Different methods for niche and fitness differences computation offer contrasting explanations of species coexistence

Jurg W. Spaak^{1,*}, Po-Ju Ke^{2,3}, Andrew D. Letten⁴,
& Frederik De Laender⁵

1. Department of Ecology and Evolutionary Biology, Cornell University, Ithaca, New York, USA

2. Institute of Ecology and Evolutionary Biology, National Taiwan University, Taipei 106, Taiwan.

3. Department of Ecology & Evolutionary Biology, Princeton University, Princeton, New Jersey, USA.

4. School of Biological Sciences, University of Queensland, Brisbane, Queensland 4072, Australia

5. University of Namur, Rue de Bruxelles 61, Belgium; Institute of Life-Earth-Environment, Namur Center for Complex Systems, Namur, Rue de Bruxelles 61, Belgium;

* Corresponding author; e-mail: j.w.spaak@gmail.com, phone-number: +1 607 2626 195

Statement of authorship: J.W.S. wrote the computer code. J.W.S. wrote the first draft, all authors discussed content and contributed to revisions.

Data accessibility: All computer code used in this manuscript will be made publicly available on figshare.

Abstract

In modern coexistence theory, species coexistence can either arise via stabilizing mechanisms that increase niche differences or equalizing mechanisms that reduce fitness differences. Having a common currency for interpreting these mechanisms is essential for synthesizing knowledge across different studies and systems. Several methods for quantifying niche and fitness differences exist, but it remains unknown to what extent these methods agree on the reasons why species coexist. Here, we apply four common methods to quantify niche and fitness differences to one simulated and two empirical data sets. We ask if different methods result in different insights into what drives species coexistence. We find that different methods disagree on the effects of resource supply rates (simulated data), and of plant traits or phylogenetic distance (empirical data), on niche and fitness differences. More specifically, these methods often do not agree better than expected by chance. We argue for (1) a better understanding of what connects and sets apart different methods, and (2) the simultaneous application of multiple methods to enhance a more complete insight into why species coexist.

Keywords: niche differences | fitness differences | resource competition | annual plant model | coexistence

Introduction

Explaining biodiversity is a central goal in ecology (Hutchinson, 1959). There are many different perspectives to address this objective. Neutral theory focuses on regional processes (Hubbell & Knapp, 2003), contemporary niche theory focuses on investigating species' limiting factors (Tilman, 1982), and modern coexistence theory focuses on separating niche and fitness differences that help and hamper coexistence (Adler *et al.*, 2007).

Ecologically equivalent species can't coexist, as stated by the competitive exclusion principle (Levine & HilleRisLambers, 2009). To stably coexist, species must differ in some aspect. For example, if species differ in their resource uptake traits, this difference may contribute to negative frequency dependence and help rescue a rare species from competitive exclusion (Letten *et al.*, 2017). Under modern coexistence theory, such differences are termed niche differences (Spaak *et al.*, 2020; Chesson, 2000), which capture all species differences that help a species to recover from low density. Mechanisms that increase niche differences are called stabilizing mechanisms. Examples include resource partitioning, different root lengths, or species-specific predators. However, species differences may not only increase niche differences, but can also increase competitive differences. For example, differences in rooting length of plant species may not only cause a species to consume resources at different depth, but also affect how much resources a species can take up (Adler *et al.*, 2007). Such differences are termed fitness differences, which capture all differences of species that are associated with competitive dominance. More precisely, fitness differences identify all differences in species performance in the hypothetical case that there are no niche differences. Mechanisms that decrease fitness differences are called equalizing mechanisms.

Modern coexistence theory assesses coexistence by computing invasion growth rates (Chesson, 2003; Barabás *et al.*, 2018; Ellner *et al.*, 2018; Schreiber, 2000). Niche and fitness differences, which are derived from the decomposition of invasion growth rates, are tools that help us understand the mechanisms underlying species coexis-

tence (Spaak *et al.*, 2021; Carmel *et al.*, 2017; Carroll *et al.*, 2011; Adler *et al.*, 2007; Song *et al.*, 2019). While the invasion growth rates often suffice to predict coexistence, niche and fitness differences can in principle offer additional understanding of why the species do or do not coexist (Grainger *et al.*, 2019; Ke & Letten, 2018; Spaak *et al.*, 2020, 2021). Without such additional understanding the effort of computing niche and fitness differences is unnecessary.

Unfortunately, computing niche and fitness differences is not a straightforward task. There are many different methods to compute niche and fitness differences (Spaak *et al.*, 2020), with new methods still being developed (Johnson, 2021; Koffel *et al.*, 2021). Some of these methods are tailored to a specific community model (Godoy & Levine, 2014; Saavedra *et al.*, 2017; Chesson, 2018), while others are model agnostic (Spaak *et al.*, 2020; Carroll *et al.*, 2011; Zhao *et al.*, 2016). Few of these methods have been applied extensively to experiments, while others have been used only few times and little is known about their practical usability (Adler *et al.*, 2007; Bimler *et al.*, 2018; Zhao *et al.*, 2016; Spaak *et al.*, 2020; Carmel *et al.*, 2017). Despite the proliferation of different methods to assess niche and fitness differences, there is very little guidance on the insights offered by different methods on the mechanisms underpinning coexistence (Spaak *et al.*, 2020, 2021; Godwin *et al.*, 2020).

Optimally, all these different methods to compute niche and fitness differences would lead to the same insights. That is, independent of which method we apply to our data, we would come to the same conclusion regarding, for instance, which traits increase niche or fitness differences (Kraft *et al.*, 2015; Pérez-Ramos *et al.*, 2019), how phylogenetic distance affects niche and fitness differences (Narwani *et al.*, 2013; Germain *et al.*, 2016), or how the presence of trophic interactions increases fitness differences (Petry *et al.*, 2018; Terry & Lewis, 2020).

However, here we show that this is not always the case. We do so by analysing three data sets (Letten *et al.*, 2017; Pérez-Ramos *et al.*, 2019; Germain *et al.*, 2016), using four different methods to compute niche and fitness differences. This exercise shows that different methods can lead to different conclusions regarding the putative

mechanisms driving coexistence. The first data set consisted of simulation output from a virtual experiment in which we investigated how changes in resource availability affect niche and fitness differences, similar to the work performed by Letten *et al.* (2017). We find that the different methods often do not agree on how resource availability affects niche and fitness differences. The second data set contains empirically measured interaction strengths among, and traits of, annual plant species. We asked which traits best predict niche and fitness differences, similar to the work of Pérez-Ramos *et al.* (2019). We find that different methods lead to different conclusions on which traits best predict niche and fitness differences. The third data set contains empirically measured interaction strengths among, and phylogeny information of, annual plants. We assess under which conditions phylogenetic distances hold information about niche or fitness differences, similar to the work of Germain *et al.* (2016). Again, we find that the method used to assess niche and fitness difference influences the answer to this question.

Methods

Methods to assess niche and fitness differences

There are as many as 13 different methods to compute niche and fitness differences (Spaak *et al.*, 2020; Koffel *et al.*, 2021; Johnson, 2021). In the current work, we focus on those methods that are most widely applied and/or are applicable to most data. The original definition of the four selected methods can be found in Chesson & Kuang (2008); Carroll *et al.* (2011); Zhao *et al.* (2016) and Spaak *et al.* (2020). Specifically, we chose these four methods as they are the only ones that can be applied to all three investigated datasets (see below).

For better comparison between the four methods we have slightly adjusted them. We changed the definitions such that a community with no niche differentiation will indeed have zero niche differences and increasing the stabilizing mechanisms will increase niche differences. We changed the definition of fitness differences such that

the two neutral species have fitness differences of 1. Here, we assigned a fitness difference to each species, i.e. $\mathcal{F}_i$, and the species with fitness difference below 1 is the competitive superior species, which would competitively exclude the other species in the absence of niche differences. In all cases this was achieved by a linear transformation of the original definitions, which does not affect the underlying conceptual ideas of the methods.

The first and original method to compute niche and fitness differences is based on a Lotka-Volterra model, which is derived from an underlying MacArthur consumer-resource model with the potential inclusion of a higher trophic level (Chesson, 1990; Chesson & Kuang, 2008; Chesson, 2013).

We assume a two-species Lotka-Volterra model $\frac{1}{N_i} \frac{dN_i}{dt} = \mu_i - \sum_j \alpha_{ij} N_j$, where N_i is the density of species i , μ_i is the intrinsic growth rate of species i and α_{ij} is the per-capita effect of species j on species i . For this model, the niche and fitness differences are defined as

$$\mathcal{N} = 1 - \sqrt{\frac{\alpha_{ij}\alpha_{ji}}{\alpha_{ii}\alpha_{jj}}} \quad (1)$$

$$\mathcal{F}_i = \frac{\mu_j}{\mu_i} \sqrt{\frac{\alpha_{ij}\alpha_{ii}}{\alpha_{ji}\alpha_{jj}}} \quad (2)$$

Usually, niche differences are described as $1 - \rho$ in the literature, where $\rho = \sqrt{\frac{\alpha_{ij}\alpha_{ji}}{\alpha_{ii}\alpha_{jj}}}$ is the niche overlap. We did not alter the definition of this method (Godoy & Levine, 2014; Letten *et al.*, 2017; Chesson, 2013). This method is originally only defined for a two species Lotka-Volterra model, however, it can be applied to other community models by reducing a given community model to Lotka-Volterra form. Specifically, Letten *et al.* (2017) have adapted this method to the Tilman consumer-resource model and Godoy & Levine (2014) have adapted this method to the Beverton-Holt annual plant model. We effectively use their computations of niche and fitness differences for the examples below. We will refer to this method as the *Original* method to assess niche and fitness differences.

