24 July, 2015

Measuring temporal turnover in ecological communities

Short title: Measuring temporal turnover

Hideyasu Shimadzu (hs50@st-andrews.ac.uk),
Maria Dornelas (maadd@st-andrews.ac.uk) and
Anne E. Magurran (aem1@st-andrews.ac.uk)
Centre for Biological Diversity and Scottish Oceans Institute, University of St Andrews

Author correspondence:
Hideyasu Shimadzu
Centre for Biological Diversity and Scottish Oceans Institute,
University of St Andrews, Dyers Brae House, St Andrews, Fife KY16 9TH, UK
Tel: +44 1334 46 3362
Fax: +61 2 6249 9920
hs50@st-andrews.ac.uk

This article has been accepted for publication and undergone full peer review but has not been through the copyediting, typesetting, pagination and proofreading process, which may lead to differences between this version and the Version of Record. Please cite this article as doi:
10.1111/2041-210X.12438

Abstract

1. Range migrations in response to climate change, invasive species and the emergence of novel ecosystems highlight the importance of temporal turnover in community composition as a fundamental part of global change in the Anthropocene. Temporal turnover is usually quantified using a variety of metrics initially developed to capture spatial change. However, temporal turnover is the consequence of unidirectional community dynamics resulting from processes such as population growth, colonisation and local extinction.

2. Here, we develop a framework based on community dynamics, and propose a new temporal turnover measure.

3. A simulation study and an analysis of an estuarine fish community both clearly demonstrate that our proposed turnover measure offers additional insights relative to spatial-context-based metrics.

4. Our approach reveals whether community turnover is due to shifts in community composition or in community abundance, and identifies the species and/or environmental factors that are responsible for any change.

Key words:

Beta diversity; Biodiversity; Population dynamics; Temporal turnover.

1 Introduction

A large number of metrics have been developed to measure community dissimilarity (Anderson et al., 2011, Jost et al., 2011, Vellend et al., 2011, Sæther et al., 2013, Legendre and Gauthier, 2014) in the century since Jaccard (1901, 1912) first proposed a method of quantifying differences in assemblage composition. This variation in the identities of species found in different sites, as measured by Jaccard dissimilarity and other turnover metrics, was dubbed β-diversity by Whittaker (1960, 1972), to distinguish it from within assemblage diversity (known as α-diversity). Although the emphasis, to date, has been on measuring β-diversity across spatial contexts, growing concern about threats to biodiversity underlines an urgent need to quantify and understand temporal turnover (Dornelas et al., 2013). Recent analyses have shown that although the α-diversity of local communities is not changing consistently through time, with many assemblages showing increasing trends and others showing decreasing trends (Vellend et al., 2013, Dornelas et al., 2014), the rate of change in community composition (temporal β-diversity) is greater than predicted by null models of baseline temporal turnover (Dornelas et al., 2014, Supp and Ernest, 2014).

One approach to quantifying temporal turnover is to use metrics such as the Jaccard and Bray-Curtis (Bray and Curtis, 1957) indices, initially developed to capture spatial change. However, temporal turnover has features, such as unidirectional change, not usually present in investigations of spatial β-diversity (Dornelas et al., 2013). Moreover, temporal turnover is the consequence of community dynamics resulting from processes such as local immigration and extinction, population growth and density dependence. Ideally, then, temporal turnover should recognise the dimensionality of temporal change and the ecological processes that lead to shifts in community composition through time.

Here we formalise the concept of temporal turnover. We present a novel framework

for measuring temporal turnover based on the dynamics of species abundance within a community. The turnover measure we propose is directly derived from a general form of a population dynamics model which describes the extent to which each population, and consequently the community to which these species belong, changes over time. Our framework highlights the fact that the concept of turnover can be dissected into two key ecological aspects: change in community composition and change in community size (or capacity) *sensu* its abundance (Brown, 1981). In addition to providing new insights into the ecological basis of measures of temporal turnover, our approach has important practical applications because it allows the user to identify which species and/or which environmental factors play a key role in the change of the species community. We demonstrate how the turnover measure performs using a simulation and by analysing an estuarine fish community time-series from the Bristol Channel, UK. We also discuss the relationship between our approach and existing methods of α- and β-diversities. For β-diversity, we identify links between our measure and those previously developed to capture the nestedness and turnover components of β-diversity (Baselga, 2010, 2012, 2013).

2 A measure of temporal turnover

The temporal changes that all ecological communities experience are a cumulative consequence of changes in the abundance of each species in the community. Population ecology describes the dynamics of species abundance by differential equations, modelling the change rate of the population size of a given species. We can consider this rate of change in abundance as, literally, the temporal *turnover* of the species in question. This assumption is the basis of our approach. The temporal turnover of the whole community can then be defined as an additive effect of turnover in each species, since the total abundance of the community is the sum of the species abundances in the community. Keeping these points

in mind, we propose a new measure of temporal turnover of communities, and identify two key ecological components of this turnover: change in community composition and change in total abundance.

Throughout the paper we consider an ecological community consisting of s species whose abundance varies over time governed by the population's background state at each time t. We use the term *expected abundance* to denote the background state, which is often described using other terms such as mean abundance, model expectation and (ideal) population size (a more precise definition will follow in Section 3). We recognise that observed abundances reflect both expected abundance and natural variability. This recognition allows us, when we have observations, to separate the effect of natural variability from the expected abundance in which features of turnover can be found. As such, here we construct a framework on the expected abundance, instead of the observations themselves.

Let $\lambda_i(t) > 0$ be the expected abundance of the i-th species at time t. Note the fact that the observed abundance can be zero even though the expected abundance is assumed to be always positive. A species with very low expected abundance (close to zero) tends to be absent in observations most of the time. We can therefore deal with the case of species absence, adopting expected abundance as the basis of our framework. The expected total-abundance of the community is then the sum of each species' expected abundance as $\lambda(t) = \sum_{i=1}^{s} \lambda_i(t)$. This fact is always valid regardless of whether the species abundances within a community are independent or not, and serves as the basis of our discussion below.

We begin by examining the dynamics of the populations that make up the community. Consider an equation for the expected abundance of the i-th species as

$$\lambda_i(t + dt) = (1 + \alpha_i(t)dt)\lambda_i(t), \tag{1}$$

where $\alpha_i(t)$ is the instantaneous change rate at time t that drives the increase or the decrease of the species abundance; the abundance increases when $\alpha_i(t) > 0$, and decreases when $\alpha_i(t) < 0$. If the equation (1) is solved, the Malthusian growth model (Malthus, 1798) specifies the expected abundance, $\lambda_i(t)$, and the instantaneous change rate, $\alpha_i(t)$, becomes the growth rate of the population. The rate can also be modelled by other processes such as those, birth, death, immigration and emigration rates, $b_i(t) > 0$, $d_i(t) > 0$, $m_i(t) > 0$ and $e_i(t) > 0$ respectively, like $\alpha_i(t) = b_i(t) - d_i(t) + m_i(t) - e_i(t)$. This of course leads $\lambda_i(t)$ to another type of model. As such, in order to keep our framework below general, we do not specify any functional form of the expected abundance, $\lambda_i(t)$, at this stage, but it will be discussed in Section 3. Note that when those rates are offset relative to each other, $\alpha_i(t) = b_i(t) - d_i(t) + m_i(t) - e_i(t) = 0$, the population's abundance is stable for the time increment dt, since its expected abundance becomes constant, $\lambda_i(t + dt) = \lambda_i(t)$.

The reasoning given above implies that change rate is a legitimate indicator to quantify, literally, the turnover of the i-th species, based on its abundance, for the time increment dt. However, as the change rate is in most cases unknown, it needs to be determined by the trajectory of the expected abundance, $\lambda_i(t) > 0$, itself as

$$\begin{aligned} \alpha_i(t)dt &= \frac{\lambda_i(t+dt) - \lambda_i(t)}{\lambda_i(t)} = \frac{d\lambda_i(t)}{\lambda_i(t)} \\ &= d\log\left(\lambda_i(t)\right). \end{aligned}$$

Here and elsewhere in the paper log refers to the natural logarithms.

We now consider the community scale, represented by the expected total-abundance of the species in the community $\lambda(t) = \sum_{i=1}^{s} \lambda_i(t)$, applying the same analogy used for the population scale. Although the continuous form of the expected abundances, $\lambda_i(t), i = 1, 2, \ldots, s$, is unknown, we can estimate it using observations at discrete time points. We

therefore rewrite Equation (1) in a discrete form representing a community dynamics model over a relatively short time interval, between times t and $t+h$ ($h > 0$), within which the difference equation is still valid. By integrating the instantaneous change rate, $\alpha_i(t)$, over the time interval, the community dynamics can be described as

$$\begin{aligned}
\lambda(t+h) &\doteq \sum_{i=1}^{s} \left(1 + \int_t^{t+h} \alpha_i(v)dv\right) \lambda_i(t) \\
&= \sum_{i=1}^{s} \left(1 + \int_t^{t+h} d\log(\lambda_i(v))\right) \lambda_i(t) \\
&= \sum_{i=1}^{s} \left(1 + \log\left(\frac{\lambda_i(t+h)}{\lambda_i(t)}\right)\right) \lambda_i(t).
\end{aligned}$$

The change rate and the instantaneous change rate may therefore hold the following relationship for the short time period h,

$$\frac{\lambda(t+h) - \lambda(t)}{\lambda(t)} \doteq \sum_{i=1}^{s} \log\left(\frac{\lambda_i(t+h)}{\lambda_i(t)}\right) p_i(t), \tag{2}$$

where $p_i(t)$ is the relative abundance of the i-th species at time t, that is determined as

$$p_i(t) = \frac{\lambda_i(t)}{\lambda(t)}.$$

The collection of the relative abundances of the community — the relative abundance distribution $\boldsymbol{p}(t) = \{p_1(t), p_2(t), \ldots, p_s(t)\}$ — satisfies the conditions that $p_i(t) > 0$ and $\sum_{i=1}^{s} p_i(t) = 1$.

Now we can formally define the turnover measure, which we call D. From Equation (2),

we define the turnover measure between times t and $u, (u > t)$ as

$$
\begin{aligned}
D(t:u) = \sum_{i=1}^{s} d_i(t:u) &= \sum_{i=1}^{s} \log\left(\frac{\lambda_i(u)}{\lambda_i(t)}\right) p_i(t) && (3)\\
&= -\sum_{i=1}^{s} \log\left(\frac{p_i(t)}{p_i(u)}\right) p_i(t) + \log\left(\frac{\lambda(u)}{\lambda(t)}\right) && (4)\\
&= D_1(\boldsymbol{p}(t):\boldsymbol{p}(u)) + D_2(\lambda(t):\lambda(u)).
\end{aligned}
$$

As shown above, temporal turnover, D, is an abundance based measure, and is additively dissected into two quantities, D_1 and D_2. Figure 1 illustrates a schematic picture of our framework for measuring the temporal turnover discussed above.

The proposed turnover measure has the following properties:

1. $-\infty < D < \infty$;
2. $D = 0 \Longleftrightarrow D_1 = 0$ and $D_2 = 0 \Longleftrightarrow \boldsymbol{p}(t) = \boldsymbol{p}(u)$ and $\lambda(t) = \lambda(u)$;
3. $D_1 \leq 0$ (*cf.* Kullback and Leibler (1951));
4. $D_2 > 0 \Longleftrightarrow \lambda(u) > \lambda(t)$, $D_2 < 0 \Longleftrightarrow \lambda(u) < \lambda(t)$.

Note that Property 2 holds when the instantaneous change rate (Equation 1) of the every species in the community is equal to zero, $\alpha_i(t) = 0, i = 1, 2, \ldots, s$.

The implications of Equation (4) are worth consideration. First, it suggests that the temporal turnover is additively decomposed into two parts: the first term (D_1) related to the amount of change in community composition, and the second term (D_2) being dependent only on the amount of change in community size *sensu* its abundance. This fact highlights two important aspects in evaluating the turnover of species community: 1) change in community composition and 2) change in total abundance.

