A 'Portfolio of Model Approximations' Approach to Understanding Invasion Success with Vector-Borne Disease

Matthew J Young[1,2] and Nina H Fefferman[1,2]

1 *National Institute for Mathematical and Biological Synthesis (NIMBioS), University of Tennessee, Knoxville, TN, USA*
2 *Department of Ecology and Evolutionary Biology, University of Tennessee, Knoxville, TN, USA*

Abstract

The central challenge of mathematical modeling of real-world systems is to strike an appropriate balance between insightful abstraction and detailed accuracy. Models in mathematical epidemiology frequently tend to either extreme, focusing on analytically provable boundaries in simplified, mass-action approximations, or else relying on calculated numerical solutions and computational simulation experiments to capture nuance and details specific to a particular host-disease system. We propose value in an approach striking a slightly different compromise in which a detailed but analytically difficult system is modeled with careful detail, but then abstraction is applied to the results of numerical solutions to that system, rather than to the biological system itself. In this 'Portfolio of Model Approximations' approach, multiple levels of approximation are used to analyze the model at different scales of complexity. While this method has the potential to introduce error in the translation from model to model, it also has the potential to produce generalizable insight for the set of all similar systems, rather than isolated, tailored results that must be started anew for each next question.

In this paper, we demonstrate this process and its value with a case study from evolutionary epidemiology. We consider a modified Susceptible-Infected-Recovered model for a vector-borne pathogen affecting two annually reproducing hosts. From observing patterns in simulations of the system and exploiting basic epidemiological properties, we construct two approximations of the model at different levels of complexity that act as two hypotheses about the behavior of the model. We compare the predictions of the approximations to the simulated results and discuss the trade-offs between accuracy and abstraction. We discuss the implications for this particular model, and in the context of mathematical biology in general.

keywords: ecological competition, biological invasion, Ross-Macdonald Models, approximation techniques

1 Introduction

Mathematical models of natural systems allow us to gain insight into patterns and causal mechanisms that are difficult, costly, or unethical to study in practice, and to consider the abstract, underlying logic of how things work. Complicated models, including more features, can better approximate biological systems and thus yield more tailored, reliable, and applicable results, but also become increasingly complex to simulate or analyze mathematically [1]. Many techniques therefore exist for analyzing and approximating complex mathematical models, allowing us to understand and predict dynamics of an approximate model rather than the underlying biological processes it was built to reflect [2]. The natural system a model captures can itself suggest methods for simplification and approximation. As a trivial explanatory example, in demographic models, a population model will stabilize only when the realized net growth rate is equal to 1. Using such knowledge can lead to a solution for the equilibrium behavior of the system without requiring any further analysis of the factors contributing to, or equations capturing the dynamics of, birth or death processes in the population.

Because biological models are simplified approximations of extremely complex systems, it is common to add small details bit by bit to increase the representational accuracy of the system [3, 4, 5]. Gains in accuracy from each addition may not be sufficiently justified by their concomitant increase in mathematical complexity. Approximations and heuristics of the model may be much simpler to adjust. Good approximations can account for features which are important relative to their complexity, while ignoring or simplifying features which are not [3]. However, the best choices to make will depend on the purpose of the model and approximation. A model which is designed to make specific predictions regarding a particular case in the real world will ideally make vastly fewer simplifying assumptions than a model designed to elucidate general trends and patterns over a class of systems. The best model is the one that most directly provides the desired insights [6].

One system of acknowledged biological importance and long history of insight generated by models of gradually increasing complexity is that of vector-borne spread of diseases [7, 8, 9, 10, 11, 12]. In a recent paper [13] we construct a modified Ross-MacDonald SIR-type model for a vector-borne disease affecting two annually reproducing hosts. We investigate a biological invasion by one host carrying a novel pathogen into an area with a native host which lacks resistance to this pathogen and so suffers greater effects from disease. Although simulating the model can be done easily and yields intuitive and visible patterns, the many interlocking variables and dynamics involving both daily and yearly timescales render it mathematically intractable to many common techniques for proving rigorous mathematical statements regarding the dynamics. We therefore propose to shift the nature of approximation, away from approximations in the representation of the biological system and towards the realized outcomes of numerical solutions to a more detailed, intractable model. To demonstrate this, we construct two approximations of this model based on reasonable assumptions

regarding the system and its biological underpinnings. The first approximation makes many simplifying assumptions, leading to simpler equations that are easy to extract general principles from. The second approximation relaxes some of these assumptions, leading to more accurate predictions, but more complicated equations that are less easily analyzed.

While the work we present here is tailored specifically to gain increasing insight into the system studied in our own earlier work, our purpose is more generally to propose a case study in model simplification and approximation that can provide broad insight into biological systems for which tailored models are difficult to analyze and interpret. Although we are certainly not the first to adopt such a protocol, we do propose that full understanding of complex biological systems at different scales of behavior may be best achieved by this "portfolio of model approximations" (PMA) approach. We believe that the presentation of this case study therefore extends in utility beyond further understanding of our own question and can help advance a broad class of evolutionary epidemiological models of vector-borne disease.

2 The Model

Note: This work is built as a further study of a model already presented in the literature. Portions of this methods section presenting and explaining the equations and model choices are therefore reproduced directly from [13] for clarity and consistency.

We construct an epidemiological model which tracks the disease dynamics and population of two species of hosts following the introduction of a pathogen: one native host (hereafter simply referred to as "type 1") which is vulnerable to the disease, but due to being well adapted to the native habitat has high fecundity when uninfected, and one invasive host (hereafter referred to as "type 2"), which has coevolved defenses to the pathogen that decrease its duration of infection and probability of death due to infection, but is not inherently as well-adapted to the habitat in the absence of infection (i.e., its intrinsic rate of growth in the new habitat is lower than that of the native). Similar systems have been considered by others, for example [14].

Our initial conditions correspond to a population of uninfected type 1 hosts with a small number of both uninfected and infected type 2 hosts, representing an invasion by a novel competitor carrying a novel pathogen into the type 1 population. We consider a vector-borne disease, and make the simplifying assumption that there is an already abundant competent vector species in the habitat.

The model couples two biological dynamics: the daily vector-borne spread of the disease among hosts, and a yearly host breeding cycle. We simulate in discrete time-steps that represent days using an SIR model taking into account the interactions between the disease, the two species of host, and the vectors. While the vectors are assumed to breed daily, the hosts reproduce as part of an assumed annual breeding season, every t_c time-steps (typically equal to 365).

2.1 Epidemiological Model

The model tracks eight variables corresponding to combinations of host species and vectors with their infection status. Hosts may be of type 1 or 2, and are either susceptible to the disease (S_1, S_2), currently infected (I_1, I_2), or recovered (R_1, R_2). We assume that recovery is complete and recovered individuals suffer no residual effects from their infection aside from a lifelong immunity to becoming reinfected. For simplicity, we model using only one stage of infection in which individuals are both infectious and symptomatic. The model also tracks the status of the vector population, which may either be susceptible (S_v) or infected (I_v). We assume that vectors do not recover from the disease, but also suffer no negative effects from being infected, acting only as carriers.

For convenience of notation, we denote the total number of hosts

$$H = S_1 + I_1 + R_1 + S_2 + I_2 + R_2$$

and the relative frequencies of infection within their respective population

$$F_1 = \frac{I_1}{H}, F_2 = \frac{I_2}{H}, F_v = \frac{I_v}{S_v + I_v}$$

which allows some equations to be written more compactly.

Table 1: Variables

Variable	Description
S_1, I_1, R_1	Susceptible/Infected/Recovered host 1
S_2, I_2, R_2	Susceptible/Infected/Recovered host 2
S_v, I_v	Susceptible/Infected vectors
H	Total hosts
$F_1, F_2, F_v,$	frequency of infection for host 1/host 2/vector

Equation 1 shows continuous ODE equations approximating the dynamics. Note that the actual model instantiates these in discrete time-steps in which t increases by 1 before updating the variables.