The first method that we have introduced is model specific, all other methods

can be applied to a larger range of community models and can all be interpreted as a decomposition of the invasion growth rates. We therefore first introduce the notion of invasion growth rates. We assume a community model of the form $\frac{1}{N_i} \frac{dN_i}{dt} = f_i(N_i, N_j)$. We assume that both species have a stable monoculture equilibrium density, denoted N_i^* and N_j^* , respectively. The invasion growth rate of species i is the growth rate when it is at low density (mathematically zero) and its competitor at their monoculture density, i.e. the invasion growth rate is denoted as $f_i(0, N_j^*)$. Another important growth rate is the intrinsic growth rate when both species are at low density, i.e. $f_i(0, 0)$.

The second method, as defined by Carroll *et al.* (2011), uses the sensitivity of each species to competition, i.e. $S_i = 1 - \frac{f_i(0, N_j^*)}{f_i(0, 0)}$. With this they define

$$\mathcal{N} = 1 - \sqrt{S_i S_j} \quad (3)$$

$$\mathcal{F}_i = \sqrt{\frac{S_i}{S_j}} \quad (4)$$

The original work uses the notation ND and RFD for niche and fitness differences, respectively. The only difference is that RFD in the original studies corresponds to the maximum of $\mathcal{F}_i$ and $\mathcal{F}_j$. Importantly, this definition yields identical results for the Lotka-Volterra model as the original method. However, it is not equivalent to the derivations of Godoy & Levine (2014) or Letten *et al.* (2017) for their respective models. We will refer to this method as the *geometrical* method, because it computes niche and fitness differences as the geometrical mean and variance of the sensitivities, S_i .

The third method, by Zhao *et al.* (2016), define niche and fitness differences as

$$\mathcal{N} = \frac{f_i(0, N_j^*) + f_j(0, N_i^*)}{2} \quad (5)$$

$$\mathcal{F}_i = 1 + \log \left(\frac{N_i^*}{N_j^*} \right) \quad (6)$$

We have adjusted their definition in the following way $\mathcal{N} = (-\rho)/2$. where ρ is their definition of niche overlap and $\mathcal{F}_i = 1 + FD'_i$, where FD'_i is the original definition by Zhao *et al.* (2016); this was done to ensure that a neutral community would indeed have zero niche differences. The method by Zhao *et al.* (2016) is not equivalent to the original method for the Lotka-Volterra model. We will refer to this method as the *arithmetical* method as they compute niche differences as the arithmetic mean of the invasion growth rates.

The last method we investigate is the method by Spaak *et al.* (2020), who defined $\mathcal{N}$ and $\mathcal{F}$ as

$$\mathcal{N}_i = \frac{f_i(0, N_j^*) - f_i(c_{ij}N_j^*, 0)}{f_i(0, 0) - f_i(c_{ij}N_j^*, 0)} \quad (7)$$

$$\mathcal{F}_i = 1 - \frac{f_i(c_{ij}N_j^*, 0)}{f_i(0, 0)} \quad (8)$$

Where c_{ij} is a conversion factor that converts densities of species j to densities of species i , while keeping the converted density constant. c_{ij} is the solution of the equation $|1 - \mathcal{N}_i| = |1 - \mathcal{N}_j|$. The current work does not include any facilitation, i.e. $\mathcal{N}_i > 1$, which implies $\mathcal{N}_i = \mathcal{N}_j$, we therefore drop the subscript i . We have adjusted their definition in the following way, $\mathcal{F}_i = 1 - \mathcal{F}'_i$, where $\mathcal{F}'_i$ is the original definition by Spaak *et al.* (2020).

Equation 7 compares the actual invasion growth rate of species i ($f_i(0, N_j^*)$) to the hypothetical invasion growth rate when the two species would not interact ($f_i(0, 0)$) and to another hypothetical invasion growth rate when the two species had no niche differences ($f_i(c_{ij}N_j^*, 0)$). Importantly, this definition yields identical results for the Lotka-Volterra model as the original method, similar to the method by Carroll *et al.*

(2011). We will refer to this method as the *species-specific* method, as niche and fitness differences are species-specific properties and not community specific properties in this method. For a more in detail explanation of all these methods we refer to the original papers (Carroll *et al.*, 2011; Zhao *et al.*, 2016; Spaak *et al.*, 2020).

Examples

We have chosen three studies from the literature, and performed similar analysis of their data, but with four instead of only one definition for niche and fitness differences. We have chosen these three studies because they belong to the core literature of modern coexistence theory. As such, we assume and hope that many readers will be familiar with their work. Importantly, we do not re-analyze the data of those three studies identically as described in the original study, but simplified (e.g., random effect structure, error distributions). As such, our results cannot be directly compared with the original material even if using the same method, but rather, can be used to compare how interpretations may differ using different estimates of niche and fitness differences. We had no a-priori expectations on the robustness of the results of these three studies, and we do not want to imply that our reevaluation of their results should cast doubt on them. In fact, the three studies are all more complex and insightful than presented here, we invite any reader not familiar with their work to read the original papers.

Theoretical example

Letten *et al.* (2017) investigated how changes of parameters in a two-species consumer-resource model affect niche and fitness differences. They used a Tilman consumer-resource model with substitutable resources with a Holling type 2 response. The

community dynamics for the substitutable resources are given by

$$\frac{1}{N_i} \frac{dN_i}{dt} = \frac{\sum_l w_{il} R_l}{k_i + \sum_l w_{il} R_l} - m_i \quad (9)$$

$$\frac{dR_l}{dt} = S_l - R_l - \sum_j u_{lj} N_j \quad (10)$$

where w_{il} is the conversion of resource l to biomass of species i , u_{li} is the utilization of resource l by species i , m_i is the mortality rate, S_l is the resource supply and k_i is half-saturation constant.

Letten *et al.* (2017) investigated how niche and fitness differences change in response to a change in all of the parameters (u_{li}, w_{il}, k_i, S_l). We will, however, only investigate how niche and fitness differences change as a function of S_l . Specifically, we assess how a change in resource supply rates S_l increases or decreases niche and fitness differences.

Additionally, we investigate whether the different methods agree on which species is the superior competitor and whether the different methods agree on the frequency dependence of the system. The superior competitor is defined as the species with $\mathcal{F}_i < 1$. The frequency dependence of the system is defined by the sign of the niche differences $\mathcal{N}$: negative niche differences correspond to *positive* frequency dependence and positive niche differences correspond to *negative* frequency dependence. $\mathcal{N} = 0$ indicates the absence of frequency dependence.

Empirical examples

Many different papers have tried to identify predictors of niche and fitness differences, such as phylogenetic distance (Narwani *et al.*, 2013; Germain *et al.*, 2016), the presence of predators (Chesson & Kuang, 2008; Petry *et al.*, 2018; Terry & Lewis, 2020) or species traits (Pérez-Ramos *et al.*, 2019; Kraft *et al.*, 2015; Gallego *et al.*, 2019). We chose to perform an analysis similar to that in Pérez-Ramos *et al.* (2019) (i.e., which trait differences are most important for predicting niche and fitness differences) and Germain *et al.* (2016) (i.e., under what circumstances does phylogenetic distance pre-

dict niche and fitness differences). We chose Pérez-Ramos *et al.* (2019) and Germain *et al.* (2016) as they offer a large and open data set with many species and many predictors of niche and fitness differences.

Pérez-Ramos *et al.* (2019) investigated whether differences in species traits translate to niche or fitness differences. Pérez-Ramos *et al.* (2019) parameterized pairwise competition models for 10 annual plant species (a total of 45 two-species annual plant communities). Additionally, Pérez-Ramos *et al.* (2019) measured functional traits for each species and correlated species' differences in traits to their niche and fitness differences.

Germain *et al.* (2016) investigated under which circumstance phylogenetic distance can be used as a predictor of niche or fitness differences. Germain *et al.* (2016) investigated 30 annual plants and fitted an annual plant model to 20 two-species annual plant communities under two different environmental conditions (wet versus dry), leading to a total of 40 annual plant community models. Half of the species pairs were found in similar habitats (sympatric) and half were found in different habitats (allopatric). Additionally, they measured the phylogenetic distance of the two competing species to assess under which conditions phylogenetic distance affects niche and fitness differences.

With each study's fully-parameterized annual plant model, we compute niche and fitness differences as defined by the original method using equations 1 and 2 and simulated the necessary growth rates for the other methods. We then perform a linear regression of the predictor (trait differences or phylogenetic distance between the two species) and the response variables, niche or fitness differences. For each linear regression we report the slope and the p-value.

Results

Theoretical example

First, we investigate two different communities of two species competing for two substitutable resources (Fig. 1, A,C). For each community, we changed the resource supply rates (from black triangle to black dot).

As observed before (Spaak *et al.*, 2020; Song *et al.*, 2019), the different methods have different quantitative values for $\mathcal{N}$ and $\mathcal{F}$. However, in addition to differences in their quantitative values, the different methods also yield qualitatively different interpretations of the change of resource supply rates.

