MR. MICHAEL KALYUZHNY (Orcid ID : 0000-0002-5574-0079)

Article type   : Research Article

Editor   : Dr. Steven Kembel

Dissimilarity-Overlap analysis of community dynamics: opportunities and pitfalls

Running title: DOC analysis of community dynamics

Authors: Michael Kalyuzhny[1], Nadav M. Shnerb[2]

Affiliations: 1. Department of Ecology, Evolution and Behavior, Institute of Life Sciences, Hebrew University of Jerusalem, Givat-Ram, Jerusalem 91904, Israel.

2. Department of Physics, Bar-Ilan University, Ramat Gan 52900. Israel.

Corresponding author: Michael Kalyuzhny (michael.kalyuzhny@mail.huji.ac.il)

Department of Evolution, Ecology and Behavior, Institute of Life Sciences,

Hebrew University of Jerusalem, Givat-Ram, Jerusalem 91904, Israel.

Phone number: +972-(0)50-8683886

Abstract

1. Do multiple different communities have similar dynamics? This novel question was recently addressed with the development of the Dissimilarity-Overlap Curve (DOC) method, designed to identify sets of communities with similar ("universal") dynamics using pairwise comparison of communities' composition. However, the theoretical properties of the method were tested using very restrictive and, in some cases, unrealistic assumptions.

2. We test the ability of the DOC method to discriminate universal from non-universal dynamics in more realistic scenarios - under various levels and regimes of stochasticity and sampling errors, in cases where species respond to environmental gradients or geographical axes and in cases when only a fraction of the communities have universal dynamics. We also examine the ability of the method to test the Neutral Theory of Biodiversity.

3. We found that multiple factors and processes that produce variability among observed communities lead in general to a spurious detection of universal dynamics. In particular, sets of communities subject to environmental gradients, differences in stochasticity and in sampling quality among communities, and cases in which only a small fraction of the communities are similar, may be detected as having universal dynamics. We suggest the DOC method as a potential for testing the symmetric interactions assumption of neutral or symmetric theories, superior to the common techniques which are based on species abundance statistics.

4. Despite the promise of the DOC method, interpreting its results may be challenging and further developments are needed to address the question of similarity in dynamics.

Introduction

Ecological communities are highly complex and dynamical systems displaying high variability in space and time. To try and disentangle the causes of this variability, indices quantifying the compositional similarity of communities (sometimes referred to as "beta diversity"), such as Jaccard and Bray-Curtis have been developed. These indices are frequently used to study the response of ecological communities to environmental gradients (e.g. Bray & Curtis 1957), the importance of determinism vs. stochasticity in shaping communities (e.g. Chase *et al.* 2011; Segre *et al.* 2014), the drivers of temporal dynamics (e.g. Kampichler & van der Jeugd 2013; Kalyuzhny, Kadmon & Shnerb 2015), and many other questions. While quantifying the compositional similarity between communities is definitely a basic tool in ecology, a more fundamental question is – do these communities have similar underlying dynamics?

This new concept of dynamical similarity may open new research horizons, and has great practical importance. From the basic science perspective, it may be possible to detect the factors that determine, increase or reduce dynamical similarity. Moreover, communities with similar dynamics may respond similarly to perturbations, management actions or (in the case of the human microbiome) medical treatments. If similar dynamics is the rule, or a common scenario, then knowledge obtained from well-researched systems can be generalized into many other, less known systems.

Despite its importance, this question has not been addressed until the recent work of Bashan *et al.* (2016). This work developed the Dissimilarity-Overlap Curve (hereafter DOC) method to detect if a set of communities have the same dynamics, using a single snapshot of every community. Briefly (see below for elaboration), to use the method one calculates, for each pair of communities in the set, the proportion of individuals belonging to the shared species (overlap) and the compositional dissimilarity among the species shared by these communities. If a negative relationship is detected between overlap and dissimilarity for high values of overlap, this suggests that the communities have the same ("universal") dynamics with interspecific interactions. If no relationship is detected, this is interpreted either as a result of no interaction between the species, or as a result of the dynamics being different for each community ("individual dynamics"). The method was applied to the human microbiome, finding that the gut and mouth microbiomes in humans indeed have universal dynamics (Bashan *et al.* 2016).

While the question of universal dynamics has not been studied as such, it seems, at least intuitively, that there is an important ecological theory that predicts universal dynamics – the Unified Neutral Theory of Biodiversity (UNTB, Hubbell 2001). UNTB has been the center of a hot debate in the past decade and a half (e.g. McGill 2003; Purves & Pacala 2005; Alonso, Etienne & McKane 2006; Clark 2009), and has been tested extensively in multiple communities of various taxa (e.g. McGill 2003; Volkov *et al.* 2007; Connolly *et al.* 2014) and in particular in the human microbiome (e.g. Venkataraman et al. 2015; Li & Ma 2016). UNTB assumes that all the species in the community compete symmetrically with each other, and therefore the dynamics is governed solely by random dispersal, speciation and drift. Since all the species are interacting and are identical, the theory predicts universal dynamics with interactions (Azaele *et al.* 2016). Therefore the DOC method should be able to test a

core assumption of UNTB that has never (to our knowledge) been tested – that the dynamics is identical for at least some sets of communities.

A major strength of the DOC method is its great simplicity and attractiveness, making us believe it may be widely applicable for many large ecological datasets in order to identify the relative importance of universal versus local factors that shape community dynamics in different locations. Not only in ecology: the method appears to be relevant to any network of interacting "species", from biochemical dynamics of protein-protein interaction (Voit 2000), to gene regulatory networks (Karlebach & Shamir 2008). Therefore, it is important to assess the performance of the DOC under realistic assumptions.

Bashan *et al.* tested their method by studying model communities. They examined deterministic Lotka-Volterra equations in the parameter regime where the interspecific interactions are weak (the full coexistence phase in the model of Kessler & Shnerb (2015)), leading to a single stable equilibrium composition in each community. The DOC was calculated for the communities at equilibrium, and the differences between communities (in the universal case) were caused only by different initial conditions, as each community was initialized with some of the species absent. This is a highly unrealistic settings, ignoring many important processes and factors that shape the observed variability in the composition of ecological communities, such as stochastic variation affecting individuals (demographic stochasticity) and entire populations (environmental stochasticity, Lande, Engen & Saether 2003; Chisholm *et al.* 2014; Kalyuzhny *et al.* 2014a), sampling errors (Knape & de Valpine 2012; Kalyuzhny *et al.* 2014b), immigration (Hanski & Gilpin 1997; Chave, Alonso & Etienne 2006), environmental gradients, geographic variation in abundance (not related to environmental factors, e.g. Svenning & Skov 2005) and many other processes and factors. Moreover, each community may be affected by different levels of each of the above

processes and factors. It is also possible that only a fraction of the communities have universal dynamics, while the other communities have highly variable dynamics.

Here we test the ability of the DOC method to correctly classify universal and non-universal dynamics in more realistic settings. We apply the DOC method to simulated communities that are subject to various levels and regimes of stochasticity and sampling errors, as well as to communities which are shaped solely by environmental gradients or geographical variation. We also examine the predictions of UNTB regarding the DOC and the utility of the DOC method in testing UTNB. We start by describing the DOC method, as suggested by Bashan *et al.* (2016), followed by an intuitive explanation of its underlying logic.

