



Universitat
de les Illes Balears

DOCTORAL THESIS
2024

**Theoretical and data-driven models in
Ecology**

Àlex Giménez Romero



Universitat
de les Illes Balears

DOCTORAL THESIS 2024

Doctoral programme in Physics

Theoretical and data-driven models in Ecology

Àlex Giménez Romero

Thesis Supervisor: Manuel A. Matías
Thesis Tutor: Cristóbal López Sánchez

Doctor by the Universitat de les Illes Balears

Supervisors:
Manuel Matías

Àlex Giménez Romero.
Theoretical and data-driven models in Ecology. ©
Palma de Mallorca, July 2024

A en Manuel Miranda
pel seu suport i ajuda
durant tots aquests anys.
Sempre estaràs amb mi.
i recordare sempre
el que em vas ensenyar.

Dr Manuel A. Matías of the Consejo Superior de Investigaciones Científicas (CSIC)

I DECLARE:

That the thesis title *Theoretical and data-driven models in Ecology*, presented by Àlex Giménez Romero to obtain a doctoral degree, has been completed under my supervision and meets the requirements to opt for an International Doctorate.

For all intents and purposes, I hereby sign this document.

Signature

Dr. Manuel A. Matías
Thesis Supervisor

Palma de Mallorca, July 2024

Acknowledgements

Acknowledgements go here.

Resum

El resum va aquí.

Resumen

El resumen va aquí.

Abstract

Abstract goes here.

Contents

I

Introduction

1	The global biodiversity crisis	0
1.1	Introduction	0
1.2	The Mass Mortality Event of <i>Pinna nobilis</i>	0
1.3	<i>Xylella fastidiosa</i> : an emerging global threat	0
1.4	The global decline of seagrasses: the case of <i>Posidonia oceanica</i>	0
1.5	Coral reefs: a global crisis	0
2	Models in Ecology	1
2.1	Introduction	1
2.2	Why do we need models?	1
2.3	Theoretical modelling	1
2.3.1	Dynamical systems	1
2.3.2	Compartmental models	1
2.3.3	Individual-based models	1
2.3.4	Reaction-difussion models	1
2.4	Data-driven modelling	1
2.4.1	Machine learning	1
2.4.2	Climate data and modelling	1
2.4.3	Species Distribution Models	1
3	Main original contributions of this thesis	3
3.1	Introduction	3
3.2	Modelling parasite-produced marine diseases	3
3.3	Modelling the spread of <i>Xylella fastidiosa</i> diseases	3
3.4	Modelling the establishment risk of <i>Xylella fastidiosa</i> diseases	3
3.5	Coral reef size and structure	3
3.6	Machine learning to reconstruct pH time series	3
3.7	Mapping <i>Posidonia oceanica</i> meadows from remote sensing data	3

4 Modelling parasite-produced marine diseases: The case of the mass mortality event of *Pinna nobilis* 7

4.1	Introduction	7
4.2	The SIRP model	9
4.2.1	Model structure and initial considerations	9
4.2.2	General SIRP model	10
4.2.3	Model reduction	13
4.3	Numerical analysis of the model	14
4.3.1	The basic reproduction number R_0	15
4.3.2	Final state of the epidemic	16
4.3.3	Maximum of infected individuals	16
4.3.4	Numerical verification of the fast-slow approximation	16
4.3.5	Numerical verification of the model approximation from the exact reduction	18
4.4	Model validation with data of the <i>Pinna nobilis</i> Mass Mortality Event	18
4.5	Conclusions	22
4.6	Appendix	23
4.6.1	Finding a conserved quantity for the SIRP model	23
4.6.2	Stability analysis of the fixed points of the SIRP model	24
4.6.3	Calculation of R_0 using the Next Generation Matrix method	25
4.6.4	Sensitivity Analysis	26
4.6.5	General rate change with temperature	26

5 Spatial effects in parasite-induced marine diseases of immobile hosts 29

5.1	Introduction	29
5.2	The SIRP spatial model	30
5.3	Results	32
5.3.1	Non-spatial limit	32
5.3.2	Approximate relation between parasites and infected hosts	33
5.3.3	Spatial threshold	35
5.3.4	Spreading speed of the infected population and time to extinction	37
5.4	Conclusions	40
5.5	Appendix	41
5.5.1	Derivation of the non-spatial equation for R_∞	41

6 Vector-borne diseases with non-stationary vector populations: the case of growing and decaying populations 45

6.1	Introduction	45
6.2	The model	46
6.2.1	Preliminary analysis of the model	47
6.3	Results	48
6.3.1	The effect of non-stationary vector populations into the epidemic threshold and disease dynamics	48
6.3.2	The basic reproduction number for non-stationary vector populations	51
6.3.3	Fast-slow approximation	53
6.3.4	Reduction to a SIR model	54