For example, in the first community (Panels A,B), we might ask whether the decrease in resource 1 is most beneficial for the blue or the green species. If the increase of resource 1 benefits the blue species, consistent with resource ratio theory (Tilman, 1982), then fitness differences of the green species, $\mathcal{F}_g$, will increase. Conversely, $\mathcal{F}_g$ will decrease if it benefits the green species. Yet, the different methods give different results. The original, geometric and species-specific methods all consider the change in resource 1 as more beneficial for the blue species. In contrast, the arithmetic method interprets this change as most beneficial for the green species.

Alternatively, we might ask whether the community is driven by positive or negative frequency dependence, which would lead to negative or positive niche differences, respectively. Again, we find that the different methods do not always lead to the same conclusion. At the new resource supply point (black dot), the original method finds that the community is driven by negative frequency dependence, as niche differences are positive. On the other hand, the geometric and arithmetic method interpret that the community is driven by positive frequency dependence. Finally, the species-specific method interprets that frequency dependence is exactly zero.

In the next community (Panels C,D), we ask about the stabilizing effect of resource change. The original method finds that the change in resource supply does

not affect niche differences (Letten *et al.*, 2017). The arithmetic and geometric method find that the change is stabilizing (increase in $\mathcal{N}$), while the species-specific method interprets the change as destabilizing (decrease in $\mathcal{N}$). Alternatively, we might ask which species is the competitive superior. Again, we obtain different answers when using different methods. The original and arithmetic method conclude that the blue species is the competitive superior ($\mathcal{F}_g > 1$), while the geometric and the species-specific methods conclude that the green species is the competitive superior ($\mathcal{F}_g < 1$, albeit with close to no fitness difference).

Figure 1: A,C: Two examples of two species competing for substitutional resources. The solid lines indicate the zero-net growth isoclines (ZNGI), at which the species has zero growth. The dashed lines indicate the consumption vectors u_{li} of the species. The black triangle and dot show the resource supply rates before and after an environmental change. B,D: Corresponding niche and fitness differences of the green species for the four investigated methods (colors). In certain cases, the different methods do not agree qualitatively on whether: 1. the community is driven by positive ($\mathcal{N} < 0$, +fd), negative ($\mathcal{N} > 0$, -fd) or no frequency dependence (Panel B, dot). 2. the blue species ($F_g > 1$) or the green species ($F_g < 1$) is the competitive superior species (Panel D, triangle). 3. the change in resource supply rates is more beneficial for the blue (higher F_g for the dot than for the triangle) or green species (lower F_g for the dot than for the triangle, Panel B); 4. the change in resource supply rates increases, decreases or has no effect on niche differences (Panel D, change in $\mathcal{N}$) This figure is conceptually similar to Figure 2 from Letten *et al.* (2017). For a more quantitative assessment of these cases see Figure 2.

Admittedly, we have selected communities to illustrate our message that different methods can give different explanations as to why species coexist. We will now show that these are not (rare) special cases and that disagreement among methods is a persistent feature. To this end, we randomly generated 1000 two-species communities competing for substitutable resources (see Appendix S1) and computed their niche

and fitness differences (Figure 2A). We then checked whether the methods agree qualitatively on frequency dependence ($\mathcal{N} > 0$ or $\mathcal{N} < 0$) and the identity of the competitive dominant species ($\mathcal{F}_i > 1$ or $\mathcal{F}_i < 1$) of the community at the initial resource supply rate. This quantitative analysis confirms our qualitative examples; often, the methods did not agree well (red dots, Figure 2A) and they did not agree much better than expected by chance (Appendix S1).

Next, we changed the resource supply rates and computed niche and fitness differences for the changed resource supply rate. We assessed whether the methods agree on the stabilizing effect (increase or decrease in $\mathcal{N}$) and on the equalizing effect (increase or decrease in $\mathcal{F}_i$) of the change in resource supply rate (Figure 2B). Again, the different methods did not agree well on how changes in resource supply rates affect niche and fitness differences.

The only case with 100% agreement was the one where the geometric and the species-specific method identify the superior competitor, i.e. $\mathcal{F}_i > 0$ or $\mathcal{F}_i < 0$. We show in the appendix why this is always the case (Appendix S3). For all other cases, the specific percentages depend on the specifics of the randomly generated communities, yet the qualitative picture, that the methods do not always agree is very consistent (Appendix S1). Changes in resource supply rates do not affect niche differences when the original method was applied to the model in equation 9 and 10 (Letten *et al.*, 2017), whereas it can affect niche differences for the other methods.

Figure 2: A: For 1000 randomly generated communities we computed niche (above diagonal, white panels) and fitness differences (below diagonal, grey panels) according to the four investigated methods (diagonal). Each dot corresponds to one randomly generated community. We say that two methods agree qualitatively if they agree on the frequency dependence ($\mathcal{N} > 0$ or $\mathcal{N} < 0$) and superior competitor ($\mathcal{F}_i > 1$ or $\mathcal{F}_i < 1$). Communities where the methods agree qualitatively are shown in blue, communities where the methods disagree are shown in red. The top-left number indicates how often the two methods agree qualitatively (fraction of blue dots). B: Similar to A, but we compare the change of niche and fitness differences after a change in resource supply rates. We say that two methods agree qualitatively when niche or fitness differences increase or decrease for both methods. A,B: In general the different methods are well correlated. However, their qualitative agreement (top left number = fraction of blue points) is often not much higher than expected by chance (see Appendix S1).

Empirical examples

An important empirical application of modern coexistence theory is to link traits to niche and fitness differences. Doing so is important to provide insight into which processes govern coexistence in nature (Kraft *et al.*, 2015; Gallego *et al.*, 2019; Germain *et al.*, 2016; Narwani *et al.*, 2013; Godoy & Levine, 2014). To do so, one typically measures the relationship between traits and niche and fitness differences. For example, if we hypothesize that species with different root length will consume resources at dif-

ferent locations (Levine & HilleRisLambers, 2009), we can test the correlation between differences in root length and niche differences. Alternatively, we might hypothesize that species with longer roots will have a competitive advantage, as they can take up more resources, which can be tested by quantifying the correlation between species differences in root length and fitness differences.

We computed niche and fitness differences for competing annual plant species with the four methods mentioned above. We then tested which trait differences were correlated to niche and fitness differences (Fig. 3). The different methods yielded different quantitative and qualitative results about the importance of different traits.

Quantitatively, different methods predicted different p-values (Fig. 3 A,C) and different slopes (Fig. 3 B,D) of the relationships between traits and niche and fitness differences. However, these differences are also qualitative and offer different answers to the question which traits drive coexistence. For example, only the geometric and species-specific method find differences in leaf dry matter content as significant predictors for niche differences. Similarly, only the original and arithmetic method find differences in peak productivity as significant predictors for fitness differences. Additionally, sometimes the slopes between the different methods are different, as can be seen by the example of the specific leaf area for niche differences. However, in the investigated dataset, we never observed two methods predicting qualitatively different slopes that were both significant.

Figure 3: The p-value (A,C) and the scaled slope (B,D) of the linear regression between the pairwise trait differences (x-axis) and the niche (A,B) and fitness differences (C,D) for each of the four investigated methods (y-axis) for the Pérez-Ramos *et al.* (2019) data set. We find many quantitative and qualitative differences among these linear regressions. For example, the original method predicts that leaf carbon content is a good predictor of fitness differences (cyan-star indicates p-value < 0.05) while none of the other methods do. Whether the experiments would have been analyzed by the original method or not will therefore affect our understanding of the system. We re-scaled the slopes of the linear regression for illustration purposes, for each method. The slopes were scaled such that the absolute value of the maximal slope is 1. Importantly, this does not qualitatively alter the slopes, as their sign and relative importance are not altered.

In our third example we investigate in which scenarios phylogenetic distance is a good predictor of niche or fitness differences. Germain *et al.* (2016) conducted experiments with four different scenarios in a two-by-two factorial design: whether the species evolved under sympatry or allopatry, and whether the environment was wet or dry.

Again, we find that the different methods yield different results. For example, phylogenetic distance of species is a good predictor of niche and fitness differences for allopatrically-evolved species, when these two components are as measured by the original and arithmetic method, but not by the others. Conversely, phylogenetic distance of species that evolved sympatrically is a good predictor of fitness differences only for the geometrical and species-specific method.

Figure 4: Similar to figure 3, we report the p-value (A,C) and the relative slope (B,D) of the linear regression between the pairwise phylogenetic distance (x-axis) and the niche (A,B) and fitness differences (C,D) for each of the four investigated methods (y-axis) for the Germain *et al.* (2016) data set. Again, we find that different methods lead to qualitatively different conclusions about in which scenario phylogenetic distance predicts niche or fitness differences. The first column (Phylogenetic distance; PD) reports the linear regression based on all communities, while the second and third column does so for communities in wet and dry conditions, respectively. The fourth and fifth column reports the results for sympatric and allopatric communities, respectively.

For the two empirical datasets investigated here, we computed how often they agree qualitatively. That is, how often do two different methods give the same prediction of whether a given trait or phylogenetic distance significantly predicts niche or fitness differences (Fig.5). On the one hand, the different methods agree fairly well (agreement ranging from 55% – 82% for niche differences and 65% – 96% for fitness differences). On the other hand, most of the methods do not agree much better than expected by chance (60% – 65% and 50% – 62% for fitness differences). In general, the methods seem to agree more for fitness differences than for niche differences. The geometric and the species-specific methods agree best in general.

