

## RESEARCH OUTPUTS / RÉSULTATS DE RECHERCHE

### Different measures of niche and fitness differences tell different tales

Spaak, Jurg; Ke, Po-Ju; Letten, Andrew; DE LAENDER, Frederik

*Published in:*  
OIKOS

*Publication date:*  
2022

*Document Version*  
Publisher's PDF, also known as Version of record

#### [Link to publication](#)

*Citation for pulished version (HARVARD):*

Spaak, J, Ke, P-J, Letten, A & DE LAENDER, F 2022, 'Different measures of niche and fitness differences tell different tales', *OIKOS*. <<https://onlinelibrary.wiley.com/doi/epdf/10.1111/oik.09573>>

#### General rights

Copyright and moral rights for the publications made accessible in the public portal are retained by the authors and/or other copyright owners and it is a condition of accessing publications that users recognise and abide by the legal requirements associated with these rights.

- Users may download and print one copy of any publication from the public portal for the purpose of private study or research.
- You may not further distribute the material or use it for any profit-making activity or commercial gain
- You may freely distribute the URL identifying the publication in the public portal ?

#### Take down policy

If you believe that this document breaches copyright please contact us providing details, and we will remove access to the work immediately and investigate your claim.

## Research article

# Different measures of niche and fitness differences tell different tales

Jurg W. Spaak, Po-Ju Ke, Andrew D. Letten and Frederik De Laender

J. W. Spaak (<https://orcid.org/0000-0001-5157-9188>) ✉ ([j.w.spaak@gmail.com](mailto:j.w.spaak@gmail.com)), Dept of Ecology and Evolutionary Biology, Cornell Univ., Ithaca, NY, USA. – P.-J. Ke (<https://orcid.org/0000-0002-8371-7984>), Inst. of Ecology and Evolutionary Biology, National Taiwan Univ., Taipei, Taiwan and Dept of Ecology & Evolutionary Biology, Princeton Univ., Princeton, NJ, USA. – A. D. Letten (<https://orcid.org/0000-0001-6436-7942>), School of Biological Sciences, Univ. of Queensland, Brisbane, QLD, Australia. – F. De Laender (<https://orcid.org/0000-0002-4060-973X>), Univ. of Namur, Namur, Belgium and Inst. of Life-Earth-Environment, Namur Center for Complex Systems, Namur, Belgium.

Oikos

2022: e09573

doi: 10.1111/oik.09573

Subject Editor: Shurong Zhou

Editor-in-Chief: Pedro Peres-Neto

Accepted 31 October 2022



In modern coexistence theory, species coexistence can either arise via strong niche differences or weak fitness differences. Having a common currency for interpreting these mechanisms is essential for synthesizing knowledge across different studies and systems. However, several methods for quantifying niche and fitness differences exist, with little guidance on how and why these methods differ. Here, we first organize the available methods into three groups and review their differences from a conceptual point of view. Next, we apply four methods to quantify niche and fitness differences to one simulated and one empirical data set. We show that these methods do not only differ quantitatively, but affect how we interpret coexistence. Specifically, the different methods disagree on how resource supply rates (simulated data) or plant traits (empirical data) affect niche and fitness differences. We argue for a better theoretical understanding of what connects and sets apart different methods and more precise empirical measurements to foster appropriate method selection in coexistence theory.

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

## 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 2005), contemporary niche theory focuses on investigating species' limiting factors (Tilman et al. 1982) and modern coexistence theory focuses on separating niche and fitness differences that help and hamper coexistence (Adler et al. 2007).

The competitive exclusion principle states that ecologically equivalent species cannot coexist (Gause 1934). 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



[www.oikosjournal.org](http://www.oikosjournal.org)

© 2022 The Authors. Oikos published by John Wiley & Sons Ltd on behalf of Nordic Society Oikos.

This is an open access article under the terms of the Creative Commons Attribution License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

are termed niche differences (Chesson 2000, Spaak and De Laender 2020) which capture all species differences that help a species to recover from low density. However, species differences may not only increase niche differences, but can also affect the competitive strength of a species (Kraft et al. 2015). 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.

While conceptually straightforward, there are many different methods to compute niche and fitness differences (Godwin et al. 2020, Spaak and De Laender 2020), with new methods still being developed for competition for mutualists (Johnson 2021), resource competition (Koffel et al. 2021), plant–soil feedbacks (Kandlikar et al. 2019), multitrophic communities (Spaak et al. 2021b), spatial explicit communities (Luo et al. 2022) or communities with evolving species (Yamamichi et al. 2022). Few of these methods have been applied extensively to experiments, while others have been rarely used (Buche et al. 2022) and little is known about their practical usability. 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 (Godwin et al. 2020, Spaak and De Laender 2020, Spaak et al. 2021b). It is therefore unclear why there are so many methods and what sets the different methods apart.

We here inventory available definitions of niche and fitness differences, and examine how they are related and differ conceptually. Next, we show that their differences are not merely semantic, but rather affect which mechanisms and traits are interpreted as most important. We do so by analysing two data sets with 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 from Letten et al. (2017) consisted of simulation output from a virtual experiment in which we investigated how changes in resource availability affect niche and fitness differences. We find that the different methods often do not agree on how resource availability affects niche and fitness differences. The second data set from Pérez-Ramos et al. (2019) 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.

We finally recognize that it is, at present, difficult to draft guidelines on which method to use to address a given research question. While we can identify general criteria on which to base said guidelines, some of these criteria require

further theoretical and empirical investigation; we lay out such future directions in the discussion.

## Review of niche and fitness differences methods

Chesson (2000) formalized the conceptual idea of niche and fitness differences based on the mathematical equations in Chesson (1990). Species coexist if niche differences are sufficiently strong to overcome fitness differences and the species with higher fitness (i.e. lower fitness differences) will out-compete its competitor in the absence of niche differences. Most methods match this verbal description of niche and fitness differences (Fig. 1, solid lined boxes).

Mathematically, niche and fitness differences were initially defined for a two-species Lotka–Volterra community model by a comparison of the interspecific interaction strength to the intraspecific interaction strength. They were first labeled as niche and fitness differences in Chesson and Kuang (2008) which investigated resource competition and apparent competition, although the mathematical ground work stems from Chesson (1990), based on an underlying MacArthur consumer–resource model and its links to Lotka–Volterra community model. While this definition is virtually undisputed, it only applies to a competitive two-species Lotka–Volterra model. In such a community, niche differences are defined

as  $\mathcal{N} = 1 - \sqrt{\frac{\alpha_{ij}\alpha_{ji}}{\alpha_{ii}\alpha_{jj}}}$  and fitness differences are defined as

$$\mathcal{F}_i = \frac{\mu_i}{\mu_j} \sqrt{\frac{\alpha_{ji}\alpha_{jj}}{\alpha_{ii}\alpha_{ij}}}, \text{ where } \alpha_{ij} \text{ are the species interaction coefficients and } \mu_i \text{ is the intrinsic growth rate.}$$

Many other methods to assess niche and fitness differences trace their origin back to this original idea (Fig. 1, red block). These equations have been applied to similar community models, by rephrasing these similar community models in the terms of the inter and intra-specific interaction coefficients  $\alpha_{ij}$  (Spaak et al. 2021a). Specifically, this method has been used in community models including competition for predators (Chesson and Kuang 2008) or mutualists (Johnson 2021), a time discrete annual plant model (Godoy and Levine 2014), resource competition models with Holling type II response (Letten et al. 2017), inclusion of facilitation between species (Bimler et al. 2018) and plant–soil feedbacks (Kandlikar et al. 2019). These methods have typically been created with a specific research question in mind and there was no alternative available to compute niche and fitness differences for that specific model/research question. That is, most of these methods have been designed because of practical necessity and not because the authors disagreed with earlier methods.