How and why the DOC works

Given a set of communities for which the relative abundances of the species are known, for each pair of communities two indices are calculated: the overlap and the dissimilarity in abundance. An example of the type of data that can be analyzed using the DOC is presented in Table 1, where the abundance of four different species, (A-D) is given for four different communities (0-3).

The overlap between two communities with relative abundance vectors $\tilde{x}$ and $\tilde{y}$ is defined as:

(1) $O(\tilde{x}, \tilde{y}) = \sum_{i \in S} \frac{\tilde{x}_i + \tilde{y}_i}{2}$,

where the summation is done over all species S that are shared between by the two communities. This index ranges between 0 (no shared species) and 1 (all the species are shared), and can be thought of as analogous to the Jaccard index weighted by abundance, or as the average proportion of individuals in the communities that belong to shared species.

Next, for the calculation of dissimilarity only the shared species are considered and their relative abundances are renormalized to sum to one, giving the vectors $\widehat{x}$ and $\widehat{y}$ respectively. Bashan *et al.* (2016) showed that multiple dissimilarity indices may be used, but they recommended specifically the root Jensen-Shannon divergence, defined as:

$$(2) \qquad D(\widehat{x}, \widehat{y}) = (\frac{D_{KL}(\widehat{x}, m) + D_{KL}(\widehat{y}, m)}{2})^{0.5},$$

where the vector m is $\frac{\widehat{x} + \widehat{y}}{2}$ and D_{KL} is the Kullback-Leibler divergence of $\widehat{x}$ and $\widehat{y}$. Accordingly, we will use this dissimilarity index throughout.

The overlaps and dissimilarities of each pair of communities in the set are then plotted in the overlap-dissimilarity plane (Fig. 1). A negative slope in the high overlap region is interpreted as indication for universal dynamics with interactions (Fig. 1a), while the absence of a relationship in this region (Fig. 1b) is interpreted as a result of either no interactions, or as communities having different dynamics ("individual dynamics").

The underlying logic of the DOC method is as follows: *if* the abundance of a species is determined solely by its interactions with itself and with other species, and *if* these interactions are universal (e.g., in each community species 1 competes with species 2, is predated by species 3, with the very same niche overlap or predation pressure) then (up to effects of noise and alternative steady states) two communities that have the same species will have the same equilibrium, i.e., the abundance of these species will be identical. Species turnover between communities (as in table 1) causes a decrease in overlap and, as the equilibrium changes, an increase in dissimilarity. This leads to the negative dissimilarity-overlap slope.

When the two conditions mentioned above are not fulfilled, e.g., when the abundance of a species is not determined by the effect of other species, or when the dynamics is non-

universal, meaning that the interaction between two specific species varies significantly between communities, one expects no relationships between species overlap and abundance similarity.

Two other aspects of the DOC method deserve attention: First, to eliminate possible artifacts, negative slopes are compared to a null model in which the richness and identities of the different species in the communities and are kept fixed but the abundance of each species in a given community is resampled from all the observed abundances of this species. Another noteworthy feature is that the renormalization of the shared species abundances may lead to the emergence of a positive slope, particularly in the low overlap region. This is easily understood if we remember that if only one species is shared between two communities, the dissimilarity D(1,1) is 0, and it may only increase from that point. For higher overlap, there is still a weak positive dependence of dissimilarity on the number of shared species (Bashan *et al.* 2016).

Methods

To study various dynamical situations, we used a multispecies Ricker model (Kilpatrick & Ives 2003) in all the non-neutral simulations of community dynamics. In the model, the expected population of species i at time $t+1$, $N_{i,t+1}$, is:

$$(3) \quad \mathrm{E}(N_{i,t+1}) = N_{i,t}\exp\left(r_i(1 - \frac{N_{i,t} - \sum_{j \neq i}\alpha_{ij}N_{j,t}}{K_i}) + \varepsilon_{i,t}\right),$$

Where r_i and K_i are the intrinsic growth rate and carrying capacity of species i, respectively. α_{ij} is the per capita effect of species j on species i and $\varepsilon_{i,t}$ is a normally distributed random variable (with zero mean and variance σ_e^2). $\varepsilon_{i,t}$ is drawn independently for each species and each time-step, representing environmental stochasticity. To consider also the effect of demographic noise, (meaning that the number of individuals is an integer, and extinction

happens when this number is zero) the actual population size at time $t+1$, $N_{i,t+1}$, is drawn from a Poisson distribution with mean $E(N_{i,t+1})$ (Chisholm *et al.* 2014; Kalyuzhny *et al.* 2014a).

In each batch of our numerical experiment (with either individual or universal dynamics) we simulated 80 sets of 100 communities each, using the parameter regimes shown in table 2. These particular parameter regimes were chosen so that for the universal dynamics there would be relatively strong interactions, magnifying their effect, and for the individual dynamics there would be relatively high overlap. All communities initially had 100 species whose initial abundance, N_0, was drawn (independently for each species in each community) from a uniform distribution, as also specified in Table 2. Simulations were run for 8000 time steps and sampled once, with stochastic or deterministic extinctions leading to non-full overlap.

We tested the DOC method by considering two types of communities:

1. **Universal dynamics**: all the communities in a set had a single combination of dynamical parameters r_i, K_i and α_{ij} that were drawn according to table 2. The noise terms were drawn independently (from the same distribution) for each community.

2. **Non-Universal dynamics**: the parameters r_i, K_i and α_{ij} were drawn independently for each community.

Given the end results of the simulations, for each set of communities the overlap and dissimilarity of each pair of communities were calculated and the slope was found by OLS regression using points with above-median overlap, as in some of the analyses of Bashan *et al*. We present for each batch of simulations (80 sets of 100 communities each) the mean slope, as well as the central 95% of the distribution of slopes that were found.

To test the effect of stochasticity on the performance of the DOC method, we considered two scenarios for environmental stochasticity: (i) identical σ_e in all the communities (hereafter "uniform noise"), and (ii) different σ_e in each community (hereafter "variable noise"), with the σ_e of each community independently drawn from a uniform distribution between 0.03 and σ_{emax}. The second scenario was introduced since the relative importance of deterministic and stochastic processes may be different in different communities (Segre *et al.* 2014). Hence, we study the effect of σ_e and σ_{emax} on the DOC slope under both universal and individual dynamics. Following the logic of the method, we expected no slope under individual dynamics and a negative slope under universal dynamics, regardless of the stochasticity.

Our next goal is to test the robustness of the method to sampling errors, which we model as an independent sampling of a number of individuals that equals a proportion P of the community. Again, we consider uniform sampling, where P is the same for each community in a set, and variable sampling, where P is uniformly drawn for each community between 0 (rounded to one individual) and P_{max}. This represents differences in the quality of sampling in different communities. We applied both sampling methods to communities with individual and universal dynamics, expecting, as before, a negative slope for universal dynamics and no slope for individual dynamics.

