A critical examination of models of annual-plant population dynamics and density-dependent fecundity

Daniel B. Stouffer

Centre for Integrative Ecology, School of Biological Sciences, University of Canterbury, Christchurch, New Zealand

Correspondence
Daniel B. Stouffer
Email: daniel.stouffer@canterbury.ac.nz

Funding information
Marsden Fund Council, Grant/Award Number: 16-UOC-008

Handling Editor: Robert Freckleton

Abstract

1. Mathematical models serve many purposes in biology. Each and every model is a necessary simplification of reality, and, as simplifications, these models are also wrong by definition. And yet there are many ways to be wrong, and some of these are a much greater concern than others. For example, a paradoxical model-based prediction may simply be puzzling whereas unphysical variables (e.g. negative amounts of time or temperatures below absolute zero) and non-biological variables (e.g. negative abundances or negative feeding rates) should be avoided altogether.

2. Here I analyse a discrete-time model of annual-plant population dynamics and three phenomenological models for density-dependent fecundity. These phenomenological models are generally interchanged solely on the basis of their statistical fit to data. On the other hand, I highlight ways in which their phenomenological basis hampers our ability to capture known aspects of annual-plant biology.

3. I then demonstrate how to specify a more flexible, biologically appropriate model and illustrate this model’s behaviour and interpretation in single-species and multi-species contexts. By constructing this generative model for annual-plant population dynamics, I can also demonstrate how and why emergent phenomena, such as negative density dependence, emerge.

4. Although my focus is on a model applied to annual plants, the biological implications extend to modelling any species with a discrete life cycle and nonoverlapping generations. More broadly, my exploration here showcases that there are many more criteria with which we could, and arguably should, ground-truth mathematical models across biology.

KEYWORDS
competition, facilitation, generative models, phenomenological models, plant reproduction, statistical models, structural validity, theoretical ecology
1 | INTRODUCTION

Mathematical models and theory play many valuable roles in biology: they can aid in interpretation, guide data collection, help develop novel hypotheses and permit explorations that go beyond the bounds of what is possible experimentally (Otto & Day, 2007; Venable, 1989). There are many factors which influence how scientists select a specific mathematical model to interrogate a specific biological question (Evans et al., 2013; Houlanhan et al., 2015; Marquet et al., 2014; Marquet et al., 2015; Otto & Rosales, 2020). Convenience or precedent may rule the day, scientists may borrow from a different domain, or they may develop a completely new approach to maintain fidelity to the focal system and its established idiosyncrasies (Holling, 1966; Levins, 1966; Otto & Day, 2007). Since ‘any model is at best a useful fiction’ (Box & Luceño, 1997), it can be difficult or impossible to completely rule out any candidate model: even a model of penguin movement that includes flight may continue to work well as long as the probability that penguins fly is vanishingly small.

Models of plant competition may be an especially apt example of useful ecological fictions. Competition between plants results primarily from interference between individuals that grow in close proximity to each other (Antonovics & Levin, 1980; Harper, 1977). Neighbouring plants can deplete local resource pools in the soil (Craine, 2006; Dybinski & Tilman, 2007; Tilman & Wedin, 1991), can change light availability (Adams et al., 2007; Falster & Westoby, 2003; Hautier et al., 2009), and can be more attractive to shared pollinators (Lanuza et al., 2018; Levin & Anderson, 1970). Each of these factors—and many others—can impact the performance of plant individuals. This notwithstanding, it is often far more practical to treat these intermediate factors implicitly and model the outcome of competitive effects phenomenologically (i.e. in terms of plant densities; Connell, 1990; Goldberg, 1990; Adler et al., 2018). Beyond practicalities, phenomenological approaches can also provide a more immediate connection between theory and experimental or observational data (Hart et al., 2018).

The top priority of phenomenological models is simply ‘to describe [a] relationship, with the assumption that the relationship extends past the measured values’ (Hilborn & Mangel, 1997). As such, it can be tempting to think that the only way to assess phenomenological models is via their statistical fit to data (Gass, 1983; Hilborn & Mangel, 1997; Houlanhan et al., 2015; Power, 1993). And yet this is not exactly true: phenomenological and nonphenomenological models may be based on untenable assumptions even when their mathematical form fits observed data well (Abrams, 2015; Arditi & Michalski, 1996; Geritz & Kisidi, 2004; Malard et al., 2020; Moisset de Espanés et al., 2021; Morozov & Petrovskii, 2013; Quinlan, 1986; Schnapp et al., 1991). Indeed, mathematical systems theory argues that a model should only be considered ‘structurally valid’ when it (i) reproduces the observed behaviour of a real system and (ii) accurately reflects the way in which that system operates (Klir, 1985; Zeigler et al., 2000). For example, a simple linear regression might provide a good fit to a predator-prey functional-response dataset, but not just any linear regression is biologically sensible: model-predicted consumption rates must always be greater than or equal to zero and must equal exactly zero when there are no prey available to be consumed. Although relaxing this constraint may not imperil fits to data, completely ignoring it could lead to nonsensical predictions if the inferred parameters subsequently get used in a population-dynamics model (Abrams, 2015; Broekman et al., 2019). Building on this notion of structural validity, my focus here is to examine the mathematical and biological underpinnings of phenomenological models of competition between annual plants.

Due to their short life span and the apparent simplicity of their life cycle (Symonides, 1988; Watkinson, 1981), annual-plant assemblages have garnered substantial attention for empirical and theoretical purposes alike (Ellner, 1985a, 1985b; Firbank & Watkinson, 1985; Godoy & Levine, 2014; Jarry et al., 1995; MacDonald & Watkinson, 1981; Matías et al., 2018; Pacala, 1986; Petry et al., 2018; Venable & Lawlor, 1980; Watkinson, 1980). With minimal variation, the core structure of the model most researchers use to study the population dynamics of annual plants has been around since the late 1960s (Cohen, 1966), and the model’s verbal description is delightfully brief: A local population of plant species $i$ has some density of viable seeds $N_i$ that lie dormant in the seed bank at the start of the year. A fraction $g_i$ of these seeds germinate and grow to maturity, and each of these plants has an average fecundity $F_i$ corresponding to seeds that return back into the seed bank. Of the original fraction $(1 - g_i)$ of seeds that did not germinate, only a fraction $s_i$ remains viable for future years.

Competitive effects in this annual-plant model are usually introduced phenomenologically such that per capita fecundity $F_i$ varies as a function of the densities of conspecific and heterospecific plants growing in the local environment (Hart et al., 2018; Levine & HilleRisLambers, 2009; Watkinson, 1980). I focus here on three phenomenological models of density-dependent fecundity and highlight key ways in which they are ‘structurally invalid’: they are incompatible with known facets of the biology of annual plants. I therefore also show how to develop an alternative model that has a sound mathematical basis while also maintaining the desired qualitative and quantitative model behaviour. In addition to providing a new way to model and describe annual-plant dynamics and density-dependent plant performance, my analysis provides a principled demonstration that—despite appearances to the contrary—not all phenomenological models are created equal.

