sample estimates of a measure of divergence between the true species abundance distribution \mathbf{p} and the perfectly even distribution \mathbf{p}^* . This measure quantifies the unevenness contained in \mathbf{p} , using a parametric form

$$I_p(\lambda) = \frac{1}{\lambda(\lambda + 1)} \sum_{i=1}^S p_i \left[\left(\frac{p_i}{p_i^*} \right)^\lambda - 1 \right]. \quad (6)$$

By varying the parameter λ , we change the weights of the contributions from different species to the index. By recording the changes in index values, we can infer information about evenness. If the data exhibit perfect evenness, the estimator is independent of λ . The greater the departure from evenness, the more pronounced are the changes in index values with λ . This is an important feature when we introduce the concept of evenness profiles below.

What makes a good index of evenness?

To evaluate the usefulness of this approach for applications, we need to look for key criteria that should be fulfilled. For unitary evenness metrics, Smith and Wilson (1996) drew up an extensive

catalogue of essential "requirements" and desirable "features". In general, we need to distinguish between properties of the evenness metric and its estimators.

Statistical properties.—As Buckland et al. (2005) pointed out in their criteria for diversity indices, the statistical properties of the estimators (here the family of goodness-of-fit statistics) are crucial if we want to draw valid inference on evenness. The theory of the family of goodness-of-fit statistics is well studied (Read and Cressie 1988), in particular its expectation and variance are known (for arbitrary λ). Analogous to the special cases of Pearson's X^2 and the likelihood ratio G , the estimator we get through the family of goodness-of-fit statistics is a maximum likelihood estimator. Although its expectation is not completely independent of sample size (due to the multiplicative factor $1/2n$), both bias and precision steadily improve (with a rate of $1/n$) as sample size increases (details in appendix C).

The range of index values.—Smith and Wilson (1996) advocate the use of indices that are constrained to lie between 0 (all p_i 's equal to 0 except one) and 1 (perfect evenness). The indices in $I_p(\lambda)$ violate this feature in two ways. First, values are not constrained to lie between 0 and 1, and second, large values of the indices, not small values, indicate large departures from perfect evenness, rather than a high degree of evenness. A transformation of $I_p(\lambda)$ that standardizes the index family (given in Appendix D) may be used should an application require a measure of diversity with a more traditional range of support.

Relationship between evenness and species richness.—Traditionally, evenness metrics were sought that are largely independent of species richness, since evenness and richness were viewed as two separate characteristics of the species abundance distribution. More recently, it has been recognized that their relationship is more complex and that they cannot be regarded as independent components of diversity (Gosselin 2006, Jost 2010): adding or reducing the number of species, without making any adjustments to the abundances of the remaining ones, will in general change the degree of evenness. For example, if species i disappears from the species pool, i.e., $n_i = 0$, traditional scalar and existing parametric diversity and evenness met-

rics will register the reduction both in number of species and in evenness at the same time. This makes it difficult to distinguish between both effects. Our approach gives the option of either retaining the same species richness S or resetting it to $S - 1$.

In the first case, evenness is then evaluated with respect to S species (species i making a negative contribution) or in the latter case, species i is not taken into account. This allows us to choose which we consider to be the most appropriate, at the cost of having to specify the assumed species pool rather than let it be chosen by the measure. A potential problem in this context lies in the fact that it is generally difficult to estimate the number of species S without bias or high uncertainty (Lande 1996, Gotelli and Colwell 2001, 2010). This issue is further discussed below when we look at related index families, and the problem of unobserved species in samples.

Evenness profiles

As functions of the free parameter λ , families such as $I_p(\lambda)$ allow us to plot index values over the range of that parameter. Thus they provide a graphical display of information contained in the species abundance distribution in addition to a quantification of diversity. We will show in this section that this is highly useful and more intuitive for conveying multidimensional information than deriving either a single or a range of scalar metrics alone. The graphical equivalent of a parametric index family has been termed a “profile” (Patil and Taillie 1982). Profile plots of existing index families are usually restricted to the positive parameter range. For most, the parameter value of zero corresponds to the number of species (or a transformation thereof). In this case, by simply counting the number of species, we assign equal weight to all species. Plots are usually a monotonic decreasing function of the parameter, and the metrics tend to be increasingly driven by the dominance of the most abundant species.

Negative parameter values would put increasingly greater weight on the rarest species, but would at the same time reverse the order when comparing distributions with the same number of species but different levels of evenness. This violates what Ricotta (2003) refers to as consis-

tency with the intrinsic order (as given by the Lorenz curve). Because this consistency is a desirable property when profile plots are used for comparison in ecological applications, negative parameter values have so far been excluded in any of the existing parametric index families. Nevertheless, they do provide valid index values and contain information on community structure. We will see that profile plots that are based on our parametric approach, and which we call “evenness profiles” in the following, do not suffer from this change in order and we can hence explore the whole parameter range.

Similar to diversity ordering, we derive these evenness profiles by plotting the family of divergence measures $I_p(\lambda)$ for a given species abundance distribution \mathbf{p} as a function of λ . This profile can be used to analyze community structure and compare species abundance distributions. The resulting graph is always continuous and has a parabolic shape. In this it differs from existing approaches, which would be sigmoid were negative parameter values included; it is this sigmoid shape which causes the ordering to reverse.

Typically, we are interested in diversity metrics as relative rather than absolute measures, to compare species abundance distributions through either space or time. The curvature of our evenness profiles provides further information to aid such comparisons. The closer a species abundance distribution is to perfect evenness, the less the index values depend on λ until in the limiting case of perfect evenness, all members of the family are equal (corresponding to a horizontal line). Hence, the degree of curvature of the profile plot reflects the degree of unevenness in the species abundance distribution (see Fig. 1).

In traditional diversity profiling, the focus shifts between the two components evenness and species richness as the parameter varies. By contrast, for evenness profiles, the focus switches between the two opposite ends (rarity/dominance) of one component (evenness). Although not completely independent of the number of species S (which must be fixed in advance), an evenness profile is equally dependent on S everywhere, allowing comparison of profile plots.

In terms of a sample, the members of the family $I_n(\lambda)$ vary in their sensitivity to more

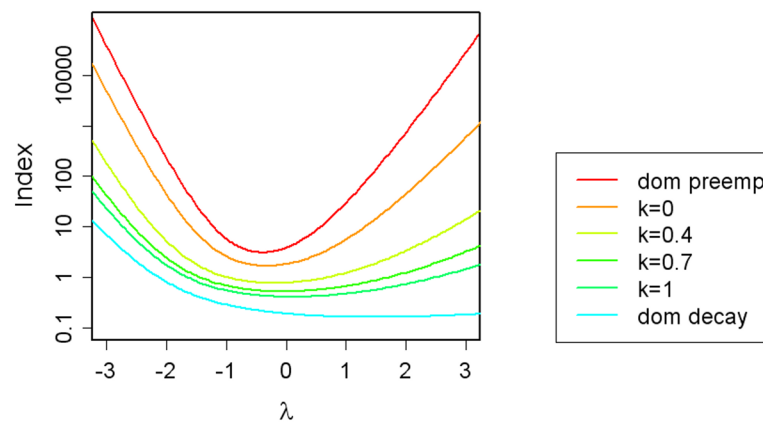


Fig. 1. Evenness profiles. Plots are showing evenness profiles for simulated species abundance distributions following Tokeshi's niche models (where species richness S is constant). High index values as well as strong curvature correspond to low evenness. The profiles differentiate clearly between the different models and order them correctly with respect to their evenness.

extreme ratios of observed to expected frequencies per class (species) for different values of λ . Large positive λ put more weight on large ratios of observed frequencies to those expected under perfect evenness, thus detecting unevenness due to pronounced dominance of a species. At the other end, large negative λ highlight observations that lie below their expected levels under perfect evenness, i.e., on rare species. Thus, these profiles are able to distinguish between a community with high dominance of a species and one where the unevenness lies mainly within the rare species, while the common species have similar abundance (see example below). This is not detected by existing approaches. For a graphical assessment of community structure, the range of $-3 \leq \lambda \leq 3$ usually provides sufficient information.

