Universitat de les Illes Balears

MASTER'S THESIS

A compartmental model for vector transmitted diseases: an application to *Xylella fastidiosa*

Rosa Flaquer Galmés

Master's Degree in Physics of Complex Systems

(Specialization/Pathway in Complex Systems)

Centre for Postgraduate Studies

Academic Year 2020/2021

A compartmental model for vector transmitted diseases: an application to *Xylella fastidiosa*

Rosa Flaquer Galmés

Master's Thesis

Centre for Postgraduate Studies

University of the Balearic Islands

Academic Year 2020/2021

Keywords:
Ecology, Epidemiology, Biological Modelling

Thesis Supervisor's Name: Dr. Manuel Alberto Matías Muriel

To all those who love and support me.

Acknowledgments

First and foremost I would like to sincerely thank my supervisor Dr. Manuel Matías for his guidance, support and all the hours invested in this project. I would also like to thank him for the empathy and patience he has showed me during all the process. A special thanks also to Alex Giménez-Romero for his insightful comments and the inestimable aid he provided to push forward this project. Without their help, this Master's thesis would not have been possible.

I would like to thank my parents and brother Miquel for the love and support they have always given me, and specially, during the writing process of this project. I am also pleased to acknowledge all my friends, the ones that have been with me since childhood, and the ones whom I have met along the way. Each one of them cheer me up in their own way, and I am really grateful for that. An affectionate thought to David for being the one who shed light on me whenever I got frustrated, from both a scientific and personal perspective.

Finally, I want to mention Martí for staying by my side in my best and worst moments, for supporting me and pushing me to be a better version of myself.

Rosa Flaquer Galmés

Abstract

In 2016 the presence of the bacterium *Xylella fastidiosa* was detected in Majorcan almond trees and was regarded as the causal pathogen agent of the mortality of these trees over the last fifteen years. *Xylella fastidiosa*, being endemic of the Americas, is an insect-transmitted bacterium which is considered one of the major threats to plants worldwide for its large number of hosts, strains and potential vectors. By 2017 the reported incidence of the bacterium over Majorcan almond trees was $79.5\% \pm 2.0$, presenting a great danger to one of the historical crops of the Balearic Islands. then, the aim of this project is to develop a first step on the study of this epidemic by means of a deterministic compartmental (mean-field) model. In particular, we present a model for the vector assisted transmission of the pathogen, performing an analytical and computational analysis. The model incorporates the specific biological and epidemiological considerations of the bacterium *Xylella fastidiosa* and its interactions with insects and hosts in Majorca Island. In our work the vector (insect) population varies with time, mimicking the field observations. One of the conclusions is that, in general, this temporal variation hinders the theoretical characterization of epidemic thresholds. We show that only in the case that the vector population is all the time in its stationary value, the epidemic threshold can be calculated using the standard techniques, and we numerically characterize the value of this threshold.

Resum

El passat 2016 es va detectar la presència del bacteri *Xylella fastidiosa* als ametllers de les Illes Balears, i es determinà com la causa de la mortaldat en aquests arbres dels darrers quinze anys. Aquest bacteri, que es transmet per via d'insectes, és endèmic del continent Americà, i és considerat un dels patògens que major risc representa per a les plantes a escala global, a causa del seu gran nombre d'hostes, ceps i vectors potencials. En data de 2017, es reportà una incidència del bacteri d'un $79.5\% \pm 2.0$ en els ametllers mallorquins, posant en greu risc un dels cultius històrics de les Balears. L'objectiu d'aquest projecte és doncs presentar una primera base per estudiar aquesta epidèmia mitjançant models compartimentals deterministes (de camp mitjà). Concretament, presentam un model matemàtic per a la transmissió de l'epidèmia, que analitzem analíticament i computacionalment, on incorporam els trets biològics i epidemiològics característics del bacteri *Xylella fastidiosa* i la seva interacció amb insectes i hostes a l'illa de Mallorca. En el nostre model la població del vector (insecte) varia amb el temps, imitant les observacions de camp. Un dels resultats d'aquest treball és que aquesta variació temporal obstaculitza la caracterització teòrica dels llindars d'epidèmia. Demostram que només en el cas que la població de vectors estigui tot el temps en el seu valor estacionari, el llindar de l'epidèmia es pot calcular utilitzant les tècniques estàndard, i caracteritzam numèricament el valor d'aquest llindar.

Abbreviations

ALSD almond leaf scorch disease

Xf *Xylella fastidiosa*

DFE Disease-Free Equilibrium

NGM Next Generation Matrix

Contents

Abstract **v**

1 Introduction **1**

2 Epidemiological modeling **4**

2.1 The SIR model 4

2.2 The basic reproduction number $\mathcal{R}_0$ 6

3 Modeling vector-borne diseases caused by *Xylella fastidiosa* **7**

3.1 Preliminary considerations 7

3.2 The model 7

3.2.1 Fixed points 10

3.3 General case, $\delta > 0$: 10

3.3.1 Linear stability analysis 10

3.3.2 Basic reproduction number 11

3.3.3 Fast-slow approximation 13

3.4 Particular case $\delta = 0$ 14

4 Numerical study of the model **16**

4.1 General case $\delta > 0$ 16

4.2 Particular case $\delta = 0$ 24

5 Conclusions **26**

A Next-Generation Matrix Method **27**

B Computation of $\mathcal{R}_0$ considering the transmission from host-vector-host as a two step process **28**

References **29**

1 Introduction

Over the last 15 years, almond trees in Majorca have experienced severe decline and mortality. The almond tree disease was preliminary associated with a complex of fungal trunk pathogens and their interactions with prolonged drought and tree aging, among other disease-predisposing factors. Nevertheless, this early diagnosis, was challenged in 2016 by the detection of the bacterium *Xylella fastidiosa*, *Xf*, in Majorcan almond trees [1]. In a plant infected with *Xf*, the bacteria multiply within its xylem vessels and a biofilm[1] is formed. The presence of this biofilm occludes the vessels, generating clots, thus, the flux of water is inhibited which eventually can block the nutrition of the plant [3]. This hydraulic stress can result in death of the infected plant which, in general, is worst for rain fed crops than irrigated ones.

In 2013 *Xf* was detected for the first time in Europe, associated with a severe disease of olive trees in southern Italy [4]. In spite of this, diseases caused by *Xf* had been known for more than a century [5] affecting a wide a range of crops and trees, some of economic importance. The pathogen is considered to be endemic of the Americas, and it was first reported in California in the beginning of the XX century [6]. *Xf* is one of the most important threats to plants worldwide, given its large host range [4]. Some examples of the different diseases known to be caused by it are Pierce's disease of grapes in California, citrus variegated chlorosis in Brazil, bacterial leaf scorch in shaded trees in North America, oleander leaf scorch in California and olive diseases in Europe [7].

Xf is an insect-transmitted plant disease, the key feature of the diseases caused by it is that its development in nature[2] on a host population is completely dependent on plant-to-plant transmission by xylem-feeding insect vectors. The complexity around *Xf* related diseases resides in the considerable diversity of host plants, bacteria and vectors. The bacterium has multiple strains, each capable of infecting distinct hosts [8]. The vectors are many, and depend on the geographic region; the main vectors in Europe being *Philaenus spumarius*, and in experimental conditions, *Neophilaenus campestris* and *Philaenus italosignus* are also confirmed to be capable of transmitting the bacteria; nevertheless, all xylem fluid-feeding insects are considered to be potential vectors [9]. As per the host range, being more than 500 plant species including both economic crops and plants in natural communities, is one of the widest for any plant pathogen [1].

In almond trees, the bacterium *Xf* causes the almond leaf scorch disease, ALSD. Infected trees show shoot and branch diebacks and/or general decline with frequent death over the following 14 years [1]. In Majorca, this disease it is transmitted mostly by meadow spittlebugs, *Philaenus spumarius* [1]. Prior to the introduction of *Xf*, *Philaenus spumarius* had never been considered an agricultural pest in Europe [10]. These insects are xylem feeders; they have sucking mouthparts with a characteristic shape, known as stylets, that allow them to reach the xylem of the host plant from where they ingest the sap [9]. It is considered that it is during the feeding process when the *Xf* cells present in the foregut of the insect are inoculated to the host's xylem vessel through egestion [11].

Philaenus spumarius has a one-year seasonal life cycle, which is schematically represented in Fig. 1. The average life of an adult spanning from April to December being in October-November the oviposition, which continues until the female dies naturally or is killed by severe frost. In Majorca the eggs hatch, typically, around March-April when the spittlebug is born as a nymph. They feed on green plants until late spring, when green plants are dry moment in which they migrate to the trees and, by then, their stylet is sufficiently developed to penetrate harder tissues [10].

This insect, *Philaenus spumarius*, has the potential to live under different environmental conditions, from moist to relatively dry, as long as the host plants it feeds on are actively growing and not subject to water stress. As for their movement, it has been recorded in Italy seasonal movement from herbaceous vegetation to the canopy of evergreen trees/shrubs on late spring-early summer. At the end of summer-beginning of autumn adults, mostly females, recolonize herbaceous vegetation looking for suitable sites for oviposition [10]. *Philaenus spumarius* has been reported to have a

[1] A biofilm is defined as structured communities of fixed microbial aggregates, confined in a self-produced polymeric matrix [2].

[2] Human intervention can also cause the disease to spread via grafting infected scions onto local rootstocks or via plant trade [1].

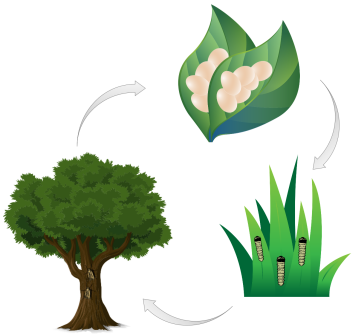


Figure 1: One-year biological cycle of *Philaenus spumarius*.

great flight performance, with a mean flight duration of 500 m in half an hour of flight and a maximum distance traveled of 5.5 km in a single 5.4 h continuous flight [12]. Moreover, there also exists passive dispersal of the insect over great distances is mediated by wind and human activities, for example due to transportation by cars [10].

The lethal outbreak of the disease affecting olive trees in Apulia, Italy, in 2013 put the Mediterranean region on the spotlight and raised alarms about the threat it posed to the Mediterranean agriculture. Since then, various *Xf* genotipes have been detected and/or introduced in Europe [1]. In the Americas, the increase of *Xf*'s threat in the last century was a consequence of an invasive and efficient exotic vector acquiring an endemic pathogen. On the other hand, in Europe, up to now, the threat arises from an endemic vector acquiring an exotic pathogen [5].

Regarding the *Xf* invasion in Majorca, the first abnormal clusters of dying almond trees were noted in Son Carrió, around 2003 and were primarily regarded as a fungal infection. Latter studies revealed the presence of *Xf* in the island and, in the study presented by Moralejo et al. in [1], it was determined that by 2017 the incidence of the ALSD disease over almond trees was $79.5\% \pm 2.0$. This high incidence, spread all across the island, suggested an old introduction of the *Xf* pathogen which had went undetected, its symptoms, leaf scorch and progressive decline, confused for the ones caused by fungal trunk infections propitiated by drought periods. Then, the presence of the fungal pathogens in ALSD infected trees, is a consequence of the stress induced by the presence of *Xf* on the tree, which wakens it, activating the pathogenic phase of the endophytic fungi. In the same study, [1], Moralejo et al. conclude that the most likely date for the arrival of *Xf* in Majorca was 1993 through material infected from California, meaning that the pathogen had went unnoticed for over 20 years in the island.

Given the high incidence of *Xf*, it is essential a better understanding of the disease in order to be able of developing informed containment measures. To this aim, our approach is to use the mathematical scheme provided by the so-called compartmental models. In such models, the population under study is divided into a finite set of non-intersecting classes where the individuals are allowed to move from one class to the other according to some rules. Each class represents the state of a sub-population of individuals and the transfer rules and connectivity between compartments are inferred depending on the specific problem that is being modeled. Thus, the main assumptions when working with compartmental models are: the nature and amount of compartments needed to represent the specific epidemiological problem under study and the transfer rules between such compartments. Compartmental models are valuable as a first approach when dealing with an epidemiological situation, nevertheless it is worth noting that they account for a mean-field description of the system, and thus, the interaction of the sub-populations in each compartment is treated as a whole, not taking into account spatial distribution of the individuals.

The mathematical framework that encompasses compartmental models has two main approaches: the deterministic, described by ODEs, and the stochastic, usually described by master equations or agent based models. In the work presented here we are going to focus on the deterministic approach in order to develop a model for the *Xf* disease in Majorca. We are going to work under the assumption that the size of the population, and thus the different compartments, is large enough

so that the mixing of the different states is homogeneous (well-mixed assumption). In the case of small populations the stochastic approach would be needed.

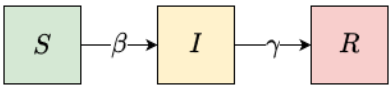


Figure 2: Graphical scheme for the SIR model.

The quintessential compartmental model in epidemic modeling is the well-known **SIR** model, represented schematically in Fig. 2. It is one of the first epidemic models and it was proposed by Kermack and McKendrick in 1927 in [13]. This model supposes that the disease is such that the three classes in which the population can be divided is the following: The *susceptible*, S, are the ones who can become infected by the disease; the *infectives*, I, standing for the ones who already have the disease and can transmit it; and the *removed/recovered* class, R, being the individuals who have had the disease and are removed/recovered.

Other compartmental models have been proposed such as the **SEIR** model, which adds the *exposed* class, E, accounting for a compartment where the disease in the individuals is latent, so they are not yet infective; and the **SI** model, where the infective do not become removed. Other possibilities include **SIRS** models, with temporary immunity on recovery from infection; **SEIS** models, with an exposed period between being infected and becoming infective; or **SIS** which describes a disease with no immunity against reinfection.

Not all infectious diseases are transmitted in the same fashion, so they can be classified and modeled accordingly [14]. **Individual-to-individual** transmitted diseases, including direct contact, touching or sexual, and indirect contact, like fluid exchange, are the ones where the pathogen is transmitted from one infected individual to another. This situations are usually described as contact processes using the SIR or SIR-like models. Airborne diseases, such as influenza or COVID-19, are also usually modeled as a contact process. **Vector-borne** or **vector transmitted** diseases, these are the ones which the pathogenic microorganism is transmitted from infected individual, *host*, to another individual via a living carrier, *vector*; these carriers are usually arthropods, they are thought to not get ill from the disease and, in general, once infected they remain infectious until the end of their lives. Finally, the last two ways of transmission are **vertical-transmission**, when the disease is transmitted from mother to child before birth, and **environmental transmission** where the pathogen is present in contaminated food or water.

The first studies of the dynamics of vector born diseases were done for malaria, at the beginning of the 20th century by Ronald Ross [15]. The mathematical modeling including dynamics of the vectors alongside the dynamics of the hosts. Since these first studies, the framework of vector borne diseases has been also used to study other types of infections in humans like Zika virus in humans [16] and in plants, the diseases caused by *Xf* among others.

In many epidemiological models there exists a threshold behavior: under a certain conditions a disease spreads into the population causing an epidemic, whereas in other cases it does not. Note that this is not always the case, for example a **SI** model (susceptible-infected model with a constant population) or the **SIS** model do not present a threshold, as the disease always spreads into the population. Then, one of the key questions to answer in any epidemic situation is whether or not the infection will proliferate. A widely used parameter to tackle this question is the *basic reproduction number* $\mathcal{R}_0$. It has been regarded "the most important quantity in infectious disease epidemiology" [17]. $\mathcal{R}_0$ represents the number of secondary infections produced by one infectious individual in a population consisting on susceptible individuals only.

