**An empiricist’s guide to modern coexistence theory for competitive communities**

**List of authors:** Feng-Hsun Chang1\*, Casey M. Godwin1,2, Bradley Cardinale,1,2

1School for Environment and Sustainability, University of Michigan, 440 Church street, Ann Arbor, Michigan, USA

2Cooperative Institute for Great Lakes Research (CIGLR), University of Michigan, 440 Church street, Ann Arbor, Michigan, USA

**\*Correspondence:**

Feng-Hsun Chang

1School for Environment and Sustainability, University of Michigan, 1040 Dana, Ann Arbor, Michigan, 48105, USA

fhchang@umich.edu

Total word count for main body of text:

Number of figures:

Number of tables:

**Abstract**

Prior to 2000, nearly all ecological theory argued that the coexistence of species in competitive communities was controlled by the magnitude of niche differences among species in space or time. In 2000, Chesson published a more general theory of coexistence in competitive communities, showing that species coexistence in most models of competition is controlled by the balance of two opposing forces: Relative fitness differences (RFD) establish competitive hierarchies among species that hasten competitive exclusion, whereas niche differentiation (ND) helps offset the effects of RFD by giving species a growth advantage when rare, which helps to stabilize negative interactions. Even since Chesson’s general framework of coexistence was established, an increasing number of empirical studies have attempted to quantify the magnitude of ND and RFD for real assemblages of plants and animals. However, because the various methods used to measure ND and RFD to date have been derived from different theoretical origins and make differing assumptions, they are not directly comparable in terms of the quantities and predictions they produce. In this review, we compare five of the most common methods that have been used to measure ND and RFD: (1) negative frequency dependence (NFD), (2) sensitivity to mutual invasibility (Sensitivity method), (3) parameterization of Lotka-Volterra competition models (Lotka-Volterra Method), (4) MacArthur’s consumer resource model (MacArthur’s CRM), and (5) Tilman’s consumer resource model (Tilman’s CRM). We show that ND and RFD can be measured using all of five methods, and that the measures correctly predict species coexistence under the specific limiting assumptions of those methods. However, the methods are not comparable to one other, which means that measures of ND and RFD estimates from different studies may not be comparable. Furthermore, several of the methods have important assumptions or limitations that empiricists need to be aware of before using them to measure ND and RFD in any real system. After reviewing the methods, we offer guidance for empiricists to determine which method(s) are most appropriate for measuring ND and RFD in their study system.

**Introduction**

Throughout most of the history of community ecology, it has been assumed that niche differentiation among species is the sole biological mechanism that can offset the negative impacts of interspecific competition and prevent competitive exclusion (Volterra 1926, Gause 1934, May and Macarthur 1972, Abrams 1990, Chesson 1991, Leibold 1995, Chase and Leibold 2003, Tilman 2004). This idea dates at least to 1928 when Volterra (Volterra 1926) introduced a dynamic model of competition that became the foundation for the competitive exclusion principle (Gause 1934). The competitive exclusion principle led to the conclusion that ecosystems should only contain as many species as there are limiting resources (or consumers); thus, species only coexist if they are limited by different resources (or consumers) at the same location and time, or if they partition resources (or consumers) in space or time (Rescigno and Richards 1965, Macarthur and Levins 1967, Macarthur 1970, Abrams 1986, Leibold 1995). Nearly all subsequent hypotheses to explain coexistence have argued that biodiversity exists because of such niche differences among species.

While niche partitioning has been the foundation for most explanations of the world's biodiversity, theories of species coexistence have recently undergone a major revision. In 2001, Hubbell published The Unified Neutral Theory of Biodiversity (Hubbell 2001), which argued that patterns of biodiversity in nature can be explained by a relatively simple model that does not invoke niche differences among species. According to Hubbell’s theory, species coexist not because they are different, but because their demographic parameters are identical, or nearly so, such that the consequences of their interactions are 'neutral' (i.e. essentially equal among all species). As such, Hubbell argued that the biodiversity we observe in nature can be explained by a series of stochastic events that give some populations the chance opportunity to rise to dominance while others exhibit random walks to extinction.

Even as Hubbell was developing his neutral theory, Chesson (Chesson 2000) was completing a ground-breaking synthesis of coexistence framework that would ultimately provide a road map for how to integrate the niche and neutral perspectives on biodiversity. Chesson’s coexistence framework was built on his insight to the invisibility criterion, i.e. whether a species can have positive growth rate when invading resident species from rare (Chesson 1994, 2000). If both species have positive growth rates as the invader, the two species will coexist. Chesson argued that, theoretically, species’ invasion growth rate can be decomposed into two general terms representing fluctuation independent and fluctuation dependent mechanisms (Chesson 2000, 2003). Fluctuation independent mechanisms are those that do not require fluctuations in population densities or environmental factors in order to facilitate coexistence. On the other hand, fluctuation dependent mechanisms occur through species’ differential responses to the variance of environmental factors, i.e. relative nonlinearities (e.g. Letten et al. 2018), or operate through the covariance between environment and species’ interactions, i.e. storage effects (Miller and Klausmeier 2017).

The fluctuation independent mechanisms in Chesson’s coexistence framework include two forces, which he called stabilizing and equalizing. Stabilizing forces represent various forms of niche difference (ND) among a pair of species, all of which cause species to limit their own growth more strongly than they limit the growth of other species (intra > interspecific competition). This can occur when species partition limiting resources in space or time, or when they experience differential consumption by shared consumers. In contrast, equalizing forces minimize what Chesson and others now refer to as relative fitness differences (RFD) among species (Adler et al. 2007, Levine and HilleRisLambers 2009, Hillerislambers et al. 2012, Turnbull et al. 2013). Chesson's definition of a ‘fitness difference’ is not the same as that used by evolutionary biologists, as he was referring to differences in competitive abilities among species, not differences in reproduction among individuals. RFDs set up competitive hierarchies among species when ND is absent, and are the result of inherent variation in biological traits such as minimum resource or consumer requirements (e.g. R\*/P\*’s; Tilman 1982), differential resistance to consumers, or differences in potential growth rates (Adler et al. 2007, Levine and HilleRisLambers 2009, Hillerislambers et al. 2012). RFDs represent differences in competitive abilities that persist irrespective of how much ND there is among species. For all fluctuation independent mechanisms of coexistence, Chesson showed it is the balance of these two forces – RFDs that establish competitive hierarchies, and NDs that prevent competitive exclusion – that ultimately determine whether species maintain non-negative long-term growth rates in competitive communities (Chesson 2000). To coexist, ND must be sufficiently large to offset and stabilize the competitive hierarchies generated by RFD’s. It has subsequently been shown that Hubbell's neutral theory represents a specific, limiting case of Chesson's coexistence theory where NDs and RFDs are both zero, causing the outcome of competition to be approximated by a random walk to extinction (Adler 2007). Thus, Chesson’s inequality provides a general framework for predicting species coexistence.

Much attention in ecology has recently turned towards the empirical measurement of niche and relative fitness differences to determine how these two forces contribute to coexistence in real communities. As a growing number of empiricists have tried to quantify ND and RFD in their individual study system, the number of different empirical approaches proposed for doing so has also grown. For example, based on empirical measurements of population dynamics, niche difference can be quantified and demonstrated to be important to maintain biodiversity in annual plant communities (Levine and HilleRisLambers 2009). Moreover, by comparing species’ population dynamics with versus without a competitor, one can quantify ND and RFD between species pairs and to predict coexistence in green algae communities (Narwani et al. 2013). However, these methods stem from very different models of species coexistence, make different assumptions, and use different experimental designs. Therefore, it is important to compare these methods and how they would be implemented in order to determine whether these methods give comparable values of ND and RFD as well as offer the same predictions regarding coexistence.

Here we provide a summary and comparison of five common methods that have been used to measure ND and RFD empirically in real communities. In part 1 of our paper, we describe the theoretical background of each method and how it could be implemented empirically, compare the methods in terms of how they relate to ND and RFD, and use simulations to ask whether they give the same estimates and predictions regarding coexistence. In Part 2 of the paper we provide a list of decision steps to help empiricists select the most appropriate method for their study system and contrast the methods in terms of the amount of empirical information (e.g. experimental treatments) required to use them. In part 3, we discuss the main advantages and disadvantages of the approaches and make some suggestions for future empirical work on coexistence theory.

**Part 1. Theoretical background for five empirical methods**

1. *Theoretical background for five empirical methods*

In Chesson’s coexistence framework, niche difference (ND) and relative fitness difference (RFD) are being used in the following inequality to assess the mutual invasibility criterion (Chesson 1990).

(1)

In this inequality, *ρ* is the niche overlap, or one minus niche difference (ND), and is the relative fitness difference (RFD). The niche difference (ND; ), relative fitness difference (RFD; ) and the inequality (equation 1) are first derived by using the classic Lotka-Volterra competition model to approximate MacArthur’s consumer resource model (Chesson 1990). Since ND, RFD and the inequality for mutual invasibility are based on the Lotka-Volterra competition model, we first describe how each method can be used to approximate the interaction coefficients in the Lotka-Volterra competition model in order to measure ND and RFD for assessing the inequality (equation 1). We then explain the empirical approach required for each method and end each method section with some of the known limitations of each method. In the end of this part, we conclude by comparing these methods in terms of whether these methods yield the same prediction for coexistence based on the mutual invasibility criteria and whether they give the same estimates of ND and RFD.