6.4	Conclusions	55
6.5	Appendix	56
6.5.1	Calculation of R_0 from standard methods	56
6.5.2	Calculation of R_0 for non-stationary vector populations	57
7	A compartmental model for <i>Xylella fastidiosa</i> diseases with explicit vector seasonal dynamics	59
7.1	Introduction	59
7.2	Materials and Methods	60
7.2.1	Epidemic model: the SEIR-V model	60
7.2.2	Basic reproductive number	62
7.2.3	Epidemiological data	62
7.2.4	Model fitting through Bayesian Inference	62
7.2.5	Sensitivity Analysis	63
7.3	Results	63
7.3.1	Model fit and parameter estimates	63
7.3.2	Global Sensitivity Analysis	67
7.3.3	Epidemic control through vector management	69
7.4	Discussion	69
7.5	Appendix	71
7.5.1	Vector population dynamics	71
7.5.2	Determination of R_0	72

IV**Modelling the risk of vector-borne plant diseases**

8	Global predictions for the risk of establishment of Pierce's disease of grapevines	77
9	Global warming significantly increases the risk of Pierce's disease epidemics in European vineyards	79
10	High-resolution climate data reveals increased risk of Pierce's Disease for grapevines worldwide	81

V**Data-driven methods for global ecological problems**

11	A comprehensive dataset on global coral reefs size and geometry	85
12	Universal spatial properties of coral reefs	87
13	pH trends and seasonal cycle in the coastal Balearic Sea reconstructed through machine learning	89
14	Mapping the distribution of seagrass meadows from space with deep convolutional neural networks	91

List of publications

1. Àlex Giménez-Romero et al. "Modelling parasite-produced marine diseases: The case of the mass mortality event of *Pinna nobilis*". In: *Ecological Modelling* 459 (2021), page 109705. ISSN: 0304-3800. DOI: <https://doi.org/10.1016/j.ecolmodel.2021.109705>. URL: <https://www.sciencedirect.com/science/article/pii/S030438002100260X>
2. Susana Flecha et al. "pH trends and seasonal cycle in the coastal Balearic Sea reconstructed through machine learning". In: *Scientific Reports* 12.1 (July 2022), page 12956. ISSN: 2045-2322. DOI: [10.1038/s41598-022-17253-5](https://doi.org/10.1038/s41598-022-17253-5). URL: <https://doi.org/10.1038/s41598-022-17253-5>
3. Àlex Giménez-Romero et al. "Spatial effects in parasite-induced marine diseases of immobile hosts". In: *Royal Society Open Science* 9.8 (2022), page 212023. DOI: [10.1098/rsos.212023](https://doi.org/10.1098/rsos.212023). eprint: <https://royalsocietypublishing.org/doi/pdf/10.1098/rsos.212023>. URL: <https://royalsocietypublishing.org/doi/abs/10.1098/rsos.212023>
4. Àlex Giménez-Romero, Rosa Flaquer-Galmés, and Manuel A. Matías. "Vector-borne diseases with nonstationary vector populations: The case of growing and decaying populations". In: *Phys. Rev. E* 106 (5 Nov. 2022), page 054402. DOI: [10.1103/PhysRevE.106.054402](https://doi.org/10.1103/PhysRevE.106.054402). URL: <https://link.aps.org/doi/10.1103/PhysRevE.106.054402>
5. Alex Giménez-Romero et al. "Global predictions for the risk of establishment of Pierce's disease of grapevines". In: *Communications Biology* 5.1 (Dec. 2022), page 1389. ISSN: 2399-3642. DOI: [10.1038/s42003-022-04358-w](https://doi.org/10.1038/s42003-022-04358-w). URL: <https://doi.org/10.1038/s42003-022-04358-w>
6. Àlex Giménez-Romero, Eduardo Moralejo, and Manuel A. Matías. "A Compartmental Model for *Xylella fastidiosa* Diseases with Explicit Vector Seasonal Dynamics". In: *Phytopathology®* 113.9 (2023). PMID: 36774557, pages 1686–1696. DOI: [10.1094/PHYTO-11-22-0428-V](https://doi.org/10.1094/PHYTO-11-22-0428-V). eprint: <https://doi.org/10.1094/PHYTO-11-22-0428-V>. URL: <https://doi.org/10.1094/PHYTO-11-22-0428-V>
7. Àlex Giménez-Romero et al. "Contrasting Patterns of Pierce's Disease Risk in European Vineyards Under Global Warming". In: *bioRxiv* (2023). DOI: [10.1101/2023.07.17.549293](https://doi.org/10.1101/2023.07.17.549293). eprint: <https://www.biorxiv.org/content/early/2023/07/18/2023.07.17.549293.full.pdf>. URL: <https://www.biorxiv.org/content/early/2023/07/18/2023.07.17.549293>

8. Àlex Giménez-Romero, Eduardo Moralejo, and Manuel A Matias. "High-resolution climate data reveals increased risk of Pierce's Disease for grapevines worldwide". In: *bioRxiv* (2024). DOI: [10.1101/2024.03.06.583743](https://doi.org/10.1101/2024.03.06.583743). eprint: <https://www.biorxiv.org/content/early/2024/03/11/2024.03.06.583743.full.pdf>. URL: <https://www.biorxiv.org/content/early/2024/03/11/2024.03.06.583743>
9. A comprehensive dataset on global coral reefs size and geometry
10. Universal spatial properties of coral reefs
11. Mapping the distribution of seagrass meadows from space with deep convolutional neural networks