The work in this thesis is structured as follows: First, we present a theoretical introduction in Section 2 where we discuss in depth the SIR model, as a basis of our future analysis, and present a brief discussion on $\mathcal{R}_0$. In Section 3 we propose and analytically discuss a deterministic model to tackle this epidemiological situation. Next, in Section 4 we perform a numerical study to validate and contrast the analytical analysis of the previous section. Finally, the conclusions and outlook are presented in Section 5.

2 Epidemiological modeling

2.1 The SIR model

Given the similarities that the SIR model has with the model discussed in this thesis, an in-depth discussion is presented following [18], [19] and [14]. Let $S(t)$, $I(t)$ and $R(t)$ represent the number of individuals in the susceptible, infective and removed class respectively and N be the total number of individuals. The assumptions about the transmission of the infection are the following:

1. An average member of the population makes contact sufficient to transmit infection with βN others per unit time, being $\beta > 0$ a constant parameter.
2. The time spent in the infectious state is $1/\gamma$, with $\gamma > 0$ a constant parameter, thus the infectives leave the infected class at a rate γI.
3. The incubation period is negligible. Thus, a susceptible in contact with an infective will be infective right away.
4. The only way to enter or leave an state is through the processes listed above. No death or birth is taken into account, and thus, the total population remains constant: $S(t)+I(t)+R(t) = N$. This hypothesis holds when the time scales of the disease are much faster than the demographic time scales.

From Item 1 above, and under the assumption that the size of the different compartments is large enough so that the mixing of the different states is homogeneous (well mixed assumption), it follows that the number of new infections per unit of time per infective is $(\beta N)(S/N)$ since the probability that a random contact by an infective is with a susceptible is S/N, giving a rate of new infections $(\beta N)(S/N)I = \beta SI$. Another way of understanding the infection process is focusing on the contacts of a susceptible, which with probability I/N will be with an infective, then, the rate of new infections per susceptible is given by $(\beta N)(I/N)$, therefore the rate of new infections is $(\beta N)(I/N)S = \beta SI$. Even if both approaches give rise to the same infection rate, it is worth noting that the underlying interpretations are different, and, depending on the situation, one point of view would be more suitable than the other. Then, the SIR model is represented with the following system of ODEs:

$$\begin{cases} \dot{S} &= -\beta SI \\ \dot{I} &= \beta SI - \gamma I \\ \dot{R} &= \gamma I \end{cases} \tag{2.1}$$

and the set of initial conditions $S(0) = S_0$, $I(0) = I_0$ and $R(0) = R_0$, in most approaches $R_0 = 0$. In general, this model does not have an analytical solution, nevertheless a great deal of information can be obtained following a qualitative approach. First, note that $S(t)$, $I(t)$ and $R(t)$ are positive functions bounded by N as they represent a population. Then, it is straightforward to see that $\dot{S} < 0\ \forall t$ always, thus $S(t)$ is a monotonically decreasing positive function and so:

$$\lim_{t\to\infty} S(t) = S_\infty,\ S_\infty \in (0, S_0] \tag{2.2}$$

In the same way, $\dot{R} > 0\ \forall t$ always too, so $R(t)$ also has a monotone behavior, in this case it is a monotonically increasing positive function. The number of removed individuals is bounded by the total size of the population, thus:

$$\lim_{t\to\infty} R(t) = R_\infty,\ R_\infty \in [R_0, N] \tag{2.3}$$

The behavior of the infective class is not so straightforward. Rewriting the equation for $\dot{I}$ as $\dot{I} = I(\beta S - \gamma)$ one can see that $I(t)$ can have either a monotonically decreasing behavior or a non

monotonic one, by first increasing to some maximum and then decreasing, depending on the sign of $(\beta S - \gamma)$. Moreover, as $S(t)$ is a monotonically decreasing positive function, the initial condition fully determines whether or not there will be an increase of the number of infective individuals, then:

$$\left.\frac{dI}{dt}\right|_{t=0} = I_0(\beta S_0 - \gamma) \Longrightarrow \begin{cases} \left.\frac{dI}{dt}\right|_{t=0} > 0 & \text{iif } S_0 > \frac{\gamma}{\beta} \\ \left.\frac{dI}{dt}\right|_{t=0} < 0 & \text{iif } S_0 < \frac{\gamma}{\beta} \end{cases} \tag{2.4}$$

Eq. (2.4) answers the question whether a disease will spread or not into the population, key in epidemic modeling. So, the SIR model gives rise to a threshold phenomenon. If $S_0 > S_c = \gamma/\beta$ there is an epidemic[3], while if $S_0 < S_c = \gamma/\beta$ there is not. This threshold defines the so-called *basic reproduction number* $\mathcal{R}_0$:

$$\mathcal{R}_0 = \frac{\beta S_0}{\gamma} \tag{2.5}$$

$\mathcal{R}_0$ defines whether there is an epidemic or not. The infection spreads if $\mathcal{R}_0 > 1$, while if $\mathcal{R}_0 < 1$ it dies out. Another important analytical result that can be derived form this model are the phase plane trajectories in the (S, I) and (S, R) space. The trajectories in the (S, I) can be found dividing the equation for $\dot{I}$ and $\dot{S}$ from Eq. (2.1):

$$\frac{dI}{dS} = -1 + \frac{\rho}{S}, \ \rho = \frac{\gamma}{\beta}, \ (I \neq 0) \tag{2.6}$$

Integrating, the phase plane trajectories are:

$$I + S - \rho \ln S = I_0 + S_0 - \rho \ln S_0 \tag{2.7}$$

In the usual case that $R_0 = 0$ then $N = I_0 + S_0$ and Eq. (2.7) can be rewritten to:

$$I + S - \rho \ln S = N - \rho \ln S_0 \Longleftrightarrow N = I + S - \rho \ln \left(\frac{S}{S_0}\right) \tag{2.8}$$

Thus, clearly $I + S < N$ if $t > 0$. From Eq. (2.8) the maxim number of infected can be obtained. The peak occurs when $\dot{I} = 0$, thus from Eq. (2.4), when $S = \frac{\gamma}{\beta} = \rho$:

$$N = I_{max} + \rho - \rho \ln \left(\frac{\rho}{S_0}\right) \Longleftrightarrow I_{max} = N - \rho + \rho \ln \left(\frac{\rho}{S_0}\right) \tag{2.9}$$

Finally, the phase plane trajectories in the (S, R) space are found by dividing $\dot{S}$ and $\dot{R}$ from Eq. (2.1):

$$\frac{dS}{dR} = -\frac{\beta}{\gamma} S = -\frac{S}{\rho} \Longrightarrow S = S_o e^{-\frac{R}{\rho}} \geq S_o e^{-\frac{N}{\rho}} > 0 \tag{2.10}$$

Therefore, Eq. (2.10) implies that $S_\infty > 0$. This result is important as it means that the epidemic does not die out because all susceptible have been infected, there will always be some individuals that escape the disease. Then, the epidemic dies out because of the lack of infective, meaning that $\lim_{t\to\infty} I(t) = I_\infty = 0$. This result can be shown by integrating the equation for $\dot{S}$ in Eq. (2.1):

[3] An epidemic is said to occur if the number of infected individuals increases with time at some point of the dynamical evolution of the system.

$$\begin{aligned}\int_0^\infty \dot{S}dt =& -\beta \int_0^\infty S(t)I(t)dt,\\ S_0 - S_\infty =& \beta \int_0^\infty S(t)I(t)dt,\\ S_0 - S_\infty \geq& \beta S_\infty \int_0^\infty I(t)dt\end{aligned} \tag{2.11}$$

The last inequality in Eq. (2.11) implies that $I(t)$ is integrable on $[0, \infty)$. Hence, $\lim_{t\to\infty} I(t) = 0$. One final remark is that, in this section, we have presented the SIR model under the mass-action formulation used in the early epidemic models following [18], [19] and [14]. This approach supposes a rate of contacts per infective proportional to the population size (the first of the assumptions listed above). An alternative, more realistic formulation, is via the standard incidence, which describes a situation where the number of contacts per infective does not depend on the population size. This alternative formulation leads to the following system of ODEs

$$\begin{cases}\dot{S} &= -\beta S\frac{I}{N}\\ \dot{I} &= \beta S\frac{I}{N} - \gamma I\\ \dot{R} &= \gamma I\end{cases} \tag{2.12}$$

The standard incidence approach is the one which is going to be used along this work.

2.2 The basic reproduction number $\mathcal{R}_0$

In Section 2.1 the concept of $\mathcal{R}_0$ has been introduced as the threshold parameter controlling whether or not an epidemic is expected. Mathematically, $\mathcal{R}_0$ is a threshold parameter, if $\mathcal{R}_0 > 1$ the infection spreads, while if $\mathcal{R}_0 < 1$ it dies out. Epidemiologically, $\mathcal{R}_0$ gives the number of secondary infections produced by one infectious individual in a population consisting on susceptible individuals only. To see this interpretation from Eq. (2.5), one has to notice that the number of new cases per unit of time produced by all infective individuals is βSI. If one single infective individual is considered $I = 1$ in a population consisting of S_0 individuals the number of secondary cases produced by it, per generation, will be βS_0. Since one infective individual remains infectious for a period $1/\gamma$, the number of secondary cases it will produce will be $\beta S_0/\gamma$ which is the expression derived for $\mathcal{R}_0$ in Eq. (2.5) [14].

It has been shown that the basic reproduction number is mathematically characterized by considering the transmission of a disease as a demographic process, where producing an offspring is not seen as giving birth in the demographic sense, but as causing a new infection. This leads to viewing the epidemiological process in terms of consecutive generations of infected individuals [17]. In some models, like the SIR, the computation of $\mathcal{R}_0$ is unambiguous and straightforward, however this is not always the case. The main reason being that there is a certain ambiguity on what is considered as causing a new infection and what is a transition between an individual state to another [20]. As an example, in vector-borne diseases, the infection from host-vector-host can be seen as a one generation process or a two generation process [16], which will result in different expressions of $\mathcal{R}_0$. However, the predicted threshold is going to be the same independently of the choice [20].

The most widely used methods to analytically determine $\mathcal{R}_0$ are the linear stability analysis of the disease free equilibrium, or the next generation matrix, NGM, approach presented by Diekmann and Heesterbeek in [21]. It has to be noted that for the computation of $\mathcal{R}_0$ via the NGM that some assumptions need to be fulfilled, in particular, the existence of a disease free equilibrium, DFE. For further details on the assumptions and the construction of NGM the method see Appendix A.

3 Modeling vector-borne diseases caused by *Xylella fastidiosa*

3.1 Preliminary considerations

As discussed in the previous sections, when posing any compartmental model, we have to make assumptions of the nature of the sub-populations and the interactions among them. In the particular case of diseases caused by *Xf*, these classes have to account not only for the population of hosts, trees, but also for the population of the vectors, insects.

Regarding the main vector in Majorca, *Philaenus spumarius*, it is considered that newborn nymphs do not have the *Xf* bacterium in their organism even if the parents had it; that is, there is no vertical transmission of the pathogen [22]. Then, the only way for a vector of becoming infected by *Xf* is via direct contact with an infected host while feeding. Additionally, it is also considered that the transmission rates between infected host to non infected vector are different than from infected vector to non infected host [11]. Moreover, we will also deem an infected vector to remain infective for all its lifetime and in addition, that the presence of the pathogen does not affect it, so there is no difference between an infected and a non infected one in terms of death rates [22]. Finally, it has been seen that a vector which has been infected with the pathogen is already infectious after a short period of time, a few hours, so a latent period is negligible compared to the rest of the timescales of the model [23].

On the other hand, as per the host population, only death by the infection is considered but not natural death. This simplification is based on the consideration that the characteristic timescale of the natural death of hosts, around 50 years in the case of almond trees, is negligible in front of both the timescale of death by the infection and the lifetime of the vectors, which is one year at most [10] (less if we consider predators and human pest control). Moreover, we do not consider recovery of infected hosts, so the only transition allowed for an infected host is to the removed, dead, compartment.

The focus of this work is to perform a preliminary analysis of a model that, in a future work, could be then integrated into a more sophisticated framework accounting for the seasonality of the vector, alongside climate effects. To this aim we start by considering a model in which the vectors are born and die at constant rates, which are going to be different, in general. This approach is less restrictive than what is usually considered in the literature, where both rates are normally imposed to be equal [19]. Nevertheless, the consideration of constant birth and death rates is a strong simplification of the one-year cycle of *Philaenus spumarius* presented in Fig. 1. This simplification is made under the assumption that the particular life cycle of the vector can be neglected given the difference in scales of the lifetime of trees and vectors. Finally, a more suitable model to include the vector seasonality will be in the limit of the birth rate going to zero. For a zero birth rate, the model presented here accounts for an annual cycle with an initial population of vectors, that die during that year, and then, the model may be iterated to account for longer periods of time.

3.2 The model

The compartmental model that we propose to tackle this problem, describing both host and vector populations is schematized in Fig. 3. It is inspired by the models presented by Brauer and Castillo-Chavez [16], with some variations. In our approach, we consider that the population is fully mixed and that it is large enough to neglect fluctuations. The model,

$$\begin{cases} \dot{S}_H &= -\beta I_v \frac{S_H}{N_H} \\ \dot{I}_H &= \beta I_v \frac{S_H}{N_H} - \gamma I_H \\ \dot{R}_H &= \gamma I_H \\ \dot{S}_v &= -\alpha S_v \frac{I_H}{N_H} - \mu S_v + \delta N_v^o \\ \dot{I}_v &= \alpha S_v \frac{I_H}{N_H} - \mu I_v \end{cases} \tag{3.1}$$

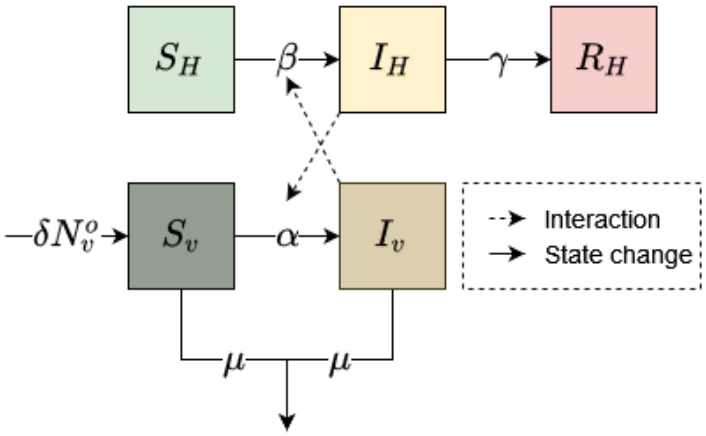


Figure 3: Graphical scheme for the 5-D model.

describes infection of susceptible hosts, S_H, at a rate β through theirinteraction with infected vectors, I_v, while susceptible vectors, S_v, infect at a rate α through interaction with infected hosts I_H. Infected hosts, I_H, move (because they die) to the R_H compartment at a rate γ, while infected vectors stay infected the rest of their life time, mimicking empirical observations (they are not afected negatively by *Xf*). We do not consider natural death of the hosts, as the typical time development of the disease, $1/\gamma$, is much faster than natural death, while vectors die at a rate μ, with a characteristic time $1/\mu$ that is shorter than $1/\gamma$, and we further assume that dead trees are not usually removed from the field, as it happens typically in the Balearic Islands, and thus $N_H(t) = N_H(0) = S_H + I_H + R_H$.

