

RESEARCH ARTICLE

Estimating competition in metacommunities: accounting for biases caused by dispersal

Liang Xu¹  | Adam T. Clark²  | Mark Rees³  | Lindsay A. Turnbull¹ 

¹Department of Plant Sciences, University of Oxford, Oxford, UK

²Institute of Biology, Karl-Franzens-Universität Graz, Graz, Austria

³Department of Animal and Plant Sciences, University of Sheffield, Sheffield, UK

Correspondence

Lindsay A. Turnbull

Email: lindsay.turnbull@plants.ox.ac.uk

Funding information

John Fell Fund, University of Oxford, Grant/Award Number: 0008262

Handling Editor: Res Altwegg

Abstract

1. Estimating the strength of interactions among species in natural communities has always been a challenge for empirical ecologists. Sessile organisms, like plants or corals, often occur in metacommunities where they compete only with their immediate neighbours but disperse propagules over a wider area. To estimate the strength of competitive interactions, ecologists often count abundances in cells on a spatial grid for at least two time-points. This data is then analysed using regression, by modelling the change in population size as a function of local densities, using cells as independent data-points: a technique known as space-for-time substitution. These analyses generate estimates of competition coefficients; however, the method ignores dispersal among cells.
2. To determine the impact of ignoring dispersal, we derived the bias that would arise when we apply regression methods to a metacommunity in which a fraction of seeds disperse beyond their natal cells but this dispersal is ignored in the model fitting process. We present results from a range of population models that make different assumptions about the nature of competition and assess the performance of our bias formulae by analysing data from simulated metacommunities.
3. We reveal that: estimates of competition coefficients are biased when dispersal is not properly accounted for; and the resulting bias is often correlated with abundance, with rare species suffering the greatest overestimation. We also provide a standardized metric of competition that allows the bias to be calculated for a broad range of other population models.
4. Our study suggests that regression methods that ignore dispersal produce biased estimates of competition coefficients when using space-for-time substitution. Our analytical bias formula allows empirical ecologists to potentially correct for biases, but it requires either tailored experiments in controlled conditions or an estimate of the average dispersal rate in a natural community, so may be challenging to apply to real datasets.

This is an open access article under the terms of the [Creative Commons Attribution](https://creativecommons.org/licenses/by/4.0/) License, which permits use, distribution and reproduction in any medium, provided the original work is properly cited.

© 2022 The Authors. *Methods in Ecology and Evolution* published by John Wiley & Sons Ltd on behalf of British Ecological Society.

KEYWORDS

bias in estimation, conspecific and heterospecific competition, metacommunities, negative density dependence, regression models, simulations, spatially explicit models

1 | INTRODUCTION

Species interactions are critical to understanding how species coexist across space and time, but measuring their strength is hard (Adler et al., 2018; Chesson, 2000; Connell, 1971; Lotka, 1926; Petermann et al., 2008; Volterra, 1926). Ecologists have spent a great deal of time debating the pros and cons of different methods; for example, in plant communities it has been noted that field-based estimates might be qualitatively different—and potentially more realistic—than those obtained in experimental settings (Adler et al., 2018; Tuck et al., 2018). However, recent papers have emphasized that estimates obtained using observational field data are fraught with other challenges (Chisholm & Fung, 2018; Detto et al., 2019; Freckleton et al., 2006; Hulsmann et al., 2021; Knappe & de Valpine, 2012; Rinella et al., 2020).

These recent studies have highlighted that the strength of within-species competition is likely to be systematically overestimated because of sampling error (Detto et al., 2019; Freckleton et al., 2006; Hulsmann et al., 2021). However, while the impact of sampling error has received detailed attention, model mis-specification is likely to be a much more widespread problem (Clark et al., 2010), as studies often pick a particular population model to represent their community with little justification for their choices. Addressing this problem is harder because it requires knowing the 'true' model under which individuals grow and compete.

To estimate competition coefficients, the chosen population model is usually rearranged so that linear regression techniques can be applied to the data collected. In a single well-mixed population, a census might be conducted each year, but it will take many years of data to create a sufficiently long time-series for robust analysis, as each transition will only yield a single data-point. But in communities of sessile organisms, adults compete locally, while propagules (whether seeds or juveniles) disperse more widely, resulting in a metacommunity. The spatial structure within metacommunities offers a practical advantage over a single well-mixed population when it comes to estimating the strength of competitive interactions: using many cells, analysis can potentially be conducted after just 2 years by employing a space-for-time substitution. Even when the system is spatially contiguous, plant ecologists often count individuals in multiple small quadrats or cells whose spatial scale is usually chosen in a somewhat arbitrary way (Hart et al., 2018; Tuck et al., 2018). If counts are obtained for two consecutive years, then the same regression techniques can be applied, with each cell providing a single data-point.

The method relies on demographic stochasticity to create variation in densities among cells, but usually ignores dispersal. However, if dispersal is not explicitly modelled then fitting a local population model to a spatially structured dataset produces biased estimates of

competition, as shown by Freckleton et al. (Freckleton et al., 2006; Figure 2) for a single-species metacommunity consisting of two cells. This makes intuitive sense, because the local population model links the per-capita growth rate to the population density within cells, but immigration from other cells uncouples this relationship. Unfortunately, while Freckleton et al. (2006) raised the problem, they did not offer any analytical solutions to deal with the resulting bias, nor did they include competitor species in their analyses, making it unclear whether estimates of between-species competition would also be biased.

In this paper, we work through a series of analytical results and simulations to explore how failing to account for dispersal within a metacommunity can bias estimates of competitive effects both within and between species. We consider three commonly used local population models (Ricker, power-law and Beverton–Holt) that link population growth rates to competitor densities using different functional forms and compare each one to its spatially structured counterpart. By using a general metric for the competition coefficient, we are able to derive formulae for the bias that will arise in estimates of the competition coefficients, if the local model is fitted without taking dispersal into account. We test the performance of the bias formulae using a simulation model of annual plants from which we census the number of individuals in grid cells in two consecutive years, to mirror the spatial datasets often collected by plant ecologists in nature (Hart et al., 2018; Rees et al., 1996; Tuck et al., 2018). We show that our analytically derived formulae capture the magnitude of the bias caused by dispersal, and suggest how this can be addressed in real datasets. These analyses make use of a general metric that compares different definitions of competition across population models and allows our method to be extended to a broad range of other potential models.

2 | MATERIALS AND METHODS

2.1 | Overview

The kinds of analyses presented in this paper are appropriate for communities of sessile organisms, like plants or barnacles, that are unable to move, and so tend to compete only with their close neighbours. However, adults produce propagules or seeds that disperse over larger distances, creating a metacommunity. These systems are usually idealized as a series of cells, within which individuals compete for limiting resources, while seeds or propagules disperse among cells.

To explore what happens when we apply a local population model to a metacommunity and ignore the spatial structure, we begin with the Ricker model, which has been widely used for modelling both animals and plants (Comita et al., 2010; Johnson et al., 2012; LaManna, Belote, et al., 2017; LaManna, Mangan, et al., 2017; Ricker, 1954). We

first consider how a single well-mixed population behaves using the deterministic version of the model. We then see how applying the model to data collected from a metacommunity leads to biased estimates and crucially gain insight into how and why these biases arise.