Table 2: Parameters for SIR dynamics

Variable	Description
β_1, β_2	Infection rate of hosts from vectors (probability of infection when bitten by infected vector)
δ_1, δ_2	Infection rate of vectors from hosts (probability of infection when biting infected host)
b_1, b_2	Bite rate of vectors on hosts (number of times bitten per day per mosquito divided among the host population)
γ_1, γ_2	Recovery rate for hosts
$\mu_{1-}, \mu_{1+}, \mu_{2-}, \mu_{2+}$	Death rate for hosts when uninfected/infected
α_v, μ_v	Birth and death rates for vectors

$$
\begin{aligned}
\frac{dS_1}{dt} &= -S_1\beta_1 b_1 I_v/H - S_1\mu_{1-} \\
\frac{dI_1}{dt} &= S_1\beta_1 b_1 I_v/H - \gamma_1 I_1 - I_1\mu_{1+} \\
\frac{dR_1}{dt} &= I_1\gamma_1 - R_1\mu_{1-} \\
\frac{dS_2}{dt} &= -S_2\beta_2 b_2 I_v/H - S_2\mu_{2-} \\
\frac{dI_2}{dt} &= S_2\beta_2 b_2 I_v/H - I_2\gamma_1 - I_2\mu_{2+} \\
\frac{dR_2}{dt} &= I_2\gamma_2 - R_2\mu_{2-} \\
\frac{dS_v}{dt} &= \alpha_v H - S_v\delta_1 b_1 F_1 - S_v\delta_2 b_2 F_2 - S_v\mu_v \\
\frac{dI_v}{dt} &= S_v\delta_1 b_1 F_1 + S_v\delta_2 b_2 F_2 - I_v\mu_v
\end{aligned}
\tag{1}
$$

Following a standard SIR model, susceptible hosts can become infected, and infected hosts become recovered, but each equation also contains a negative term corresponding to deaths. Thus, the total population of hosts is strictly decreasing in this time-frame. We assume that the vectors breed on a much shorter timescale than hosts, so we include a term for their births here, while host births are implemented by a yearly breeding event. We assume that the daily birthrate for each vector depends on access to hosts, and is thus equal to $\frac{\alpha_v H}{S_v+I_v}$. A population with a larger number of hosts will be able to sustain a larger number of vectors.

For a population with a constant number of hosts, the equilibrium vector

population will be proportional to the number hosts: aH where $a = \frac{\alpha_v}{\mu_v}$ is the equilibrium vector density. The population of vectors will asymptotically approach this value over time.

2.2 Breeding Event

Every t_c days (typically 365), a breeding event occurs according to the following algorithm:

Table 3: Parameters for breeding event

Variable	Description
$\alpha_{1-}, \alpha_{1+}, \alpha_{2-}, \alpha_{2+}$	Birth rate for hosts when uninfected/infected
κ	Carrying capacity
t_c	Number of days between each breeding cycle

Let

$$\Delta S_1 = t_c\alpha_{1-}(S_1 + R_1) + t_c\alpha_{1+}I_1$$
$$\Delta S_2 = t_c\alpha_{2-}(S_2 + R_2) + t_c\alpha_{2+}I_2$$

be the number of new host offspring of each type born this generation. Let H be the current total number of hosts. Let

$$c = \begin{cases} 0 & \text{if } H \geq \kappa \\ 1 & \text{if } H + \Delta S_1 + \Delta S_2 \leq \kappa \\ \frac{\kappa - H}{\Delta S_1 + \Delta S_2} & \text{otherwise} \end{cases}$$

be the proportion of offspring that survive to adulthood.
Then

$$S_1 + c\Delta S_1 \to S_1$$
$$S_2 + c\Delta S_2 \to S_2$$

We assume there is no vertical disease transmission, so all new hosts and vectors begin in the susceptible category. We assume that the host population is iteroparous, such that the new offspring and the existing adult population both carry over to the next generation. If the new population would exceed the carrying capacity, we assume the limited space or supplies reduces the number

of successful offspring so that the population exactly reaches the carry capacity by reduction in juvenile survival rather than population-wide competition that could also reduce the adult population.

The carrying capacity is therefore what drives the interspecific host competition. Because births of both species are summed and then normalized by the total number of births, the higher the birthrate of one host, the larger a fraction of the available space it will capture during the breeding event. Similarly, the lower the death-rate of a host, the less space it frees up for the next breeding event. Even if one host species would be able to sustain a stable population on its own, the presence of a more fit competitor can lead to the extinction of the less fit type by driving its effective birth rate down.

2.3 Immune-reproductive Trade-offs and Boundary Conditions

To implement the scenario in which type 2 has increased resistance to the disease at the expense of overall fecundity, we implement the following boundary conditions:

$$\begin{aligned}
&\beta_1 > \beta_2 \\
&0 = \gamma_1 < \gamma_2 \\
&\mu_{1-} = \mu_{2-} < \mu_{2+} < \mu_{1+} \\
&\alpha_{1-} > \alpha_{2-} > \alpha_{2+} > \alpha_{1+}
\end{aligned}$$

Type 2 hosts are less likely to contract the disease, and are able to recover from it, while type 1 lack the immunological strength to eradicate it completely. Additionally, while both types of host are weakened by the disease, type 2 suffer fewer negative effects. However, this stronger immune response comes at the cost of reducing their birth rate when compared to healthy type 1 hosts.

3 Results

Before we can present the approximations which are the primary contribution of this work, we must first present the outcomes from the model itself.

The long-term behavior of the model is sensitive to parameter values, but does not depend on the initial conditions, provided the starting size for each population is nonzero. Thus, we focus on presenting an analysis of the parameter space in the competition between hosts, rather than sensitivity to initial conditions.

We classify outcomes for the system into one of four categories:

1. Failure to Establish: Both the pathogen and the invading host 2 population asymptotically go to zero, while the host 1 population remains near the carrying capacity.

2. Coexistence: Both host types survive at a stable level without going extinct.

3. Competitive Exclusion: The host 1 population decreases asymptotically to zero and is replaced completely by type 2 hosts.

4. Extinction: Introduction of infection alters the system such that both host populations asymptotically go to zero.

We define a set of parameters that lead to coexistence, which we refer to as the 'default parameters', shown on Table 4. All figures and numerical results are made using the default values for each parameter except when otherwise specified.

Table 4: Default Parameters

	Host 1		Host 2		Vector
Transmission rate	0.008		0.005		0.05
Recovery rate	0		0.003		
	Uninfected	Infected	Uninfected	Infected	
Death rate	0.001	0.0025	0.001	0.0011	0.02
Birth rate	0.002	0.0003	0.0018	0.0014	0.02

Additionally, we set the carrying capacity $\kappa = 15000$, days per year $t_c = 365$, bite rate $b_j = 1$, and as initial conditions set $S_1 = 14000, S_2 = 1300, I_2 = 200, S_v = 14000$, and all other initial populations to 0. Although in general the vector transmission rate from the host types, δ_1 and δ_2, need not be equal, for simplicity here we set them both equal to 0.05.

Host populations in the model fluctuate over time due to the annual breeding cycle. To observe the longer-term trends, we sample data points once each year immediately after the breeding event (Figure 1), thereby smoothing out the yearly cycles in the population. Under this default scenario, the initial infection grows into an epidemic which reduces the host 1 population, which then causes the outbreak to recede. This in turn allows the host 1 population to recover until it triggers another smaller epidemic, again reducing their population. These oscillations gradually decrease in magnitude and the population approaches a stable equilibrium. Because the total host population reaches the carrying capacity after each breeding event under these parameters, the host 2 population varies inversely with the host 1 population.

Similar behavior is observed over a wide range of parameters, with the equilibrium frequency of host 1 depending primarily on parameters that influence the spread of infection. Figure 2 shows the 200 year projected results for simulations using default parameters for every parameter except α_v, which we multiply by a constant that differs for each simulation, thereby allowing us to choose any vector density.