Second, it is interesting that the Kullback–Leibler divergence (Kullback and Leibler, 1951) is subsequently derived from the definition (Equation 3) as D_1. Kullback–Leibler divergence

quantifies the difference between two relative abundance distributions of the community at different time occasions, $\boldsymbol{p}(t)$ and $\boldsymbol{p}(u)$. Thus, the first term, D_1, is interpreted as the part evaluating the compositional change of the community.

The basis of our framework, equation (2), highlights a key feature of temporal turnover — the arrow of time. If $\lambda(t+h)$ and $\lambda(t)$ are swapped, the instantaneous change rate $\alpha_i(t)$ takes a different form. In other words, the interpretation of the temporal turnover measure is asymmetrical to time — a feature that contrasts with spatial turnover. Although the proposed turnover measure, D, could be applied to spatial contexts, by assigning t and u to different locations, it is important to note that its interpretation will differ from the temporal one. This is because of the asymmetrical structure inherent in temporal turnover whereas symmetricity is usually assumed when studying spatial turnover.

3 Specifying the expected abundance $\lambda_i(t)$

We introduced the idea of the expected abundance earlier, but have not specified it yet. Since we can only observe species abundances and never observe the expected abundance in ecological investigations, we need to determine the expected abundance from the observations. Here, we denote the observed abundance of the i-th species at time t as $n_i(t)$, lower case, and treat it as a realisation of a random variable $N_i(t)$, upper case, which follows the Poisson distribution with the time varying mean parameter $\lambda_i(t)$. We adopt the convention of using lower and upper cases for non random and random variables respectively. The expected abundance is then specified as the expectation (or the mean) of the random variable as $\lambda_i(t) = \mathrm{E}\left[N_i(t)\right]$. A range of approaches are available for estimating the expected abundance, $\lambda_i(t)$, from observations, $\{n_i(t)\}$; the choice is largely dependent on the researcher and data characteristics, such as the number of observations. We give here a short summary of common approaches, widely used in population and community ecology.

Many methods of estimating the expected abundance rely on the maximum likelihood principle. For instance, recalling the assumption we made, an approach widely used in the ecological literature is to use the observation itself as $\hat{\lambda}_i(t) = n_i(t)$. Here, the hat sign means its estimate. Using the observation as the estimate of the expected abundance can be regarded as the maximum likelihood estimate (MLE) when we have only one observation obtained for estimating the expected abundance at time t. Note that this may violate some properties of the turnover measure discussed in Section 2, when $n_i(t) = 0$ is involved. Another choice is taking the average of the observations by assuming the expected abundance to be constant, $\lambda_i(t) = \lambda_i$, over the observation period, $t = 1, 2, \ldots, t_n$, the MLE of which choice is then given as $\hat{\lambda}_i = \sum_{t=1}^{t_n} n_i(t)/t_n$.

An alternative approach is to model expected abundance, $\lambda_i(t)$, in relation to relevant environment factor(s). Although this approach requires more data than those outlined above, a range of regression type models are available. Examples include generalised linear models (GLMs; McCullagh and Nelder, 1989) and generalised additive models (GAMs; Hastie and Tibshirani, 1990), that delineate the relationship between species abundance and environmental factors as well as other species' abundances (Kedem and Fokianos, 2002). This approach also entails specifying the link function and the distribution function that the abundance, $N_i(t)$, follows, with common examples including Poisson, Negative Binomial and others, see Zuur et al. (2009), Hilbe (2011) for details. This type of modelling approach offers more detailed insights in terms of interpreting community change as is discussed in the next section. We stress that the choice of estimation approach will depend on the extent to which information is available for study, and that $n_i(t) = 0$ does therefore not necessary mean $\hat{\lambda}_i(t) = 0$ unless the observation is used as the estimate. Many other approaches are, of course, applicable and not limited to those discussed above. For example, determining the trajectory of abundance, $N_i(t)$, by (stochastic) differential equations, as did Ives et al.

(2003), Mutshinda et al. (2009, 2011), is also a sensible approach.

4 Identifying the influential drivers

An advantage of modelling species expected abundance, $\lambda_i(t)$, is that it becomes possible to identify drivers that influence the turnover measure, D. Equation (3) reveals how this can be done. We illustrate this using a common model class, GLMs with the log link function, such as Poisson and Negative Binomial distributions, as an example.

Suppose that the expected abundance of the i-th species, $\lambda_i(t)$, is modelled by a GLM with the log link as

$$\log(\lambda_i(t)) = \sum_{j=1}^{m} \beta_{ij} x_j(t), \tag{5}$$

where $\{x_j(t)\}$ are environmental variables, and $\{\beta_{ij}\}$ are the parameters to be estimated. From Equations (3) and (5), we have the turnover measure, D, described in an additive form as

$$D(t:u) = \sum_{j=1}^{m} \sum_{i=1}^{s} \beta_{ij}(x_j(u) - x_j(t))p_i(t). \tag{6}$$

Since the contribution of the species and the environment factors are all additive, putting

$$\begin{aligned} d_i(t:u) &= \sum_{j=1}^{m} \beta_{ij}(x_j(u) - x_j(t))p_i(t) \text{ and} \\ d_j(t:u) &= \sum_{i=1}^{s} \beta_{ij}(x_j(u) - x_j(t))p_i(t), \end{aligned}$$

the contribution ratio of the i-species and of the j-th environment variable to the turnover

measure, D, is respectively defined as

$$r_i(t:u) = \frac{|d_i(t:u)|}{\sum_{i=1}^{s}|d_i(t:u)|}, \quad (7)$$

$$r_j(t:u) = \frac{|d_j(t:u)|}{\sum_{j=1}^{m}|d_j(t:u)|}. \quad (8)$$

These quantities show what proportion each factor contributes to the absolute amount of the turnover. Whether this type of additive decomposition is available is largely dependent on how the species expected abundance, $\lambda_i(t)$, has been modelled. Although our example makes use of a GLM, as long as the right-hand-side of model (5) is additive in terms of the environment variables, as it is in GAMs, the additive decomposition (Equation 6) is available, and consequently the contribution ratios, Equations (7) and (8), can be obtained. A key fact is that the turnover measure D can be additive in terms of species i *and* environment factor j, given an appropriate model form and a link function. A similar idea can be found in calculating the contribution of individual species based on the Bray–Curtis index (for example, `simper` of an R package 'vegan', Oksanen et al., 2014) although, calculating the contribution of environment variables may not be straightforward since the Bray–Curtis index is, in general, non additive with respect to the environment variables.

5 Simulation study

To illustrate how our turnover measure works, and the kind of information that can be obtained, we perform a simulation study, and compare the results with those produced using two popular turnover metrics: Jaccard (Jaccard, 1901) and Bray–Curtis (Bray and Curtis, 1957). We simulate an ecological community within which each species hold inter-species relationships keeping the zero-sum community condition (Hubbell, 2001), apart from an imposed abrupt change to represent a disturbance as per Dornelas (2010).

5.1 Simulated data

In the simulation study, we assume an ecological community of 30 species ($s = 30$), within which species abundances at time t, $\boldsymbol{N}(t) = (N_1(t), N_2(t), \ldots, N_s(t))^\top$, are dependent on each other in a sense of zero-sum community (Hubbell, 2001). The abundances distribute as a multinomial distribution, $\boldsymbol{N}(t) \sim \mathsf{Mn}(n(t), \boldsymbol{p}(t))$, and the covariance between any pair of species, i and i', is given as $\mathrm{Cov}\left[N_i(t), N_{i'}(t)\right] = -n(t)p_i(t)p_{i'}(t)$ for $i \neq i'$. We generate series of abundance values (200 time steps for each of the 30 species), $\{\boldsymbol{N}(t) = \boldsymbol{n}(t) : 1 \leq t \leq 200\}$, for the whole community with an artificial change taking a place at $t = 100$ as

$$\Pr\left(\boldsymbol{N}(t) = \boldsymbol{n}(t); n, \boldsymbol{p}\right) = \begin{cases} \mathsf{Mn}(n, \boldsymbol{p}) & t < 100, \\ \mathsf{Mn}(n', \boldsymbol{p}') & t \geq 100, \end{cases}$$

where n and n' are the total abundance of the community and $\boldsymbol{p}$ and $\boldsymbol{p}'$ are the relative species abundances. In other words, the simulated community is a zero-sum community whose total size is respectively n and n' before and after the time point $t = 100$. The initial values of the relative species abundances, $\boldsymbol{p}$, are obtained using the program provided by Dornelas (2010). We consider three types of changes in community size at time t: stable ($n = n'$), an increase ($n < n'$) and a decrease ($n > n'$). For the compositional change, we consider two cases: no-alteration ($\boldsymbol{p} = \boldsymbol{p}'$) and alteration ($\boldsymbol{p} \neq \boldsymbol{p}'$) in the relative species abundances. The scenarios to be considered are specified by a combination of those states, but we have omitted the trivial case, stable and no-alteration ($n = n'$ and $\boldsymbol{p} = \boldsymbol{p}'$), so that there are five scenarios to be examined. See Appendix C for the detailed explanation.

5.2 Calculating turnover metrics

For each of our five simulated community time series, $\{n_i(t) : 1 \leq t \leq 200, i = 1, 2, \ldots, 30; n, n', \boldsymbol{p}, \boldsymbol{p}'\}$ $(n = 200, n' = 180, 200, 230)$, the three turnover metrics: our measure D, Jaccard and Bray–Curtis indices, are calculated in two different ways. One is based on a GLM and the other is based on the simulated raw data, $\hat{\lambda}_i(t) = n_i(t)$.

First, we fit a GLM with a Poisson distribution as

$$\log(\lambda_i(t)) = \theta_{i1} + \theta_{i2} I(t \geq 100), \tag{9}$$

where θ_{i1} and θ_{i2} are the parameters to be estimated by maximising the likelihood. The fitted model (Equation 9) is exact re-parametrisation of the simulation data to capture the abrupt change introduced at time $t = 100$. Once the expected abundance is estimated, $\hat{\lambda}_i(t)$, the estimated relative abundance $\hat{p}_i(t)$ can be calculated as $\hat{p}_i(t) = \hat{\lambda}_i(t)/\hat{\lambda}(t)$, where $\hat{\lambda}(t) = \sum_{i=1}^{s} \hat{\lambda}_i(t)$. The turnover metrics: our turnover measure, Jaccard and Bray–Curtis are then calculated. However, these indices are originally defined on observations, $\{n_i(t)\}$, so we introduce a model-based version of Jaccard (J_λ) and Bray–Curtis (BC_λ) indices for this study. See Appendix B.

Second, we also calculated the three turnover metrics based on the simulated data, in other words the observation $n_i(t)$ itself, $\hat{\lambda}_i(t) = n_i(t)$. In doing so, we adopt the convention $\log(0/p_i(u))0 = 0$ for D_1 of Equation (4) when the i-th species is newly observed in the community at time u, and also we omit the case where the i-th species becomes absent at time u, $\log(p_i(t)/0)p_i(t) = \infty$, although this treatment may conceal the cases $D_1 = -\infty$, it is still useful for illustration purposes. Alternatively, an arbitrary very small value, $\varepsilon \approx 0$, could be used for $\hat{\lambda}_i(t)$ when $n_i(t) = 0$ to avoid numerical difficulties.

5.3 Result

For illustration purposes, here we show the result (Figure 2) of the scenario, a decrease in the total abundance ($n = 200 > n' = 180$) and an alteration ($\boldsymbol{p} \neq \boldsymbol{p}'$) in the relative species abundances. See Figures A1–A4 in Appendix C for other scenarios. Since we are interested in documenting cumulative changes, all turnover metrics are calculated against the first point of the time series (i.e. the baseline). To reduce sensitivity to the exact nature of the baseline chosen, taking an average over some time points could also be a possible option.