2.2 | The Ricker model

The **Ricker model** (Ricker, 1954) describes the population size ($N_{i,t+1}$) of the i th species in a community of n competitors as:

$$N_{i,t+1} = N_{i,t} e^{r_i - \sum_{j=1}^n \alpha_{ij} N_{j,t}}, \quad (1)$$

where r_i is the intrinsic population growth rate and α_{ij} quantifies the effect of species j on species i . Following long-standing notation, a positive value of α_{ij} indicates competition while a negative value denotes facilitation. Rearranging Equation 1 and taking logs of both sides gives a linear expression for the per capita population growth rate:

$$\log \frac{N_{i,t+1}}{N_{i,t}} = r_i - \sum_{j=1}^n \alpha_{ij} N_{j,t}, \quad (2)$$

where the slopes correspond to the interaction coefficients (the α_{ij} terms) and the intercept corresponds to the intrinsic population growth rate (r_i).

The Ricker model exhibits a wide range of behaviours, including limit cycles and chaos. For example, for intermediate values of $1 < r_i < 2$ in a single-species model the population exhibits damped oscillations (Figure 1a). This behaviour occurs because the population initially overshoots the carrying capacity, inducing over-compensating density dependence (sensu Myers, 2001) once $N_{i,t}$ exceeds $1/\alpha_{ii}$ (Figure 1b). Whether or not natural populations are likely to exhibit this kind of over-compensating competition is debated, so in subsequent sections (*Other population models*), we consider two further widely used population models: the power-law (Kellner & Hubbell, 2018; Wright et al., 2005) and the Beverton–Holt (Beverton & Holt, 1981) models.

The Ricker model can be parameterized by fitting Equation 2 to counts of individuals over time. In a well-mixed population, a long time-series is required to obtain good estimates; however, in principle, all values of $\log(N_{i,t+1}/N_{i,t})$ fall on the regression line, as there is no stochasticity, (Figure 1b), so recovering the true strength of competition (the α_{ij} terms) does not depend on knowing whether or not a population has reached equilibrium. However, while the estimates obtained from a single well-mixed population should be unbiased – no matter where the population lies on its trajectory – the same is not true once we move to a metacommunity with demographic stochasticity.

2.3 | The spatially structured Ricker model

One simple way to represent dispersal within a metacommunity is to assume that some fixed fraction (hereafter “the stay rate”, or δ) of the seeds or juveniles stay in the natal cell while the rest ($1 - \delta$)

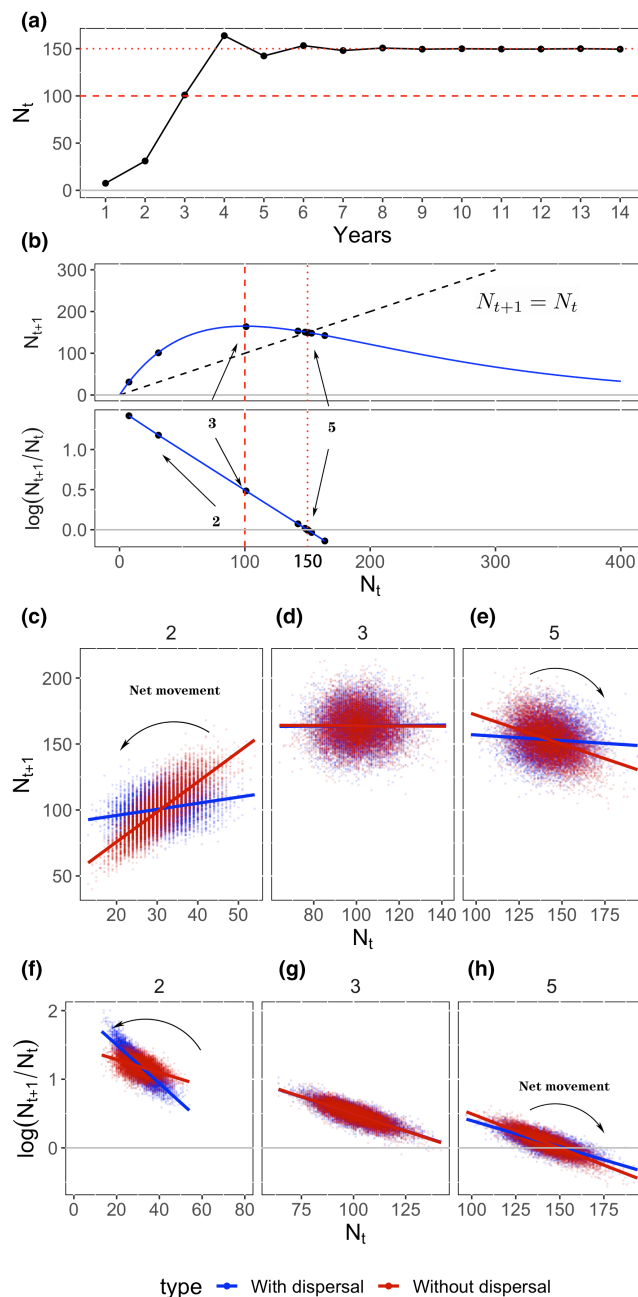


FIGURE 1 Behaviour of the Ricker model. (a) Trajectory (b) future population size (N_{t+1}) and per-capita population growth rate as a function of current density (N_t) for the deterministic Ricker model (a single well-mixed population). The dotted red line denotes the equilibrium population size, and the dashed red line denotes the point at which N_{t+1} is maximized, which occurs when $N_{i,t} = 1/\alpha_{ii}$. (c–h) Future population size (N_{t+1}) and the per-capita population growth rate as a function of current density (N_t) in a spatially structured version of the Ricker model at three time points. Data points are drawn from simulations of a metacommunity consisting of 10,000 cells, both with dispersal (blue dots and lines) and without dispersal (red dots and lines). The density variation among cells in the spatially structured model (c–h) is due to demographic stochasticity. Parameter values are $r = 1.5$, $\alpha_{11} = 0.01$, $\delta = 0.2$.

disperse over the community, for example, by joining a global seed rain. Under these circumstances, the population dynamics within a local cell i is given by:

$$N_{i,j,t+1} = \delta N_{i,j,t} e^{r_i - \sum_{j=1}^n \alpha_{ij} N_{j,t}} + I_{i,j,t}, \quad (3)$$

where $I_{i,j,t}$ is the number of immigrating seeds or juveniles (see [Supplementary Material](#) for a detailed derivation). Note, that although [Equation 3](#) contains competition coefficients (α_{ij}) their precise mathematical definition is different to those in the local Ricker model ([Equation 1](#)).

To derive an expression for the competition coefficients that would be obtained ($\hat{\alpha}_{ij}$) if we fitted the Ricker model to the spatially structured dataset, but ignored dispersal among cells, we need to directly compare the competition coefficients in the two versions of the model ([Equations 1](#) and [3](#)). To do this, we first need to define a general metric of competition (α_{ij}^G) that allows such interconversions. This takes the form:

$$\alpha_{ij}^G = \frac{\partial \log \frac{N_{i,t+1}}{N_{i,t}}}{\partial N_{j,t}}. \quad (4)$$

which is mathematically convenient for the Ricker model, as the terms correspond to the interaction coefficients (the α_{ij} terms), but also allows us to perform the same operation for other widely used local population models ([Table 1](#)). Biologically, the general metric of competition (α_{ij}^G) represents the change in the growth rate of a focal species i with respect to the change in the density of competitor species j .