At difference with the model in [16], we do not consider a Exposed compartment, according to empirical observations, as mentioned above. In addition, the constant vector population assumed in [16] is not realistic in our case, as the vector population varies seasonally, and, thus, we introduce a source term for the susceptible vector population, S_v, that is not balanced with the death term as in [16].This term is of the form δN_v^o, where N_v^o defines the scale of the stationary total vector population, what can be seen by adding the last 2 equations in Eq. (3.1),

$$\dot{N}_v = \delta N_v^o - \mu N_v \tag{3.2}$$

where $N_v = S_v + I_v$, so that N_v has a stationary solution (a fixed point) given by $N_v(\infty) = \delta/\mu N_v^o$.

The case $\delta = \mu$ is bit special, as one should use $N_v^o = N_v(0)$ to be in agreement with the standard view [16] that the total vector population is identical to the initial condition (as happens also for the hosts). We notice that using the ansatz δN_v in a direct generalization of [16] would lead to exponential growth, $\dot{N}_v = (\delta - \mu)N_v$, with asymptotic values $N_v(\infty)$ of 0 or ∞, depending on the sign of $\delta - \mu$, what would not be suitable in our study. On the other hand, in the particular case of $\delta = \mu$ defining an asymptotic value $N_v(\infty)$ of N_v^o is meaningless if $N_v^o \neq N_v(0)$ because being equal the birth and death rates, δ and μ, respectively, this implies that N_v is constant along the dynamics (i.e. a conserved quantity). It is also important to recall that the rates of infection for hosts and vectors, β and α will be different, in general. If the biting rate of the vectors and the transmission probability are known experimentally, it could be useful to relate α and β through a balance relation [16].

The model Eq. (3.1) considers that the hosts are trees, so they are fixed on the ground, the moving agents of this system are the vectors, the number of new infected hosts per unit of time will be given by:

$$\beta I_v \frac{S_H}{N_H} \tag{3.3}$$

The transmission rate from vectors to hosts being β. Following an analogous reasoning, the number of new infected vectors per unit of time is:

$$\alpha S_v \frac{I_H}{N_H} \tag{3.4}$$

The conservation of N_H enables an straightforward reduction from this 5-D system to a 4-D one, as one of the equations for the hosts is redundant. In the following sections we are going to disregard the differential equation for $\dot{R}_H$ and we are going to consider the following 4D system:

$$\begin{cases} \dot{S}_H &= -\beta I_v \frac{S_H}{N_H} \\ \dot{I}_H &= \beta I_v \frac{S_H}{N_H} - \gamma I_H \\ \dot{S}_v &= -\alpha S_v \frac{I_H}{N_H} - \mu S_v + \delta N_v^o \\ \dot{I}_v &= \alpha S_v \frac{I_H}{N_H} - \mu I_v \end{cases} \tag{3.5}$$

being $R_H = N_H - S_H - I_H$. Whereas N_H is a conserved quantity of the system, N_v is time dependent in general, as seen in Eq. (3.2). The explicit time dependence being

$$N_v(t) = N_v^o \frac{\delta}{\mu} + \left(N_v(0) - N_v^o \frac{\delta}{\mu} \right) e^{-\mu t} \tag{3.6}$$

For the analysis of the model it is important to recall that the parameters α, β, γ, μ, and δ are non negative. In fact, we will consider α, β, μ and $\gamma \in (0, \infty)$ as $\alpha = \beta = 0$ would mean that the disease does not spread by means of this vector, and a $\mu = 0$, apart from being biologically non realistic, we have verified that it would yield in a non-threshold model where the disease spreads inevitably into the population, contrary of what is used in the literature [5]; finally $\gamma = 0$ would mean that the hosts do not die from the infection, which is also not realistic. For δ we consider it to be $\delta \in [0, \infty)$ and we are going to discuss the cases in which $\delta > 0$ and $\delta = 0$. Regarding S_H, I_H, R_H, S_v and I_v, as represent populations, they are positive, bounded upwards by N_H in the case of the hosts. Regarding the vectors they fulfill that $N_v(t) = I_v + S_v$ and, as μ is a positive non-zero parameter, it follows that $N_v(t) \in [N_v(0), N_v^o \delta / \mu]$.

A preliminary qualitative analysis, inspired by the one done in Section 2 for the SIR model, can be made for this case. Through all this work, we are going to consider that initially $R_H(0) = I_v(0) = 0$ while the initial population of infected hosts is small but non zero, because we study the case of the introduction of the disease via infected hosts, but not via infected vectors (as in Majorca it is thought that the disease was introduced through hosts [1]). Then, regardless of the value of the parameters, when $t = 0$

$$\begin{aligned} \left. \frac{dI_H}{dt} \right|_{t=0} &= -\gamma I_H(0) < 0 \\ \left. \frac{dI_v}{dt} \right|_{t=0} &= \alpha S_v(0) \frac{I_H(0)}{N_H} > 0 \end{aligned} \tag{3.7}$$

So, initially, the number of infected hosts is always a decreasing function of time while the number of infected vectors is an increasing one, at the beginning of the dynamics. Then, as the number of infected vectors will always increase from its initial value, two distinct situations may happen for I_H:

1. $\dot{I}_H = (\beta I_v \frac{S_H}{N_H} - \gamma I_H) < 0 \quad \forall t$ the term γI_H is always bigger than the term $\beta I_v \frac{S_H}{N_H}$. In this case, the number of infected hosts always decreases and no epidemic occurs.
2. $\dot{I}_H = (\beta I_v \frac{S_H}{N_H} - \gamma I_H) > 0 \quad t \in [t_1, t_2]$. In this case, there is an interval of times so that $\beta I_v \frac{S_H}{N_H} > \gamma I_H$. This is the case when there is an epidemic.

Note that, contrary of what happens for the SIR model, the condition that determines if an epidemic takes place, $\dot{I}_H > 0$, is not at $t = 0$. Thus, it may happen that the peak value of $I_H^{max} < I_H(0)$, but we still consider that an epidemic has taken place in this situation, as we define the threshold for epidemic the change in behavior of I_H, form a monotonically decreasing function $\dot{I}_H < 0 \; \forall t$, to a non monotonical one, $\exists\, t : \dot{I}_H > 0$. This is going to be the criterion used in Section 4 to determine whether or not an epidemic has occurred. As per the case of the infected vectors, note that there will always be an initial increase. We do not consider this behavior in order to define the epidemic as we are interested in the state of the hosts.

3.2.1 Fixed points

When dealing with any dynamical system, one of the first questions is to find its fixed points and their stability. The nullclines for this system are:

$$\begin{cases} \dot{S}_H &= -\beta I_v \frac{S_H}{N_H} = 0 \Rightarrow \begin{cases} I_v &= 0 \\ S_H &= 0 \end{cases} \\ \dot{I}_H &= \beta I_v \frac{S_H}{N_H} - \gamma I_H = 0 \Rightarrow \begin{cases} I_v, I_H &= 0 \\ S_H, I_H &= 0 \\ I_H &= \frac{\beta}{\gamma} I_v \frac{S_H}{N_H} \end{cases} \\ \dot{S}_v &= -\alpha S_v \frac{I_H}{N_H} - \mu S_v + \delta = 0 \Rightarrow S_v = \frac{N_v^o \delta}{\alpha \frac{I_H}{N_H} + \mu} \\ \dot{I}_v &= \alpha S_v \frac{I_H}{N_H} - \mu I_v = 0 \Rightarrow \begin{cases} S_v, I_v &= 0 \\ I_H, I_v &= 0 \\ I_H &= \frac{\mu}{\alpha} I_v \frac{N_H}{S_H} \end{cases} \end{cases} \tag{3.8}$$

From the analysis of the nullclines one can determine the fixed points of the system, by finding their intersection. From the nullclines of S_H and I_H it is straightforward to see that in the fixed point $I_H = 0$ and therefore from $\dot{I}_v = 0$ it implies that $I_v = 0$. So the fixed point of the system is:

$$(S_H = S_H^*, I_H = 0, S_v = N_v^o \frac{\delta}{\mu}, I_v = 0) \tag{3.9}$$

From Eq. (3.9) we see that in the host subspace, the line $(S_H, I_H = 0)$ is a line of non-isolated fixed points, as there is no constrain for the value of S_H^*. In the following sections we would be interested primarily in the point when $S_H = N_H$, that is, the pre-epidemic stage. We are going to focus in two cases depending on the values of δ. The first scenario when $\delta > 0$ while in the second one we analyze the extreme case in which $\delta = 0$.

3.3 General case, $\delta > 0$:

3.3.1 Linear stability analysis

The goal of this analysis is to determine the conditions for which the epidemic model exhibits a threshold. This is customary done for a Disease Free Equilibrium (DFE) before the epidemic starts (i.e. pre-epidemic) where there are no infected individuals in the system, $I_v = I_H = 0$. However, one can easily notice that any pre-epidemic state is not, in general, an equilibrium (fixed point) of the model Eq. (3.5). We can easily see this from the expression of $\dot{N}_v$ Eq. (3.2). $\dot{N}_v = 0$ at $t = 0$ only if $N_v(0) = S_v(0) = N_v^o \delta/\mu$, where the identity at $t = 0$ of N_v and S_v results from the initial condition $I_v(0) = 0$. This means that in this system one can only deduce the position of the epidemic threshold if the initial condition is chosen exactly as $S_v(0) = N_v^o \delta/\mu$, that is, if the vector population is in its stationary state [24]. Then the DFE is defined by $(S_H = N_H, I_H = 0, S_v = N_v^o \delta/\mu, I_v = 0)$. Notice that this discussion has a parallel to the one we are going to present in Section 3.3.2 on the application of the NGM method to this model, as the NGM is nothing more that a (clever) method to obtain $\mathcal{R}_0$ directly from the analysis of the DFE, so, we stress that a linear stability analysis

of the DFE (and also the obtention of $\mathcal{R}_0$ from the NGM in Sec. Section 3.3.2) is only possible for the particular initial condition $N_v(0) = \delta/\mu N_v^o$. For the linear stability analysis, the Jacobian matrix of the system is computed

$$J = \begin{pmatrix} -\beta\frac{I_v}{N_H} & 0 & 0 & -\beta\frac{S_H}{N_H} \\ \beta\frac{I_v}{N_H} & -\gamma & 0 & \beta\frac{S_H}{N_H} \\ 0 & -\alpha\frac{S_v}{N_H} & -\alpha\frac{I_H}{N_H} - \mu & 0 \\ 0 & \alpha\frac{S_v}{N_H} & \alpha\frac{I_H}{N_H} & -\mu \end{pmatrix} \tag{3.10}$$

Which, evaluated at the DFE:

$$J|_{DFE} = \begin{pmatrix} 0 & 0 & 0 & -\beta \\ 0 & -\gamma & 0 & \beta \\ 0 & -\alpha\frac{N_v^o}{N_H}\frac{\delta}{\mu} & -\mu & 0 \\ 0 & \alpha\frac{N_v^o}{N_H}\frac{\delta}{\mu} & 0 & -\mu \end{pmatrix} \tag{3.11}$$

so that the eigenvalues of this matrix are:

$$\begin{aligned} det(J|_{DFE} - \lambda\mathbb{I}) = -\lambda\Big[-(\mu+\lambda)^2(\gamma+\lambda) + (\mu+\lambda)\beta\alpha\frac{N_v^o}{N_H}\frac{\delta}{\mu}\Big] = 0 \Rightarrow \\ \lambda_0 = 0 \\ \lambda_\mu = -\mu \\ \lambda_\pm = -\frac{(\gamma+\mu)}{2} \pm \frac{1}{2}\sqrt{(\gamma-\mu)^2 + 4\beta\alpha\frac{N_v^o}{N_H}\frac{\delta}{\mu}} \end{aligned} \tag{3.12}$$

It is straightforward to see that all eigenvalues are real, as all parameters are positive, then, we need to determine whether they are positive or negative reals. $\lambda_\mu = -\mu < 0$ as μ is defined positive, so in order to discuss the stability of this fixed point, we need to study the $\lambda_\pm$ eigenvalues. λ_- is always negative, but λ_+ changes sign depending on the values of the parameters. The threshold condition $\lambda_+ = 0$ leads to:

$$\lambda_+ = 0 \;\Rightarrow\; \frac{\beta\alpha}{\gamma\mu}\frac{N_v^o}{N_H}\frac{\delta}{\mu} = 1 \tag{3.13}$$

So, for $\frac{\beta\alpha}{\gamma\mu}\frac{N_v^o}{N_H}\frac{\delta}{\mu} < 1 \Rightarrow \lambda_+ < 0$ the fixed point is stable and for $\frac{\beta\alpha}{\gamma\mu}\frac{N_v^o}{N_H}\frac{\delta}{\mu} > 1 \Rightarrow \lambda_+ > 0$ a perturbation will grow in the direction of the eigenvector associated to λ_+. In Section 3.3.2 we are going to show that this threshold retrieved by means of a linear stability analysis is in fact the basic reproduction number $\mathcal{R}_0$.

3.3.2 Basic reproduction number

As we have discussed on previous sections, the computation of $\mathcal{R}_0$ is key when dealing with an epidemiological situation. For the case of vector transmitted diseases, the definition of $\mathcal{R}_0$ being the number of secondary infections caused by introducing a single infected individual into a susceptible population is ambiguous. The spread from host to vector to host can be understood as a two step process, or it can be considered as one generation. In our case, following [16], we consider the spread from host to vector to host as being one generation. This choice is motivated by the fact that we understand the vector as a means of transmitting the disease, but the main object of interest are the hosts. An analysis of the case when the spread from host to vector to host is understood as a two step process is done in Appendix B.