Connection to other parametric diversity indices

There are numerous measures for diversity, some of which gave rise to evenness indices (Smith and Wilson 1996). If we want to compare our approach to existing (parametric) methods, we are faced with a variety of number sets to choose from. For parametric diversity indices it has been shown that Hill's numbers not only perform best in terms of distinguishing species abundance distributions, but also that most other parametric approaches can be derived from them (Tothmeresz 1995). More recently, a new trans-

formation of them, termed "inequality factors" has been introduced (Jost 2010). This set of numbers claims to respect the interdependence and correctly separate the effect of richness from evenness. We show that our approach is closely connected to both these families of diversity measures, before we then compare their performance when they are applied to distinguish between abundance models with different degrees of evenness.

Hill's diversity numbers.—Hill's (1973) one parameter family of "diversity numbers" is given by $N(\alpha) = (\sum p_i^\alpha)^{1/(1-\alpha)}$ where $\alpha \geq 0$. For specific values of α it corresponds to several classical diversity measures, namely species richness $N(0) = S$, the exponential of Shannon's entropy $N(1) = e^H$, and the inverse Simpson index $N(2) = 1/D$.

We can rewrite our index family $I_p(\lambda)$ in terms of Hill's diversity numbers as

$$I_p(\lambda) = \frac{1}{\lambda(\lambda + 1)} \left[\left(\frac{N(0)}{N(\lambda + 1)} \right)^\lambda - 1 \right]. \quad (7)$$

Hence the parametric family considered here is a transformation of quotients of Hill's numbers. Hill himself suggested the use of quotients $N(a)/N(b)$ of two diversity numbers to quantify evenness, giving a value of 1 when the underlying distribution is completely even (Hill 1973). In theory, any two real numbers a, b can be chosen. Commonly quotients involving $N(0) = S$ are used to render the measure independent of species

richness (Pielou 1969, Heip 1974, Camargo 1993, Smith and Wilson 1996). Nonetheless, Hill discourages the reader from using $N(0)$ because of the difficulty of determining S . As our measure involves quotients based on $N(0) = S$, we need to justify why it is sensible to base an evenness measure on the usually unknown number of species.

Inequality factors.—Traditional evenness measures have typically used division by S with the aim of achieving independence from species richness. Recently, Jost (2010) argued that richness and evenness cannot be independent and instead proposes factorizing richness into a diversity and an evenness component where diversity is quantified via Hill's diversity number $N(a)$. These evenness components are called "inequality factors" and given by the quotients $N(0)/N(a)$, which (from the equation above) moves them close to the index family considered here. Jost discusses different transformations of these evenness components to derive relative evenness as well as logarithmic relative evenness which measure evenness relative to the minimum and maximum possible for a given number of species.

Should an evenness measure be based on the number of species?—Any approach that is—like ours—based on $N(0) = S$ should either give reasons why this is sensible given the difficulty of determining the number of species, or look for ways of avoiding this problem. Hill's solution was to regard quotients with $N(2)$ in the nominator instead of $N(0)$. This could be done for the method suggested here, too, but would result in breaking the connection to goodness-of-fit statistics which provides a natural interpretation of the measure and a direct sample based estimate. On the other hand, quotients $N(2)/N(a)$ are difficult to interpret in most cases and show ambiguous behavior (Peet 1974, Jost 2010). They no longer monotonically increase with evenness and hence can show similar values for highly even and for uneven species abundance distributions (see example below). In applied studies, the problem is usually avoided by assessing evenness of a sample (i.e., only taking into account observed species) rather than drawing inference on the species abundance distribution. In this case, the measures do not reflect a characteristic of the population but of the sample

itself (Hill 1997). Parametric indices on the other hand have been defined and studied theoretically for given (and fully known) species abundance distributions. Both approaches are on their own unsatisfactory, the first because we are interested in characteristics of the population, not the sample, and the latter because we have to rely on samples to draw inferences on these characteristics.

Instead, we here assume that the number of species S in a study area is fixed (and known) at least over the period of time we are interested in. We think it reasonable to base an evenness measure on S for two reasons. First, theoretical studies (e.g., Bulla 1994), are often implicitly built on the assumption of a fixed number of species. These studies provide valuable insight in the general behavior and properties of diversity indices. By stating this assumption on S explicitly, we set a defined framework in which our analyses and conclusions hold. Second, in the context of monitoring temporal change in biodiversity within a site or region, it is self-consistent to define a list of potential species, and to estimate species proportions as zero for those species not recorded at a given time point. Disappearance of species will then be reflected by a decline in evenness. If unanticipated species are recorded in later surveys, the list of potential species may be revised, and the entire sequence of biodiversity measures updated.

The comparison of two completely independent communities on the other hand is more complex. Because of the multidimensionality of diversity as a concept, the question of which of two or more such communities is the more diverse will rarely have a single answer. Yet, if we can assume that either species richness is comparable or that more or less the same catalogue of species holds across communities (this can always be achieved just by artificially pooling species), our family of evenness measures offers an instrument of comparison. As it is sensitive to unobserved or missing species, it will pick up differences between these communities.

UNOBSERVED SPECIES AND SAMPLING EFFECTS

We rely on observations at sample locations within the study area to estimate evenness profiles based on $I_n(\lambda)$, which then allow us to

Table 1. Overview of Tokeshi's niche models.

Model	Specifications	Degree of evenness
Dominance decay	Newly colonizing species take a random part of the largest current niche	high
Power-fraction	Niche is chosen randomly, depending on its size (degree of dependence regulated via a model parameter k), random break point; $k = 0$ (random fraction) completely random niche choice, $k = 1$ (McArthur fraction) largest niche is chosen	variable from $k = 0$ (low evenness) to $k = 1$ (high evenness)
Dominance preemption	Arriving species take more than 50% of the remaining niche space	low

draw inference on the community. However, as pointed out above, species may be missing from samples (Colwell and Coddington 1994, Yoccoz et al. 2001, Chao and Shen 2003, MacKenzie et al. 2005). Some species may not be observed despite being present, either because of their low probability of detection or their scarceness in the survey area. Others might be part of the assumed species catalogue in general but absent at the time of the survey (e.g., migrants).

In contrast to other diversity indices, the approach to evenness considered here is sensitive to unobserved species in a sample. Zeros in the sample vector \mathbf{n} are treated as observations which lie below what would be expected and therefore increase departure from evenness. Often, there will be a valid reason to consider missing species as part of the community. Our method then allows us to include these and thus goes beyond a simple assessment of characteristics of a sample. In particular, this enables us to register cycles in diversity patterns due to migration or the change in evenness, if a species genuinely disappears from the population. However, given the sensitivity of the goodness-of-fit approach to zeros, a careful and educated decision about the species catalogue is needed in order to avoid biased results.

Even if we know the exact number of species, zeros in the vector of sample proportions \mathbf{q} cause problems. Evenness profiles which are estimated through $I_{\mathbf{n}}(\lambda)$ are no longer continuous at $\lambda = -1$ if $q_i = 0$ for some i . There are three ways to overcome this problem and regain the evenness profiles, at least partially. The usual approach is to discard zeros and calculate diversity indices or construct diversity profiles solely based on the species observed in the sample. As discussed above, this accounts for the properties of the sample but not the community. If we want to

retain the full profiles, a common solution is to add small quantities ε to achieve non-zero values for all q_i . Alternatively we can restrict the analysis to positive parameter values, taking into account the zeros but without the explicit focus on rare species of the negative parameter values. The latter will be a good solution if sampling intensity is low, since it discounts the part of the sampled distribution that carries higher uncertainty. All three approaches will be looked at when we apply evenness profiling techniques to simulations from Tokeshi's models in the next section.