Using [Equation 4](#) we can now produce the following approximations for the Ricker model:

$$\hat{\alpha}_{ij} \approx \delta \alpha_{ij} + (1 - \delta) \frac{1}{\bar{N}_{i,t}}, \quad (5)$$

$$\hat{\alpha}_{ij} \approx \delta \alpha_{ij}. \quad (6)$$

where $\bar{N}_{i,t}$ is the average population size observed across all cells (the expressions are approximations that we believe would hold in a homogeneous environment; see [Supplementary Material](#) for more details). The expressions reveal that the coefficients obtained when fitting the local model to the spatially structured dataset ($\hat{\alpha}_{ii}$ and $\hat{\alpha}_{ij}$) are biased, as in neither case are they equal to the true parameter values (α_{ii} and α_{ij}). In particular, estimates of both coefficients depend on the fraction of seeds or juveniles that stay within the natal cell (δ), although the

estimated value of the conspecific coefficient ($\hat{\alpha}_{ii}$) also depends on the average population size ($\bar{N}_{i,t}$).

Thus, if we collect data from multiple quadrats within a meta-community and treat that data as if each cell were isolated from the rest, then we will obtain biased estimates of competitive effects. The bias in the conspecific coefficient is expected to change as the population size changes, while the bias in the heterospecific coefficient is expected to be independent of population size.

2.4 | Understanding the bias in the Ricker model

Theoretically, the bias formula ([Equation 5](#)) for the conspecific competition coefficient indicates that the biased estimate ($\hat{\alpha}_{ii}$) falls between the true value (α_{ii}) and the inverse of the mean density, averaged across all cells ($1/\bar{N}_{i,t}$). This immediately gives us a criterion by which to decide whether we expect to see overestimation ($\hat{\alpha}_{ii} > \alpha_{ii}$) or underestimation ($\hat{\alpha}_{ii} < \alpha_{ii}$) of the strength of conspecific competition. When $1/\bar{N}_{i,t} > \hat{\alpha}_{ii}$, then $1/\bar{N}_{i,t} > \hat{\alpha}_{ii} > \alpha_{ii}$ ([Equation 5](#)), leading to overestimation of the strength of conspecific competition. In contrast, when $1/\bar{N}_{i,t} < \hat{\alpha}_{ii}$, then $1/\bar{N}_{i,t} < \hat{\alpha}_{ii} < \alpha_{ii}$, and the strength of conspecific competition is underestimated. The switching between over and under-estimation corresponds to the switch from under- to over-compensating competition in the non-spatial version of the Ricker model, that is, at the point where $\bar{N}_{i,t} = 1/\bar{\alpha}_{ii} = 1/\alpha_{ii}$ ([Figure 1c–h](#)).

Our analysis suggests that positive bias (over-estimation of the strength of conspecific competition) is more likely when the average density of the focal species is low ($\bar{N}_{i,t} < 1/\alpha_{ii}$); for example, at the early stages of population growth or when the species is suppressed by strong competition in a mixed community. This occurs because, when the average density is low, the Ricker model defines a positive relationship between total seed production and adult plant density across cells (cells with high plant densities produce more seeds than cells with low plant density [[Figure 1c](#)]). In turn, this means that there is net movement of seeds from high-density cells to low-density cells, and this movement steepens the slope in the regression relationship ([Figure 1f](#)) and leads to overestimation of the strength of conspecific competition.

In contrast, a negative bias (under-estimation of the strength of conspecific competition) is more likely to occur when the average

Scenarios	Interaction matrix	Stay rate	Population models
1 species	$\alpha_{11} = 0.01$	0.2, 0.4, 0.8	Ricker
2 species	$\alpha_{11} = 0.005, \alpha_{12} = 0.0015, \alpha_{21} = 0.0036, \alpha_{22} = 0.01$	0.2, 0.4, 0.6, 0.8, 1	Ricker
2 species	$\alpha_{11} = 0.01, \alpha_{12} = 0.0089, \alpha_{21} = 0.009, \alpha_{22} = 0.01$	0.2, 0.4, 0.6, 0.8, 1	Ricker
1 species	$\alpha_{11} = 0.5$	0.2, 0.4, 0.6, 0.8, 1	Power-law
2 species	$\alpha_{11} = 0.005, \alpha_{12} = 0.0015, \alpha_{21} = 0.0036, \alpha_{22} = 0.01$	0.2, 0.4, 0.6, 0.8, 1	Beverton–Holt

TABLE 1 Parameters used in simulation studies. The growth rates of all species in simulations of the Ricker model and the Beverton–Holt model, $r = 1.5$, while for the Power-law model, $r = 10$

density of the focal species is high ($\bar{N}_{i,t} > 1/\hat{\alpha}_{ij} > 1/\alpha_{ij}$); for example, when the focal species is close to equilibrium in monoculture or when it is numerically dominant in a mixed community. This occurs because, when plant density is high ($\bar{N}_{i,t} > 1/\alpha_{ij}$), the Ricker model describes a negative relationship between total seed production and adult plant density (Figure 1e). Thus, when seeds disperse from cells where seed production is high to cells where seed production is low, the net flow of seeds is actually from low plant density to high plant density. This process flattens the slope in the regression relationship (Figure 1h), resulting in underestimation of the strength of conspecific competition. Importantly, there is a critical point where the bias does not occur and density dependence is almost perfectly compensating (Figure 1d,g). This result suggests a potential empirical approach whereby biases can be avoided without applying a correction formula (see Discussion: Section 4.2).

The bias formula also indicates that the strength of heterospecific competition in the Ricker model is underestimated by a factor of δ (Equation 6). The bias occurs because seed output declines as the number of heterospecific competitors increases, so the net movement of seeds is from cells with lower densities of competitors to cells with higher densities of competitors. This reduces the negative relationship between the per capita growth rate and the density of competitors, leading to underestimation of the strength of heterospecific competition (see Figure S1).

Taken together, we can conclude that if dispersal among cells is ignored when using space-for-time substitutions, then biased estimates of competition coefficients will result. The direction of the bias depends on the density of the population and the identity of competitors (conspecific vs. heterospecific).

2.5 | The power-law model

To test the generality of our results, we also explored the bias in estimates of the strength of competition for two other population models – the power-law and the Beverton–Holt model. The **power-law model**:

$$N_{i,t+1} = rN_{i,t}^{\alpha_{ij}}, \quad (7)$$

is often applied to single-species populations, particularly when modelling plants (Kellner & Hubbell, 2018; Wright et al., 2005). The nature of the density-dependence changes with the value of α_{ij} : when $0 < \alpha_{ij} < 1$, the number of recruits in the following year ($N_{i,t+1}$) increases monotonically with density ($N_{i,t}$); hence, there is no over-compensating density dependence; however, over-compensation (sensu Myers, 2001) is seen when $\alpha_{ij} < 0$. The relationship is linear on a log scale, giving:

$$\log(N_{i,t+1}) = \log(r) + \alpha_{ij}\log(N_{i,t}), \quad (8)$$

where the slope yields an estimate of α_{ij} . Using the method described above for the Ricker model yields the following bias formula (see Supplementary Material for derivation details):

$$\hat{\alpha}_{ij} \approx \delta\alpha_{ij}. \quad (9)$$

As for the Ricker model, the expressions are approximations that we believe would hold in a homogeneous environment (see Supplementary Material for more details). Equation 9 reveals that the conspecific competition coefficient, α_{ij} , is underestimated by a factor equal to the stay rate, δ , meaning that the strength of conspecific competition will be overestimated regardless of population size when $0 < \alpha_{ij} < 1$ and underestimated when $\alpha_{ij} < 0$. This bias occurs because there is a net flow of seeds from cells with high plant density to cells with low plant density, which strengthens the negative relationship between the number of seeds and plants across cells when $0 < \alpha_{ij} < 1$ and weakens the negative relationship when $\alpha_{ij} < 0$ (see Figure S2, S12 and also Figure 1c).