Other publications

1. Clara Lago et al. "Degree-day-based model to predict egg hatching of *Philaenus spumarius* (Hemiptera: Aphrophoridae), the main vector of *Xylella fastidiosa* in Europe". In: *Environmental Entomology* 52.3 (Apr. 2023), pages 350–359. ISSN: 0046-225X. DOI: [10.1093/ee/nvad013](https://doi.org/10.1093/ee/nvad013). eprint: <https://academic.oup.com/ee/article-pdf/52/3/350/50615564/nvad013.pdf>. URL: <https://doi.org/10.1093/ee/nvad013>

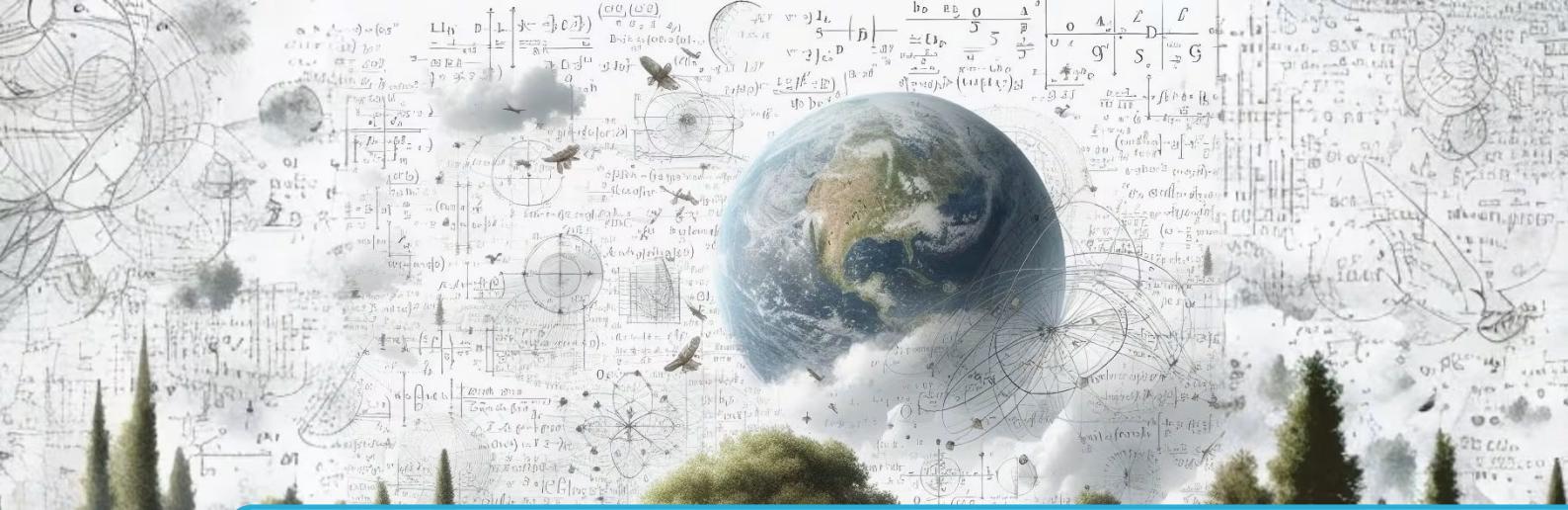
Introduction

1	The global biodiversity crisis	0
1.1	Introduction	0
1.2	The Mass Mortality Event of <i>Pinna nobilis</i>	0
1.3	<i>Xylella fastidiosa</i> : an emerging global threat	0
1.4	The global decline of seagrasses: the case of <i>Posidonia oceanica</i> 0	0
1.5	Coral reefs: a global crisis	0
2	Models in Ecology	1
2.1	Introduction	1
2.2	Why do we need models?	1
2.3	Theoretical modelling	1
2.4	Data-driven modelling	1
3	Main original contributions of this thesis	3
3.1	Introduction	3
3.2	Modelling parasite-produced marine diseases	3
3.3	Modelling the spread of <i>Xylella fastidiosa</i> diseases	3
3.4	Modelling the establishment risk of <i>Xylella fastidiosa</i> diseases	3
3.5	Coral reef size and structure	3
3.6	Machine learning to reconstruct pH time series	3
3.7	Mapping <i>Posidonia oceanica</i> meadows from remote sensing data	3



1. The global biodiversity crisis

- 1.1 Introduction
- 1.2 The Mass Mortality Event of *Pinna nobilis*
- 1.3 *Xylella fastidiosa*: an emerging global threat
- 1.4 The global decline of seagrasses: the case of *Posidonia oceanica*
- 1.5 Coral reefs: a global crisis



2. Models in Ecology



2.1 Introduction

2.2 Why do we need models?

2.3 Theoretical modelling

2.3.1 Dynamical systems

2.3.2 Compartmental models

2.3.3 Individual-based models

2.3.4 Reaction-diffusion models

2.4 Data-driven modelling

2.4.1 Machine learning

2.4.2 Climate data and modelling

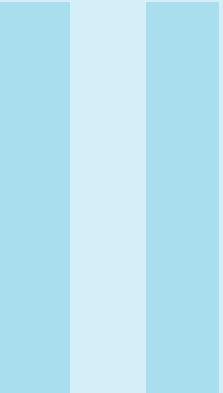
2.4.3 Species Distribution Models



3. Main original contributions of this thesis



- 3.1 Introduction**
- 3.2 Modelling parasite-produced marine diseases**
- 3.3 Modelling the spread of *Xylella fastidiosa* diseases**
- 3.4 Modelling the establishment risk of *Xylella fastidiosa* diseases**
- 3.5 Coral reef size and structure**
- 3.6 Machine learning to reconstruct pH time series**
- 3.7 Mapping *Posidonia oceanica* meadows from remote sensing data**