We observe that the host outcome is strongly dependent on vector density. Low vector density leads to the Failure to Establish outcome. As vector density increases, Coexistence occurs, with the frequency of each host population

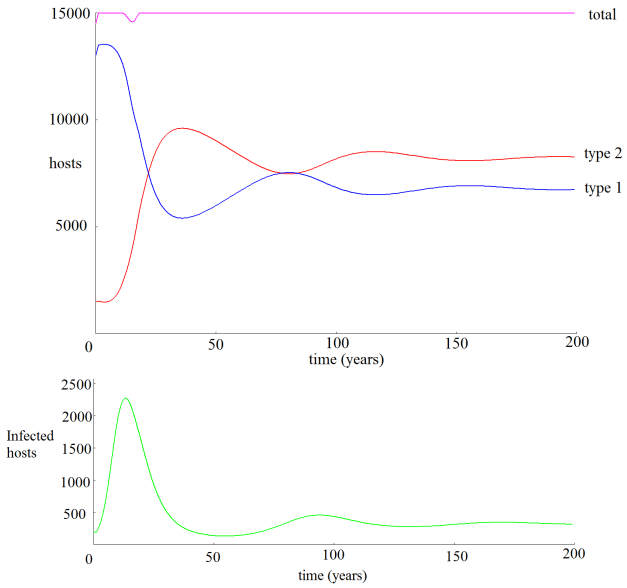

Figure 1: A smoothed, longer-term projection of the host populations (blue and red curves) over 200 years using default parameters. Under this longer time frame, we observe damped oscillations in the infection prevalence (green curve) before the populations stabilize to Coexistence.

changing continuously with vector density. For high density, we observe the Competitive Exclusion outcome.

Extinction never occurs under the range of parameters shown in these figures, since even if every host becomes infected, the host 2 population can replace itself faster than it dies. Extinction can occur under different birth or death rates that do not guarantee demographic replacement for host 2. When populations are below the carrying capacity, the birth and death equations are proportional to the current population size, so the host populations will grow approximately exponentially, assuming the infection level is stable. If the exponent is positive, the population will increase until it reaches the carrying capacity. If the exponent is negative, the population will asymptotically approach 0.

In order to better show how the epidemic interacts with the host frequency equilibrium and extinction, we allow the birth rates of both host types to vary. In particular, we multiply the default values for α_1 and α_2 by the same fixed value Ψ, and construct a plot that shows the population state for a simulation after 200 years, as Ψ and vector density vary (Figure 3).

All four possible outcomes are achievable through different combinations of Ψ and vector density (Figure 3). When birth rates are low, Extinction can occur, regardless of vector density, though the threshold for survival is seen to be vector density dependent. When birth rates are high and vector density is low,

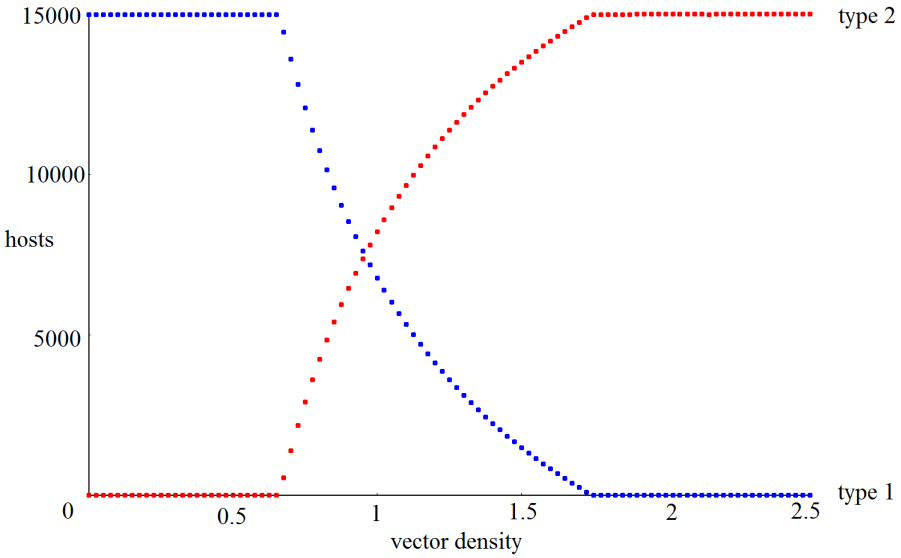

Figure 2: Hosts populations after 200 years as a function of vector density. Because of the similar role of several variables in determining the force of infection, a nearly identical result would occur if the x axis instead presented multiples of any of the pairs of parameters $\{\beta_1, \beta_2\}, \{\delta_1, \delta_2\}, or \{b_1^2, b_2^2\}$.

no epidemic occurs and we see Failure to Establish. Under intermediate values of vector density, we see Coexistence between both host types. Additionally, we observe a continuous gradient of host frequencies, with more type 2 hosts as vector density increases. When density is high, we see Competitive Exclusion, with only type 2 hosts surviving.

We also observe a phase transition between extinction and reaching carrying capacity as the birth rate varies, with a small transition region between them. This occurs since populations near the phase transition will exponentially grow or decay with an exponent very close to 0, so the time required to reach equilibrium will exceed the 200 year horizon presented in Figure 3. As the time horizon increases, the boundary between the extinction (white) and non-extinction (colored) regions in the figure becomes increasingly sharp (not shown).

4 Analytical Approximations

Having presented the results of the initial model, we can now focus on abstracting the behavior of the system to provide better understanding. In order to explain the behavior of the model to understand why each outcome occurs under each set of parameters, we construct an approximation of the model which

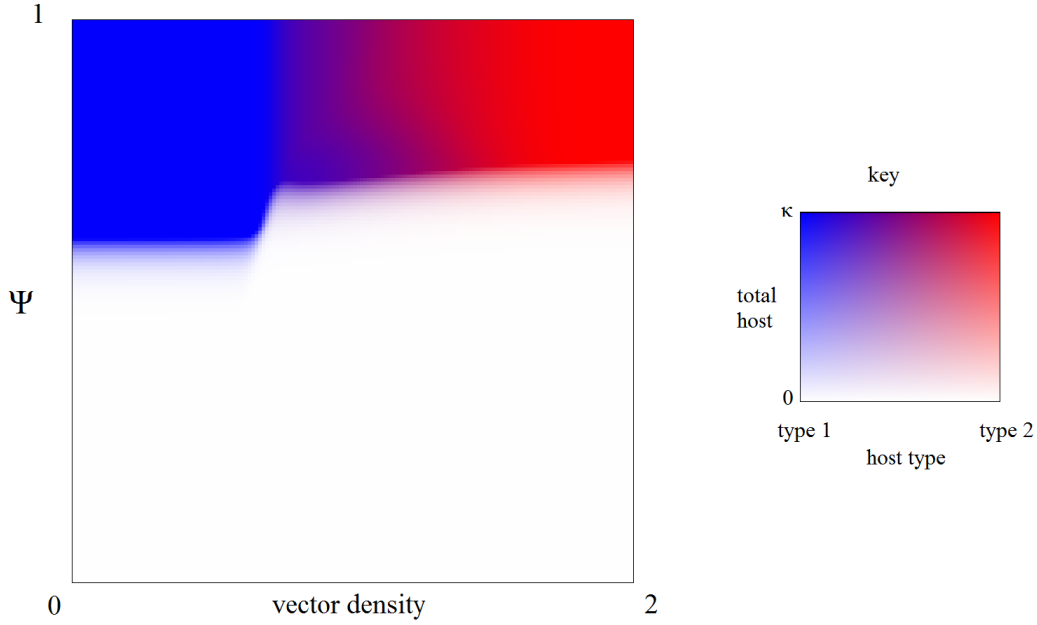

Figure 3: The representation of the two types of hosts after 200 years as the factor by which we multiply the intrinsic birth rate of both hosts, Ψ, and vector density vary

acts as a hypothesis for the description of its behaviors. We begin with a simple approximation that avoids much of the nuance so that we may also avoid most of the complexity, and compare its predictions to the model's outcomes. We then modify it to increase its accuracy, despite the cost of increased complexity, in order to meet a targeted trade-off goal.