To further test if our qualitative results are robust to initial conditions, we re-run our analyses also in a model incorporating immigration from a regional species pool.

To test the predictions of UNTB on the DOC slope, we simulated the spatially implicit version of UNTB (Hubbell 2001) with local community size $J = 10^4$. We tested the effect of regional diversity as manifested by the "biodiversity number" (θ) and immigration (m) by simulating a set of 200 communities with each of the 12 combinations of $\theta = 20$, 50 and 100 and $m = 10^{-4}$, 10^{-3}, 10^{-2} and 10^{-1}. The simulations were run for 5000 generations (each

consisting of J time steps) and sampled once, with the slope calculated by OLS regression using points with above-median overlap. Since in UNTB there is no additional variability in the parameters there was no need to generate multiple sets of simulations for each parameter regime. Because the species are (symmetrically) competing and the dynamics is universal (identical in all communities), we expected the DOC slope to be negative.

The effect of environmental gradients and/or geographical axes on the DOC slope was tested using a toy model of 200 species that may exist in 500 sites that lie along a gradient/axis. The abundance of each species in the model is a linear function of the site index (Fig 5a). The slope of this function is drawn from a standard normal distribution and the intercept from a normal distribution with mean 0 and S.D. of 50 for each species, and a negative abundance is rounded to 0. Hence, species respond independently to the site (whether sites represent different environmental conditions or geographical coordinate). We calculated the DOC for this model, expecting to get a negative slope because as species undergo turnover along the gradient, the abundance of the shared species also changes.

Our last goal was to test the performance of the DOC method when some of the communities have universal dynamics, while others have individual dynamics. To that end we simulated 80 sets of 100 communities each undergoing universal dynamics with the baseline parameters (see Table 2), and 80 sets of 100 communities undergoing individual dynamics. To ensure that the individual dynamics resemble the universal dynamics, we generated these communities using the baseline parameters of the universal dynamics, and then added to each α_{ij} a normally distributed random number with mean 0 and S.D. of 0.15 and to each K_i a normally distributed random number with mean 0 and S.D. of $1.5 \cdot 10^4$, independently for each community. Then, we generated 80 mixed sets of 100 communities, in which P_{univ} of the communities are drawn without replacement from the universal set, and 1- P_{univ} are drawn without replacement from the corresponding individual set. As before, we calculated the

DOC slope and presented its distribution. Since the slope is determined by the high overlap communities, we expect that sets with some levels of mixing would show negative slopes despite the non-universality of all the communities.

In all cases when a negative slope was found, we compared the slope to the null model to rule out artifacts. The full code for all the analyses is provided in appendix S2.

Results

The effect of stochasticity is presented in Fig. 2 and Fig. S1. Under universal dynamics, an increase in the strength of the stochasticity (whether uniform or variable between different communities) caused an increase in the slope, up to the point where the slopes may became primarily slightly positive (Fig. 2a,c). Such positive slopes may be a result of comparing communities with different richness, because the dissimilarity has a weak positive dependence on richness (see "how the DOC works" above and Bashan *et al.* 2016). When the dynamics is individual with uniform noise (Fig. 2b), the slopes are small and primarily positive, with a weak dependence on the strength of the noise. Surprisingly, we found that if the dynamics is individual with variable levels of stochasticity (Fig 2d), the slopes may become strongly negative, mimicking the behavior under universal dynamics with weak noise. Similar results were obtained by leaving the noise, σ_e, fixed, and varying the strength of the density regulation, the r_is (Fig S2).

Fig. 3 and Fig. S3 show the effect of uniform and variable sampling regimes under universal and individual dynamics on the DOC slope. For universal dynamics (Fig. 3a,c) Both uniform and variable sampling increased the DOC slope up to positive values. This is likely due to the decrease in overlap and because such sampling eliminates the rare species, which are most affected by the interactions, masking their effect. Under individual dynamics uniform sampling had no effect on the DOC slope (Fig 3b), but variable sampling had remarkable

effects (Fig 3d) – under intense sampling levels ($P_{max} = 10^{-1}$), the slopes became highly variable, and then intermediate sampling ($P_{max} = 10^{-2}$) turned the slopes to negative, contrary to expectations. Scarce sampling ($P_{max} = 10^{-3}$) finally turned the slopes to be positive, likely due to the strong decrease in overlap in this regime.

Qualitatively very similar results were obtained when incorporating weak immigration into the models that test the effects of sampling and noise on the DOC – see appendix S1 and Figures S4-S5 therein.

We were further surprised to find out that neutral dynamics, although universal, never leads to negative slopes, but primarily to positive slopes of the DOC (Fig. 4). This is also the case for high values of dispersal, despite the significant attraction of the local community to the regional composition in this case. Similar results were obtained using simulations of the Ricker model in which all species had $\alpha_{ij} = 0.5$, representing symmetric interactions where intra-specific competition is stronger than inter-specific competition (Fig. S6), despite the fact that here the local dynamics is supporting an isolated stable fixed point (Pigolotti & Cencini 2013).

As expected, communities responding to environmental gradients exhibited a DOC with a clear and strong negative slope (Fig. 5b).

Finally, as we expected, we obtained negative slopes for mixed universal and individual dynamics (Fig. 6 and Fig. S7), and generally, the higher P_{univ}, the more negative the slopes became. Surprisingly, however, even when only 20 percent of the communities are universal, all the slopes turned out to be negative, indicating that even if a very small proportion of communities are universal, the set of communities would be classified as "universal dynamics".

Discussion

Often, studies on methodology examine the sensitivity (type II error rate) and specificity (type I error rate) of a given method. In this work we focused on a more fundamental issue – the interpretation of the DOC method output, by testing its results under realistic ecological scenarios which we controlled through simulations.

We found that many realistic scenarios of individual dynamics or non-interactive species lead to a false detection of "universal dynamics". Specifically, environmental gradients, variation along geographical axes, variations in the importance of deterministic and stochastic processes (manifested by differences between communities in the magnitudes environmental stochasticity or the growth rate), and different sampling quality (manifested by difference between communities in the proportion of detected individuals), all led to negative slopes under non-universal dynamics. These surprising results could be understood as follows: Both sampling and stochasticity affect the presence of species in the (observed) community – high detection or weak stochasticity lead to high overlap of species, while low detection and strong stochasticity lead to lower overlap through extinctions or missing species. Low detection and strong stochasticity in some communities also increase the dissimilarity in species that are shared with other communities through noise. Hence, a set of communities with some communities having low detection or strong stochasticity and some having high detection or weak stochasticity may show a negative dependence of dissimilarity on overlap, despite having individual dynamics.

We think this enables us to attempt a generalization – every process that would affect both the presence of species and their abundance, (either in the real community or in the sample) in may lead to a negative DOC slope, regardless of the universality of interactions. To test this idea, we used a very general and phenomenological model of an abstract process that

randomly replaces species by other species (reducing overlap) and alters the abundance of species. If this process affects different communities at different levels, a negative DOC slope is observed – see Appendix S3 for results and details.

