**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, (2) mutual invasibility, (3) parameterization of Lotka-Volterra competition models, (4) MacArthur’s consumer resource model, and (5) Tilman’s consumer resource model. 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 cannot be directly compared. 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 argues 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 argues 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 is developing his neutral theory, Chesson (Chesson 2000) is 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 is 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 argues 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 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. As a result, there is potential for different methods yield different values of ND and RFD, as well as divergent predictions. Moreover, if values of ND and RFD quantified using different methods were compared to each other, such as what could happen if some researcher in the future performed a data synthesis, such comparisons could lead to incorrect conclusions.

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 compare these method in terms of their algebraic equivalence and ability to correctly predict species coexistence. In Part 2 of the paper we summarize when, why, and how each method should be used, and provide a decision-tree that will help empiricists select the most appropriate method for their study system. In part 3, we provide a few warnings for empiricists that will help prevent future confusion and miscalculations, as well as some suggestions that will help maximize the success of this exciting new field of coexistence theory.

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

1. *Theoretical background for five empirical methods*

The following five empirical methods are proposed to measure fluctuation independent mechanisms, i.e. niche difference (ND) and relative fitness difference (RFD). That being said, these methods are critically premised on (1) negligible environmental fluctuations so that (2) linear approximation of species interactions is accurate enough to assess mutual invasibility criterion (Barabás et al. 2018). Based these core assumptions, Chesson by linearly approximate the MacArthur’s consumer resource model with the classic Lotka-Volterra model (Chesson 1990) to express the mutual invasibility by the following inequality.

(1)

In Chesson’s inequality, *ρ* is the niche difference (ND) and is the relative fitness difference (RFD).

Following the same logic, we first describe the underlying rationale of each method by relating the calculations of ND and RFD to the Lotka-Volterra model that underlies Chesson’s inequality (equation 1). We then explain the experiment required for each method and end each method section with the limitations of each method. In the end of this part, we conclude by comparing these methods in terms of whether these methods yield qualitatively the same prediction for coexistence based on the mutual invasibility criteria.

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

The negative frequency dependence method d 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 dependency of species’ growth rate on its relative frequency is then used to assess the mutual invasibility criterion.

The NFD method relies on a key assumption that the community must be saturated with respect to density or biomass (Adler et al. 2007). When the community is assumed to be saturated, all resources or niches are occupied by either the focal species *i* or its competitor *j*. Under this assumption, increasing the frequency of individuals for focal species *i*, means individuals of species *i* will compete with individuals of their kind for resources, and 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. More negative NFD slope represents larger degree of stabilization, i.e. niche difference (ND; Adler et al. 2007). In addition, Adler et al. 2007 argue that when a species’ frequency approaches zero, difference between species’ growth rate and zero, i.e. the NFD intercept, can be used to quantify the relative fitness difference (RFD; Figure 1 in Adler et al. 2007). The NFD method thus does not require estimating any specific species interaction coefficients to predict coexistence. As long as the relationship between a species’ frequency and its growth rate is linear, knowing the slope of that relationship and the growth rate at any intermediate frequency could allow an empiricist to extrapolate and predict the intercept of the relationship and determine whether both species are mutually invasible.

* + 1. *Empirical approaches*

The NFD method is attractive for empirical work in that it can be used to assess whether species can coexist without assuming any underlying population growth model or mechanism of species interactions. Frequency dependency of a species can be quantified by experimentally manipulating the relative frequency of species in a community, and then plotting the per capita growth rate versus the corresponding relative frequency. Several empirical studies have used the NFD method to assess whether species can coexist by assuming a linear frequency dependence (e.g. Levine and HilleRisLambers 2009, Yenni et al. 2012, 2017). For example, one experimental approach that is particular to annual plants is to seed a plot with varying densities of the two species and then use seed production and survival as a surrogate for per capita growth rates between years. Frequency dependency could, in principle, also be constructed using observational data 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.

* + 1. *Limitations*

Despite some of the desirable aspects of the NFD method, there is a major caveat of this method. 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). We also show that the non-linear frequency dependence can lead to incorrect predictions about species coexistence (Supplement 1), although the slope of NFD has been used to represent ND for annual plant communities (Yenni et al. 2012, 2017). When the NFD slope is not constant, the NFD method can only have correct predictions only 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).

* 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 with empirical data. 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.