4.1 Competition Outcomes Model

Figure 4 shows a flowchart of the Competition Outcomes (CO) Model. Essentially, we first determine the severity of the epidemic, which determines the competitive landscape between the two host types, and we then check the birth rates of the relevant host type to see if they will survive or go extinct.

To derive this algorithm, we first consider the ability of the disease to spread in the population. Due to the heterogeneous population, there is ambiguity in defining R_0 for the disease. To resolve this, we define several related values. Let R_0^j be the R_0 of the disease in a homogeneous population of type j hosts: the average number of type j hosts infected (indirectly, through vectors) from a single infected host in a population consisting entirely of type j hosts. An infected host will remain infected for an average of $1/(\mu_{j+} + \gamma_j)$ days until it recovers or dies, and will infect an average of $a\delta_j b_j$ vectors each day. Each of

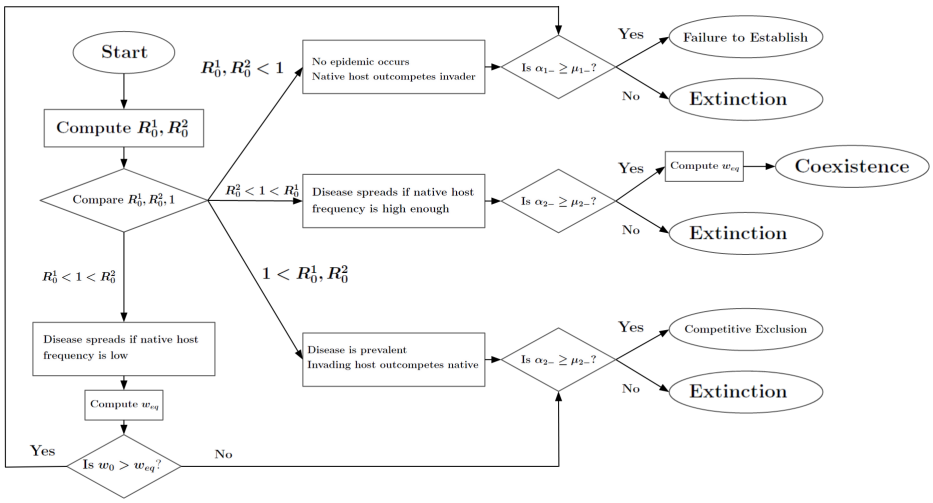

Figure 4: Algorithm for the Competition Outcomes Model

those vectors will survive for an average of $1/\mu_{j+}$ days, and infect an average of $\beta_j b_j$ new hosts each day. Multiplying all of these together yields

$$R_0^j = \frac{\delta_j \beta_j a (b_j)^2}{\mu_v(\mu_{j+} + \gamma_j)} \tag{2}$$

We can also define R_0^v: the average number of vectors that an infected vector will spread the disease to, though this will depend on the frequency of each type of host. Let w be the fraction of living hosts that are type 1 (and $1 - w$ is the fraction of type 2 hosts). Then

$$R_0^v = R_0^1 w + R_0^2 (1 - w)$$

When we compute the average number of hosts that will be infected by a randomly chosen infected host in a mixed population, it also ends up being equal to R_0^v, so this acts as the overall R_0 for the disease as a whole. If $R_0^v > 1$, the infection will spread throughout the population. If $R_0^v < 1$, the infection will exponentially decay.

Note that these R_0 values only accurately describe dynamics during an initial outbreak. In general the effective R_0 of a system is multiplied by the fraction of hosts which are susceptible. If some of the hosts are infected or recovered, then there will be fewer interactions involving susceptible hosts. The frequency of susceptible hosts will differ between type 1 and 2 hosts, but each will be usually be close to 1. The CO model assumes that they equal 1 for simplicity, though we relax this in our second approximation model. Note also that R_0^v can change over time if the frequency of each host type changes, but since it is a weighted average of R_0^1 and R_0^2, it will be bounded between them.

Let ϵ_i be the average fraction of type i hosts that are infected in the long-term equilibrium of the system. Likewise, let ϵ_v be the average fraction of vectors that are infected. If we ignore temporal fluctuations in infection levels, then the death rate of host type i will be approximately the weighted average of the death rate of infected and uninfected hosts.

$\mu_i = \mu_{i+}\epsilon_i + \mu_{i-}(1-\epsilon_i)$

Further, if we ignore the decrease in population throughout the year, the inherent birthrate of each host type can be approximated in the same way

$\alpha_i = \alpha_{i+}\epsilon_i + \alpha_{i-}(1-\epsilon_i)$.

However near the carrying capacity, the effective birthrate of each host will depend on the total death rate creating space for more hosts, and the total birthrate attempting to fill that space.

$\alpha_i' = \alpha_i(w\mu_1 + (1-w)\mu_2)/(w\alpha_1 + (1-w)\alpha_2)$

Treating the infection as static and creating differential equations corresponding to these birth/death rates yields the result that whichever host has a higher ratio α_i/μ_i will eventually reach fixation while the other goes extinct. Given the restrictions in our boundary conditions, this means that host type 1 will dominate when infection rates are consistently low, host type 2 will dominate when infections rates are consistently high, and coexistence will only occur given a careful balance of infection that keeps these ratios roughly equal.

Case 1: Suppose $R_0^1, R_0^2 < 1$, then any initial infection will exponentially decay over time. As a result, $\alpha_1 > \alpha_2$ and $\mu_1 < \mu_2$, and host type 1 will outcompete type 2 in the long term. Thus, the predicted outcome is failure to establish, or extinction.

Case 2: If $R_0^2 < 1 < R_0^1$, then the infection will readily spread in populations iff there are enough type 1 hosts. In particular there exists a frequency of type 1 hosts w_{eq} such that $R_0^v = 1$.

$$w_{eq} := \frac{1 - R_0^2}{R_0^1 - R_0^2}$$

Any time the frequency of type 1 hosts, w, is below this value, R_0^v will be less than 1, reducing infection levels until the population of type 1 hosts can recover. Any time the frequency of type 1 is above this value, R_0^v will be greater than 1, causing the infection to spread until the type 2 hosts can outcompete the type 1 hosts. We note that the birth rates for hosts affect the speed of these effects, but not the equilibrium value of w. As long as the boundary conditions are met, type 1 hosts can replace type 2 near the carrying capacity given low prevalence of the pathogen, and type 2 hosts can replace type 1 given high prevalence. Thus, the predicted outcome is coexistence or extinction.

Case 3: If $1 < R_0^2, R_0^1$, the infection will spread in any population, and type 2 will outcompete type 1 hosts. Thus, the predicted outcome is competitive exclusion, or extinction.

To determine extinction, we first determine which case the host competition falls under. For the CO model, we also ignoring the fact that the decrease in population throughout the year will change the number of births at the

breeding event for the purposes of determining extinction. Instead, we simply check whether $\alpha_i > \mu_i$ for the relevant population, and assume the population survives if it is, and goes extinct if it is not. Additionally, we assume that equilibrium infection rates are sufficiently low such that using the uninfected birth and death rates, α_{i-}, μ_{i-} are close enough approximations for this check.

In case 1, type 1 hosts dominate, so it suffices to check only the survival of host type 1. In the event the type 1 hosts go extinct in the absence of disease, then type 2 hosts will as well given their lower birth rates. Similarly, in case 3, it suffices to check only type 2 hosts.

In case 2 we check only the extinction of type 2 hosts under the assumption type 1 hosts will go to extinction *iff* type 2 hosts do in this case. In one direction, $\alpha_{1-} < \alpha_{2-}$ implies that type 1 hosts meeting our criteria for extinction necessarily implies that type 2 hosts will. In the other direction, type 2 hosts going extinct prevents the negative feedback that keeps the disease in balance from occurring, meaning the disease will rise significantly and likely wipe out type 1 hosts as well. In practice this is not guaranteed to occur, and this assumption is relaxed in the next approximation model, but for simplicity we ignore it here.