2 | MATERIALS AND METHODS

2.1 | Phenomenological models of plant competition

The mathematical form of the ‘standard’ annual-plant competition model (Cohen, 1966) is

$$N_{i,t+1} = N_{i,t} - g_i N_{i,t} - (1 - s_i) N_{i,t} + g_i N_{i,t} F_{i,t}$$ (1)
which simplifies to
\[
\frac{N_{i,t+1}}{N_{i,t}} = (1 - g_i) s_i + g_i F_i.t.
\] (2)

As introduced previously, \(N_i\) is the density of viable seeds from plant species \(i\) that lie dormant in the seed bank at the start of year \(t\), \(g_i\) is seed germination rate, \(s_i\) is survival rate of nongerminated seeds and \(F_i.t\) is the realised per capita fecundity of germinated plants from species \(i\) in year \(t\). Year-to-year changes in seed populations therefore occur via one of two pathways (Figure 1a). First, the \((1 - g_i)N_i\) seeds that do not germinate survive with probability \(s_i\) and remain in the seed bank for subsequent generations. Second, the \(g_iN_i\) seeds that do germinate produce new seeds as determined by their realised fecundity \(F_i.t\).

There is an extremely large collection of phenomenological models which have been proposed over the years in order to describe density-dependent fecundity \(F_i.t\) (see, e.g. Table 1 in Law & Watkinson, 1987; Supplementary Table 2 in Levine & HilleRisLambers, 2009; or Appendix Table A1 in Hart et al., 2018). Here I concentrate on three of these models because their downsides are sufficiently representative of common issues that can arise when employing such phenomenological models. The first, and mathematically simplest, model is a linear Lotka–Volterra-like model given by
\[
F_i.t = \lambda_i \left( 1 - \sum_j a_{ij} g_j N_j.t \right)
\] (3)

where \(\lambda_i\) is intrinsic fecundity (i.e. seed production in the absence of density-dependent interaction effects) and \(a_{ij}\) captures the strength of an interaction between a neighbouring species \(j\) on the fecundity of each seed-producing individual of species \(i\). Values of \(a_{ij} > 0\) imply competitive effects and values of \(a_{ij} < 0\) imply facilitative effects. These interaction coefficients are multiplied by \(g_jN_j\) because the model-predicted density of plants in year \(t\) is given by the product of the density of viable seeds and those seeds’ germination rate. The second model is the reciprocal relationship (Watkinson, 1980; Weiner, 1982) given mathematically by
\[
F_i.t = \frac{\lambda_i}{1 + \sum_j a_{ij} g_j N_j.t}
\] (4)

**FIGURE 1** Annual-plant population dynamics of a single species. (a) The solid line and arrows show a conceptualisation of seed bank population dynamics in discrete time. Every year, seeds in the seed bank decrease because some proportion germinate \((g)\) and a fraction of the remainder die \(((1 - g)(1 - s))\). The seed bank is replenished by seed production derived from the seeds that germinated \((sF)\). For simplicity, this figure indicates that germination and seed mortality occur prior to the growing season and all seeds enter into the seed bank simultaneously; in reality, these processes will be more homogeneously spread across time. The discrete-time model (Equation 2) merely aims to capture their appropriate average annual rates. (b, c and d) The dashed lines indicate equivalent predictions made using a population-dynamics model within which seed germination, seed mortality, plant mortality and plant biomass growth are modelled as continuous-time processes governed by Equations 6–8. During the first period of the year of duration \(T_1 = 0.5\), seed germination rate \(g_i = 1.50\), seed mortality rate \(\mu_i = 0.75\), plant mortality rate \(\nu_i = 0\), germinant biomass \(\beta_i = 0.02\) and biomass growth rate \(r_i = 0\). During the second period of the year of duration \(T_2 = 0.5\), seed germination rate \(g_i = 0\), seed mortality rate \(\mu_i = 0.75\), plant mortality rate \(\nu_i = 0.10\), germinant biomass \(\beta_i = 0.02\), biomass growth rate \(r_i = 12\) and biomass carrying capacity \(K_i = 100\). At the end of the growing season, the conversion of biomass to seeds occurs at a rate of \(\phi_i = 1.00\).
where all parameters maintain the same underlying meaning. For reference, this reciprocal form is also referred to as the Beverton–Holt model owing to a similar model used to describe population growth in fisheries (Beverton & Holt, 1957). The third model is the Ricker model (Ricker, 1954a, 1954b), given by

\[ F_{ij} = \lambda_i e^{-\sum_j a_{ij} N_{ij}}, \]

where all parameters again maintain the same underlying meaning. These three models are also linked mathematically: Equation 3 is the first-order Taylor series approximation of both Equations 4 and 5 when the sum \( \sum_j a_{ij} N_{ij} \) is small.

When combined with empirical data, all three of these models can capture the foundational pattern of negative density dependence that occurs between co-occurring plant species. For reasons that I will elaborate on in the following section, the linear model (Equation 3) is rarely used in practice; I include it here predominantly as a third basis for comparison. In recent years, both the reciprocal model (Equation 4) and Ricker model (Equation 5)—or modifications thereof—have been adopted in studies that statistically infer the strength of interactions between co-occurring annual plants (e.g. Bimler et al., 2018; Godoy et al., 2014; Godoy & Levine, 2014; Hart et al., 2018; Mathias et al., 2018; Mayfield & Stouffer, 2017; Stouffer et al., 2018). However, the reciprocal model is by far the most widely used. It was originally derived for self-thinning single-species plant populations based on the assumption that, as the number of individual plants in a stand increases, their total yield does not increase beyond a certain limit (Holliday, 1960; Kira et al., 1953; Shinozaki & Kira, 1956). To the best of my knowledge, Watkinson (1980) was the first to propose how total yield could be converted into total seed production and incorporated into an annual-plant population-dynamics model. In contrast, the Ricker model was originally developed in a fisheries context (Ricker, 1954a, 1954b) although it also appears in plant-focused studies that compare fits across large sets of candidate models (e.g. Law & Watkinson, 1987; Levine & HilleRisLambers, 2009). Its application to studying density-dependent fecundity likely also relates to the growing use of generalised linear models (GLMs) across ecology. Fecundities are discrete counts, and the most common statistical approach for analysing such response variables is via Poisson or negative binomial GLMs (O’Hara & Kotze, 2010; Rao & Toutenberg, 2010). Both of these use a log link function, which implicitly imposes a Ricker (or Ricker-like) structure on the model being fit (Mayfield & Stouffer, 2017; Stouffer et al., 2018).

### 2.2 Issues with these phenomenological models

If my own personal experience is any indication, many researchers have not selected a particular phenomenological model based on how or why that model was originally derived. Instead, they have tended to choose on the basis of three primary criteria (with the first two heavily outweighing the third): (a) whether or not the model is widely used across the prevailing literature, (b) the model’s statistical fit to observed data and (c) the model’s ability to accurately capture the phenomena under study. Regarding the less frequently evoked last reason, two main questions researchers have focused on when selecting between models are ‘How likely is the model to make a nonbiological prediction?’ and ‘Does the model work for all potentially observable interaction types?’ The answers to both of these intersect in different but important ways across the aforementioned three phenomenological models.