In this simulation, the Jaccard index of dissimilarity, calculated on the raw simulated data ($1 - J$) and on the model ($1 - J_\lambda$), reveals no turnover of the community structure. In contrast, the Bray–Curtis index, BC (original metric as in Bray and Curtis (1957)) and BC_λ (as in Appendix B) does detect a shift in turnover. This contrast reflects the fact that species richness of the simulated community has remained relatively constant throughout, the dominant species are observed constantly but the low abundance species are observed occasionally, whereas the abundances of the species within the community shifted at the perturbation point, $t = 100$. This analysis also highlights differences between the metrics based on the observations and those based on the model. The model-based ones tend to be robust against the variation due to observations, a feature that could be useful for detecting this kind of abrupt change.

Our new turnover measure offers additional insights relative to the Jaccard and Bray–Curtis indices discussed above. D_1 (the green line) quantifies the change in the community composition. With D_1 the absolute divergence from zero indicates change has occurred in the relative abundance distribution of the community. For example, the addition of an invasive species to a community, leading to the local extinction of a native species would be detected by this metric. The change in community abundance is indicated by D_2 (the blue line). When D_2 is positive the community abundance has increased, and when

it is negative community abundance has decreased. For those scenarios involving decreased community abundance, $n > n'$ (Figures 2 and A1 in Appendix C), D_2 is actually negative suggesting the community abundance decreased, whereas the Bray–Curtis index struggles to show such a difference, whether the community abundance is increased or decreased, because of it accounts for absolute differences. Once again the model-based measure of turnover, D and D_2, provide a clear indication of the perturbation. A mass mortality event associated with a chemical spill, for example, would be detected by this metric.

6 Application

6.1 Data

We further explore the performance of our new turnover measure using an exceptionally complete estuarine community time series. Sampling took place at Hinkley Point 'B' power station on the southern bank of the Bristol Channel in Somerset, UK (51°14′14.05″N, 3°8′49.71″W). Monthly quantitative sampling of a fish community commenced in January 1981, and more than 80 species have been observed over the last three decades (1981 – 2012) (Henderson and Holmes, 1991, Henderson and Seaby, 2005, Henderson, 2007, Henderson and Bird, 2010, Henderson et al., 2011). Ambient environmental factors, namely water temperature and tide height, have also been recorded. See Henderson and Bird (2010) for details of the survey and its methodology.

6.2 Model

Following Shimadzu et al. (2013), we fit the same GAMs to the 45 core species that are consistently present in the assemblage (Magurran and Henderson, 2003). The GAM for each

i-th abundant species is

$$\log(\lambda_i(t)) = \beta_{i0} + f_{i1}(\texttt{Year}) + f_{i2}(\texttt{Tide.height}) + f_{i3}(\texttt{Water.temp}) + f_{i4}(\texttt{Month}),$$

where the β_{i0} is a constant and $f_{ij}(\cdot)$ is a smoothing spline function. For the remaining 36 species that occur infrequently we fit a constant model as the simplest model,

$$\log(\lambda_i(t)) = \beta_{i0}.$$

Once the expected abundance is estimated, $\hat{\lambda}_i(t)$, the estimated relative abundance $\hat{p}_i(t)$ can then be calculated as $\hat{p}_i(t) = \hat{\lambda}_i(t)/\hat{\lambda}(t)$, where $\hat{\lambda}(t) = \sum_{i=1}^{s} \hat{\lambda}_i(t)$.

6.3 Result

For illustration purposes, we present the turnover result for the month August over the period (Figure 3). The top panel (Figure 3a) shows the extent to which the fish community changes each year relative to the first observation, August 1981. The turnover measure D (the red line) shows a reasonably stable fluctuation a little below the zero level over the period, and also suggests cyclical variation, since the late 1990s. The two components of the turnover measure, one related to the change in community composition D_1 (the green line) and the other related to the change in community abundance D_2 (the blue line), tell a more detailed story. The green line (D_1) shows the changes in the relative abundance distribution of the community, quantified as the departure from zero. The continuous decrease indicates that the species composition of the community has been gradually changing since 1981. On the other hand, the blue line (D_2) resembles the cyclic variation of the red line (D) but with an increase over the period. This indicates that the community size has been getting slightly larger, relative to the 1981 level, since the late 1990s.

In Figure 3b, the contribution ratios of some species resemble the cyclical variation in D and D_2. The middle panel illustrates each species' contribution ratio defined by Equation (7). Each coloured area represents the proportion of contribution to the change of the turnover measure D made by the species. A closer look reveals at least five species to be influential in the community: *Trisopterus minutus*, *Pomatoschistus minutus*, *Merlangius merlangius*, *Solea solea* and *Sprattus sprattus* (see the legend in Figure 3b). The key role of these five species is reflected in their large contribution ratio (Equation 7), inducing the departure of the turnover measure D from the zero level (Figure 3a). The large area in Figure 3b assigned to *Trisopterus minutus* means that it has been particularly influential over the period. Interestingly, *Pomatoschistus minutus* and *Merlangius merlangius* appear to be offset — as one increases in contribution ratio the other decreases, and their cyclical variations seem to correspond to the turnover measure D and D_2 in the late 1990s. *Sprattus sprattus*, a known dominant species in the community, has a largely constant contribution since the late 1980s.

The bottom panel (Figure 3c) shows the contribution ratio of the three environmental factors: year (blue), tide height (green) and water temperature (yellow). The influence of each factor varies through time. The peaks in the year effect match with those of D and D_1, implying the fact that this community has been under the strong influence by some particular speices whose abundance fluctuates in a relatively regular cycle.

7 Relationship with α-diversity measures

The framework that we have proposed is conceptually linked with α-biodiversity measures. Measuring biodiversity using Shannon's entropy is equivalent to examining community change rate, but between two particular occasions: the reference time and the time when the community composition has become uniform. Consider a case where the community consisting of s species at the reference time t and at a particular time $u, (u > t)$ when the

relative abundance distribution of the community is homogeneous, $\boldsymbol{p}(u) = 1/\boldsymbol{s}$. The turnover measure D (Equation 2) is then reduced to

$$D(t:u) = -\log(s) - \sum_{i=1}^{s} p_i(t)\log(p_i(t)) + \log\left(\frac{\lambda(u)}{\lambda(t)}\right).$$

If the community abundance stays the same, $\lambda(t) = \lambda(u)$, the turnover measure is further simplified as

$$D(t:u) = D_1(t:u) = -\log(s) - \sum_{i=1}^{s} p_i(t)\log(p_i(t)).$$

This is exactly Shannon's entropy with a constant shift, $-\log(s)$. In other words, the two terms on the right-hand-side are, respectively, a species richness measure and a heterogeneity measure. Interestingly, this links back to the introduction of Shannon's entropy for measuring biodiversity by Margalef (1957, 1958). Combining the above reasoning with our framework discretising differential equations, it is clear that Margalef's idea of using Shannon's entropy bridges theoretical and empirical work within biodiversity research. Theoretical population ecology has modelling of the change rate of populations as its heart. As we have explained above, empirical studies that calculate Shanon's entropy are doing something that is mathematically equivalent to investigating the change rate, but using field data.

There have been many attempts to achieve generality in the heterogeneity component of biodiversity, although by doing so the theoretical link with the community change rate as mentioned earlier is lost. Hill (1973) is one of the first researchers to notice that community heterogeneity can be measured by changing the emphasis on species dominance in a community. In terms of the turnover measure, D, a generalisation equivalent to Hill (1973) can further be made on the change in community composition, D_1. Since Hill's number is the natural log scale of Rényi's entropy (Rényi, 1961), Rényi's divergence that encompasses Kullback–Leibler divergence as a special case may be used as a measure of the

change in community composition, D_1. There have also been a variety of divergence measures examined in information theory, for example see Kawada (1987), Read and Cressie (1988), Cichocki and Amari (2010).

β-diversity is often discussed in the context of partitioning of diversity with other types of diversity measures, such as γ- and α-diversities (Whittaker, 1960, Lande, 1996, Jost, 2007, Chao et al., 2012). In this context Marcon et al. (2012) have derived a β-diversity measure, based on Shannon's entropy and its decomposition, using weighted Kullback–Leibler divergence. Reeve et al. (2014) have also studied diversity partitioning in the spatial context, and have discussed a β-diversity measure that considers species similarity by taking the exponent of Kullback–Leibler divergence. However, we stress that our turnover measure, D, has been derived directly from the community dynamics. Moreover the derivation of D is independent of biodiversity partitioning.

8 Bray–Curtis index and its decomposition

Temporal β-diversity (temporal turnover) represents the change in biodiversity between two occasions. The Bray–Curtis index is an abundance-based metric of community dissimilarity, initially developed to measure spatial change but now increasingly used in the temporal context. Although it is, to be precise, defined using observed abundances, not on the expected abundance, there is a direct link between the turnover measure we propose and the Bray–Curtis index (see Appendix B for the original definition). In fact, the Bray–Curtis index, BC, can be rewritten as a function of our turnover measure, D, as

$$BC \doteq \frac{\sum_{i=1}^{s} |d_i|}{2 + D}, \tag{10}$$

where $|d_i|$ is the absolute value of $d_i = \log(\lambda_i(u)/\lambda_i(t))p_i(t)$, and recall that the turnover measure D is the sum of these d_i's over the species i ($i = 1, 2, \ldots, s$) which make up the community, so that $D = \sum_{i=1}^{s} d_i$ (Equation 3). Figure 4a illustrates this functional relationship (Equation 10). The vertical striped area shows where the value of Bray–Curtis index (BC) can lie, given the value of the turnover measure (D), as the numerator of Equation (10) varies depending on the extent to which the species in the community increase or decrease. The black solid line represents the lower bound of Bray–Curtis index, the case where the abundances of all species have increased. The lower bound clearly shows that the Bray–Curtis index is asymmetrical against the community abundance change. In other words, the Bray–Curtis index tends to amplify the effect of decreases in community abundance compared to its increase.

Baselga (2010, 2012, 2013) studied a decomposition of β-biodiversity metrics, such as Jaccard, Sørensen (Sørensen, 1948) and Bray–Curtis indices, and showed that they can be additively decomposed into two components. In particular, the Bray–Curtis index can be dissected into balanced variation in abundance and the abundance gradient components (Baselga, 2013). We note here that our decomposition (Equation 4) has parallels with but is not equivalent to these components. In fact, the balanced variation in abundance, d_{BC-bal}, and the abundance gradient, d_{BC-gra}, are functions of the turnover measure, $D = D_1 + D_2$.

The balanced variation in abundance is described as

$$d_{BC-bal} \doteq \begin{cases} D^+ e^{-D_2}, & (D < 0) \\ \\ -D^-, & (D > 0) \end{cases} \tag{11}$$

where $D^+ = \sum_{i=1}^{s}(\lambda_i(u) - \lambda_i(t))_+/\lambda(t)$ and $D^- = \sum_{i=1}^{s}(\lambda_i(u) - \lambda_i(t))_-/\lambda(t)$. Here $(\cdot)_+$

and $(\cdot)_-$ respectively denotes its positive or negative part; note that $D = D^+ + D^-$ but clearly $D_1 \neq D_2 \neq D^+$ or D^-. This means that the balanced variation in abundance can be interpreted as the proportion of abundance increased or decreased. Figure 4b illustrates that there is little relationship between the balanced variation in abundance, d_{BC-bal}, and the change in the species composition, D_1.

The abundance gradient is also described as

$$d_{BC-gra} \doteq \begin{cases} \dfrac{|D|}{2+D} q e^{-D_2}, & (D < 0) \\ \\ \dfrac{D}{2+D} q, & (D > 0) \end{cases} \tag{12}$$

where $q = \sum_{i=1}^{s} \min(\lambda_i(u), \lambda_i(t))/\lambda(t)$. This is proportional to Bray–Curtis index itself, and is its lower bound. Figure 4c illustrates the relationship between the abundance gradient, d_{BC-gra}, and the change in the total abundance, D_2. For the detailed derivations of Equations (11) and (12) from Baselga (2013) paper, see Appendix A.