APPLICATIONS

Evenness profiles for Tokeshi's niche models

Tokeshi (1990, 1996) introduced a group of stochastic models (dominance pre-emption, power fraction, dominance decay) to predict species abundance distributions based on niche apportionment. These models are used by ecologists to generate simulated datasets with varying degrees of evenness (see Table 1). This gives us means to test the behavior of the evenness profiles.

The following simulations were implemented in the statistical language R (R Development Core Team 2010; code details can be found in the supplementary material). For each model, 500 random realizations were generated by a discrete algorithm where available niche space was fixed by setting the total abundance $N = 50000$. The expected species abundance distributions were then derived by averaging over these. The power fraction model was simulated for model parameter $k \in \{0, 0.4, 0.7, 1\}$. We consider the following two scenarios.

In scenario 1, we construct evenness profiles based on the true species abundance distributions for all the Tokeshi models for a fixed

number of species $S = 100$. These are compared with both Hill's scaled diversity numbers and Jost's evenness measures to assess how well each approach is able to distinguish between the different distributions.

For scenario 2, we look at profiles for one sample from each of the Tokeshi models. This allows us to study the effects of sampling and the problem of unobserved species. Different ways of dealing with unobserved species in a sample are applied. Assuming various values for underlying species richness, we investigate the effects of mis-specification on inference from the profile plots.

Ordering of Tokeshi's models (scenario 1).—Evenness profiles put Tokeshi's models in the correct order (see Fig. 1). Curvature decreases corresponding to higher degrees of evenness. For the dominance decay model, only the negative parameter range enables us to detect departure from evenness as evenness is reduced by the rare species while common species are highly homogeneous in numbers of individuals. Overall, the similar curvature in the negative parameter range reveals that it is the degree of dominance of the abundant species that distinguishes between Tokeshi's models, rather than differences in occurrence of rare species.

For comparison, profiles based on Hill's diversity numbers are shown in Fig. 2, with the usual restriction to positive parameters. While high index values are associated with departure from evenness in our evenness profiles, they refer to high degrees of evenness in these plots. As only positive parameter values are usually considered, a quotient close to 1 can either be because the underlying species abundance distribution is truly close to evenness or simply due to homogeneity in abundance within the more dominant species while the tail of rare species is neglected.

Plots based on $N(a)/N(2)$ show the aforementioned lack of consistency: for higher parameter values, the highly even dominance decay model is placed closest to the most uneven dominance pre-emption model.

The last of the profile plots in Fig. 2 shows a logarithmic transformation termed "logarithmic relative evenness" by Jost (2010). While profiles using untransformed quotients of Hill's numbers have difficulties in differentiating between the different Tokeshi models, logarithmic evenness is

as effective as the evenness profiles with respect to the dominance structure. However, because it is restricted to the positive parameter range, it does not reveal that the models are similar with respect to the tail of rare species. In particular, high evenness of the dominance decay model is picked up, but the remaining unevenness due to rare species is undetected.

Sampling effects and unobserved species (scenario 2).—To examine the performance of the evenness profiles when estimated from a sample, we look at one sample from each of Tokeshi's power fraction models with $k = 0, 0.4$, and 1, to cover a range of uneven, intermediate and even distributions. To estimate the true evenness profile based on this sample, we need to assume a value for the number of species S . We are interested in to what extent deliberate underestimation (by taking just the number of species observed in the sample) versus overestimation of S affects our inference on the true community structure. For this, we plot sample profiles for a range of assumed values of S along with the true profile (Fig. 3). Where a higher number of species was assumed than was observed, a small ε is added to species counts in order to achieve a full profile plot including negative values for λ .

For all profile plots in Fig. 3, we see that in general we do better by taking unobserved species into account. This is relevant if we are interested in the contribution of or changes in rare species, visible through the negative parameter range of the evenness profiles. Provided we do not grossly underestimate the number of missing species, the picture we get for rare species is quite robust to mis-specification of our sample. There is a slight drawback: while the curves are largely similar, the bias from the true profile depends on the value of ε (for more details see Appendix E). For the positive parameter range, the sample evenness profiles are not much affected by the choice of ε as the dominance structure is unaffected.

If we assess evenness by looking at curvature of the evenness profiles in Fig. 3 rather than absolute values, the dependence on ε is in general less of an issue. By contrast, if we omit unobserved species, profiles are almost flat, giving a false indication of evenness of the community. This holds no matter which evenness measure we choose. Even if we are only

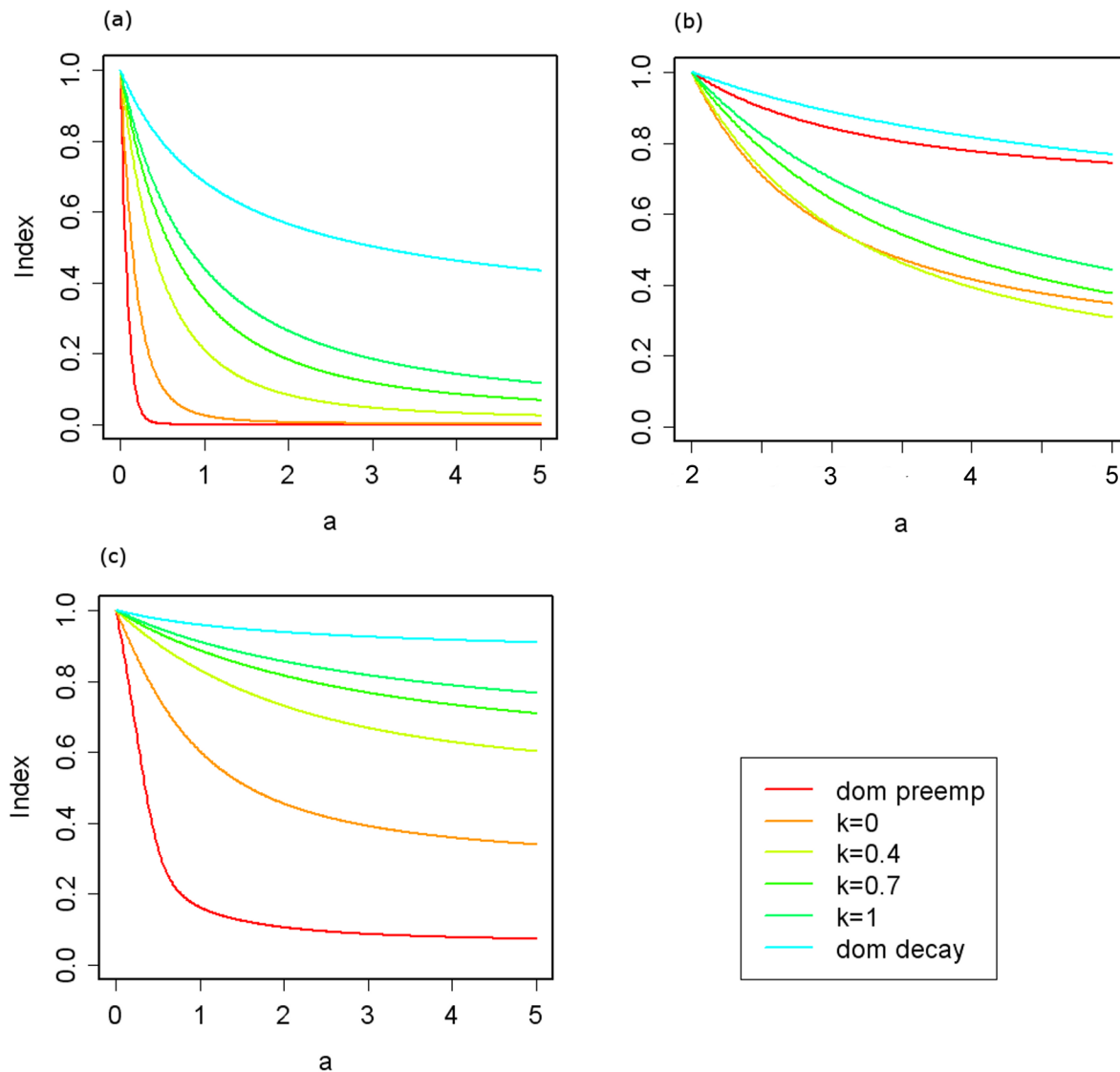


Fig. 2. Profiles for alternative evenness measures. Plots show profiles for Tokeshi's models for transformations of Hill's diversity numbers, namely (a) $(N(a)/N(0))^2$, (b) $(N(a)/N(2))^2$ and (c) relative logarithmic evenness. Division by $N(0) = S$ just scales the diversity numbers to $[0, 1]$, whereas scaling by $N(2)$ results in an inconsistent ordering of the abundance distributions. Logarithmic relative evenness separates the profiles more clearly by spanning the entire range between 0 and 1.