To meet the mutual invasibility criteria for any two species (e.g. *i* and *j*), both intra-specific competition coefficients must be greater than the inter-specific competition coefficients, i.e. *αii* > *αij* and *αjj* > *αji*. After obtaining the intra- and inter-specific competition coefficients, ND and RFD can be calculated following Chesson’s definition, i.e. and , and access Chesson’s coexistence inequality.

* + 1. *Empirical approaches*

To use the Lotka-Volterra model to empirically predict coexistence for species *i* and *j*, one must first estimate six different parameters that are used in the Lotka-Volterra model: intrinsic per capita growth rate of each species (*ri* and *rj*), per capita intra-specific competition coefficients (*αii* and *αjj*), and per capita inter-specific competition coefficients (*αij* and *αji*). In principle, this could be accomplished using maximum likelihood method from a single co-culture time-series dataset, where both species are introduced at low density and allowed to grow to steady-state. However, in practice, it is difficult to parameterize all eight variables from a single time-series. So the most likely 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, and one time-series representing a co-culture of the two species (Fig. 2).

* + 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 and relative frequencies. 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. Although this assumption has been shown to be violated in practice (Schoener 1974, Abrams 1980), the competition coefficients can still be used to accurately predict coexistence so long as these parameter values are measured close to conditions for mutual invasibility.

* 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, a species’ sensitivity (*Si*) to competition is defined as the amount by which its per capita growth rate is reduced when invading a steady-state population of a competitor (*μij*) relative to the growth rate achieved during exponential growth in monoculture (*μi*):

(3)

A species’ sensitivity is jointly determined by the niche difference (ND) and the relative fitness difference (RFD) (Chesson 2000, Adler et al. 2007). Specifically, greater ND between the two species reduces the sensitivity of both species’ to competition (Adler et al. 2007), 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 (Narwani et al. 2013). 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 theoretically valid. We can also derive an inequality for coexistence expressed with sensitivity metrics, , which is in the same form as in Chesson’s coexistence framework (appendix B). In conclusion, the sensitivity measurement (*Si*) is not directly equal to the 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.

* + 1. *Empirical approaches*

The sensitivity measurement is developed specifically for mutual invasibility experiments using small organisms that are relatively easy to manipulate, like green algae (Narwani et al. 2013). To do the mutual invasibility experiments, one would need to grow each species to its carrying capacity on its own and then invade the other species from rare. If small organisms like green algae were used, one can perform chemostat experiment and track species *i*’ density or biomass through time. From the resulting growth curve, one can measure the per capita growth rate of species *i* when growing alone from rare (*μi*). When species *i* reach steady-state, e.g. when species *i*’ density change is not significantly different from 0 for three successive days, the second species *j* would be introduced at low density, e.g. 0.01% of the resident species. After introducing species *j*, one continues to monitor the density of species *j*. The density of invading species will then be fit to exponential growth curve to calculate the per capita growth rate of species *j* when its competitor (species *i*) is at its carrying capacity (*μji*). Similarly, by invading species *i* to the steady-state of species *j*, one can measure *μi* and *μji*. Consequently, species’ sensitivity to competition (*Si*) can then be calculated by applying equation 3. Finally, following the deduction in appendix B, ND and RFD can also be calculated and Chesson’s inequality (equation 1) can be assessed from the sensitivity metric (*Si*).

* + 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 are easy to measure, such as for small and fast-growing organisms (bacteria, algae, invertebrates, etc.). Mutual invasibility experiments are harder to apply to organisms that require long periods of time to measure their population growth rates, like perennial plants or mammals.

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

Parameterizing MacArthur’s consumer resource model (MacArthur 1969, 1970) and reorganize it into a Lotka-Volterra form is another method to measure ND and RFD and assess Chesson’s inequality. By doing so, empirically measured parameters in MacArthur’s consumer resource model can be translated into parameters in Lotka-Volterra model, and thus 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). Chesson has shown how such reorganization can be done by applying time scale separation technique (Chesson 1990, 2000). Finally, Chesson’s coexistence inequality can be accessed when the parameters in MacArthur’s consumer resource model is obtained.

* + 1. *Empirical approaches*

Parameterizing MacArthur’s consumer resource model requires feeding experiments in which each consumer species is grown across a gradient of resource species density (inset figure of Fig. 4). From the feeding experiments, the empiricist measures (1) the per capita consumption of the focal species on all the required resource species and (2) Conversion efficiency from each resource species to the focal consumer species. Finally, the per capita growth rate and carrying capacity of each resource also need to be measured in the absence of consumer species.