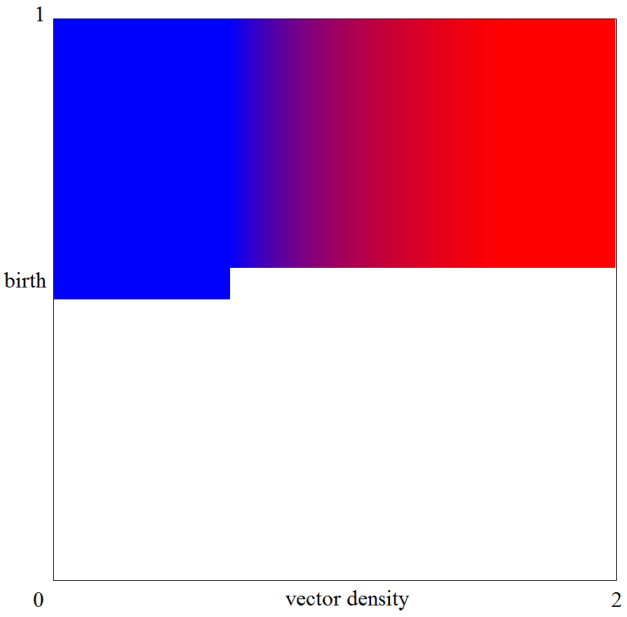

Figure 5: Population state predicted by CO model under same conditions as Figure 3

There is an additional possible case where $R_0^1 < 1 < R_0^2$. This case requires more deviations from our default parameters (so has not been observed in any of the numerical results presented so far) and displays more complex behavior.

We therefore present it separately in section 4.3.

Figure 5 shows the population state predicted by the CO model using the same parameters as Figure 3. We observe the same general qualitative behavior, with the same colored regions in approximately the same locations on the graph. However, it lacks the smooth transition between some regions, and it predicts survival at lower birthrates than the actual model exhibits. The general similarity suggests that the majority of the model's long-term behavior is described by the interactions between the two monoculture R_0 values and the uninfected birth/death parameters. Notably, any change in parameters that don't affect these, such as doubling both β_1 and μ_{1+}, which keeps R_0^1 constant, will have little to no impact on the model's long-term behavior.

4.2 Equilibrium Infection Rates Model

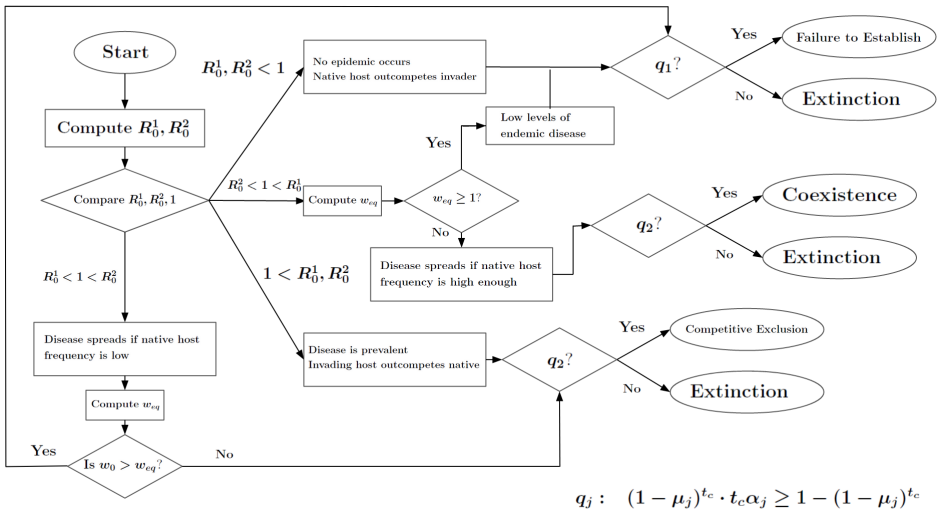

Figure 6: The Equilibrium Infection Rates Model

While the CO model succeeds in trading accuracy for efficiency, it may go too far for some purposes. Many features of the original model fail to be captured by it. We can construct a similar model that is more accurate by relaxing some of the assumptions made by the CO model. This will increase the accuracy of its predictions at the cost of increasing the complexity of the algorithm, increasing the time to compute and decreasing its legibility. This purposeful acceptance of increased complexity (rather than obligate complexity stemming from a model meant to capture the full reality of the biological system) allows us to select the point in the abstraction-realism trade-off that best provides insight into our driving questions.

Figure 6 shows a flowchart of the Equilibrium Infection Rates (EIR) model.

The general structure of the EIR model is mostly the same as the CO model, with three changes.

1: We estimate the average infection level in the host population and use this to compute the effective R_0 values resulting from the fraction of the population that is susceptible. This gives us a more accurate estimate for the equilibrium levels of each host type in the coexistence outcome.

2: We also use this adjusted R_0 to predict an additional case where a small infection persists in type 1 hosts, but at levels too low to require coexistence with type 2 hosts.

3: We update the extinction criteria to account for the fact that host populations are at their lowest during the breeding event.

Let

$$R_{eff} = u_v(R_0^1 u_1 w + R_0^2 u_2(1-w))$$

where u_1, u_2 denote the fraction of type 1 and type 2 hosts which are susceptible, and u_v denotes the fraction of vectors which are susceptible. These values will tend to be close to but less than 1. R_{eff} is the effective R_0 in the population after accounting for the slow in the disease due to already infected and recovered hosts and vectors. Each infected infected vector will infect enough hosts which, on average, will infect R_{eff} vectors before recovering. And similarly, the average infected host will infect enough vectors to infect R_{eff} new hosts. Thus, the disease will increase or decrease in prevalence when $R_{eff} > 1$ or $R_{eff} < 1$ respectively. However, R_{eff} is not fixed from the parameters, but depends also on the current state of the population, and so can change over time. The EIR model will estimate the average level of infection in the long term in order to compute a value for R_{eff}.

If both R_0 values are less than 1, the disease will not spread as in the CO model.

However, when considering the possibility of a mixed population of type 1 and 2 hosts maintaining $R_{eff} = 1$, we no longer have a linear equation. The composition of host types will affect the level of infection and thus the values for u_1 and u_2.

Suppose the population is at carrying capacity and a stable mixture of type 1 and type 2 hosts exists (i.e., each cycle, the same number of type i hosts are born as die). As opposed to the more sophisticated extinction criteria used in the complex version, here we will use the simplifying assumptions that (a) the total number of deaths is proportional to the death rate, and (b) the total number of births is proportional to the equilibrium number of individuals. For the most part, the errors caused by these assumptions will cancel between the type 1 and type 2 hosts. Recall the definitions of α_i and μ_i:

$$\alpha_i = \epsilon_i \alpha_{i+} + (1-\epsilon_i)\alpha_{i-}$$
$$\mu_i = \epsilon_i \mu_{i+} + (1-\epsilon_i)\mu_{i-}$$