* 1. *Negative frequency dependence (NFD)*
     1. *Theoretical background*

The negative frequency dependence method quantifies the relationship between a species’ frequency in a community (individuals of species 1 / total individuals of all species) and its per capita growth rate (Adler et al. 2007, Levine and HilleRisLambers 2009). The relationship between a species’ growth rate and its relative frequency has been used to assess the mutual invasibility criterion because the magnitude of the frequency dependency should reflect the difference between intra- versus inter-specific competition. The magnitude of the frequency dependency can be used to reflect the difference between intra- versus inter-specific competition relies on a key assumption that the community is saturated with respect to density or biomass (Adler et al. 2007). Under this assumption, all resources or niches are occupied by either the focal species *i* or its competitor(s). Increasing species *i*’s frequency means the individuals of species *i* will compete more with individuals of its own kind than with individuals of other kinds, and will thus experience more intra- than inter-specific competition. Therefore, if intra-specific competition is greater than inter-specific competition, the dependency of species’ growth rate on its relative frequency should be negative, i.e. the NFD slope should be negative. According to Adler et al. 2007, more negative NFD slope thus represents a stronger stabilizing force, i.e. niche difference (ND). In addition, Adler et al. 2007 argue that the difference between species’ growth rate in the absence of stabilizing forces is the equalizing force, i.e. relative fitness difference (RFD; Figure 1 in Adler et al. 2007), and the RFD is proportional to the NFD intercept. As long as the relationship between a species’ frequency and its growth rate is linear, the NFD slope and intercept could theoretically be used to evaluate Chesson’s inequality (equation 1) and be used to predict whether both species will have positive growth rates when rare, thus meeting the mutual invasibility criterion.

* + 1. *Empirical approaches*

It is important to note that the slope of NFD has not been used literally as a measure of ND, nor has the intercept of NFD been used as a literal measure of RFD. Instead, papers have used estimates of inter and intraspecific competition coefficients from a Lotka Volterra-like model to quantify ND and then experimentally remove ND through manipulation to illustrate its effect as a stabilizing force (Levine and HilleRisLambers 2009). Other studies have measured NFD as evidence for the importance of stabilizing forces, but do not directly interpret the slope as ND or the intercepts as RFD (Yenni et al. 2017a). Nonetheless, the NFD method could be applied empirically to predict coexistence based on the mutual invasibility principle.

Levine et al [Levine and HilleRisLambers 2009 Nature] demonstrated how the NFD method can be implemented empirically. In their experiment with 10 species of grassland plants, they manipulated the relative frequency of each focal species by varying the proportion of seeds belonging to the focal species versus all other species. At the end of the growing season, they quantified the growth rate of each species in the plots by enumerating the number of seeds belonging to each species and the proportion of those seeds that were viable the next year. They then quantified the slope of NFD by plotting the growth rate of each species against its frequency in the initial community. Although the slope of NFD is not equal to Chesson’s ND and the difference in intercepts is not equal to RFD, they show how the impact of niche differences can be displayed by eliminating this mechanism. Specifically, they show that the effect of niche differences on growth rates can be removed by experimentally maintaining species densities at non-equilibrium levels.

Additionally, frequency dependency could, in principle, also be constructed using observational data from natural ecosystems. Doing this approach allows an empiricist to construct frequency dependency on species that are not easy to manipulate their growth rate. As long as the per capita growth rate and the relative frequency of the species can be estimated, and the community is assumed to be saturated with respect to density or biomass, it would be possible to parameterize the relationship between growth rate and frequency. Although this approach has not been applied empirically to make predictions regarding coexistence (but see Yenni et al. 2012, 2017), it is one of only two methods reviewed here that do not require manipulative experiments.

* + 1. *Limitations*

Despite some of the desirable aspects of the NFD method, there is a major caveat of this method. Namely, this method requires that the relationship between growth rate and frequency is linear, or that an empiricist can adequately describe the relationship to account for the non-linearity in predict growth rates when invading. From numerical simulations that mimic invasion experiments, we show that the species’ per capita growth rate is non-linearly dependent on its frequency, i.e. the NFD slope is not constant across frequencies even when the total community biomass is saturated (Fig. 1). The NFD slope is only constant when the community density is not only saturated but is always fixed regardless of the species’ frequency (appendix A). According to Appendix A, even though the NFD slope is constant, the NFD slope and intercept cannot be used to calculate ND and RFD. We also show that the non-linear frequency dependence can lead to incorrect predictions about species coexistence when applied to systems with non-linear relationships between species’ growth rates and densities (Supplement 1). When the NFD slope is not constant, the NFD method can only have correct predictions when evaluated at frequency approaching zero for each species. This finding means that for an empiricist to use the NFD method, they would need to either 1) measure the growth rate of each species across the full range of frequencies to establish that the growth rate of each species is linearly related to its frequency or 2) evaluate the growth rate of each species when rare (i.e. directly demonstrate mutual invasibility). Both of these options would dramatically increase the effort required but may be necessary in systems where only observational studies are possible.

* 1. *Parameterizing the Lotka-Volterra competition model*
     1. *Theoretical background*

The second empirical method to measure ND and RFD is based on parameterizing the Lotka-Volterra model empirical data from experiments or time series of species in natural ecosystems. In the classic Lotka-Volterra model, the *per capita* growth rate of species *i* can be described by equation 2:

(2)

In equation 2, *αii* is the *per capita* intra-specific competition coefficient, which describes the *per capita* effect of species *i* on the *per capita* growth rate of species *i*. *Ni*, is the density of species *i* scaled to its carrying capacity, and *ri* is the intrinsic per capita growth rate. Following Chesson’s deduction (Chesson 1990, 2000), ND can be calculated as and RFD can be calculated as. After obtaining ND and RFD, Chesson’s equality (equation 1) can be assessed for predicting species coexistence.

* + 1. *Empirical approaches*

To use the Lotka-Volterra model to empirically predict coexistence requires that the empiricist assume that the per capita strength of species interactions is independent of species densities, i.e. the competition coefficient can be linearly modeled. Under this assumption, there are six different parameters that must be quantified in order to populate equation 2: intrinsic per capita growth rate of each species (*ri* and *rj*), per capita intra-specific competition coefficients ( and ), and per capita inter-specific competition coefficients ( and ). The simplest way to parameterize the model would be to use three datasets for each species pair: a time-series of each species grown alone as a monoculture to obtain intrinsic per capita growth rate of each species as well as per capita intra-specific competition coefficients, and one time-series representing a co-culture of the two species to obtain per capita inter-specific competition coefficients (Fig. 2). From each monoculture time series, the empiricist needs to measure the maximum per capita growth rate of each species (*ri*), which occurs as the species’ density approaches zero. Then, the empiricist needs to measure the steady-state population size of each species in monoculture (*Ki*). The quotient is equal to the intra-specific interaction term in the Lotka-Volterra model (Equation 2 but *Nj* equals to zero). After quantifying the intra-specific interaction terms, the empiricist must perform two competition experiments in which the density and growth rate of each species is measured when each species is near their monoculture equilibrium density (See Appendix 2, Figure S2). Using these time series from competition, the empiricist can quantify both inter-specific competition coefficients ( and ) by substituting the previously quantified intraspecific interaction terms into equation 2. Having estimated all four interaction coefficients, the empiricist can calculate ND and RFD for assessing Chesson’s inequality.

* + 1. *Limitations*

An inherent limitation associated with the Lotka-Volterra method is that it assumes that the intra- and inter-specific competition coefficients are constant with respect to population sizes of either species. In other words, the species’ first individual and the last individual have the same per capita effect on the growth rates of its own kind or its competitor. However, intra- and inter-specific competition coefficients are easily subject to species’ density (e.g. Smith-Gill and Gill 1978), so that constant intra- and inter-specific competition coefficients might be unlikely occur in practice (Abrams 1980). Errors in estimating intra- and inter-specific competition coefficients can lead to incorrect calculation of ND and RFD. One solution is to make the assumption that intraspecific competition coefficients (*αii*) is fixed at 1/*Ni\**, where *Ni\** is the equilibrium density of the focal species. Supplement figure S2 shows that when applied to the numerical simulations based on Tilman’s CRM, the method using assumed intraspecific competition coefficients accurately predicts coexistence whereas intraspecific coefficients measured in monoculture near equilibrium lead to inaccurate predictions regarding coexistence.

* 1. *Sensitivity measurement in mutual invasibility experiments*
     1. *Theoretical background*

A third empirical method for measuring ND and RFD is to conduct mutual invasibility experiments that evaluate whether each species can invade a steady-state population of the other from rarity (Fig. 3). In the mutual invasibility experiment, species *i*’s sensitivity to competition () is calculated by its per capita maximum growth rate in monoculture () and its invasion per capita growth rate () when being introduced to the steady-state biomass of the competitor (species *j*) with the following equation.

(3)

The sensitivity to competition () is thus a measure of per capita growth rate reduction caused by a steady-state biomass of the competitor. A species’ sensitivity to competition is jointly determined by ND and RFD (Carroll et al. 2011, Narwani et al. 2013). Specifically, greater ND between the two species reduces the sensitivity of both species’ to competition, so that *Si* will approach zero. Greater RFD, on the other hand, causes species to be asymmetrically affected by competition such that one species’ sensitivity increases while the other’s decreases. Further, Carroll et al. (2011) verbally argue that ND can be calculated as and RFD as , where *Si* is the sensitivity of the species with the greater sensitivity. In appendix B, we show that such argument is valid and derive an inequality for coexistence expressed with the sensitivity metrics (*Si*), , which is in the same form as in Chesson’s coexistence framework (appendix B).