* + 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 Tilman’s two-species consumer resource model to a Lotka-Volterra form. By doing so, parameters in Tilman’s consumer resource model can be used to estimate the intra- and inter-specific competition coefficients, which are consistent with Chesson’s equations for ND and RFD as well as the inequality for species coexistence (equation 1).

* + 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 green algae species (*Asterionella formosa* and *Cyclotella meneghiniana*) compete for two essential resources (silicate and phosphate). In general, measuring parameters in Tilman’s resource ratio model can be more easily done with freshwater phytoplankton in chemostat experiments (e.g. Tilman 1980, 1981), though other types of organisms are applicable as well (Miller et al. 2005). Figure 5 shows the hypothetical results from chemostat experiments to parameterize the Tilman’s resource ratio model.

* + 1. *Limitations*

Reorganizing Tilman’s consumer resource model is similar to conducting a hypothetical mutual invasibility experiment, which means the growth of each species is evaluated at the resource condition that would result from a steady-state population of the other species. However, this should be theoretically valid and should not affect ND and RFD calculation for assessing mutual invasibility, because mutual invasibility is to assess whether each species invades a steady-state population of the other.

1. *Do the methods yield qualitatively similar predictions for coexistence?*

We showed that with the exception of the NFD method, the parameters or measurements of the other four methods can all be used to calculate ND and RFD, and to assess Chesson’s inequality for coexistence. Although these four methods are developed under different theoretical origins and make differing assumptions, these four methods thus give qualitatively the same predictions for coexistence.

We further showed that the NFD slope from the NFD method can only be used to predict mutual invasibility when the negative frequency is linear, i.e. the NFD slope is constant, and when the NFD slope is measured close to zero frequency of the focal species. Therefore, the condition for the NFD method to correctly predict coexistence is the same as conducting mutual invasibility experiment and directly demonstrate mutual invasibility. In addition, the fact that the NFD method does not give estimates of ND and RFD limits the comparisons that can be made to the other methods.

**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 four 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 how many different experimental settings that would be required for the method. Finally, the section ‘Outputs’ explains the types of predictions that the method can make.

* 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 (equation 1). If a method cannot be directly related back to Chesson’s inequality, that method might generate different prediction of coexistence due to its own assumptions or attributes. Question 1 isolates the negative frequency dependence method from all others. While the NFD method can accurately quantify mutual invasibility, it is the only method that cannot be used to directly quantify the magnitude of ND and RFD in order to predict species coexistence using Chesson’s inequality. While this decision step separates the NFD method from other methods, it is worth noting that the NFD method is the only one that does not require that species are all grown alone as monocultures. This could be an advantage for study systems where experimental manipulations are not feasible (e.g. long-lived species, protected habitats). In addition, the negative frequency dependence method is one of only two methods that has been applied to observational data, which might be necessary for some study systems.

The next question in the decision step is whether the empiricist knows the resource(s) the species are competing for in their study system, and how species are competing for these resources. This question separates the four methods for estimating ND and RFD into two separate groups: phenomenological and mechanistic methods. The phenomenological methods are the two (the Lotka-Volterra and Sensitivity methods) that are informed by quantifying species interactions, but which make no assumptions about what resource(s) the species compete for, and which resources define their niche differences. In many cases it will be possible for an empiricist to choose phenomenological methods over mechanistic methods, because knowing exactly what resource(s) species are competing for might require exhaustive prior knowledge about the study system. Choosing the phenomenological methods thus offers an advantage that ND and RFD can still be quantified even if an empiricist does not have a good understanding of which resource(s) species are competing for, and thus, which resources axes define their niche.