Figure 5: We report how often the two methods agree on whether a certain trait (A,B) or phylogenetic distance (C,D) is significant for niche (A,C) or fitness (B,D) differences. We compare how often they agree (below diagonal) to how often we would expect them to agree based on chance (above diagonal). While the methods, generally, agree fairly well, they do not agree much better than expected by chance.

Discussion

We have compared four different methods to compute niche and fitness differences on one theoretical and two empirical datasets. We found that the different methods can lead to different qualitative insights about the underlying ecology of the system.

For example, the methods did not agree on whether a change in resource supply rates is stabilizing or destabilizing. Similarly, they did not agree on which traits are most important for niche and fitness differences. We have shown that these differences are not special cases, but rather the norm. We found that the different methods agree only slightly better than expected by chance (Fig. 2 and 5).

It is well known that these two key concepts of modern coexistence theory (niche and fitness difference) are not defined universally. The presented results show that this variation can lead to different interpretations, confirming results by Song *et al.* (2019). These authors showed that stabilizing and equalizing mechanisms have two different meanings in modern coexistence theory, depending on whether one investigates a two-species community or a multi-species community. We have generalized their work and showed that even for simple two species communities different methods can produce different assessments of stabilizing and equalizing mechanisms. There are other methods to assess niche and fitness differences which we have not examined. We expect that including other methods would increase this variation as different methods have different requirements and capture different aspects of coexistence.

Perspectives on quantifying niche and fitness differences

Intuitively, the different methods lead to different interpretations of coexistence because they capture different aspects of ecology. This is mirrored by which information is included into the calculation of niche differences (Appendix S3). The original method includes inter- and intra-specific species interactions; the arithmetic method includes invasion growth rates; the geometric method includes invasion and intrinsic growth rates; and the species-specific method includes invasion, intrinsic and no-niche growth rates. These conceptual differences show that different information is included in the computation of niche and fitness differences, which lead to qualitatively and quantitatively different interpretations of coexistence.

For example, a change in resource supply rates that increases invasion growth

rates, e.g. increasing resource supply rates at fixed ratios, will increase niche differences according to the arithmetic method. However, this change in resource supply rates will likely also affect the other growth rates, e.g. increase intrinsic growth rate and decrease no-niche growth rate (Appendix S2). Whether niche differences according to the geometric and species-specific method also increase depends on the magnitude and direction of the change in these growth rates. How different methods interpret changes in resource supply rates therefore mirror how the intrinsic, invasion and no-niche growth rates are affected by these resource supply rate changes, as well as whether this information is included. A similar understanding of the empirical data-sets can explain how different methods interpret different traits or phylogenetic distances as predictors of niche and fitness differences (Appendix S2).

We found that different methods yield different niche and fitness differences, and different values for stabilizing and equalizing mechanisms. We argue that identifying which method is the superior one is both impossible and unwarranted. Instead, we feel that the different methods considered merely reflect different interpretations of these notoriously slippery concepts. The objective is therefore not to seek hierarchy, but understanding across methods, much like what is done in other research fields. Indeed, variation of interpretation, and the resulting methodological variation, is found back in other research fields as well. For instance, concepts like stability and biodiversity can be translated in various ways into mathematical formulae, depending on what one considers a “stable” or “diverse” ecosystem, or which aspects one wishes to highlight (Montoya *et al.*, 2018; Donohue *et al.*, 2016; McGill *et al.*, 2015). In stability research, this variation has been embraced and coined the dimensionality of stability (Radchuk *et al.*, 2019; Donohue *et al.*, 2016). Understanding why these methods differ and when they correspond, i.e. when dimensionality is reducible, is a major objective in stability research (Radchuk *et al.*, 2019; Arnoldi *et al.*, 2018; Clark *et al.*, 2021; Carpentier *et al.*, 2021). We argue that similar efforts are needed in modern coexistence theory, if we want to gain synthetic understanding across published results adopting a variety of methods.

Additionally, the best-suited method to quantify certain aspects of niche and fitness differences may depend on the particular empirical system Godwin *et al.* (2020). Of the methods examined here, some require parameterizing a demographic model (Chesson & Kuang, 2008; Chesson, 2013) whereas others require measuring species' invasion growth rate (Carroll *et al.*, 2011; Zhao *et al.*, 2016; Carmel *et al.*, 2017). While the latter methods are potentially model-independent, they are only better suitable when the monoculture equilibrium can be ensured and the invasion growth rate can be quantified within a feasible time frame; these criteria likely explain the prevalence of such methods when studying the coexistence of microorganisms (e.g. Narwani *et al.*, 2013; Grainger *et al.*, 2019; Li, 2017). On the other hand, quantifying niche and fitness differences for long-lived organisms, e.g. plants, were mainly achieved by parameterizing demographic models, e.g. Godoy & Levine (2014); Chu & Adler (2015). Although labour-intensive, we would argue that parameterizing a demographic model might be the most flexible approach when the goal is to assess niche and fitness differences with multiple methods. Indeed, we were able to calculate multiple metrics from the Pérez-Ramos *et al.* (2019) and Germain *et al.* (2016) data set precisely because they provided a fully-parameterized annual plant model, from which the original method can be calculated and invasion growth rates required by other methods can be simulated.

We therefore propose that future work in modern coexistence theory should consider applying multiple methods to assess niche and fitness differences. Often, one can with little to no additional effort compute multiple methods of niche and fitness differences with the same data (or with an underlying model). Computing multiple niche and fitness differences on the same data will help the individual researcher as well as the field of modern coexistence theory.

We see the two following benefits for the field of modern coexistence theory. First, we start to understand under which conditions the dimensionality of niche and fitness differences is high. This advances our understanding of when the different methods agree well and might increase our understanding of when intrinsic growth

rates or no-niche growth rates are important for coexistence. Second, new methods to compute niche and fitness differences are continuously being developed (Spaak *et al.*, 2020, 2021; Koffel *et al.*, 2021; Johnson, 2021), however, these are rarely applied empirically. Applying multiple methods to one data set will allow us to understand the differences and potential benefits of newer methods earlier.

Additionally, each individual researcher benefits from applying multiple methods to their data. First, the different methods highlight different aspects of coexistence. The correctness of a certain hypothesis might depend on which part of coexistence is highlighted. By applying multiple methods, and therefore highlighting different aspects, one is more likely to highlight the correct aspect as well. Of course, this should not be misused for p-hacking or similar practices. Second, by highlighting different aspects and understanding under which scenario a hypothesis is correct one will also obtain a deeper understanding of coexistence and of one's own data. For example, by which growth rate is the correctness of a hypothesis driven? Third, niche and fitness differences are difficult to grasp and each of the currently available methods describes them from a different perspective. By applying multiple methods we may get a better grasp of them. For example, phylogenetic distance in sympatric evolution is a good predictor of niche differences according to three different methods (Fig. 4 A). We can therefore be more certain that phylogenetic distance in sympatric evolution indeed affects niche differentiation.

Acknowledgements

J.W.S received support from Schweizerischer Nationalfonds Early Post-doc mobility (P2SKP3.194960) F.D.L. received support from grant of the Fund for Scientific Research, FNRS (PDR T.0048.16). P.-J. K. received support from the Yushan scholar program of Taiwan MOE (NTU-110VV010). We thank Rachel Germain and Oscar Godoy for the shared data on the annual plant communities. We thank Rachel Germain and Oscar Godoy for comments on earlier versions of this manuscript.

References

- Adler, P.B., HilleRisLambers, J. & Levine, J.M. (2007). A niche for neutrality. *Ecology Letters*, 10, 95–104.
- Arnoldi, J.F., Bideault, A., Loreau, M. & Haegeman, B. (2018). How ecosystems recover from pulse perturbations: A theory of short- to long-term responses. *Journal of Theoretical Biology*, 436, 79–92.
- Barabás, G., D’Andrea, R. & Stump, S.M. (2018). Chesson’s coexistence theory. *Ecological Monographs*, 88, 277–303.
- Bimler, M.D., Stouffer, D.B., Lai, H.R. & Mayfield, M.M. (2018). Accurate predictions of coexistence in natural systems require the inclusion of facilitative interactions and environmental dependency. *Journal of Ecology*, 106, 1839–1852.
- Carmel, Y., Cornell, S.J., Belmaker, J., Suprunenko, Y.F., Kent, R., Kunin, W.E. & Bar-Massada, A. (2017). Using exclusion rate to unify niche and neutral perspectives on coexistence. *Oikos*, 126, 1451–1458.
- Carpentier, C., Barabás, G., Spaak, J.W. & De Laender, F. (2021). Reinterpreting the relationship between number of species and number of links connects community structure and stability. *Nature Ecology & Evolution*.
- Carroll, I.T., Cardinale, B.J. & Nisbet, R.M. (2011). Niche and fitness differences relate the maintenance of diversity to ecosystem function. *Ecology*, 92, 1157–1165.
- Chesson, P. (1990). MacArthur’s consumer-resource model. *Theoretical Population Biology*, 37, 26–38.
- Chesson, P. (2000). Mechanisms of maintenance of speciesdiversity. *Annual Review of Ecology, Evolution, and Systematics*, 31.
- Chesson, P. (2003). Quantifying and testing coexistence mechanisms arising from recruitment fluctuations. *Theoretical Population Biology*, 64, 345–357.