In this section we present the determination of $\mathcal{R}_0$ using the next generation matrix approach. This method is connected to the linear stability analysis of the system as presented in Section 3.3.1, as it relays on the linearization of the system about the DFE. For the construction of the NGM, the linearized infection subsystem is used, that is, the subset of equations describing the production of new infected and changes in the states of already existing infected, in this case, the equations for $\dot{I}_H$ and $\dot{I}_v$ from Eq. (3.5). These equations are taken linearized by the DFE state, which has to be an equilibrium state of the system with no presence of infected. Again, the pre-epidemic state has to be considered, $S_H = N_H$ for the hosts, and the stationary value for the susceptible vectors, $S_v(0) = N_v^o \frac{\delta}{\mu}$, as discussed in the previous section. From these linearized equations, all epidemic events that lead to new infections are incorporated in the transmission matrix F, which includes the production of new infections (F_{ij}: being the rate at which individuals in infected state j give raise to individuals in infected state i). All the other events are included in the transition part via the V matrix, which describes changes in state. The NGM K is then $K \equiv FV^{-1}$ (see Appendix A for a more in depth analysis on the construction of the NGM). The NGM for this system is:

$$K = FV^{-1} = \begin{pmatrix} \frac{\beta\alpha}{\gamma\mu}\frac{N_v^o}{N_H}\frac{\delta}{\mu} & \frac{\beta}{\mu} \\ 0 & 0 \end{pmatrix} \tag{3.14}$$

With:

$$F = \begin{pmatrix} 0 & \beta\frac{N_H}{N_H} \\ 0 & 0 \end{pmatrix} \qquad V = \begin{pmatrix} \gamma & 0 \\ -\alpha\frac{N_v^o}{N_H}\frac{\delta}{\mu} & \mu \end{pmatrix} \Rightarrow V^{-1} = \begin{pmatrix} \frac{1}{\gamma} & 0 \\ \frac{\alpha}{\gamma\mu}\frac{N_v^o}{N_H}\frac{\delta}{\mu} & \frac{1}{\mu} \end{pmatrix} \tag{3.15}$$

The basic reproduction number $\mathcal{R}_0$ is the spectral radius of this matrix, so:

$$det(K - \sigma\mathbb{I}) = 0 \implies \begin{vmatrix} \frac{\beta\alpha}{\gamma\mu}\frac{N_v^o}{N_H}\frac{\delta}{\mu} - \sigma & \frac{\beta}{\mu} \\ 0 & -\sigma \end{vmatrix} = (-\sigma)\left(\frac{\beta\alpha}{\gamma\mu}\frac{N_v^o}{N_H}\frac{\delta}{\mu} - \sigma\right) = 0 \implies$$

$$\sigma = \frac{\beta\alpha}{\gamma\mu}\frac{N_v^o}{N_H}; \quad \sigma = 0 \tag{3.16}$$

Therefore, the basic reproduction number $\mathcal{R}_0$:

$$\mathcal{R}_0 = \frac{\beta\alpha}{\gamma\mu}\frac{N_v^o}{N_H}\frac{\delta}{\mu} \tag{3.17}$$

We then recover the same threshold than in Eq. (3.13), for $\frac{\beta\alpha}{\gamma\mu}\frac{N_v^o}{N_H}\frac{\delta}{\mu} > 1$ there will be initial exponential growth, and for $\frac{\beta\alpha}{\gamma\mu}\frac{N_v^o}{N_H}\frac{\delta}{\mu} < 1$ the epidemic will die out. It is worth noting that for $\delta = 0$ Eq. (3.17) does not make sense, even if it is formally valid, as the population of vectors, at the stationary state is identically zero. Had we considered a more general case instead of the pre-epidemic state, for a generic $S_H(0)$, i.e. if $S_H(0) \neq N_H$, the basic reproduction number would have been:

$$\mathcal{R}_0^{'} = \frac{\beta\alpha}{\gamma\mu}\frac{N_v^o}{N_H}\frac{\delta}{\mu}\frac{S_H(0)}{N_H} \tag{3.18}$$

This expression of $\mathcal{R}_0$ would be more suitable than Eq. (3.17) in the eventuality that there are already some removed hosts in the pre-epidemic state. Nevertheless, as the usual procedure is to consider all population susceptible before the epidemic, the expression of $\mathcal{R}_0$ that we are going to use during this thesis is Eq. (3.17).

3.3.3 Fast-slow approximation

In general, 4-D systems as the one in Eq. (3.5) are difficult to analyze, for the high dimensionality of their phase space. Then, in this section a reduction from 4-D to 3-D considering a separation of time scales is proposed. In the fast-slow reduction we consider that the death rate of the vectors, μ, is a large parameter. In biological terms, it makes sense as an adult vector dies from natural death in less than half a year, and dies even faster if predators, human intervention or adverse climate conditions are taken into account. Then, from $\dot{I}_v$ of Eq. (3.5) one can write:

$$\epsilon \dot{I}_v = \frac{\alpha}{\mu} S_v \frac{I_H}{N_H} - I_v \tag{3.19}$$

Being $\epsilon = 1/\mu \to 0$ an small parameter. Then, in the case in which ϵ can be disregarded in front of the other parameters, and so while α/μ remains finite one gets

$$I_v \approx \frac{\alpha}{\mu} S_v \frac{I_H}{N_H} \tag{3.20}$$

Which leads to the following expression for the 3-D system:

$$\begin{cases} \dot{S}_H &= -\beta \frac{\alpha}{\mu} S_v \frac{I_H}{N_H} \frac{S_H}{N_H} \\ \dot{I}_H &= \beta \frac{\alpha}{\mu} S_v \frac{I_H}{N_H} \frac{S_H}{N_H} - \gamma I_H \\ \dot{S}_v &= -\alpha S_v \frac{I_H}{N_H} - \mu S_v + \delta N_v^o \end{cases} \tag{3.21}$$

It is worth noting that now, with the equations written as in Eq. (3.21) one can analyze $\dot{I}_H$ at $t = 0$ like in the SIR model, as done in Section 2:

$$\left.\frac{dI_H}{dt}\right|_{t=0} = I_H(0)\left(\beta \frac{\alpha}{\mu} \frac{S_v(0)}{N_H} \frac{S_H(0)}{N_H} - \gamma\right) \tag{3.22}$$

It is straightforward to see that when the system is at the DFE at $t = 0$ we recover the expression for $\mathcal{R}_0$ in Eq. (3.17). For the approximated system then, the condition for an epidemic outbreak is given at the beginning of the dynamics not after some time t_1 as discussed for the complete model. So, at the beginning of the dynamics, there will always be some discrepancy between the approximation and the complete system. This is so because, when performing this reduction we have substituted a differential equation for an algebraic expression of one of the variables. Now, the ODEs are also known as problems of initial values, then when we have done the approximation Eq. (3.20) we have lost one of the initial conditions of the system. In fact, we may notice that this expression for I_v does not allow for the initial condition $I_v(0) = 0$, as Eq. (3.20) needs either $S_v(0) = 0$ or $I_H(0) = 0$. Then, we conclude that we can expect the approximation to be good after an initial transient of the dynamics.

Eq. (3.21) can be further reduced to a 2-D system via the explicit expression for $N_v(t)$ Eq. (3.6), leading to equations only for hosts. From an experimental point of view this reduction is quite useful as the epidemiological state of the vectors is not easy to obtain. So writing S_v in terms of $N_v(t)$ and I_v one arrives to

$$\begin{cases} \dot{S}_H &= -\beta \frac{\alpha N_v(t)}{\mu N_H + \alpha I_H} S_H \frac{I_H}{N_H} = -\zeta(t) S_H \frac{I_H}{N_H} \\ \dot{I}_H &= \beta \frac{\alpha N_v(t)}{\mu N_H + \alpha I_H} S_H \frac{I_H}{N_H} - \gamma I_H = \zeta(t) S_H \frac{I_H}{N_H} - \gamma I_H \end{cases} \tag{3.23}$$

with the explicit expressions for S_v and I_v being

$$S_v = N_v(t) \frac{\mu N_H}{\alpha I_H + \mu N_H}, \quad I_v = N_v(t) \frac{\alpha I_H}{\alpha I_H + \mu N_H} \tag{3.24}$$

The expression in Eq. Eq. (3.23) is quite remarkable as it resembles of the equations for an standard SIR model, for the host population, but with a coefficient $\zeta(t) \equiv \beta\alpha N_v(t)/(\mu N_H + \alpha I_H)$ that depends on time via $N_v(t)$ and $I_H(t)$. A SIR model with time-dependent coefficients has been discussed in [25], where the authors indicate how to obtain the epidemic thresholds, i.e, how to define the right $\mathcal{R}_0$ that predicts them. In the particular case that the initial number of vectors is already the stationary value of $N_v(t)$: $N_v(0) = N_v^o \frac{\delta}{\mu}$ the explicit time dependence disappears and Eq. Eq. (3.23) is then an autonomous system, and the "usual techniques" can be applied. It is clear that this fact mirrors our discussions in Section 3.3.1 in which we discussed that the epidemic thresholds can only be obtained from the pre-epidemic DFE, i.e. if the initial condition for the vectors is already in the steady state, $S_v(0) = N_v^o \frac{\delta}{\mu}$, and in Section 3.3.2 when we discussed that the NGM can only be applied in this very same situation, for which [25] become the standard, autonomous, SIR method, in which $\mathcal{R}_0$ can be obtained in the usual way. This calls [25] for a generalization of the NGM method, obtained by solving a functional equation, instead of the usual NGM matrix method of Appendix A if one wishes to study generic initial conditions $S_v(0)$. Thus, we can state that $\mathcal{R}_0$ will capture the threshold behavior of the system when the initial population of vectors is equal o close its stationary state. For other initial conditions it is not guaranteed that $\mathcal{R}_0$, as expressed in Eq. (3.23) captures the behavior of the system. In Section 4 we further discuss this point, with numerical simulations.

3.4 Particular case $\delta = 0$

The case in which $\delta = 0$ is an special case of this model, as in the fixed point the population of vectors is zero ($S_H = S_H^*, I_H = 0, S_v = 0, I_v = 0$). Via the usual methods of computing $\mathcal{R}_0$ with the NGM approach at an equilibrium pre-epidemic state, or via a linear stability analysis, one gets that this situation is a non-threshold phenomenon where the disease cannot spread. This is so as, in the equilibrium state of the system, the number of vectors is zero and so there is no way for the disease to spread. Nevertheless, numerically we see that this analysis of the fixed point is not sufficient to determine whether or not an epidemic will occur, as we find situations in which there is an outbreak. The usual NGM approach as presented in Appendix A fails in this case because it assumes a pre-epidemic disease-free state to be an stable equilibrium state of the system, which, in the case of $\delta = 0$, it can only be the state with no vectors. Then, other tools different form the ones presented in this work are needed in order to study this case.

Even if $\mathcal{R}_0$ could not be determined analytically by means of the NGM method used in Section 3.3.2, the analysis of this particular case is also interesting. For the case $\delta = 0$ the system has a non trivial conserved quantity, which can be derived from Eq. (3.5) starting from the equation for $\dot{S}_v$ and the relation between $\dot{S}_v$ and $\dot{I}_v$:

$$\begin{cases} \dot{S}_v = -\alpha S_v \frac{I_H}{N_H} - \mu S_v \implies I_H = -N_H \frac{\dot{S}_v}{\alpha S_v} - N_H \frac{\mu}{\alpha} \\ \dot{I}_v + \dot{S}_v = -\mu(I_v + S_v) \implies -\mu = \frac{\dot{I}_v + \dot{S}_v}{I_v + S_v} = \frac{d}{dt}(\ln(S_v + Iv)) \end{cases}$$

Now if we sum the equations for $\dot{I_H}$ and $\dot{S_H}$:

$$\begin{aligned} &\dot{S}_H + \dot{I}_H = -\gamma I_H = \frac{\gamma}{\alpha} N_H \frac{\dot{S}_v}{S_v} + \frac{\mu}{\alpha}\gamma N_H = \frac{\gamma}{\alpha} N_H \frac{\dot{S}_v}{S_v} - \frac{\gamma}{\alpha} N_H \frac{\dot{I}_v + \dot{S}_v}{I_v + S_v} \implies \\ &\dot{S}_H + \dot{I}_H - \frac{\gamma}{\alpha} N_H \frac{d}{dt}(\ln(S_v)) + \frac{\gamma}{\alpha} N_H \frac{d}{dt}(\ln(S_v + Iv)) = 0 \implies \\ &\frac{d}{dt}\left(S_H + I_H + \frac{\gamma}{\alpha} N_H \ln\left(1 + \frac{I_v}{S_v}\right)\right) = 0 \implies \\ &S_H + I_H + \frac{\gamma}{\alpha} N_H \ln\left(1 + \frac{I_v}{S_v}\right) = S_H(0) + I_H(0) + \frac{\gamma}{\alpha} N_H \ln\left(1 + \frac{I_v(0)}{S_v(0)}\right) \end{aligned}$$

We will name this conserved quantity as C:

$$C \equiv S_H + I_H + \frac{\gamma}{\alpha} N_H \ln\left(1 + \frac{I_v}{S_v}\right) \tag{3.25}$$

For the usual initial conditions used in this thesis, $S_H(0) = N_H$ and $I_H(0) = I_v(0) = 0$, one can relate C to R_H:

$$\begin{cases} N_H &= S_H + I_H + R_H \\ C &= S_H(0) + I_H(0) + \frac{\gamma}{\alpha} N_H \ln\left(1 + \frac{I_v(0)}{S_v(0)}\right) = N_H \;\Rightarrow\; N_H = S_H + I_H + \frac{\gamma}{\alpha} N_H \ln\left(1 + \frac{I_v}{S_v}\right) \end{cases}$$
$$\Rightarrow\; R_H = \frac{\gamma}{\alpha} N_H \ln\left(1 + \frac{I_v}{S_v}\right) \tag{3.26}$$

Now, the presence of this extra conserved quantity enables for an exact reduction of the dimensionality of the system. From the expression of C one can write:

$$1 + \frac{I_v}{S_v} = e^{\frac{\alpha}{\gamma N_H}(C - S_H - I_H)} \;\Rightarrow\; I_v = S_v\left(e^{\frac{\alpha}{\gamma N_H}(C - S_H - I_H)} - 1\right) \tag{3.27}$$

Thus,

$$\begin{cases} \dot{S}_H &= -\beta \frac{S_H}{N_H} S_v\left(e^{\frac{\alpha}{\gamma N_H}(C - S_H - I_H)} - 1\right) \\ \dot{I}_H &= \beta \frac{S_H}{N_H} S_v\left(e^{\frac{\alpha}{\gamma N_H}(C - S_H - I_H)} - 1\right) - \gamma I_H \\ \dot{S}_v &= -\alpha S_v \frac{I_H}{N_H} - \mu S_v \end{cases} \tag{3.28}$$

One could use again the expression for $N_v(t)$ Eq. (3.6) in order to further reduce the system, to a 2-D non-autonomous one:

$$\begin{cases} \dot{S}_H &= -\beta S_H \frac{N_v(0)e^{-\mu t}}{N_H}\left(1 - e^{-\frac{\alpha}{\gamma N_H}(C - S_H - I_H)}\right) \\ \dot{I}_H &= \beta S_H \frac{N_v(0)e^{-\mu t}}{N_H}\left(1 - e^{-\frac{\alpha}{\gamma N_H}(C - S_H - I_H)}\right) - \gamma I_H \end{cases} \tag{3.29}$$

with

$$I_v = N_v(0)e^{-\mu t}\left(1 - e^{-\frac{\alpha}{\gamma N_H}(C - S_H - I_H)}\right), \qquad S_v = N_v(0)e^{-\mu t}e^{-\frac{\alpha}{\gamma N_H}(C - S_H - I_H)} \tag{3.30}$$

Note that this exact reduction is only valid when $\delta = 0$ as the conserved quantity C is not exact for the general case $\delta \neq 0$. Thus, for the case $\delta = 0$ the system allows for two reductions, one exact using the conserved quantity C and the other via the fast-slow approximation presented in Section 3.3.3. Particularizing Eq. (3.23) for $\delta = 0$:

$$\begin{cases} \dot{S}_H &= -\beta \frac{\alpha N_v(0)e^{-\mu t}}{\mu N_H + \alpha I_H} S_H \frac{I_H}{N_H} \\ \dot{I}_H &= \beta \frac{\alpha N_v(0)e^{-\mu t}}{\mu N_H + \alpha I_H} S_H \frac{I_H}{N_H} - \gamma I_H \end{cases} \tag{3.31}$$

Comparing Eq. (3.29) and Eq. (3.31) one sees that, for $\delta = 0$ and in the regime where the fast slow approximation is valid, the following relation must be fulfilled

$$\left(1 - e^{-\frac{\alpha}{\gamma N_H}(C - S_H - I_H)}\right) = \frac{\alpha I_H}{\mu N_H + \alpha I_H} \tag{3.32}$$

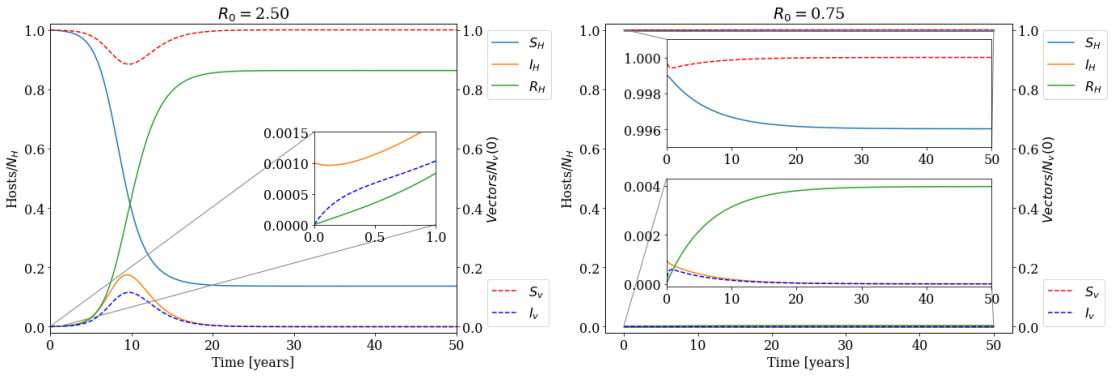
4 Numerical study of the model

4.1 General case $\delta > 0$

We start the numerical analysis by illustrating the validity of the threshold value found analitically by means of numerically solving the system of differential equations of Eq. (3.5). We expect an outbreak of the epidemic when $\mathcal{R}_0 > 1$, that is the existence of a peak value of I_H, and a nonexistence of epidemic for $\mathcal{R}_0 < 1$, I_H showing a monotonically decreasing behavior. In order to illustrate this phenomenon, the following parameters and initial conditions are used:

$$N_H = 100, \quad N_v^o = 1000 \quad S_v(0) = N_v^o \frac{\delta}{\mu}, \quad I_H(0) = 0.001 N_H, \quad I_v(0) = R_H(0) = 0, \quad \mu = 4,$$
$$\gamma = 0.7, \quad \alpha = 0.75\mu \quad \text{and} \quad \delta = \mu \tag{4.1}$$

Then, $\beta = 0.233$, $\beta = 0.07$ so that $\mathcal{R}_0 = 2.5$, $\mathcal{R}_0 = 0.75$ respectively. Note that these initial conditions refer to the stationary state of the vector population.