On the other hand, the other two mechanistic methods are based on consumer resource models in which the resource(s) species compete for are assumed to be known. The advantage of the mechanistic methods is that these methods requires less experimentation efforts especially when encountering novel combinations of species and environmental conditions. Choosing mechanistic methods allows an empiricist can predict coexistence of any species pair based solely on experiments that take measurements from specie grown alone in monoculture. For example, when designing experiments to predict coexistence among three hypothetical specie, e.g. A, B and C, applying mechanistic methods require three experiments: species A alone, B alone, and C alone. These experiments can be used to estimate the parameters of a consumer resource model, e.g. consumption rates on known resources, for calculating ND and RFD as well as predicting coexistence of any pairwise combination of these three species. On the other hand, applying phenomenological methods would require 6 experiments: species A, alone, B, alone, C alone, and three pairwise species combinations. In addition, the resulting metrics for predicting coexistence, e.g. species’ sensitivity to competition (*Si*) only apply to the particular conditions under which the experiments are conducted. Because this decision step is so influential, the remaining steps are particular to either the phenomenological or mechanistic methods.

Deciding between the two phenomenological methods, highlighted in green, depends on the answers to whether the method must work for observational datasets (question 3), whether it is experimentally possible to obtain a species monoculture that is at steady-state (question 4), and whether the method can be used to predict pairwise coexistence among more than two species (question 7). The two phenomenological methods differ in terms of whether they can predict species coexistence among more than two species: the Lotka-Volterra model can be parameterized to obtain all pairwise competition coefficients for a pool of species, but the sensitivity method requires that all non-focal species are considered in aggregate (e.g. species *i* invading a community of *j+k+l*).

Deciding between the two mechanistic methods, highlighted in blue, depends on whether the resource is abiotic and governed by a constant rate of supply (e.g. inorganic nutrients consumed by plants) or biotic with its own population dynamics (decision step 7). 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.

* 1. *Experimental Requirements*

In addition to the decision steps outlined in Table1, there are important practical differences for the number and types of experiments required to quantify ND and RFD for each method. The major difference is between the phenomenological and the mechanistic methods. For the phenomenological methods, the number of new experiments required for all pairwise combinations of species increases linearly or exponentially with each additional species being considered. Since these methods require long-time series and/or the growth of species at steady-state, empirical studies to assess pairwise coexistence among several species could be arduous. In contrast, adding additional species to consumer-resource method requires only as many additional experiments as the number of resources species are known to compete for. This is because the methods based on consumer-resource models do not require any direct competition experiments in order to estimate competition coefficients (*αii* or *αjj*), while all of the phenomenological methods require at least one co-culture of each species pair in order to quantify the competition coefficients.

* 1. *Method Outputs*

Having shown previously that the NFD method cannot give estimates of ND and RFD, there remain distinctions about what the methods can and cannot be used to predict. Importantly, only the consumer resource models have the ability predict the potential for coexistence among combinations of species without growing species together simultaneously (as discussed previously in the 2.1 Decision steps section). Moreover, consumer resource models can be used to predict ND and RFD under limited sets of different environmental conditions. For instance, Letten et al. 2017 showed that the Tilman R\* model can be used to predict the ND and RFD at different nutrient supply rates or dilution rates (Letten et al. 2017), but if for example, temperature were changed, the model could not be used to make predictions. In contrast, the outcome of pairwise LV or sensitivity experiments cannot be used to predict coexistence for novel pairs of species or under any changes to environmental conditions.

**Part 3. Cautions and future directions**

Having described and compared the methods for implementing modern coexistence theory, we offer several cautions for this area of research as more empirical studies are performed and, eventually, outcomes are synthesized from different experiments.

*Caution 1: Using only ND and RFD to assess mutual invasibility assumes negligible environmental fluctuations*

The five methods reviewed here only apply to fluctuation independent mechanisms of coexistence so that environmental dependent mechanisms, i.e. relative nonlinearity and storage effects, are assumed to be negligible. Empiricists should be able to neglect fluctuation dependent mechanisms and still correctly predict species coexistence, when conducting well-controlled experiments. However, it has been shown that environmental fluctuation can be important to mediate species coexistence in some empirical systems where environmental fluctuations cannot be negligible (Jiang and Morin 2007, Angert et al. 2009). When applying these five methods to predict coexistence, one should justify that environmental fluctuations can be negligible.