Multiple processes that affect both presence and abundance which we did not test may be important factors shaping ecological communities, such as different levels of habitat heterogeneity (i.e. environmental variability in space, Rosenzweig 1995; Allouche *et al.* 2012), different levels of productivity or disturbance (Kondoh 2001), different species pools (Ricklefs 1987; Zobel 1997), differential effect of keystone species and predators (Menge *et al.* 1994; Chase *et al.* 2009), shared response to environmental forcing (Moran 1953) and many more. Moreover, we found that even if a very small fraction of communities are indeed universal and the other communities are very different – universal dynamics may be detected nevertheless. All these issues pose a severe challenge to the interpretation of the DOC results when applied to realistic communities.

Some of these issues may be addressed by comparing the Relative Abundance Distributions (RADs) of the communities (not only of the shared species). If the dynamics is indeed universal, we would expect relatively similar RADs (but the opposite, similar RADs implying similar dynamics, is generally not true (Purves & Pacala 2005)). On the other hand, the effects mentioned above should change the RADs. Hence, if both conditions, similarity in the RADs at different overlaps, and negative DOC slope are met, this may be considered a somewhat stronger evidence for universal dynamics of at least some of the communities.

The problem of species response to environmental gradients or geographical axes may be addressed by statistically controlling for the effect of environmental and geographic variables that are known to affect species abundances (as acknowledged and done by Bashan *et al.*, who performed the analysis separately for different grouping of the microbiome communities

according to host-specific factors). However, since niche space may be highly multidimensional, special care must be taken to detect the important variables, and this may be a formidable task. Moreover, since differences in interspecific interactions are often shaped by differences in the environment, this correction, if successful, may actually "correct" the underlying individual dynamics by eliminating the causes of inter-community variability in dynamics, leading to the detection of universal dynamics. In the extreme case where all the environmental heterogeneity between communities is somehow corrected for, it is quite trivial that the communities may have the same dynamics, even if the set of communities has highly variable dynamics. We conclude that more research is needed on the validation of the DOC method in various realistic scenarios.

We also found that symmetric universal interactions (identical for all species in all communities), such as in UNTB, do not result in a negative DOC slope. Although the species are interacting, their identities and composition have no real importance, and as a result overlap does not affect dissimilarity. Therefore, if communities have symmetric interactions or are stabilized primarily by immigration (Chave, Alonso & Etienne 2006; Condit, Chisholm & Hubbell 2012; Azaele *et al.* 2016), the DOC method would not detect their universality. On the other hand, if a reliable negative slope is detected, this may be a strong rejection of symmetric interactions and stabilization by immigration. It is noteworthy, however, that although we considered two classic neutral models, there is a possibility that other neutral models may show different behaviors. Most tests of Neutral theory using observational data focused primarily on RADs (McGill 2003; Volkov *et al.* 2007; Connolly *et al.* 2014; Li & Ma 2016). However, different Neutral theories may produce a variety of RADs (Kessler & Shnerb 2014; Kalyuzhny, Kadmon & Shnerb 2015), and there is a growing consensus that RADs may not provide a strong test of neutrality (McGill *et al.* 2007; Chisholm *et al.* 2014; Kalyuzhny *et al.* 2014b; Al Hammal *et al.* 2015). Hence, the DOC method indeed offers a

unique test of the core assumptions of neutrality – symmetric interactions between species in a set of communities. In particular, the neutrality of the human microbiome has been debated in recent studies (e.g. Venkataraman *et al.* 2015; Li & Ma 2016), and the DOC results of (Bashan *et al.* 2016) provide support for non-neutral processes acting in the gut and mouth.

In our view, the case of the DOC method is typical of many ecological studies. Many interesting questions in population and community ecology deal with the dynamical processes that shape populations and communities. These questions are often addressed by looking at static snapshots of these populations and communities and trying to infer the underlying processes (e.g. Hanski 1994; Volkov *et al.* 2005; Chave, Alonso & Etienne 2006; Kefi *et al.* 2007; Li & Ma 2016). However, since multiple processes and factors often lead to the same static patterns, this approach may be misleading and lead to contrasting conclusions (e.g. Hubbell 2001; Purves & Pacala 2005; Volkov *et al.* 2005; Chave, Alonso & Etienne 2006; Kefi *et al.* 2007; Seri, Maruvka & Shnerb 2012). We believe that the best way to use observational data to understand ecological processes is to explicitly examine dynamical data and to test the correspondence between dynamical models and the observed time-series of populations and communities (Chisholm & O'Dwyer 2014). In this spirit, we hope that new methods will be developed that would allow addressing the novel and interesting question of similarity in dynamics by explicitly examining ecological dynamics.

Acknowledgements

Data accessibility

This work does not use any data.

Author contribution statement

M.K. and N.M.S planned the study, M.K performed the study and M.K. and N.M.S wrote the paper.

References

Al Hammal, O., Alonso, D., Etienne, R.S. & Cornell, S.J. (2015) When Can Species Abundance Data Reveal Non-neutrality? *Plos Computational Biology,* **11**.

Allouche, O., Kalyuzhny, M., Moreno-Rueda, G., Pizarro, M. & Kadmon, R. (2012) Area–heterogeneity tradeoff and the diversity of ecological communities. *Proceedings of the National Academy of Sciences,* **109,** 17495-17500.

Alonso, D., Etienne, R.S. & McKane, A.J. (2006) The merits of neutral theory. *Trends in Ecology & Evolution,* **21,** 451-457.

Azaele, S., Suweis, S., Grilli, J., Volkov, I., Banavar, J.R. & Maritan, A. (2016) Statistical mechanics of ecological systems: Neutral theory and beyond. *Reviews of Modern Physics,* **88**.

Bashan, A., Gibson, T.E., Friedman, J., Carey, V.J., Weiss, S.T., Hohmann, E.L. & Liu, Y.Y. (2016) Universality of human microbial dynamics. *Nature,* **534,** 259-+.

Bray, J.R. & Curtis, J.T. (1957) An Ordination of the Upland Forest Communities of Southern Wisconsin. *Ecological Monographs,* **27,** 326-349.

Chase, J.M., Biro, E.G., Ryberg, W.A. & Smith, K.G. (2009) Predators temper the relative importance of stochastic processes in the assembly of prey metacommunities. *Ecology Letters,* **12,** 1210-1218.

Chase, J.M., Kraft, N.J.B., Smith, K.G., Vellend, M. & Inouye, B.D. (2011) Using null models to disentangle variation in community dissimilarity from variation in alpha-diversity. *Ecosphere,* **2**.

Chave, J., Alonso, D. & Etienne, R.S. (2006) Theoretical biology - Comparing models of species abundance. *Nature,* **441,** E1-E1.

Chisholm, R.A., Condit, R., Abd Rahman, K., Baker, P.J., Bunyavejchewin, S., Chen, Y.-Y., Chuyong, G., Dattaraja, H.S., Davies, S., Ewango, C.E.N., Gunatilleke, C.V.S., Gunatilleke, I.A.U.N., Hubbell, S., Kenfack, D., Kiratiprayoon, S., Lin, Y., Makana, J.-R., Pongpattananurak, N., Pulla, S., Punchi-Manage, R., Sukumar, R., Su, S.-H., Sun, I.F., Suresh, H.S., Tan, S., Thomas, D. & Yap, S. (2014) Temporal variability of forest communities: empirical estimates of population change in 4000 tree species. *Ecology Letters,* **17,** 855-865.