Legendre (2014) has recently reviewed approaches to partitioning β-diversity, and identified two major classes: the Baselga family that we have discussed above, and the Podani family (Podani and Schmera, 2011, Carvalho et al., 2013); see also Legendre and De Cáceres (2013) for the discussion on partitioning β-diversity. Since the Podani family is mathematically related to the the Baselga one, it can also be described as a function of our turnover D. The R package 'BAT' (Cardoso et al., 2014) can be used to partition β-diversity.

9 Conclusion

We have developed a framework for measuring the temporal turnover of species communities based on community dynamics resulting from processes, such as local immigration, extinction

and population growth. Our framework decomposes turnover into two classical ecological rules (Brown, 1981): change in the community composition associated with allocation rules and change in community abundance due to capacity rules. The new temporal turnover measure combines these two aspects in an additive manner. The formal study of the framework highlights the links with other widely used methods of measuring α- and β-diversity.

The performance of the new turnover measure was examined by a simulation study and the analysis of an estuarine fish community. A comparison with other common indices, namely the Jaccard and Bray–Curtis indices, demonstrated that our measure offers new insights into the ecological basis of community change. In addition the simulation study has clearly shown that the approach we have described in this paper can provide useful information in investigating the community change based on observational data. Furthermore, we have demonstrated that our framework offers an important practical advantage in allowing researchers to identify the species and/or environmental factors that play important roles in community change through time.

Taken together, these features of this new approach suggest that it will be a useful framework for quantifying and interpreting temporal turnover in ecological communities at a time when the natural world is facing unprecedented threats (Butchart et al., 2010).

Acknowledgements

Many thanks go to the Associate Editor and two anonymous reviewers for their valuable comments that have improved the paper greatly. We acknowledge support from the European Research Council (project BioTIME 250189) and the Royal Society. MD acknowledges funding from the Marine Alliance for Science and Technology Scotland (MASTS). MASTS is funded by the Scottish Funding Council (grant reference HR09011) and contributing

institutions. We are grateful to Peter Henderson for allowing us to use the Hinkley Point data, and to Oscar Gaggiotti for his helpful comments on the manuscript.

The authors declare that they have no conflict of interest.

Data Accessibility

The Hinkley Point data we used are archived by Pisces Conservation, `http://consult.pisces-conservation.com/hinkdataset.html`. These data are the property of Dr Peter Henderson and Pisces Conservation Ltd. and is not to be used for commercial purposes and in any reports which are produced under a commercial contract to a third party without the explicit permission of Pisces Conservation Ltd.

References

Anderson, M. J., Crist, T. O., Chase, J. M., Vellend, M., Inouye, B. D., Freestone, A. L., Sanders, N. J., Cornell, H. V., Comita, L. S., Davies, K. F., Harrison, S. P., Kraft, N. J. B., Stegen, J. C. and Swenson, N. G. (2011). Navigating the multiple meanings of β diversity: a roadmap for the practicing ecologist, *Ecology letters* **14**(1): 19–28.

Baselga, A. (2010). Partitioning the turnover and nestedness components of beta diversity, *Global Ecology and Biogeography* **19**(1): 134–143.

Baselga, A. (2012). The relationship between species replacement, dissimilarity derived from nestedness, and nestedness, *Global Ecology and Biogeography* **21**(12): 1223–1232.

Baselga, A. (2013). Separating the two components of abundance-based dissimilarity: balanced changes in abundance vs. abundance gradients, *Methods in Ecology and Evolution* **4**: 552–557.

Bray, J. and Curtis, J. (1957). An ordination of upland forest communities of southern Wisconsin, *Ecological Monographs* **27**: 325–349.

Brown, J. H. (1981). Two decades of homage to santa rosalia: Toward a general theory of diversity, *American Zoologist* **21**(4): 887–888.

Butchart, S. H. M., Walpole, M., Collen, B., van Strien, A., Scharlemann, J. P. W., Almond, R. E. A., Baillie, J. E. M., Bomhard, B., Brown, C., Bruno, J., Carpenter, K. E., Carr, G. M., Chanson, J., Chenery, A. M., Csirke, J., Davidson, N. C., Dentener, F., Foster, M., Galli, A., Galloway, J. N., Genovesi, P., Gregory, R. D., Hockings, M., Kapos, V.,

Lamarque, J.-F., Leverington, F., Loh, J., McGeoch, M. A., McRae, L., Minasyan, A., Morcillo, M. H., Oldfield, T. E. E., Pauly, D., Quader, S., Revenga, C., Sauer, J. R., Skolnik, B., Spear, D., Stanwell-Smith, D., Stuart, S. N., Symes, A., Tierney, M., Tyrrell, T. D., Vié, J.-C. and Watson, R. (2010). Global biodiversity: Indicators of recent declines, *Science* **328**(5982): 1164–1168.

Cardoso, P., Rigal, F. and Carvalho, J. C. (2014). BAT — biodiversity assessment tools, an R package for the measurement and estimation of alpha and beta taxon, phylogenetic and functional diversity, *Methods in Ecology and Evolution*.

Carvalho, J. C., Cardoso, P., Borges, P. A. V., Schmera, D. and Podani, J. (2013). Measuring fractions of beta diversity and their relationships to nestedness: a theoretical and empirical comparison of novel approaches, *Oikos* **122**(6): 825–834.

Chao, A., Chiu, C. and Hsieh, T. C. (2012). Proposing a resolution to debates on diversity partitioning, *Ecology* **93**: 2037–2051.

Cichocki, A. and Amari, S. (2010). Families of Alpha- Beta- and Gamma- divergences: Flexible and robust measures of similarities, *Entropy* **12**: 1532–1568.

Dornelas, M. (2010). Disturbance and change in biodiversity, *Philosophical Transactions B* **365**: 3719–3727.

Dornelas, M., Gotelli, N. J., McGill, B., Shimadzu, H., Moyes, F., Sievers, C. and Magurran, A. E. (2014). Assemblage time series reveal biodiversity change but not systematic loss, *Science* **344**(6181): 296–299.

Dornelas, M., Magurran, A. E., Buckland, S. T., Chao, A., Chazdon, R. L., Colwell, R. K., Curtis, T., Gaston, K. J., Gotelli, N. J., Kosnik, M. A., McGill, B., McCune, J. L., Morlon, H., Mumby, P. J., Øvreås, L., Studeny, A. and Vellend, M. (2013). Quantifying temporal change in biodiversity: challenges and opportunities, *Proceedings of the Royal Society B: Biological Sciences* **280**(1750): 20121931.

Hastie, T. J. and Tibshirani, R. J. (1990). *Generalized Additive Models*, Chapman & Hall/CRC.

Henderson, P. A. (2007). Discrete and continuous change in the fish community of the Bristol Channel in response to climate change, *Journal of the Marine Biological Association of the UK* **87**: 589–598.

Henderson, P. A. and Bird, D. J. (2010). Fish and macro-crustacean communities and their dynamics in the severn estuary, *Marine Pollution Bulletin* **61**(1–3): 100–114.

Henderson, P. A. and Holmes, R. H. A. (1991). On the population dynamics of dab, sole and flounder within bridgwater bay in the lower severn estuary, England, *Netherlands Journal of Sea Research* **27**: 337–344.

Henderson, P. A. and Seaby, R. M. (2005). The role of climate in determining the temporal variation in abundance, recruitment and growth of sole, *solea solea* in the Bristol Channel, *Journal of the Marine Biological Association of the UK* **85**: 197–204.

Henderson, P., Seaby, R. M. H. and Somes, J. R. (2011). Community level response to climate change: the long-term study of the fish and crustacean community of the Bristol Channel, *Journal of Experimental Marine Biology and Ecology* **400**: 78–89.

Hilbe, J. M. (2011). *Negative Binomial Regression*, 2nd edn, Cambridge University Press.

Hill, M. O. (1973). Diversity and evenness: A unifying notation and its consequences, *Ecology* **54**(2): 427–432.

Hubbell, S. P. (2001). *The Unified Neutral Theory of Biodiversity and Biogeography*, Princeton University Press.

Ives, A., Dennis, B., Cottingham, K. and Carpenter, S. (2003). Estimating community stability and ecological interactions from the time-series data, *Ecological Monographs* **73**: 301–330.

Jaccard, P. (1901). Distribution de la flore alpine dans le bassin des Dranses et dans quelques régions voisines, *Bulletin de la Société Vaudoise des Sciences Naturelles* **37**: 241–272.

Jaccard, P. (1912). The distribution of the flora in the alpine zone, *New Phytologist* **11**: 37–50.

Jost, L. (2007). Partitioning diversity into independent alpha and beta components, *Ecology* **88**(10): 2427–2439.

Jost, L., Chao, A. and Chazdon, R. L. (2011). Compositional similarity and β (beta) diversity, *Biological Diversity*, Oxford University Press.

Kawada, Y. (1987). Information and statistics, *Proceedings of the Institute of Statistical Mathematics* **35**(1): 1–57.

Kedem, B. and Fokianos, K. (2002). *Regression Models for Time Series Analysis*, Wiley.

Kullback, S. and Leibler, R. (1951). On information and sufficiency, *The Annals of Mathematical Statistics* **22**(1): 79–86.

Lande, R. (1996). Statistics and partitioning of speciesd iversity and similarity among multiple communities, *Oikos* **76**: 5–13.

Legendre, P. (2014). Interpreting the replacement and richness difference components of beta diversity, *Global Ecology and Biogeography* **23**: 1324–1334.

Legendre, P. and De Cáceres, M. (2013). Beta diversity as the variance of community data: dissimilarity coefficients and partitioning, *Ecology Letters* **16**(8): 951–963.

Legendre, P. and Gauthier, O. (2014). Statistical methods for temporal and space–time analysis of community composition data, *Proceedings of the Royal Society of London B: Biological Sciences* **281**: 20132728.

Magurran, A. E. and Henderson, P. A. (2003). Explaining the excess of rare species in natural species abundance distributions, *Nature* **422**: 714–716.

Malthus, T. R. (1798). *An Essay on the Principle of Population*, J. Johnson, London.

Marcon, E., Hérault, B., Baraloto, C. and Lang (2012). The decomposition of Shannon's entropy and a confidence interval for beta diversity, *Oikos* **121**: 516–522.

Margalef, R. (1957). La teoría de la información en ecología, *Memoirs of the Real Academia de Ciencias y Artes de Barcelona* **32**: 373–449.

Margalef, R. (1958). Information theory in ecology, *General Systems* **3**: 36–71.

McCullagh, P. and Nelder, J. (1989). *Generalized Linear Models*, 2nd edn, Chapman and Hall/CRC.

Mutshinda, C. M., O'Hara, R. B. and Woiwod, I. P. (2009). What drives community dynamics?, *Proceedings of the Royal Society of London B: Biological Sciences* **276**: 2923–2929.

Mutshinda, C. M., O'Hara, R. B. and Woiwod, I. P. (2011). A multispecies perspective on ecological impacts of climatic forcing, *Journal of Animal Ecology* **80**(1): 101–107.

Oksanen, J., Blanchet, F. G., Kindt, R., Legendre, P., Minchin, P. R., O'Hara, R. B., Simpson, G. L., Solymos, P., Stevens, M. H. H. and Wagner, H. (2014). *vegan: Community Ecology Package*. R package version 2.2-0.

Podani, J. and Schmera, D. (2011). A new conceptual and methodological framework for exploring and explaining pattern in presence-absence data, *Oikos* **120**(11): 1625–1638.

Podani, J., Ricotta, C. and Schmera, D. (2013). A general framework for analyzing beta diversity, nestedness and related community-level phenomena based on abundance data, *Ecological Complexity* **15**: 52–61.

Read, T. and Cressie, N. (1988). *Goodness-of-Fit Statistics for Discrete Multivariate Data*, Springer-Verlag, New York.

Reeve, R., Matthews, L., Cobbold, C., Leinster, T., Thompson, J. and Brummitt, N. (2014). How to partition diversity, *ArXiv e-prints*.

Rényi, A. (1961). On measures of entropy and information, *Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability* **1**: 547–561.