* + 1. *Empirical approaches*

Like the NFD method, the Sensitivity method does not require an empiricist to assume a specific model of species interactions or define the resources that species compete for. Rather, this method uses a combination of monocultures and pairwise invasion experiments to quantify the reduction in each species’ growth rate caused by a steady-state biomass of the other species. The experiment by Narwani et al. (Narwani et al. 2013) provides an example for how to implement the Sensitivity method empirically. Their experimental system involved species of freshwater green algae growing under controlled conditions in the laboratory, where each species can be grown either as a monoculture or in combination. In their invasion experiment, they first grew each species as a monoculture in fresh growth medium from low biomass to its steady-state, where the population biomass did not have significant changes through time. From these time series, they quantified the per capita maximum growth rate of each species as a monoculture at low biomass ( and ). After species reached the steady-state biomass, they introduced the other species from low biomass, e.g. 0.01% of the resident species, and quantified the per capita invasion growth rate of the invader species ( and ). Finally, for each pair of species, they used these growth rates to calculate the sensitivity metrics () by equation 3. The are then used to evaluate Chesson’s inequality. Note that Narwani et al. 2013 ordered the sensitivity metrics () in order to simplify Chesson’s inequality.

It is important to note that the invasion growth rate ( and ) should be measured when the invader population biomass approaches zero, i.e. intra-specific competition is negligible, and the resident species’ biomass is still at steady-state. If the growth rate of the invader species were measured at larger biomass of the invader species (i.e. long after invasion), then the will be affected by both intra- and inter specific competition. The resulting predicting regarding species coexistence would not be correct. It is also important to note that the Sensitivity method is only applicable to competitive interactions, under which Chesson’s inequality is only applicable. Finally, when conducting invasion experiments to measure sensitivity () and to calculate ND and RFD, it is not necessary to quantify species’ steady-state density or biomass as in the Lotka-Volterra Method. It is also not required to monitor the growth rate of the resident species after the other species is introduced.

* + 1. *Limitations*

Using the sensitivity metric (*Si*) to quantify the impact of one species on another species requires one to perform mutual invasibility experiments, which is only practical for organisms whose population growth rates can be measured over tractable periods of time (i.e. days for fast-growing organisms like bacteria, algae, invertebrates, etc.). Mutual invasibility experiments are harder to apply to organisms that grow slowly since measuring their per capita growth rates would take a long time (e.g. trees). But, those are good estimates of growth rates since the densities are low. Also, one could use the seed-bank approach here.

* 1. *Parameterizing MacArthur’s consumer resource model*
     1. *Theoretical background*

Another approach is to first parameterize a mechanistical model that determines the mechanisms by which species interact and then reorganize the parameters to estimate ND and RFD for assessing Chesson’s inequality (equation 1). A mechanistical model that has been reorganized is MacArthur’s consumer resource model (equation 4 to 5), which describes how consumer species consume and thus compete for prey resource (MacArthur 1969, 1970).

(4)

(5)

In equation 4 and 5, is population density of the consumer species , is population density of the prey resource , is the per capita consumption rate of consumer species on prey resource , is the yield of species ’s population density or mass relative to each prey resource consumed, is the mortality of species , and and are the per capita maximum growth rate and carrying capacity of prey resource .

Chesson has showed that, by implementing time scale separation technique, parameters in MacArthur’s consumer resource model can be used to estimate inter and intraspecific interaction coefficients for calculating ND and RFD in Chesson’s inequality (Chesson 1990, 2000). Specifically, the following two equations show how to calculate ND and RFD from parameters in MacArthur’s consumer resource model.

(6)

(7)

Through equation 6 and 7, empirically measured parameters in MacArthur’s consumer resource model can be used to calculate niche difference (ND) that describe how similar two species are with respective to using resources (Fig. 4) and relative fitness difference (RFD). Finally, Chesson’s coexistence inequality can be accessed when the parameters in MacArthur’s consumer resource model is obtained.

* + 1. *Empirical approaches*

Using MacArthur’s consumer resource model to estimate ND and RFD for a pair of species requires quantifying 1) the per capita consumption rate of each consumer species on each prey resource (, 2) the per capita maximum growth rate and carrying capacity of each prey resource when no consumers are present ( and ), and 3) the yield of consumer population density or mass relative to each prey item consumed (). Since there are no examples of empirical studies that have parameterized the MacArthur model for the purpose of estimating ND and RFD, we describe the experimental approach that would be required. First, the experimentalist would need to identify or define the prey resources that are available to the consumer species. Each prey resource would need to be inoculated or planted at low density into an environment free of other prey resources and consumers, then the population density would be measured over time to estimate the per capita maximum growth rate ( , which occurs as the prey population density approaches zero) and the carrying capacity (, which occurs when the prey growth rate approaches zero). Next, the experimentalist would need to introduce each consumer species into different densities of each prey resource monoculture. Under different prey resource densities, the experimentalist would measure the per capita consumption rate of prey resource by the consumer species (; inset figure of Fig. 4) and the yield of consumer density or mass per unit prey resource consumed (). For example, for two consumer species and three prey resources, the hypothetical experiment amounts to 18 parameters that must be quantified: 3 different maximum per capita growth rates and 3 carrying capacities of the prey resources ( and , = 1 to 3), 6 per capita consumption rates (; = 1 and 2, = 1 to 3), and 6 yields (; = 1 and 2, = 1 to 3). These parameters can then be used in Equation 6 and 7 to assess Chesson’s inequality.

* + 1. *Limitations*

Parameterizing MacArthur’s consumer resource model not only gives a more mechanistic understanding of species’ competition but also allows one to predict coexistence for other new pairs of species. However, these desirable aspects come with greater experimental effort compared to other methods. In particular, this method requires as many feeding experiments as there are resources, each of which involves measuring consumption rates at a range of resource species densities (Figure 4 inset). While this constraint does not impact the ability of the method to predict coexistence under defined conditions, it does limit the extent to which those predictions can be applied to different environmental circumstances.

* 1. *Parameterizing Tilman’s consumer resource model*
     1. *Theoretical background*

Similar to parameterizing MacArthur’s consumer resource model, Tilman’s consumer resource model (Tilman 1977) can also be empirically parameterized and reorganized to a Lotka-Volterra form (Letten et al. 2017). Recently, Letten et al. 2017 reorganize the following Tilman’s two-species consumer resource model for two essential and non-substitutable resources (equation 8 and 9) to a Lotka-Volterra form.

(8)

(9)

In equation 8 and 9, and is the population density of focal species and resource respectively. Replacing with lead to the dynamics of species and resource . () is the maximum per capita growth rate of species (), () is the half saturation constant for species limited by resource (), and () represents the constant mortality of the consumer (). *D* is the turnover rate of resources and the density-independent loss rate for both species and () is the supply concentration for resource (). Finally, is the consumption term of consumer species on resource , so it also contains a parameter that represent the yield of consumer species per resource .

According to Letten et al 2017, the above parameters can be used to calculate ND and RFD as following.

(10)

(11)

With equation 10 and 11, Chesson’s inequality (equation 1) can be assessed from Tilman’s consumer resource model for predicting species coexistence.

* + 1. *Empirical approaches*

There are few empirical attempts to parameterize Tilman’s consumer resource model. For example, in the paper where Tilman first proposed the model, Tilman had measured the parameters describing how two algae species (*Asterionella formosa* and *Cyclotella meneghiniana*) compete for two essential resources (silicate and phosphate) that are delivered to the experimental system at a constant supply rate that matches the death rate (Tilman 1980, 1981). We use this chemostat system to describe how this approach could be implemented empirically. First, the empiricist would need to quantify the dependence of each species on each resource using batch cultures. For instance, Tilman inoculated each species as a monoculture into growth medium containing a range of concentrations of the limiting resource (either silicate or phosphate) with all other resource in excess. From these time series of population densities, he quantified the growth rate of the two species at each concentration of the limiting resource and used non-linear regression to fit the Monod function to these data (Equation 8, Figure 5). This yields estimates of the 4 half saturation constants () and 2 maximum per capita growth rates of species () for each combination of species ( = 1 and 2) and the limiting resource ( = 1 and 2). Separately, Tilman quantified the 4 yields () of each species on each resource by measuring the elemental content of a known number of cells ( = 1 and 2 and = 1 and 2). By following the same approach as Tilman, an empiricist can quantify the 10 parameters needed for the two-species consumer resource model for two essential and non-substitutable resources (equation 8 and 9). One important note is that the elemental content of algae is not constant (Goldman et al. 1979, Finkel et al. 2010) and can vary under different nutrient limitation (Bertilsson et al. 2003) or light condition (Sterner et al. 1997, Urabe et al. 2002).

As shown by Letten et al [2017], these parameters described above can be used to predict coexistence under different resource supply ratios and dilution rates in a chemostat. However, the form the equation used to get ND and RFD depends upon knowing how the resource supply ratio relates to the R\* values for each species. Specifically, the equations depend upon whether or not the supply ratio is outside the ratio of and for the two species, i.e. resource supply conditions where the resource ratio theory predicts that they will not coexist.

* + 1. *Limitations*

Reorganizing Tilman’s consumer resource model requires an empiricist to know which resources are relevant for competition and species’ yield on resources, which could raise substantial efforts. Additionally, the resource supply must be constant and equal to the density-independent loss rate, which could be achieved by some experimental settings like chemostat experiment.