2.6 | The Beverton–Holt model

The **Beverton–Holt model** allows estimation of both conspecific and heterospecific effects. It is widely used in fisheries and takes the form:

$$N_{i,t+1} = rN_{i,t} / \left(1 + \sum_k \alpha_{ik} N_{k,t}\right). \quad (10)$$

The model cannot be linearised using a log-transformation, but non-linear least-squares methods can be used to fit the model to appropriate data. Using the method described above for the Ricker model yields the following expressions for the bias formulae (see also Supplementary Material):

$$\hat{\alpha}_{ij} \approx \alpha_{ij} + \frac{(1-\delta)(1 + \sum_{k=1}^n \alpha_{ik} \bar{N}_{k,t})}{\delta \bar{N}_{i,t}}, \quad (11)$$

$$\hat{\alpha}_{ij} \approx \alpha_{ij}. \quad (12)$$

The bias formulae predict that the strength of conspecific competition will always be overestimated given positive α_{ij} (although these expressions are approximations that we believe would hold in a homogeneous environment; see Supplementary Material for more details; Figure S3). Again, rare species with lower mean density suffer greater overestimation from the bias caused by dispersal. Surprisingly, and in contrast to the Ricker model, the heterospecific competition coefficient is unbiased by dispersal (see Supplementary Material for proof). This difference is due to the different definition of the competition coefficients in the Beverton–Holt model compared to the Ricker model, which we return to in the Discussion (Section 4.1).

2.7 | Simulations

To test whether our theoretical predictions are broadly correct, we simulated selected metacommunities of annual plants and compared the observed bias in estimated coefficients to the relevant bias formulae. Each community consists of a grid of 10,000 cells on a torus

in which annual plants grow and compete within cells according to a local competition model (Ricker, power-law or Beverton–Holt) before setting seed (Figure 2).

We modelled seed dispersal in two ways: either by assuming that some fraction of locally produced seeds ($1 - \delta$) enters a global seed rain (Equation 3), that is, a uniform distribution; or by assuming a more realistic pattern of dispersal and using a Gaussian dispersal kernel (normalized on a discrete grid) to distribute seeds to neighbouring cells. We began each simulation by drawing the initial number of seeds for each cell from a random uniform distribution, ranging between 5 and 10.

When using a uniform dispersal kernel, the number of local seeds is drawn from a Poisson distribution with mean equal to the number of seeds that remain in the natal cell (determined by the stay rate, δ), whereas the number of immigrant seeds follows a Poisson distribution with mean equal to $(1 - \delta)$ times the average seed production. As such, the sum of these two numbers is still a Poisson distribution with the mean determined by local competition. This approach is equivalent to generating a Poisson distribution for seed production and then using a binomial distribution to determine the number that disperse. For the Gaussian dispersal kernel, we first calculate the expected number of seeds by summing over all sites weighted by the Gaussian dispersal kernel and then generate a Poisson number of seeds with this mean. Hence, even once the model has reached equilibrium, variation in densities will always persist among cells. Indeed, as outlined above, the space-for-time substitution relies on this spatial variation to apply the regression method to a single transition in time.

To examine how the bias in estimates of competitive effects changes through time, we simulated all populations for 15 years, which was sufficient for them to reach equilibrium. The number of plants in all cells was counted at each time-point. Depending on the local population model, we then applied the relevant regression method (Equations 2, 8 or 10) to a single transition to obtain estimates of competition coefficients for each year. We then compared these estimates to those predicted using the relevant bias formulae.

We carried out the most extensive simulations with the Ricker model. We simulated both monocultures and combinations of two

species with uniform and Gaussian dispersal for a range of dispersal rates. For mixtures of two species, we considered a case in which the two species had very different competitive abilities and a second case where the two species were more equivalent. These parameter combinations were chosen to bracket a range of scenarios and we did not attempt an exhaustive exploration (see Table 1 for parameter values). All regression analyses were carried out using the `glm` function in R with a Poisson link (version 4.0.5). To avoid fitting N_t on both sides of Equation 2 (Dennis & Taper, 1994), we used the following rearrangement:

$$\log(N_{i,t+1}) = r_i - \sum_{j=1}^n \alpha_{ij} N_{j,t} + \log(N_{i,t}), \quad (13)$$

and fitted N_t as an offset. Note that there is demographic stochasticity in fecundity (see Supplementary Material for R code).

For the power-law model, we carried out simulations of a single-species population with uniform and Gaussian dispersal for a range of dispersal rates. As for the Ricker model, all regression analyses were carried out using the `glm` function in R with a Poisson link (version 4.0.5). Finally, for the Beverton–Holt model, we simulated a two-species mixture comprising two unequal competitors with both uniform and Gaussian dispersal for a range of dispersal rates. Competition coefficients were estimated using the function `nls` fitted to untransformed densities in R (see Supplementary Material for details of starting values and other settings). See Table 1 for details of parameter values used in all simulations.

3 | RESULTS

3.1 | The Ricker model

We first report the results of simulations which use a uniform distribution to distribute dispersing seeds. For a monoculture of one species, the number of individuals in the metapopulation follows a

Simulation model

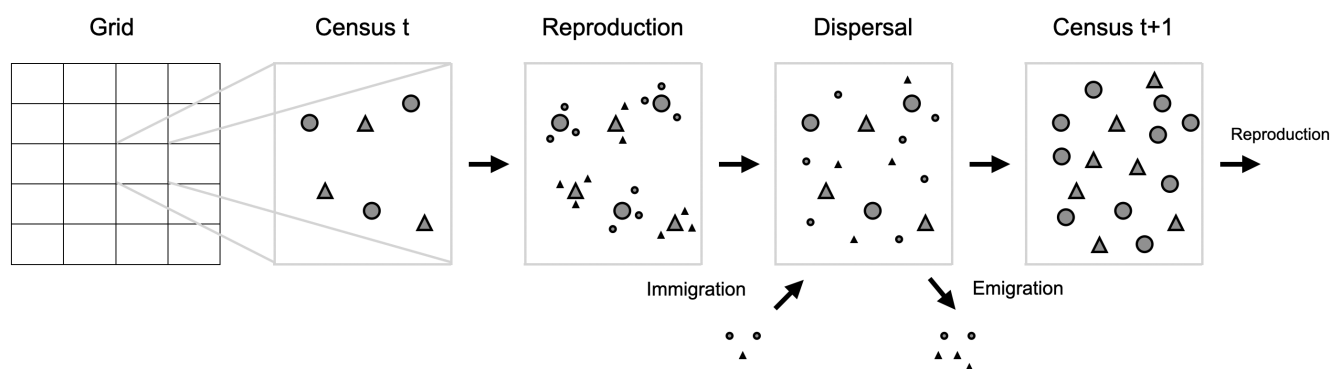


FIGURE 2 Schematic of the simulation model. Each cell hosts a number of adult plants (large symbols) which are censused each year. Adults produce seeds (small symbols) following a local population model with competition occurring only within cells. Seeds then disperse to other cells, following a dispersal kernel. Existing adults die and seeds grow into new adults that are censused at the following time step.

similar logistic pattern as the equivalent non-spatial version of the model. The dispersal rate ($1 - \delta$; Figure 3) has little impact on the overall growth curve (but see Figure S2) but there is significant bias in the estimates of the conspecific competition coefficients. The extent of this bias depends on the dispersal rate, which is captured well by the bias formula (Equation 5; Figure 3). In the early stages of growth, when population density is low ($t < 3$), there is positive bias, indicating that the strength of competition is over-estimated (as predicted by Equation 5). But, as the population grows, the bias declines and actually becomes negative for $t \geq 3$, indicating that the strength of competition is under-estimated (compare with Figure 1). In addition, as the stay rate declines (and hence the dispersal rate increases) the extent of the bias increases (Figure 3; Table 2).