Ruzicka, M. (1958). Awndung matematisch-statishticher methoden in der geobotanik (syshetische bearbutung von autnahme), *Biologia, Bratislava* **13**: 647–661.

Sæther, B.-E., Engen, S. and Grøtan, V. (2013). Species diversity and community similarity in fluctuating environments: parametric approaches using species abundance distributions, *Journal of Animal Ecology* **82**(4): 721–738.

Shimadzu, H., Dornelas, M., Henderson, P. A. and Magurran, A. E. (2013). Diversity is maintained by seasonal variation in species abundance, *BMC Biology* **11**: 98.

Sørensen, T. (1948). A method of establishing groups of equal amplitude in plant sociology based on similarlity of species content and its application to analyses of the vegetation on Danish commons, *Biologiske Skrifter* **5**: 1–34.

Supp, S. R. and Ernest, S. K. M. (2014). Species-level and community-level responses to disturbance: a cross-community analysis, *Ecology* **95**: 1717–1723.

Vellend, M., Baeten, L., Myers-Smith, I. H., Elmendorf, S. C., Beauséjour, R., Brown, C. D., De Frenne, P., Verheyen, K. and Wipf, S. (2013). Global meta-analysis reveals no net change in local-scale plant biodiversity over time, *Proceeding of the National Academy of Science of the United States America* **110**(48): 19456–19459.

Vellend, M., Cornwell, W. K., Magnuson-Ford, K. and Mooers, A. Ø. (2011). Measuring phylogenetic biodiversity, *Biological Diversity*, Oxford University Press.

Whittaker, R. H. (1960). Vegetation of the Siskiyou mountains, Oregon and California, *Ecological Monographs* **30**(3): 279–338.

Whittaker, R. H. (1972). Evolution and measurement of species diversity, *Taxon* **21**(2): 213–251.

Zuur, A., Ieno, E., Walker, N., Saveliev, A. and Smith, G. (2009). *Mixed Effects Models and Extensions in Ecology with R*, Springer-Verlag.

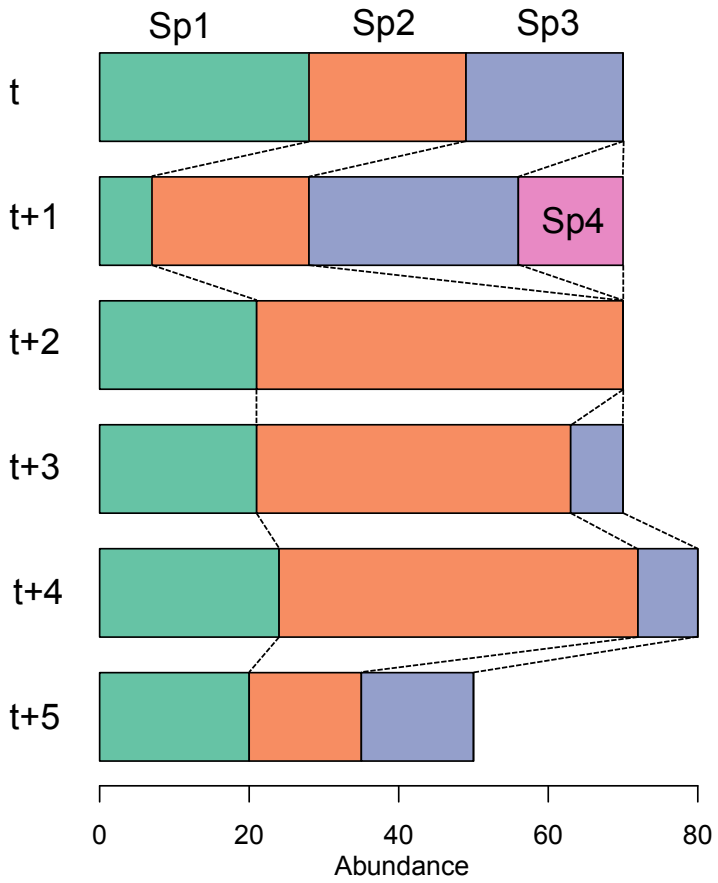

Figure 1: A schematic figure of the framework for measuring temporal turnover with four species (Sp1, Sp2, Sp3 and Sp4) in a community at times $t, t+1, \ldots, t+5$. For clarity, the x-axis represents observed-abundance, but note the fact that zero observed-abundance does not necessarily mean zero expected-abundance in the framework (for details see Section 3). The proposed turnover measure $D = D_1 + D_2$ (Equation 4) quantifies the turnover of the species community due to two key ecological processes: change in community composition (D_1) and change in total abundance (D_2). For times t to $t+3$ — a new species Sp4 comes in at $t+1$, both Sp3 and Sp4 drop out at $t+2$ and only Sp3 comes back at $t+3$ — the turnover is quantified by D_1 as there are changes in community composition but not in total abundance ($D_2 = 0$). Between times $t+3$ and $t+4$, the situation is opposite: there are changes in total abundance but not in community composition. Hence, the turnover of this period is quantified by D_2 as $D_1 = 0$. Both community composition and total abundance change between $t+4$ and $t+5$, the turnover of which is quantified by D_1 and D_2 both.

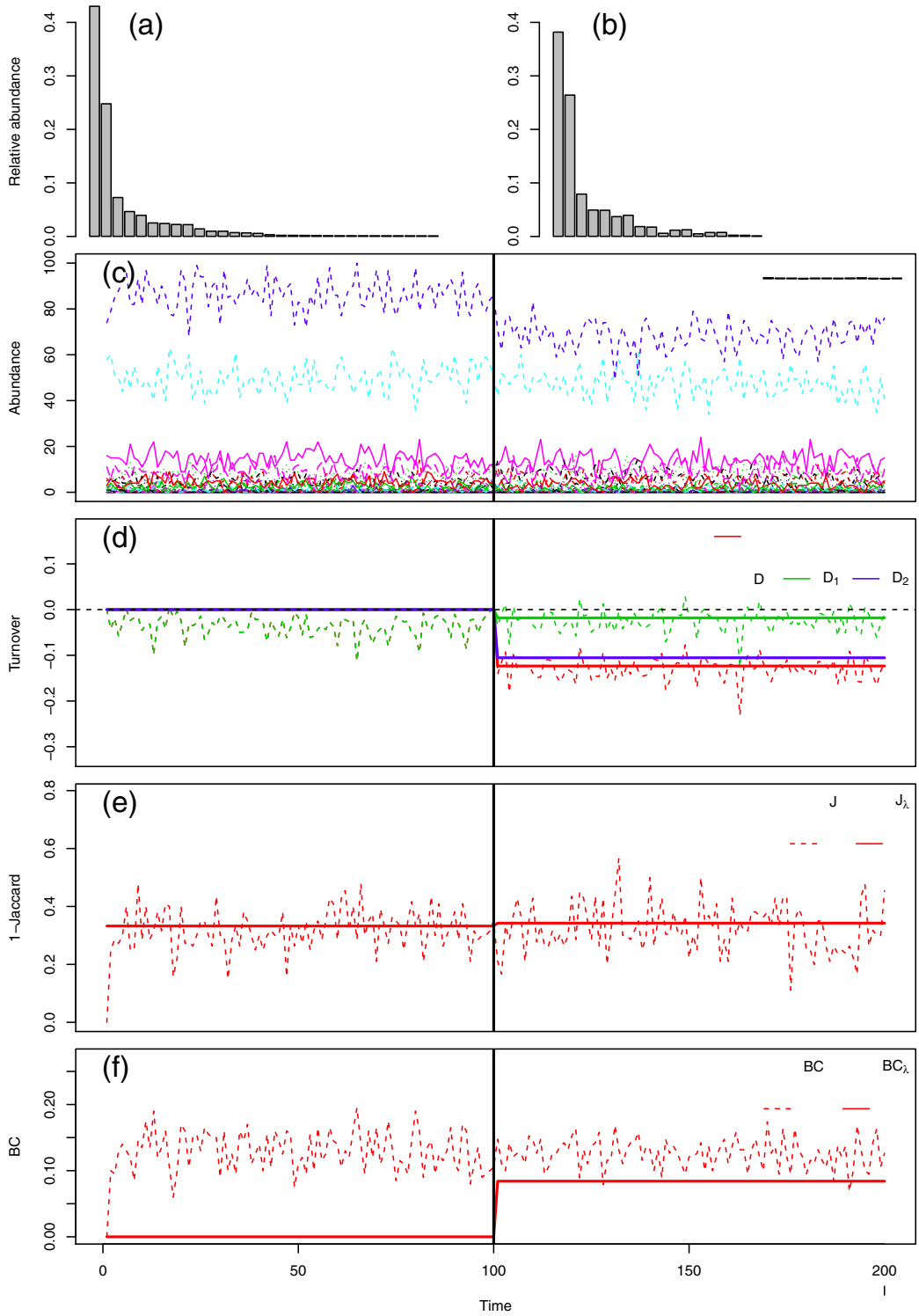

Figure 2: Simulation results of the scenario with a decrease ($n = 200 > n' = 180$) and alteration ($\boldsymbol{p} \neq \boldsymbol{p}'$). All turnover measures are calculated against the first time point. The top row shows the relative abundances before (a) and after (b) the artificial change introduced at $t = 100$; (c) the simulated respective abundance series over 200 time steps; (d) the new turnover measure, D (red), D_1 (green) representing the change in the community composition and D_2 (blue) representing the change in the community abundance. The solid lines are calculated based on the model (Equation 9), and the dashed lines are based on the simulated data; (e) the outcome of Jaccard index, $1 - J$ (dashed line, simulated data based) and $1 - J_\lambda$ (solid line, model based); (f) the outcome of Bray–Curtis index, BC (dashed line, simulated data based) and BC_λ (solid line, model based).