A substantial downside to the linear model is that it can predict nonbiological negative fecundities if competitive effects are too strong (i.e. \( \sum_j a_{ij} N_{ij} \) is too large and positive). As long as all interactions are competitive (i.e. \( a_{ij} > 0 \)), neither the Ricker model nor the reciprocal model form exhibit this ‘flaw.’ However, the reciprocal model has the peculiar behaviour that it can predict negative fecundities if enough interactions are facilitative (i.e. \( a_{ij} < 0 \)) and the corresponding

| Parameter | Meaning | Dimensions |
|-----------|---------|------------|
| \( N_{ij} \) | Seed density of species \( i \) at time \( t \) | seeds length\(^{-2} \) |
| \( p_{ij} \) | Plant density of species \( i \) at time \( t \) | plants length\(^{-2} \) |
| \( B_{ij} \) | Biomass density of species \( i \) at time \( t \) | mass length\(^{-2} \) |
| \( T_1 \) | Duration of nongrowing season | time |
| \( T_2 \) | Duration of growing season | time |
| \( \gamma_i \) | Germination rate of species \( i \) | 1 time \(^{-1} \) |
| \( \mu_i \) | Seed mortality rate of species \( i \) | 1 time \(^{-1} \) |
| \( \nu_i \) | Plant mortality rate of species \( i \) | 1 time \(^{-1} \) |
| \( r_i \) | Biomass growth rate of species \( i \) | 1 time \(^{-1} \) |
| \( K_i \) | Biomass carrying capacity of species \( i \) | mass length\(^{-2} \) |
| \( \beta_i \) | Biomass of early germinants of species \( i \) | mass plant |
| \( \phi_i \) | Biomass to seed conversion rate of species \( i \) | mass seeds |
| \( a_{ij} \) | Effect of species \( j \) on biomass growth rate of species \( i \) | Dimensionless |
neighbours are sufficiently abundant such that $\sum a_{ij}N_{ij} < -1$. Due to interactions appearing in the exponential, the Ricker model overcomes this drawback of the linear and reciprocal models since it can never predict negative fecundities regardless of how large $\sum a_{ij}N_{ij}$ becomes. On the other hand, both the Ricker and linear models suffer from unchecked and unbounded increases in predicted fecundities when neighbour densities increase and interactions are facilitative (i.e. $a_q < 0$). Even in situations where facilitative effects make ecological sense at low neighbour densities, as an individual plant’s local neighbourhood becomes overwhelmingly dense (i.e. as $\sum N_{ij} \to \infty$) any structurally valid model for per capita fecundity should either (i) decrease towards 0 or (ii) approach an asymptotic value. None of Equations 3–5 can accomplish this; this leaves the phenomenological model in search of an alternative mathematical form that could, and existing attempts generally come with their own additional caveats or complexities in terms of additional parameter constraints (e.g. Martorell & Freckleton, 2014; Mayfield & Stouffer, 2017).

Sticking to a phenomenological perspective can also complicate researchers’ ability to include additional covariates into models of density-dependent fecundity. For example, incorporating changing environmental conditions (Bimler et al., 2018; Matías et al., 2018) or interactions with nonplant species like pollinators (Lanuza et al., 2018) could ostensibly require ‘sub-models’ for any or all of $g_i$, $s_i$, $\lambda_i$ and $a_{iq}$ across Equations 1–5. There are no clear rules of thumb for what these sub-models could or should look like mathematically and how the parameters might also need to be constrained. Phenological differences between plant species are also an important predictor of competitive outcomes in annual-plant communities (Godoy & Levine, 2014). In nature, species germinate and senesce at different times. Plants that grow for longer can potentially access more resources allowing them to accumulate more biomass and produce more seeds. It is equally unclear, however, whether and how phenological differences should best be translated into differences in the different parameters of purely phenomenological models.

Conceptually, the above phenomenological models— and any alterations thereof—strive to capture the outcome of the myriad biological variables and processes thought to impact the production of seeds post-germination. The aforementioned issues often arise, however, because of constraints imposed by the models’ different mathematical forms. Rather than specify or select a model on the basis the final outcome (i.e. seed production), I argue here that a useful alternative approach is to study generative models (Geritz & Kisdi, 2004), the solutions to which can then provide ‘phenomenological’ predictions (Hagihara, 1999; Quintian, 1986; Schnapp et al., 1991). In what follows, I will show what one such model could look like and describe how it helps overcome some of the downsides of the fully phenomenological approaches.

### 2.3 Single-species generative seed-production models

For simplicity, I will initially show how to develop a generative model for single-species monocultures before extending the same arguments to multi-species polycultures. Since seed production is actually a process that occurs and accumulates over a period of time, I will specifically consider a continuous-time generative model.

The first step to formalise a generative process for the annual-plant competition model is to define continuous-time models for the two previously introduced state variables: the density of seeds $N_i$ and the density of plants $p_i$. This creates an immediate contrast to the phenomenological treatment from earlier that focused on choosing between alternative models for per capita fecundity $F_{it}$ while assuming that germination and seed survival (i.e. $(1 - g_i)s_i$) could be left alone. At any given time during the year, we can describe the dynamics of the two state variables as

$$\frac{dN_i}{dt} = -\gamma_i N_i - \mu_i N_i$$

and

$$\frac{dp_i}{dt} = \gamma_i N_i - \nu_i p_i,$$

where $\gamma_i$ is the continuous-time germination rate of seeds of species $i$, $\mu_i$ is the continuous-time mortality rate of seeds of species $i$, and $\nu_i$ is the continuous-time mortality rate of plants of species $i$. Like in the discrete-time model, the density of viable seeds in the seed bank can either germinate or undergo mortality. Unlike in the discrete-time model, the density of plants at any given moment of time will be a function of the seeds that germinate and plant losses due to mortality. Making time explicit and continuous opens up the possibility that germination and mortality occur at fixed or varying rates throughout the year.

In order to also make a prediction about the number of new seeds that get produced, we need at least one further state variable and hence one further equation. This is because fecundity is not simply a deterministic product of the density of plants but also depends on those plants’ condition, access to resources, etc. Rather than explicitly model newly produced seeds (i.e. $F_{ij}$), I will instead argue that it is more appropriate to define a continuous-time model of total biomass density $B_i$, of species $i$, for multiple reasons. First, biomass is the most widespread currency for ecological population-dynamics models. Second, it vastly simplifies the decision of the model’s initial condition: early germinants undeniably have measurable biomass, but they do not yet have measurable seeds (in the conventional sense). On this basis, we can describe the dynamics of $B_i$ as

$$\frac{dB_i}{dt} = B_i \left(1 - \frac{B_i}{K_i}\right) + \beta_i \gamma_i N_i - \nu_i p_i \frac{B_i}{p_i}.$$

where $r_i$ is the intrinsic growth rate of species $i$, $K_i$ is the local carrying capacity of species $i$ (something that I will elaborate on below) and $\beta_i$ is the biomass of a germinant of species $i$. The three main components of Equation 8 therefore capture biomass growth, biomass addition via the appearance of new germinants and biomass loss via the mortality of maturing plants. Since this model does not track biomass growth of
each plant separately, note that the mortality term reflects that loss of a single plant is equivalent to losing the biomass of an average plant (i.e. $B_i/p_i$). Plant mortality also introduces an Allee effect for biomass growth whenever $v_i > |r_i(1 - B_i/K_i) + \beta_i N_i/B_i|$.