Alternatively, instead of using model specific parameters some methods define niche and fitness differences via model independent criteria of the community, most prominently the invasion growth rates (Fig. 1 blue block). Many of these methods define niche differences as a sort of average

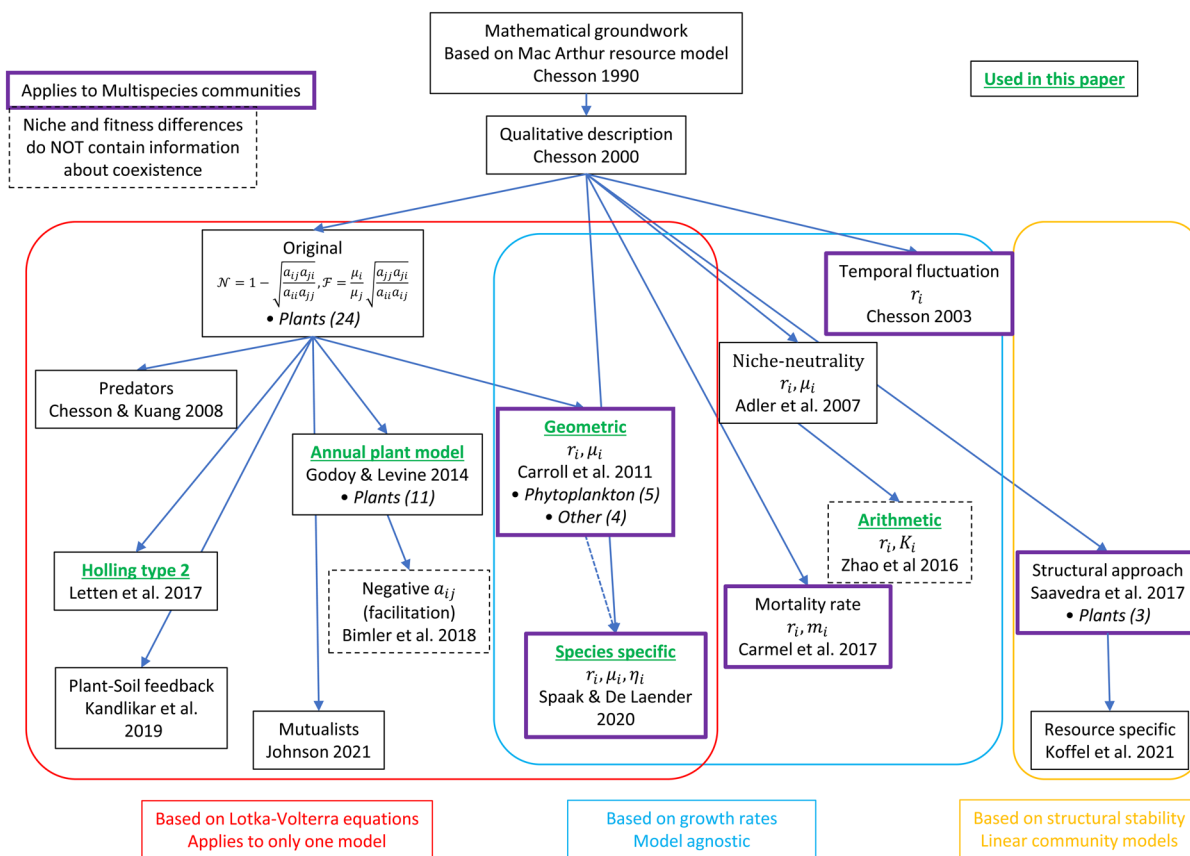


Figure 1. The available methods to compute niche and fitness differences can be loosely grouped into three conceptually different groups. First, all methods which track their definition back to the original two-species Lotka–Volterra model based on a comparison of interspecific interaction strength to intraspecific interaction strength (red box). Typically, these methods have been designed to be applied to a specific community model, often because no alternative was available. Second, all methods which are based on specific growth rates (blue box). Typically, these methods have been designed because of a different conceptual interpretation of how niche and fitness differences should be measured. They differ by which information they include into the computation of niche and fitness differences, namely invasion growth rate  $r_i$ , intrinsic growth rate  $\mu_i$ , no-niche growth rate  $\eta_i$ , mortality rate  $m_i$ , or carrying capacity  $k_i$ . Third, the methods which are defined on the structural stability and the feasibility of the community (yellow box). Only four of these methods have been applied empirically more than once as of December 2020 (Buche et al. 2022), indicated by the bullet point below the methods. The outline of the box indicates whether the method applies to multi-species communities (purple) and whether the method allows inference about coexistence (solid line). The methods used in our current paper are written in green (annual plant model, Holling type II, geometric, species specific and arithmetic). See the Supporting information for structured overview on how the methods differ.

of invasion growth rates, and fitness differences as a variation around this mean (Chesson 2003, Carroll et al. 2011, Zhao et al. 2016, Carmel et al. 2017). The methods differ in which additional information is included into the computation of this mean. The creation of these methods were typically not driven by a specific application, but rather the creation of the method is a key aspect of those papers. Often, these authors disagreed with the previous methods on how niche and fitness differences should be defined in general.

For example, Chesson (2003), Adler et al. (2007) and Zhao et al. (2016) assumed that invasion growth rates  $r_i$  contain sufficient information to compute niche differences. For these three methods, niche differences increase with increasing invasion growth rates, but differ in the functional dependence of niche differences on invasion growth rates. Importantly, there is no lower or upper bound of niche and

fitness differences. However, Carroll et al. (2011) argued that 1 should be a natural upper bound for niche differences indicating non-interacting species, i.e. completely different niches. To achieve this, they compare the invasion growth rates  $r_i$  to the intrinsic growth rate  $\mu_i$  of the species. Specifically, if  $\mu_i = r_i$  then the residents do not affect the growth rate of the invader, i.e. they do not interact, in this case niche differences  $\mathcal{N}_i$  is set to its maximum 1. On the other hand, Carmel et al. (2017) argued that 0 should be a natural lower bound of niche differences. To this end, they compared the invasion growth rates to the mortality rate of the species, as this would be the smallest possible growth rate. Minimal niche differentiation, i.e.  $\mathcal{N} = 0$  is achieved if the invasion growth rate of all species equals their mortality rate. We note, however, that this interpretation of  $\mathcal{N} = 0$  differs from interpretations in most other definitions where neutral species have

$\mathcal{N} = 0$  (Chesson 1990, Adler et al. 2007, Zhao et al. 2016, Spaak and De Laender 2020). Finally, Spaak and De Laender (2020) argue that both values, 0 and 1, are important for the definition of niche differences, where niche difference is equal to 1 for non-interacting species and niche difference is 0 for species with identical niches. To achieve this, they compare the invasion growth rates to both the invasion and the no-niche growth rate to assess the effect of interspecific and intraspecific interactions respectively. However, additionally, they show that  $\mathcal{N} > 1$  corresponds to facilitation, while  $\mathcal{N} < 0$  is indicative of positive frequency dependence in agreement with Ke and Letten (2018). In all these methods, fitness differences are defined as the competitive advantage a species would have in absence of niche differences, the fitness differences therefore differ across the different methods because of varying underlying niche differences definitions.