Then, the net change in hosts of each type each generation will be zero when

$$wt_c\mu_1 = wt_c\alpha_1 c$$
$$(1-w)t_c\mu_2 = (1-w)t_c\alpha_2 c$$

where c is the proportion of individuals that survive to adulthood, defined in the breeding event. Cancelling redundant terms on each side of the equations results in

$$\mu_1 = \alpha_1 c$$
$$\mu_2 = \alpha_2 c$$

Taking the ratio of these equations gives

$$\frac{\mu_1}{\mu_2} = \frac{\alpha_1}{\alpha_2}$$

If we replace $\mu_1, \mu_2, \alpha_1, \alpha_2$ with their definitions, rearrange terms, and define $\mu_{i'} := \mu_{i+} - \mu_{i-}$, and $\alpha_{i'} := \alpha_{i+} - \alpha_{i-}$ we can rewrite this as

$$\begin{aligned} &\epsilon_1\epsilon_2(\mu_{1'}\alpha_{2'} - \mu_{2'}\alpha_{1'}) \\ &+\epsilon_1(\mu_{1'}\alpha_{2-} - \mu_{2-}\alpha_{1'}) + \epsilon_2(\mu_{1-}\alpha_{2'} - \mu_{2'}\alpha_{1-}) \\ &+(\mu_{1-}\alpha_{2-} - \mu_{2-}\alpha_{1-}) = 0 \end{aligned} \tag{3}$$

we can write ϵ_i as

$$\epsilon_i = \frac{\epsilon_v a\beta_i b_i}{\mu_{i+}} u_i$$

Since $u_i = 1 - \epsilon_i = 1 - \frac{\epsilon_v b_i \beta_w}{\mu_{i+}}$, we can substitute this into the equation and simplify to get

$$\epsilon_i = \epsilon_v \frac{a\beta_i b_i}{\mu_{i+}} \cdot \frac{1}{1 + \frac{\epsilon_v \beta_i b_i}{\mu_{i+}}}$$

Each ϵ_i is almost, but not quite, linear with respect to ϵ_v. We first approximate the ϵ_v in the denominators as 0, and treat these as linear terms with some coefficient, which we can plug into equation 3 to create a quadratic equation with respect to ϵ_v. Within our boundary conditions, this always yields two real solutions, one of which is within the range $[0, 1]$, which we use as an estimate for ϵ_v. We can then increase the accuracy of this estimate by recursively plugging the previous estimate of ϵ_v into the denominators for ϵ_1 and ϵ_2, making linear terms with slightly different coefficients. We can then solve for ϵ_v repeatedly, which yields a converging sequence of estimates. In practice, we find that computing the second element of this sequence (one recursive backstep) yields a good approximation.

Once a satisfactory approximation for ϵ_v is computed, it can be used to compute an estimate of ϵ_1 and ϵ_2 as well, using the above formulas.

Since the type 1 recovery rate, γ_1 is zero, there are no recovered type 1 hosts, so we get $u_1 = 1 - \epsilon_1$. For simplicity, we assume that the number of recovered type 2 hosts will be small, and thus approximate $u_2 = 1-\epsilon_2$. Plugging these into the formula for R_{eff} and setting it equal to 0 allows us to solve for w, computing the estimated frequency of type 1 hosts that will lead to an equilibrium. We label this value w_{eq} for the purposes of this model. Note that this value of w_{eq} may not be in the range $[0, 1]$, indicating that such a coexistence equilibrium is not possible.

This computation of w_{eq} assumes that the frequencies of each type of host, as well as the level of the infection, are constant throughout the year. In practice, the infection spreads quickly following each breeding event and the resulting influx of susceptible hosts, and then gradually decreases as type 2 hosts recover, decreasing R_{eff}. But the approximated value of R_{eff} is close to the time-averaged value.

Case 1: Suppose $R_0^1, R_0^2 < 1$. Since $u_1, u_2, u_v \leq 1$, this implies that $R_{eff} < 1$, so the infection will not spread in the initial population. The type 1 population will therefore outcompete type 2, leading to Failure to Establish or Extinction.

Case 2: if $R_0^2 < R_0^1$, and $w_{eq} > 1$, the resulting value for w is 1. The infection begins to spread, but because R_0^1 is so close to 1, a small decrease in the susceptible population is enough to stabilize it. In particular, this will occur when $u_1 = 1/R_0^1$, thus $\epsilon_1 = 1 - 1/R_0^1$. This case occurs if this infection level is sufficiently low that type 1 hosts can still outcompete type 2. Thus, the predicted outcome is either Failure to Establish or Extinction.

Case 3: if $R_0^2 < R_0^1$, and $w_{eq} \in [0, 1]$, then the equilibrium assumptions hold true, assuming birth rates are high enough to avoid extinction. The predicted outcome is Coexistence with host frequencies w_{eq} and $1 - w_{eq}$, or Extinction.

Case 4: if $R_0^2 < R_0^1$, and $w_{eq} < 0$, the resulting value for w is 0. The infection spreads no matter what the population composition is, so type 2 hosts always outcompete type 1. The outcome is either Competitive Exclusion, or Extinction.

To improve the accuracy of Extinction predictions, we note that deaths throughout the year will diminish the number of hosts that live to the next breeding event. Assuming constant infection rates, the proportion of hosts of type i that survive from one breeding event to the next will be approximately $(1-\mu_i)^{t_c}$. This means that the the effective birthrate each year when below the carrying capacity will be approximately $\alpha_i t_c (1 - \mu_i)^{t_c}$, and the effective death rate will be $1 - (1 - \mu_i)^{t_c}$. It follows that host type i will go extinct if

$$\alpha_i t_c (1 - \mu_i)^{t_c} < 1 - (1 - \mu_i)^{t_c} \tag{4}$$

In Cases 1 and 2, we check this condition only for host type 1, since it is dominant. In case 4 we check this for host type 2 since it is dominant. In case 3 we have the same negative feedback equilibrium as in case 2 for the CO model. Again, we only check the extinction criteria for host type 2, under the

assumption that host type 2 going extinct fails to allow this feedback loop and is likely to drive type 1 extinct, while host type 1 going extinct necessarily implies host type 2 will as well given their relative birth and death parameters. For simplicity, in cases 1,2, and 3 we assume infection rates are low enough that we can use α_{i-} and μ_{i-} for this check.

In case 4, this assumption tends to give less accurate results since the infection can spread without being balanced by changes in host composition, stopping only when the susceptible host rate drops low enough to force it to slow. For simplicity, we assume that in this case, infectivity is high enough that all type 2 hosts become infected relatively early in their life. The average time an infected individual will remain infected is approximately $t_i = \frac{1}{\gamma_2+\mu_{2+}}$, with a chance of surviving of $p = \frac{\gamma_2}{\gamma_2+\mu_{2+}}$. Afterwards, the time the recovered individual will survive uninfected is $t_u = \frac{1}{\mu_{2-}}$. Thus, the average fraction of infected individuals will be

$$\epsilon_2 = \frac{t_i}{pt_u + t_i} = \frac{\mu_{2-}}{\gamma_2 + \mu_{2-}}$$

If R_0^2 is only slightly larger than 1, then this will overestimate the number of infected, as some individuals will never become infected. Additionally, if R_0^2 is very large compared to $1/\gamma_2$, then this will overestimate the number of infected individuals during the breeding event, since most individuals will become infected immediately and have plenty of time to recover, so the number of infections during the breeding event will be much lower than the average number that ϵ_2 represents. Although even more EIR model could be constructed to include more detail on the precise infection rates over time and smooth the transition between cases 3 and 4, we find this simplification provides a reasonable estimate.

We then use this estimate for ϵ_2 to compute the average values for α_2 and μ_2, and test for extinction.

As in the CO model, these cases are not quite exhaustive, leaving out the possibility of $R_0^1 < 1 < R_0^2$, which we address in section 4.3.

Figure 7 shows a plot of the outcomes predicted by the EIR model over the same range of parameters as Figures 3 and 5. Although it doesn't stitch together continuously due to the piece-wise definitions each using different approximations, it still captures the overall behavior of the simulation, and gives more accurate results compared to the CO model.

See the Appendix for a quantified analysis regarding the accuracy of both approximation models over a wider range of parameters.