Modelling parasite-produced marine diseases

4	Modelling parasite-produced marine diseases: The case of the mass mortality event of <i>Pinna nobilis</i>	7
4.1	Introduction	7
4.2	The SIRP model	9
4.3	Numerical analysis of the model	14
4.4	Model validation with data of the <i>Pinna nobilis</i> Mass Mortality Event	18
4.5	Conclusions	22
4.6	Appendix	23
5	Spatial effects in parasite-induced marine diseases of immobile hosts	29
5.1	Introduction	29
5.2	The SIRP spatial model	30
5.3	Results	32
5.4	Conclusions	40
5.5	Appendix	41



4. Modelling parasite-produced marine diseases: The case of the mass mortality event of *Pinna nobilis*

4.1 Introduction

Marine organisms, like their terrestrial counterparts, can serve as hosts for a diversity of parasites and pathogens present in the ecosystem, which are directly responsible for disease outbreaks. Disease induced mortality affects not only the host population, but can cascade through the whole ecosystem, altering its structure and functioning (169). Furthermore, climate change can increase the spread range and impact of parasites and pathogens (29). In fact, marine infectious diseases are recently increasing due to climate change and other anthropogenic pressures, like pollution and overfishing (104). This, in turn, threatens many valuable ecologically habitats and can also result in substantial economic losses in e.g. aquaculture (105). Analysing the impact of these events at appropriate scales (spatial and temporal) and biological organisation levels (species, populations and communities) is crucial to accurately anticipate future changes in marine ecosystems and propose adapted management and conservation plans (128). Thus, there is a strong need to address

the mechanism of disease propagation in marine organisms.

However, the state of the art of epidemiological studies in marine ecosystems lags behind that of terrestrial ecosystems (84). Contact and vector-borne based infectious diseases of terrestrial vertebrates and their epidemiology are typically studied using variations of the classical formulation of Kermack and McKendrick (101, 102, 103), the SIR model. Among other things, this formalism allows to understand why epizootics spread and stop, as the propagation of a disease is a threshold phenomenon (5), regulated by the now commonplace R_0 dimensionless number. Within this framework the initiation of epidemic transmission occurs when an infected individual is in close contact with a susceptible host or through a transmission vector, as typically pathogens can only survive for a very limited time outside the host in an aerial environment. On the other hand, as air is typically a much harsher medium for pathogens than water, the sea is expected to host a large number of pathogens (viruses, bacteria and parasites) for a relatively long time. The longer life span of pathogens in a water medium, together with the increased buoyancy arising from the different physical properties of seawater and air, coupled to the existence of marine currents that can transmit pathogens for long distances away, allows diseases to spread faster and reach further distances in marine environments compared to epidemics in terrestrial systems (33). As a result, the possible long-term transmission of parasites by currents in marine environments make them more prone to suffer from persistent zoonotics compared to terrestrial ecosystems, where for an epidemic outbreak to occur the presence of an initial infected host (or vector) is necessary within a susceptible population. Until quite recently, marine zoonotics were mostly studied using different models compared to terrestrial diseases and it was not even clear whether the same tools could be applied (118). The abundance of pathogens in marine ecosystems is one of the reasons why proliferation models, that do not focus on transmission and assume a widespread occurrence of the pathogen and a rapid transmission problem, have

been most popular in the field (133). In fact, compartmental models are starting to be used only recently in the study of marine epizootics (18).

An important subset of marine organisms are sessile, e.g. bivalves, which means that they can not move. In the case of sessile terrestrial organisms disease transmission occurs mostly through vectors, insects that transmit the pathogens causing the disease. Instead, in marine ecosystems disease transmission is most often waterborne, in particular in passive water filtering feeders, as is the case of bivalves. Recently, some compartmental models considering the pathogen population have been proposed to study particular bivalve epidemics (16, 17, 18). In the present work we analyse a model that is aimed to describe disease transmission from an infected immobile host to a susceptible one of the same species through waterborne parasites, that are explicitly described. The model is closely related to the SIP model introduced in (18). In this first study we analyse in detail the properties of the mean-field version of the model, that aims to describe spatially homogeneous (i.e. well mixed) populations. The well mixed approximation will be valid whenever the mean distance among hosts is smaller than the mixing length of the parasites before they get inactivated or absorbed. The model is written such that waterborne transmission is the only mechanism by which one infected immobile host can infect a healthy one, and, thus, does not describe infection through direct contact. It is also assumed that the infected hosts, as invertebrates, do not have immune memory, and that the probability that an infected individual recovers is small and can be neglected. Thus, the model is not adequate to study infection of highly aggregated molluscs (like some mussels) or other passive filters like corals, as for these hosts one should also include the possibility of infection through direct contact. A first very relevant question is whether the model, describing infection of immobile (sessile) hosts through waterborne parasites can be reduced to a simpler version in which the parasite compartment is not needed. One exact and two approximate reductions are presented. We believe that the model can be most useful in the rapid characterisation of emergent marine epidemics if the right data from a well mixed system are available.