Carroll et al. (2011) and Spaak and De Laender (2020) form a bridge between these two conceptually different approaches to define niche and fitness differences (red and blue block). They are both defined over growth rates and are therefore model agnostic but, importantly, if applied to a Lotka–Volterra model they recover equivalent definitions as the original definition for this specific model (Supporting information). However, this is generally not the case for other models or methods, as we show below. For example, other growth rate based methods do not recover the original equations for a Lotka–Volterra community model. Similarly, Carroll et al. (2011), Godoy and Levine (2014) Spaak and De Laender (2020) do not agree with each other on niche differences in the annual plant model.

Finally, two methods define a structural equivalent to niche and fitness differences (Fig. 1, yellow box) and differ from the other methods in two aspects. First, as their name suggests, they investigate structural stability of the community, and not dynamical stability as the methods based on Lotka–Volterra equations or growth rates. That is, they ask how the community model responds to a change in model parameters, not species densities (Meszéna et al. 2006). Second, the methods based on Lotka–Volterra equations or growth rates define niche and fitness differences based on a specific parametrization of the community model, i.e. niche and fitness differences are computed for given  $\mu_i$  and  $\alpha_{ij}$ . The structural approaches compute niche differences based on some fixed parameters and some unfixed, i.e. niche and fitness differences are computed for given  $\alpha_{ij}$  and all feasible values for  $\mu_i$ . Similar to general methods (blue box), the creation of these methods is a main focus of these papers.

## Methods

### Niche and fitness differences metrics

We focus on five of the fifteen methods to assess niche and fitness differences as they are the ones which can be directly applied to our two examples, we use two from the Lotka–Volterra based methods and three from the growth rates

based methods. We will refer to both Lotka–Volterra based methods as ‘original’ method in the rest of the manuscript, as they use the same underlying mathematical concept. The first description of these selected methods can be found in Carroll et al. (2011), Godoy and Levine (2014), Zhao et al. (2016), Letten et al. (2017) and Spaak and De Laender (2020).

We first introduce the original method to assess niche and fitness differences for a two-species Lotka–Volterra model, which is derived from an underlying MacArthur consumer–resource model (Chesson 1990, 2013). The growth rates of the species are given by  $\frac{1}{N_i} \frac{dN_i}{dt} = \mu_i - \sum_j \alpha_{ij} N_j$ , where  $N_i$  is the density of species  $i$ ,  $\mu_i$  is the intrinsic growth rate of species  $i$  and  $\alpha_{ij}$  is the per-capita effect of species  $j$  on species  $i$ . For this model, the niche and fitness differences are defined as:

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

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

Often, 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. This method is originally only defined for a two species Lotka–Volterra model but can be applied to other community models as well by reducing a given community model to Lotka–Volterra form. Specifically, we use the adaptations by Letten et al. (2017) to the Tilman consumer–resource model and the adaptation by Godoy and Levine (2014) to the Beverton–Holt annual plant model in the examples below, corresponding to the methods used by Letten et al. (2017) and Pérez-Ramos et al. (2019), respectively. We will refer to this method as the original method to assess niche and fitness differences.

The other three methods are based on different growth rates to assess niche and fitness differences. We therefore first introduce the different growth rates, for this 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. Then, the intrinsic growth rate  $\mu_i$  of species  $i$  is the growth rate when both species are at low density (mathematically zero) i.e.  $\mu_i = f_i(0, 0)$ . The invasion growth rate  $r_i$  of species  $i$  is the growth rate when it is at low density (mathematically zero) and its competitor at their monoculture density, i.e.  $r_i = f_i(0, N_j^*)$ . The no-niche growth rate  $\eta_i$  of species  $i$  is the growth rate when species  $i$  invades species  $j$  in the hypothetical scenario where the species have no-niche differences. Mathematically this is defined as the growth rate at the scaled carrying capacity



of its competitor, i.e.  $\eta_i = f_i(c_{ij}N_j^*, 0)$  where  $c_{ij}$  is a conversion factor ensuring that both species consume the same total amount of resources.

Given this notation, Carroll et al. (2011) defined niche and fitness differences based on the intrinsic and invasion growth rates, which define the sensitivity of each species to competition, i.e.  $\frac{\mu_i - r_i}{\mu_i}$  and defined

$$\mathcal{N} = 1 - \sqrt{\frac{\mu_i - r_i}{\mu_i} \frac{\mu_j - r_j}{\mu_j}} \quad (3)$$

$$\mathcal{F}_i = \sqrt{\frac{\frac{\mu_i - r_i}{\mu_i}}{\frac{\mu_j - r_j}{\mu_j}}} \quad (4)$$

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,  $\frac{\mu_i - r_i}{\mu_i}$ .

The method by Zhao et al. (2016) includes the invasion growth rate and the mono-culture equilibrium densities and defines niche and fitness differences as

$$\mathcal{N} = \frac{r_i + r_j}{2} \quad (5)$$

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

We have adjusted their definition in the following way  $\mathcal{N} = (1 - \rho)/2$ , where  $\rho$  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 linear change was done to ensure that a neutral community would have zero niche differences and fitness differences of 1 (similar to the other methods). 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 and De Laender (2020), who defined  $\mathcal{N}$  and  $\mathcal{F}$  as

$$\mathcal{N}_i = \frac{r_i - \eta_i}{\mu_i - \eta_i} \quad (7)$$

$$\mathcal{F}_i = 1 - \frac{\eta_i}{\mu_i} \quad (8)$$

Equation 7 compares the actual invasion growth rate  $r_i$  of species  $i$  to the hypothetical invasion growth rate when the two species would not interact, i.e.  $\mu_i$ , and to another hypothetical invasion growth rate when the two species had no niche differences, i.e.  $\eta_i$ . 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 and De Laender (2020) to ensure that a neutral community will have fitness difference equal to 1 and the competitive superior species has lower fitness differences. 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. However, since the current work focuses on competition we always have  $\mathcal{N}_i = \mathcal{N}_j$ , and so we define  $\mathcal{N} := \mathcal{N}_i = \mathcal{N}_j$ .

## Examples

We have chosen two examples from the literature, and let four hypothetical scientists analyse the data using a different method each. We then investigated whether these different hypothetical scientists would have come to a different conclusion about their systems.

## Theoretical example

Letten et al. (2017) investigated how changes of parameters in a two-species consumer–resource model affect niche and fitness differences. Letten et al. (2017) described a specific method (Fig. 1) and applied their method as described below. We repeat part of their work with multiple different methods. They used a Tilman consumer–resource model with substitutable resources with a Holling type II 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 sign of 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 and Kuang 2008, Petry et al. 2018, Terry et al. 2021) or species traits (Kraft et al. 2015, Gallego et al. 2019, Pérez-Ramos 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). We chose Pérez-Ramos et al. (2019) 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.

Pérez-Ramos et al. (2019) fitted the community parameters of the annual plant model to their empirical data. Given these community parameters we computed the intrinsic, invasion and no-niche growth rates to compute the niche and fitness differences. We then performed a linear regression of trait differences and the response variables, niche or fitness differences. For each linear regression we report the coefficient of correlation.