1. *Do the methods give the same prediction regarding coexistence? Do the methods yield the same values of ND and RFD?*

Although each of the five methods can be used to predict coexistence based on mutual invasibility, the experimental approaches to parameterizing those method are different, and it is not clear that these methods would yield the same predictions (or values of ND and RFD) if applied to a given study system. Here we use numerical simulations to investigate whether four of those methods, when implemented as shown in Figures 1-4, lead to the same prediction regarding coexistence and give the same estimates of ND and RFD. We chose to use numerical simulation for this demonstration since we are unaware of any experimental dataset that has been, or could be, analyzed using more than two of the methods. The numerical simulations were based on Tilman’s consumer-resource model [Tilman 1977] with two species of phytoplankton competing for two essential resources (phosphate and silicate), which were the only parameters varied among simulations (See Supplement). For each set of resource conditions, we performed numerical simulations that represent four methods: fitting the Lokta-Volterra model to monocultures and a co-culture, the Sensitivity method applied to an invasion experiment, the method using Tilmans’s CRM, and the NFD method.

Figure 6 shows that all four methods made the same prediction about coexistence and that these predictions match the outcome based on the equilibrium condition from simulation. Across the different resource conditions that we explored, the two species were predicted to coexist when the resource supply conditions caused each species to be limited by a different resource [Tilman 1977]. This agreement among the methods was conditional on how the Lotka Volterra and NFD methods were parameterized. The Lotka Volterra method only matched the predictions from the other methods when we assumed that intraspecific competition coefficients were equal to the inverse of the carrying capacity (Figure S2). When we estimated the intraspecific coefficients from the time series as the monocultures approached their carrying capacity, the method overestimated the range of parameter space allowing for coexistence. The NFD method only matched the other methods when we evaluated the slope when species densities were approaching zero and by using both the slope and the intercept to predict the growth rate when invading. Unless these conditions were met, the NFD method tended to over- or under-estimate the region of resource conditions that allow for coexistence.

Although the methods gave the same predictions regarding coexistence, Figure 6 (f and g) shows that the methods do not consistently yield the same values of ND and RFD when applied to the same study system. The Lotka Volterra method (using ) and the Sensitivity method gave identical estimates of ND and RFD across the range of resource conditions used, but these estimates differed from the method based on Tilman’s consumer resource model. This disparity is explained by the fact that the Lotka Volterra and Sensitivity methods assume that per capita inter- and intraspecific interaction coefficients are independent of species densities. Although this assumption is likely to be violated when species population dynamics are affected by mechanisms that produce non-linearity between population densities and growth rates, using the assumption that allows for accurate predictions regarding coexistence. In contrast, the method based on Tilman’s consumer resource model does not assume that interaction coefficients are independent of species densities, but instead quantifies both inter- and intraspecific interaction coefficients only at the steady-state density for monocultures. This means that the interaction coefficients, and thus ND and RFD, measured according to either of the phenomenological methods (Figures 2 and 3) are unlikely to match the values predicted from a mechanistic method.

This comparison of the methods highlights an important caution for the field of modern coexistence theory, namely that estimates of ND and RFD obtained by different methods are not always comparable. For example, future syntheses or meta-analyses cannot include studies that measured ND and RFD by different methods. Even within a single method (e.g. the Lotka Volterra method) there can be substantial differences in the estimates of ND and RFD depending on the experimental design and how the interaction coefficients were parameterized. Nonetheless, it appears that each of these methods is capable of correctly predicting coexistence based on the criterion of mutual invasibility.

**Part 2. An empiricist’s guide to when, why, and how each method should be used to predict coexistence and estimate ND and RFD**

Having described and compared the foundation of each empirical method, here in Part II of the paper we now to help empiricists determine 1) which method(s) are most appropriate for their study system, 2) how much experimental effort is required for a given method, and 3) what types of predictions the method can deliver. To aid our discussion, we have summarized the methods in Table 1, which is organized into three sections. The section ‘Decision Steps’ is a decision tree that allows an empiricist to identify the most appropriate method for their study system. The ‘Method’ section directs the empiricist to the key literature for each method. The section ‘Experimental Requirements’ outlines key aspects of the experiments that are required to use the method.

* 1. *Decision Steps - deciding which method to use*

The first section of Table 1 (Decision Steps) uses a sequence of questions about the study system that help an empiricist identify the most appropriate method for their work. The first question asks whether the method must yield estimates of ND and RFD that can be directly related back to Chesson’s inequality (eq. X). This distinction is important since the outputs from this method should not be compared to the other four methods. Thus, Question 1 isolates the negative frequency dependence method from all others. However, the negative frequency dependence method can accurately predict coexistence based on mutual invasibility and, depending on the answers to the remaining questions, it could be the most appropriate method for some study systems. In particular, the NFD method is the only option that does not require that the empiricist grow each species alone as a monoculture (Question 2). This could be an advantage for study systems where experimental manipulations are not feasible (e.g. long-lived species, protected habitats). Several empirical studies have implemented the NFD method in experiments [Levine and HilleRisLambers 2009; Chung and Rudgers 2016] and other similar studies have used NFD in observational studies [Adler et al 2010].

The next question in the decision sequence is whether the empiricist knows which specific resources the species are competing for and can quantify the dependence of each species’ population dynamics on those resources (Question 3). This question separates the four methods for estimating ND and RFD into two separate groups. The phenomenological methods (Lotka-Volterra and Sensitivity method) are those that are informed by quantifying species interactions, but which make no assumptions about the resources that species are competing for (highlighted in green). The mechanistic methods (MacArthur’s CRM and Tilman’s CRM) are based on consumer resource models in which species are assumed to interact only by competing for shared resources (highlighted in blue). To use either of the mechanistic methods, an empiricist needs to know which resources determine niche differences. In certain cases, it will not be possible for an empiricist to answer ‘yes’ to Question 3, because the resources required for species to grow are not known or cannot be readily quantified. When one cannot answer yes to Question 3, then the Lotka-Volterra and Sensitivity methods are appropriate because they can still quantify ND and RFD even if the empiricist does not have a good understanding of which resource(s) species are competing for, and thus, which resources define their niche axes. Because this decision step is so influential, the remaining steps are particular to either the phenomenological or mechanistic methods.

Deciding between the Lotka-Volterra method and the Sensitivity method (phenomenological methods, highlighted in green), depends on the answers to whether the method must work for observational datasets (Question 4), whether it is necessary to experimentally grow each species as at steady-state near its carrying capacity (Question 5), and whether the method can be used to predict coexistence among 3 or more species simultaneously (Question 7). An empiricist working with long-lived species or in protected habitats would likely answer ‘yes’ to Question 4, eliminating the Sensitivity method. In this case, the empiricist would need to decide whether it is essential to obtain values of ND and RFD compatible with the other four methods (requiring the Lotka Voltera Method) or whether the negative frequency method could be employed to predict mutual invasibility and thus coexistence. Similarly, if an empiricist is unable to answer ‘yes’ to Question 5, she/he would be forced to use either the Lotka Volterra Method or the negative frequency dependence method. The distinction in Question 5 could be particularly important for slowly-growing species where it is possible to parameterize the carrying capacity term from a time series of species densities, but it would take too long for the species to approach the carrying capacity to justify beginning an invasion by the other species. Lastly, the two phenomenological methods differ in terms of whether they can predict species coexistence among three or more species simultaneously (Question 7). While the Lotka-Volterra model can be parameterized to obtain all pairwise competition coefficients for a pool of species, it has not been applied to predicting coexistence of more than two species simultaneously. The Sensitivity method can be used beyond pairwise species interactions; however, this is limited to situations where all non-focal species can be considered in aggregate (e.g. species invading a community of ).

Deciding between the MacArthur and Tilman CRM methods (mechanistic methods, highlighted in blue), is straightforward and depends on whether the resources that the species compete for are abiotic and governed by a constant rate of supply (e.g. inorganic nutrients consumed by plants) or biotic with their own population dynamics (Question 6). It is worth noting that Tilman’s R\* concept has been extended to include competition mediated by predators (e.g. P\*, Tilman 1982). However, to date, models including predation have not been related to Chesson’s ND and RFD. Additionally, use of MacArthur’s consumer model can theoretically work for more than two species at a time, but this has not been demonstrated for the method based on Tilman’s consumer-resource model (Question 7).

Using this decision tree, an empiricist can determine which method(s) are appropriate for their study system. Due to the study system or experimental constraints, an empiricist may have multiple options for which method to use. In these cases, it can be useful to consider the experimental requirements of each method (Table 1, *Experimental Requirements*) and the tradeoffs among the methods in terms of their utility as discussed in Part 3.

* 1. *Experimental Requirements*

In addition to the decision steps outlined in Table 1, there are important practical differences for the experimental or observational studies required to quantify ND and RFD for each method. The most important difference in study design among these methods is whether or not they require species to be grown together in order to make a prediction about coexistence. The negative frequency dependence method and the two phenomenological methods require each pair of species to be grown together in at least one co-culture, but the mechanistic methods do not require the species to be grown together as co-cultures. This distinction means that only the mechanistic methods can yield predictions about coexistence of species without the need for pairwise competition experiments. For example, consider a typical competition experiment involving a pool of three species (A, B, and C). The mechanistic methods can make predictions about all pairwise combinations of the species (A+B, A+C, and B+C) based solely on information about each species when grown individually. The phenomenological methods, however, require at least one co-culture for each pairwise combination of species, which means that information from pairs A+B and A+C cannot be used to make a prediction about coexistence for the pair B+C. The need for species to be grown together in co-culture has important implications for the total number of experimental treatments that would be required to quantify ND and RFD.