interested in the (true) dominance structure of a community, disregarding unobserved species leaves a bias in all the profiles. Bias in the estimated profiles increases slowly as our estimate of S moves further from the true number of species, in either direction. The exception is Hill's evenness quotient based on $N(2)$ which appears invariant to the assumed S . However, bias and

precision of the latter are highly sensitive to the choice of ε (see Appendix E). We also need to keep in mind that Hill's evenness quotients generally fail to display the difference in the models correctly due to the discontinuity noted above. This is also evident in Fig. 3 with little difference in the quotients for the power-fraction models with $k = 0$ and $k = 0.4$.

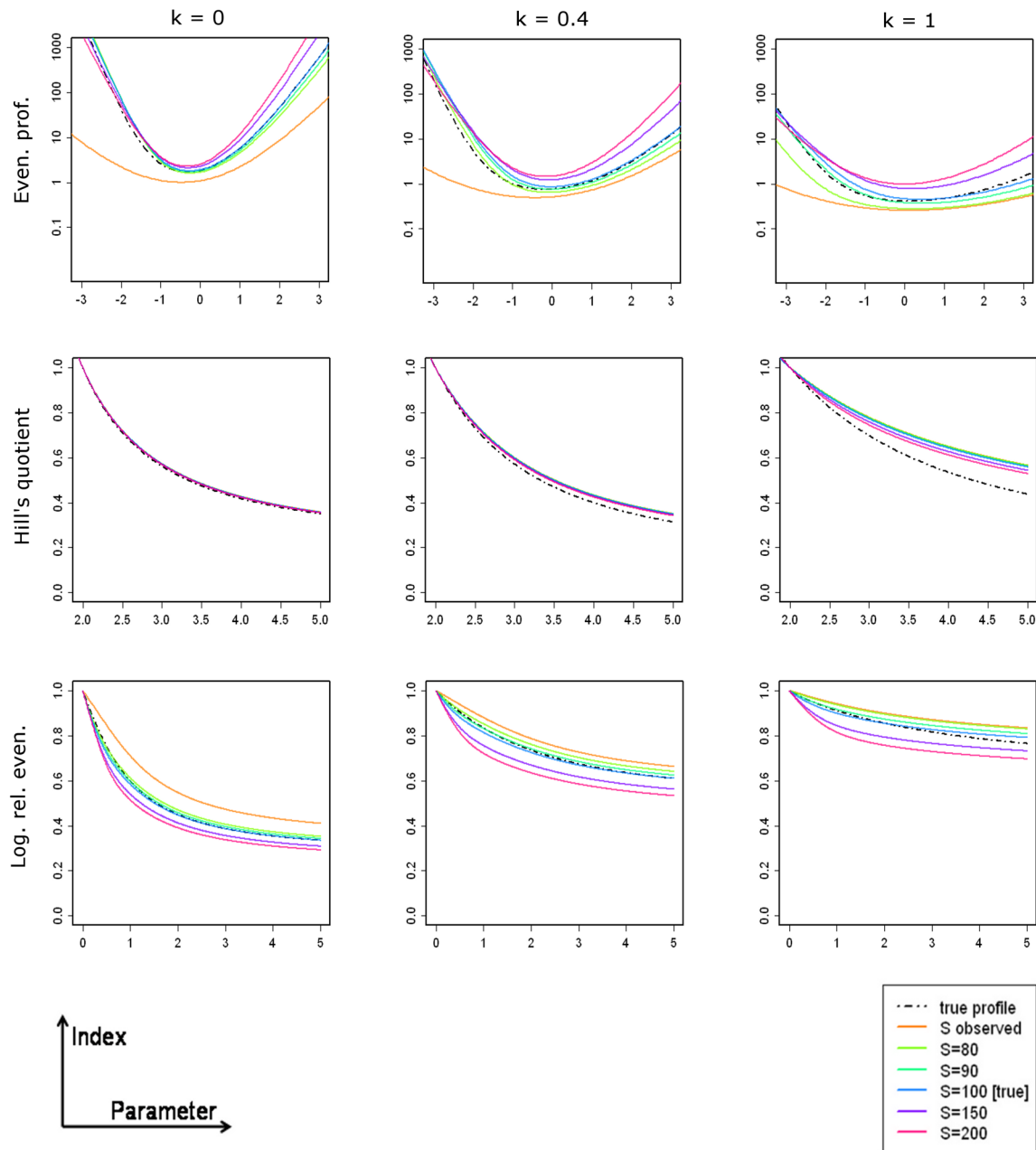


Fig. 3. Sampling effects and unobserved species. Estimated profiles are derived for one sample from each of three Tokeshi's power fraction models (columns of the panel correspond to $k=0, 0.4$, and 1). These are compared with the true underlying abundance distribution ($S=100$, black dashed line). Rows show evenness profiles, Hill's quotients $(N(a)/N(2))^2$ and logarithmic relative evenness (from top to bottom). Either underestimating or overestimating S , a range of assumed values for S is considered. Where this number exceeds the observed number of species ($S_{obs} = 43, 69, 78$ for the sample from each of model), a small quantity is added to species counts, to allow evaluation of the evenness profiles into the negative parameter range.

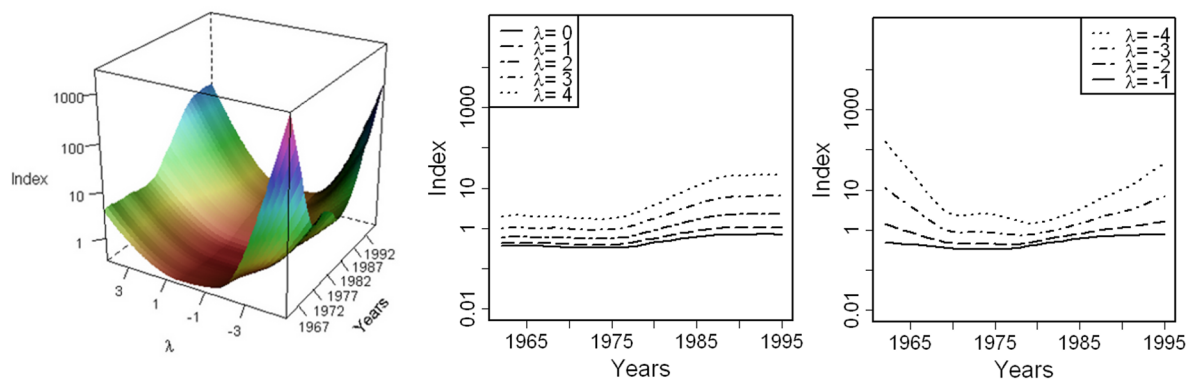


Fig. 4. Changes in diversity over time (CBC data). Evenness profiles are constructed for a group of 13 farmland bird species for every year from 1962 to 1995, leading to a (smoothed) 3D surface plot as shown on the left. Classical indices correspond to the middle parameter range, Shannon's at $\lambda = 0$ and Simpson's at $\lambda = 1$. They give no indication of a pronounced trend, other than a moderate increase in unevenness from the late 1970s through the 1980s. In addition, trends over time are shown separately for a selection of negative parameter values $\lambda = -1, -2, -3, -4$ (left) and non-negative parameter values $\lambda = 0, 1, 2, 3, 4$ (right).