### 2.4 Multi-species generative seed-production models

Building off the single-species case, we can describe the continuous-time dynamics for $k$ species in a multi-species community with a related set of $3 \times k$ continuous-time models of the form

\[
\frac{dN_i}{dt} = -\gamma_i N_i - \mu_i N_i, \tag{9}
\]

\[
\frac{dp_i}{dt} = r_i N_i - v_i p_i, \tag{10}
\]

\[
\frac{dB_{ij}}{dt} = B_ir_i \left(1 - \frac{(B_i + \sum_{j \neq i} a_{ij} B_j)}{K_i}\right) + \beta_{ij} N_i - v_{pi} \frac{B_i}{p_i}, \tag{11}
\]

where the new parameters $a_{ij}$ capture interspecific interactions. When $a_{ij} > 0$, neighbouring heterospecific plants from species $j$ exert net competitive effects that reduce biomass growth rates, and when $a_{ij} < 0$ neighbouring heterospecific plants from species $j$ exert net facilitative effects that increase biomass growth rates.

### 2.5 Generating model predictions

Given an initial density of seeds, an initial density of plants and an initial biomass density, the solutions of Equations 6–8 allow us to make single-species continuous-time dynamic predictions that are analogous to what the different phases of the discrete-time model attempt to capture (Figure 1). Given all three initial densities across all $k$ species in Equations 9–11, one can achieve the same for multi-species continuous-time dynamic predictions. As discussed below in the Results, it is possible to derive analytical solutions for some limiting cases in the single-species context. In general, however, explorations of many single-species and almost all ecologically interesting multi-species scenarios with these models will need to be simulated numerically as routinely occurs in various other ecological contexts (Rosenbaum & Rall, 2018; Stouffer & Novak, 2021).

As in the original discrete-time models described by Cohen (1966) and Watkinson (1980), I will maintain the notion here that population dynamics within each year can be separated into two distinct phases. During the first phase, biomass growth is minimal and the key dynamics relate to seed germination and seed mortality. Although both germination and seed mortality may continue throughout the year, the dynamics of the second phase are instead dominated by plant biomass growth and production of new seeds. Mathematically, generating predictions for a single species across a full year works as follows.

At the very start of the year, there are viable seeds in the seed bank ($N_{i,0} > 0$) but no plants nor plant biomass ($p_{i,0} = 0$ and $B_{i,0} = 0$). If we know all model parameters during the first phase ($\theta_{T_1}$), then we can solve Equations 6 and 7 to determine the density of seeds $N_{i,T_1}$ and density of plants $p_{i,T_1}$ at the end of the first phase (i.e. at $t = T_1$). Since plants germinate but do not acquire additional biomass during the first phase, $B_{i,T_1} = \beta_i p_{i,T_1}$; this corresponds to the biomass of each germinant multiplied by the density of germinants. If we know all model parameters during the second phase ($\theta_{T_2}$), then we can repeat the process with the new initial conditions of $N_{i,T_1}, p_{i,T_1}$ and $B_{i,T_1}$ and solve Equations 6–8 to determine the density of seeds $N_{i,T_2}$, density of plants $p_{i,T_2}$, and total biomass density $B_{i,T_2}$ at the end of the vegetative growth phase (i.e. at $t = T_2$).

To mirror what occurs in the discrete-time model, three more things occur at the very end of the plant growth phase. First, the final total biomass $B_{i,T_2}$ is converted to seeds at a rate $\phi_i$; this implies that a total of $\phi_i B_{i,T_2}$ new seeds are added to the seed bank. This mirrors the equivalent process proposed by Watkinson (1980). The initial seed density at the start of the following year is therefore $N_{i,1} = N_{i,T_2} + \phi_i B_{i,T_2}$. Second, all plants die and hence the initial plant density at the start of the following year is $p_{i,1} = 0$. Third, at senescence those plants’ biomass is also lost to the system; therefore, the initial biomass density at the start of the following year is $B_{i,1} = 0$.

The R script `simulate.two.species.model.R` available from the associated code repository provides a two-species example of this process. For simplicity, that code and the examples highlighted below in the Results assume identical phenologies for all species. In the general multi-species context, there is no requirement that the first and second phases occur at the same time across all species. In principle, each species can exhibit a unique phenology in terms of the start, end and duration of the different phases and their species-specific parameter values within each.

### 2.6 Key features of the single-species and multi-species generative models

The carrying capacity $K_i$ is best understood as a population-wide constraint on total biomass density. It therefore acts as a combination of resource limitation and any self-limitation that might still occur due to physical/physiological constraints on size without any effect of neighbours (Rees et al., 2010). This can be seen in particular by considering what occurs in the model if plant density $p_{i,T_2}$ corresponded to exactly one plant. The biomass growth or size of this single plant should still eventually be limited, even when locally available resources are infinite, and the parameter $K_i$ will indeed retard that single plant’s growth. When multiple plants are present, $K_i$ caps each individual plant’s maximum possible size and ostensibly includes locally available resources that contribute to these maxima. In contrast to other ecological contexts where it is possible to separate intraspecific effects into self-effects and effects of conspecifics (e.g.
Cervantes-Loreto et al., 2021; Skalski & Gilliam, 2001; Stouffer & Novak, 2021), differentiating between self-limitation and intraspecific competition would require a more complex model than the one described here.

Related to the interpretation of carrying capacity $K_i$, intraspecific effects in the model always act to limit biomass growth (i.e. are net competitive) as is consistent with a majority of empirical evidence (Adler et al., 2018). When interspecific coefficients $a_{ij}$ are facilitative in the multi-species model, these can lead to increased instantaneous biomass growth rates. However, when combined with intraspecific effects, these facilitative interactions will always attenuate over time as any population’s biomass $B_i$ approaches, or exceeds, its carrying capacity $K_i$. The model therefore overcomes the aforementioned issues of ‘runaway facilitation’ that arise in various phenomenological models.

Note that both the single-species and multi-species models defined above assume that only biomass growth rates are density dependent; however, plant mortality rates $v_i$ might also depend on local neighbourhoods (e.g. Chu & Adler, 2015; Watkinson, 1980). For zoo- philic plants, the number of seeds produced and hence the parameter $\phi_i$ might depend on both biomass growth and pollinator visitation rates (e.g. Lanuza et al., 2018). Moreover, germination rates $g_i$ and biomass growth rates $r_i$ might both depend on previous years’ local neighbourhoods, for example due to plant–soil feedbacks (Brandt et al., 2013; Teste et al., 2017). The ‘semi-mechanistic’ nature of the generative model arguably makes the question of where in the model one could incorporate such added biological processes more accessible. I will nevertheless leave the inclusion of these and other potential factors for future studies.

3 | RESULTS

3.1 | Model predictions after the seed germination and seed mortality phase

As noted above, we can assume that there are viable seeds in the seed bank ($N_{i,0} > 0$) at the start of each year but neither plants nor plant biomass ($p_{i,0} = 0$ and $B_{i,0} = 0$). During the first phase, plant growth is also considered to be negligible; this implies that $r_i \approx 0$, and also eliminates all potential interspecific interactions from the multi-species model as currently specified. Importantly, this means that the predictions at the end of this phase will be identical for the single- and multi-species cases. It also represents a situation for which an analytical solution is possible. If this phase lasts for a time period of duration $T_1$, we will end up with

$$N_{i,T_1} = e^{-(r_i+v_i)T_1}N_{i,0},$$

$$p_{i,T_1} = \frac{\gamma_i e^{-\gamma_i T_1}(1-e^{-(r_i+v_i)T_1})}{\gamma_i + v_i - v_i}N_{i,0},$$

and

$$B_{i,T_1} = \beta_i p_{i,T_1},$$

for viable seeds remaining in the seed bank, plants and plant biomass respectively. In a multi-species case, equivalent solutions for other species $j$ can be obtained by swapping all $i$’s for $j$’s. Intuitively, Equation 12 indicates that viable seeds in the seed bank will decrease with increasing germination or increasing seed mortality rates. Equations 13 and 14 indicate that the density of plants and those plants’ biomass density will increase with increasing germination rates or decrease with increasing seed mortality or increasing plant mortality rates.