- Chesson, P. (2013). Species Competition and Predation. In: *Ecological Systems*. Springer New York, pp. 223–256.
- Chesson, P. (2018). Updates on mechanisms of maintenance of species diversity. *Journal of Ecology*, 106, 1773–1794.
- Chesson, P. & Kuang, J.J. (2008). The interaction between predation and competition. *Nature*, 456, 235–238.
- Chu, C. & Adler, P.B. (2015). Large niche differences emerge at the recruitment stage to stabilize grassland coexistence. *Ecological Monographs*, 85, 373–392.
- Clark, A.T., Arnoldi, J.F., Zelnik, Y.R., Barabas, G., Hodapp, D., Karakoç, C., König, S., Radchuk, V., Donohue, I., Huth, A., Jacquet, C., de Mazancourt, C., Mentges, A., Nothaaß, D., Shoemaker, L.G., Taubert, F., Wiegand, T., Wang, S., Chase, J.M., Loreau, M. & Harpole, S. (2021). General statistical scaling laws for stability in ecological systems. *Ecology Letters*, pp. 1–13.
- Donohue, I., Hillebrand, H., Montoya, J.M., Petchey, O.L., Pimm, S.L., Fowler, M.S., Healy, K., Jackson, A.L., Lurgi, M., McClean, D., O'Connor, N.E., O'Gorman, E.J. & Yang, Q. (2016). Navigating the complexity of ecological stability. *Ecology letters*, 19, 1172–1185.
- Ellner, S.P., Snyder, R.E., Adler, P.B. & Hooker, G. (2018). An expanded modern coexistence theory for empirical applications. *Ecology Letters*.
- Gallego, I., Venail, P. & Ibelings, B.W. (2019). Size differences predict niche and relative fitness differences between phytoplankton species but not their coexistence. *ISME Journal*.
- Germain, R., Weir, J. & Gilbert, B. (2016). Species coexistence: macroevolutionary patterns and the contingency of historical interactions. *PRSB*, 283, 20160047.
- Godoy, O. & Levine, J.M. (2014). Phenology effects on invasion success: Insights from coupling field experiments to coexistence theory. *Ecology*, 95, 726–736.

- Godwin, C.M., Chang, F.H. & Cardinale, B.J. (2020). An empiricist's guide to modern coexistence theory for competitive communities. *Oikos*.
- Grainger, T.N., Letten, A.D., Gilbert, B. & Fukami, T. (2019). Applying modern coexistence theory to priority effects. *Proceedings of the National Academy of Sciences*, 116, 6205–6210.
- Hubbell, S.P. & Knapp, S. (2003). *The Unified Neutral Theory of Biodiversity and Biogeography*. vol. 110. Princeton University Press, New Jersey.
- Hutchinson, G.E. (1959). HOMAGE TO SANTA ROSALIA or WHY ARE THERE SO MANY KINDS OF ANIMALS?* G. *The american Naturalist*, XCIII.
- Johnson, C.A. (2021). How mutualisms influence the coexistence of competing species. *Ecology*, 102, 1–8.
- Ke, P.J. & Letten, A.D. (2018). Coexistence theory and the frequency dependence of priority effects. *bioRxiv*.
- Koffel, T., Daufresne, T. & Klausmeier, C.A. (2021). From competition to facilitation and mutualism: a general theory of the niche. *Ecological Monographs*.
- Kraft, N.J.B., Godoy, O. & Levine, J.M. (2015). Plant functional traits and the multidimensional nature of species coexistence. *Proceedings of the National Academy of Sciences*, 112, 797–802.
- Letten, A.D., Ke, P.J. & Fukami, T. (2017). Linking modern coexistence theory and contemporary niche theory. *Ecological Monographs*, 87, 161–177.
- Levine, J.M. & HilleRisLambers, J. (2009). The importance of niches for the maintenance of species diversity. *Nature*, 461, 254–257.
- Li, J. (2017). Assessing the accuracy of predictive models for numerical data : Not r nor r² , why not ? Then what ? *Plos one*, pp. 1–16.

- Mcgill, B.J., Dornelas, M., Gotelli, N.J. & Magurran, A.E. (2015). Fifteen forms of biodiversity trend in the Anthropocene. *Trends in Ecology & Evolution*, 30, 104–113.
- Montoya, J.M., Donohue, I. & Pimm, S.L. (2018). Planetary Boundaries for Biodiversity: Implausible Science, Pernicious Policies. *Trends in Ecology and Evolution*, 33, 71–73.
- Narwani, A., Alexandrou, M.A., Oakley, T.H., Carroll, I.T. & Cardinale, B.J. (2013). Experimental evidence that evolutionary relatedness does not affect the ecological mechanisms of coexistence in freshwater green algae. *Ecology Letters*, 16, 1373–1381.
- Pérez-Ramos, I.M., Matías, L., Gómez-Aparicio, L. & Godoy, Ó. (2019). Functional traits and phenotypic plasticity modulate species coexistence across contrasting climatic conditions. *Nature Communications*, 10, 1–11.
- Petry, W.K., Kandlikar, G.S., Kraft, N.J., Godoy, O. & Levine, J.M. (2018). A competition–defence trade-off both promotes and weakens coexistence in an annual plant community. *Journal of Ecology*, 106, 1806–1818.
- Radchuk, V., Laender, F.D., Cabral, J.S., Boulangeat, I., Crawford, M., Bohn, F., Raedt, J.D., Scherer, C., Svenning, J.C., Thonicke, K., Schurr, F.M., Grimm, V. & Kramer-Schadt, S. (2019). The dimensionality of stability depends on disturbance type. *Ecology Letters*, 22, 674–684.
- Saavedra, S., Rohr, R.P., Bascompte, J., Godoy, O., Kraft, N.J. & Levine, J.M. (2017). A structural approach for understanding multispecies coexistence. *Ecological Monographs*, 87, 470–486.
- Schreiber, S.J. (2000). Criteria for robust permanence. *Journal of Differential Equations*, 162, 400–426.
- Song, C., Barabás, G. & Saavedra, S. (2019). On the consequences of the interdependence of stabilizing and equalizing mechanisms. *The American Naturalist*, pp. 000–000.

- 576 Spaak, J.W., De Laender, F. & Laender, F.D. (2020). Intuitive and broadly ap-
 577 plicable definitions of niche and fitness differences. *Ecology Letters*, p. doi:
 578 <https://doi.org/10.1101/482703>.
- 579 Spaak, J.W., Godoy, O. & De Laender, F. (2021). Mapping species niche and fitness
 580 differences for communities with multiple interaction types. *Oikos*, pp. 1–13.
- 581 Terry, J.C.D. & Lewis, O.T. (2020). Finding missing links in interaction networks.
 582 *Ecology*, 0, 1–13.
- 583 Tilman, G.D. (1982). *Resource Competition and Community Structure*. Princeton univer-
 584 sity press.
- 585 Zhao, L., Zhang, Q.G. & Zhang, D.Y. (2016). Evolution alters ecological mechanisms
 586 of coexistence in experimental microcosms. *Functional Ecology*, 30, 1440–1446.

S1 Quantitative comparisons of methods for resource limitation

To create Figure 2 from the main text we created 1000 two-communities competing for two substitutional resources, given by the following per-capita growth rates

$$\frac{1}{N_i} \frac{dN_i}{dt} = \frac{\sum_l w_{il} R_l}{k_i + \sum_l w_{il} R_l} - m_i \quad (\text{S11})$$

$$\frac{dR_l}{dt} = S_l - R_l - \sum_j u_{lj} N_j \quad (\text{S12})$$

The species parameters were chosen randomly as given by table S1. For each two species community we computed niche and fitness differences in two different environmental settings. The two settings differed only in the resource supply rates S_l , the distribution of the resource supply rates are also given in table S1. The parameter values were chosen randomly, there's nothing special about the given parameter values and they do not necessarily represent realistic parameter values. We performed similar analysis with other parameter values and got qualitative similar results.

There are three qualitatively different effects the change of resource supply rates can have on niche or fitness differences: increase, decrease or have no effect. Figure 2 B show how often two different methods agree on the sign of the change in niche or fitness differences. To compute how often we would expect them to agree by chance we computed the probabilities $P(x|\text{method})$ where x is the sign of the change of niche or fitness differences ($x \in \{-1, 0, +1\}$) and method is any of the four investigated methods (Fig. S2). The base probability that the two methods agree is given by $\sum_{x \in \{-1, 0, 1\}} P(x|\text{method}_1)P(x|\text{method}_2)$. For change in niche differences this probability was usually around 1/3, which is what we would expect naively, as there are 3 different possible outcomes. For change in fitness differences this probability was quite exactly 1/2. We never observed that a change in resource supply rates does not affect fitness differences, leaving effectively only two possible outcomes.