For simulations involving two species, we began with a pair of unequal competitors: species 1 is competitively superior, and hence numerically dominant at equilibrium, while species 2 is a weaker competitor that persists at lower density. Both species increased their population sizes initially, but then the weaker competitor declined as the population of species 1 approached equilibrium (Figure 4a). For the conspecific competition coefficients, the bias pattern for the better competitor (species 1) declines from positive to negative as the population grows and crosses the threshold $1/\alpha_{11}$ (Equation 5). However, for the weaker competitor (species 2), the population only

exceeds $1/\alpha_{22}$ at $t = 3$ and drops below $1/\alpha_{22}$ immediately so the bias is mostly positive, although the bias is negative at $t = 3$. This means that at equilibrium, the conspecific competition coefficient is *underestimated* for the dominant competitor and *overestimated* for the competitively inferior species. The heterospecific competition coefficients (α_{12} and α_{21}) are consistently underestimated at all time points. As predicted, when the stay rate is low (and hence dispersal high) the bias is greater. In all cases, the bias formulae (Equations 5 and 6) give good agreement with the linear regression method.

In the simulation where the two species were more closely matched in terms of their competitive abilities, both populations increase from low density and reach similar abundances. Because their abundances are so similar, and are both below $1/\alpha_{ij}$ ($i = 1, 2$), they both suffer from positive bias (overestimation of the conspecific competition coefficients). As in the previous scenario, the heterospecific competition coefficients are, therefore, also both underestimated (Figure S4).

When using a Gaussian distribution to disperse seeds, we found little impact on the performance of the bias formula (Figures S5 and S6). This is because, if we rearrange Equations 5 and 6, we see that the bias, that is, the difference between the regression estimates and the true values, only depends on the stay rate (δ) and the mean density ($\bar{N}_{i,t}$) given the nature of competition (fixed α_{ij}), and is not

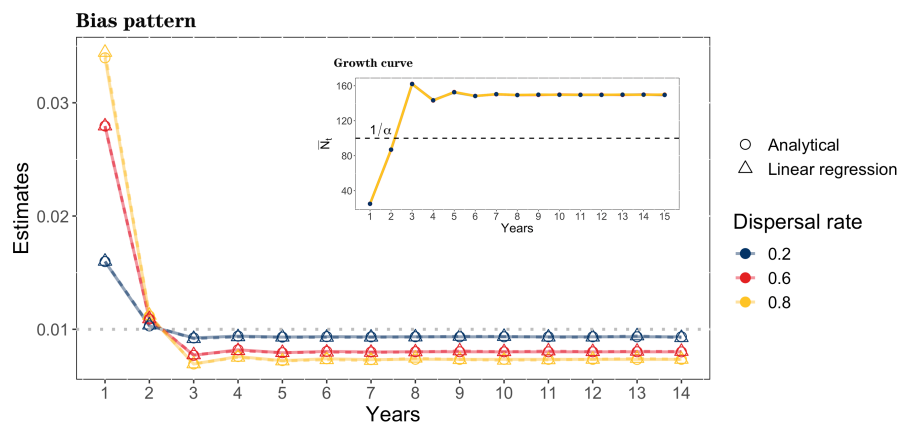


FIGURE 3 The change in bias during population development in simulations of monocultures with three different dispersal rates. The dotted grey line denotes the true value, $\alpha = 0.01$. Inset panel: Growth curves for each dispersal rate, which are near-identical. Dashed line indicates $\bar{N}_t = 1/\alpha$, at which point the bias switches from positive to negative (as predicted by Equation 5). Main panel: Estimates of conspecific competition using the linear regression method (dashed lines) and the analytical bias formula (solid lines). Intrinsic growth rate is set to $r = 1.5$ in all simulations.

TABLE 2 Specification of competition in six major population models and their relationship to the general metric defined in Equation 4 (α_{ij}^G)

Name	Model	α_{ij}^G	Reference
Ricker	$N_{i,t+1} = N_{i,t} e^{r_i - \sum_k \alpha_{ik} N_{k,t}}$	$-\alpha_{ij}$	Ricker (1954)
Power law	$N_{i,t+1} = r N_{i,t}^\alpha$	$(\alpha - 1) / N_{i,t}$	Wright et al. (2005)
Beverton-Holt	$N_{i,t+1} = r N_{i,t} / (1 + \sum_k \alpha_{ik} N_{k,t})$	$-\alpha_{ij} / (1 + \sum_k \alpha_{ik} N_{k,t})$	Beverton and Holt (1981)
Quadratic	$N_{i,t+1} = N_{i,t} (1 + r(1 - \sum_k \alpha_{ik} N_{k,t}))$	$-r\alpha_{ij} / (1 + r(1 - \sum_k \alpha_{ik} N_{k,t}))$	May (1976)
θ -Ricker	$N_{i,t+1} = N_{i,t} e^{r_i - (\sum_k \alpha_{ik} N_{k,t})^\theta}$	$-\theta\alpha_{ij} (\sum_k \alpha_{ik} N_{k,t})^{\theta-1}$	Bellows (1981)
Hassell	$N_{i,t+1} = r N_{i,t} / (1 + \sum_k \alpha_{ik} N_{k,t})^c$	$-c\alpha_{ij} / (1 + \sum_k \alpha_{ik} N_{k,t})$	Hassell (1975)