Chisholm, R.A. & O'Dwyer, J.P. (2014) Species ages in neutral biodiversity models. *Theoretical Population Biology,* **93,** 85-94.

Clark, J.S. (2009) Beyond neutral science. *Trends in Ecology & Evolution,* **24,** 8-15.

Condit, R., Chisholm, R.A. & Hubbell, S.P. (2012) Thirty Years of Forest Census at Barro Colorado and the Importance of Immigration in Maintaining Diversity. *Plos One,* **7**.

Connolly, S.R., MacNeil, M.A., Caley, J., Knowlton, N., Cripps, E., Hisano, M., Thibaut, L.M., Bhattacharya, B.D., Benedetti-Cecchi, L., Brainard, R.E., Brandt, A., Bulleri, F., Ellingsen, K.E., Kaiser, S., Kroncke, I., Linse, K., Maggi, E., O'Hara, T.D., Plaisance, L., Poore, G.C.B., Sarkar, S.K., Satpathy, K.K., Schuckel, U., Williams, A. & Wilson, R.S. (2014) Commonness and rarity in the marine biosphere. *Proceedings of the National Academy of Sciences of the United States of America,* **111,** 8524-8529.

Hanski, I. (1994) A practical model of metapopulation dynamics. *Journal of Animal Ecology,* **63,** 151-162.

Hanski, I. & Gilpin, M.E. (1997) *Metapopulation biology: ecology, genetics, and evolution*. Academic press San Diego.

Hubbell, S.P. (2001) *The unified neutral theory of biodiversity and biogeography*. Princeton University Press.

Kalyuzhny, M., Schreiber, Y., Chocron, R., Flather, C.H., Kadmon, R., Kessler, D.A. & Shnerb, N.M. (2014a) Temporal fluctuation scaling in populations and communities. *Ecology,* **95,** 1701-1709.

Kalyuzhny, M., Seri, E., Chocron, R., Flather, C.H., Kadmon, R. & Shnerb, N.M. (2014b) Niche versus Neutrality: A Dynamical Analysis. *The American naturalist,* **184,** 439-446.

Kalyuzhny, M.K., Kadmon, R. & Shnerb, N.M. (2015) A neutral theory with environmental stochasticity explains static and dynamic properties of ecological communities. *Ecology Letters,* **18,** 572-580.

Kampichler, C. & van der Jeugd, H.P. (2013) Determining patterns of variability in ecological communities: time lag analysis revisited. *Environmental and Ecological Statistics,* **20,** 271-284.

Karlebach, G. & Shamir, R. (2008) Modelling and analysis of gene regulatory networks. *Nature Reviews Molecular Cell Biology,* **9,** 770-780.

Kefi, S., Rietkerk, M., Alados, C.L., Pueyo, Y., Papanastasis, V.P., ElAich, A. & de Ruiter, P.C. (2007) Spatial vegetation patterns and imminent desertification in Mediterranean arid ecosystems. *Nature,* **449,** 213-U215.

Kessler, D.A. & Shnerb, N.M. (2014) Neutral-like abundance distributions in the presence of selection in a continuous fitness landscape. *Journal of Theoretical Biology,* **345,** 1-11.

Kessler, D.A. & Shnerb, N.M. (2015) Generalized model of island biodiversity. *Physical Review E,* **91,** 042705.

Kilpatrick, A.M. & Ives, A.R. (2003) Species interactions can explain Taylor's power law for ecological time series. *Nature,* **422,** 65-68.

Knape, J. & de Valpine, P. (2012) Are patterns of density dependence in the Global Population Dynamics Database driven by uncertainty about population abundance? *Ecology Letters,* **15**.

Kondoh, M. (2001) Unifying the relationships of species richness to productivity and disturbance. *Proceedings of the Royal Society of London Series B-Biological Sciences,* **268**.

Lande, R., Engen, S. & Saether, B.-E. (2003) *Stochastic population dynamics in ecology and conservation*. Oxford University Press.

Li, L. & Ma, Z. (2016) Testing the Neutral Theory of Biodiversity with Human Microbiome Datasets. *Scientific Reports,* **6,** 31448.

McGill, B.J. (2003) A test of the unified neutral theory of biodiversity. *Nature,* **422,** 881-885.

McGill, B.J., Etienne, R.S., Gray, J.S., Alonso, D., Anderson, M.J., Benecha, H.K., Dornelas, M., Enquist, B.J., Green, J.L., He, F., Hurlbert, A.H., Magurran, A.E., Marquet, P.A., Maurer, B.A., Ostling, A., Soykan, C.U., Ugland, K.I. & White, E.P. (2007) Species abundance distributions: moving beyond single prediction theories to integration within an ecological framework. *Ecology Letters,* **10,** 995-1015.

Menge, B.A., Berlow, E.L., Blanchette, C.A., Navarrete, S.A. & Yamada, S.B. (1994) THE KEYSTONE SPECIES CONCEPT - VARIATION IN INTERACTION STRENGTH IN A ROCKY INTERTIDAL HABITAT. *Ecological Monographs,* **64,** 249-286.

Moran, P. (1953) The statistical analysis of the Canadian Lynx cycle. *Australian Journal of Zoology,* **1,** 291-298.

Pigolotti, S. & Cencini, M. (2013) Species abundances and lifetimes: From neutral to niche-stabilized communities. *Journal of Theoretical Biology,* **338,** 1-8.

Purves, W., Drew & Pacala, W., Stephen (2005) Ecological drift in niche-structured communities: neutral pattern does not imply neutral process. *Biotic Interactions in the Tropics: Their Role in the Maintenance of Species Diversity* (eds D. Burslem, M. Pinard & S. Hartley), pp. 108-138. Cambridge University Press.

Ricklefs, R.E. (1987) Community diversity - relative roles of local and regional processes. *Science,* **235,** 167-171.

Rosenzweig, M.L. (1995) *Species diversity in space and time.* Cambridge University Press, Cambridge.

Segre, H., Ron, R., De Malach, N., Henkin, Z., Mandel, M. & Kadmon, R. (2014) Competitive exclusion, beta diversity, and deterministic vs. stochastic drivers of community assembly. *Ecology Letters,* **17,** 1400-1408.

Seri, E., Maruvka, Y.E. & Shnerb, N.M. (2012) Neutral Dynamics and Cluster Statistics in a Tropical Forest. *American Naturalist,* **180,** E161-E173.

Svenning, J.C. & Skov, F. (2005) The relative roles of environment and history as controls of tree species composition and richness in Europe. *Journal of Biogeography,* **32,** 1019-1033.

Venkataraman, A., Bassis, C.M., Beck, J.M., Young, V.B., Curtis, J.L., Huffnagle, G.B. & Schmidt, T.M. (2015) Application of a Neutral Community Model To Assess Structuring of the Human Lung Microbiome. *Mbio,* **6**.

Voit, E.O. (2000) *Computational analysis of biochemical systems: a practical guide for biochemists and molecular biologists.* Cambridge University Press.