A very timely case study of such emerging epidemics is the noble fan mussel (or pen shell) *Pinna nobilis*. This fan mussel is the largest endemic bivalve in the Mediterranean Sea, and is under a serious extinction risk due to a Mass Mortality Event (MME) that has occurred throughout the whole Mediterranean basin very recently (68, 166, 176). Right before this MME, it was distributed across a wide type of habitats including coastal and paralic ecosystems at depths between 0.5 to 60m (31, 134). In open coastal waters, the distribution of the species is mainly associated with seagrass meadows, typically of *Posidonia oceanica*, which has been indicated as its optimal habitat (86). Its lifespan is up to 50 years in favourable conditions and its size can get up to 1.2m, placing it among the largest bivalves of the world (32). These fan mussels play a crucial ecological role in their habitat, as *P. nobilis* individuals filter water, thus retaining a large amount of organic matter from suspended detritus, contributing to water clarity (161). Furthermore, it is a habitat-forming species, because its shell provides a hard-surface within a soft bottom ecosystem, which can be colonised by different benthic species, augmenting biodiversity (32). In addition, at very dense populations, the species can function as an ecosystem engineer, creating biogenic reefs (99).

Despite *P. nobilis* populations have greatly declined due to anthropogenic activities in the 20th century (166), the ongoing MME is the most worrying and widespread threat to *P. nobilis* throughout the Mediterranean Sea. As a consequence, the species has been declared as critically endangered (91). Although different aetiological agents have been proposed, including Mycobacteria and other bacteria (34, 147, 149), there is evidence that the main cause of this mortality is the protozoan *Haplosporidium pinnae* (21, 38, 54), a new species that belongs to the genus Haplosporidium, one of the four genera of the protist order Haplosporida, where it has been found that other Haplosporidian parasites are behind the extensive mortality of several oyster species (8, 30). Life stages include uninucleate and binucleate cells, plasmodia, and spores. A group of experts following up the event predicted a high risk that the disease would be spread by marine currents through the Mediterranean basin, which could cause the extinction of the species (32) as it is endemic. This has helped to better understand the spread of the disease, and identified surface currents as the main factor influencing local dispersion, whereas environmental factors influence the disease expression, which seems to be favoured by temperatures above 13.5 °C and a salinity range between 36.5 and 39.7 psu.

In summary, we introduce and study in detail the properties of the mean-field version of

a general compartmental model to study marine epidemics for bivalve populations, namely passive filtering sessile invertebrate hosts infected through waterborne parasites. There are two main hypotheses, the first one that a population level description (i.e. without the consideration of spatial effects) is able to describe well the dynamics of the epidemic in a relatively dense population in small bounded regions. A second assumption is that the host becomes infected with some probability, but that there is not a critical parasite load in the infection process. After presenting the full SIRP model, then three different reductions are discussed, one exact, an approximate reduction of the former and a third reduction based on a timescale approximation. The study is closed with a validation with the available experimental data for the infection process of *Pinna nobilis* kept in tanks. We wish to point out, that being a highly endangered and protected species, the reported data correspond to an *unintended experiment* that cannot be repeated, and maybe these data represent the only opportunity to estimate the fundamental parameters of the model. In addition, the setup in which the *Pinna nobilis* were kept in tanks, represent themselves the ideal implementation of the conditions under which the mean-field model SIRP is valid.

4.2 The SIRP model

4.2.1 Model structure and initial considerations

In this work we analyse the SIRP model, a deterministic multi-compartmental mean-field model, continuous in time and unstructured in spatial or age terms, to study infection in bivalve populations. In particular, we stress that the model as it is written describes spatially-homogeneous populations. Compartmental models are the most frequently used class of models in terrestrial epidemiology (58), and originated in the classic study of SIR models by Kermack and McKendrick (102). The use of compartmental models in the study of infectious processes in marine systems is quite rare until very recently (84). As already advanced in Section 4.1, there are relevant features in the description of epidemic processes in marine ecosystems that are different with respect to the case of terrestrial ecosystems (118), and their study in marine environments is dominated so far by so called proliferation models (133), which do not address the transmission of the pathogen. See (18) for a discussion of several compartment models for the study of marine epizootics.

Compartmental models of diseases in terrestrial ecosystems caused by microparasites (i.e. viruses, bacteria and protozoans) do not consider a compartment to describe the dynamics of the parasite (117), describing just the different stages of the host. Infection typically occurs in 2 ways: i) as a contact process, in which the microparasite is transmitted directly from a an infected host, I, by contact or through air in close proximity, to a susceptible host, S; ii) through a vector, that has acquired the microparasite by biting an infected host, I, and passes the microparasite to a susceptible host, S. In the first case one can describe the infection process through some probability that the individuals come close, while in the second it is very relevant to describe the vector mobility, and at least 2 compartments, susceptible and infected vector, are typically needed. Once the microparasite enters the host, it proliferates inside it, so the infection process can be described by using compartments for susceptible individuals, S, infected individuals, I, and possible exposed individuals, E. In particular, transmission in terrestrial sessile organisms (e.g. plants) is generically vector-borne. In the case of marine ecosystems, infection typically occurs through water-borne parasites, in particular in filter-feeders sessile organisms, while vector-transmitted diseases are much less frequent. Parasites may be transported by diffusion, sea currents, or even active motion (i.e. if they have flagella). In any case, infection between sessile hosts is not through direct contact, but instead through the production and excretion of parasites by infected individuals and the assimilation by filtering of parasites by a healthy (susceptible) host. So, parasites are produced and excreted to the marine medium, in which they stay infective until they become deactivated (i.e die) or are absorbed by hosts. In a way, in parasite transmitted marine diseases parasites have a dual role: they are not only agents that induce infection but also act as vectors that transmit disease from an immobile infected host to a susceptible one.