3.2 | Model predictions after the plant growth and seed production phase

Given the densities at the end of the first phase ($N_{i,T_1}, p_{i,T_1}$ and $B_{i,T_1}$ from Equations 12–14), we can next determine the model predictions at the end of the second ‘plant growth/seed production’ phase. Here, I will first highlight predictions under two single-species scenarios for which explicit analytical solutions can be obtained. I will then move on to one specific single-species scenario and the general multi-species scenario that must be solved numerically.

3.2.1 | Single-species predictions

Scenario 1: In first instance, we will assume that seeds do not germinate ($g_i \approx 0$) and there is no plant mortality ($v_i \approx 0$). This scenario closely matches what is expected under a well-controlled response–surface experimental design (Gibson et al., 1999; Hart et al., 2018; Inouye, 2001; Law & Watkinson, 1987) where we have a known, fixed of plants, each of which starts with average initial biomass $\beta_i$. If this first set of conditions lasts for a time period of duration $T_2$, the predicted density of viable seeds remaining in the seed bank will be $N_{i,T_2} = e^{-\kappa_i T_2}N_{i,T_1}$. The density of plants will stay constant—implying $p_{i,T_2} = p_{i,T_1}$. If we start from $B_{i,T_1} = \beta_i p_{i,T_1}$, the predicted final plant biomass density will be

$$B_{i,T_2} = \frac{\beta_i e^{\gamma_i T_2} p_{i,T_1}}{1 + \frac{\beta_i}{\kappa_i}(e^{\gamma_i T_2} - 1)p_{i,T_1}},$$

and hence the predicted per capita fecundity of surviving plants is

$$F_{i,T_2} = \frac{\phi_i \beta_i e^{\gamma_i T_2}}{1 + \frac{\phi_i}{\kappa_i}(e^{\gamma_i T_2} - 1)p_{i,T_2}}.$$

Note that Equation 15 is the solution to the well-known logistic-growth equation. Intriguingly, Equation 16 also looks mathematically similar to a reparameterisation of the reciprocal model (Equation 4). That said, the mathematical form of Equation 15 makes it clear that the parameter combinations what would phenomenologically be treated as intrinsic fecundity (the numerator) and as the interaction strength (the parameter combination in the denominator that is multiplied by $p_{i,T_1}$) are not independent of each other since they both depend on $\beta_i$ and $r_i$. While this contrasts to the phenomenological models in which...
these are usually fit as separate free parameters, it is ecologically sensi
tive. To produce a large number of seeds, a plant should be able to grow
large—and the faster it can do so the better it can avoid competition
from its neighbours—and/or convert a large proportion of its biomass
into seeds; however, the larger and faster a population of plants grows,
the quicker it will reach any local limit imposed by carrying capacity,
constraining its ability to grow further and the abilities of any neigh-
bouring conspecific plants (Figure 2a–d).

Scenario 2: In second instance, we will assume that seeds do not
germinate (\(y_t \approx 0\)) but there is plant mortality (\(v_t > 0\)). Even
under the best controlled experimental settings, it is not uncom-
mon for some plants to die over the course of an experiment. If this
second set of conditions lasts for a time period of duration \(T_2\), the
predicted number of viable seeds remaining in the seed bank will
again be \(N_{i,T_2} = e^{-\gamma T_2} N_{i,T_1}\), but the predicted number of plants will be
\(p_{i,T_2} = e^{-\gamma T_2} p_{i,T_1}\). The final predicted plant biomass density will be
\[
B_{i,T_2} = \frac{\beta_i e^{(r_i-1)T_2} p_{i,T_1}}{1 + \frac{\beta_i}{K_i} (e^{(r_i-1)T_2} - 1) p_{i,T_1}}
\]
and hence the predicted per capita fecundity of surviving plants is
\[
F_{i,T_2} = \frac{\phi_i \beta_i e^{r_i T_2}}{1 + \frac{\beta_i}{K_i} (e^{r_i T_2} - e^{r_i T_2}) p_{i,T_2}}.
\]
which again bear strong resemblance to the solution to logistic-growth
equation and reciprocal model respectively. Perhaps not surprisingly,
the most immediate consequences of including plant mortality are de-
creases in final plant density and in those plants’ final total biomass
at the end of the growing season (Figure 2e–g). When combined to-
gether, these two outcomes create a decrease in predicted per capita
fecundity of the surviving plants and hence an appearance of stronger
density dependence (Figure 2h).

Scenario 3: In third instance, we will assume that seeds do ger-
minate (\(y_t > 0\)) and there is plant mortality (\(v_t > 0\)). Unfortunately,
although this scenario is likely the most biologically realistic in the
field, it is one for which there is no explicit analytical solution to the
corresponding initial-value problem for final predicted biomass den-
sity and hence per capita fecundity. Numerical simulations of this
scenario indicate that mortality of more mature plants is unlikely
to be compensated for by the appearance of smaller germinants
(Figure 2i–l). Therefore plant mortality plus prolonged seed ger-
mation will most likely lead to decreases in predicted final total bio-
mass density at the end of the plant growth/seed production phase
(Figure 2k). The changes to predicted per capita fecundity, however,
are much more substantial and deviate considerably from the quali-
itative patterns we have seen so far (Figure 2l). First, predicted per
capita fecundity always decreases compared to Scenarios 1 and 2,
which is reasonable given that similar biomass is shared between a
greater number of plants. Second, even small amounts of germina-
tion can produce nonmonotonic density dependence relationships,
giving rise to what could be described visually as facilitative intra-
specific effects at low density that eventually shift to competitive
intraspecific effects as densities increase. If someone fit the phe-
no
nomenological models to the data in Figure 2l, the result ‘flatten-
ning’ of the relationship between per capita fecundity and final plant
density could also give the false impression of comparatively weak
density dependence.

3.2.2 Multi-species predictions

Because the dynamics of viable seeds in the seed bank \(N_i\) and of
plant density \(p_{i,t}\) are independent for each species, the predicted
multi-species dynamics of both of these state variables match the
single-species cases described above. On the other hand, we must
solve the coupled initial value problem for all species simultaneously
to determine the final predicted biomass densities and predicted per
capita fecundities in a multi-species system given by Equations 9–
11. Even without seed germination, these nonlinear equations (i.e.
quadratic in \(B\)) fall into a class of well-studied first-order ordinary
differential equations known as Riccati matrix differential equations
(Darling, 1997; Reid, 1946, 1972). For such equations, analytical
solutions may not always exist (Califér & Willems, 1981), and even
when they do they are likely complex (Kučera, 2011; Ledyaev, 2011).