Assessing trends in evenness over time in British farmland birds (Common Bird Census data)

As discussed in the previous section, focusing solely on the more dominant species and neglecting unevenness caused by the rare species may give a false impression of homogeneity. The following example shows that this is particularly true for assessing trends in diversity over time if the proportions of dominant species remain fairly stable while there is a change in some of the rare species. The data come from the UK Common Bird Census (CBC). This long-term survey of breeding birds organized by the British Trust for Ornithology was run over 38 years (1962–2000). Over this time period, changes in agricultural practice had major effects on farmland birds in the UK (Gregory et al. 2004).

The survey protocol follows a territorial mapping approach. Fewster et al. (2000) used generalized additive modeling to smooth the time series of territory counts for a set of 13 farmland species during 1962–1995, which yields predicted counts for every species for each site in each year, whether or not a site was surveyed in a given year.

To look at time trends in evenness of these farmland birds, we plot a profile over the parameter range $[-5, 5]$ for every year of the survey based on these smoothed estimates of individual counts. The results are compared with diversity trends estimated by taking the geomet-

ric mean of the 13 species-specific relative abundance estimates obtained by Buckland et al. (2005).

We see that the intermediate parameter range, which covers the transformations of Shannon's and Simpson's indices, does not display any substantial change in evenness over time, aside from an increase (reduced biodiversity) from the late 1970s to the late 1980s (see Fig. 4). This is in accordance with Buckland et al. (2005) who remarked that these classical measures register a decline in diversity after 1975, but show no trend in the first half of the survey. Our 3D surface plot not only confirms this, but makes it apparent that this is true for all measures in the power divergence family which concentrate on dominance of a species. The extended parameter range however reveals further changes in evenness. For more clarity, profiles for selected parameter values are plotted in addition to the surface plot in Fig. 4. For non-negative $\lambda = 0, 1, 2, 3, 4$, they look fairly similar over time, displaying the same qualitative information as Shannon's or Simpson's index. Profiles based on negative parameter values on the other hand show a decrease in the corresponding indices (i.e., increased evenness) within the group of rare species during the early years, followed by an increase (i.e., decreased biodiversity) from the late 1970s, which continued to the end of the time period. Buckland et al. (2005) also detected the

increasing trend in biodiversity in the early years, using the geometric mean of relative abundances and attributed this to a recovery after the severe winter of 1962–63. While a geometric mean considered the community as a whole, our analysis reveals that this primarily affected the rarer species.

Overall, this leads to the conclusion that biodiversity in British farmland birds increased during the 1960s and early 1970s, followed by a decline from 1975 to the late 1980s. Further decline occurred amongst the rare species, but not amongst the abundant species, until the end of the time period (1995). Traditional indices do not reflect the changes amongst the rarer species.

SYNTHESIS AND FUTURE DIRECTIONS

The Convention on Biological Diversity's 2010 and 2020 targets were a response to large-scale loss of biodiversity (Butchart et al. 2010). They reinforced the need for reliable methods to assess change in ecological communities. Loss in biodiversity often affects rare species first, while the dominance structure of a community might change only gradually (Gotelli et al. 2010). However, most methods are limited in their ability to detect change amongst rare species (Colwell and Coddington 1994). We have proposed a parametric approach to quantify evenness which includes explicit information on rare species. In particular, it allowed us to separate changes in rare species from those in common species for British farmland birds.

A single parameter family of evenness measures

The method suggested is based on a correspondence between divergence measures and a family of goodness-of-fit statistics (Read and Cressie 1988). In general, a high degree of evenness or uniformity of the species abundance distribution is equated with high biodiversity. The idea of this method is that, while ecological communities will never be perfectly even, evenness can serve as a “null model” (Gotelli and Graves 1996) and we can measure the departure of the species abundance distribution from perfect evenness to gain insight into the structure of a community. Goodness-of-fit statistics provide us with genuine insight into the properties of the community while quantifying this depar-

ture on a sample level. By combining these statistics in a single family, we capture more information on the species abundance distribution (as a multivariate object) than can a single scalar index. We have shown that these measures are closely connected to Hill's numbers, which have been the subject of recent interest as diversity metrics (Jost 2006, Chao et al. 2010, Jost 2010).

Performance of evenness profiles

Following the idea of diversity profiling of plotting parametric index families as functions of their parameter, we constructed “evenness profiles”. The parameter range of traditional diversity profiles is restricted to the positive values, which focuses attention on the most abundant species. In contrast, evenness profiles can be sensibly evaluated for both positive and negative parameter values. These profile plots do not only give index values, but display information on evenness via their curvature. A horizontal line would correspond to perfect evenness. This feature is especially valuable for comparison.

Simulated species abundance distributions from Tokeshi's (1990, 1996) models allowed us to examine the performance of evenness profiles based on our approach as well as explicitly addressing sampling effects. The main advantage of this method was illustrated by the example of time trends in British farmland bird diversity, to reveal that changes beyond a general decline during the late 1970s and mid-1980s are mainly due to fluctuations in the abundance of rare species. With the full parameter range at hand, these evenness profiles allow exploration in both directions, focusing on dominant species for positive and on rare species for negative parameter values. Thus they detect that Tokeshi's dominance decay model, despite its highly even allocation of niche space, still contains some unevenness—as every natural community would do. In the example of British farmland birds, classical indices, which are represented by the positive parameter range of the evenness profiles, only exhibit changes in the second half of the survey period following changes in agricultural practices. The effects of recovery from a harsh winter in 1962/63 are however only visible along the negative parameter range of the profile plots.

Including unobserved species

The goodness-of-fit approach is sensitive to zeros in the species proportions. This allows us to include species in the diversity assessment even if they are not detected in our samples. As with any other approach, but perhaps more obviously here, careful consideration of the assumed list of species to include in an analysis is required. In reality the reasons for missing species are manifold, and we might want to distinguish between a true absence, a temporary absence, and rarity or difficulty of detection of a species that is present (Gotelli et al. 2010). Assessing underlying evenness based on a sample from the Tokeshi models showed that omitting unobserved species gives a false impression of higher evenness for both rare and common species, for any of the evenness metrics considered. However, as long as we do not greatly underestimate the true number of species, mis-specification has little effect on inference for rare species. Analysis based on the curvature is even less sensitive to the number of species. Some bias for the common species remains; this is shared by the alternative approaches we considered.

To draw evenness profiles that extend to unobserved species, some small quantity must be added to counts. Some caution is needed when this quantity is chosen, as the negative parameter range is sensitive to this choice. Thus if the information about rare species is not essential for the objectives of a study or if uncertainty about rare species is too high, profiles are better plotted for the positive parameter range only. In this case, dominance is still displayed relative to all the species and not only the sampled ones.

Assessing time trends in diversity

A particular question of interest is the decline or increase in diversity over time (Magurran and Dornelas 2010). Monitoring and conservation programs often focus on rare species (Thomas 2002). Our example of the CBC data for British farmland birds showed that the proposed evenness profiles are able to separate changes that are mostly within the rare species from those that affect common species.

In general, the method we consider in this paper allows us to include species that are not observed at every time point. Valid quantitative

and qualitative conclusions on time trends can be drawn as long as we can establish a representative catalogue of species which we assume form the community over the period of time we are interested in. The loss of biodiversity should a species truly disappear from the community would be registered as long as the species is included in the assumed catalogue of species in the community. If it is found necessary to revise this list as more data become available, we can re-evaluate the index family over the entire time period, to improve our knowledge of the dominance and rarity characteristics of the community. A more detailed statistical analysis would also include precision of the evenness profiles and establish points in time which show a significant change in trend. This can be achieved by considering bootstrap resamples of survey sites as in Buckland et al. (2005).