Experiments using the phenomenological methods can require more experimental treatments to predict pairwise coexistence among a pool of species than the mechanistic methods do. For the phenomenological methods, the number of experimental treatments required for all pairwise combinations of species increases exponentially with each additional species being considered. In contrast, for the mechanistic methods the total number of experimental treatments required increases linearly with the number of species being considered. This is because the methods based on consumer-resource models do not require any direct competition experiments in order to estimate competition coefficients ( and/or ), while all of the phenomenological methods require at least one co-culture of each species pair in order to quantify the competition coefficients. As a result, the relative efficiency of the phenomenological versus mechanistic methods depends upon both the number of species being considered and also the number of resources. When the number of species being considered is small and the number of limiting resources is few, the difference in experimental effort can be modest. For example, to predict pairwise coexistence among a pool of four species, using the Sensitivity method requires 16 experimental treatments (time series): 4 monocultures to parameterize both maximum growth rate and carrying capacity and 12 invasions to parameterize sensitivity (A invading B, B invading A, etc.). In contrast, using either of the consumer resource models (two limiting resources) would require two experiments per species for a total of 8 experiments. If the mechanistic methods require parameterizing four or more limiting resources, then the phenomenological methods may be more efficient for a pool of four species. However, for larger pools of species the difference can be substantial. Obtaining pairwise estimates of ND and RFD for a pool of 10 species requires between 55 and 180 treatments for the phenomenological methods but as few as 20 treatments for mechanistic methods.

In addition to the number of experimental treatments required for each method, it is important to consider the amount of effort and time required for each treatment. Specifically, the NFD, Lotka Volterra, and Sensitivity method require time series of species densities in the experimental or observation study. In the case of the NFD and Sensitivity method, these time series may be short in duration (i.e. at least one generation) and focused only on population dynamics when species densities are very low or near the steady-state density of monocultures. However, the Lotka Volterra method requires longer time series in order to parameterize both the interaction coefficients and carrying capacities. Longer time series in monoculture and co-culture are more easily attainable for quickly-growing species like microbes and invertebrates, but even short time series could be prohibitively arduous for slowly growing species like trees.

Ultimately, the total effort and resources required for a study is jointly determined by the method, number of species, number of limiting resources (if applicable), length of time series, level of replication, and any other design elements. Using Table 1 as a guide, an empiricist should be able to select a method and begin to design a study that satisfies their aims.

**Part 3.**

In Part 1 of this paper, we explained how each of five models/methods can be used to quantify niche differences (ND) and relative fitness differences (RFD), which can then be used to predict coexistence among pairs of species. We further showed that while the methods are not mathematically identical, nor do they always give comparable values of ND and RFD, they all qualitative predict the outcome of species interactions, and correctly predict whether or not species coexist. Then, in Part 2 of the paper we offered a decision-making framework that can be used by empiricists to select the most appropriate method for their particular study system, specific aims, and experimental capabilities. Now, in Part 3, we end the paper by offering some advice for empiricists about how to navigate tradeoffs among the methods, how to compare and synthesize measurements of ND and RFD from different methods, and lastly, key future directions for implementing modern coexistence theory empirically.

* 1. *Tradeoffs Between Phenomenological and Mechanistic Methods*

Given the substantial differences in experimental design requirements and effort that is needed to execute the five methods described in Part 1 of the paper, it is highly likely that empiricists will face choices that require tradeoffs in selection of a particular model or method for their study system. The most obvious and important tradeoffs occur among the phenomenological methods and the mechanistic methods, which differ in three important ways: First, the phenomenological methods (i.e. the NFD, Lotka-Volterra model and the Sensitivity method) make no assumptions about the resources that species compete for. This could be beneficial for empiricists who can still measure ND and RFD even if they lack detailed information about the biological resources that species compete for. But the trade-off for this lack of knowledge is the experiments that need to be performed to quantify ND and RFD are more elaborate and cumbersome. Indeed, a key disadvantage of all three phenomenological methods is that they require each pair of species to be grown together in competition, which causes the total effort to increase exponentially as more species are considered. In addition, the results of phenomenological experiments are specific to each pair of species tested and cannot be generalized to interactions beyond that pair. They are also specific to the particular environmental conditions, like resource density or resource supply rates, used in that experiment and cannot be generalized.

But for those who can identify the resources that species compete for, use of the mechanistic models allows for simpler, less complex experiments that are more easily generalized to predict coexistence among all species in the focal species pool. Indeed, an empiricist who is able to answer ‘yes’ to Question 3 in Table 1 could use a mechanistic model to predict coexistence (or not) not just of a single species pair of interest, but of any and all species pairs of interest based solely on experiments that take measurements from each species grown alone in monoculture. Importantly, the mechanistic experiments also offer the ability to make predictions about species coexistence under different environmental conditions. For example, Letten et al. showed that the Tilman consumer resource model can be used to predict the ND and RFD at different nutrient supply concentration or dilution rates (Letten et al. 2017). The ability of the mechanistic methods to handle some changes to environmental context, while limited, could be useful for predicting how anthropogenic stressors (e.g. nutrient pollution) are likely to affect species coexistence.

The ability to make predictions about combinations of species without the need to perform all pairwise competition experiments has already been touted as a benefit of the mechanistic models [Tilman 1982], and it could be useful for addressing certain ecological questions that don’t always lend themselves well to manipulative experiments (e.g. invasions by introduced species, coexistence of rare or endangered species). The key point here is that there is a trade-off between the simplicity and generality of results, and how much one knows about the focal species and what they compete for.

* 1. *Comparing and synthesizing measurements of ND and RFD*

To date, only three of the five methods described for measuring niche and relative fitness differences have been used empirically. No one, to our knowledge has used the MacArthur Resource Model or Tilman’s Consumer Resource Model to quantify ND and RFD in any real system, despite publications showing these can be used to do so. That means that most of our inferences about ND and RFD that have been measured empirically stem from the phenomenological models of coexistence. Furthermore, we are unaware of any empirical study that has applied more than one of model/method to the same empirical study system. As such, we have no way to compare the performance of the methods/models empirically. Therefore, we believe an important avenue for future research is to focus on measuring ND and RFD using more mechanistic models, and for studies that measure ND and RFD using different methods in the same study system so that we can compare results and attempt to demonstrate equivalence of these methods.

Even as we call for more mechanistic experiments and comparative studies, we caution against the urge that will eventually rise to synthesize measures of ND and RFD in an informal data synthesis or more form meta-analysis. Although we have shown that all five existing methods should correctly predict the qualitative outcome of coexistence, the methods are by no means mathematically equivalent. There is no reason to expect, a prior, that the quantitative values of ND (or RFD) measured for a particular group of organisms for one method will produce quantitatively similar values of ND (or RFD) for that same group of organisms using a different method. As such, the methods are not directly comparable, and the measurements they produce should not be mixed-and-matched to produce some synthesized estimate of the niche or fitness difference for, say, grassland plants.

* 1. *Future directions for implementing modern coexistence theory*

In our view, there are at least three important new directions that work on species coexistence must go if Chesson’s modern theory of coexistence is to become more general and practical. First, it needs to move beyond prediction of pair-wise species interactions. Several authors have recently emphasized that modern coexistence theory is under-developed for multi-species systems (Carroll et al. 2011, Levine et al. 2017). In theory, competitive hierarchy between species versus and versus might not directly translate to species and , when these species are engaged in intransitive competition or higher-order interaction (Levine et al. 2017). In fact, none of the three phenomenological methods (the NFD, Lotka-Volterra model and the Sensitivity method) can deal with intransitive competition or higher-order interactions. Importantly, the emphasis to date on pairwise interactions and experimentation means that intransitive competitive interactions and higher-order interactions, if present, are unaccounted for. Chesson’s coexistence framework is an informative synthesis so far, and how to expand this framework to multi-species system is a direction worth pursuing.

Second, to also include fluctuation dependent mechanisms of coexistence, we need to expand the consideration of the five methods reviewed here or even develop new empirical methods. To our knowledge, there is limited empirical studies that explicitly quantify the fluctuation dependent mechanisms, i.e. relative nonlinearities and storage effects (but see Angert et al. 2009, Letten et al. 2018). Even so, it is well-known that environmental fluctuations mediate species coexistence in some empirical systems where environmental fluctuations cannot be negligible (e.g. Cáceres 1997, Jiang and Morin 2007) and any modern theory of coexistence is incomplete without them.

Lastly, we think a fruitful avenue for future work is to focus on the application of modern coexistence theory to key environmental problems. By applying the modern coexistence theory to study how biodiversity affects ecosystem functions, a theoretical study shows that niche difference can increase the biomass of a community (Carroll et al. 2011). However, rarely have empirical studies quantified the effects of niche difference on community biomass. Besides community biomass, it is also not clear how other ecosystem functions or services, like biotic or abiotic resource exploitation efficiency or food and material productions, are influenced by both niche and fitness differences under the modern coexistence framework. In the light of modern coexistence theory, not only can we understand how species coexist to form a community, but we should also be able to understand how species affect the functions and services the we human rely on.

**Appendix A**

In this appendix, we show that for the NFD slope to be constant, the community density must be not only saturated but also fixed. To do so, we attempt to derive the NFD slope and intercept from a two species Lotka-Volterra competition model (), where *Ni* is scaled on its carrying capacity but does not mean two species have the same carrying capacity. In a Lotka-Volterra model, there is no variable representing species’ frequency, so that we have to assume a fixed community density, *B*. This assumption also satisfies the assumption of the NFD method that the community density is always saturated. Fixing the community density makes the density dependency, *αij*, equivalent to frequency dependency (Adler et al. 2007), and allows species’ frequency to be calculated by *Ni*/*B*. The two-species Lotka-Volterra competition model can then be rewritten as follows

(A1)

, where *B* is the fixed community density and one unit decrease of *Ni* will lead to one unit increase of *Nj*. Note that *B* is an arbitrarily defined constant describing the equilibrium community density at any given density of *Ni*. From equation A1, we derive the NFD slope and intercept in the following equations.