Variable	Distribution
w_{12}, w_{21}	$[2, 10]$
w_{11}, w_{22}	$[6, 30]$
$u_{11}, u_{12}, u_{21}, u_{22}$	$[0.001, 0.01]$
$k_1 = k_2$	$[1, 2]$
$m_1 = m_2$	$[0.1, 0.3]$
S_1^I, S_2^I S_1^F, S_2^F	$[0.2, 2]$

Table S1: All distributions were assumed to be uniform distributions. We assumed that species 1 is a better competitor for resource 1 and species 2 is a better competitor for resource 2, we therefore chose different distributions for w_{11} and w_{22} than for w_{12} and w_{21} . Additionally, we assumed that both species have identical mortality rate and identical half-saturation constants. This ensures that the zero net growth isoclines of the two species cross. All parameter were chosen arbitrarily. We investigated other parameter settings as well as competition for essential resources and got qualitatively similar results.

Similarly we computed the frequency dependence of the community (positive, negative or no frequency dependence) and the superior competitor (species 1, species 2 or equal fitness) for the initial resource supply rates. No-frequency dependence was only ever observed with the species-specific method, which happens for very unbalances resource supply rates, as both species drive one resource to extinction in mono-culture, effectively competing only for one resource. The other methods can lead to no change in niche differences, but the probability for no niche difference is zero. Similarly, we never observed equal fitness of both competitors, which is theoretically possible, but has zero probability of occurring. We again compared how often the different methods agreed and how often we would expect them to agree based on chance.

Figure S1: Similar to figure 5. A, C: How often two methods agree on how changes in resource supply rates affect niche differences (A) and fitness differences (C). Entries below diagonal show actual agreement, entries above diagonal show how much we would expect them to agree based on chance. B,D: How often two methods agree on the frequency dependence (B) and the superior competitor (D) for the initial resource supply point. Entries below diagonal show actual agreement, entries above diagonal show how much we would expect them to agree based on chance.

Figure S2: Base probabilities $P(x|method)$ for the different methods for niche and fitness differences investigated for change in niche differences (A), frequency dependence (B), change in fitness differences (C) and superior competitor (D). Colors indicate the three different possible outcomes (blue, orange and green correspond to: A and C: decrease, no change, increase, B: positive, zero and negative, D: species 1, similar fitness and species 2). Panel A corresponds to above diagonal of figure 2B, Panel B to above diagonal 2A, Panel C to below diagonal of Figure 1B, and Panel D to below diagonal of Figure 2A. A: For the original method no changes in resource supply rates affected niche differences, as proven by (Letten *et al.*, 2017). For the geometrical method we never observed no-change in niche differences. B: Only the species-specific predicts that some communities have no niche differentiation, this is the case when both species drive one resource to extinction in monoculture and are effectively limited by only on resource. C and D: We never observed that changes in resource supply rates have no effect on fitness differences or that both species have equal fitness. The differences from 0.5 probabilities stem from the randomness of the simulations.

621 S2 Growth rates and niche and fitness differences

Figure S3: A-C: Histograms of the mean intrinsic (A), invasion (B) and no-niche (C) growth rates for three different resource supply rate scenarios, corresponding to the reference case (red), increased resource supply at fixed ratios (blue) and changed resource supply rates but same total supply rates (yellow). Increasing resource supply rates has a strong positive effect on intrinsic growth rate, a smaller positive effect on invasion growth rates and a negative effect on the no-niche growth rate. Changing resource ratios has no effect on the intrinsic growth rates, but small effects invasion and no-niche growth rates. D-F: Histograms of the niche differences according to the geometric (D), arithmetic (E) and species-specific method (F). Different resource scenarios have different effects on the growth rates and therefore different effects on the niche differences according to the different methods. D: Increasing resource supply rates has a positive effect on both, intrinsic and invasion growth rates. However, the effect on the intrinsic growth rate is stronger, such that niche differences according to the geometric method are decrease with increasing resource supply rates. E: For the arithmetic method, niche differences depend only on the invasion growth rates, such that increasing resources increases niche differences. F: For the species-specific method, niche differences are lowest for the altered resource scenario, as this increases the no-niche growth rates and (slightly) decreases invasion growth rates. A-F: Vertical bars indicate the mean of the corresponding distribution.

Figure S4: Equivalent to figure 3, the p-value (A,C) and the relative slope (B,D) of the linear regression between the pairwise trait differences (x-axis) and the niche (A,B) and fitness differences (C,D) for each of the four investigated methods (y-axis) for the Pérez-Ramos *et al.* (2019) data set (above the black horizontal line). Below the black horizontal line we show the linear regressions for the mean of the intrinsic growth rate μ_i , the invasion growth rate r_i and the no-niche growth rate η_i . Intuitively, if a trait correlates with a growth rate, then we would expect that this trait also correlates with niche differences that include this growth rate. For example, the arithmetic method depends only on the invasion growth rate, consequentially, every trait that correlates with the invasion growth rates also correlates with the niche differences as computed by the arithmetic method. However, the different growth rates are also strongly correlated, which complicates the issue. For example, root density does not correlate with any growth rate, yet it does correlate with niche differences according to the species-specific method. Similarly, leaf nitrogen content correlates with all three growth rates, yet it does not correlate with the original, geometric or species-specific method.

S3 Theoretical comparison of the different methods

In the main text we reported that the different methods lead to different interpretations of niche and fitness differences and coexistence in general. We here give a short insight into where these differences stem from. We first analyse two simple and well

known community models (Lotka-Volterra model and Beverteon-Holt annual plant model) and then give some limited insight into the general case without a specific community model.

S3.1 Lotka-Volterra model

As in the main text, we assume a Lotka-Volterra community model of the form

$$\frac{1}{N_i} \frac{dN_i}{dt} = \mu_i - \alpha_{ij}N_j - \alpha_{ii}N_i \quad (\text{S13})$$

Niche and fitness differences according to the original method are given in equation 1 and 2. Carroll *et al.* (2011) have shown that the geometrical method gives identical values for niche and fitness differences as the original method for the Lotka-Volterra model. Similarly, Spaak *et al.* (2020) have shown that the species-specific method gives identical values for niche differences and fitness differences.

The arithmetic method differs from all these methods. To compute niche and fitness differences according to the arithmetic method we have to compute the monoculture equilibrium densities $N_i^* = \frac{\mu_i}{\alpha_{ii}}$ and the invasion growth rates $f_i(0, N_j^*) = \mu_i - \alpha_{ij} \frac{\mu_j}{\alpha_{jj}}$. With this we can compute

$$\mathcal{N} = \frac{\mu_i - \alpha_{ij} \frac{\mu_j}{\alpha_{jj}} + \mu_j - \alpha_{ji} \frac{\mu_i}{\alpha_{ii}}}{2} \quad (\text{S14})$$

$$\mathcal{F}_i = 1 + \log \left(\frac{\mu_i / \alpha_{ii}}{\mu_j / \alpha_{jj}} \right) \quad (\text{S15})$$

From this we see that the niche by the arithmetic method depend on the intrinsic growth rates μ_i and μ_j , while the niche differences as defined by the original method does not, they are therefore qualitatively different. Similarly, the fitness differences by the arithmetic method do not depend on interspecific interaction coefficients α_{ij} and α_{ji} .

S3.2 Annual plant model

We here investigate how the different methods compare for the annual plant model, another simple but widely applied community model. For simplicity we assume that the annual plants do not have a seedbank, i.e. we use the following community model:

$$\frac{N_i(t+1)}{N_i(t)} = \frac{\lambda_i}{1 + \sum_j \alpha_{ij} N_j} \quad (\text{S16})$$

, where $N_i(t)$ is the density of species i in year t , λ_i is the basic reproduction rate and α_{ij} is the per-capita effect of species j on species i . The annual plant model is time discrete, to apply the niche and fitness differences metrics we first have to convert the growth rates into continuous form by computing the natural logarithm of the growth rates.

The intrinsic growth rates, monoculture growth rates and the invasion growth rates are given by:

$$f_i(0,0) = \log(\lambda_i) \quad (\text{S17})$$

$$N_i^* = \frac{\lambda_i - 1}{\alpha_{ii}} \quad (\text{S18})$$

$$f_i(0, N_j^*) = \log \left(\frac{\lambda_i}{1 + \alpha_{ij} \frac{\lambda_j - 1}{\alpha_{jj}}} \right) \quad (\text{S19})$$

With this we can compute the different methods

For the original method we have

$$\mathcal{N} = 1 - \sqrt{\frac{\alpha_{ij}\alpha_{ji}}{\alpha_{ii}\alpha_{jj}}} \quad (\text{S20})$$

$$\mathcal{F}_i = \frac{\lambda_j - 1}{\lambda_i - 1} \sqrt{\frac{\alpha_{ii}\alpha_{ij}}{\alpha_{ji}\alpha_{jj}}} \quad (\text{S21})$$