*Caution 2: Be aware to the limitation of each method*

Although we have shown that the empirical approaches for predicting coexistence among species can give the same qualitative outcome, each method has its own assumptions or limitations. For instance, the Lotka-Volterra method assumes that the per capita effect of species *i* on species *j* is independent of the density of either species *i* or species *j*. However, there are cases where this assumption is clearly not met (Schoener 1974, Abrams 1980). A hypothetical situation is that, if species were limited by resources (e.g. nutrients), a positive saturating relationship between the availability of resources and per-capita growth rate would mean that density-dependence is weak at low population sizes and stronger at higher population densities. Thus, both the inter and intraspecific competition coefficients would appear to be very small if measured at low population densities and very high if measured at population densities approaching the steady-state biomass. In designing and interpreting experiments, it would be unclear which value to select for the interaction coefficients. Consequently, it is important to apply the empirical method at the conditions that follow the assumptions. When those assumptions are not recognized and justified, any of these five methods can give misleading predictions.

*Caution 3: Limits to the applicability of CRM.*

As shown in table 1, using either of the methods based on consumer-resource models requires that the empiricist knows what the resources that species are competing for. This is more easily achieved for certain experimental systems (e.g. microbes grown under laboratory conditions) than others (e.g. mammals). However, in reality, it is hard to know for sure which resources or factors govern population dynamics. While these experiments are useful for isolating the *mechanism* of competition, they require detailed knowledge about natural history of the organisms, which in many cases is unknown.

*Future directions*

We have shown that five methods for empirically evaluating niche difference (ND) and relative fitness difference (RFD) are all capable of predicting coexistence, but there are important distinctions among these methods that make each empirical method suitable for different conditions. We offer future directions that could benefit the development of modern coexistence theory. First, the modern coexistence theory is under-developed for multi-species systems (but see Carroll et al 2011). In theory, competitive hierarchy between species *i* versus *j* and *j* versus *k* might not directly translate to species *i* and *k*, when these species are engaged in intransitive competition or higher-order interaction (Levine et al. 2017). In fact, none of these three phenomenological methods (NFD, LV, and Sensitivity) 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, we need to empirically demonstrate equivalence of these methods. To date, we are unaware of any empirical studies that have applied more than one of these methods to the same study system. According to what we have mathematically shown in part 1, all methods except the negative frequency dependency method can be reduced to the same algebra to calculate ND and RFD, and to assess Chesson’s inequality. We argue that ND and RFD derived from the NFD method should not be comparable to the other methods. However, as we show here, a few papers have applied empirically-derived parameter values obtained from one method to show that two methods are comparable (Levine and HilleRisLambers 2009, Letten et al. 2017), but these ad hoc tests do not reflect the differences in experimental design, assumptions, and calculations that are outlined in Table 2. Without knowing whether these methods are empirically equivalent, we cannot make any synthesis and make inference on the coexistence mechanisms from empirical perspective. We thus advocate studies to empirically demonstrate the equivalence of these methods. As the number of empirical studies using these methods grows, it is critical that differences among these methods are acknowledged.

**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. 2012, 2017). 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.