Volkov, I., Banavar, J.R., He, F.L., Hubbell, S.P. & Maritan, A. (2005) Density dependence explains tree species abundance and diversity in tropical forests. *Nature,* **438,** 658-661.

Volkov, I., Banavar, J.R., Hubbell, S.P. & Maritan, A. (2007) Patterns of relative species abundance in rainforests and coral reefs. *Nature,* **450,** 45-49.

Zobel, M. (1997) The relative role of species pools in determining plant species richness. An alternative explanation of species coexistence? *Trends in Ecology & Evolution,* **12,** 266-269.

Supporting information

Additional Supporting Information may be found online in the supporting information tab for this article:

Appendix S1. Model results with immigration

Appendix S2. ZIP file with code for reproducing all the results and figures, including pre-calculated results in some cases to save runtime

Appendix S3. Abstract process affecting presence and abundance

Figure S1. Example of DOCs from Fig. 2.

Figure S2. Effect of variable growth rates on the DOC slope, calculated for communities undergoing individual dynamics.

Figure S3. Example of DOCs from Fig. 3.

Figure S6. DOC under a symmetric Ricker model.

Figure S7. Example of DOCs from Fig. 6.

i. Dynamics with no interactions						
Community (k)	Sp. A	Sp. B	Sp. C	Sp. D	O(0,k)	D(0,k)
0	5	10				
1	5	10			1	0
2	5	10	3		0.9167	0
3	5	10	4	7	0.7885	0
ii. Universal Dynamics with interactions						
Community (k)	Sp. A	Sp. B	Sp. C	Sp. D	O(0,k)	D(0,k)
0	5	10				
1	5	10			1	0
2	4	9	3		0.9062	0.0194
3	2	6	4	7	0.7885	0.0649

Table 1. Numerical example of the DOC for (i) - dynamics with no interactions, where the introduction of new species does not affect the abundance of the shared species. Hence, as overlap decreases, the dissimilarity does not change. (ii) - Universal dynamics with interactions, where new species affect the present species in a similar manner across communities, leading to an increase of dissimilarity with species turnover. The leftmost column shows the numbering of communities, while the two rightmost columns present the overlap and dissimilarity, respectively, of each community compared to community 0.

| Baseline parameters:					
	K_i	r_i	α_{ij}	N_0	Baseline σ_e
Individual dynamics	~normal(m= 1000, σ =20)	~uniform(0.05, 0.2	~normal(m= 0, σ =0.02)	~uniform(0, 100)	0.1
Universal dynamics	10^5	~uniform(0.02, 0.1	~|normal(m= 0.1, σ =0.15)|	~uniform(0, 10^4)	0.05
Different scenarios:					
Uniform noise:	σ_e = 0, 0.02, 0.05, 0.1, 0.2				
Variable noise:	σ_{emax}=0.03, 0.05, 0.08, 0.13, 0.23				
Uniform sampling:	P = 1, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4} (only for universal dynamics).				
Variable sampling:	P_{max}=1, 10^{-1}, 10^{-2}, 10^{-3}, 10^{-4} (only for universal dynamics).				