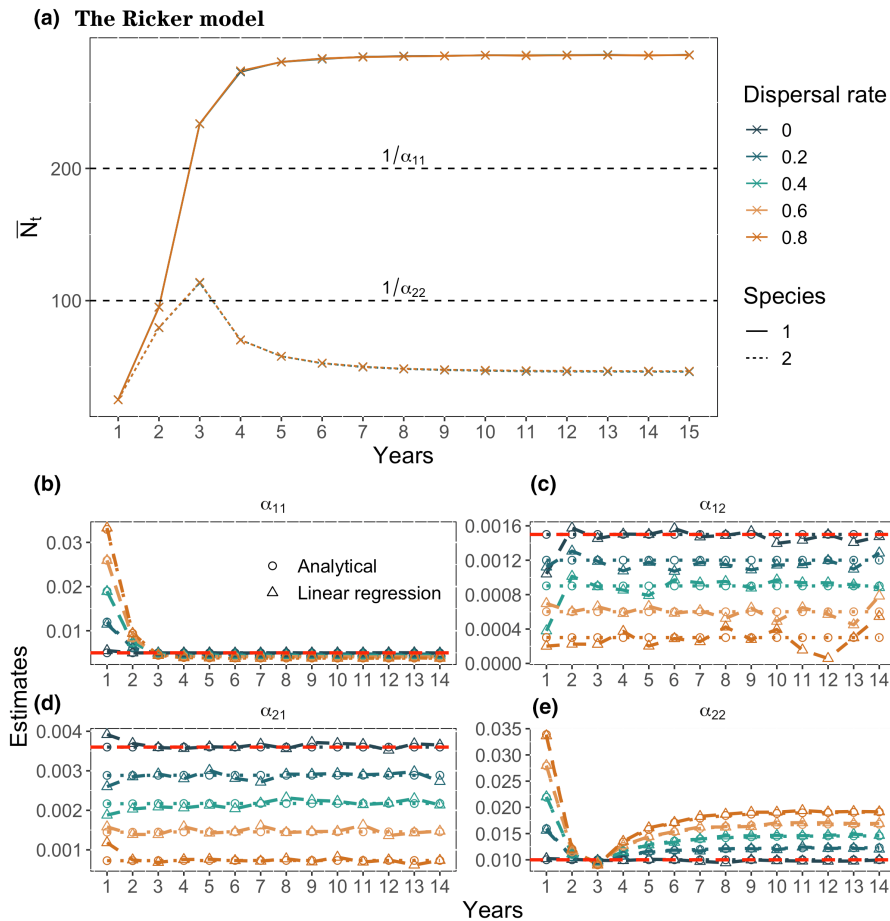


FIGURE 4 The change in bias during population development in simulations of a two-species mixture with five different dispersal rates in the Ricker model. (a) Growth curves of the two species (does not depend on dispersal rate). Dashed lines indicate the points where the bias in α_{11} and α_{22} are predicted to change sign. (b–e) Estimates of competition coefficients using the linear regression method (dashed lines) and the analytical bias formula (dotted lines) described in the main text. Red dashed lines denote the true values. In panel (c) the high variability from the linear regression method is due to very low spatial variation when dispersal is high, making estimation difficult. The interaction matrix is given by $\alpha_{11} = 0.005$, $\alpha_{12} = 0.0015$, $\alpha_{21} = 0.0036$, $\alpha_{22} = 0.01$ and the growth rate (r) for both species is 1.5.

affected by the precise way that seeds are distributed across the remaining cells:

$$\hat{\alpha}_{ij} - \alpha_{ij} \approx (1 - \delta) \left(\frac{1}{N_{it}} - \alpha_{ij} \right), \quad (14)$$

$$\hat{\alpha}_{ij} - \alpha_{ij} \approx (\delta - 1) \alpha_{ij}. \quad (15)$$

The bias formula can, therefore, be applied as long as the stay rate can be empirically estimated and does not depend on being able to quantify the details of the dispersal kernel.

3.2 | Power-law and Beverton–Holt

Under uniform dispersal for the power-law model, simulations of monocultures with different stay rates revealed that the extent of the bias increases as the stay rate declines and more seeds disperse beyond their natal cells (Figure 5a). The analytical formula accurately captured the extent and nature of this bias (Figure 5a). When seeds dispersed according to a Gaussian distribution the performance of the bias formula was unaffected (Figure S7).

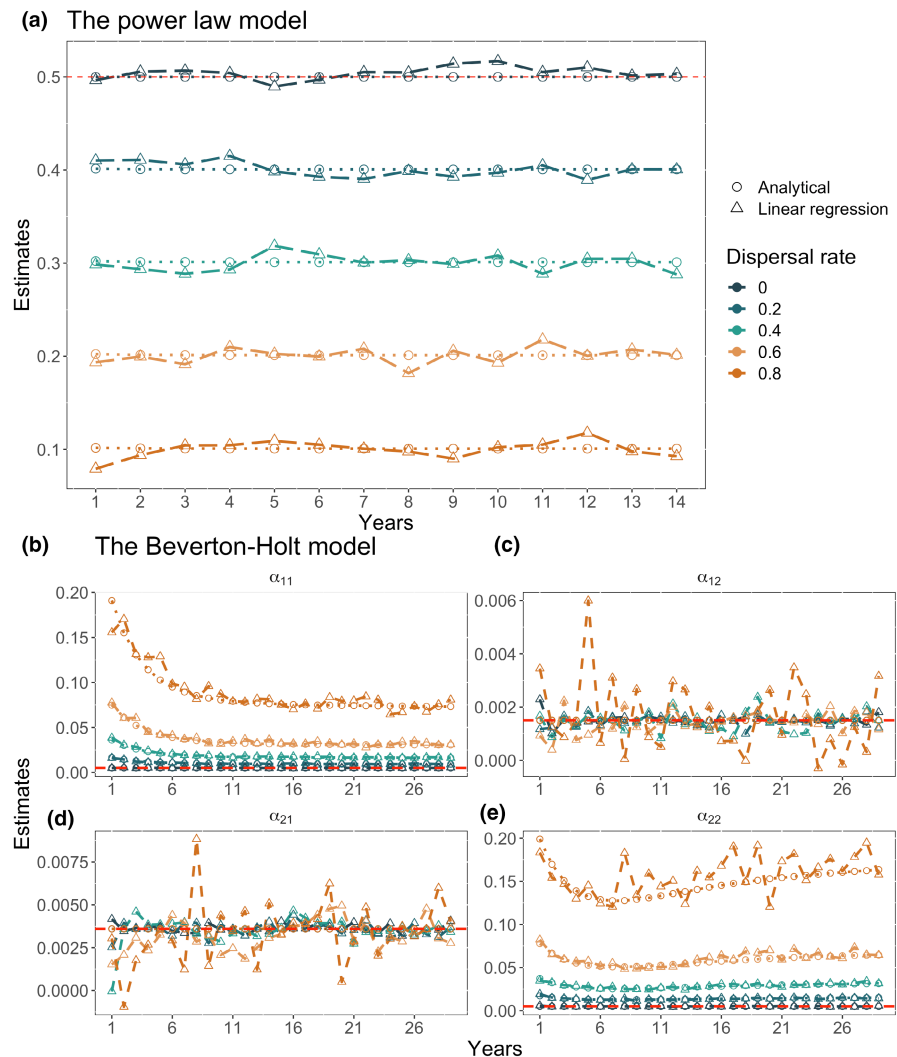
Under uniform dispersal for the Beverton–Holt model, the simulations of two-species communities verified that the strength of conspecific competition was overestimated, but

the estimates of heterospecific competition fluctuated slightly around the true values, although never in a consistent direction. When the stay rate was low (and the dispersal rate high) there were larger fluctuations in the estimates of heterospecific competition because there was little variation in density among cells, making it difficult to fit the regression model (Figure 5b). As with the previous two models, when seeds dispersed according to a Gaussian distribution the performance of the bias formula was unaffected (Figure S8).

4 | DISCUSSION

We have demonstrated that estimates of the strength of competition within spatially structured communities of sessile organisms are generally biased if dispersal is not accounted for when using space-for-time substitution. The extent and nature of the bias depends on the fitted model. The analytical bias formulae that we derive for three widely used population models show that the bias can depend on both population size and the identity of competitors (whether con- or heterospecific). These differences stem from different definitions of competitive interactions that are used in different models, but by providing a general metric of competition, we show how bias formulae can be derived for a broad range of potential models.

FIGURE 5 (a) Estimates of the conspecific competition coefficient in the power-law model using the bias formula (Equation 9) and the linear regression method for five different dispersal rates. Parameters: $r = 10, \alpha = 0.5$. (b–e) Estimates of competition coefficients in the Beverton–Holt model in a two-species mixture using the bias formula (Equations 11 and 12) and the linear regression method. Parameters: $r = 1.5, \alpha_{11} = 0.005, \alpha_{12} = 0.0015, \alpha_{21} = 0.0036, \alpha_{22} = 0.01$. Red dashed lines denote the true values.



4.1 | Unifying bias corrections across models

From a mathematical perspective, the difference in the extent and nature of the observed bias that we see in the three population models is due to different definitions and interpretations of the competition metrics (Hulsmann et al., 2021). In the power-law model, the strength of conspecific competition (α_{ij}) is the slope of the relationship between $\log(N_{i,t+1})$ and $\log(N_{i,t})$, while in the Ricker model, the strength of conspecific competition (α_{ij}) is the slope of the relationship between $\log(N_{i,t+1}/N_{i,t})$ and $N_{i,t}$. Finally, in the Beverton–Holt model, the same coefficient has an even more complex definition (see Supplementary Material). While this variation is rarely discussed, it could be the reason why we see so much variation in the bias formulae for different population models.