(a) (b)

Figure 4: Different realizations of the model. The parameters used are $N_H = 100$, $N_v^o = 1000$, $\mu = 4$, $\gamma = 0.7$, $\alpha = 0.75\mu$ and $\delta = \mu$ and the initial conditions $S_v(0) = N_v^o\delta/\mu$, $I_H(0) = 0.001N_H$, $I_v(0) = R_H(0) = 0$. (a) above the threshold with $\beta = 0.233$ so that $\mathcal{R}_0 = 2.5$; (b) below the threshold with $\beta = 0.07$ so that $\mathcal{R}_0 = 0.75$.

Fig. 4 illustrates the theoretical results for the threshold value. In Section 3 we have characterized the threshold by the change in the evolution of I_H, below it I_H has a monotonically decreasing behavior Fig. 4(b), while above threshold, there is an initial decrease followed by an increase of the infected host population Fig. 4(a). We have checked that both panels in Fig. 4 yield the same results if $\delta > \mu$ and $\delta < \mu$ for the same values of $\mathcal{R}_0$ (β was varied acordingly). Next, in order to give a more quantitative approach for the threshold validation of the system, we study the behavior of the system by varying every parameter independently. In order to do so, from the expression of $\mathcal{R}_0$ in Eq. (3.17) we derive an expression of each of the parameters at the critical value $\mathcal{R}_0 = 1$:

$$\alpha_c = \frac{\gamma\mu^2}{\beta\delta}\frac{N_H}{N_v^o}, \quad \beta_c = \frac{\gamma\mu^2}{\alpha\delta}\frac{N_H}{N_v^o}, \quad \gamma_c = \frac{\alpha\beta\delta}{\mu^2}\frac{N_v^o}{N_H}, \quad \mu_c = \sqrt{\frac{\alpha\beta\delta}{\gamma}\frac{N_v}{N_H}}, \quad \delta_c = \frac{\gamma\mu^2}{\beta\alpha}\frac{N_H}{N_v^o} \tag{4.2}$$

Then, setting $N_H = N_v^o = 100$ and $\alpha = \beta = \gamma = \mu = \delta = 1$ the critical parameters become $\alpha_c = \beta_c = \gamma_c = \mu_c = \delta_c = 1$. The strategy followed to numerically validate the critical parameters is to focus at one parameter at the time and vary it, and check if I_H^{max} exists. Then, the first parameter

value for which $I_H^{max} = 0$ is considered the critical value of such parameter. A consideration to bear in mind when doing this analysis is that the initial population of vectors needs to be near its stationary value $N_v = N_v^o \delta/\mu$ for $\mathcal{R}_0$ to be meaningful, as discussed previously in Section 3 and as we are going to show later in this section. For this reason, when either δ of μ are varied the initial condition $N_v(0)$ will vary accordingly to start in the stationary value of $N_v(\infty) = N_v^o \delta/\mu$. The results of this analysis are presented in Fig. 5. We see that the numerical results obtained for each parameter are in good agreement with the theoretical result expected, each with a relative error of 0.5%.

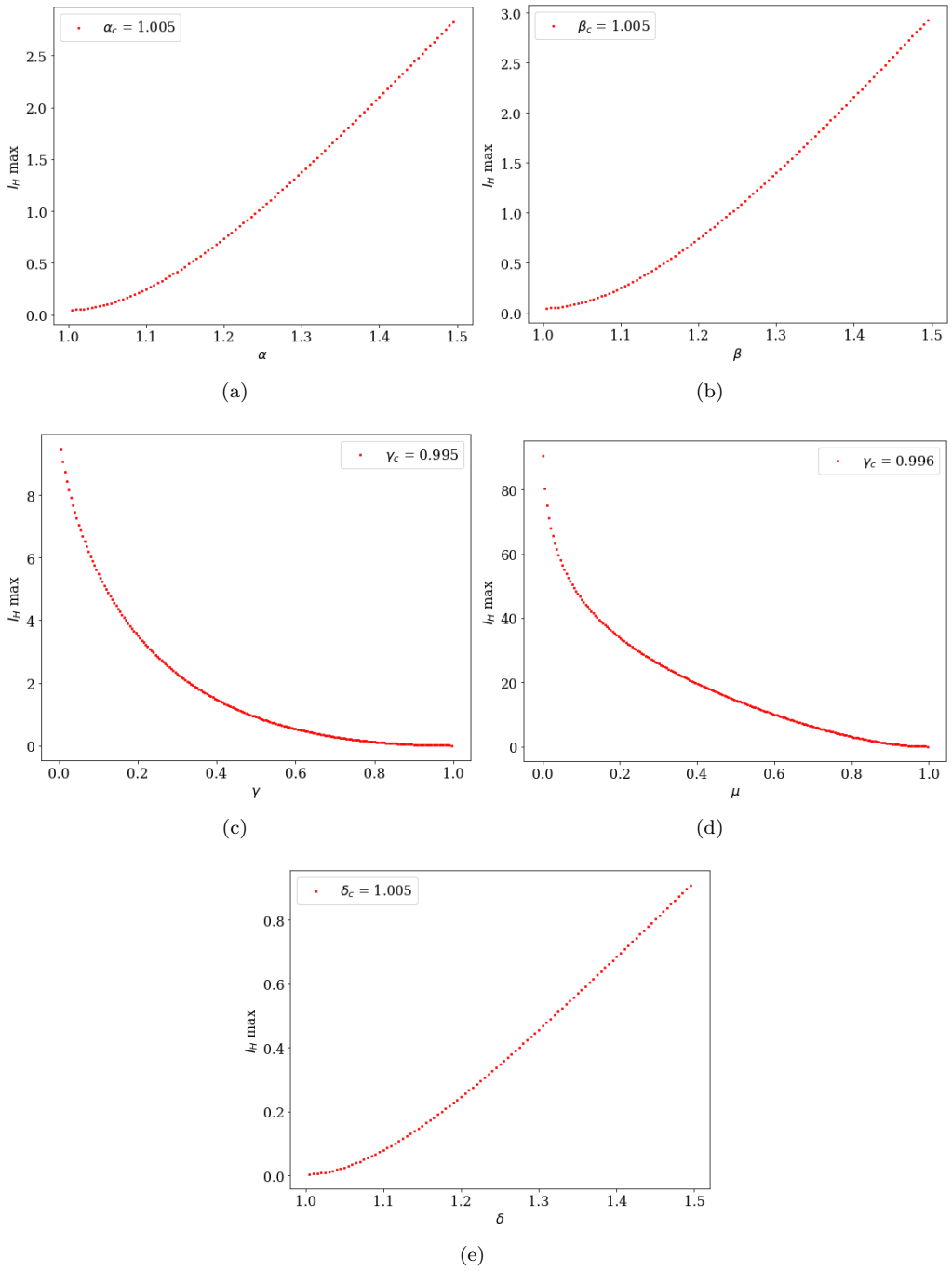


Figure 5: Numerical determination of the critical parameters. For $N_H = N_v^o = 100$ and $\alpha = \beta = \gamma = \mu = \delta = 1$ so that the critical parameters become $\alpha_c = \beta_c = \gamma_c = \mu_c = \delta_c = 1$. The initial conditions used being $I_H(0) = 0.001 N_H$, $I_v(0) = R_H(0) = 0$ and $S_v(0) = N_v^o \delta/\mu$.

As a part of this parameter analysis, we analyze explicitly the effect of having distinct values for μ and δ. To this aim, we present a validation of $\mathcal{R}_0$, using the similar strategy than we did in Fig. 5. We fix all parameters but β and δ. For every value of δ, we vary β and compute the time evolution of the system. Then, we check whether or not I_H is monotonically decreasing, and, for the first β that there is a change in behavior, we compute the $\mathcal{R}_0$ associated to it. We repeat this procedure for three distinct values of μ. The values used for this computation are $\gamma = \alpha = 1$, $N_H = N_v^o = 100$. From the three first panels of Fig. 6 we see that the results are consistent for the different values of μ and that the $\mathcal{R}_0$ obtained is in good agreement with the theoretical one Eq. (3.17), with a relative error of less than 1% for every panel.

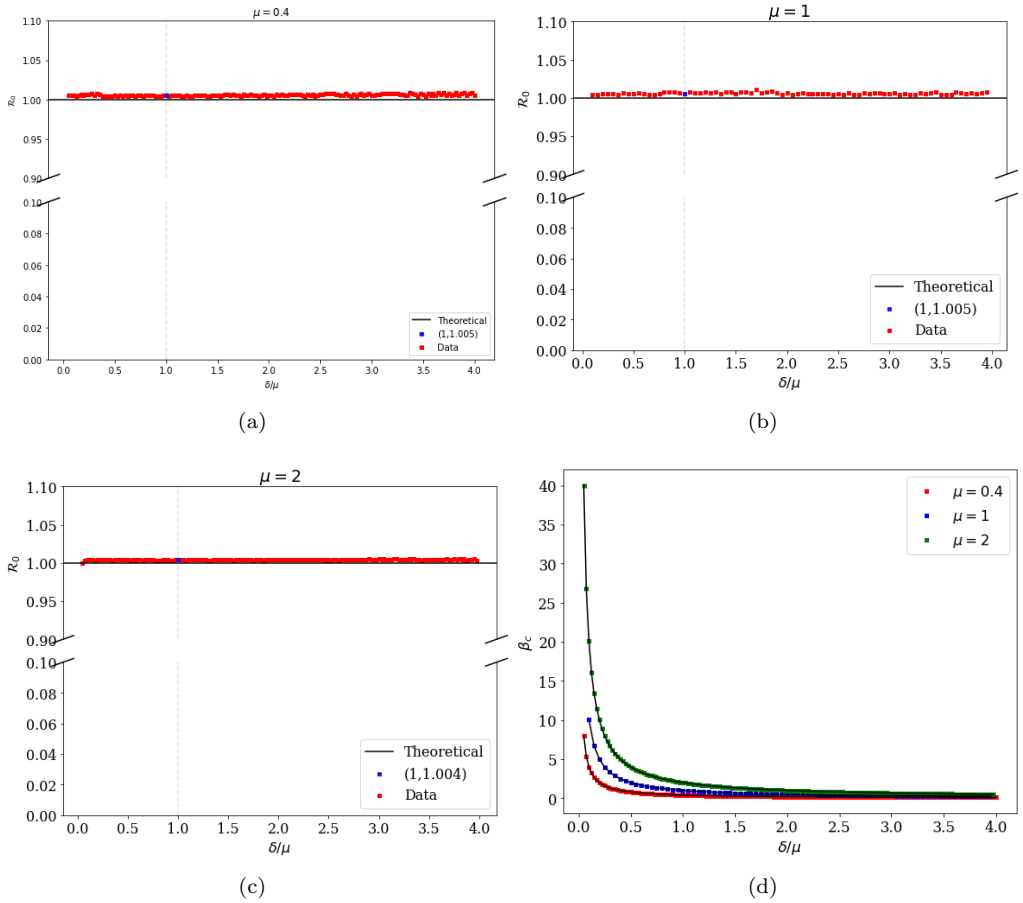
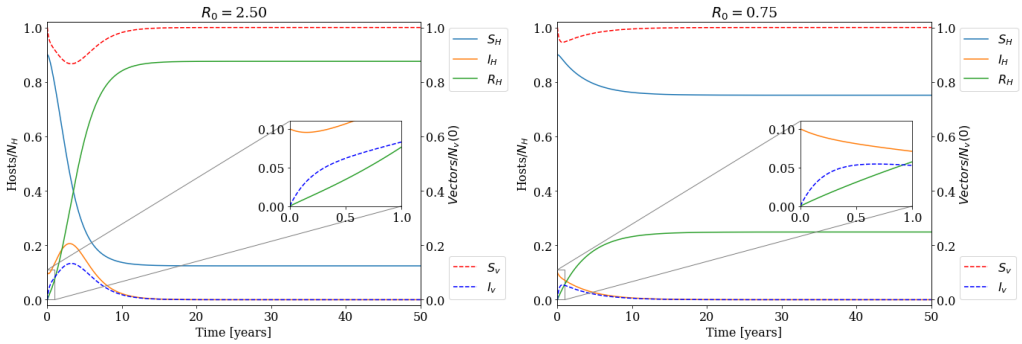


Figure 6: Numerical check of the threshold for $\delta \neq \mu$ where β is varied and $\gamma = \alpha = 1$, $N_H = N_v^o = 100$ and the initial conditions used being $I_H(0) = 0.001N_H$, $I_v(0) = R_H(0) = 0$ and $S_v(0) = N_v^o\delta/\mu$. (a), (b) and (c) are different realizations for $\mu = 0.4$, 1 and 2 respectively. (d) represents the value of β_c of the previous figures.