Like the single-species Scenario 3 above, we can nevertheless
numerically simulate the predictions of the multi-species model in
order to understand the consequences of the model param-
eters. As a representative example, consider a two-species sce-
nario where seeds from species \(j\) continue to germinate during the
growing season (\(y_j \geq 0\)) and plants of species \(j\) experience mortality

FIGURE 2 Simulated predictions of Equations 6–8 for three distinct single-species scenarios. (a and b) Model-predicted plant density (\(p_{i,t}\)) and plant biomass (\(B_{i,t}\)) during a growing season of duration \(T_2 = 0.5\) starting with 1, 10 or 50 initial plant densities (\(p_{i,0}\); solid, dashed and dotted lines) and with neither continued seed germination (\(y_t = 0\)) nor plant mortality (\(v_t = 0\)). (c) Model-predicted final plant biomass (\(B_{i,t}\)) at the end of the growing season as a function of the final plant density (\(p_{i,T_2}\)). (d) Model-predicted final per capita fecundity (\(F_{i,T_2}\)) at the end of the growing season as a function of the final plant density (\(p_{i,T_2}\)), i.e. \(f, g, h\) and \(i\) corresponding model-predicted plant density, plant biomass, final plant biomass and final per capita fecundity with no continued seed germination (\(y_t = 0\)) but some plant mortality (\(v_t = 1.5\)). Note that one must start with many more germinants at the beginning of the growing season to achieve similar variation in final plant densities in \(g\) and \(h\). \((i, j, k, l)\) Corresponding model-predicted plant density, plant biomass, final plant biomass and final per capita fecundity with some continued seed germination (\(y_t = 0.5\) and \(N_{i,T_1} = 15\)) and some plant mortality (\(v_t = 1.5\)). Note that continued germination makes it harder to observe low final plant densities in \(k\) and \(l\). The dashed grey lines in \(g\) and \(k\) correspond to the prediction in \(c\), and the dashed grey lines in \(h\) and \(l\) correspond to the prediction in \(d\); all are presented to facilitate visual comparison. All other parameters are seed mortality rate \(y_t = 0\),
germinant biomass \(\beta_i = 0.02\), biomass growth rate \(r_i = 12\), biomass carrying capacity \(K_i = 100\) and biomass to seed conversion rate \(\phi_i = 5\).
**γ_j = 0 and ν_i = 0**

- Plant density ($p_i$)

**γ_j = 0 and ν_i > 0**

- Plant density ($p_i$)

**γ_j > 0 and ν_i > 0**

- Plant density ($p_i$)

**Plant biomass ($B_i$)**

- 1 plant
- 10 plants
- 50 plants

**Final plant density ($p_{i,T}$)**

**Final plant biomass ($B_{i,T}$)**

**Per capita fecundity ($F_{i,T}$)**
(\(v_j > 0\)), but where the number of plants of species \(i\) remains constant throughout the growing season (\(y_i = 0\) and \(v_i = 0\)). Even if germination rates and mortality rates of species \(j\) are such that the number of plants \(j\) at the end is near identical to the number at the beginning (i.e. \(N_{j,T} = N_{j,T_0}\)), the model predicts considerable qualitative changes to the emergent way that fecundity depends on neighbour densities (Figure 3). Chief among these is nonmonotonic density-dependent behaviour that could not be captured by any of the phenomenological models discussed previously. These deviations are strongest at low abundance because this is where the dynamic consequences of continued germination and plant mortality outweigh the broader limits imposed by local conditions and local carrying capacities.

4 | DISCUSSION

I highlight here the value of digging a bit deeper into the mechanisms that underpin phenomenological models of annual-plant population dynamics in general, and of density-dependent fecundity in particular. I also show the ease with which a purely phenomenological perspective can lead to the adoption of models that may be incompatible with relevant biological phenomena (e.g. facilitation). To overcome these inconsistencies, I outline a novel continuous-time generative model for population dynamics of annual plants with a seed bank. Importantly, this model captures much of the same core ecology of interest as previously employed phenomenological models while also being structurally valid ‘under the hood’.

**FIGURE 3** Final total plant biomass and per capita fecundities as a function of the densities of two annual plants growing in a local environment, as predicted by numerical simulations of Equations 9-11. During the growing season of duration \(T_2 = 0.5\), seeds from species \(i\) do not germinate (\(y_i = 0\)) and plants from species \(i\) do not experience mortality (\(v_i = 0\)). In contrast, an initial \(N_{j,T_0} = 25\) seeds from species \(j\) do germinate (\(y_j = 0.15\)) and plants of species \(j\) do experience mortality (\(v_j = 0.20\)). Species \(j\) exerts stronger competitive effects on species \(i\) than does \(i\) on \(j\) (\(x_j = 0.75\); \(x_i = 0.5\)). (a) Isoclines of constant final plant biomass of species \(i\) (\(B_i,T_2\)), indicating the combinations of final plant densities \((p_{i,T_2}\) and \(p_{j,T_2}\)) at which species \(i\) achieves equivalent final biomass growth (\(B_i,T_2\)). (b) Isoclines of constant final per capita fecundity of species \(i\) (\(F_i,T_2\)). Note that per capita fecundity decreases monotonically as density of either species increases. (c and d) Isoclines of constant final plant biomass (\(B_j,T_2\)) and constant final per capita fecundity (\(F_j,T_2\)) of species \(j\) respectively. Even though final biomass decreases monotonically as both densities increase, note that per capita fecundity varies nonmonotonically as density of either species increases. All other parameters are identical for the two species: germinant biomass \(\tilde{p}_i = \tilde{p}_j = 0.02\), biomass growth rates \(r_i = r_j = 12\), biomass carrying capacities \(K_i = K_j = 100\) and biomass to seed conversion rates \(\phi_i = \phi_j = 5\).
My examination here has a few immediate consequences for how researchers could study annual-plant population dynamics going forward. First, it requires only a bit more work to simulate the predictions of the continuous-time model during any given year as to project the discrete-time model forward in time. By making time during the year explicit, we can obtain a more faithful representation of the actual dynamics at play (Figure 1) as well as more closely match experimental setups that manipulate the duration of the growing season (e.g. Matías et al., 2018) or that change neighbour abundance over the course of an experiment (e.g. Wainwright et al., 2019). Second, using generative models to study how density-dependent fecundity arises reveals a subtle inconsistency with how many phenomenological models are routinely fit to data: Equations 16 and 18 imply that the conspecific predictor variable should be the total density of conspecific plants in a given location—including the focal plant—and not just the density of conspecific neighbours. Third, if researchers continue to prefer to adopt a phenomenological approach for a single-species scenario, the reciprocal model (a.k.a. the Beverton–Holt model) is without a doubt the most consistent. Of course, this is only true when seed germination has finished and features the caveat that intrinsic fecundities and interaction strengths should not be regarded as independent free parameters. In contrast to the single-species case, the multi-species solution to the model presented here (Equations 9–11) will not be mathematically identical to the widely employed multi-species reciprocal model (Equation 4).

Fitting the models developed here to data is a greater challenge than for the phenomenological models (Box 1). This is especially true since many phenomenological models map nicely onto well-established generalised linear models (Bimler et al., 2018; Mayfield & Stouffer, 2017; Rao & Tottong, 2010). Nevertheless, robust methods exist for fitting nonstandard nonlinear models to data (Bolker, 2008; Glass, 1967; Paine et al., 2012). Moreover, one can infer model parameters via simulation even when lacking closed-form mathematical expressions for the resulting density-dependent fecundities (Rosenbaum & Rall, 2018; Stouffer & Novak, 2021). It should therefore be more than feasible for empiricists and theoreticians alike to make use of and expand upon this current study.