(A2)

(A3)

From equation A2, after we fix the community density at *B*, the NFD slope becomes constant, which means the species’ per capita growth rate linearly depends on its frequency. In addition, from equation A2 and A3, we argue that both NFD intercept and slope should be used with caution in evaluating Chesson’s inequality. First, the NFD intercept represents whether species can invade the carrying capacity of its competitor from rare, so it can be used to accurately assess mutual invasibility. However, neither do the difference nor the ratio of two species’ NFD intercept (equation A2) yield the same calculation of relative fitness difference by Chesson (). NFD slope (equation A3) is also not equivalent to the equation calculating niche difference (), although the slope of NFD has been used to represent ND for annual plant communities (Yenni et al. 2012b, 2017b). Thus, while the negative frequency dependency method can correctly predict mutual invasibility, the NFD intercept and slope should not be interpreted as RFD and ND in order to evaluate Chesson’s inequality.

**Appendix B**

In this appendix, we show that sensitivity measure actually describes the impact of the entire competitor population level, not of the *per capita* level, on the per capita growth rate of focal species. We derive the sensitivity metric (*Si*) from the Lotka-Volterra competition model (equation 3). The *μi* in equation 3 is the maximum growth rate in monoculture, so it is *ri* in equation 2. The *μij* is the invasion growth rate, so that we can replace with species *j*’s carrying capacity, , and replace with 0, so that . Therefore,

(B1)

From equation B1, we see that sensitivity (*Si*) is the equilibrium density of species *j* () times the *per capita* competition coefficient (*αij*). The sensitivity thus is a measurement of the overall population effect of species *j* on focal species i but not the “*per capita*” effect of species *j*. This correction allows for measured sensitivity (*Si*) to be converted to inter-specific competition coefficients (*αij*). In addition, species’ density at the equilibrium (*N\**) is actually , so . The niche difference (ND) and relative fitness difference (RFD) can be calculated from the sensitivity metric (*Si*) because  and RFD of species *j* over species *i* is .

With the above derivation, we can derive Chesson’s inequality for coexistence (equation 1) using the sensitivity metric. From the Lotka-Volterra model, intra-specific competition (*αii*) must be greater than the inter-specific competition (*αij*) to guarantee stable coexistence. Therefore, we can have the following deduction.

, so (B2)

, so (B3)

Combining equation B2 and B3, we have an inequality for coexistence expressed with sensitivity metrics, , which is in the same form as in Chesson’s coexistence framework. As a results, we show that the sensitivity measurement (*Si*) is not directly equal to the per capita competition coefficient (*αij* in the Lotka-Volterra model), but due to its mathematic attributes, it can be used to calculate ND and RFD and correctly predict coexistence. However, as

**Table**



Table 1. The table showing the equivalent of a decision tree that helps empiricists determining when, why, and how each method should be used (Section “Decision steps”). Experimental requirements and outputs of each empirical experiment are also listed.

**Figures**

Figure 1

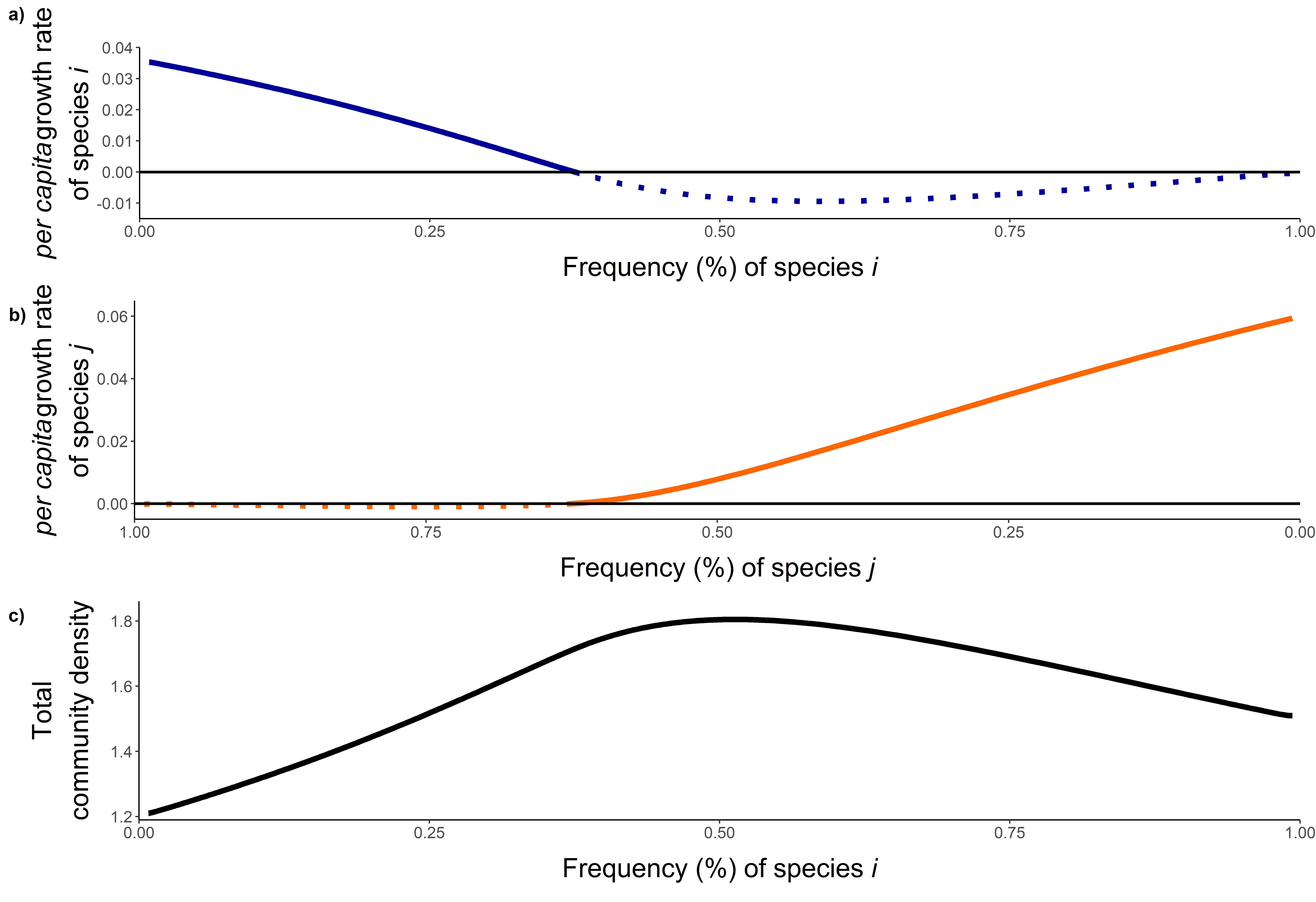


Figure 1. Simulation results of species’ per capita growth rate versus its frequency (panel a, b) from two-species Lotka-Volterra model ( and how the community density change with species’ frequency (panel c). The numerical simulations are performed with arbitrarily defined parameter values that allows species to coexist (, , , , , , ). The simulations are to mimic invasion experiment so that we set species *i* at its equilibrium density with species *j* invading from rare and set species *j* at its equilibrium density with species *i* invading from rare. From these two simulations we calculate pairs of each species’ frequency from the time series of population density. We also plugged in species’ density to the model to calculate the corresponding per capita growth rate at a given species’ frequency. In panel a and b, the species’ frequency and their corresponding per capita growth rates were used to construct frequency dependency plots. The solid parts of the line indicate where species’ density is below the bi-culture equilibrium, so that per capita growth rate is positive; otherwise species’ per capita growth rate is negative (dotted parts). In panel c, we show that community density is not fixed across species *i*’s frequency. From these simulations we see that either non-fixed community density will lead to non-liner frequency dependency.

Figure 2.

A close up of a map

Description automatically generated

Figure 2. An example plot showing the estimation of αij from fitting Lotka-Volterra model to time series. The points represent the density of species i (open circles) and j (solid dots) respectively. The dashed lines are the two fitted growth curves. The species densities were generated by a Lotka- Volterra model and added some random noise.

Figure 3



Figure 3. An example plot showing the estimation of sensitivity (*Sij*) of species *i* invading species *j* (panel a.) or the reverse (panel b.). In both panels, solid dots are the growth curve species when growing alone and the open circles are the growth curve of species when invading the carrying capacity of the other species. To estimate the sensitivity of the focal species (e.g. species *i*), the competing species (species *j*) is fixed at the equilibrium and invade the focus species to estimate the invading growth rate of the focal species.

Figure 4



Figure 4. An example plot demonstrating the idea of niche difference (ND; *ρ*) in MacArthur’s con- sumer resource model. Consumption of species *j* on resource *l* (*cjl*) are plotted against consumption of species *i* on resource *l* (*cil*). The closeness between *cjl* and *cjl* are the closeness between these points to the 1:1 line and is expressed as *ρ*. For example, if all *cjl* are equal to *cil*, which means all points are on the 1:1 line, *ρ* is 1 and ND between species *i* and species *j* is 1−*ρ* = 0. The inset plot is an example plot showing the data required to measure consumption of species *j* on resource *l* (*cjl*).

Figure 5



Figure 5. Example plots showing the empirical data required to estimate *R\** in Tilman’s consumer resource model for species *i* and *j* on resource *i* and *j*. To obtain data one would need to grow the focal species (*i* and *j*) under different resource level and measure the corresponding *per capita* growth rate, which is indicated by the blue and yellow curves in each panel. *ri* (*ri*) is the per capita maximum growth rate of species *i* (*j*), *kij* is the half saturation constant for species *i* limited by resource *j*. means the minimum level of resource *j* that allows species *i* to have positive per capita growth rate equals to the dilution rate of the experiments. The *R\**s can be used to calculate *αii* and *αij* (*αjj* and *αji*) because they determine how sensitivity species *i* will be affected by the resource (resource *j*) that limits the growth rate of species *i*.