4.3 Multiple Equilibria

We now address the case where $R_0^1 < 1 < R_0^2$. In this case, type 2 hosts spread the pathogen effectively, while type 1 do not. This causes the fixed point at w_{eq} to be unstable rather than stable. If $w > w_{eq}$, then $R_{eff} < 1$ and the disease will recede, causing type 1 hosts to dominate and increase w even higher. If $w < w_{eq}$ then $R_{eff} > 1$ and the disease will spread, causing type 2 hosts to

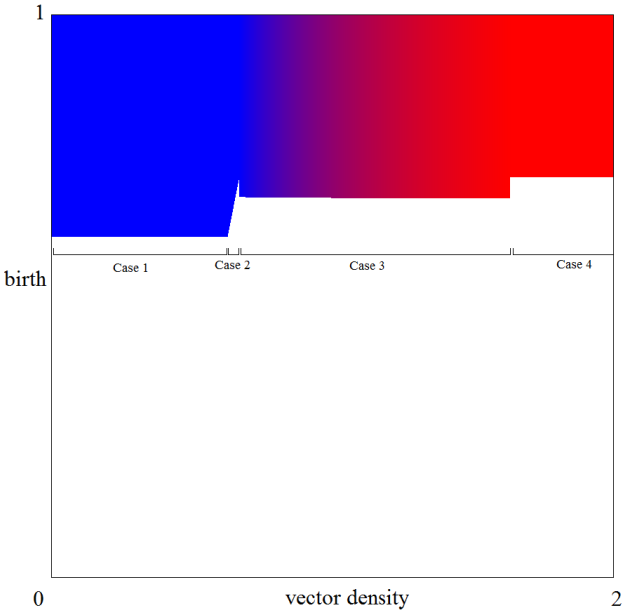

Figure 7: Population state predicted by the EIR model under same conditions as in Figures 3 and 5

dominate and decrease w. Therefore, there are two stable equilibria, one with $w = 0$ and one with $w = 1$, and which one is reached will depend on the initial conditions.

Thus, we can amend both the CO and EIR models above by adding the following case:

Case 0: If $R_0^1 < 1 < R_0^2$, let w_0 be the value of w at time $t = 0$.

Compute w_{eq} using the CO or EIR model, as appropriate.

Case 0-1: If $w_0 > w_{eq}$, w will go to 1. Determine whether the outcome is Failure to Establish, or Extinction, using the same criteria as in Case 1 for the appropriate model.

Case 0-2: If $w_0 < w_{eq}$, w will go to 0. Determine whether the outcome is Competitive Exclusion, or Extinction, using the same criteria as in Case 3 for the CO model, and Case 4 for the EIR model.

Note that this adds another level of approximation on top of those made previously in the models, since the behavior of w with respect to w_{eq} is predicted based on long term behavior. In practice, simulations that have an initially low infection rate will see an increase in w in the first few generations before the pathogen has a chance to increase to equilibrium rates and exert selective pressure on the hosts. Similarly, simulations with an initially high infection rate will see a decrease in w in the first few generations before the pathogen decreases to equilibrium rates. Thus, some sets of initial conditions will be miscategorized

by this model if w begins close to w_{eq} and crosses it early, before the long-term dynamics come into play. Nevertheless, it will correctly categorize most sets of initial conditions.

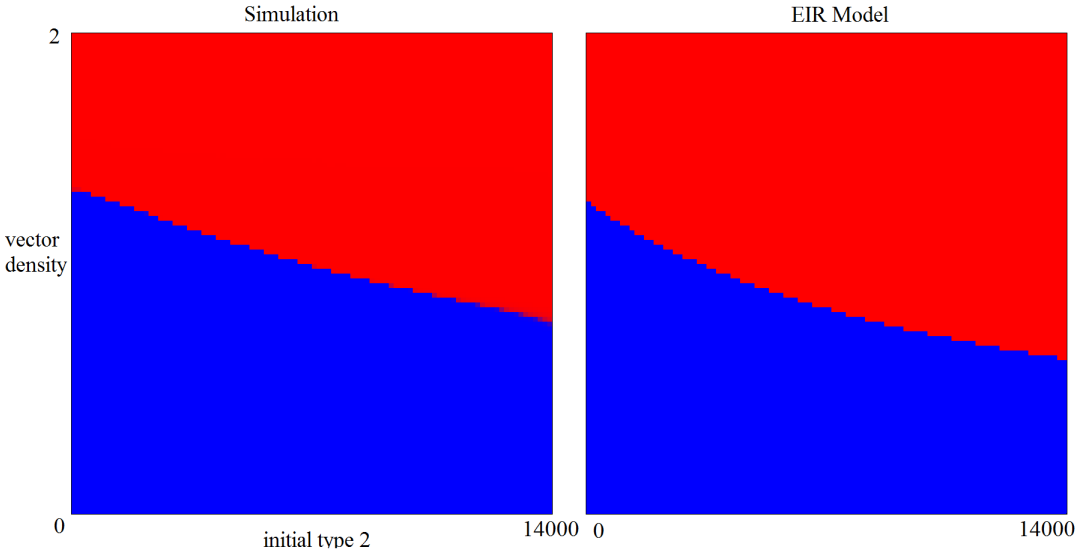

Figure 8: Outcome as a function of vector density (y) and initial number of type 2 hosts (x). Simulation run for 400 years, $\beta_2 = 0.015$

To explore this case, we ran simulations using default parameters except β_2 was multiplied by 3, causing $R_0^1 < R_0^2$. We then varied the vector density and initial frequency of type 2 hosts. Figure 8 shows the outcome for the simulations after 400 years (left) and the predicted equilibrium using the modified EIR model (right). In particular, $R_0^1 = 0.8a$, and $R_0^2 = 1.85a$, so models with a in the range $(0.54, 1.25)$ will fall in case 0, though the actual bounds in the simulation will be slightly lower because $R_{eff} < R_0^v$. We observe also that the model is biased towards the Competitive Exclusion outcome compared to the simulation, which can be explained by the initial infection frequency we chose being less than the equilibrium infection frequency.

5 Discussion

5.1 Specific Example Insights Generated

The Portfolio of Models we constrcut provide explanatory power for the behavior of the system and the role each variable plays. Uninfected birth and death rates occur in the Extinction conditions of both models, but not in the R_0 formulas, therefore they play a minimal role in the competition between host types, at least within the boundary conditions. It is because $\alpha_{1-} > \alpha_{2-}$ that type 1 thrives

with low levels of disease, and if we had $\alpha_{1-} < \alpha_{2-}$ then Competitive Exclusion would replace all non-extinction outcomes. However once this condition is met, the specific amount of each type is primarily determined by disease prevalence.

Another thing the models demonstrate is that, under certain conditions, higher values of R_0^2 increase the frequency of type 2 hosts, assuming extinction does not occur. The type 2 hosts have a higher relative fitness compared to the type 1 hosts the more prevalent the disease is, and thus benefit as a species by furthering ongoing transmission. Although each individual will be worse off by being more likely to be infected and remain infected longer, it will give them the opportunity to spread the disease to more type 1 hosts, reducing their competitor's population and freeing resources for their own offspring, the future generations of type 2 hosts.

It is worth noting that this does not affect the transition between Failure to Establish and Coexistence, only the frequency of type 2 hosts in the Coexistence outcome, or its transition to Competitive Exclusion. Critically, however, it means we can classify possible disease-resistant hosts into two primary categories. A host population whose advantage comes from reduced infection rates, and thus has $R_0^2 < 1$, is effective at suppressing the spread of the disease. This allows the type 1 hosts to survive, leading to Coexistence. A host population whose advantage comes instead from reduced harm from being infected, still having $R_0^2 > 1$, does not suppress the spread of the disease. Instead, they simply survive despite the presence of disease while the native hosts go extinct. In this way, we can see how a model with an invasion by disease-resistant hosts with multiple phenotypes may may see dynamics similar to a public goods dilemma among the invaders, where being infectious is analogous to cooperating. An invading host that is easily infected will decreases its own average fitness, but it increases the overall fitness of the invasive species by eliminating competitors even more. An invading host that is difficult to infect will have higher individual fitness, but will contribute less to the inter-species competition. Future research should consider such dynamics and other possible examples of host-disease symbiosis.