## Results

### Theoretical example

Because of the dependence on different growth rates it is generally not possible to convert niche and fitness differences from one method to another (Buche et al. 2022) and the different methods react differently to changes in underlying community parameters (Spaak and De Laender 2020). As an example, we show the effect of changes in productivity (resource availability for the resource competition model with Holling type II response and fecundity for the annual plant model) on niche and fitness differences in the resource competition model with Holling type II response and the annual plant model (Fig. 2). Niche differences in the original method only depend on intra and interspecific interactions.

Consequently, changes in productivity do not affect them (orange line). Niche differences in the arithmetic method depend only on the invasion growth rate, which increase with increased productivity, which therefore increases niche differences (red). Finally, increased productivity also increases the intrinsic growth rate and affects no-niche growth rates. This combined effect leads to non-monotonic effects of productivity on niche differences according to the species-specific and geometric method (purple and magenta). Generally, the effects on fitness differences are more complex, as changes in productivity affect both species (Fig. 2C–D).

We now let the different scientists investigate two different communities of two species competing for two substitutable resources (Fig. 3A, C). For each community, they changed the resource supply rates (from black triangle to black dot).

As observed before (Song et al. 2019, Spaak and De Laender 2020), 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 (Fig. 3A, B), the scientists might ask whether the decrease in resource 1 is most beneficial for the blue or the green species. If the decrease of resource 1 benefits the blue species, consistent with resource ratio theory (Tilman et al. 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, the scientists might ask whether the community is driven by positive or negative frequency dependence, which would lead to negative or positive niche differences, respectively. Again, 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 (C, D), the scientists 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, the scientists might ask which species is the competitive superior. Again, they obtain different answers when using different methods. The original and arithmetic method conclude that the blue species is the competitive superior ( $\mathcal{F}_g > 1$ ) at the black triangle supply point, 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).

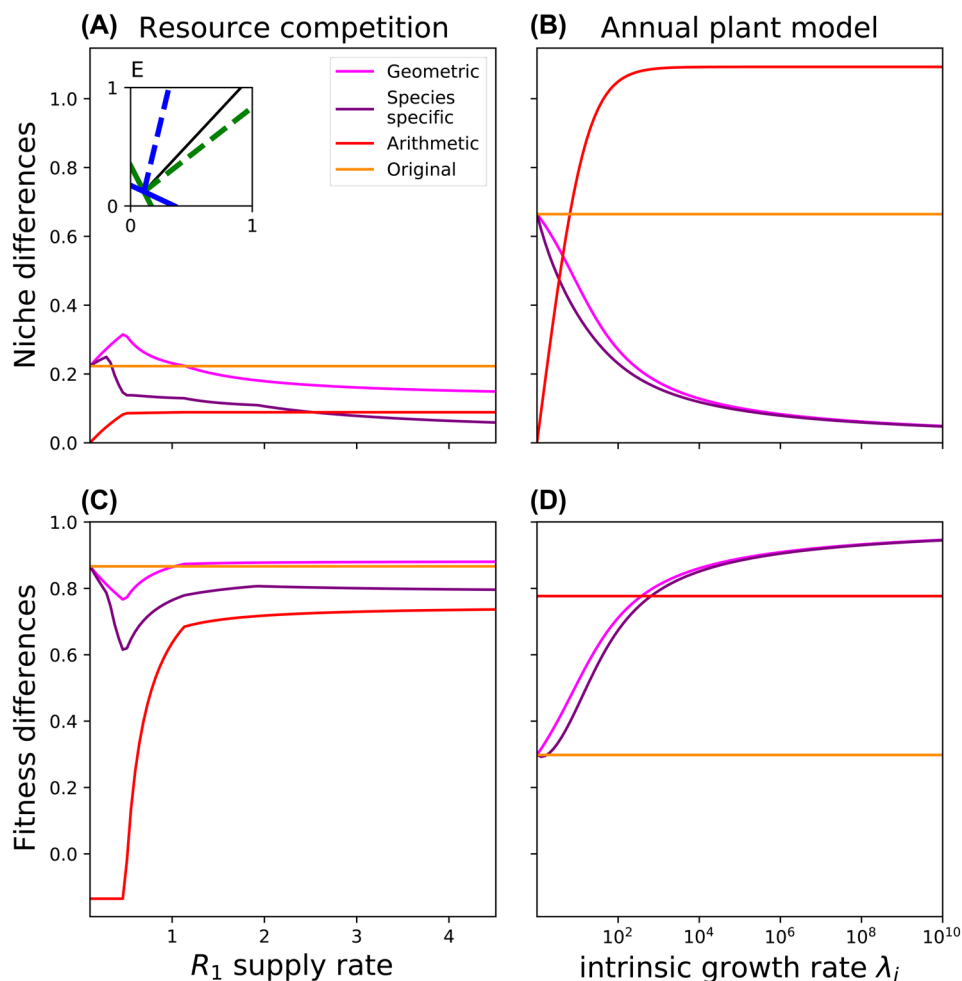


Figure 2. For the resource competition model (A, C) and the annual plant model (B, D) we investigate the effect of increasing productivity (x-axis) on niche (A, B) and fitness differences (C, D) for the different methods. To increase productivity for two species competing for resources (panel E) with zero-net-growth inclines as shown by solid lines we increased the resource supply rate along the black line. For the annual plant model we increased the fertility parameter  $\lambda_r$ . Increasing productivity also increases other metrics of the community, such as invasion growth rate or equilibrium density. Niche differences for the original method depend only on the inter- and intraspecific interaction strength and are therefore not affected by changes in productivity. Increasing the productivity increases many parameters, including the intrinsic growth rate  $\mu_i$ , the equilibrium density  $N_i^*$  and the invasion growth rate  $r_i$ . Therefore, niche differences in the arithmetic method increase with increasing productivity, as they only depend on the invasion growth rate. Conversely, niche differences for the geometric and species specific depend on multiple of these growth rates and the effect of the increased intrinsic growth rate counteract the increased invasion growth rates. In the limit case where productivity approaches 0 the geometric and species specific method approach the original method for both niche and fitness differences (Supporting information). However, niche differences for the arithmetic method approach 0.

Admittedly, we have searched (trial and error) for 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 (Supporting information) and computed their niche and fitness differences (Fig. 4A). 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 examples; often, the methods did not agree well (red dots, Fig. 4A) and they did

not agree much better than expected by chance (Supporting information).

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 their effect on  $\mathcal{N}$  and on their effect on  $\mathcal{F}_i$  of the change in resource supply rate (Fig. 4B). 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. which one has  $\mathcal{F}_i < 0$ , as both methods to assess competitive dominance are equivalent to a comparison of the sensitivities of the species to competition, i.e.