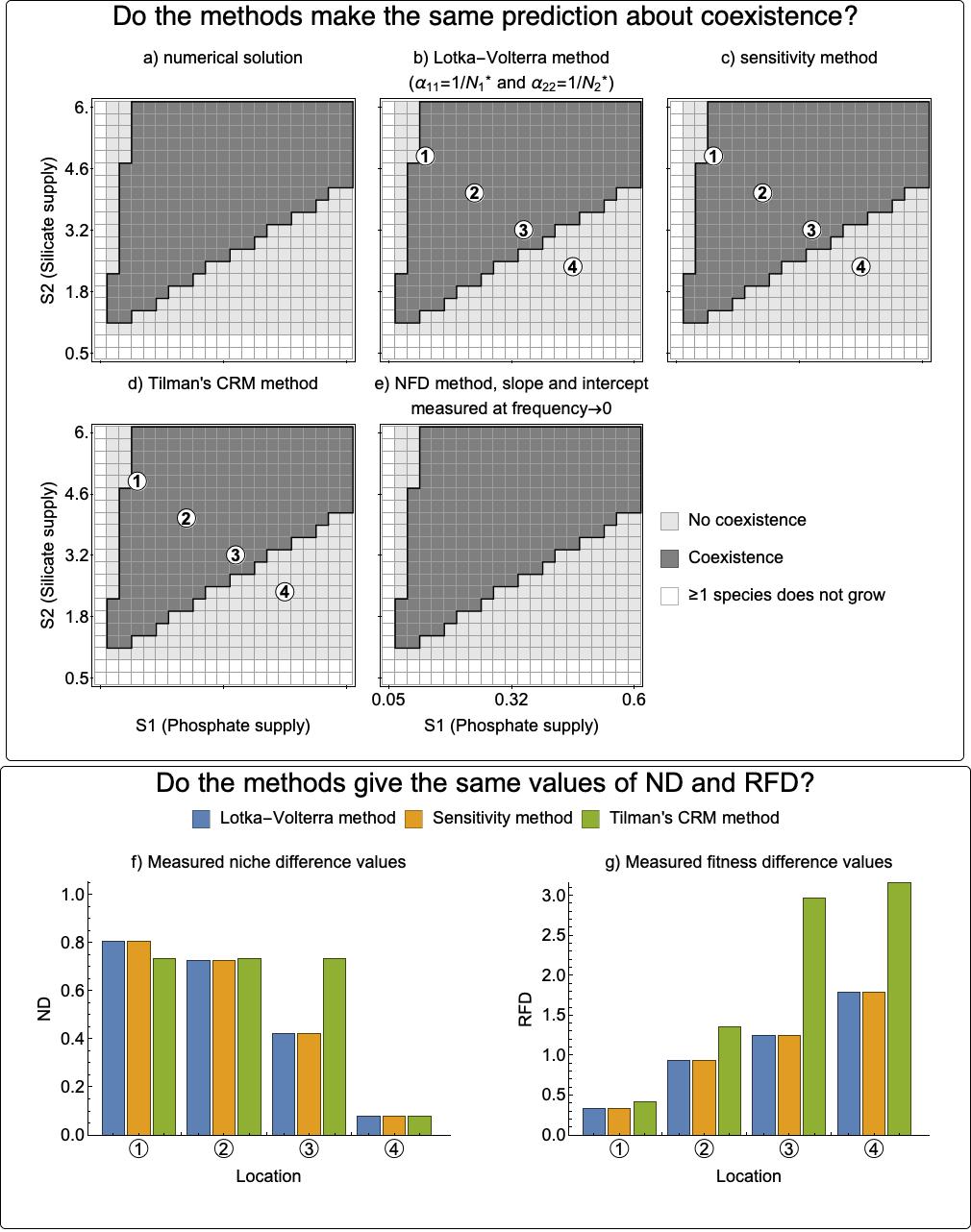


Figure 6. Predictions from the methods for coexistence under different resource supply conditions (a-e). In each panel, the predictions for coexistence are compared against the coexistence outcome based on numerical equilibrium. Three of the methods give the correct predictions regarding coexistence across this region (Lotka-Volterra, Sensitivity method, Tilman’s CRM) and the NFD method can give the correct prediction depending upon how it is parameterized and used. The methods did not give the same values for niche differences and relative fitness differences (f and g). The labeled locations in panels f and g correspond to marked locations in panels b-d and show that the disagreement among the methods is smaller toward the center of the parameter space that allows for coexistence. The raw RFD values from the Sensitivity method were converted to the same ordering as used in the other methods (species in the denominator rather than the species with the greater sensitivity). Because the NFD method cannot be used to produce values of ND and RFD that are comparable with the other four methods, only the predictions regarding coexistence are plotted.

**Supplement 1**

In this supplement, we perform numerical simulation to show that that per capita growth rate does not linearly depend on species’ frequency, i.e. the NFD slope is not constant, under the assumptions that the community biomass is saturated and that the two species are capable of mutual invasion. We then show that depending on the frequency at which the NFD slope is evaluated, predictions of species coexistence is different and can be misleading. To show again that the NFD slope can be not constant, we performed numerical simulation using Tilman’s consumer-resource model with two species of phytoplankton. The equations and parameter values are taken exactly from Tilman 1977. Annotated code for the simulations is provided in a supplemental file. Simulations were performed using the function NDSolve in Mathematica 11.2 (Wolfram Research), employing a variable step size. For each set of resource supply concentrations, we performed four simulations: (1) species 1 as a monoculture growing from rare to near its equilibrium density; (2) species 2 as a monoculture growing from rare to near its equilibrium density; (3) species 1 at its equilibrium density with species 2 invading from rare; (4) and species 2 at its equilibrium density with species 1 invading from rare. From simulations 3 and 4, we used the time series to obtain pairs of each species frequency and its growth rate. These values were used to construct the relationships depicted in panel a and b. From panel a and b of Figure S1, we see that the frequency dependency is nonlinear for both species, i.e. the NFD slope is non-constant. The non-constant slope can lead to incorrect prediction of species coexistence. For example, under the supply ratio of panel a and b of Figure S1, two species can coexist. However, when evaluating the NFD slope at the frequency in the red regions, species are not predicted to coexist.

We subsequently show that when the NFD slope is not constant, using NFD slope can be problematic in practice. To further show that non-constant NFD slope is problematic in predicting species coexistence, we again perform simulations using Tilman’s consumer-resource model with two species of phytoplankton. We the show whether different methods, including the Sensitivity method, Letten’s method, fitting Lokta-Volterra method and the NFD method, give the same prediction under a limited range of resource supply concentrations. We perform an additional numerical simulation (“competition experiment”, panel a of Figure S2) to determine whether the species can coexist when both are introduced at low densities and allowed to reach equilibrium biomass. For each simulation, the parameter values were taken directly from Tilman 1977 and only the supply concentration of phosphate and silicate were varied. For each set of resource conditions, we performed four different numerical simulations: (1) species 1 as a monoculture growing from rare to near its equilibrium density; (2) species 2 as a monoculture growing from rare to near its equilibrium density; (3) species 1 at its equilibrium density with species 2 invading from rare; (4) and species 2 at its equilibrium density with species 1 invading from rare. Note that for the NFD method (panel e-j of Figure S2), because the NFD slope depends on species’ frequency, we used the NFD slope evaluated at three frequencies, including near 0% (panel e and h of Figure S2), 50% (panel f and i of Figure S2) and near 100% (panel g and j of Figure S2), and each complimented with (panel e-g of Figure S2) or without (panel h-j of Figure S2) extrapolated vertical intercept that represent the expected growth rate when at rare. For the NFD methods, we see that only when using the NFD slope evaluated at near 0% frequency and complimented with extrapolated vertical intercept, the predictions of species coexistence based on NFD method match the those of the other methods. While certain intermediate frequencies of the two species can be used to make accurate predictions (Figure S2), an empiricist would not know these frequencies without performing the competition experiments or examining frequency dependence across the entire range of frequencies.

Figure S1

A close up of a map

Description generated with very high confidence

Figure S1. Figures showing the simulation results of per capita growth rate versus frequency of species 1 and 2 (panel a and b) and the growth rate dependency across different resource ratio (panel c) At any frequency of the two species, the NFD method requires that we use the slope to extrapolate and estimate the growth rate when a frequency approaching zero (the extrapolated vertical intercept). For frequencies where this method predicts mutual invasibility for both species, i.e. species can coexist, the lines are blue. For frequencies of the two species where the method leads to the incorrect prediction, the lines are red. Both species have positive growth rates when their frequency approaches zero, indicating that they are mutually invasible. The coexistence equilibrium occurs at a frequency of 0.88 for species 1 and 0.12 for species 2. At point A, corresponding to frequency of 0.5 for each species, this method does not predict coexistence because while species 1 would be expected to have a positive growth rate when rare (also has negative NFD slope), species 2 is predicted to have a negative growth rate when rare (positive slope of NFD). If an empiricist did their work between species 1 frequency of ~0.05 to ~0.85, and used the slope of NFD, they would incorrectly predict that the species will not coexist. At point B, which is close to their equilibrium frequencies, both species have negative NFD slope and should have positive growth rates when rare based on this method. Panel c shows the fidelity of prediction from the NFD method as a function of Si:P ratio and the frequency of the two species. The vertical dashed line represents the slice depicted in panels a and b. For all of the Si:P ratios shown in panel c, the species are mutually invasible and will coexist. This plot indicates that using NFD will often predict that the species will not coexist, when in fact they do coexist. This is important because without doing the full range of species frequencies in an experiment, one would not know whether and where the relationship between frequency and growth rate is non-linear.