Ultimately, there is no unique answer to the question of how diversity is best assessed. The method developed here allows us to concentrate on the evenness aspect of the species abundance distribution. The evenness profiles extend the range of visible information and display contents of the abundance distribution differently from existing index families. Thus they offer a new perspective which can lead to further insights. Their explicit focus on rare species may be of special interest when examining the impact of anthropogenic disturbance on diversity (Dornelas 2010).

As with any other method, the results may be biased if evenness profiles are drawn without regard to the sampling scheme (Colwell and Coddington 1994, Gotelli and Colwell 2001, 2010). Being sensitive to the number of species that are assumed to be observable, evenness profiles depend on additional information and a deliberate decision on the species that are included in the analysis. While for some taxa, like birds, a complete list of species and therefore exact knowledge of the population of interest is more easily achieved, there are many taxonomic groups for which we will never be able to compile a (nearly) complete list, as for example tropical insects (Longino et al. 2002). Evenness profiles allow us to draw inference even if we are not certain about the actual number of species. Provided we have a rough estimate of the number of species, we can include rare, unob-

served species by extending the sample vector. This should reduce bias in inference. Qualitatively, information on evenness is contained in the curvature of the profiles. Curvature seems to be less sensitive to the assumed number of species than the evenness profiles themselves. Hence even if we disregard the additional information contained in the plots because of potential bias, curvature should still give us an accurate qualitative result.

ACKNOWLEDGMENTS

We would like to thank the British Trust for Ornithology for providing the CBC data. The CBC has been funded by the U.K. Department of the Environment, and by a contract with the Joint Nature Conservation Committee on behalf of the Countryside Council for Wales, English Nature, Scottish National Heritage, and the Department of the Environment for Northern Ireland. We thank the numerous volunteers who devoted time and effort to collect the data over the many years of the survey. The work presented in this paper was supported by a studentship funded by the University of St Andrews through the National Centre for Statistical Ecology (NCSE). We want to thank two anonymous reviewers. This article gained from their valuable comments.

LITERATURE CITED

- Buckland, S. T., A. E. Magurran, R. E. Green, and R. M. Fewster. 2005. Monitoring change in biodiversity through composite indices. *Philosophical Transactions of the Royal Society B* 360:243–254.
- Bulla, L. 1994. An index of evenness and its associated measure. *OIKOS* 70:167–171.
- Butchart, S. H. M. et al. 2010. Global Biodiversity: Indicators of Recent Declines. *Science* 28:1164–1168.
- Camargo, J. A. 1993. Must dominance increase with the number of subordinate species in competitive interactions? *Journal of Theoretical Biology* 161:537–542.
- Chao, A. and T.-J. Shen. 2003. Nonparametric estimation of Shannon's index of diversity when there are unseen species. *Environmental and Ecological Statistics* 10:429–443.
- Chao, A., C. H. Chiu, and L. Jost. 2010. Phylogenetic diversity measures based on Hill's numbers. *Philosophical Transactions of the Royal Society B* 365:3599–3609.
- Colwell, R. K. and J. A. Coddington. 1994. Estimating terrestrial biodiversity through extrapolation. *Philosophical Transactions of the Royal Society B* 345:101–118.
- Cressie, N. and T. R. C. Read. 1984. Multinomial Goodness-of-Fit Tests. *Journal of the Royal Statistical Society Series B* 46:440–464.
- Dornelas, M. 2010. Disturbance and change in biodiversity. *Philosophical Transactions of the Royal Society B* 365:3719–3727.
- Dornelas, M., M. Soykan, and K. Ugland. 2010. Biodiversity and disturbance. Pages 237–251 in A. E. Magurran and B. J. McGill, editors. *Biological Diversity: Frontiers in Measurement and Assessment*. Oxford University Press, Oxford, UK.
- Drobner, U., J. Bibby, B. Smith, and J. B. Wilson. 1998. The Relation between Community Biomass and Evenness: What Does Community Theory Predict, and Can These Predictions Be Tested? *OIKOS* 82:295–302.
- Fewster, R. M., S. T. Buckland, G. M. Siriwardena, S. R. Baillie, and J. D. Wilson. 2000. Analysis of population trends for farmland birds using generalized additive models. *Ecology* 81:1970–1984.
- Gosselin, F. 2006. An assessment of the dependence of evenness indices on species richness. *Journal of Theoretical Biology* 242:591–597.
- Gotelli, N. J. and R. K. Colwell. 2001. Quantifying biodiversity: Procedures and pitfalls in the measurement and comparison of species richness. *Ecology Letters* 4:379–391.
- Gotelli, N. J. and R. K. Colwell. 2010. Estimating species richness. Pages 39–54 in A. E. Magurran and B. J. McGill, editors. *Biological Diversity: Frontiers in Measurement and Assessment*. Oxford University Press, Oxford, UK.
- Gotelli, N. J., R. M. Dorazio, A. M. Ellison, and G. D. Grossman. 2010. Detecting temporal trends in species assemblages with bootstrapping procedures and hierarchical models. *Philosophical Transactions of the Royal Society B* 365:3621–3631.
- Gotelli, N. J., and G. R. Graves. 1996. *Null Models in Ecology*. Smithsonian Institution Press, Washington, D.C., USA.
- Gregory, R. D., D. G. Noble, and J. Custance. 2004. The state of play of farmland birds: population trends and conservation status of lowland farmland birds in the United Kingdom. *Ibis* 146(Suppl.2):1–13.
- Heip, C. H. R. 1974. A new index measuring evenness. *Journal of the Marine Biological Association UK* 54:555–557.
- Hill, M. O. 1973. Diversity and evenness: A unifying notation and its consequences. *Ecology* 54:427–432.
- Hill, M. O. 1997. An evenness statistic based on abundance-weighted variance of species proportions. *OIKOS* 79:413–416.
- Jost, L. 2006. Entropy and diversity. *OIKOS* 113:363–375.
- Jost, L. 2010. The relation between evenness and diversity. *Diversity* 2:207–232.

- Lande, R. 1996. Statistics and partitioning of species diversity, and similarity among multiple communities. *OIKOS* 76:5–13.
- Longino, J., J. A. Coddington, and R. K. Colwell. 2002. The ant fauna of a tropical rainforest: estimating species richness three different ways. *Ecology* 83:689–702.
- MacKenzie, D. I., J. D. Nichols, N. Sutton, K. Kawanishi, and L. L. Bailey. 2005. Improving inferences in population studies of rare species that are detected imperfectly. *Ecology* 86:1101–1113.
- Magurran, A. E. 2004. *Measuring biological diversity*. Blackwell Science, Oxford, UK.
- Magurran, A. E. and M. Dornelas. 2010. Biological Diversity in a changing world. *Philosophical Transactions of the Royal Society B* 365:3593–3597.
- McGill, B. J. et al. 2007. Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework. *Ecology Letters* 10:995–1015.
- Patil, G. P. and C. Taillie. 1982. Diversity as a concept and its measurement. *Journal of the American Statistical Association* 77:548–561.
- Peet, R. K. 1974. The measurement of species diversity. *Annual Review of Ecology, Evolution and Systematics* 5:285–307.
- Pielou, E. C. 1969. *An Introduction to Mathematical Ecology*. Wiley Interscience, New York, USA.
- R Development Core Team. 2010. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.
- Read, T. R. C. and N. Cressie. 1988. *Goodness-of-Fit Statistics for Discrete Multivariate Data*. Springer, New York, USA.
- Ricotta, C. 2003. On parametric evenness measures. *Journal of Theoretical Ecology* 222:189–197.
- Shannon, C. E. 1948. A mathematical theory of information. *Bell System Technical Journal* 27:379–423.
- Simpson, E. H. 1949. Measurement of diversity. *Nature* 163:688.
- Smith, B. and J. B. Wilson. 1996. A consumer's guide to evenness indices. *OIKOS* 76:70–82.
- Thomas, C. D. 2002. Extinction, Colonization, and Metapopulations: Environmental Tracking by Rare Species. *Conservation Biology* 8:373–378.
- Tokeshi, M. 1990. Niche apportionment or random assortment: species abundance patterns revisited. *Journal of Animal Ecology* 59:1129–1146.
- Tokeshi, M. 1996. Power fraction: a new explanation of relative abundance patterns in species-rich assemblages. *OIKOS* 75:543–550.
- Tothmeresz, B. 1995. Comparison of different methods for diversity ordering. *Journal of Vegetation Science* 6:283–290.
- Yoccoz, N. G., J. D. Nichols, and T. Boulinier. 2001. Trends in Ecology and Evolution 16:446–453.