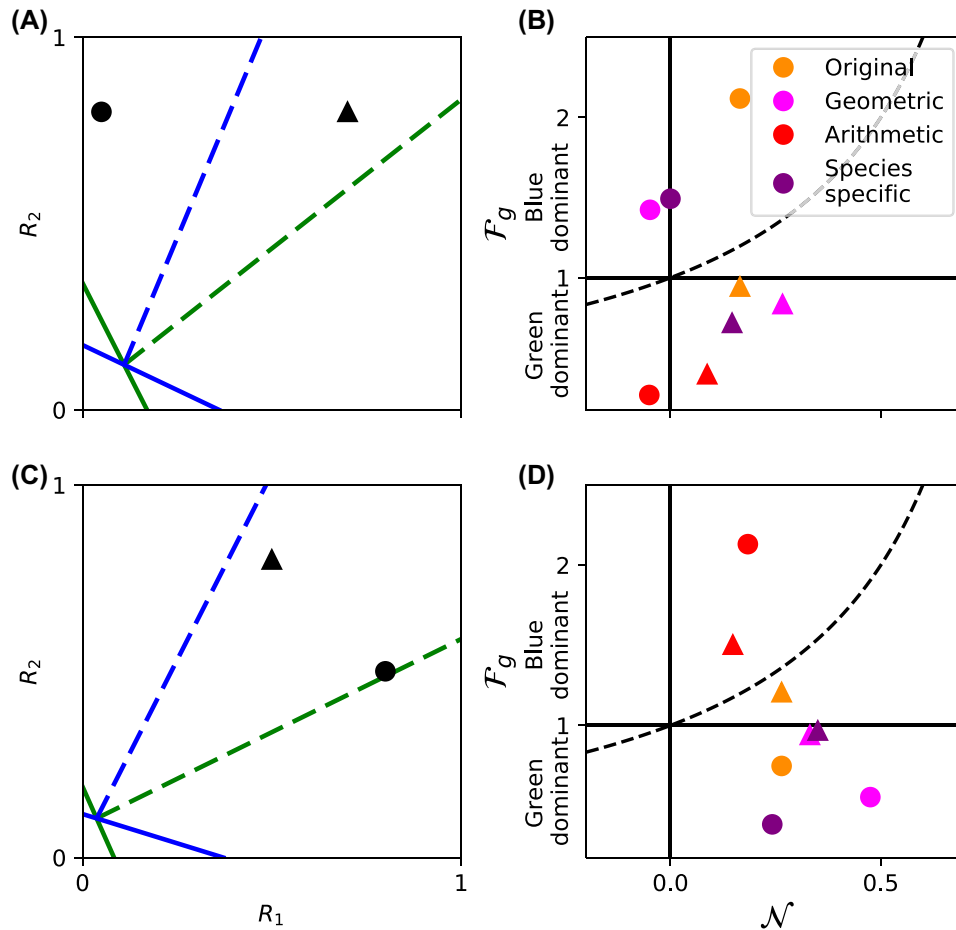


Figure 3. (A, C) Two examples of two species competing for substitutional resources. The solid lines indicate the zero-net growth isoclines (ZNGI), at which the species has zero growth. The dashed lines indicate the consumption vectors  $u_i$  of the species. The black triangle and dot show the resource 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) for the two different resource supply rates (dot and triangle). In certain cases, the different methods do not agree qualitatively on whether: 1) The community is driven by positive ( $\mathcal{N} < 0$ ), negative ( $\mathcal{N} > 0$ ) or no frequency dependence (B, dot). 2) The blue species ( $\mathcal{F}_g > 1$ ) or the green species ( $\mathcal{F}_g < 1$ ) is the competitive superior species (D, triangle). 3) The change in resource supply rates is more beneficial for the blue (higher  $\mathcal{F}_g$  for the dot than for the triangle) or green species (lower  $\mathcal{F}_g$  for the dot than for the triangle, B). 4) The change in resource supply rates increases, decreases or has no effect on niche differences (D, change in  $\mathcal{N}$ ) This figure is conceptually similar to Fig. 2 from (Letten et al. 2017). The black dashed line denotes the persistence line, i.e. if the marker is below this line the green species will persist (not for the arithmetic method). For a more quantitative assessment of these cases see Fig. 4. For parameter settings see the Supporting information.

comparing  $\frac{\mu_i - r_i}{\mu_i}$  and  $\frac{\mu_j - r_j}{\mu_j}$  (Supporting information).

For all other cases, the specific percentage values depend on the specifics of the randomly generated communities, yet the qualitative picture, that the methods do not always agree is very consistent (Supporting information). Changes in resource supply rates do not affect niche differences when the original method was applied to the model in Eq. 9 and 10 (Letten et al. 2017), whereas it can affect niche differences for the other methods.

### 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 (Narwani et al. 2013, Godoy and Levine 2014, Kraft et al. 2015, Germain et al. 2016, Gallego et al. 2019). 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 different locations (Levine and 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 again let the scientists compute niche and fitness differences for competing annual plant species with their respective method mentioned above. The scientists then tested

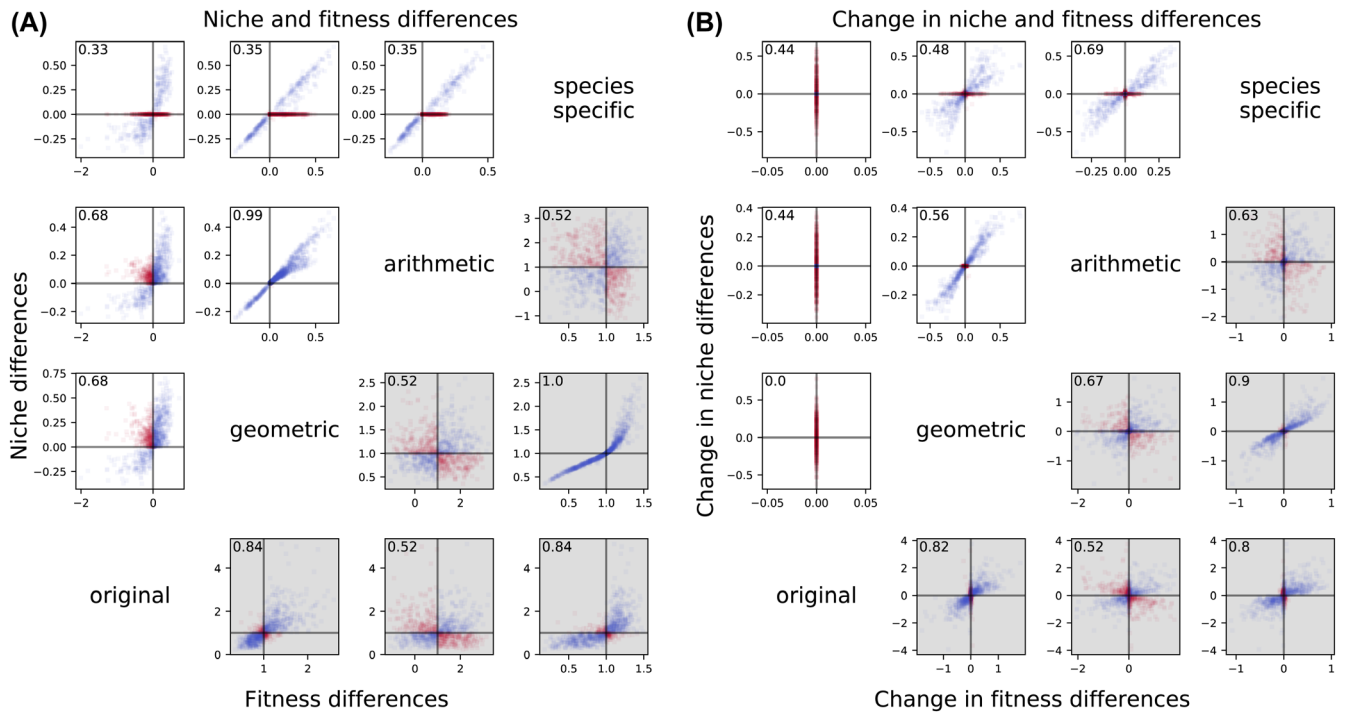


Figure 4. (A) For 1000 randomly generated communities we computed niche (above diagonal, white panels) and fitness differences (below diagonal, grey panels) according to the four investigated methods (diagonal). Each dot corresponds to one randomly generated community. We say that two methods agree qualitatively 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 (Supporting information).

which trait differences were correlated to niche and fitness differences (Fig. 5). The different methods yielded different quantitative and qualitative results about the importance of different traits.