To test this idea, we can use the generalized competition metric (Equation 4) to interconvert the competition coefficients defined across different models and hence compare them directly. For example, applying the general metric to both the power law model and the Ricker model means that we can directly relate the competition coefficient in the Ricker model (left hand side) to the power-law model (right hand side):

$$\alpha_{ii,R} = -(\alpha_{ii,PL} - 1) / \bar{N}_{i,t}. \quad (16)$$

Furthermore, we can convert the bias formula of the Ricker model (Equation 5) to the bias formula of the power law model (Equation 9) simply by replacing $\alpha_{ii,R}$ in Equation 5 with $-(\alpha_{ii,PL} - 1) / \bar{N}_{i,t}$. This derivation demonstrates that the difference in the behaviour of the bias among population models stems from the different definitions of conspecific competition. These differences also explain why the bias in the heterospecific competition coefficients varies among models (see Supplementary Material for mathematical details).

More broadly, the general competition coefficient can also be applied to a broad range of other population models, for example, the Skellam model (Skellam, 1991), the quadratic model (May, 1976), the Hassell model (Hassell, 1975) and many others (Brännström & Sumpter, 2005; see Table 2).

4.2 | Empirical bias correction

Our analysis reveals that knowledge of the stay rate (δ) is key to correcting biases that arise because of spatial structure in meta-communities. However, stay rates (or dispersal rates) are difficult

to measure in practice (see Bullock et al., 2006) for a guide to measuring dispersal) although it may be of some comfort to empiricists, that estimating precise dispersal kernels is not necessary. If estimation of the stay rate is impossible, then might also possible to take advantage of the behaviour of the bias formula in Equation 5, if the Ricker model is thought to provide an appropriate description of local competition. As the population develops, we know that the bias will switch direction (from positive to negative or vice versa) as it crosses the threshold $1/\alpha_{ij}$. So, if we have data from consecutive years, and can estimate $1/\bar{N}_{ij}$ directly and $\hat{\alpha}_{ij}$ by using the regression method, then we can determine at which point the threshold is crossed. Using data collected from observations around this switching point (i.e. where $1/\bar{N}_{ij} = \alpha_{ij}$) it is, therefore, possible to estimate the intraspecific competition coefficient with minimal bias (see [Supplementary Material](#) for details and [Figure S8](#)). Although it may be unlikely that such data would be available in a natural field setting, this method could potentially be applied under more controlled conditions, for example, where it is possible to directly manipulate abundances or where large amounts of data from natural settings are available.

4.3 | Effects of bias in published studies

Our finding that estimation bias can be affected by population size has widespread ramifications. For example, several authors have investigated the relationship between the strength of conspecific competition and species abundances (Comita et al., 2010; LaManna et al., 2016; Yenni et al., 2017). In a study of 180 tropical tree species on Barro Colorado Island, Panama (Comita et al., 2010), the authors found that rare species suffered greater conspecific competition than common species. However, our results suggest that the strength of conspecific competition is more likely to be overestimated for rare species than for common species when the bias induced by dispersal is not properly accounted for. We, therefore, suggest that including estimates of dispersal ability in such analyses would be a useful next step, which might come from seed traps or other methods (Muller-Landau et al., 2008).

Estimation bias may also affect the robustness of other patterns that have been found in natural systems. For example, LaManna et al. (2016) used an observational dataset of 38 woody-plant species in a temperate forest to show that species experience more intense conspecific competition in resource-rich environments than in resource-poor environments (LaManna et al., 2016), while a similar study of more than 3000 species across 24 forest plots worldwide found that rare species suffer greater conspecific competition in tropical latitudes than in temperate latitudes (LaManna, Mangan, et al., 2017). These patterns were also uncovered by fitting local population models which did not account for dispersal, and hence might also be due to differences in dispersal or typical population sizes in the environments compared. Indeed, any studies that relate the strength of species interactions to biotic or abiotic gradients need to be cautious that the nature and extent of dispersal may differ across such

gradients and hence drive spurious correlations (Comita et al., 2010; Harms et al., 2000; LaManna et al., 2016, 2018; LaManna, Mangan, et al., 2017; Stump & Comita, 2018).

5 | CONCLUSIONS

Fitting local population models to data by employing space-for-time substitution in communities of sessile organisms is likely to cause systematic bias in the estimated strength of competition. This type of bias is independent from other sources of bias recently identified, such as sampling error (Detto et al., 2019) and omitted variables (Rinella et al., 2020). Our analytical bias formulae suggest that an estimate of the fraction of seeds that disperse away from the local cell is required to correct for bias when estimating the strength of competition but not the full details of dispersal kernels.

AUTHOR CONTRIBUTIONS

Adam T. Clark, Mark Rees and Lindsay A. Turnbull conceived the idea for the study. Liang Xu, Adam T. Clark, Mark Rees and Lindsay A. Turnbull designed the model, Liang Xu developed the code and performed simulations. Liang Xu derived the analytical results and Liang Xu, Adam T. Clark, Mark Rees and Lindsay A. Turnbull analysed the results. Liang Xu wrote the first draft of the manuscript, while Liang Xu, Adam T. Clark, Mark Rees and Lindsay A. Turnbull contributed to revisions.

ACKNOWLEDGEMENTS

We thank the John Fell Fund grant 0008262 for financial support. We thank Axel Rossberg, Peter Adler, Steve Ellner for the initial discussion. This research was funded by John Fell Fund grant 0008262 from the University of Oxford.

CONFLICT OF INTEREST

No conflict of interest is involved in this study.

PEER REVIEW

The peer review history for this article is available at <https://publons.com/publon/10.1111/2041-210X.14022>.

DATA AVAILABILITY STATEMENT

The simulation code and the plotting function code are available in the Github repository: <https://github.com/xl0418/PlantSim/tree/v1.0.0>, and are archived on Zenodo: <https://doi.org/10.5281/zenodo.7250630> (Liang, 2022). The data used in this study can be quickly generated using the code scripts.