Although the models developed here introduce a third state variable, biomass, they still leave out any and all resources over which plants are presumably competing. That is, interactions limit species' abilities to reach their carrying capacity, but what determines carrying capacity is left implicit. This notwithstanding, the majority of the parameters across Equations 6–11 map onto measurable physical quantities (Box 1). Moreover, the interaction parameters (e.g. \( a_q \) and \( a_p \) in Equation 11) are arguably more ‘tangible’ than their phenomenological equivalents since they are also linked to measurable physical quantities—the maximum amount of biomass that could be produced in a local environment—and hence can be directly probed empirically. In contrast, it is often hard to determine what the precise biological meaning is, for example, of statistical variation in phenomenological \( \alpha \)-coefficients along environmental gradients (Bimler et al., 2018; Matías et al., 2018). Moreover, by separating biomass growth and seed production into distinct components, it may be easier to tease apart exactly how different drivers influence observed seed production (e.g. floral visitors and soil conditions; Lanuza et al., 2018), especially if there are a priori reasons why one may expect one driver to impact the upper bound imposed by \( K \) and another to alter plant mortality \( \nu \), or the conversion rate \( \phi \) of biomass into seeds. Even when no prior hypothesis exists, researchers should be able to allow the various scenarios to ‘compete’ via common model-comparison approaches (Burnham & Anderson, 2002; Murtaugh, 2009).

Although it is not always the case, the model developed here does predict some patterns that are not radically different from those predicted by the fully phenomenological model (Figures 1 and 2). It may therefore seem reasonable to wonder whether or not the core issues are too mathematically esoteric to be of legitimate concern. Indeed, can the added complexity of a generative model make it worse than no generative model at all? From my perspective, there are two compelling arguments against such thinking. First, models serve many purposes, and chief among these are description and prediction (Hilborn & Mangel, 1997; Houlahan et al., 2015; Shmueli, 2010). Unfortunately, there is a long and growing list of examples where similarly descriptive models can produce wildly different predictions (Aldebert & Stouffer, 2018; Coblentz & DeLong, 2020; Fussmann & Blasius, 2005; Iles & Novak, 2016; Myerscough et al., 1996), a phenomenon dubbed ‘structural sensitivity’ (Adamson & Morozov, 2013; Cordoleani et al., 2011; Wood & Thomas, 1999). Second, common advice given to novice modellers is to extend pre-existing models to explore the effect of their additional biological phenomenon of interest (Otto & Day, 2007; Wilson & Collins, 2019). But when that extended model produces counter-intuitive behaviour, it can be difficult to isolate the reason(s) why (Levins, 1966). Was it simply because of how the model was altered? Or was the pre-existing model structurally invalid and thus the researchers were unknowingly like the foolish person who built their house on sand instead of rock (Everyman’s Library, 1998)?

My exploration here also highlights some subtle but crucial details regarding how researchers assess the validity of biological models. In particular, there are various ways in which a model can be wrong (Box, 1976; Levins, 1966; Wimsatt, 1987). On the one hand, there are assumptions which we know are untrue but can rationalise around: dynamics of some generalist consumers may be adequately captured by a single-species model even when they are embedded in many-species communities (Murdoch et al., 2002); if we wish to explore community behaviour near equilibrium, a linear model tailored after a Taylor series expansion may be suitable (MacArthur, 1970); or if plant–plant interactions are mostly driven by how they use shared resources, we can constrain them all to be net competitive (Matías et al., 2018). On the other hand, there are model components that we know to be false and those falsehoods violate the underlying biology; obligate mutualists should go extinct in the complete absence of their mutualistic partners (Janzen, 1979); a predator should go locally extinct (or migrate...
A key contrast between the generative models studied here and the purely phenomenological annual-plant competition model is the number of parameters required to describe the dynamics of annual plants with a seed bank (Table 1). To simulate single-species population dynamics under the phenomenological approach, there are only four key parameters: $g_i$, $s_i$, $i$, and $a_i$. In the most complex single-species generative model, there are nine key parameters: $T_1$, $T_2$, $r_i$, $a_i$, $v_i$, $r_j$, $K_i$, $\beta$, and $\phi$. For every additional species $j$, we get up to nine more species-specific parameters ($T_{1j}$, $T_{2j}$, $r_j$, $a_j$, $v_j$, $r_j$, $K_j$, $\beta_j$, and $\phi_j$) and all pairwise interaction coefficients ($a_{ij}$ and $a_{ji}$). Given this added model complexity, it is important to determine the extent to which these generative models can still mapped onto empirical data, and what additional information (if any) researchers could track to facilitate their empirical evaluation in the future.

At first blush, it appears reasonable to expect researchers to at least have a coarse estimate of the duration of the ‘growing’ season $T_2$ during the year, and hence also the duration of the ‘nongrowing’ period $T_1$. As noted elsewhere, these may be common across species or may be species specific. The discrete-time germination proportion $g_i$ is an analogue to germination rate $r_i$, and both of these can be estimated empirically with existing experimental methods (as in Godoy & Levine, 2014). The same is true for seed survival proportion $s_i$ and its analogue seed mortality rate $\mu_i$ (as in Petry et al., 2018).

In order to estimate plant mortality rate $v_i$, we require initial and final plant abundances over the course of the growing season. Unfortunately, to the best of my knowledge, empirical studies of density-dependent fecundity tend to only have records of the final neighbourhood surrounding a focal plant at the end of the growing season (as in Mayfield & Stouffer, 2017) but not equivalent information at the start of the growing season. In observational contexts, this is perhaps understandable due to the considerable challenge of correctly determining the species identities of small germinants at the start of the growing season; it would nonetheless represent valuable information to collect in future studies. In manipulative experiments, it may be easier to collect these data or to assume/know that $v_i \approx 0$; however, realistic knowledge of this parameter is still valuable when extrapolating from controlled situations to real-world predictions, and even cross-species averages would provide a useful starting point.

Of the four remaining single-species parameters (all of which relate directly to biomass growth, seed production, and per capita fecundity), $\phi$ is perhaps the easiest to estimate empirically since all it requires researchers to estimate per capita fecundity (by collecting and counting seeds produced) and per capita biomass (by harvesting and weighing individual plants) to determine the average conversion rate between the two. Unfortunately, it also appears rare for researchers to measure both of these outcomes within the same experiment. Alternatively, one could conceptualise $\phi = \sigma_i / \alpha_i$, where $\sigma_i \in [0, 1]$ is the average proportion of total plant biomass allocated to seeds and $\alpha_i$ is the mass of an average seed. The proportion $\sigma_i$ is closely related to the Harvest Index (HI) which is the ratio of the yield of grain to total shoot dry matter (Donald et al., 1976; Pomeroy et al., 2020). In contrast to $\phi$, values of $\sigma_i$ may actually be available for many species within existing databases and therefore could already help make constrained predictions about $\phi$.

Estimation of $\beta_i$ is also superficially straightforward since it involves estimating average individual plant biomass at the start of the growing season. However, doing so is complicated by the aforementioned issue of species identification; moreover, measuring per capita biomass is a destructive process that would require estimating $\beta_i$ and total biomass $B_i T_1$ or per capita fecundity $F_i T_2$ in separate experiments/plots and hence could lead to notable increases in replication. On the other hand, larger seeded species tend to have larger, fitter seedlings (Giles, 1990; Moles & Westoby, 2006). Therefore any information obtained about $\sigma_i$ to inform $\phi_i$ might help provide a rough guide to the expected value of $\beta_i$ as well.