Quantitatively, different methods predicted different correlations for relationships between traits and niche and fitness differences (Fig. 5A–B). 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 sign of the correlation between the different methods are different, as can be seen by the example of the specific leaf area or leaf nitrogen content for niche differences. However, in the investigated dataset, we never observed two methods predicting qualitatively different slopes that were both significant.

## Discussion

We have reviewed the different methods for niche and fitness differences and grouped them loosely into three groups. First,

Lotka–Volterra based methods, which have been created out of necessity for an application to a new community model, which is often similar to the original Lotka–Volterra community model; second, growth rate based methods, which have been created out of a different interpretation on how to properly define these concepts; third, structural stability based methods, which also have been created based on a different conceptual interpretation of what niche and fitness differences ought to capture. Next, we have compared different methods to compute niche and fitness differences on a theoretical and an empirical dataset. 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 (Supporting information). Consequently, one would draw different conclusions of which traits affect niche and fitness differences and how, depending on the method used.

It is 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

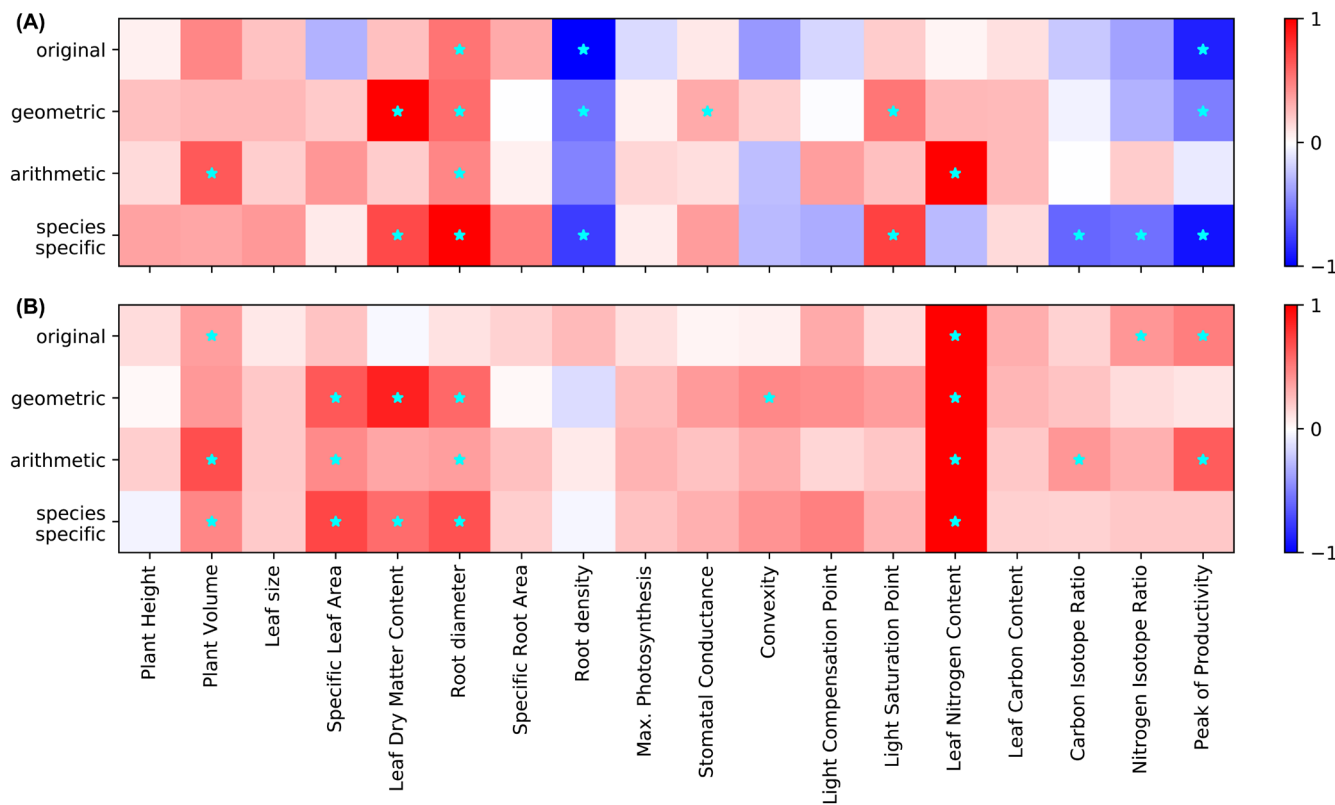


Figure 5. The correlation of the pairwise trait differences (x-axis) and the niche (A) and fitness differences (B) for the Pérez-Ramos et al. (2019) data set. We find many quantitative and qualitative differences among these correlations. For example, the original method finds that specific leaf area is not strongly correlated with fitness differences (cyan-star indicates  $p$ -value  $< 0.05$ ) while all 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. Locations with a star indicate significant correlations according to a Hochberg–Benjamini correction for multiple linear regressions.

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 expanded their work and showed that, even for simple two species communities, different methods can produce different assessments of stabilizing and equalizing mechanisms.

While the overall trends are similar among most methods (Fig. 4, 5), there is no exact correspondence between any two methods. That is, for a given system, any two methods are likely to produce different results. For example, the different methods do not agree on the frequency dependence of resource competition (Fig. 3B, dots) or on the superior competitor (Fig. 3D, triangles). Thus, comparing niche or fitness differences across studies adopting different methods is not recommended, even when the study system is identical across studies.

### Given the differences, is there one optimal method?

Intuitively, the different methods lead to different interpretations of coexistence because they capture different aspects of ecology. This is reflected by which information is included

into the calculation of niche differences and which concepts they attempt to capture (lower and/or upper bounds for niche differences, Supporting information). These conceptual differences show that different information is included in the computation of niche and fitness differences, which leads 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 (Supporting information). 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 respond to changes of resource supply rates therefore reflect 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 into the computation of a specific method. A similar understanding of the empirical data-sets can explain how different methods interpret different traits as predictors of niche and fitness differences (Supporting information).

We argue that identifying which method is the superior one is both difficult and unnecessary. 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. Doing so will permit selecting the appropriate method to address specific research questions. Indeed, variation of interpretation, and the resulting methodological variation, is found 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 (McGill et al. 2015, Donohue et al. 2016, Montoya et al. 2018). In stability research, this variation has been embraced and coined the dimensionality of stability (Donohue et al. 2016, Radchuk et al. 2019). Understanding why these methods differ and when they correspond, i.e. when dimensionality is reducible, is a major objective in stability research (Arnoldi et al. 2018, Radchuk et al. 2019, Carpentier et al. 2021, Clark 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.

### Which method should I use?

Given these many different methods, which all lead to potentially different results, a natural question is: 'Which one should I use in my research?' From a conceptual standpoint, the best method will depend on the specific question(s) one aims to address. However, we could at present not decide which method is best to answer the questions asked in the used studies (Letten et al. 2017, Pérez-Ramos et al. 2019). We argue that more theoretical research is needed to fully understand the concepts behind the various definitions of niche and fitness differences, i.e. to tease out what they really mean.