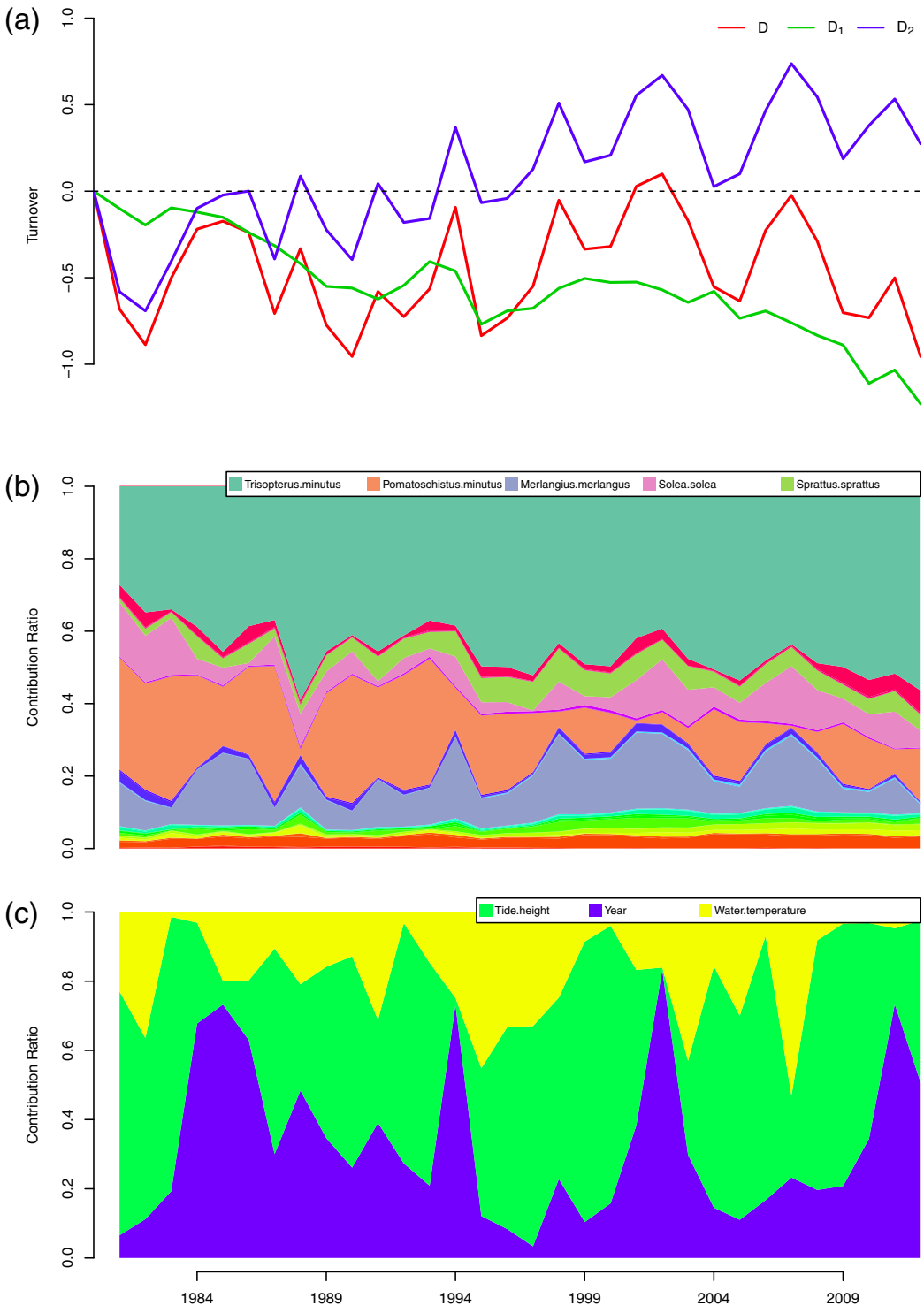

Figure 3: Illustrative results of the turnover analysis on the fish community at Hinkley Point in August over the period (1981 – 2012). Each year is compared against the first observation, August 1981. (a) The turnover measure (D: the red line) and its components, the composition change (D_1: the green line) and the community size change (D_2: the blue line); (b) The contribution ratios of each species in the community. Top five species with high contribution ratio are listed in the legend, and other colour/shading represent(s) the rest of 76 species; (c) the contribution ratios of each environmental factors: year (blue), tide height (green) and water temperature (yellow)..

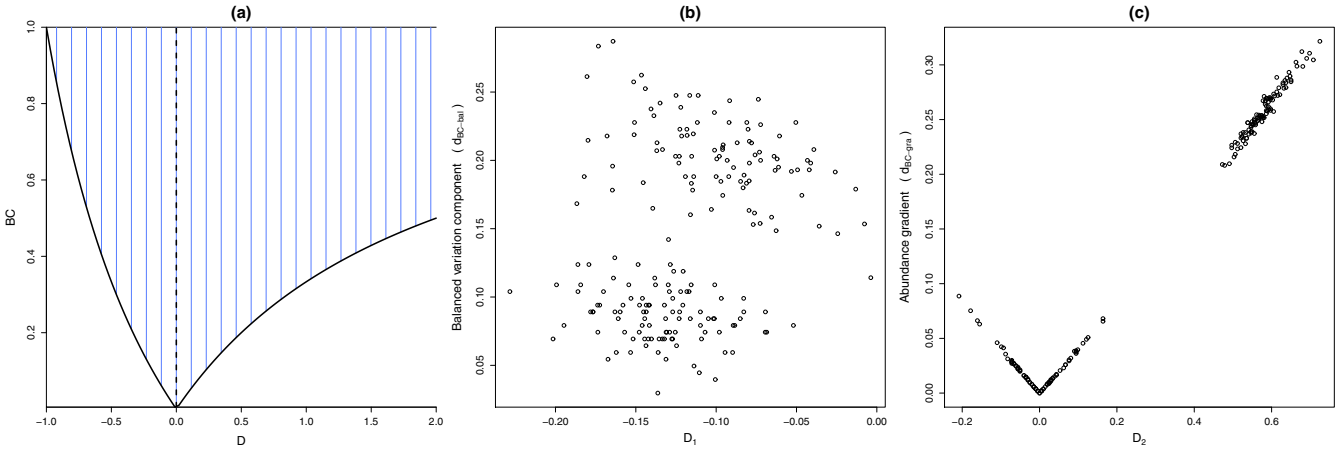

Figure 4: (a) Functional relationship of the turnover measure, D, and Bray–Curtis index, BC. The vertical striped area, $\{(D, BC) : \sum_{i=1}^{s} |d_i|/(2+D)\}$, is the values of the Bray–Curtis index can take, given the value of the turnover measure. The boundary (the black solid line) represents $BC = |D|/(2 + D)$ as $|D| \leq \sum_{i=1}^{s} |d_i|$; (b) The relationship between the balanced variation in abundance, d_{BC-bal}, and the change in the species composition, D_1; (c) The relationship between the abundance gradient, d_{BC-gra}, and the the change in the community abundance, D_2.

Appendix to: Measuring temporal turnover in ecological communities

Hideyasu Shimadzu Maria Dornelas Anne E. Magurran

Centre for Biological Diversity and Scottish Oceans Institute, University of St Andrews, UK

A The derivation of Equations (11) and (12)

Putting $\lambda_i(u) = x_{ij}$ and $\lambda_i(t) = x_{ik}$, we have eqn 1 – 3 in Baselga (2013) as

$$\begin{aligned} A &= \sum_{i=1}^{s} \min(\lambda_i(u), \lambda_i(t)) \\ B &= \sum_{i=1}^{s} \lambda_i(u) - \min(\lambda_i(u), \lambda_i(t)) = \sum_{i=1}^{s} (\lambda_i(u) - \lambda_i(t))_+ \\ C &= \sum_{i=1}^{s} \lambda_i(t) - \min(\lambda_i(u), \lambda_i(t)) = -\sum_{i=1}^{s} (\lambda_i(u) - \lambda_i(t))_- \end{aligned}$$

where $(\cdot)_+$ and $(\cdot)_-$ respectively denotes its positive or negative part. The Bray–Curtis index is then described as

$$\begin{aligned} d_{BC} &= \frac{\sum_{i=1}^{s} |\lambda_i(u) - \lambda_i(t)|}{\lambda_i(u) + \lambda_i(t)} = \frac{B + C}{2A + B + C} \\ &= \frac{\min(B, C)}{A + \min(B, C)} + \frac{|B - C|}{2A + B + C} \frac{A}{A + \min(B, C)} \\ &= d_{BC-bal} + d_{BC-gra}. \end{aligned}$$

The balanced variation for each case is respectively given as

Case I: $\min(B, C) = B$ $(B < C \iff D < 0)$

$$d_{BC-bal} = \frac{B}{A + B} = \frac{\sum_{i=1}^{s} (\lambda_i(u) - \lambda_i(t))_+}{\lambda(u)} = \frac{\sum_{i=1}^{s} (\lambda_i(u) - \lambda_i(t))_+}{\lambda(t)} \frac{\lambda(t)}{\lambda(u)} \doteq D^+ e^{-D_2},$$

Case II: $\min(B, C) = C$ $(C < B \iff D > 0)$

$$d_{BC-bal} = \frac{C}{A + C} = -\frac{\sum_{i=1}^{s} (\lambda_i(u) - \lambda_i(t))_-}{\lambda(t)} \doteq -D^-.$$

The abundance gradient for each case is respectively given as

Case I: $\min(B, C) = B$ $(B < C \Longleftrightarrow D < 0)$

$$\begin{aligned} d_{BC-gra} = \frac{|B-C|}{2A+B+C}\frac{A}{A+B} &= \frac{|\lambda(u)-\lambda(t)|}{\lambda(u)+\lambda(t)}\frac{\sum_{i=1}^{s}\min(\lambda_i(u),\lambda_i(t))}{\lambda(u)} \\ &= \frac{|\lambda(u)-\lambda(t)|}{\lambda(t)}\frac{\lambda(t)}{\lambda(u)}\frac{\sum_{i=1}^{s}\min(\lambda_i(u),\lambda_i(t))}{\lambda(u)+\lambda(t)} \\ &\doteq |D|e^{-D_2}\frac{q}{2+D}, \end{aligned}$$

Case II: $\min(B, C) = C$ $(C < B \Longleftrightarrow D > 0)$

$$\begin{aligned} d_{BC-gra} = \frac{B-C}{2A+B+C}\frac{A}{A+C} &= \frac{\lambda(u)-\lambda(t)}{\lambda(u)+\lambda(t)}\frac{\sum_{i=1}^{s}\min(\lambda_i(u),\lambda_i(t))}{\lambda(t)} \\ &\doteq D\frac{q}{2+D}. \end{aligned}$$

B Model-based Jaccard and Bray–Curtis indices

For the simulation study, we calculate two common indices: Jaccard (Jaccard, 1901) and Bray–Curtis (Bray and Curtis, 1957) indices for comparison. As these indices are originally defined on observations, $\{n_i(t)\}$, we introduce a model-based version of Jaccard and Bray–Curtis indices as follow.

Jaccard index (J)

$$J = \frac{\sum_{i=1}^{s} I(n_i(t) > 0)I(n_i(u) > 0)}{\sum_{i=1}^{s} I(n_i(t) + n_i(u) > 0)}$$

Model-based Jaccard index (J_λ)

$$J_\lambda = \frac{\sum_{i=1}^{s} P(N_i(t) > 0)P(N_i(u) > 0)}{\sum_{i=1}^{s} P(N_i(t) + N_i(u) > 0)} = \frac{\sum_{i=1}^{s}(1-\exp(-\lambda_i(t)))(1-\exp(-\lambda_i(u)))}{s - \sum_{i=1}^{s}\exp(-\lambda_i(t)-\lambda_i(u))}$$

Bray–Curtis index (BC)

$$BC = \frac{\sum_{i=1}^{s}|n_i(t) - n_i(u)|}{n(t) + n(u)},$$

where $n(t) = \sum_{i=1}^{s} n_i(t)$.

Model-based Bray–Curtis index (BC_λ)

$$BC_\lambda = \frac{\sum_{i=1}^{s}|\lambda_i(t) - \lambda_i(u)|}{\lambda(t) + \lambda(u)},$$

where $\lambda(t) = \sum_{i=1}^{s} \lambda_i(t)$.

Podani et al. (2013) note a link between Jaccard and Ruzicka (Ruzicka, 1958) indices, taking $n_i(t) = 1$ for species presence and $n_i(t) = 0$ for its absence. The same link also exists between Sørensen and Bray–Curtis indices, but not between those, Jaccard and Bray–Curtis indices that we have used. This type of continuity is not crucial in our study since the relative ordering is, whether Jaccard or Sørensen (say L) index is used, retained because of their monotonic relationship, $L = 2J/(J + 1)$.

C Simulation results

We have studied the five simulation scenarios with the three states in changes of community size: stable ($n = n'$), an increase ($n < n'$) and a decrease ($n > n'$); and of community composition: no-alteration ($\boldsymbol{p} = \boldsymbol{p}'$) and alteration ($\boldsymbol{p} \neq \boldsymbol{p}'$), when an artificial change taking a place at time $t = 100$. The scenarios are specified by a combination of those states, but we have omitted the trivial case, neutral and no-alteration ($n = n'$ and $\boldsymbol{p} = \boldsymbol{p}'$), so that there are five scenarios to be examined. The scenario 3 is presented in the main text.

Scenario	Abundance change	Composition change
1	$n > n'$	$\boldsymbol{p} = \boldsymbol{p}'$
—	$n = n'$	$\boldsymbol{p} = \boldsymbol{p}'$
2	$n < n'$	$\boldsymbol{p} = \boldsymbol{p}'$
3	$n > n'$	$\boldsymbol{p} \neq \boldsymbol{p}'$
4	$n = n'$	$\boldsymbol{p} \neq \boldsymbol{p}'$
5	$n < n'$	$\boldsymbol{p} \neq \boldsymbol{p}'$

For example, the interpretation of Scenario 1 here is that the total community size (abundance) decreases n to n' by an artificial change taking a place at time $t = 100$, but the community composition, the relative species abundances $\boldsymbol{p}$, stays as the same ($\boldsymbol{p} = \boldsymbol{p}'$).