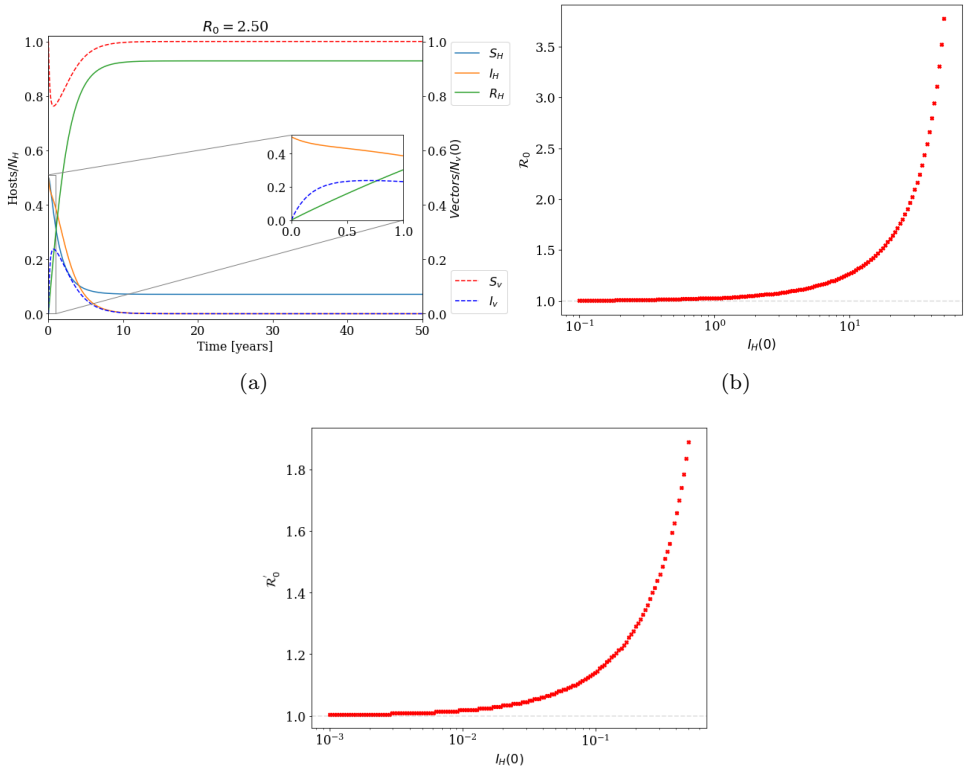
Along Section 3 we have stressed the importance of the DFE for the computation of $\mathcal{R}_0$. So, next, we are going to numerically check the relevance of the initial condition of hosts and vectors. We start by repeating Fig. 4 for a non negligible population of infected hosts $I_H(0) = 0.1N_H$ as initial condition. From Fig. 7 one may note that for this case, where $I_H(0) = 0.1N_H$, there is a non negligible number of removed individuals even if $R_0 < 1$, so there is a quantitative difference between both cases, $I_H(0) = 0.001N_H$ and $I_H(0) = 0.1N_H$. However, the qualitative behavior of I_H during the time evolution is the same, monotonically decreasing below the threshold, and a peak above it. For larger values of $I_H(0)$ the qualitative behavior of I_H, and thus the threshold behavior of the system, may not longer be well represented by $\mathcal{R}_0$ as computed in Eq. (3.17). This is so because the basic reproduction number has been computed linearizing by the DFE, at which the number of infected infected hosts is zero. Therefore, when increasing $I_H(0)$, one is getting farther

from such equilibrium and so, the linearization may no longer be valid. As an example, for the case $I_H(0) = 0.5N_H$, which is displayed in Fig. 8(a), it can be seen that no peak appears for I_H even if $\mathcal{R}_0 > 1$ (it has been numerically checked that there is not a peak that the figure resolution could be hiding).



(a) (b)

Figure 7: Different realizations of the model for $I_H(0) = 0.1N_H$ and the rest of the parameters and initial conditions used are the same than Fig. 4. (a) above the threshold with $\beta = 0.233$ so that $R_0 = 2.5$; (b) below the threshold with $\beta = 0.07$ so that $R_0 = 0.75$.



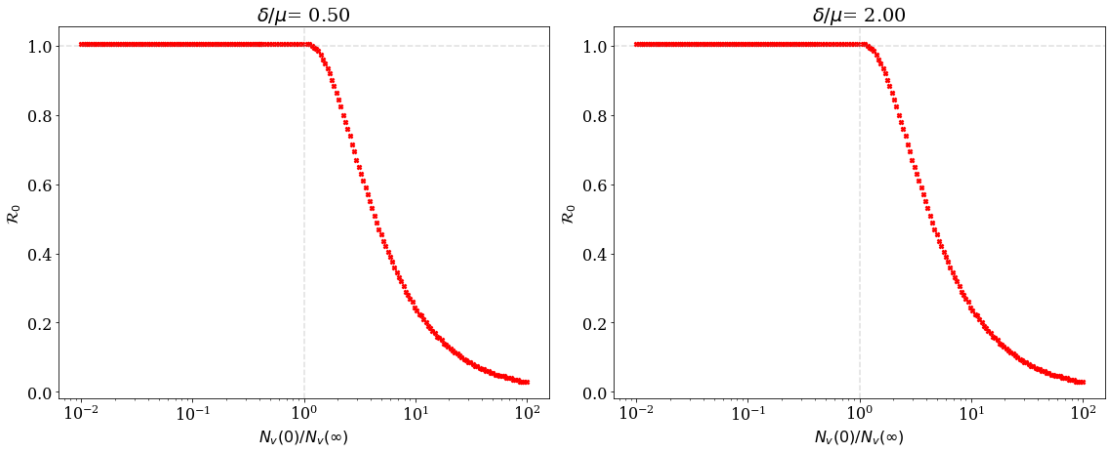
(a) (b)

(c)

Figure 8: Effect of the initial condition $I_H(0)$. All parameters used are the same than for Fig. 4 unless the contrary is specified. (a) realizations of the model for $I_H(0) = 0.5N_H$ with $\beta = 0.233$ so that $R_0 = 2.5$; (b) log-lin plot of the effective $\mathcal{R}_0$ vs $I_H(0)$, where β has been varied to obtain the different values of $\mathcal{R}_0$. (c) log-lin plot of the effective $\mathcal{R}_0'$ vs $I_H(0)$, where β has been varied to obtain the different values of $\mathcal{R}_0'$ as defined in Eq. (3.17).

In order to illustrate the previous statement we have done realizations of the model for $I_H(0) \in [0.001, 0.5]N_H$ and we have computed the value of $\mathcal{R}_0$ as it is expressed in Eq. (3.17) that is needed to see an actual outbreak for the epidemic, the results are presented in Fig. 8(b). We see that, as I_H increases the effective $\mathcal{R}_0$ needed for an outbreak increases. This phenomenon can be understood taking into account that $\dot{I}_H = \beta I_v \frac{S_H}{N_H} - \gamma I_H$, and that $I_v(0) = 0$ so, if $I_H(0)$ is big, a greater value of β is needed to match the decreasing term $-\gamma I_H$ and be able to change the decreasing behavior of I_H. One may argue that, as the number of initial infected hosts is increased then the number of susceptible hosts at $t = 0$ deviates from being N_H and, thus, the expression Eq. (3.18) would be more suitable. In fact, we have also checked that with the general expression for $\mathcal{R}_0$ Eq. (3.18) the threshold behavior deviates from $\mathcal{R}_0^{'} = 1$ when $I_H(0)$ is large. Therefore, we conclude that is the distance from the disease free equilibrium that causes the deviation, and not the value of $S_H(0)$.

A more interesting situation is for the case in which the initial population of vectors, considered all susceptible at $t = 0$, is different from its stationary value $N_v^o \delta/\mu$. This analysis accounts for an scenario in which the vector population has not reached the equilibrium at the beginning of the dynamics, which in fact, is a biologically relevant situation, as the population of vectors may not be in the stationary state in general.



(a) (b)

Figure 9: Effect of the initial condition $N_v(0)$. The parameters used are $N_H = 100$, $N_v^o = 1000$, $\mu = 1$, $\gamma = 0.7$, and $\alpha = 3$ and the initial conditions $I_H(0) = 0.001 N_H$, $I_v(0) = R_H(0) = 0$, β has been varied to obtain each $\mathcal{R}_0$.(a) for $\delta = 0.5$ so $\beta_c = 7/150$, (b) for $\delta = 2$ so $\beta_c = 7/600$.

Fig. 9 discusses the effect on the epidemic threshold of the initial condition of the vector population, $N_v(0) = S_v(0)$, as we assume that they are not infected vectors initially. The quantity $\mathcal{R}_0$ that is plotted represents the "apparent" value of $\mathcal{R}_0$, because the non-stationary vector population is altering the location of the threshold. One could write that the real threshold is $\mathcal{R}_0^* = \mathcal{R}_0 p$, where $\mathcal{R}_0$ is calculated using Eq. (3.17) and p is a correction factor accounting for a constant vector population that the NGM is not accounting properly. With the proper quantity the threshold should be defined by $\mathcal{R}_0^* = 1$, implying that the apparent $\mathcal{R}_0$ is different than 1 if the correction factor $p \neq 1$. One can see that the numerically determined threshold, the apparent $\mathcal{R}_0$, coincides with the theoretical value given by Eq. (3.17) for $N_v(0)/N_v(\infty) \lesssim 1$ while it clearly diverges for large values of this ratio, and this does not change much for the two values of δ/μ represented in Fig. 9. We can understand better this behavior by looking at the time evolution of the model for small and large values of $N_v(0)/N_v(\infty)$. In turn, Fig. 10, plots the population of the compartments versus time, for the parameters of Fig. 10 and 2 initial conditions for each of the 2 ratios δ/μ studied in Fig. 9. For $N_v(0) < N_v(\infty)$, $N_v(0)/N_v(\infty) = 0.5$, and for both values of δ/μ we can see in Fig. 10(a) and Fig. 10(c) that the the population of susceptible vectors starts below and then converges to the asymptotic value, $N_v^o \delta/\mu$. The fact that the apparent $\mathcal{R}_0$ in Fig. 9 is 1 can only be interpreted recognizing that the location of the threshold is determined by the largest value of

$N_v(t)$, namely the asymptotic value $\delta/\mu N_v^o$, that is the largest value that $S_v(t)$ can take, and $R_0 = 1$ (cf. Fig. 9). Then, the epidemic threshold is determined by the asymptotic value, that enters in Eq. (3.17), that is the largest possible value that S_v can take. In turn, the situation is different in the case that $N_v^o/N_v(\infty)$ is large enough. From Fig. 10(b) and Fig. 10(d), $N_v(0)/N_v(\infty) = 4$, we can see for a substantial time, S_v is larger than the asymptotic value. This implies that the epidemic threshold can no longer be determined by $\delta/\mu N_v^o$ as it appears in Eq. (3.17). However, $\mathcal{R}_o^*$ appears not be determined by the largest value, $S_v(0)$, that is 4 times the asymptotic value in these examples, as the observed value at the threshold is $R_0 \approx 0.52$, what implies that the correction factor $p \approx 1/0.52 = 1.92$, not 4 that is the ratio of $S_v(0)$ to the asymptotic value. So, it appears that p is determined by a suitable average of $S_v(t)$ for $N_v(0)/N_v(\infty) \gtrsim 1$. The lack of a suitable theoretical approach to the determination of the epidemic threshold of the time-varying vector population does not allow a full understanding of the observed asymmetry in these 2 cases. We remark that the transition between these 2 observed behaviors, the apparent value of $\mathcal{R}_0 = 1$ for $N_v(0)/N_v(\infty)$ small and large, does not occur when this ratio is 1, but between 1 and 2, and this cannot be explained because we lack a proper theoretical analysis of the epidemic threshold.

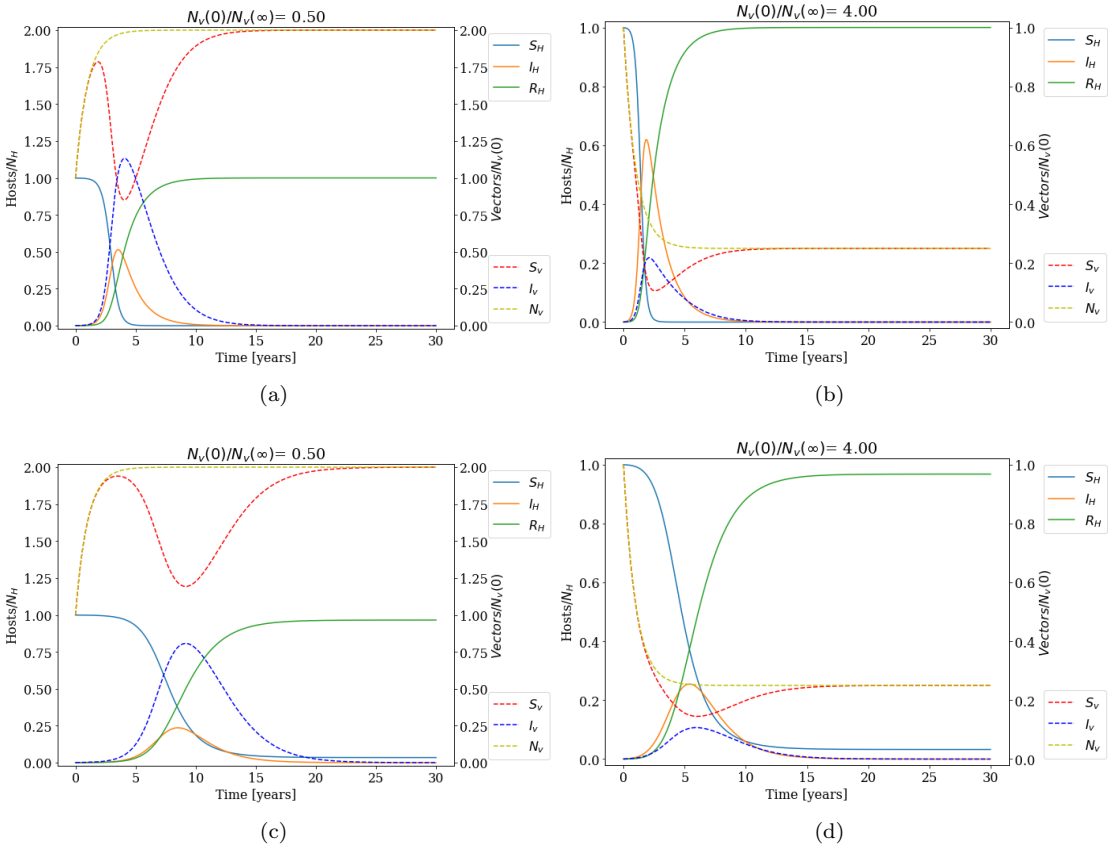


Figure 10: Effect of the initial condition $N_v(0)$. The parameters used are the same than Fig. 9 with $\beta = 0.23$. (a) for $\delta/\mu = 2$, $N_v(0)/N_v(\infty) = 0.5$. (b) for $\delta/\mu = 2$, $N_v(0)/N_v(\infty) = 4$. (c) for $\delta/\mu = 0.5$, $N_v(0)/N_v(\infty) = 0.5$. (d) for $\delta/\mu = 0.5$, $N_v(0)/N_v(\infty) = 4$.

By means of the previous discussion we have studied the validity region of $\mathcal{R}_0$ with respect of the initial conditions of the system. We have seen that the basic reproduction number obtained via the traditional methods of the NGM is not sufficient in all epidemiological situations. Specifically we have seen that for initial conditions far for the stationary state of the system $\mathcal{R}_0$ fails to predict the threshold behavior. This results are in line with the discussion presenting in Section 3.3.3, following [25], stating that a generalization of the method to compute $\mathcal{R}_0$ is needed for the cases in which the vector population is not in the stationary state for $t = 0$.

Finally, we present an analysis of the time scale approximation done in Eq. (3.23) to reduce the dimensionality of the system. To do so, we perform a comparative analysis between the results obtained by numerically solving the complete system Eq. (3.5) and the approximated one for distinct values of μ. We perform this analysis for two values of $\mathcal{R}_0 = 2.5, 0.75$, above and below threshold respectively, and for different values of δ/μ.