Apart from conceptually matching the research question, we argue a method should also fulfil a number of more objective criteria. We propose six criteria, three of which are theoretical (1–3) and the other three are related to their empirical applicability (4–6). Below, we elaborate on these criteria and discuss the difficulty associated with assessing them. In the Supporting information we summarize all methods based on how well they satisfy these criteria.

From a theoretical viewpoint, first, the niche and fitness differences of a method should inform us whether the two species coexist (Chesson 2000), if one species excludes the other, or if the community is determined by priority effects (Ke and Letten 2018). More generally, in a multi-species community, it should inform us which species will be present and which will be competitively excluded (Spaak et al. 2021b). Second, niche and fitness differences should inform us about the frequency dependence and the competitive hierarchy of the system, respectively (Adler et al. 2007). Third, the niche and fitness differences would optimally be given

by a simple formula that allows intuitive understanding of the system.

The first theoretical criterion is easy to verify and is illustrated in Fig. 1. However, the second theoretical criterion is very challenging to verify objectively. That is, it is difficult to independently assess frequency dependence without the use of niche and fitness differences (Spaak et al. 2021a). For example, in one of our examples the different methods yield different frequency dependence for the same community (Fig. 3B, dots), but we have no objective correct answer with which we can assess the correctness of the method. The only exception to this is that coexisting two-species communities should have negative frequency dependence and two-species communities driven by priority effects should have positive frequency dependence (Ke and Letten 2018). However, competitive exclusion can be accompanied by either positive or negative frequency dependence and we have no general method to assess its objective correctness (Spaak and De Laender 2020, Spaak et al. 2021a). The third theoretical criterion is relatively subjective, although we argue that one should be wary of methods which give overly simple expressions of niche and fitness differences for complex community models.

From an empirical viewpoint, our fourth criterion is that a method should be applicable to as many different empirical systems as possible, including species rich communities, communities with positive interactions (facilitation) and potentially even multi-trophic communities (Spaak et al. 2021b). Fifth, if the method applies to a system, then the method should be easy to apply and require as few experiments as possible (Godwin et al. 2020). Sixth, niche and fitness differences should be robust against measurement errors (Bowler et al. 2022).

The fourth criterion is again easy to verify and illustrated in Fig. 1, whereas the fifth criterion has been reviewed by Godwin et al. 2020). Generally, the Lotka–Volterra and structural stability based methods require fitting a community model to the empirical system, which implies that one needs sufficient prior knowledge of the system to choose a suitable community model. Conversely, the growth rate based models require measuring a few specific growth rates, most notably the invasion growth rate. However, in order to ensure that the invasion growth rate is indeed measured at the equilibrium of the resident community, one generally has to observe the system over time, which might limit its applicability to certain empirical systems (e.g. fast-growing microorganisms, Narwani et al. 2013, Li 2017, Grainger et al. 2019). We currently do not have much evidence for the sixth criterion since only few empirical measurements of niche and fitness differences have included uncertainty of the empirical measurements into their computation of niche and fitness differences (Grainger et al. 2019, Pérez-Ramos et al. 2019, Terry et al. 2021, Bowler et al. 2022). The few publications which include uncertainty typically show a very large uncertainty of the resulting niche and fitness differences. Having a robust method to assess niche and fitness differences is



very desirable since the theoretical differences between the methods might be insignificant compared to the empirical noise in these measurements.

Currently, no method satisfies all criteria perfectly (Supporting information). Of these criteria, we know the least about the correctness of frequency dependence and the robustness to empirical measurement errors. However, these are essential criteria since niche and fitness differences are precisely defined to inform us about the frequency dependence or the superior competitor. Consequentially, if a method of niche and fitness differences does not give us the correct answer to this question, or does not give us a precise answer due to measurement errors, we would lack strong motivation to calculate these methods.

## Conclusion

To conclude, we propose to choose the method by considering which of the above six criteria are most important for the specific research question according to our six criteria (Supporting information). Moreover, we argue that future studies should put effort into characterizing the above criteria for different methods. Specifically, we encourage more studies investigating the uncertainty of niche and fitness differences, such as Bowler et al. (2022) or Terry et al. (2021), as well as studies investigating the frequency dependence of the methods, such as Spaak and De Laender (2020) and Spaak et al. (2021a). Thereby, we can arrive at a more general and concrete guidelines for choosing suitable methods for the research question in hand.

*Acknowledgements* – We thank 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. Open Access funding enabled and organized by Projekt DEAL.

*Funding* – JWS received support from Schweizerischer Nationalfonds Early Post-doc mobility (P2SKP3\_194960) FDL and JWS received support from grant of the Fund for Scientific Research, FNRS (PDR T.0048.16). P-JK received support from the Yushan scholar program of Taiwan MOE (NTU-110VV010).

## Author contributions

**Jürg W. Spaak:** Conceptualization (equal); Formal analysis (lead); Software (lead); Visualization (lead); Writing – original draft (lead); Writing – review and editing (equal). **Po-Ju Ke:** Conceptualization (equal); Writing – review and editing (equal). **Andrew D. Letten:** Conceptualization (equal); Writing – review and editing (equal). **Frederik De Laender:** Conceptualization (equal); Visualization (equal); Writing – review and editing (equal).

## Data availability statement

Data are available from the Dryad Digital Repository: <<https://doi.org/10.5061/dryad.9cnp5hqnr>> (Spaak et al. 2022).

## Supporting information

The Supporting information associated with this article is available with the online version.

## References

- Adler, P. B., HilleRisLambers, J. and Levine, J. M. 2007. A niche for neutrality. – *Ecol. Lett.* 10: 95–104.
- Arnoldi, J. F., Bideault, A., Loreau, M. and Haegeman, B. 2018. How ecosystems recover from pulse perturbations: a theory of short- to long-term responses. – *J. Theor. Biol.* 436: 79–92.
- Bimler, M. D., Stouffer, D. B., Lai, H. R. and Mayfield, M. M. 2018. Accurate predictions of coexistence in natural systems require the inclusion of facilitative interactions and environmental dependency. – *J. Ecol.* 106: 1839–1852.
- Bowler, C. H., Weiss-Lehman, C., Towers, I. R., Mayfield, M. and Shoemaker, L. G. 2022. Accounting for demographic uncertainty increases predictions for species coexistence: a case study with annual plants. – *Ecol. Lett.* 25: 1618–1628.
- Buche, L., Spaak, J. W., Jarillo, J. and De Laender, F. 2022. Niche differences, not fitness differences, explain predicted coexistence across ecological groups. – *J. Ecol.* 1365–2745.
- Carmel, Y., Suprunenko, Y. F., Kunin, W. E., Kent, R., Belmaker, J., Bar-Massada, A. and Cornell, S. J. 2017. Using exclusion rate to unify niche and neutral perspectives on coexistence. – *Oikos* 126: 1451–1458.
- Carpentier, C., Barabás, G., Spaak, J. W. and De Laender, F. 2021. Reinterpreting the relationship between number of species and number of links connects community structure and stability. – *Nat. Ecol. Evol.* 5: 1102–1109.
- Carroll, I. T., Cardinale, B. J. and Nisbet, R. M. 2011. Niche and fitness differences relate the maintenance of diversity to ecosystem function. – *Ecology* 92: 9.
- Chesson, P. 1990. MacArthur's consumer–resource model. – *Theor. Popul. Biol.* 37: 26–38.
- Chesson, P. 2000. Mechanisms of maintenance of species diversity. – *Annu. Rev. Ecol. Syst.* 31: 343–366.
- Chesson, P. 2003. Quantifying and testing coexistence mechanisms arising from recruitment fluctuations. – *Theor. Popul. Biol.* 64: 345–357.
- Chesson, P. 2013. Species competition and predation. – In: Chesson, P. (ed.), *Ecological systems*. Springer, New York, NY, 2013. pp. 223–256.
- Chesson, P. and Kuang, J. J. 2008. The interaction between predation and competition. – *Nature* 456: 235–238.
- Clark, A. T., Arnoldi, J., Zelnik, Y. R., Barabás, G., Hodapp, D., Karakoç, C., König, S., Radchuk, V., Donohue, I., Huth, A., Jacquet, C., Mazancourt, C., Mentges, A., Nothhaaf, D., Shoemaker, L. G., Taubert, F., Wiegand, T., Wang, S., Chase, J. M., Loreau, M. and Harpole, S. 2021. General statistical scaling laws for stability in ecological systems. – *Ecol. Lett.* 24: 1474–1486.
- 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. and Yang, Q. 2016. Navigating the complexity of ecological stability. – *Ecol. Lett.* 19: 1172–1185.
- Gallego, I., Venail, P. and Ibelings, B. W. 2019. Size differences predict niche and relative fitness differences between phytoplankton species but not their coexistence. – *ISME J.* 13: 1133–1143.