**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

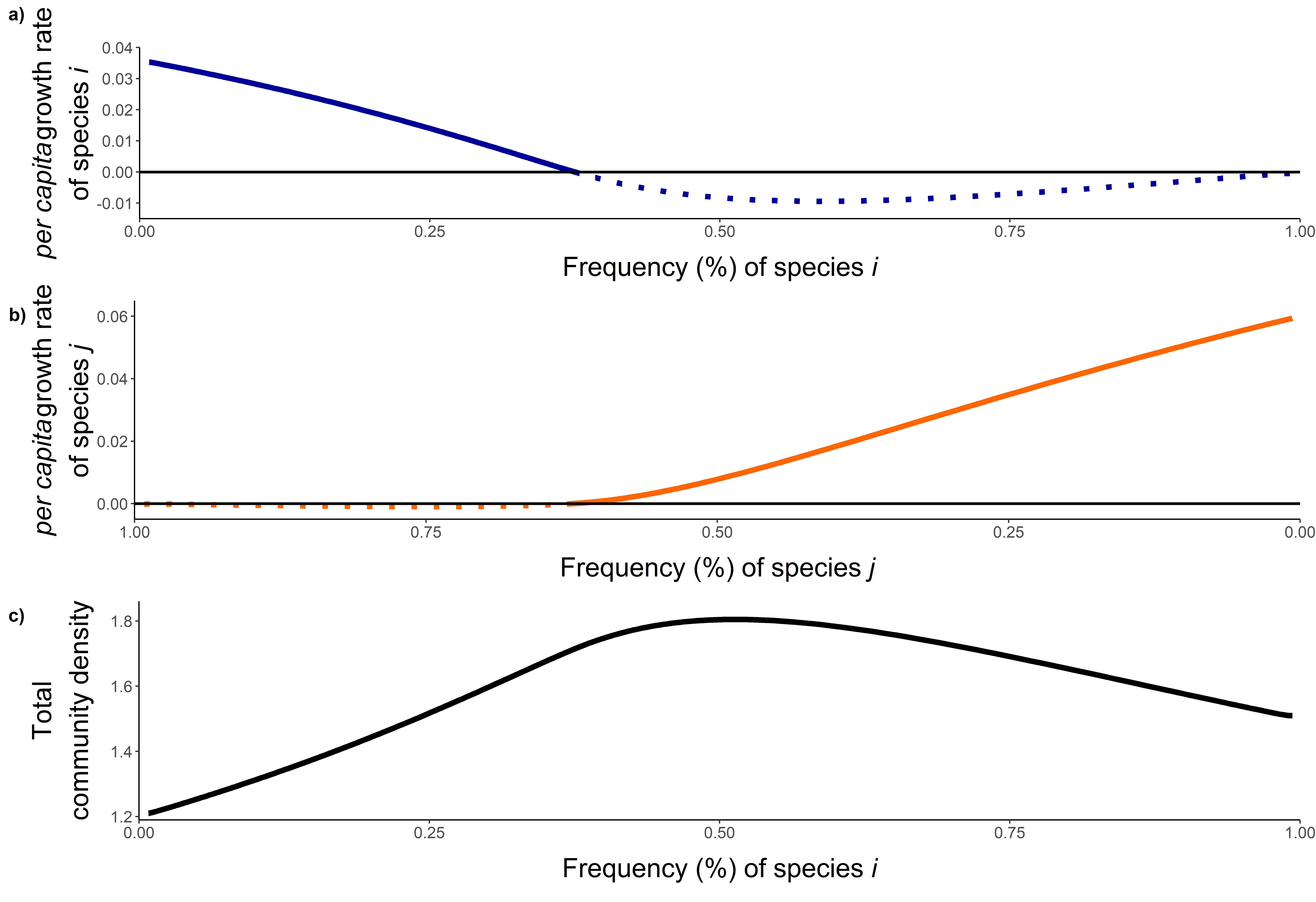


Figure 1. Simulation results of species’ per capita growth rate versus its frequency (panel a, b) from two-species Lotka-Volterra model ( and how the community density change with species’ frequency (panel c). The numerical simulations are performed with arbitrarily defined parameter values that allows 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. means the minimum level of resource *j* that still allows species *i* to have positive per capita growth rate. The *R\**s are those being 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*.

**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.

Figue S3.