Table 2. Parameter regimes used in the Ricker model simulations. For the individual dynamics, the parameters were drawn independently for each community, while for the universal dynamics identical parameters were drawn for all the communities in a set. The upper part shows the baseline used in all simulations, while the lower part shows the particular levels of parameters used in the different scenarios of sampling errors and noises. We used the absolute value of the α_{ij}s for the universal scenario to avoid the case of a single species taking over the community.
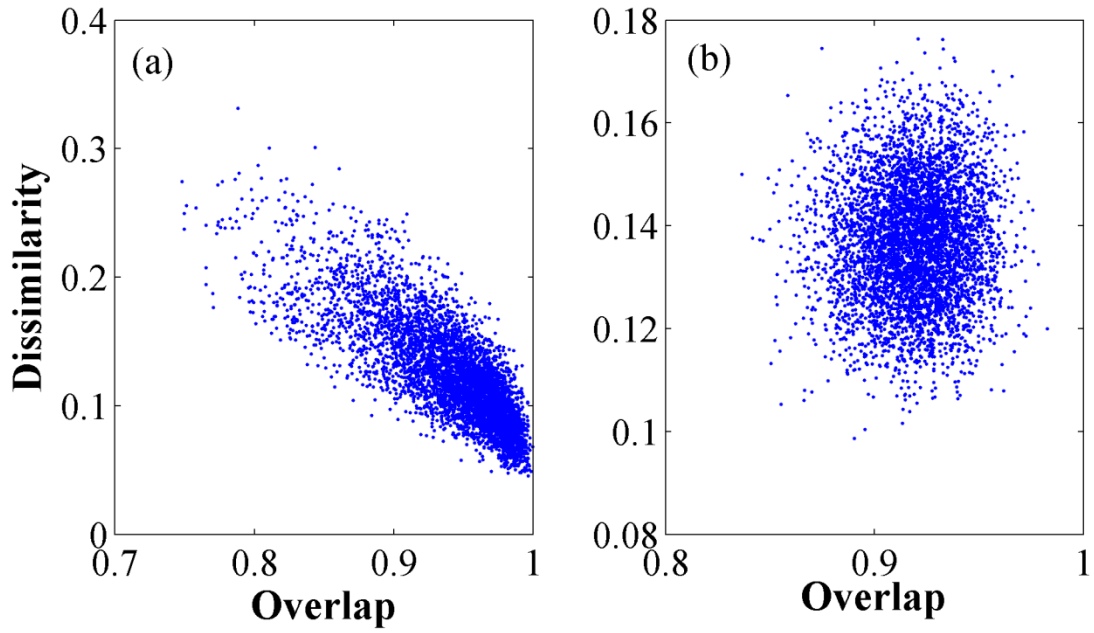
Fig. 1. Example of DOC curves for universal (a) and individual (b) dynamics; this figure corresponds to the basic findings of Bashan *et al* (2016). Each analysis was performed on 100 communities simulated using a Ricker model with the parameters in Table 2, with the exception that in (a) σ_e is 0. In the universal dynamics case there is a pronounced negative slope in the DOC, while the slope is zero for non-universal dynamics. Bashan *et al*. (2016) suggested that this is a generic feature: universal dynamics leads to a negative slope, non-universal dynamics leads to no slope.
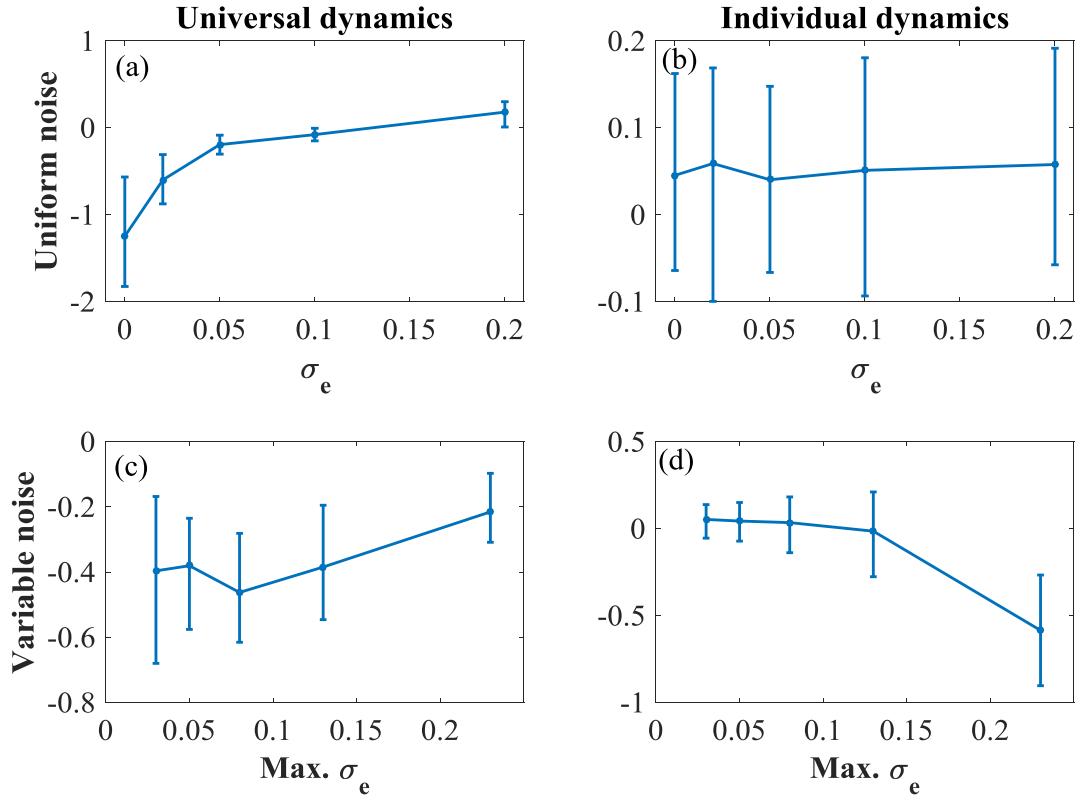
Fig. 2. Effect of uniform (a,b) and variable (c,d) stochasticity on the DOC slope calculated for communities undergoing universal (a,c) and individual (b,d) dynamics. In all panels we present the mean and the central 95% of the distribution of slopes. Fig S1 shows examples of specific DOCs (the same kind of data presented in Fig. 1) obtained for these systems. While the weak stochasticity regime is consistent with the DOC prediction – negative slopes for universal dynamics, zero for non-universal dynamics, noise may cause universal systems to develop zero and even positive DOC slope, while it decreases the slope in the non-universal systems and may even take it to the negative slope regime. The parameters for the simulations are given in Table 2.
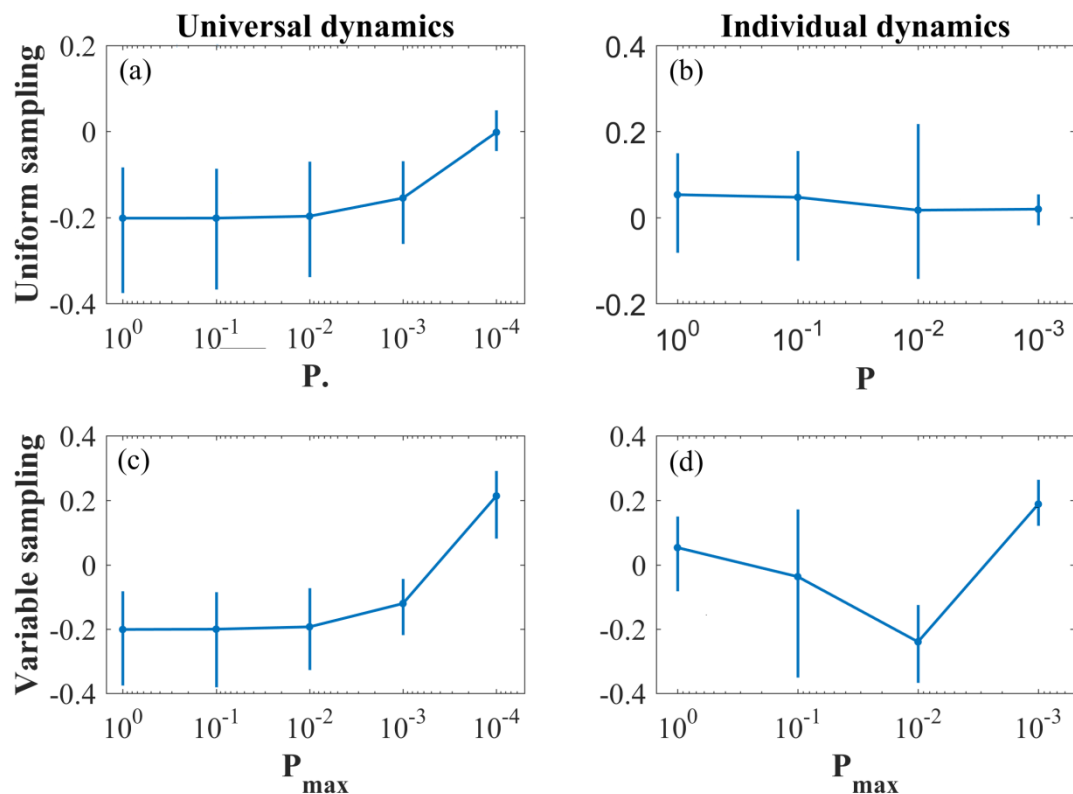