5.2 Trade-off Decisions and Pathways in Selecting Algorithms

Taking a PMA approach assists in deriving these conclusions by disambiguating the levels of effect. The CO model makes it easier to detect and explain broad patterns in behavior with fewer distractions from minor details. Each level of complexity and refinement added are then easier to explain after having already analyzed the simpler models and carrying their insights forward.

In particular, The CO model contains the core insights regarding the role of R_0^1 and R_0^2 in determining the balance between the two host types. Each parameter in the model matters primarily insofar as it influences these two values. Further, we see that the type 2 population is increasing with respect to each R_0, and how that plays out in each region above and below the threshold at 1. Not only do we see how type 2 hosts benefit from the pathogen, but

we can also extract insights about how the pathogen benefits from the type 2 hosts in different ways. We can approximate a monoculture population of type 1 hosts by decreasing α_{2-} and α_{2+} close to 0. In such a case, the pathogen will go extinct regardless of the other parameters. If $R_0^1 < 1$ then obviously the pathogen will go extinct, but because the pathogen kills type 1 hosts so quickly, it simply drives the hosts extinct and itself along with them, resulting in case 2b in the CO model. Thus, the existence of type 2 hosts benefits the pathogen by allowing it to survive. Interestingly, this benefit can take two different forms. In cases 3 and 0-2, the type 1 hosts go extinct, and the type 2 hosts provide a surviving host that the pathogen can spread within without killing off. In case 2, type 2 hosts benefits the pathogen by slowing it down and preventing an uncontrolled epidemic that drives the type 1 hosts extinct. In this case, the pathogen is not stable in a monoculture of either host type type. In a population of only type 1 the pathogen would drive them extinct. But since $R_0^2 < 1$, in a population of only type 2 the pathogen would fail to spread. However, in a mixed population the pathogen is able to use type 1 hosts to spread, while type 2 hosts act as a dampener to prevent it from spreading too aggressively and allow it to maintain a sustainable level. This also suggests that the existence of type 2 hosts can be beneficial to type 1 hosts, since it allows coexistence outcomes when there would otherwise be extinction. In essence, type 2 hosts act as a mini-ecosystem keystone species, preserving both the other host type and the pathogen by stabilizing the capacity for coexistence (as in [15]).

In some ways, of course, this goes too far by suggesting that no pathogen can spread in the type 1 hosts without driving them extinct. The sharp transition when R_0^1 crosses 1 lacks nuance. The EIR model accounts for this via its case 2. We can approximate the case in which type 2 hosts don't exist by decreasing α_{2-} and α_{2+} to near 0, in which case the majority of parameters will fall under case 2 in the EIR model. Thus, the disease will spread in the type 1 hosts up to the point where the frequency of infection lowers R_0^1 to 1. For high vector densities, this will dramatically raise the birthrate required for survival, and will likely lead to extinction instead. But for values of R_0^1 close to 1, the type 1 hosts will survive. Thus we see that, in some circumstances, the presence of type 2 hosts is beneficial to type 1 hosts by allowing coexistence rather than extinction, but in some circumstances the type 1 hosts would survive anyway and the coexistence outcome is worse for them. Further, the approximation allows us to derive these results from basic principles, rather than simply observing them in modelling outcomes.

The basic insights and patterns explained by the CO model are all present in the EIR model in some form, but they are obscured by many more steps and equations that go into details about particular terms in the equations. However, the EIR model contains a more nuanced and more accurate description of the system and can be used to describe interactions that the CO model does not capture. Furthermore, once the patterns from the CO model are understood, it is easier to recognize them in the EIR model than it would have been to deduce them from it in isolation. By first considering the CO model, we can understand the broad insights it captures relatively easily, and then can carry

them over to the EIR model and have an easier time noticing the new features that it contains and separating them from the ones that already existed.

Simplified models are useful for understanding complicated real world systems. Similarly, approximations of models are useful for understanding more complex dynamics. Here, we have presented an example case in which selecting model approximations focused on different scales of within-system complexity allowed us to gain different, more nuanced insight into the biologically relevant behaviors of the system. We suggest that having multiple approximations of the same model at differing levels of complexity is a useful tool for gaining a more thorough understanding of the system.

6 Declarations

Funding: This work was supported by the National Science Foundation under awards 1717498 and 2001213.

Conflicts of interest/Competing interests: The authors declare that they have no conflict of interest.

Availability of data and material: Not Applicable

Code availability: Project code is available on the code hosting platform GitHub at https://github.com/kazarraha/SIRVectorModel

Authors' contributions: Not Applicable

Ethics approval: Not Applicable

Consent to participate: Not Applicable

Consent for publication: All authors consent to the publication of this work.

References

[1] In Jae Myung. The importance of complexity in model selection. *Journal of mathematical psychology*, 44(1):190–204, 2000.

[2] Yolanda Guerrero-Sánchez, Muhammad Umar, Zulqurnain Sabir, Juan LG Guirao, and Muhammad Asif Zahoor Raja. Solving a class of biological hiv infection model of latently infected cells using heuristic approach. *Discrete & Continuous Dynamical Systems-S*, 14(10):3611, 2021.

[3] John Matthewson and Michael Weisberg. The structure of tradeoffs in model building. *Synthese*, 170(1):169–190, 2009.

[4] Cory Merow, Mathew J Smith, Thomas C Edwards Jr, Antoine Guisan, Sean M McMahon, Signe Normand, Wilfried Thuiller, Rafael O Wüest, Niklaus E Zimmermann, and Jane Elith. What do we gain from simplicity versus complexity in species distribution models? *Ecography*, 37(12):1267–1281, 2014.

[5] Ulf Johansson, Cecilia Sönströd, Ulf Norinder, and Henrik Boström. Trade-off between accuracy and interpretability for predictive in silico modeling. *Future medicinal chemistry*, 3(6):647–663, 2011.

[6] John Matthewson. Trade-offs in model-building: A more target-oriented approach. *Studies in History and Philosophy of Science Part A*, 42(2): 324–333, 2011.

[7] Tara Sadeghieh, Lisa A Waddell, Victoria Ng, Alexandra Hall, and Jan Sargeant. A scoping review of importation and predictive models related to vector-borne diseases, pathogens, reservoirs, or vectors (1999–2016). *PloS one*, 15(1):e0227678, 2020.

[8] Yu-Han Kao and Marisa C Eisenberg. Practical unidentifiability of a simple vector-borne disease model: Implications for parameter estimation and intervention assessment. *Epidemics*, 25:89–100, 2018.

[9] David J Rogers. Models for vectors and vector-borne diseases. *Advances in parasitology*, 62:1–35, 2006.

[10] Kbenesh Blayneh, Yanzhao Cao, and Hee-Dae Kwon. Optimal control of vector-borne diseases: treatment and prevention. *Discrete & Continuous Dynamical Systems-B*, 11(3):587, 2009.

[11] David L Smith, Katherine E Battle, Simon I Hay, Christopher M Barker, Thomas W Scott, and F Ellis McKenzie. Ross, macdonald, and a theory for the dynamics and control of mosquito-transmitted pathogens. *PLoS pathogens*, 8(4):e1002588, 2012.

[12] Chris M Stone, Samantha R Schwab, Dina M Fonseca, and Nina H Fefferman. Human movement, cooperation and the effectiveness of coordinated vector control strategies. *Journal of the Royal Society Interface*, 14(133): 20170336, 2017.

[13] Matthew J Young and Nina H Fefferman. The dynamics of disease mediated invasions by hosts with immune reproductive tradeoff. *Scientific Reports*, 12(1):1–12, 2022. doi: https://doi.org/10.1038/s41598-022-07962-2.

[14] Tammi L Johnson, Erin L Landguth, and Emily F Stone. Modeling relapsing disease dynamics in a host-vector community. *PLoS neglected tropical diseases*, 10(2):e0004428, 2016. doi: https://doi.org/10.1371/journal.pntd.0004428.

[15] Richard S Inouye. Stabilization of a predator-prey equilibrium by the addition of a second" keystone" victim. *The American Naturalist*, 115(2): 300–305, 1980.