For the geometrical method we have

$$S_i = 1 - \frac{\log\left(\frac{\lambda_i}{1 + \alpha_{ij} \frac{\lambda_j - 1}{\alpha_{jj}}}\right)}{\log(\lambda_i)} \quad (\text{S22})$$

$$= \frac{\log(\lambda_i) - \log\left(\frac{\lambda_i}{1 + \alpha_{ij} \frac{\lambda_j - 1}{\alpha_{jj}}}\right)}{\log(\lambda_i)} \quad (\text{S23})$$

$$= \frac{\log\left(1 + \alpha_{ij} \frac{\lambda_j - 1}{\alpha_{jj}}\right)}{\log(\lambda_i)} \quad (\text{S24})$$

$$\mathcal{N} = 1 - \sqrt{\frac{\log\left(1 + \alpha_{ij} \frac{\lambda_j - 1}{\alpha_{jj}}\right)}{\log(\lambda_i)} \frac{\log\left(1 + \alpha_{ji} \frac{\lambda_i - 1}{\alpha_{ii}}\right)}{\log(\lambda_j)}} \quad (\text{S25})$$

$$\mathcal{F}_i = \sqrt{\frac{\log(\lambda_j) \log\left(1 + \alpha_{ij} \frac{\lambda_j - 1}{\alpha_{jj}}\right)}{\log(\lambda_i) \log\left(1 + \alpha_{ji} \frac{\lambda_i - 1}{\alpha_{ii}}\right)}} \quad (\text{S26})$$

For the arithmetic method we have

$$\mathcal{N} = \frac{\log\left(\frac{\lambda_i}{1 + \alpha_{ij} \frac{\lambda_j - 1}{\alpha_{jj}}}\right) + \log\left(\frac{\lambda_j}{1 + \alpha_{ji} \frac{\lambda_i - 1}{\alpha_{ii}}}\right)}{2} \quad (\text{S27})$$

$$= \frac{1}{2} \log\left(\frac{\lambda_i}{1 + \alpha_{ji} \frac{\lambda_i - 1}{\alpha_{ii}}} \frac{\lambda_j}{1 + \alpha_{ij} \frac{\lambda_j - 1}{\alpha_{jj}}}\right) \quad (\text{S28})$$

$$= -\log\left(\sqrt{\frac{\left(1 + \alpha_{ij} \frac{\lambda_j - 1}{\alpha_{jj}}\right) \left(1 + \alpha_{ji} \frac{\lambda_i - 1}{\alpha_{ii}}\right)}{\lambda_i \lambda_j}}\right) \quad (\text{S29})$$

$$\mathcal{F}_i = 1 + \log\left(\frac{(\lambda_j - 1)\alpha_{ii}}{(\lambda_i - 1)\alpha_{jj}}\right) \quad (\text{S30})$$

For the species-specific method we have

$$\mathcal{N}_i = \frac{\log\left(\frac{\lambda_i}{1+\alpha_{ij}\frac{\lambda_j-1}{\alpha_{jj}}}\right) - \log\left(\frac{\lambda_i}{1+\alpha_{ii}c_{ij}\frac{\lambda_j-1}{\alpha_{jj}}}\right)}{\log(\lambda_i) - \log\left(\frac{\lambda_i}{1+\alpha_{ii}c_{ij}\frac{\lambda_j-1}{\alpha_{jj}}}\right)} \quad (\text{S31})$$

$$= \frac{\log\left(1 + \alpha_{ii}c_{ij}\frac{\lambda_j-1}{\alpha_{jj}}\right) - \log\left(1 + \alpha_{ij}\frac{\lambda_j-1}{\alpha_{jj}}\right)}{\log\left(1 + \alpha_{ii}c_{ij}\frac{\lambda_j-1}{\alpha_{jj}}\right)} \quad (\text{S32})$$

$$= 1 - \frac{\log\left(1 + \alpha_{ij}\frac{\lambda_j-1}{\alpha_{jj}}\right)}{\log\left(1 + \alpha_{ii}c_{ij}\frac{\lambda_j-1}{\alpha_{jj}}\right)} \quad (\text{S33})$$

$$\mathcal{F}_i = 1 - \frac{\log\left(\frac{\lambda_i}{1+\alpha_{ii}c_{ij}\frac{\lambda_j-1}{\alpha_{jj}}}\right)}{\log(\lambda_i)} \quad (\text{S34})$$

$$= \frac{\log\left(1 + \alpha_{ii}c_{ij}\frac{\lambda_j-1}{\alpha_{jj}}\right)}{\log(\lambda_i)} \quad (\text{S35})$$

We use $\mathcal{N}_i$ with subscript i in the appendix, as this is needed for clarity and to solve the necessary equations. Unfortunately, we were not able to obtain closed form expressions for the quantities c_{ij} .

These equations show that the mathematical expressions for niche and fitness differences differ for the various methods. However, they give us little intuitive understanding as the equations are quite complex.

Qualitatively, all niche differences depend positively on α_{ii} and α_{jj} and negatively on α_{ji} and α_{ij} . However, their functional responses differ. Conversely, the different methods have a different qualitative dependence on the parameters λ_i and λ_j (Fig. S5). We investigate here the extreme cases $\lambda_i, \lambda_j \rightarrow 1$ and $\lambda_i \rightarrow \infty$. The original method is not affected by change in intrinsic growth rate λ_i .

Figure S5: Dependence of niche differences on λ_i for the different methods. Niche differences by the original method are independent of λ_i and λ_j . Niche differences by the geometric and species-specific method decrease with increasing λ_i and λ_j . In the limit case of $\lambda_i = \lambda_j = 1$ they agree with the original method. In the other limit case of $\lambda_i = \lambda_j \rightarrow \infty$ they tend to zero. Finally, the arithmetic method increases with increasing λ_i .

For the geometric method:

$$\lim_{\lambda_j \rightarrow 1} \frac{\log\left(1 + \alpha_{ij} \frac{\lambda_j - 1}{\alpha_{jj}}\right)}{\log(\lambda_j)} \stackrel{DH}{=} \lim_{\lambda_j \rightarrow 1} \frac{\left(1 + \alpha_{ij} \frac{\lambda_j - 1}{\alpha_{jj}}\right)^{-1} \frac{\alpha_{ij}}{\alpha_{jj}}}{\frac{1}{\lambda_j}} \quad (S36)$$

$$= \frac{\alpha_{ij}}{\alpha_{jj}} \quad (S37)$$

Where $\stackrel{DH}{=}$ indicates the de L'Hopital rule. This proves that $\lim_{\lambda_i, \lambda_j \rightarrow 1} \mathcal{N}_i = 1 - \sqrt{\frac{\alpha_{ii}\alpha_{ji}}{\alpha_{ii}\alpha_{jj}}}$ for the geometric method, i.e. in the limit case the geometric and the original

Method	$\mathcal{N}$	$\lambda_i, \lambda_j \rightarrow 1$	$\lambda_i, \lambda_j \rightarrow \infty$	Dependence
Original	$1 - \sqrt{\frac{\alpha_{ij}\alpha_{ji}}{\alpha_{ii}\alpha_{jj}}}$	$1 - \sqrt{\frac{\alpha_{ij}\alpha_{ji}}{\alpha_{ii}\alpha_{jj}}}$	$1 - \sqrt{\frac{\alpha_{ij}\alpha_{ji}}{\alpha_{ii}\alpha_{jj}}}$	-
Geometrical	$1 - \sqrt{\frac{\log\left(1 + \alpha_{ij}\frac{\lambda_j-1}{\alpha_{jj}}\right)}{\log(\lambda_i)} \frac{\log\left(1 + \alpha_{ji}\frac{\lambda_i-1}{\alpha_{ii}}\right)}{\log(\lambda_j)}}$	$1 - \sqrt{\frac{\alpha_{ij}\alpha_{ji}}{\alpha_{ii}\alpha_{jj}}}$	0	Decrease
Arithmetical	$-\log\left(\sqrt{\frac{\left(1 + \alpha_{ij}\frac{\lambda_j-1}{\alpha_{jj}}\right)\left(1 + \alpha_{ji}\frac{\lambda_i-1}{\alpha_{ii}}\right)}{\lambda_i\lambda_j}}\right)$	0	$-\log\left(\sqrt{\frac{\alpha_{ij}\alpha_{ji}}{\alpha_{ii}\alpha_{jj}}}\right)$	Increase
Species specific	$1 - \frac{\log\left(1 + \alpha_{ij}\frac{\lambda_j-1}{\alpha_{jj}}\right)}{\log\left(1 + \alpha_{ii}\frac{\lambda_j-1}{\alpha_{jj}}\right)}$	$1 - \sqrt{\frac{\alpha_{ij}\alpha_{ji}}{\alpha_{ii}\alpha_{jj}}}$	0	Decrease

Table S2: Niche and fitness differences according to the different methods for the annual plant method. $\mathcal{N}$ gives the mathematical expression for the niche differences. $\lambda_i, \lambda_j \rightarrow 1$ gives the limit case for when the intrinsic growth rates approach 1. $\lambda_i, \lambda_j \rightarrow \infty$ gives the limit case for when the intrinsic growth rates approach ∞ . Dependence: describes how $\mathcal{N}_i$ depends on the intrinsic growth rates for the case of $\sqrt{\frac{\alpha_{ij}\alpha_{ji}}{\alpha_{ii}\alpha_{jj}}} < 1$. However, this dependence might not be monotone.

method are identical. For the other extreme case:

$$\lim_{\lambda_j \rightarrow \infty} \frac{\log\left(1 + \alpha_{ij}\frac{\lambda_j-1}{\alpha_{jj}}\right)}{\log(\lambda_j)} = \lim_{\lambda_j \rightarrow \infty} \frac{\log(\lambda_j - 1) + \log\left(\frac{1}{\lambda_j-1} + \frac{\alpha_{ij}}{\alpha_{jj}}\right)}{\log(\lambda_j)} \quad (\text{S38})$$

$$= 1 \quad (\text{S39})$$

662 Which proves that $\mathcal{N} \rightarrow 0$ for increasing $\lambda_i, \lambda_j \rightarrow \infty$. In general, increasing intrinsic
663 growth λ_i, λ_j therefore bring niche differences closer to 0 in the geometric method.