Fig. 3. Effect of uniform (a,b) and variable (c,d) sampling on the distribution of DOC slopes calculated for communities undergoing universal (a,c) and individual (b,d) dynamics. The mean and the central 95% of the distribution of slopes are shown against the sampling proportion P (for uniform sampling) and the maximal sampling proportion P_{max} (for variable sampling). All other parameters of the simulations are given in Table 2. Fig S3 shows examples of specific DOCs in this analysis. Under weak uniform sampling universal dynamics may yield zero and even positive DOC slopes, and non-universal dynamics may produce negative slopes for highly variable sampling intensities.
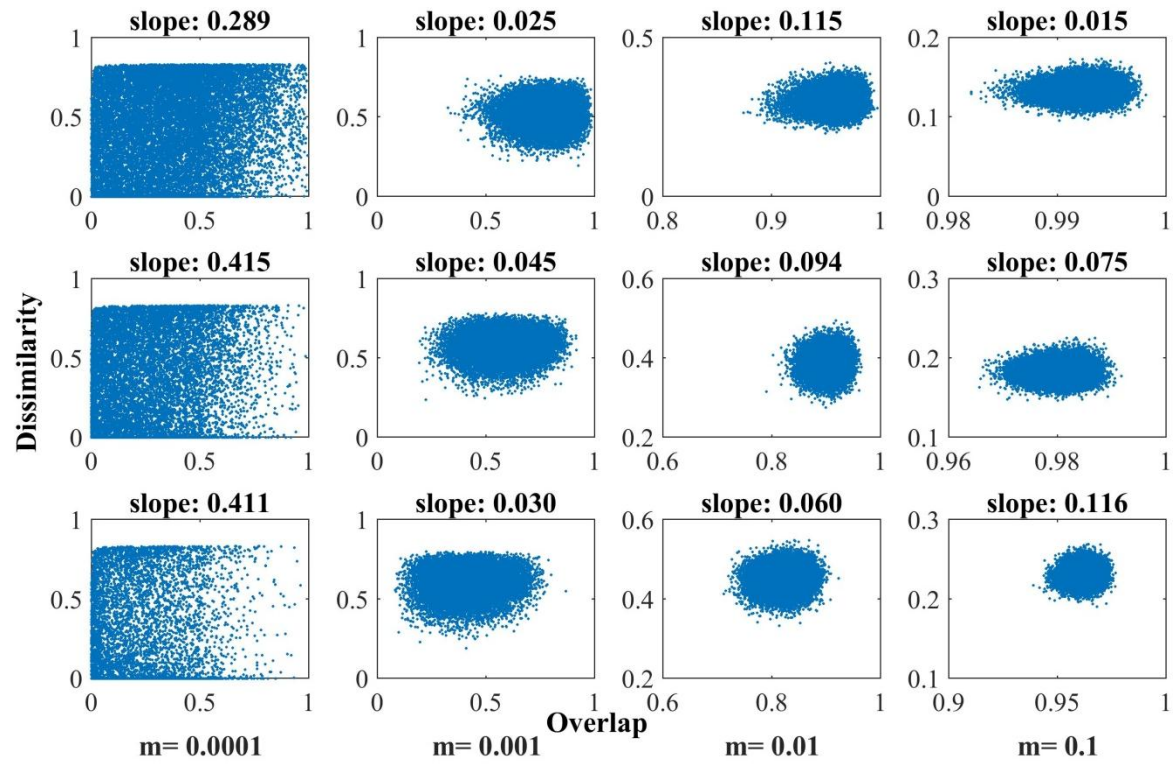
Fig. 4. DOC under UNTB with various levels of *m* and *θ*. Each plot presents the DOC of 200 communities with a specific combination of the biodiversity parameter of the regional community *θ* and dispersal level *m*. The slope, calculated by applying OLS to points with above-median overlap, is presented above each plot.
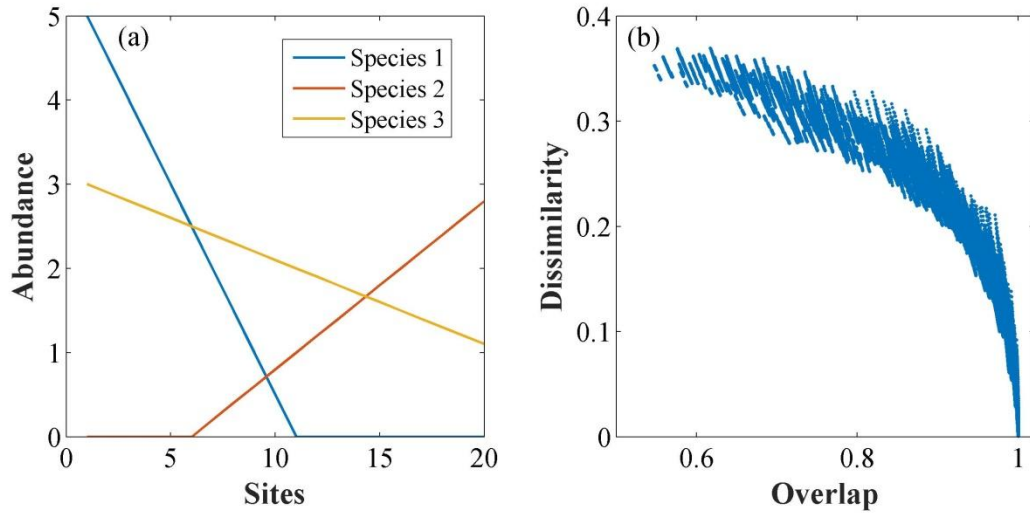
Fig 5. DOC in a scenario where communities are located along an environmental gradient or geographical axis. (a) depicts the toy-model which we use to create the gradients/axes – independent linear response of species abundance to the site index. (b) shows the DOC for this scenario. Here the DOC admits a pronounced negative slope although the dynamics is non-universal and non-interacting, since communities that are closer geographically or environmentally tend to have similar species compositions. i.e., since both overlap and similarity independently reflect the effect of a third factor: the local environment or location.
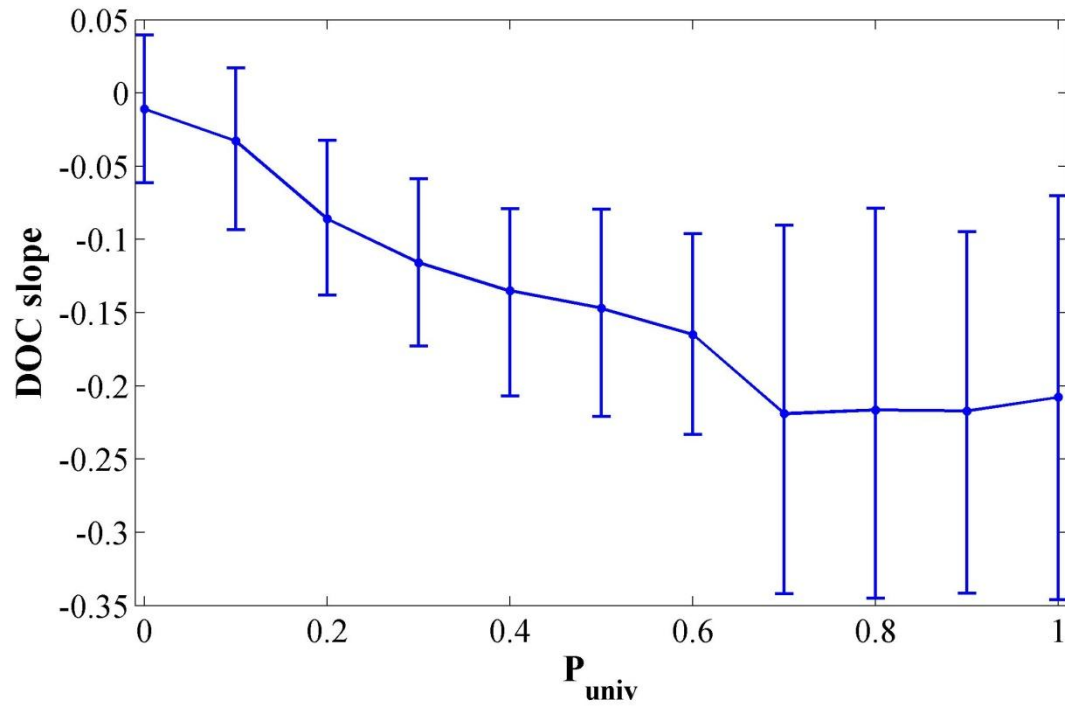
Fig. 6. Effect of P_{univ}, the proportion of communities with universal dynamics, on the distribution of DOC slopes. We present the mean and the central 95% of the distribution of slopes. Note that even for $P_{univ} = 0.2$, all slopes are negative. Fig S5 shows examples of specific DOCs in this analysis.