- Gause, G. F. 1934. Experimental demonstration of Volterra's periodic oscillations in the numbers of animals. – *Science* 79: 16–17.
- Germain, R. M., Weir, J. T. and Gilbert, B. 2016. Species coexistence: macroevolutionary relationships and the contingency of historical interactions. – *Proc. R. Soc. B* 283: 20160047.
- Godoy, O. and 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. and Cardinale, B. J. 2020. An empiricist's guide to modern coexistence theory for competitive communities. – *Oikos* 129: 1109–1127.
- Grainger, T. N., Letten, A. D., Gilbert, B. and Fukami, T. 2019. Applying modern coexistence theory to priority effects. – *Proc. Natl Acad. Sci. USA* 116: 6205–6210.
- Hubbell, S. P. 2005. Neutral theory in community ecology and the hypothesis of functional equivalence. – *Funct. Ecol.* 19: 166–172.
- Hutchinson, G. E. 1959. Homage to Santa Rosalia or why are there so many kinds of animals? – *Am. Nat.* 93: 145–159.
- Johnson, C. A. 2021. How mutualisms influence the coexistence of competing species. – *Ecology* 102: e03346.
- Kandlikar, G. S., Johnson, C. A., Yan, X., Kraft, N. J. B. and Levine, J. M. 2019. Winning and losing with microbes: how microbially mediated fitness differences influence plant diversity. – *Ecol. Lett.* 22: 1178–1191.
- Ke, P. J. and Letten, A. D. 2018. Coexistence theory and the frequency-dependence of priority effects. – *Nat. Ecol. Evol.* 2: 1691–1695.
- Koffel, T., Daufresne, T. and Klausmeier, C. A. 2021. From competition to facilitation and mutualism: a general theory of the niche. – *Ecol. Monogr.* 91: e01458.
- Kraft, N. J. B., Godoy, O. and Levine, J. M. 2015. Plant functional traits and the multidimensional nature of species coexistence. – *Proc. Natl Acad. Sci. USA* 112: 797–802.
- Letten, A. D., Ke, P. J. and Fukami, T. 2017. Linking modern coexistence theory and contemporary niche theory. – *Ecol. Monogr.* 87: 161–177.
- Levine, J. M. and 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 R<sup>2</sup>, why not? Then what? – *PLoS One* 12: e0183250.
- Luo, M., Wang, S., Saavedra, S., Ebert, D. and Altermatt, F. 2022. Multispecies coexistence in fragmented landscapes. – *Proc. Natl Acad. Sci. USA* 119: e2201503119.
- McGill, B. J., Dornelas, M., Gotelli, N. J. and Magurran, A. E. 2015. Fifteen forms of biodiversity trend in the Anthropocene. – *Trends Ecol. Evol.* 30: 104–113.
- Mészána, G., Gyllenberg, M., Pásztor, L. and Metz, J. A. 2006. Competitive exclusion and limiting similarity: a unified theory. – *Theor. Popul. Biol.* 69: 68–87.
- Montoya, J. M., Donohue, I. and Pimm, S. L. 2018. Planetary boundaries for biodiversity: implausible science, pernicious policies. – *Trends Ecol. Evol.* 33: 71–73.
- Narwani, A., Alexandrou, M. A., Oakley, T. H., Carroll, I. T. and Cardinale, B. J. 2013. Experimental evidence that evolutionary relatedness does not affect the ecological mechanisms of coexistence in freshwater green algae. – *Ecol. Lett.* 16: 1373–1381.
- Pérez-Ramos, I. M., Matías, L., Gómez-Aparicio, L. and Godoy, O. 2019. Functional traits and phenotypic plasticity modulate species coexistence across contrasting climatic conditions. – *Nat. Commun.* 10: 2555.
- Petry, W. K., Kandlikar, G. S., Kraft, N. J. B., Godoy, O. and Levine, J. M. 2018. A competition-defence trade-off both promotes and weakens coexistence in an annual plant community. – *J. Ecol.* 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. and Kramer-Schadt, S. 2019. The dimensionality of stability depends on disturbance type. – *Ecol. Lett.* 22: 674–684.
- Saavedra, S., Rohr, R. P., Bascompte, J., Godoy, O., Kraft, N. J. B. and Levine, J. M. 2017. A structural approach for understanding multispecies coexistence. – *Ecol. Monogr.* 87: 470–486.
- Song, C., Barabás, G. and Saavedra, S. 2019. On the consequences of the interdependence of stabilizing and equalizing mechanisms. – *Am. Nat.* 194: 627–639.
- Spaak, J. W. and De Laender, F. 2020. Intuitive and broadly applicable definitions of niche and fitness differences. – *Ecol. Lett.* 23: 1117–1128.
- Spaak, J., Millet, R., Ke, P. J., Letten, A. D. and De Laender, F. 2021a. The effect of non-linear competitive interactions on quantifying niche and fitness differences. – *Ecology*.
- Spaak, J. W., Godoy, O. and De Laender, F. 2021b. Mapping species niche and fitness differences for communities with multiple interaction types. – *Oikos* 130: 2065–2077.
- Spaak, J. W., Ke, P.-O., Letten, A. D. and De Laender, F. 2022. Data from: Different measures of niche and fitness differences tell different tales. – *Dryad Digital Repository*, <<https://doi.org/10.5061/dryad.9cnp5hqnr>>.
- Terry, J. C. D., Chen, J. and Lewis, O. T. 2021. Natural enemies have inconsistent impacts on the coexistence of competing species. – *J. Anim. Ecol.* 90: 2277–2288.
- Tilman, D., Kilham, S. S. and Kilham, P. 1982. Phytoplankton community ecology: the role of limiting nutrients. – *Annu. Rev. Ecol. Syst.* 13: 349–372.
- Yamamichi, M., Gibbs, T. and Levine, J. M. 2022. Integrating eco-evolutionary dynamics and modern coexistence theory. – *Ecol. Lett.* 25: 2091–2106.
- Zhao, L., Zhang, Q. and Zhang, D. 2016. Evolution alters ecological mechanisms of coexistence in experimental microcosms. – *Funct. Ecol.* 30: 1440–1446.