APPENDIX A

CONNECTION TO SHANNON'S AND SIMPSON'S INDICES

Both statistics, the likelihood ratio G and Pearson's X^2 , can be related to traditional diversity indices, namely Shannon's entropy $H = -\sum p_i \log p_i$ and $D = \sum p_i^2$, which is used in the different versions of Simpson's index:

$$G = 2n \left[\sum_{i=1}^S q_i \log q_i + \log S \right] = 2n(\log S - \hat{H}) \\ = 2n(H^* - \hat{H})$$

$$X^2 = n \left[\sum_{i=1}^S q_i^2 S - 1 \right] = n(\hat{D}S - 1) = n \left(\frac{\hat{D}}{D^*} - 1 \right),$$

where \hat{H} , \hat{D} are the sample-based estimates for Shannon's and Simpson's indices and $H^* = \log S$, $D^* = 1/S$ are the index values if the species abundance distribution is completely even ($\mathbf{p} =$

\mathbf{p}^*). Note that the estimator \hat{D} is not corrected for sampling bias here (Lande 1996).

APPENDIX B

G AND X^2 AS PART OF THE FAMILY $I_n(\lambda)$ OF GOODNESS-OF-FIT STATISTICS

The classic goodness-of-fit statistics, Pearson's X^2 and the likelihood ratio G , are part of the family $I_n(\lambda)$ for $\lambda = 0$ and $\lambda = 1$, respectively:

$$\lim_{\lambda \rightarrow 0} I_n(\lambda) = 2 \sum_{i=1}^S n_i \log \left(\frac{n_i}{n/S} \right) = G$$

$$I_n(1) = \sum_{i=1}^S n_i \left(\frac{n_i}{n/S} - 1 \right) \\ = \sum_{i=1}^S \frac{(n_i - n/S)^2}{n/S} + \sum_{i=1}^S (n_i - n/S) = X^2,$$

using $\lim_{x \rightarrow h} (x^h - 1)/h = \log x$.

APPENDIX C

ASYMPTOTICS AND MOMENTS FOR $I_n(\lambda)$

The asymptotic behavior and in particular the first and second moments of the family of goodness-of-fit $I_n(\lambda)$ is known (Read and Cressie 1988). If the underlying species abundance distribution were indeed perfectly even, all members of the family have the same asymptotic χ^2_{S-1} -distribution. Hence

$$E_{\mathbf{p}=\mathbf{p}^*}[I_n(\lambda)] \approx S - 1,$$

$$V_{\mathbf{p}=\mathbf{p}^*}[I_n(\lambda)] \approx 2(S - 1).$$

In reality, the species abundance distribution is never perfectly even. We are therefore interested in the distributional properties when $\mathbf{p} \neq \mathbf{p}^*$. In this case, the asymptotic equivalence of the statistics in the family is lost. However, they follow an approximate normal distribution where the moments depend on λ :

$$E_{\mathbf{p} \neq \mathbf{p}^*}[I_n(\lambda)] \approx \frac{2n}{\lambda(\lambda + 1)} \sum_{i=1}^S p_i [(p_i S)^\lambda - 1]$$

$$V_{\mathbf{p} \neq \mathbf{p}^*}[I_n(\lambda)] \approx \frac{4n}{\lambda^2} \left[\sum_{i=1}^S (p_i S)^{2\lambda} p_i - \left[\sum_{i=1}^S (p_i S)^\lambda p_i \right]^2 \right].$$

For large n , we have

$$E\left[\frac{1}{2n} I_n(\lambda)\right] \rightarrow I_{\mathbf{p}}(\lambda).$$

Therefore, the ML-estimator

$$\hat{I}_{\mathbf{p}}(\lambda) = \frac{1}{2n} I_n(\lambda)$$

is consistent and asymptotically unbiased (in the setting of multinomial sampling).

APPENDIX D

A STANDARDIZED VERSION OF $I_{\mathbf{p}}(\lambda)$

For $\lambda > -1$, we can define

$$\tilde{I}_{\mathbf{p}}(\lambda) = 1 - \frac{I_{\mathbf{p}}(\lambda)}{\max_{\mathbf{p}} I_{\mathbf{p}}(\lambda)}$$

where $\max_{\mathbf{p}} I_{\mathbf{p}}(\lambda) = 1/[\lambda(\lambda + 1)](S^\lambda - 1)$, i.e., the value of $I_{\mathbf{p}}(\lambda)$ for the least even abundance vector $\mathbf{p} = (1, 0, \dots, 0)$. All members of the standardized index family

$$\tilde{I}_{\mathbf{p}}(\lambda)$$

are constrained to lie in $[0, 1]$.

Application to simulated species abundance distributions (from Tokeshi's niche models) showed that the standardized index only differentiates well for parameter values $-1 < \lambda \leq 1$ (see Fig. D1).

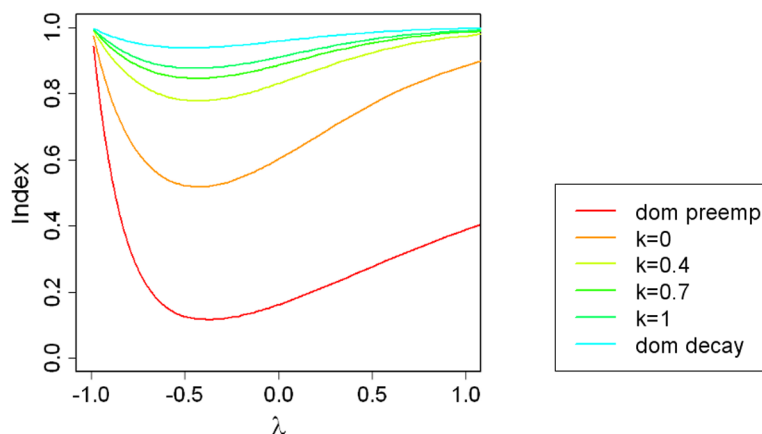


Fig. D1. Standardized version of $I_{\mathbf{p}}(\lambda)$. The plot shows the standardized version of the family of evenness indices for the Tokeshi models (analogously to scenario 1, Fig. 1). Values range between 0 and 1, where 1 corresponds to high evenness.

APPENDIX E

BIAS AND PRECISION IN SAMPLE PROFILES
(SENSITIVITY TO ε): ADDITIONAL PLOTS

Evenness profiles are insensitive to the choice of ε for the positive parameter range, which is where they are comparable with the other profiles. However, they show sensitivity to ε in the negative parameters, for which the other methods cannot be used. The value of ε reflects how “rare” rare species are. As this is what the goodness-of-

fit measures are picking up, the sensitivity cannot be avoided. If we choose a value that is too small, we artificially make these species “rarer” while a value that is too large causes the estimated profile to indicate that the distribution is more even than it actually is. The optimal value seems to depend on the fraction of the population sampled as well as the underlying model. In general, values of 0.05 to 0.1 work reasonably well. Hill’s quotients are biased whatever ε is chosen, and the degree of bias depends on the underlying model (Figs. E1 and E2).