As it can be seen in both Fig. 11 and Fig. 13 the approximated and the exact system give very similar results, both above and below the threshold, in the region of validity of the approximation, i. e. when $1/\mu \to 0$. In Fig. 12 we verify that these results are consistent for different ratios of δ/mu: $\delta = 100 = \mu$ and $\delta = 150 > \mu$ for $\mu = 100$.

Moreover, from Fig. 13 one can see that for cases below threshold the approximation also works quite well even for smaller values of μ, so $1/\mu \to 0$ is not quite fulfilled. However, in this case, as the system is below threshold, what happens is that $\dot{I}_v \approx 0$ and thus, the relation Eq. (3.20) is still fulfilled. Another result that can be seen in figures Fig. 11, Fig. 12 and Fig. 13 is that the approximation works better for larger values of t than at the beginning of the dynamics. This is an expected result, as discussed in Section 3.3.3, where we have argued that an initial condition of the system is lost in the approximation and, so, that we expect an initial transient in which the approximation and the exact system differ.

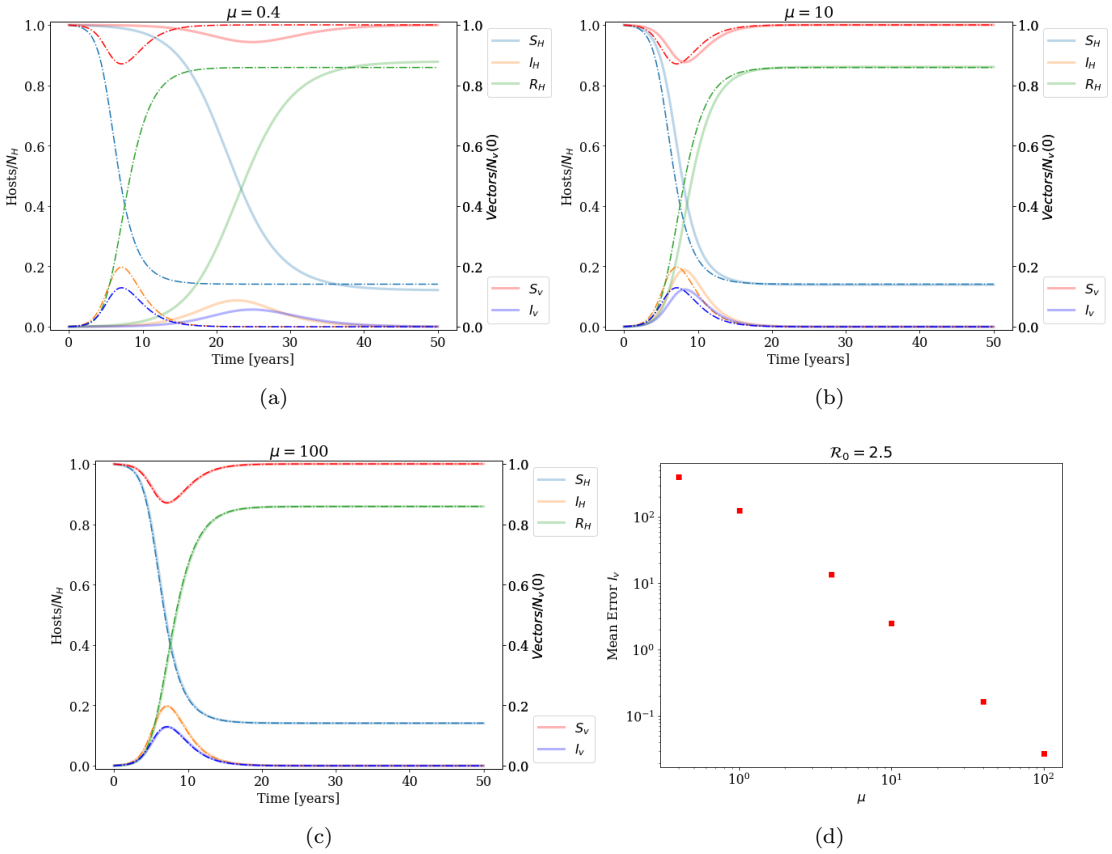
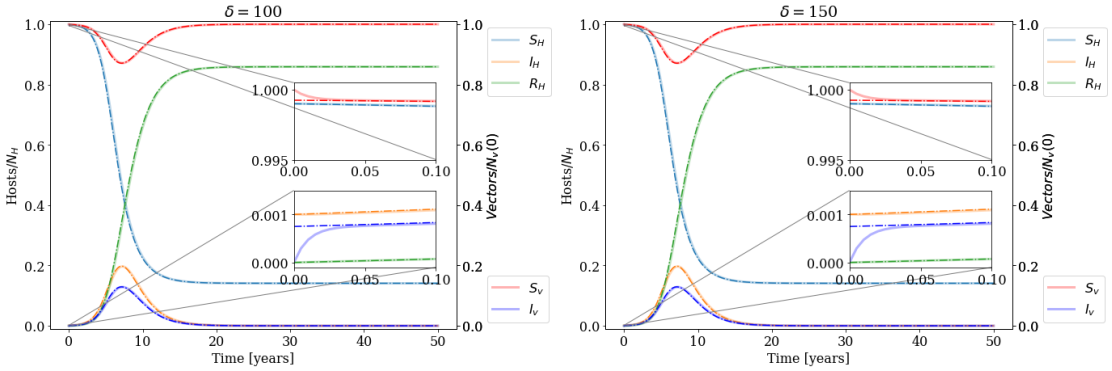
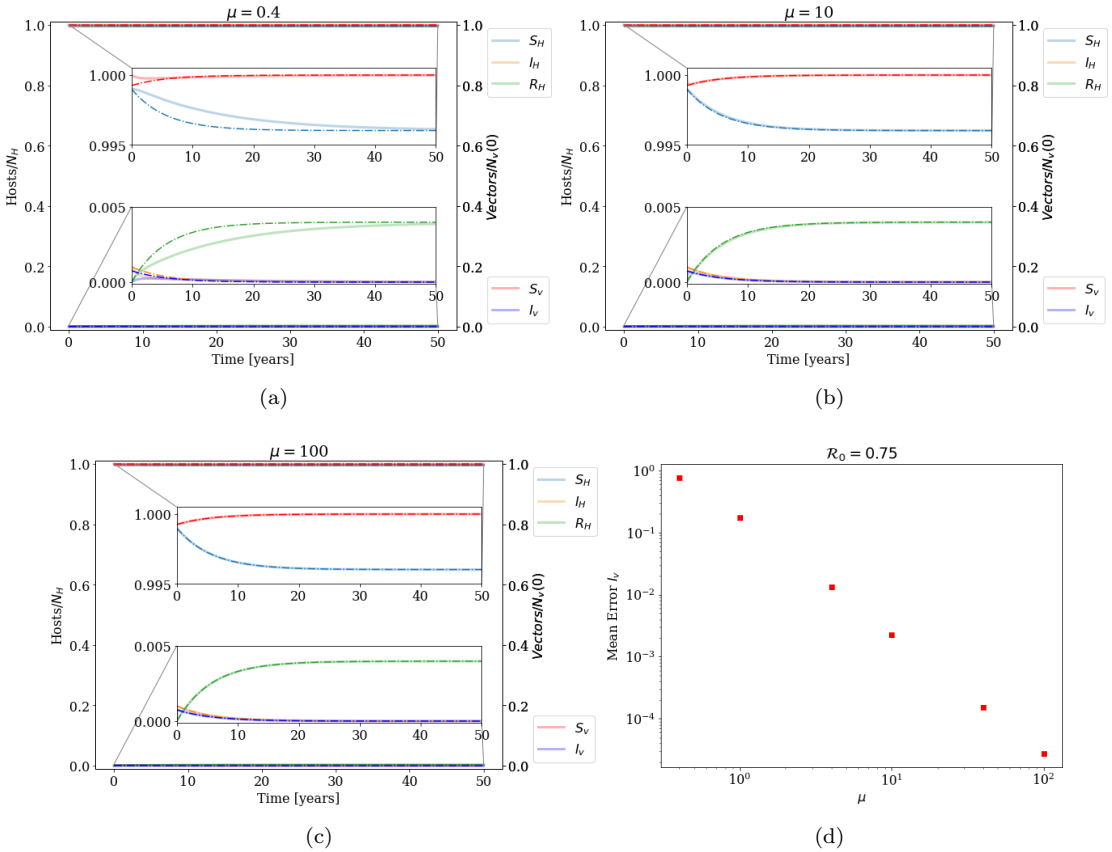


Figure 11: Numerical check of the approximated model reduction for $R_0 = 2.5$, the parameters used are $N_H = 100$, $N_v^o = 1000$, $\delta = 5$, $\gamma = 0.7$, $\alpha = 0.75\mu$ and the initial conditions $S_v(0) = 1000\delta/\mu$, $I_H(0) = 0.001N_H$, $I_v(0) = R_H(0) = 0$. (a), (b) and (c) are for $\mu = 0.4$, 10 and 100 thus $\beta = 0.01867$, 0.4667 and 4.667 respectively. The solid line represents the exact model and the dashed line represents the approximated 2-D model. In (d) the mean error between the approximate and exact solutions for increasing μ.



(a) (b)

Figure 12: Numerical check of the approximated model reduction for $R_0 = 2.5$ and $\mu = 100$, the rest of the parameters are the same than the ones used in Fig. 11 unless the contrary is specified. (a) for $\delta = 100 = \mu$ and $\beta = 0.233$. (b) for $\delta = 150 > \mu$ and $\beta = 0.1556$.



(a) (b) (c) (d)

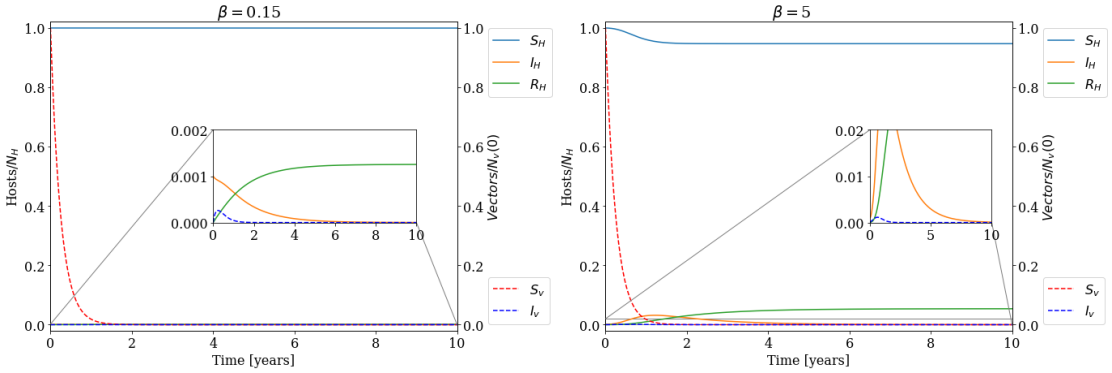
Figure 13: Numerical check of the approximated model reduction for $R_0 = 0.75$, the parameters used are $\delta = 5$, $\gamma = 0.7$, $\alpha = 0.75\mu$ and the initial conditions $S_v(0) = 1000\delta/\mu$, $I_H(0) = 0.001N_H$, $I_v(0) = R_H(0) = 0$. (a), (b) and (c) are for $\mu = 0.4$, 10 and 100 thus $\beta = 0.0056$, 0.14, and 1.4 respectively. The solid line represents the exact model and the dashed line represents the approximated 2-D model. In (d) the mean error between the approximate and exact solutions for increasing μ.

4.2 Particular case $\delta = 0$

The case $\delta = 0$ is an special case as the stationary population of vectors is zero. As we have discussed in Section 3.4, the computation of $\mathcal{R}_0$, via usual NGM method, does not make sense in this case, as in the stationary state there is no vector population and the disease cannot spread. In fact, if we try to apply the expression for the basic reproduction number we obtained for the general case, Eq. (3.17) it yields that $\mathcal{R}_0 = 0$, so there is not threshold for the epidemic. It is so because, as we have discussed, the disease cannot propagate in the stationary state of the system. As a consequence, the methods used in this work, based on linear stability analysis with respect to the the disease free equilibrium state, are not informative of an epidemic outbreak, and other techniques should be used to analyze this case. In Fig. 14 we present two different realizations of the model, one with an outbreak and another one without, illustrating the fact that, there can be epidemic outbreaks in this case. The parameter values used being:

$$N_H = 100, \quad S_v(0) = 1000, \quad I_H(0) = 0.001N_H, \quad I_v(0) = R_H(0) = 0, \quad \mu = 4, \\ \gamma = 0.7, \quad \alpha = 0.75\mu \quad \text{and} \quad \delta = 0 \tag{4.3}$$

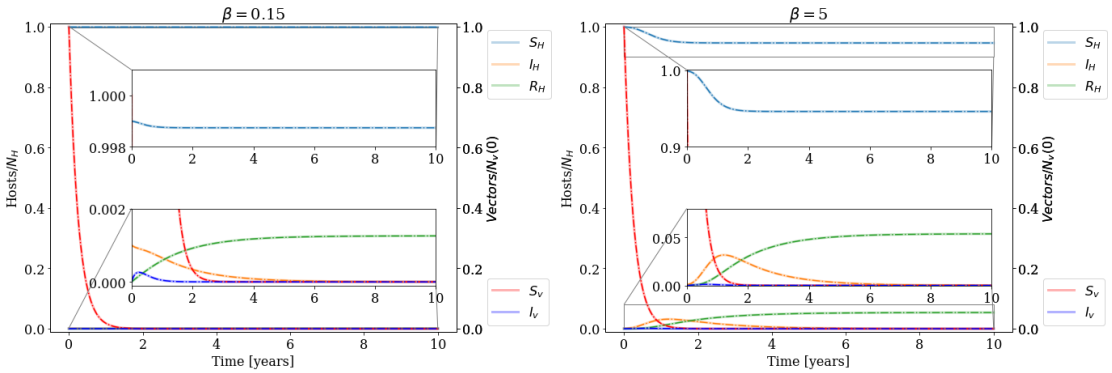
with $\beta = 0.15, 5$ respectively.



(a) (b)

Figure 14: Different realizations for the case in which $\delta = 0$. The insets focus on the initial evolution of the I_H and I_v compartment. The parameters used are: $N_H = 100$, $N_v^o = 1000$, $I_H(0) = 0.001N_H$, $I_v(0) = R_H(0) = 0$, $S_v(0) = N_v^o$, $\mu = 4$, $\gamma = 0.7$, $\alpha = 0.75\mu$ and $\delta = 0$.

For the case $\delta = 0$ it has been also derived an exact reduction to a 2D model Eq. (3.29). In Fig. 15 different realizations are presented, for the same values of β than Fig. 14 . We see that there is good agreement between the complete system and the reduced one. The good agreement between the curves computed with the exact system and the ones computed with the reduced one implies the conservation of the quantity C, so Fig. 15 is also a validation of this conserved quantity.



(a) (b)

Figure 15: Different realizations for the numerical check of the exact 2-D reduction the case in which $\delta = 0$. The solid line represents the exact model and the dashed line represents the reduced exact 2-D model. The parameters used are the same than in Fig. 14.

5 Conclusions

In this work, we aimed to study the disease caused by the bacterium *Xylella fastidiosa* in the almond trees of Majorca. The analysis has been done by means of the mathematical framework of compartmental models, and a model has been proposed and analyzed, both from an analytical and computational point of view.