The SIRP model is a general mean-field compartmental model to describe epidemic transmission through water-borne parasites, that we think is specially adequate to describe epidemic

transmission in sparsely located passive filter feeders, like many bivalves. We exclude the case of colonies in which individuals are in close proximity, e.g. mussels, corals, etc, in which direct contact could be relevant and should be included in the model. In the SIRP model hosts are described through 3 different compartments, as in the SIR model, that represent different evolution stages of the disease: a susceptible class of healthy individuals that may contract the disease, S , an infected class of individuals that may pass the disease through excretion of the parasite, I , and a class of removed (namely dead) individuals, R , that cannot be infected any more and that cannot transmit the disease plus an extra compartment, P , for the parasite population in the medium. It is important to note that invertebrates do not develop long-term immunity in the mammalian sense (133), and so, no compartment of individuals “recovered with immunity” is considered. However, bivalves have a first line of defense with hemocytes being able to fight parasites and reduce their internal population. Nevertheless, available evidence indicates that the number of individuals that can achieve a full recovery is usually small and can be neglected, and so it is not necessary to consider a process in which individuals in the I compartment return to the S compartment at some rate, like in the SIS model. Instead, the population’s long-term response to disease, when it occurs, is through natural selection for genotypes characterised by increased resistance to or tolerance for the pathogen. As already advanced, the SIRP model includes a fourth compartment that represents the parasite population in the water medium, whose population needs to be described explicitly. An explicit compartment allows to model the situation in which the population of parasites may evolve dynamically in time, although in Sec. 4.2.3.3 we will consider the case in which the parasite populations accommodates almost instantaneously to the infected host population, and the description of the parasite can be simplified.

Infection occurs when the host enters in contact with the parasite in the marine medium. It involves a process of filtration of water by the bivalve, and although a detailed representation has been discussed in the literature (16), in the current SIRP model it is represented in an effective way. Infection is modelled through a nonlinear term, typical in compartmental models, but that now depends on the parasite population and not on the population of the infected compartment, I . In terrestrial epidemiology there are two alternative ways to represent infection (116); i) mass action incidence, in which infection grows as the population gets larger, βSI ; ii) standard incidence, in which the infection is bounded as the population grows, $\beta SI/N$, where N is the total (host) population. One must look at these two choices as limit cases, with the possibility that in reality the system is best described by an intermediate form, closer to one of the limit cases, for example the modified SIR model in (23) in which the infection term has $S + I$ in the denominator instead of the total population N , because the R compartment is removed. Modelling infection with an explicit representation of the parasite population encounters the same basic dilemma about whether the incidence grows as the (host) population increases or is bounded. We will model infection as $\bar{\beta}SP$, where the two possibilities are equivalent to the two different incidences just mentioned: i) $\bar{\beta} = \beta$; ii) $\bar{\beta} = \beta/N$, where β , the disease transmission rate, is a constant, that can depend at most on external parameters, like temperature and salinity, but not on the variables defining the model (populations of host compartments or parasites). The model is valid considering both types of incidence, and in the case study we will see which incidence seems more adequate in this case.

4.2.2 General SIRP model

In this section we will write a mean-field compartmental model to describe epidemics of immobile (sessile) hosts in a marine medium through infection by a water-borne parasite. Being a mean-field model implies that the model is compartmental and does not include an explicit space dependence, and so, describes a well mixed system. The mean-field model describes a spatially homogeneous system, but we hope that it will be the basis for spatially inhomogeneous situations, by adding suitable terms accounting for the mobility of the parasite.

It is also assumed that the hosts become infected with some probability when exposed to a parasite, i.e., that there is not a critical parasite load needed for infection. The model is defined according to the following reaction processes,



which is graphically summarised in Fig. 4.1.

According to the scheme in Fig. 4.1, we consider the host in 3 possible states: susceptible, S , infected by the parasite, I and removed (dead), R . Then we introduce the parasite population in the medium (water), P . In the model, the $\bar{\beta}$, the disease transmission rate parameter regulates the infection rate of susceptible hosts and accounts, among other mechanisms, for the parasite intake rate, γ the mortality of infected hosts, being the inverse of the typical mean time for an infected host to die; λ the production rate of parasites from infected hosts, and μ the inverse of the typical life time of the parasite. μ can be related to several processes, like biological deactivation (or survival time) or other general losses, like dilution due to renewal of water in a closed experiment, natural losses in open ecosystems or absorption by other filter feeders. We are not considering the possibility of spontaneous parasite gain, i.e. immigration, in this version of the model. A summary of the model parameters can be found in Table 4.1. We do not consider vital dynamics for the hosts, and this implies that the sum of the 3 host subpopulations is constant, $N = S + I + R$, as the time scale of the disease evolution is much faster than the typical life cycle of fan mussels. The model is similar to the SIP presented in (18), except for an extra term in \dot{P} , $-\bar{\beta}PS$, accounting for the fact that when a parasite infects a host it is absorbed by it. The conditions under which the SIRP model can be simplified to the SIP model are discussed in Section 4.3.5.