For the arithmetic method:

$$\lim_{\lambda_j \rightarrow 1} \frac{1 + \alpha_{ij}\frac{\lambda_j-1}{\alpha_{jj}}}{\lambda_j} = 1 \quad (\text{S40})$$

, which gives $\mathcal{N} \rightarrow 0$ for the arithmetic method. Conversely,

$$\lim_{\lambda_j \rightarrow \infty} \frac{1 + \alpha_{ij}\frac{\lambda_j-1}{\alpha_{jj}}}{\lambda_j} = \lim_{\lambda_j \rightarrow \infty} \frac{1}{\lambda_j} + \alpha_{ij}\frac{1 - \frac{1}{\lambda_j}}{\alpha_{jj}} = \frac{\alpha_{ij}}{\alpha_{jj}} \quad (\text{S41})$$

664 Which gives $\mathcal{N} \rightarrow -\log\left(\sqrt{\frac{\alpha_{ij}\alpha_{ji}}{\alpha_{ii}\alpha_{jj}}}\right)$ for increasing $\lambda_i, \lambda_j \rightarrow \infty$. In general, decreasing
665 intrinsic growth λ_i, λ_j therefore bring niche differences closer to 0 in the arithmetic

method, unlike in the geometric method.

For the species-specific method

$$\lim_{\lambda_j \rightarrow 1} \frac{\log \left(1 + \alpha_{ij} \frac{\lambda_j - 1}{\alpha_{jj}} \right)}{\log \left(1 + \alpha_{ii} c_{ij} \frac{\lambda_j - 1}{\alpha_{jj}} \right)} =^{DH} \lim_{\lambda_j \rightarrow 1} \frac{\left(1 + \alpha_{ij} \frac{\lambda_j - 1}{\alpha_{jj}} \right)^{-1} \frac{\alpha_{ij}}{\alpha_{jj}}}{\left(1 + \alpha_{ii} c_{ij} \frac{\lambda_j - 1}{\alpha_{jj}} \right)^{-1} \frac{\alpha_{ii}}{\alpha_{jj}} \left(\frac{(\lambda_j - 1) \partial c_{ij}}{\partial \lambda_j} + c_{ij} \right)} \quad (S42)$$

$$= \frac{\alpha_{ij}}{\alpha_{ii} c_{ij}} \quad (S43)$$

Similarly, we get $\lim_{\lambda_i \rightarrow 1} \mathcal{N}_i = 1 - \frac{\alpha_{ji}}{\alpha_{jj} c_{ji}}$. By definition we have $\mathcal{N}_i = \mathcal{N}_j$ which yields

$$c_{ij} = \sqrt{\frac{\alpha_{ij} \alpha_{jj}}{\alpha_{ii} \alpha_{ji}}} \text{ and } \mathcal{N}_i = 1 - \frac{\alpha_{ij} \alpha_{ji}}{\alpha_{ii} \alpha_{jj}}.$$

For increasing λ_j we have:

$$\lim_{\lambda_j \rightarrow \infty} \frac{\log \left(1 + \alpha_{ij} \frac{\lambda_j - 1}{\alpha_{jj}} \right)}{\log \left(1 + \alpha_{ii} c_{ij} \frac{\lambda_j - 1}{\alpha_{jj}} \right)} = \lim_{\lambda_j \rightarrow \infty} \frac{\log(\lambda_j - 1) + \log \left(\frac{1}{\lambda_j - 1} + \frac{\alpha_{ij}}{\alpha_{jj}} \right)}{\log(\lambda_j - 1) + \log \left(\frac{1}{\lambda_j - 1} + \frac{\alpha_{ii} c_{ij}}{\alpha_{jj}} \right)} \quad (S44)$$

$$= 1 \quad (S45)$$

Which yields $\mathcal{N}_i \rightarrow 0$ for $\lambda_j \rightarrow \infty$. Consequently, similar to the geometric framework the species-specific framework approaches the original method for decreasing λ_i, λ_j . For increasing λ_j , $\mathcal{N}_i$ approaches 0. However, for the geometric method both λ_i and λ_j must approach ∞ for zero niche differences, while for the species-specific method only one of the two has to approach ∞ .

S3.3 Model independent

Unfortunately, little can be said for a general community model. We focus here on the arithmetic, geometric and species-specific method, as these can be applied to any community model.

Qualitatively the methods agree on the sign of niche differences for a community driven by coexistence or priority effects. For a community driven by coexistence we have $f_i(0, N_j^*), f_j(0, N_i^*) > 0$, which leads to $S_i, S_j < 1$. Consequently, we have $\frac{1}{2} (f_i(0, N_j^*) + f_j(0, N_i^*)) > 0$, $1 - \sqrt{S_i S_j} > 0$. Finally, Spaak *et al.* (2020) have shown

coexisting species are driven by negative frequency dependence and therefore have positive niche differences. The case of priority effects is equivalent with all inequality signs exchanged. These three methods therefore qualitatively disagree only for the case of competitive exclusion.

Next, we show that the species-specific and geometric method always agree on the competitive superior species. We assume without loss of generality that species i is the superior competitor by the species-specific method, we therefore have

$$1 - \frac{f_i(c_{ij}N_j^*, 0)}{f_i(0, 0)} < 1 \quad (\text{S46})$$

$$\Leftrightarrow -f_i(c_{ij}N_j^*, 0) < 0 \quad (\text{S47})$$

$$\Leftrightarrow f_i(c_{ij}N_j^*, 0) > 0 = f_i(N_i^*, 0) \quad (\text{S48})$$

$$\Leftrightarrow c_{ij}N_j^* < N_i^* \quad (\text{S49})$$

$$\Leftrightarrow N_j^* < \frac{1}{c_{ij}}N_i^* = c_{ji}N_i^* \quad (\text{S50})$$

$$\Leftrightarrow f_j(N_j^*, 0) = 0 > f_j(c_{ji}N_i^*, 0) \quad (\text{S51})$$

$$\Leftrightarrow 1 < 1 - \frac{f_j(c_{ji}N_i^*, 0)}{f_j(0, 0)} \quad (\text{S52})$$

Where we used the monotony of the per-capita growth rates from S48 to S49 and from S50 to S51, and we used the definition of $c_{ij} = \frac{1}{c_{ji}}$ from S49 to S50. We therefore indeed have only one species with a fitness difference exceeding 1. And only for this species (by assumption species i) the no-niche growth rate $f_i(c_{ij}N_j^*, 0)$ is positive.

Rearranging this we get

$$f_i(c_{ij}N_j^*, 0) > 0 \quad (\text{S53})$$

$$\Leftrightarrow f_i(0, 0) - f_i(c_{ij}N_j^*, 0) < f_i(0, 0) \quad (\text{S54})$$

$$\Leftrightarrow \frac{1}{f_i(0, 0) - f_i(c_{ij}N_j^*, 0)} > \frac{1}{f_i(0, 0)} \quad (\text{S55})$$

$$\Leftrightarrow \frac{f_i(0, 0) - f_i(0, N_j^*)}{f_i(0, 0) - f_i(c_{ij}N_j^*, 0)} > \frac{f_i(0, 0) - f_i(0, N_j^*)}{f_i(0, 0)} = S_i \quad (\text{S56})$$

where the left hand side of S56 is the niche overlap as defined by the species-specific

method and the right hand side of S56 is the sensitivity of species i as defined by the method of Carroll. As species i was the competitive dominant by the species-specific method we therefore must have

$$\frac{f_j(0,0) - f_j(0, N_i^*)}{f_j(0,0) - f_j(c_{ji}N_i^*, 0)} < \frac{f_j(0,0) - f_j(0, N_i^*)}{f_j(0,0)} = S_j \quad (\text{S57})$$

Additionally, we have by the definition of the species-specific method

$$S_i < \frac{f_i(0,0) - f_i(0, N_j^*)}{f_i(0,0) - f_i(c_{ij}N_j^*, 0)} = \frac{f_j(0,0) - f_j(0, N_i^*)}{f_j(0,0) - f_j(c_{ji}N_i^*, 0)} < S_j \quad (\text{S58})$$

$$S_i < S_j \quad (\text{S59})$$

$$\sqrt{\frac{S_i}{S_j}} > 1 \quad (\text{S60})$$

690 Therefore the species-specific and the geometric method always agree on the compet-
691 itive dominant species.

692 S4 Quantitative comparison for the annual plant data

Figure S6: In the main text we focused on the qualitative differences between the methods. Here we show a more quantitative comparison for the data from Pérez-Ramos *et al.* (2019). We plot the niche (panels above diagonal) and fitness differences (panels below diagonal) of each species by each method and compare them. The fitness differences have been log-transformed. For example, the top-right plot shows the niche differences as computed by the species-specific method on y-axis and the niche differences as computed by the arithmetic method on the x-axis. Similarly, for the other plots. We performed a linear regression (red line) and report its R^2 value. Some relationships were not linear, so we also report the Kendall-Tau correlation for non-parametric regression.

Figure S7: Similar to figure S6 but for the data from Germain *et al.* (2016).