We have presented a five dimensional model of differential equations, coupling hosts and vectors. We have considered that the hosts could be susceptible, infective and removed, while for the vectors two states have been considered, susceptible and infective. Moreover, the host population has been taken as constant while birth and death of vectors has been assumed. For this model, the basic reproduction number has been computed. We have argued that the validity of this expression only near the disease free equilibrium of the system, when the number of hosts has reached its stationary value. We have discussed that the presented methods to compute $\mathcal{R}_0$ suppose an existence of an stable disease free equilibrium, DFE, and the proximity of the system to such state. Yet, the assumption of the existence of such equilibrium to determine the stability properties of any epidemiological situation may not be sufficient in all scenarios, as noted in [26]. In particular, in vector-borne diseases with seasonal population of vectors, a generalization of $\mathcal{R}_0$ is needed to tackle the problem, as the vector population is not in equilibrium even in the disease free state [27].

Next, we have discussed reductions of the dimensionality of the system to a reduced model accounting only for the hosts. This reductions have been done via a fast-slow approximation for the general case, and through a conserved quantity, that has been derived and used for the reduction, in the particular case in which $\delta = 0$. Remarkably, this reduced model resembles a SIR model for the hosts with time dependent coefficients. This reduction is experimentally relevant because in it only appears the epidemiological state of the hosts. Then, having a reduced system for the hosts, would enable an easier comparison with experimental data, as typically the state of the vectors is more difficult to access.

Finally, we have performed a numerical analysis of the model presented, and we have seen that the results are in agreement with the ones expected from the analytical analysis. Moreover, we have also showed numerically the limitation of $\mathcal{R}_0$ when dealing with epidemiological situations that are not near the DFE. In particular, we have seen that an initial population of vectors larger than the stationary value $N_v^o\delta/\mu$ leads to an smaller value of $\mathcal{R}_0$ and finally, for the case $\delta = 0$, we have numerically seen that there can be outbreak even if the basic reproduction number does not predict it.

A more in depth future analysis may incorporate the vector seasonality via repetitions of one year annual cycle. For this model climatic effects need to be also incorporated in such model as the population of vectors at the beginning of those cycles depends on it. Moreover, as we have discussed along this work, a generalization of a method to compute $\mathcal{R}_0$ will be needed to tackle this situation, as by definition of the model, the vector population would not be in a stationary state, and so the standard NGM method does not provide the framework to predict the threshold behavior of the system in those cases. The generalization of the method to compute $\mathcal{R}_0$ would also provide a basis to compare with experimental data, as given the seasonal characteristics of the vector, the population in the field is not in a stationary state.

A Next-Generation Matrix Method

We introduce the method of the NGM following Chapter 5.3 of [14] that allows to study the cases in which a DFE can be defined. Given a compartmental model for a disease transmission, assume that there are n infected compartments and m non infected compartments, being $n+m$ the total number of independent variables. Let $x \in \mathbb{R}^n$ be the vector of infected variables and $y \in \mathbb{R}^m$ be the vector of variables in the non infected compartments. Lets arrange the equations in the original system so that the first n compartments of the ODE correspond to the infected compartments, then:

$$\begin{cases} \dot{x}_i &= f_i(x,y), \quad i = 1, \cdots, n \\ \dot{y}_j &= g_j(x,y), \quad j = 1, \cdots, m \end{cases} \tag{A.1}$$

Next, assume that the right hand side in the infected compartments can be split in the following way:

$$\begin{cases} \dot{x}_i = \mathfrak{F}_i(x,y) - \mathfrak{V}_i(x,y), & i = 1, \cdots, n \\ \dot{y}_j = g_j(x,y), & j = 1, \cdots, m \end{cases} \tag{A.2}$$

where $\mathfrak{F}_i(x,y)$ is the rate of appearance of new infections in compartment i and $\mathfrak{V}_i(x,y)$ incorporates the remaining transitional terms, such as births, deaths, disease progression, and recovery. Note that the decomposition into infected and non infected compartments and the splitting into $\mathfrak{F}$ and $\mathfrak{V}$ may not be unique, as it depends on the interpretation of the disease process of the model. This gives rise to different expressions for the NGM and therefore, different expressions of $\mathcal{R}_0$ as discussed in Section 2.2. Nevertheless, all decompositions must satisfy the following properties:

- $\mathfrak{F}_i(0,y) = 0$ and $\mathfrak{V}_i(0,y) = 0$ for $y \geq 0$ and $i = 1, \cdots, n$. The first condition imposes that all new infections are secondary infections arising from infected hosts. The second condition states that there is no immigration of susceptible individuals into the disease compartments.
- $\mathfrak{F}_i(x,y) \geq 0 \, \forall \, x, y \geq 0$.
- $\mathfrak{v}_i(x,y) \leq 0 \, \forall \, x_i = 0$, $i = 1, \cdots, n$. When the compartment is empty the net flow in the compartment needs to be inflow.
- $\sum_{i=1}^{n} \mathfrak{v}_i(x,y) \geq 0 \, \forall \, x, y \geq 0$. The total outflowof all infected compartments is positive.

Lets assume that the disease-free system:

$$\dot{y}_j = g_j(0,y) \tag{A.3}$$

has an unique DFE $\mathfrak{E}_0 = (0, y_0)$ which is a local asymptotically stable solution for initial conditions of the form $(0,y)$, which approach $(0,y_0)$ as $t \to \infty$. Then, define the matrices F and V as:

$$F = \left[\frac{\partial \mathfrak{F}_i(0,y_0)}{\partial x_j} \right] \qquad \text{and} \qquad V = \left[\frac{\partial \mathfrak{V}_i(0,y_0)}{\partial x_j} \right] \tag{A.4}$$

Both matrices F and V appear from the linearization of the system around the DFE. The linearizated system for the infected compartments is then:

$$\dot{x}_j = (F - V)x \tag{A.5}$$

Finally the NGM is defined as:

$$K = FV^{-1} \tag{A.6}$$

being the basic reproduction number $\mathcal{R}_0$ its spectral radius.

B Computation of $\mathcal{R}_0$ considering the transmission from host-vector-host as a two step process

Along this work we have considered the transmission from host-vector-host to be a one step process following [16]. In Section 2.2 we have state that in the case of considering the transmission from host-vector-host as a two step process, the expression for $\mathcal{R}_0$ will be different, however the threshold value would be the same. Here we illustrate the case in which a two step process is considered. In this case, the transmission matrix F is:

$$F = \begin{pmatrix} 0 & \beta \\ \alpha \frac{N_v^o}{N_H} \frac{\delta}{\mu} & 0 \end{pmatrix} \tag{B.1}$$

and the transition part V:

$$V = \begin{pmatrix} \gamma & 0 \\ 0 & \mu \end{pmatrix} \tag{B.2}$$

Thus:

$$V^{-1} = \begin{pmatrix} \frac{1}{\gamma} & 0 \\ 0 & \frac{1}{\mu} \end{pmatrix} \tag{B.3}$$

Therefore, the next generation matrix:

$$K \equiv FV^{-1} = \begin{pmatrix} 0 & \beta \\ \frac{\alpha}{\gamma} \frac{N_v^o}{N_H} \frac{\delta}{\mu} & 0 \end{pmatrix} \tag{B.4}$$

The with its spectral radius being:

$$det(K - \sigma \mathbb{I}) = 0 \implies \begin{pmatrix} -\sigma & \frac{\beta}{\mu} \\ \frac{\alpha}{\gamma} \frac{N_v^o}{N_H} \frac{\delta}{\mu} & -\sigma \end{pmatrix} = \sigma^2 - \frac{\beta\alpha}{\gamma\mu} \frac{N_v^o}{N_H} \frac{\delta}{\mu} = 0 \implies$$

$$\sigma = \pm \sqrt{\frac{\beta\alpha}{\gamma\mu} \frac{N_v^o}{N_H} \frac{\delta}{\mu}} \tag{B.5}$$

$$\mathcal{R}_0 = \sqrt{\frac{\beta\alpha}{\gamma\mu} \frac{N_v^o}{N_H} \frac{\delta}{\mu}} \tag{B.6}$$

Then, as expected, the expression for $\mathcal{R}_0$ is different because a different interpretation of the disease process has been used, but regardless the same threshold than Eq. (3.17) is recovered. For $\frac{\beta\alpha}{\gamma\mu} \frac{N_v^o}{N_H} \frac{\delta}{\mu} > 1$ there will be initial exponential growth, and for $\frac{\beta\alpha}{\gamma\mu} \frac{N_v^o}{N_H} \frac{\delta}{\mu} < 1$ the epidemic will die out. However, the value of the theoretical growth rate, $\mathcal{R}_0$, depends on whether one considers the spread from host to vector to host a one or two step process. Interestingly, in this approach one can interpret the two nonzero entries of the NGM K as the host transmission reproduction number $\mathcal{R}_H = \frac{\beta}{\mu}$ and the vector transmission reproduction number $\mathcal{R}_v = \frac{\alpha}{\gamma} \frac{N_v^o}{N_H} \frac{\delta}{\mu}$. The first, $\mathcal{R}_H$, accounting for the number of secondary infections of vectors when placing one infected host in an population of susceptible vectors, and similarly the second, $\mathcal{R}_v$, accounting for the infected hosts produced by placing an infected vector in a population of susceptible hosts.

References

[1] E. Moralejo *et al.*, "Phylogenetic inference enables reconstruction of a long-overlooked outbreak of almond leaf scorch disease (*Xylella fastidiosa*) in Europe," *Communications Biology*, vol. 3, no. 1, p. 560, 2020.

[2] L. L. R. Marques, H. Ceri, G. P. Manfio, D. M. Reid, and M. E. Olson, "Characterization of biofilm formation by *Xylella fastidiosa* in vitro," *Plant Disease*, vol. 86, no. 6, pp. 633–638, 2002.

[3] M. Brunetti, V. Capasso, M. Montagna, and E. Venturino, "A mathematical model for *Xylella fastidiosa* epidemics in the mediterranean regions. promoting good agronomic practices for their effective control.," *Ecological Modelling*, vol. 432, p. 109204, 2020.

[4] R. P. P. Almeida, L. De La Fuente, R. Koebnik, J. R. S. Lopes, S. Parnell, and H. Scherm, "Addressing the new global threat of *Xylella fastidiosa*," *Phytopathology*, vol. 109, no. 2, pp. 172–174, 2019.

[5] M. Jeger and C. Bragard, "The epidemiology of *Xylella fastidiosa*; a perspective on current knowledge and framework to investigate plant host–vector–pathogen interactions," *Phytopathology*, vol. 109, no. 2, pp. 200–209, 2019.

[6] M. Nicoletti, *Insect-Borne Diseases in the 21st Century*, ch. Three scenarios in insect-borne diseases, pp. 99–251. Academic Press, 2020.

[7] EFSA, "Update of the *Xylella* spp. host plant database," *EFSA Journal*, vol. 16, no. 9, p. e05408, 2018.

[8] P. Baldi and N. La Porta, "Xylella fastidiosa: Host range and advance in molecular identification techniques," *Frontiers in Plant Science*, vol. 8, p. 944, 2017.

[9] EFSA, "Update of the scientific opinion on the risks to plant health posed by *Xylella fastidiosa* in the EU territory," *EFSA Journal*, vol. 17, no. 5, p. e05665, 2019.

[10] D. Cornara, D. Bosco, and A. Fereres, "Philaenus spumarius: when an old acquaintance becomes a new threat to European agriculture," *Journal of Pest Science*, vol. 91, no. 3, pp. 957–972, 2018.

[11] D. Cornara, M. Marra, M. Morente, E. Garzo, A. Moreno, M. Saponari, and A. Fereres, "Feeding behavior in relation to spittlebug transmission of *Xylella fastidiosa*," *Journal of Pest Science*, vol. 93, no. 4, pp. 1197–1213, 2020.

[12] C. Lago, E. Garzo, A. Moreno, L. Barrios, A. Martí-Campoy, F. Rodríguez-Ballester, and A. Fereres, "Flight performance and the factors affecting the flight behaviour of *Philaenus spumarius* the main vector of *Xylella fastidiosa* in Europe," *Scientific Reports*, vol. 11, no. 1, p. 17608, 2021.

[13] W. O. Kermack, A. G. McKendrick, and G. T. Walker, "A contribution to the mathematical theory of epidemics," *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character*, vol. 115, no. 772, pp. 700–721, 1927.

[14] M. Martcheva, *An Introduction to Mathematical Epidemiology*. Springer US: Boston, MA, 2015.

[15] D. L. Smith, K. E. Battle, S. I. Hay, C. M. Barker, T. W. Scott, and F. E. McKenzie, "Ross, Macdonald, and a theory for the dynamics and control of mosquito-transmitted pathogens," *PLOS Pathogens*, vol. 8, no. 4, pp. 1–13, 2012.

[16] F. Brauer, C. Castillo-Chavez, A. Mubayi, and S. Towers, "Some models for epidemics of vector-transmitted diseases," *Infectious Disease Modelling*, vol. 1, no. 1, pp. 79–87, 2016.

[17] O. Diekmann, J. A. P. Heesterbeek, and M. G. Roberts, "The construction of next-generation matrices for compartmental epidemic models," *Journal of The Royal Society Interface*, vol. 7, no. 47, pp. 873–885, 2010.

[18] J. D. Murray, *Mathematical Biology: I. An Introduction*, ch. Dynamics of Infectious Diseases: Epidemic Models and AIDS, pp. 315–394. Springer: New York, 2002.

[19] F. Brauer in *Mathematical Epidemiology* (F. Brauer, P. van den Driessche, and J. Wu, eds.), ch. Compartmental Models in Epidemiology, pp. 19–79, Springer: Berlin, Heidelberg, 2008.

[20] J. Cushing and O. Diekmann, "The many guises of R0 (a didactic note)," *Journal of Theoretical Biology*, vol. 404, pp. 295–302, 2016.

[21] O. Diekmann, J. Heesterbeek, and J. Metz, "On the definition and the computation of the basic reproduction ratio R0 in models for infectious diseases in heterogeneous populations," *Journal of Mathematical Biology*, vol. 28, no. 4, p. 365—382, 1990.

[22] A. Purcell and A. Finlay, "Evidence for noncirculative transmission of Pierce's disease bacterium by sharpshooter leafhoppers," *Phytopathology*, vol. 69, no. 4, p. 393—395, 1979.

[23] EFSA, "Scientific opinion on the risks to plant health posed by *Xylella* fastidiosa in the EU territory, with the identification and evaluation of risk reduction options," *EFSA Journal*, vol. 13, no. 1, p. 3989, 2015.

[24] H.-M. Wei, X.-Z. Li, and M. Martcheva, "An epidemic model of a vector-borne disease with direct transmission and time delay," *Journal of Mathematical Analysis and Applications*, vol. 342, no. 2, pp. 895–908, 2008.

[25] N. Bacaër and M. G. M. Gomes, "On the final size of epidemics with seasonality," *Bulletin of Mathematical Biology*, vol. 71, no. 8, p. 1954, 2009.

[26] I. G. Laukó, "Stability of disease free sets in epidemic models," *Mathematical and Computer Modelling*, vol. 43, no. 11, pp. 1357–1366, 2006.

[27] N. Bacaër and S. Guernaoui, "The epidemic threshold of vector-borne diseases with seasonality: The case of cutaneous leishmaniasis in Chichaoua, Morocco," *Journal of Mathematical Biology*, vol. 53, pp. 421–36, 2006.