In order to build the deterministic model we consider that the population is large enough to neglect fluctuations and that it is well mixed, so that spatial effects can be neglected. In this situation, we consider the infection process to be proportional to the number of parasites in the medium, so that the average number of contacts between susceptible fan mussels and the average parasite population is given by PS , and, thus, the change in the number of susceptible fan mussels takes the form $\dot{S} = -\bar{\beta}PS$, where the dot over a variable indicates a differentiation with respect to time: $\dot{S} = dS/dt$.

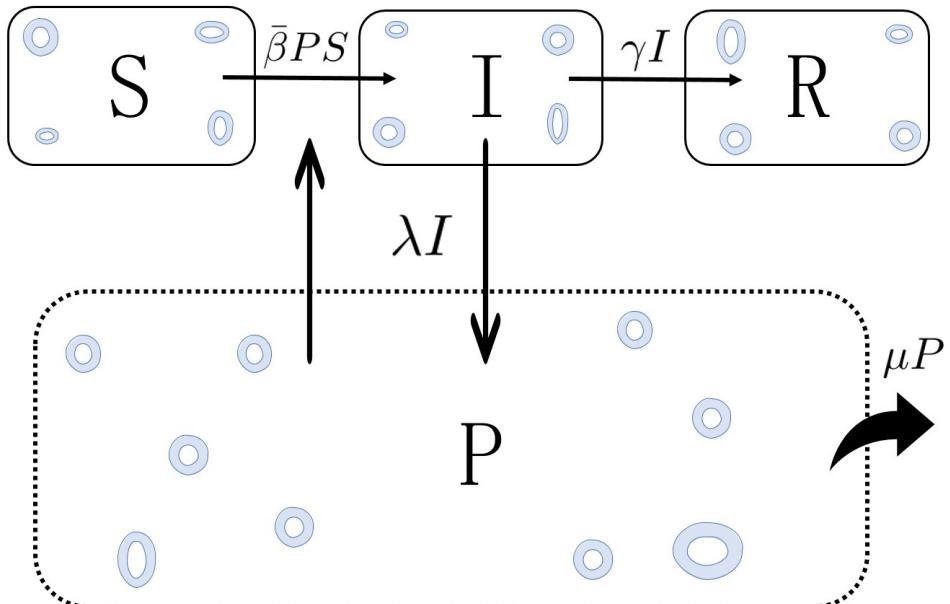


Figure 4.1: SIRP model flow diagram. The model variables are represented by capital letters: susceptible hosts (S), infected hosts (I), dead hosts (R) and the population of parasite (P). Arrows represent the processes in the model with their rates indicated next to them and blue rings represent parasites. The flow follows the scheme in Eq. (4.1), that leads to the system of differential equations in Eq. (4.2).

Table 4.1: Model parameters description

Variable	Definition	Parameter	Definition
S	Susceptible host	β	Disease transmission rate
I	Infected host	γ	Host mortality rate
R	Removed host	λ	Production rate of parasites by infected hosts
P	Parasite in the medium	μ	Parasite deactivation/dilution rate

Following this argumentation, the scheme in Eq. (4.1) and Fig. 4.1, one can write the evolution equations of the SIRP model,

$$\begin{aligned}\dot{S} &= -\bar{\beta}PS \\ \dot{I} &= \bar{\beta}PS - \gamma I \\ \dot{R} &= \gamma I \\ \dot{P} &= \lambda I - \bar{\beta}PS - \mu P.\end{aligned}\tag{4.2}$$

Model (Eq. (4.2)) lives in the 4-dimensional (S,I,R,P) phase space, representing the variables the populations of individuals in the susceptible, infected and removed host compartments and of parasites, respectively. These variables could be redefined so that S , I and R represent proportions of hosts in each compartment and P the population of parasites per host.

The fixed points of Eq. (4.2) are determined by the conditions¹ $I = P = 0$, to be fulfilled simultaneously. We will study the stability of the fixed point defined by $S(0), I(0) = P(0) = 0$ and $R(0) = N - S(0)$. A linear stability analysis of this fixed point reveals that it has two null eigenvalues, that stem from the condition $N = S + I + R$ and the conserved quantity of 4.6.1. The first condition, $S + I + R = N$, implies that it is enough to consider two of the host populations, e.g. S and I , as the third one can be obtained from the other two. The implications of the conserved quantity reported in 4.6.1 are more subtle, as it implies that fixed points are not isolated, as it happens in ordinary dissipative dynamical systems, and there is an infinite number (a line of) fixed points for the final state of the epidemic, depending on the initial conditions. This also implies that the phase space is foliated by the conserved quantity, C of Eq. (5.13), and every initial condition, S_0 , with a different value of C leads to a different asymptotic condition, S_∞ , just as shown in (126) for the SIR model (cf. Fig. 10.1 in *op. cit.*). The third eigenvalue, that is the largest of the two non-zero eigenvalues, can be positive if $\beta S_0 \lambda > \gamma(\beta S_0 + \mu)$ and negative if the inequality is reversed, defining the conditional stability of the fixed point. The fourth eigenvalue is always negative and all the eigenvalues are always real (cf. 4.6.2). The instability of the fixed point along the third eigenvalue drives the beginning of the epidemic.