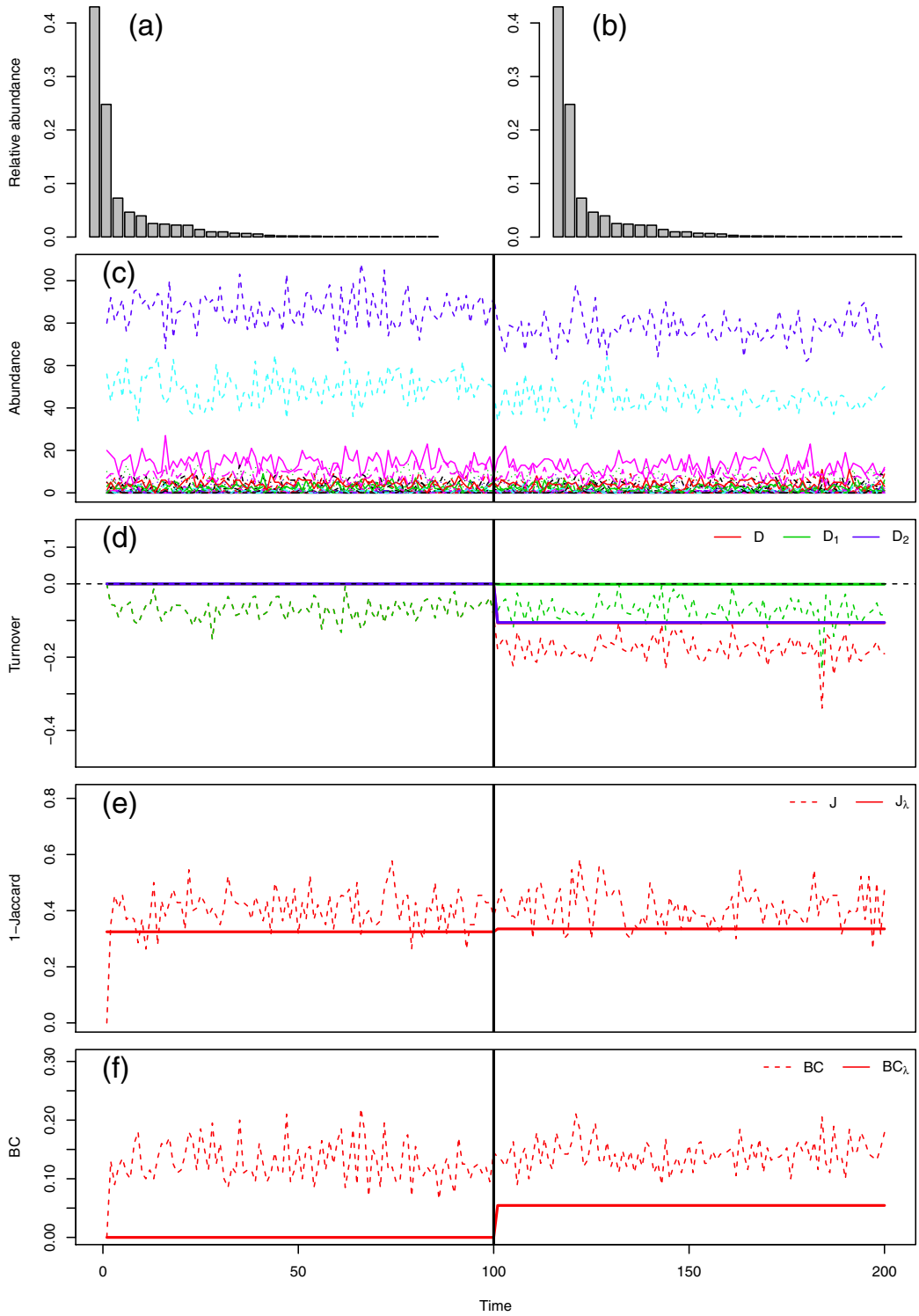

Figure A1: Simulation results of the scenario with a decrease ($n = 200 > n' = 180$) and no-alteration ($\boldsymbol{p} = \boldsymbol{p}'$). All turnover measures are calculated against the first time point. The top row shows the relative abundances before (a) and after (b) the artificial change introduced at $t = 100$; (c) the simulated respective abundance series over 200 time steps; (d) the new turnover measure, D (red), D_1 (green) representing the change in the community composition and D_2 (blue) representing the change in the community abundance. The solid lines are calculated based on the model (Equation 9), and the dashed lines are based on the simulated data; (e) the outcome of Jaccard index, $1 - J$ (dashed line, simulated data based) and $1 - J_\lambda$ (solid line, model based); (f) the outcome of Bray–Curtis index, BC (dashed line, simulated data based) and BC_λ (solid line, model based).

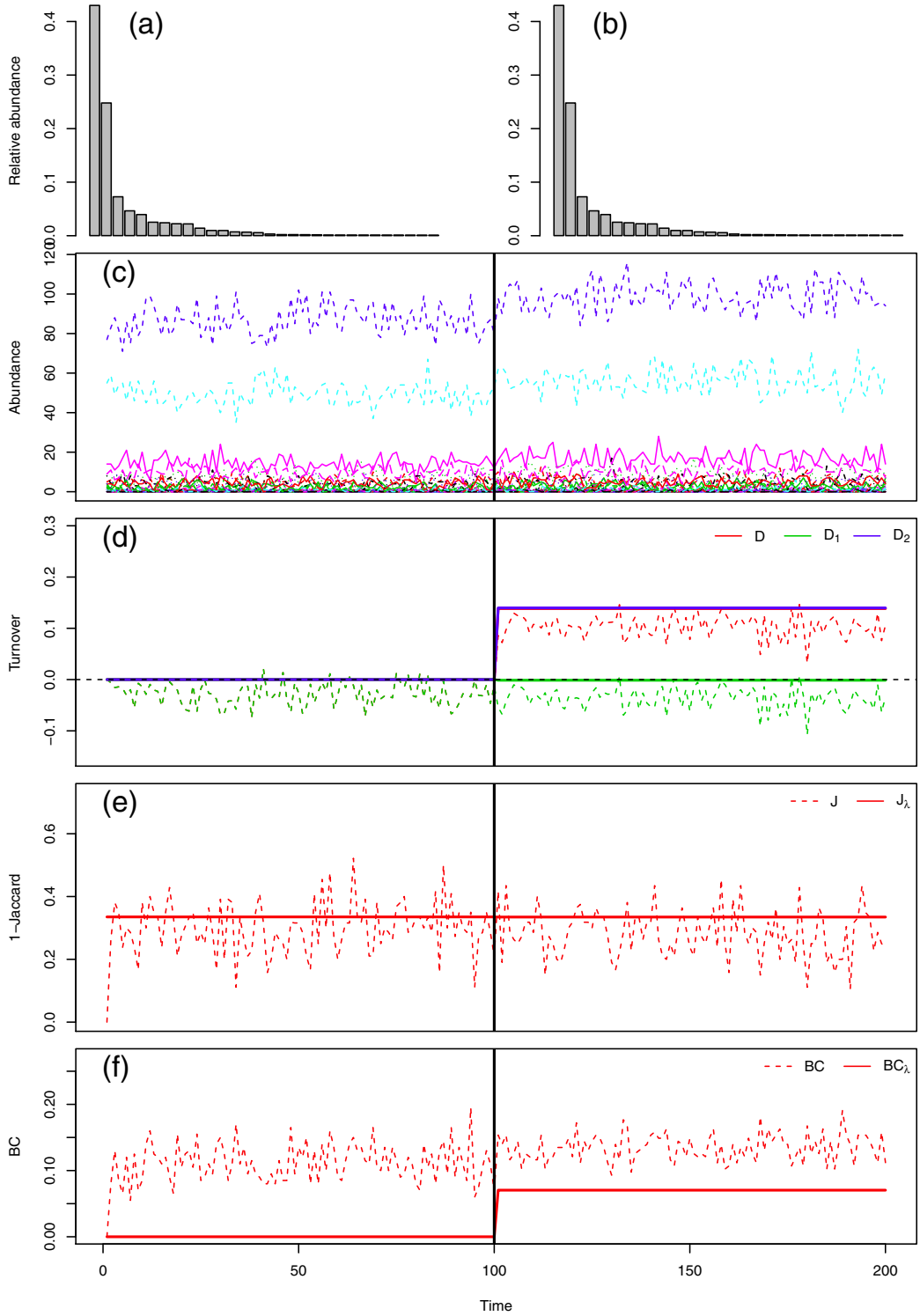

Figure A2: Simulation results of the scenario with an increase ($n = 200 < n' = 230$) and no-alteration ($\boldsymbol{p} = \boldsymbol{p}'$). All turnover measures are calculated against the first time point. The top row shows the relative abundances before (a) and after (b) the artificial change introduced at $t = 100$; (c) the simulated respective abundance series over 200 time steps; (d) the new turnover measure, D (red), D_1 (green) representing the change in the community composition and D_2 (blue) representing the change in the community abundance. The solid lines are calculated based on the model (Equation 9), and the dashed lines are based on the simulated data; (e) the outcome of Jaccard index, $1 - J$ (dashed line, simulated data based) and $1 - J_\lambda$ (solid line, model based); (f) the outcome of Bray–Curtis index, BC (dashed line, simulated data based) and BC_λ (solid line, model based).