Figure S2.

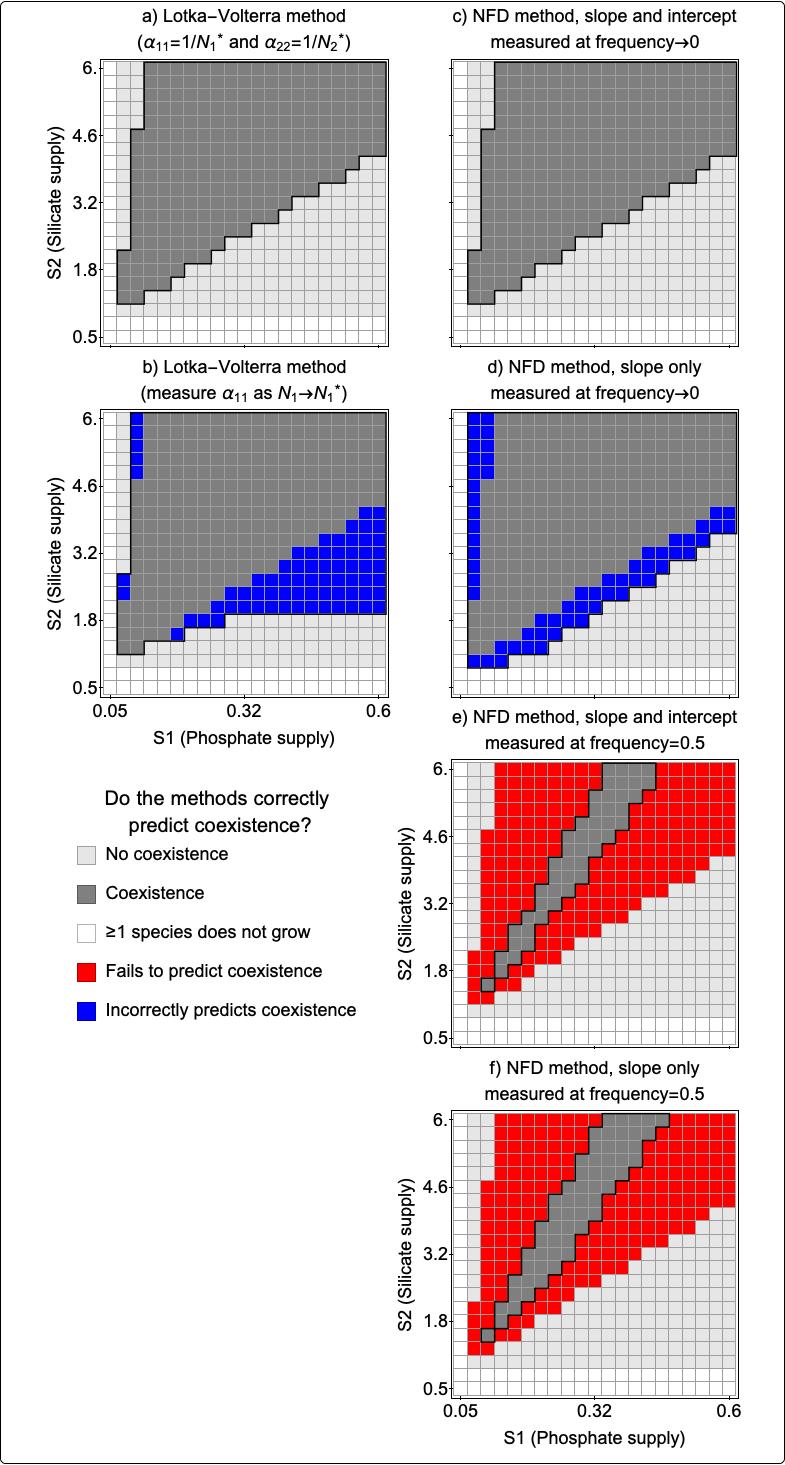


Figure S2. Simulation results showing the parameter space in which two species are predicted to stably coexist using different methods. In each panel, the outcome is compared against the equilibrium from numerical simulation.

**Reference**

Abrams, P. 1980. Are Competition Coefficients Constant? Inductive Versus Deductive Approaches. The American Naturalist 116:730–735.

Adler, P. B., J. HilleRislambers, and J. M. Levine. 2007. A niche for neutrality. Ecology Letters 10:95–104.

Angert, A. L., T. E. Huxman, P. Chesson, and D. L. Venable. 2009. Functional tradeoffs determine species coexistence via the storage effect. Proceedings of the National Academy of Sciences 106:11641 LP-11645.

Bertilsson, S., O. Berglund, D. M. Karl, and S. W. Chisholm. 2003. Elemental composition of marine Prochlorococcus and Synechococcus: Implications for the ecological stoichiometry of the sea. LIMNOLOGY AND OCEANOGRAPHY 48:1721–1731.

Cáceres, C. E. 1997. Temporal variation, dormancy, and coexistence: A field test of the storage effect. Proceedings of the National Academy of Sciences 94:9171 LP-9175.

Carroll, I. T., B. J. Cardinale, and R. M. Nisbet. 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. 1994. Multispecies Competition in Variable Environments. Theoretical Population Biology 45:227–276.

Chesson, P. 2000. Mechanisms of maintenance of species diversity. Annual Review of Ecology and Systematics 31:343–366.

Chesson, P. 2003. Quantifying and testing coexistence mechanisms arising from recruitment fluctuations. Theoretical Population Biology 64:345–357.

Finkel, Z. V., J. Beardall, K. J. Flynn, A. Quigg, T. A. V Rees, and J. A. Raven. 2010. Phytoplankton in a changing world: Cell size and elemental stoichiometry. Journal of Plankton Research 32:119–137.

Goldman, J. C., J. J. McCarthy, and D. G. Peavey. 1979. Growth rate influence on the chemical composition of phytoplankton in oceanic waters. Nature 279:210–215.

Hillerislambers, J., P. B. Adler, W. S. Harpole, J. M. Levine, and M. M. Mayfield. 2012. Rethinking Community Assembly through the Lens of Coexistence Theory. Annu. Rev. Ecol. Evol. Syst 43:227–48.

Jiang, L., and P. J. Morin. 2007. Temperature fluctuation facilitates coexistence of competing species in experimental microbial communities. Journal of Animal Ecology 76:660–668.

Letten, A. D., M. K. Dhami, P.-J. Ke, and T. Fukami. 2018. Species coexistence through simultaneous fluctuation-dependent mechanisms. Proceedings of the National Academy of Sciences 115:6745 LP-6750.

Letten, A. D., P. J. Ke, and T. Fukami. 2017. Linking modern coexistence theory and contemporary niche theory. Ecological Monographs 87:161–177.

Levine, J. M., J. Bascompte, P. B. Adler, and S. Allesina. 2017. Beyond pairwise mechanisms of species coexistence in complex communities. Nature 546:56.

Levine, J. M., and J. HilleRisLambers. 2009. The importance of niches for the maintenance of species diversity. Nature 461:254–7.

Miller, E. T., and C. A. Klausmeier. 2017. Evolutionary stability of coexistence due to the storage effect in a two-season model. Theoretical Ecology 10:91–103.

Narwani, A., M. A. Alexandrou, T. H. Oakley, I. T. Carroll, and B. J. Cardinale. 2013. Experimental evidence that evolutionary relatedness does not affect the ecological mechanisms of coexistence in freshwater green algae. Ecology Letters 16:1373–1381.

Smith-Gill, S. J., and D. E. Gill. 1978. Curvilinearities in the Competition Equations: An Experiment with Ranid Tadpoles. The American Naturalist 112:557–570.

Sterner, R. W., J. J. Elser, E. J. Fee, S. J. Guildford, and T. H. Chrzanowski. 1997. The Light: Nutrient Ratio in Lakes: The Balance of Energy and Materials Affects Ecosystem Structure and Process. The American Naturalist 150:663–684.

Tilman, D. 1977. Resource competition between plankton algae: An experimental and theoretical approach. EcologyEcology 58:338–348.

Tilman, D. 1980. Resources: A Graphical-Mechanistic Approach to Competition and Predation. The American Naturalist 116:362–393.

Tilman, D. 1981. Tests of Resource Competition Theory Using Four Species of Lake Michigan Algae. Ecology 62:802–815.

Turnbull, L. A., J. M. Levine, M. Loreau, and A. Hector. 2013. Coexistence, niches and biodiversity effects on ecosystem functioning. Ecology Letters 16:116–127.

Urabe, J., M. Kyle, W. Makino, T. Yoshida, T. Andersen, and J. J. Elser. 2002. Reduced light increases herbivore production due to stoichiometric effects of light/nutrient balance. ECOLOGY 83:619–627.

Yenni, G., P. B. Adler, and S. K. M. Ernest. 2012a. Strong self-limitation promotes the persistence of rare species. Ecology 93:456–461.

Yenni, G., P. B. Adler, and S. K. M. Ernest. 2012b. Strong self-limitation promotes the persistence of rare species. Ecology 93:456–461.

Yenni, G., P. B. Adler, and S. K. M. Ernest. 2017a. Do persistent rare species experience stronger negative frequency dependence than common species? Global Ecology and Biogeography 26:513–523.

Yenni, G., P. B. Adler, and S. K. M. Ernest. 2017b. Do persistent rare species experience stronger negative frequency dependence than common species? Global Ecology and Biogeography 26:513–523.