Given variation in the density of co-occurring conspecific plants $p_{ij}$, as in the widely employed response–surface experimental designs (Hart et al., 2018; Inouye, 2001), it would then be possible to infer both of the remaining parameters $r_i$ and $K_i$ via nonlinear regression and/or simulation (Bolker, 2008; Rosenbaum & Rall, 2018). Variation in the density of co-occurring heterospecific plants $p_{ji}$ enables additional inference of $\alpha_{ij}$. With sufficient knowledge of the species or system under study, it may also be possible to provide an alternative empirical estimate of $K_i$ by leaning on its explicit meaning in Equations 8 and 11: the maximum amount of total biomass that a population of species $i$ is ever expected to obtain. From similar arguments, $\phi_i K_i$ is maximum total fecundity that all local individuals of species $i$ are ever expected to produce. Both of these upper bounds could be estimated via community surveys without the need for more laborious response–surface designs.

With few exceptions and independent of whether the study is manipulative or observational, a typical dataset consists solely of (i) estimates of the fecundities of focal individuals and (ii) identities and abundances of conspecific and heterospecific plants in each focal individual’s local ‘neighbourhood’ (e.g. Wainwright et al., 2019). What then can be done about the four unknown parameters in these less than ideal situations? Unfortunately, in this situation only two of these last four parameters will ever be unequivocally identifiable, implying that the generative model is over-parameterised for such data. This over-parameterisation is due to (i) nonindependence between $r_i$ and $K_i$ and (ii) nonindependence between $\beta_i$ and $\phi_i$. (Not coincidentally, this is why the single-species phenomenological models have exactly two parameters.) Therefore, it is necessary to fix two of the values prior to model fitting as a form of nondimensionalisation. In the absence of a better alternative, I would recommend fixing carrying capacity to $K_i = 1$ and germinant biomass $\beta_i$ to any small value such that $\beta_i, p_{i1} \ll K_i$. This implies that early stage germinants represent a very small proportion of a
if all of its prey have done the same (Allesina et al., 2009); or a population-dynamics model should never predict negative biomasses (Gilarranz et al., 2016). Distinguishing between these cases can be highly nontrivial—it involves carefully scrutinising emergent model predictions, model components, the origin of those model components, etc. In our increasingly computational era, it is all too easy to devote most of our time and effort on disentangling model outputs. I therefore hope this study serves as a reminder of the vital importance of checking model inputs as well.

ACKNOWLEDGEMENTS

I thank Peter Abrams, Steve Ellner, and Mark Novak for feedback on a very early draft; and Alba Cervantes Loreto, Shawn Leroux, Andrew Letten, Anne McLeod, Robert Freckleton and two anonymous reviewers for suggestions and comments on later versions. I am also grateful for the support of the Marsden Fund Council, from New Zealand Government funding (grant 16-UOC-008).

CONFLICT OF INTEREST

I have no conflict of interest to declare.

PEER REVIEW

The peer review history for this article is available at https://publons.com/publon/10.1111/2041-210X.13965.

DATA AVAILABILITY STATEMENT

Code necessary to recreate all analysis presented in this manuscript and providing examples of simulating the two-species population-dynamics model and fitting the model to simulated data based on a two-species response–surface experimental design is available at https://github.com/stoufferlab/annual-plant-dynamics/ and archived on Zenodo (Stouffer, 2022, https://doi.org/10.5281/zenodo.6946484).

ORCID

Daniel B. Stouffer @ https://orcid.org/0000-0001-9436-9674

REFERENCES

Abrahams, P. A. (2015). Why ratio dependence is (still) a bad model of predation. Biological Reviews, 90, 794–814.

Adams, T. P., Purves, D. W., & Pacala, S. W. (2007). Understanding height-structured competition in forests: Is there an R* for light? Proceedings of the Royal Society of London B, 274, 3039–3048.

Adamson, M. W., & Morozov, A. Y. (2013). When can we trust our model predictions? Unearthing structural sensitivity in biological systems. Proceedings of the Royal Society of London A, 469, 20120500.

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.

Aldebert, C., & Stouffer, D. B. (2018). Community dynamics and sensitivity to model structure: Towards a probabilistic view of process-based model predictions. Journal of the Royal Society Interface, 15, 20180741.

Allesina, S., Bodini, A., & Pascale, M. (2009). Functional links and robustness in food webs. Philosophical Transactions of the Royal Society of London B, 364, 1701–1709.

Antonovics, J., & Levin, D. A. (1980). The ecological and genetic consequences of density-dependent regulation in plants. Annual Review of Ecology and Systematics, 11, 411–452.

Arditi, R., & Michalski, J. (1996). Nonlinear food web models and their responses to increased basal productivity. In G. A. Polis & K. O. Winemiller (Eds.), Food webs: Integration of patterns and dynamics (pp. 122–133). Chapman & Hall.

Beverton, R. J. H., & Holt, S. J. (1957). Dynamics of exploited fish populations, volume XIX of fishery investigations series II. Her Majesty’s Stationery Office.

Bimler, M. D., Stouffer, D. B., Lai, H. R., & Mayfield, M. M. (2018). Accurate predictions of coexistence in natural systems require the inclusion of facilitative interactions and environmental dependency. Journal of Ecology, 106, 1839–1852.

Bolker, B. M. (2008). Ecological models and data in R. Princeton University Press.

Box, G. E. P. (1976). Science and statistics. Journal of the American Statistical Association, 71, 791–799.

Box, G. E. P., & Luceño, A. (1997). Statistical control by monitoring and adjustment. John Wiley & Sons, Inc.

Brandt, A. J., de Kroon, H., Reynolds, H. L., & Burns, J. H. (2013). Soil heterogeneity generated by plant–soil feedbacks has implications for species recruitment and coexistence. Journal of Ecology, 101, 277–286.

Broekman, M. J. E., Muller-Landau, H. C., Visser, M. D., Jongejans, E., Wright, S. J., & de Kroon, H. (2019). Signs of stabilisation and stable coexistence. Ecology Letters, 22, 1957–1975.

Burnham, K. P., & Anderson, D. R. (2002). Model selection and multimodel inference. Springer-Verlag.

Callier, F., & Willems, J. (1981). Criterion for the convergence of the solution of the Riccati differential equation. IEEE Transactions on Automatic Control, 26, 1232–1242.

Cervantes-Loreto, A., Ayers, C. A., Dobbs, E. K., Brosi, B. J., & Stouffer, D. B. (2021). The context dependency of pollinator interference: How environmental conditions and co-foraging species impact floral visitation. Ecology Letters 24,1443–1454. ISSN 1461-023X.

Chu, C., & Adler, P. B. (2015). Large niche differences emerge at the recruitment stage to stabilize grassland coexistence. Ecological Monographs, 85, 373–392.

Coblentz, K. E., & DeLong, J. P. (2020). Predator-dependent functional responses alter the coexistence and indirect effects among prey that share a predator. Oikos, 129, 1404–1414.

Cohen, D. (1966). Optimizing reproduction in a randomly varying environment. Journal of Theoretical Biology, 12, 119–129.