ORCID

Liang Xu  <https://orcid.org/0000-0002-5528-1298>

Adam T. Clark  <https://orcid.org/0000-0002-8843-3278>

Mark Rees  <https://orcid.org/0000-0001-8513-9906>

Lindsay A. Turnbull  <https://orcid.org/0000-0002-8035-2879>

REFERENCES

- Adler, P. B., Smull, D., Beard, K. H., Choi, R. T., Furniss, T., Kulmatiski, A., Meiners, J. M., Tredennick, A. T., & Veblen, K. E. (2018). Competition and coexistence in plant communities: intraspecific competition is stronger than interspecific competition. *Ecology Letters*, 21, 1319–1329.
- Bellows, T. S. (1981). The descriptive properties of some models for density dependence. *Journal of Animal Ecology*, 50, 139–156.
- Beverton, R. J. H., & Holt, S. J. (1981). *On the dynamics of exploited fish populations*. Springer.
- Brännström, Å., & Sumpter, D. J. T. (2005). The role of competition and clustering in population dynamics. *Proceedings of the Royal Society B: Biological Sciences*, 272, 2065–2072.
- Bullock, J. M., Shea, K., & Skarpaas, O. (2006). Measuring plant dispersal: An introduction to field methods and experimental design. *Plant Ecology*, 186, 217–234.
- Chesson, P. (2000). Mechanisms of maintenance of species diversity. *Annual Review of Ecology, Evolution, and Systematics*, 31, 343–366.
- Chisholm, R. A., & Fung, T. (2018). Comment on “Plant diversity increases with the strength of negative density dependence at the global scale”. *Science*, 360, eaar4685.
- Clark, F., Brook, B. W., Delean, S., Reşit Akçakaya, H., & Bradshaw, C. J. A. (2010). The theta-logistic is unreliable for modelling most census data. *Methods in Ecology and Evolution*, 1, 253–262.
- Comita, L. S., Muller-Landau, H. C., Aguilar, S., & Hubbell, S. P. (2010). Asymmetric density dependence shapes species abundances in a tropical tree community. *Science*, 329, 330–332.
- Connell, J. H. (1971). On the role of natural enemies in preventing competitive exclusion in some marine animals and in rain forest trees. In *Dynamics of populations*. Centre for Agricultural Publishing and Documentations.
- Dennis, B., & Taper, M. L. (1994). Density dependence in time series observations of natural populations: Estimation and testing. *Ecological Monographs*, 64, 205–224.
- Detto, M., Visser, M. D., Wright, S. J., & Pacala, S. W. (2019). Bias in the detection of negative density dependence in plant communities. *Ecology Letters*, 22, 1923–1939.
- Freckleton, R. P., Watkinson, A. R., Green, R. E., & Sutherland, W. J. (2006). Census error and the detection of density dependence. *The Journal of Animal Ecology*, 75, 837–851.
- Harms, K. E., Wright, S. J., Calderón, O., Hernández, A., & Herre, E. A. (2000). Pervasive density-dependent recruitment enhances seedling diversity in a tropical forest. *Nature*, 404, 493–495.
- Hart, S. P., Freckleton, R. P., & Levine, J. M. (2018). How to quantify competitive ability. *Journal of Ecology*, 106, 1902–1909.
- Hassell, M. P. (1975). Density-dependence in single-species populations. *The Journal of Animal Ecology*, 44, 283.
- Hulsmann, L., Chisholm, R. A., & Hartig, F. (2021). Is variation in conspecific negative density dependence driving tree diversity patterns at large scales? *Trends in Ecology & Evolution*, 36, 151–163.
- Johnson, D. J., Beaulieu, W. T., Bever, J. D., & Clay, K. (2012). Conspecific negative density dependence and forest diversity. *Science*, 336, 904–907.
- Kellner, J. R., & Hubbell, S. P. (2018). Density-dependent adult recruitment in a low-density tropical tree. *Proceedings of the National Academy of Sciences of the United States of America*, 115, 11268–11273.
- Knape, J., & de Valpine, P. (2012). Are patterns of density dependence in the Global Population Dynamics Database driven by uncertainty about population abundance? *Ecology Letters*, 15, 17–23.
- LaManna, J. A., Belote, R. T., Burkle, L. A., Catano, C. P., & Myers, J. A. (2017). Negative density dependence mediates biodiversity-productivity relationships across scales. *Nature Ecology and Evolution*, 1, 1107–1115.
- LaManna, J. A., Mangan, S. A., Alonso, A., Bourg, N. A., Brockelman, W. Y., Bunyavejchewin, S., Chang, L. W., Chiang, J. M., Chuyong, G. B., Clay, K., Condit, R., Cordell, S., Davies, S. J., Furniss, T. J., Giardina, C. P., Gunatilleke, I. A. U. N., Gunatilleke, C. V. S., He, F., Howe, R. W., ... Myers, J. A. (2017). Plant diversity increases with the strength of negative density dependence at the global scale. *Science*, 356, 1389–1392.
- LaManna, J. A., Mangan, S. A., Alonso, A., Bourg, N. A., Brockelman, W. Y., Bunyavejchewin, S., Chang, L. W., Chiang, J. M., Chuyong, G. B., Clay, K., Cordell, S., Davies, S. J., Furniss, T. J., Giardina, C. P., Gunatilleke, I. A. U. N., Gunatilleke, C. V. S., He, F., Howe, R. W., Hubbell, S. P., ... Myers, J. A. (2018). Response to Comment on “Plant diversity increases with the strength of negative density dependence at the global scale”. *Science*, 360, eaar5245.
- LaManna, J. A., Walton, M. L., Turner, B. L., & Myers, J. A. (2016). Negative density dependence is stronger in resource-rich environments and diversifies communities when stronger for common but not rare species. *Ecology Letters*, 19, 657–667.
- Liang Xu. (2022). xI0418/PlantSim: PlantSim package (v1.0.0). Zenodo. <https://doi.org/10.5281/zenodo.7250630>
- Lotka, A. J. (1926). Elements of physical biology. *Science Progress in the Twentieth Century (1919-1933)*, 21, 341–343.
- May, R. M. (1976). Simple mathematical models with very complicated dynamics. *Nature*, 261, 459–467.
- Muller-Landau, H. C., Wright, S. J., Calderón, O., Condit, R., & Hubbell, S. P. (2008). Interspecific variation in primary seed dispersal in a tropical forest. *Journal of Ecology*, 96, 653–667.
- Myers, R. (2001). Stock and recruitment: Generalizations about maximum reproductive rate, densitydependence, and variability using meta-analytic approaches. *ICES Journal of Marine Science*, 58, 937–951.
- Petermann, J. S., Fergus, A. J. F., Turnbull, L. A., & Schmid, B. (2008). Janzen-Connell effects are widespread and strong enough to maintain diversity in grasslands. *Ecology*, 89, 2399–2406.
- Rees, M., Grubb, P. J., & Kelly, D. (1996). Quantifying the impact of competition and spatial heterogeneity on the structure and dynamics of a four-species guild of winter annuals. *The American Naturalist*, 147, 1–32.
- Ricker, W. E. (1954). Stock and Recruitment. *Journal of the Fisheries Research Board of Canada*, 11, 559–623.
- Rinella, M. J., Strong, D. J., & Vermeire, L. T. (2020). Omitted variable bias in studies of plant interactions. *Ecology*, 101, e03020.
- Skellam, J. G. (1991). Random dispersal in theoretical populations. *Bulletin of Mathematical Biology*, 53, 135–165.
- Stump, S. M., & Comita, L. S. (2018). Interspecific variation in conspecific negative density dependence can make species less likely to coexist. *Ecology Letters*, 21, 1541–1551.
- Tuck, S. L., Porter, J., Rees, M., & Turnbull, L. A. (2018). Strong responses from weakly interacting species. *Ecology Letters*, 21, 1845–1852.
- Volterra, V. (1926). Fluctuations in the abundance of a species considered mathematically. *Nature*, 118, 558–560.
- Wright, S. J., Muller-Landau, H. C., Calderón, O., & Hernández, A. (2005). Annual and spatial variation in seedfall and seedling recruitment in a neotropical forest. *Ecology*, 86, 848–860.
- Yenni, G., Adler, P. B., & Ernest, S. K. M. (2017). Do persistent rare species experience stronger negative frequency dependence than common species? *Global Ecology and Biogeography*, 26, 513–523.

SUPPORTING INFORMATION

Additional supporting information can be found online in the Supporting Information section at the end of this article.

How to cite this article: Xu, L., Clark, A. T., Rees, M., & Turnbull, L. A. (2023). Estimating competition in metacommunities: accounting for biases caused by dispersal. *Methods in Ecology and Evolution*, 14, 291–301. <https://doi.org/10.1111/2041-210X.14022>