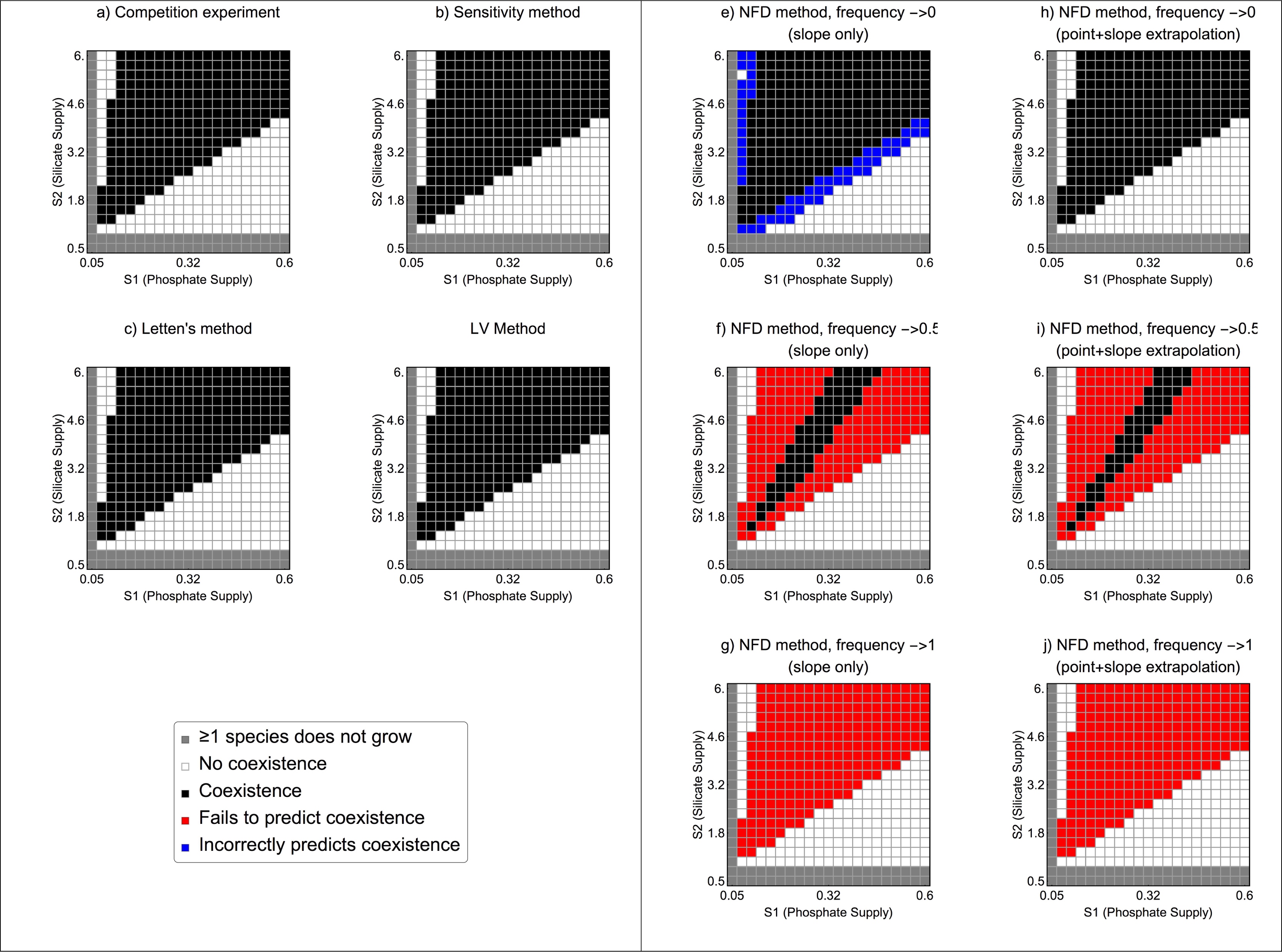


Figure S2. Simulation results showing the parameter space in which two species are predicted to stably coexist using different methods. Panel a (the “competition experiment”) is the simulation of both species growing from rare and allowed to reach equilibrium biomass. Panel b shows the predictions from Letten’s method, which is not based on numerical simulation, but uses the parameter values and resource supply concentrations to predict the inter- and intra-specific interaction coefficients as specified by the original authors. Panel c shows the predictions based on the sensitivity method, where growth rates were determined from simulations 1-4 and used to compute sensitivities as directed by Carroll et al 2011. Panel d shows the prediction based on computing Lotka-Volterra interaction coefficients. In this method, the per - capita impact of an invader on the resident' s per - capita growth rate was calculated using the time series from simulations 3 and 4. Additionally, the intraspecific interaction coefficient was determined as the per-capita impact of each species on its own growth rate when the population size approaches 99% of its equilibrium value (simulations 1 and 2). Panels e-j show the predictions based on the NFD method, evaluated at different frequencies of the two species and both using only the slope of NFD (e-g) and the extrapolation to estimate growth rates when rare (h-j).

**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.

Barabás, G., R. D’Andrea, and S. M. Stump. 2018. Chesson’s coexistence theory. Ecological Monographs 88:277–303.

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.

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. Ke, and T. Fukami. 2018. Species coexistence through simultaneous fluctuation-dependent mechanisms 115:6745–6750.

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.

MacArthur, R. 1969. Species packing, and what competition minimizes. Proceedings of the National Academy of Sciences of the United States of America 64:1369–71.

MacArthur, R. 1970. Species packing and competitive equilibrium for many species. Theoretical Population Biology 1:1–11.

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.

Miller, T. E., J. H. Burns, P. Munguia, E. L. Walters, J. M. Kneitel, P. M. Richards, N. Mouquet, and H. L. Buckley. 2005. A critical review of twenty years’ use of the resource-ratio theory. The American Naturalist 165:439–448.

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.

Schoener, T. W. 1974. Some Methods for Calculating Competition Coefficients from Resource-Utilization Spectra. The American Naturalist 108:332–340.

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.

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

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