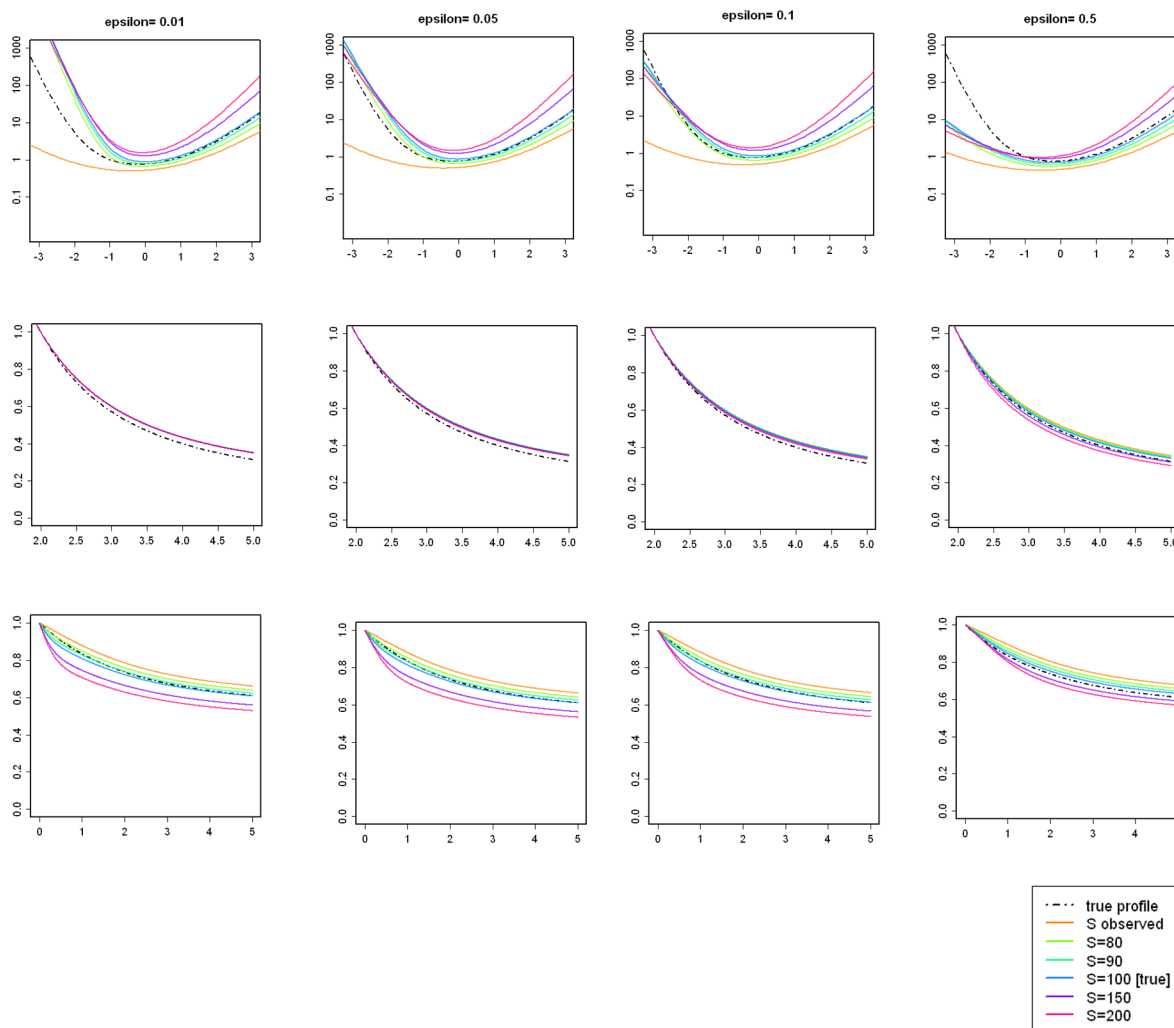


Fig. E1. Bias and precision in sample profiles (sensitivity to ε). Different ε -corrections are applied to the sample profiles for scenario 2 (for an assumed $S = 100$). The true underlying profile is generated from Tokeshi’s power fraction model with $k = 0.4$. Rows from top to bottom show (a) evenness profiles, (b) $(N(a)/N(2))^2$ as well as (c) logarithmic evenness.

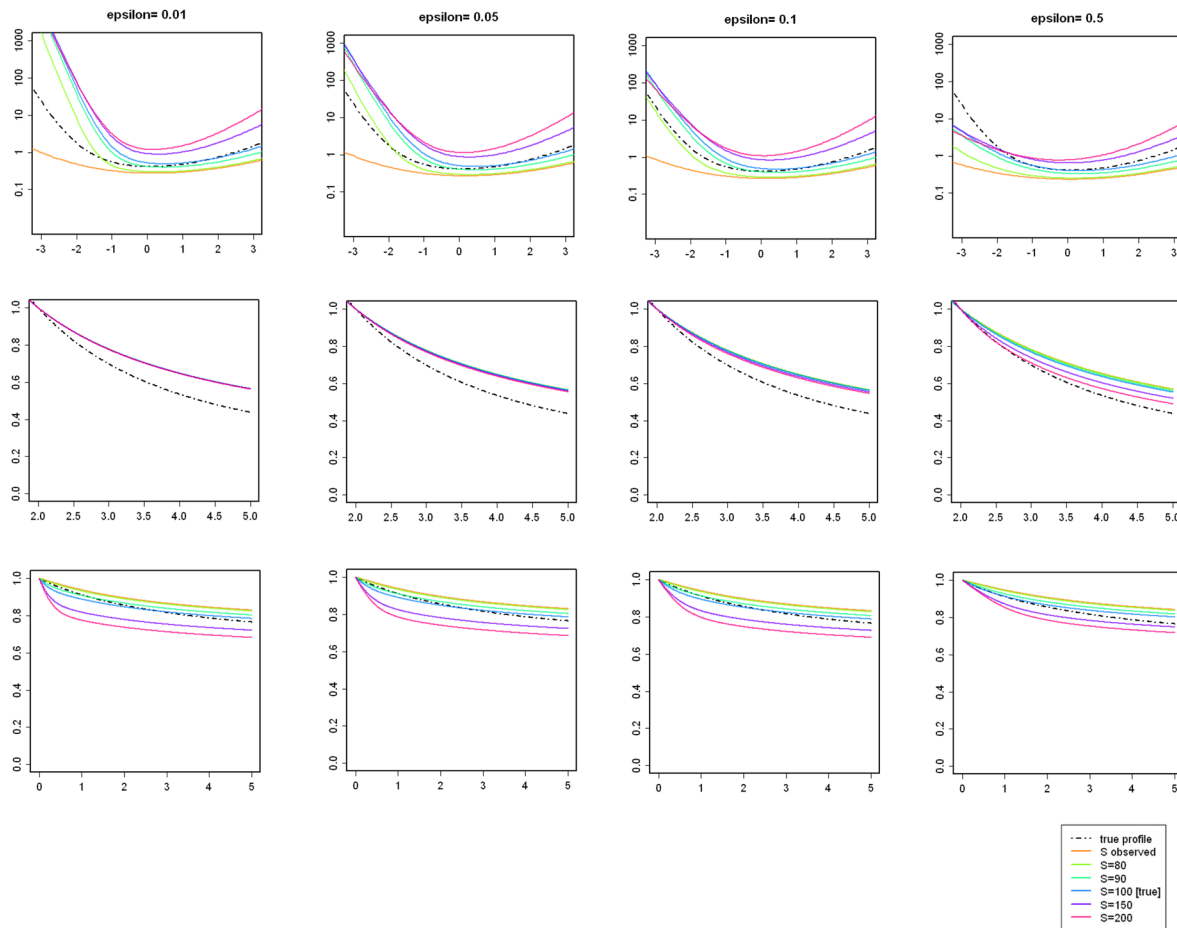


Fig. E2. Bias and precision in sample profiles (sensitivity to ε). Different ε -corrections are applied to the sample profiles for scenario 2 (for an assumed $S = 100$). The true underlying profile is generated from Tokeshi's power fraction model with $k = 1$. Rows from top to bottom show (a) evenness profiles, (b) $(N(a)/N(2))^2$ as well as (c) logarithmic evenness.

SUPPLEMENT

R code for simulations of Tokeshi's niche models (*Ecological Archives* C002-001-S1).