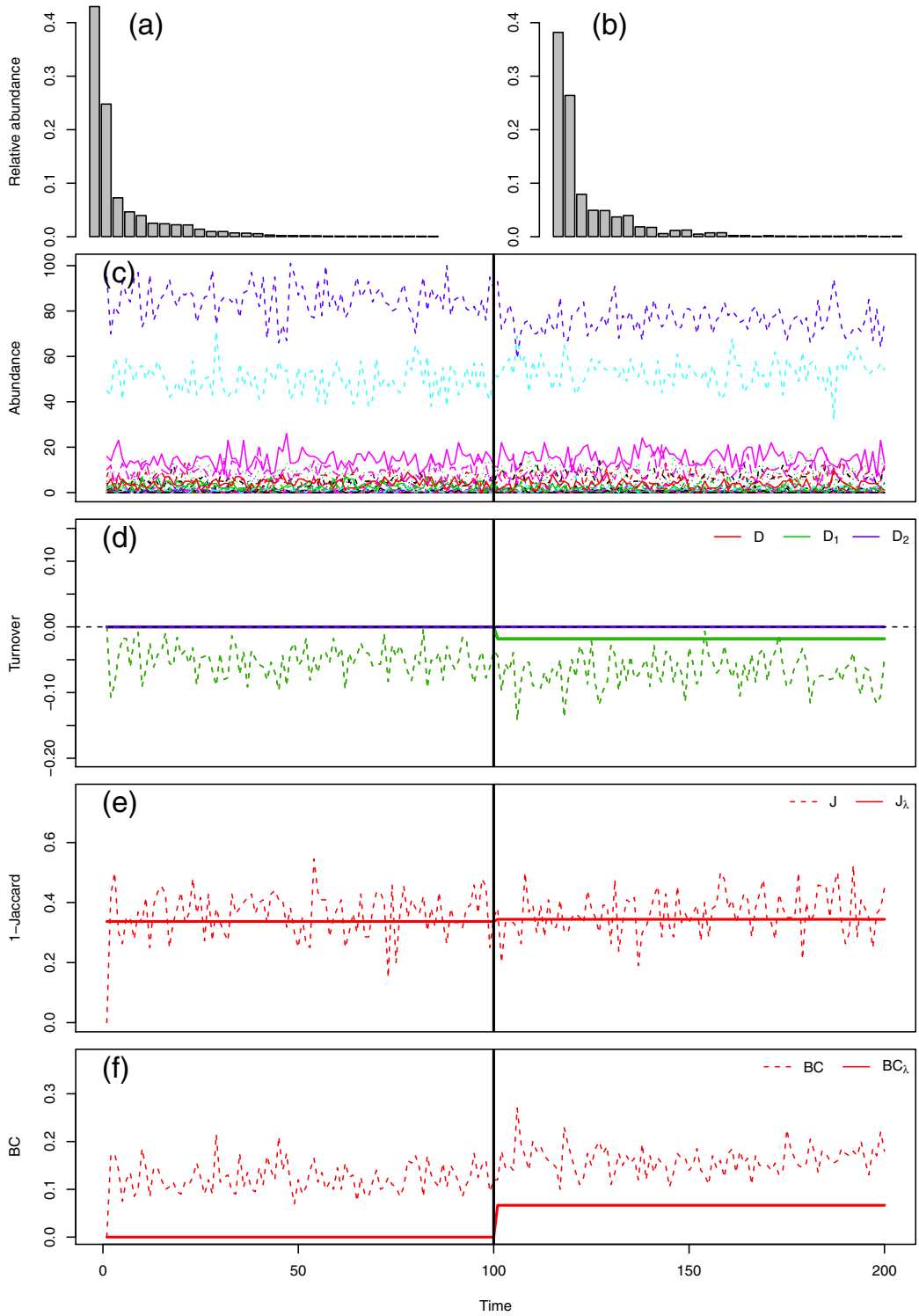

Figure A3: Simulation results of the scenario with being stable ($n = n' = 200$) and alteration ($\boldsymbol{p} \neq \boldsymbol{p}'$). All turnover measures are calculated against the first time point. The top row shows the relative abundances before (a) and after (b) the artificial change introduced at $t = 100$; (c) the simulated respective abundance series over 200 time steps; (d) the new turnover measure, D (red), D_1 (green) representing the change in the community composition and D_2 (blue) representing the change in the community abundance. The solid lines are calculated based on the model (Equation 9), and the dashed lines are based on the simulated data; (e) the outcome of Jaccard index, $1 - J$ (dashed line, simulated data based) and $1 - J_\lambda$ (solid line, model based); (f) the outcome of Bray–Curtis index, BC (dashed line, simulated data based) and BC_λ (solid line, model based).

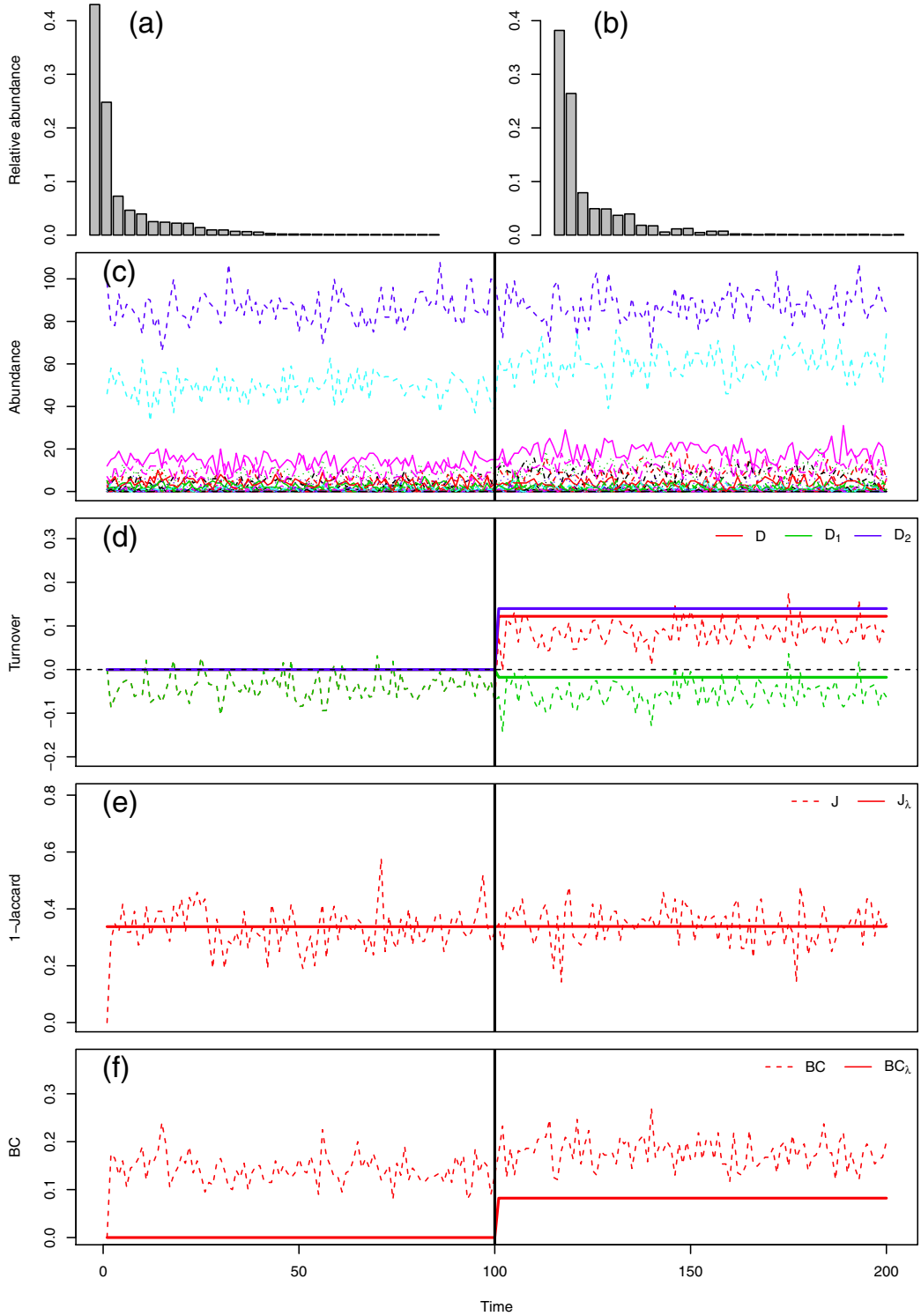

Figure A4: Simulation results of the scenario with an increase ($n = 200 < n' = 230$) and alteration ($\boldsymbol{p} \neq \boldsymbol{p}'$). All turnover measures are calculated against the first time point. The top row shows the relative abundances before (a) and after (b) the artificial change introduced at $t = 100$; (c) the simulated respective abundance series over 200 time steps; (d) the new turnover measure, D (red), D_1 (green) representing the change in the community composition and D_2 (blue) representing the change in the community abundance. The solid lines are calculated based on the model (Equation 9), and the dashed lines are based on the simulated data; (e) the outcome of Jaccard index, $1 - J$ (dashed line, simulated data based) and $1 - J_\lambda$ (solid line, model based); (f) the outcome of Bray–Curtis index, BC (dashed line, simulated data based) and BC_λ (solid line, model based).

D Some R codes

We provide some R codes to calculate our turnover measure and to produce the some figures presented in the main text.

DD

Calculates the turnover measure, D, D_1 and D_2 against any reference time point (the default is the first time point).

Usage

```
DD(x)
DD(x, ref.t = 1, zero.rm = FALSE)
```

Arguments

`x`	a data frame (t, i) that contains the estimated expected abundances, $\hat{\lambda}_i(t)$.
`ref.t`	reference time point, t, the turnover measure is calculated against. The default value is `ref.t=1`, the first time point.
`zero.rm`	a logical value indicating whether $p_i(u) = 0$ should be stripped before the computation proceeds. The default value is `zero.rm=FALSE`. To avoid any `NaN` produced when observations are used as the estimated expected abundances, try `zero.rm=TRUE`.

Values

A data frame that contains the turnover measure D, D_1 and D_2 in each column.

Example

```
results <- list()
for(1 in 1:n) results[[i]] <- glm(y ~ x1 + x2, family="poisson")
x <- do.call(cbind, lapply(results, predict, type="response"))
DD(x)
```

Code

```
DD <- function(x, ref.t=1, zero.rm=FALSE){
        lmb <- apply(x, 1, sum)
        D2 <- log(lmb/lmb[ref.t])

       x.p <- t(apply(x, 1, function(z)z/sum(z)))
       Pt <- x.p[ref.t,]
       if(zero.rm==FALSE){
         D1 <- -t(apply(x.p, 1, function(z)ifelse(Pt==0, 0, log(Pt/z)))) %*% Pt
       }else{
         D1 <- -t(apply(x.p, 1, function(z)ifelse(Pt==0|z==0, 0, log(Pt/z)))) %*% Pt
       }

       D <- D1 + D2
       data.frame(D, D1, D2)
      }
```

cr.cal

Calculates the contribution ratios of the species $r_i(t : u)$ or the contribution ratios of the environment factors $r_j(t : u)$ from a list object, each branch of which must contain a GLM or a GAM object.

Usage

```
cr.cal(x)
cr.cal(x, id=1, ref.t=1)
```

Arguments

`x`	a list contains a GLM or a GAM output.
`id`	a numerical value indicating whether the contribution ratios of the species, $r_i(t : u)$, (`id=1`) or the contribution rations of the environmental factors, $r_j(t : u)$, (`id=2`) to be calculated. The default value is `id=1`.
`ref.t`	reference time point, t, the turnover measure is calculated against. The default value is `ref.t=1`, the first time point.

Values

A data frame containing the cumulative contribution ratios that can be passed to the function `cr.plot` to draw a contribution ratio diagram.

Example

```
x <- list()
for(1 in 1:n) x[[i]] <- glm(y ~ x1 + x2, family="poisson")
cr.cal(x)
```

Code

```
cr.cal <- function(x, id=1, ref.t=1){
  lmb <- do.call(cbind, lapply(x, predict, type="response"))
  lmb.p <- t(apply(lmb, 1, function(z)z/sum(z)))

  lmb <- lapply(x, function(z)as.matrix(predict(z, type="terms")))
  lmb <- lapply(lmb, function(z)t(t(z)-z[ref.t,]))
  lmb <- array(unlist(lmb), dim=c(nrow(lmb[[1]]), ncol(lmb[[1]]), length(lmb)))

  if(id==1){
    absDi <- abs(apply(apply(lmb, c(1,3), sum), 1, "*", lmb.p[ref.t,])) # sp
  }else{
    absDi <- t(abs(apply(lmb, c(1,2), "%*%", lmb.p[ref.t,]))) # env
  }

  ri <- apply(absDi, 2, function(z)z/sum(z))
  ri.cum <- t(apply(ri, 2, cumsum))

  data.frame(ri.cum)
}
```

cr.plot

Plots a contribution ratio diagram (eg. Figures 3b and 3c).

Usage

```
cr.plot(Y)
cr.plot(Y, x=NULL, col.pal=NULL, k=5, ...)
```

Arguments

`Y`	a data frame (t, i) or (t, j) that contains the contribution ratios of each species, $r_i(t : u)$, or each environmental factor, $r_j(t : u)$.
`x`	the coordinates of points in the plot.
`col.pal`	the colour palette to be used. The length needs to be the same as `ncol(Y)`. The default choice is `rainbow`.
`k`	the number of species or environment factors to be listed in the legend of the figure. The default value is `k=5`.
`...`	further arguments passed to or from other methods.

Example

```
Y <- cr.cal(x)
names(Y) <- col.names
cr.plot(Y)
```

Code

```
cr.plot <- function(Y, x=NULL, col.pal=NULL, k=5, ...){
  k <- min(ncol(Y), k)
  if(is.null(col.pal)){col.pal <- rainbow(ncol(Y))}
  Y <- data.frame(0, Y)
  if(is.null(x)) x <- c(1:nrow(Y))

  matplot(x, Y, type="n", ylab="Contribution Ratio", ...)
  for(j in 2:ncol(Y)){
    polygon(c(x, rev(x)), c(Y[,j-1], rev(Y[,j])), col=col.pal[j-1], border=NA, ...)
    }

  lgd.list <- apply(apply(Y, 1, diff, na.rm=T), 1, sum, na.rm=T)
  legend("topright", legend=names(rev(sort(lgd.list)))[1:k],
         col=col.pal[order(lgd.list, decreasing=T)][1:k],
         pch=15, pt.cex=2, bg="white", ...)
}
```