An extremely important result in epidemiology is the so-called *basic reproduction number*, R_0 , a dimensionless number which represents the number of secondary infections produced by a primary infection in a fully susceptible population. $R_0 = 1$ defines the threshold for epidemic propagation: an epidemic will occur when $R_0 > 1$, and the number of infected individuals will grow, at an exponential rate in the early phases of the epidemic (37), while if $R_0 < 1$ the infection will wane naturally. This quantity can be formally obtained making use of the Next Generation Matrix (NGM) method (57, 59). Applying this formal method to our system of ordinary differential equations (ODE's) one obtains the following relation for the basic reproduction number (cf. 4.6.3),

$$R_0 = \frac{\lambda}{\gamma \left(1 + \frac{\mu}{\bar{\beta}S(0)} \right)}.\tag{4.3}$$

The threshold condition provided by R_0 (Eq. (4.3)) is equivalent to the linear stability condition for the third eigenvalue of the initial, pre-epidemic, fixed point, as $\bar{\beta}S(0)\lambda > \gamma(\beta S(0) + \mu)$ implies that this eigenvalue is positive and the disease-free equilibrium state unstable being this equivalent to $R_0 > 1$ (cf. 4.6.2). Thus, if $R_0 > 1$ the fixed point is unstable, and an epidemic will ensure if infected hosts, I , or parasites, P appear in the system. An epidemic will propagate until the system reaches an stable fixed point, that signals the end of the epidemic (cf. 4.6.2).

¹We do not consider the trivial fixed point $S = I = R = P = 0$ that would imply $N = 0$ and $P = 0$ at all time.

4.2.3 Model reduction

The SIRP model lives in a 4-dimensional phase space and depends on 4 parameters, what makes difficult to confront it with experimental data. Thus, we will discuss here three alternative ways of reducing the model. The first involves an exact reduction of the model, based on the conserved quantity derived in 4.6.1. The second reduction consists of an approximation to the previous exact reduction, that turns out to be equivalent to an exact reduction of a slightly simplified model (without the $-\bar{\beta}SP$ term in the equation of \dot{P}). The third one is based on an approximation valid if the system parameters fulfil certain conditions.

4.2.3.1 Exact reduction of the SIRP model

From the conserved quantity derived in 4.6.1, it is possible to write the parasite population in the SIRP model as a function of the host states as follows,

$$P(S, I) = -\frac{\lambda}{\gamma}(S + I) + \frac{\mu}{\bar{\beta}} \ln(S) + S + C(0), \quad (4.4)$$

$$\text{where } C(0) = P(0) + \frac{\lambda}{\gamma}(S(0) + I(0)) - \frac{\mu}{\bar{\beta}} \ln(S(0)) - S(0).$$

Substituting Eq. (4.4) into the general SIRP model of Eq. (4.2) we obtain the following nonstandard SIR model,

$$\begin{aligned} \dot{S} &= \frac{\lambda \bar{\beta}}{\gamma} S(S + I) - \mu S \ln(S) - \bar{\beta} S^2 - S \bar{\beta} C(0) \\ \dot{I} &= -\frac{\lambda \bar{\beta}}{\gamma} S(S + I) + \mu S \ln(S) + \bar{\beta} S^2 + S \bar{\beta} C(0) - \gamma I \\ \dot{R} &= \gamma I. \end{aligned} \quad (4.5)$$

Although using the conserved quantity yields an exact reduction from a 4D dynamical system to a 3D one, the number of independent parameters and initial conditions remain unchanged, i.e. they still depend on 4 parameters and 4 initial conditions. Thus, although useful, (4.5) is not ideal when trying to fit experimental data, and this is the reason for trying a further approximation to Eq. (4.5) to be discussed next.

4.2.3.2 Further approximation to the exact reduction

A further approximation to Section 4.2.3.1, that is less restrictive and expected to be valid in a broader parameter range than Section 4.2.3.3 is possible. This approximation reduces the number of free parameters by one, what is useful in fitting available data. The approximation consists of neglecting the S term in Eq. (4.4), what is possible if $\lambda/\gamma \gg 1$ and also $\mu \ln N / (\bar{\beta} N) \gg 1$, as $S(t)$ decreases monotonically with time and is, at most, N at the initial time. Interestingly, this approximation is equivalent to the simplification of the equation for \dot{P} in (4.2) so that the $-\bar{\beta}SP$ is skipped, what yields exactly the SIP model of Ref. (18). This reduced model has an exact conserved quantity, \mathcal{C} , that differs from that of the SIRP model in the linear S term (cf. 4.6.1). Using this approximation one can write,

$$\begin{aligned} \dot{S} &= \frac{\lambda'}{\gamma} S(S + I) - \mu S \ln(S) - S \bar{\mathcal{C}}(0) \\ \dot{I} &= -\frac{\lambda'}{\gamma} S(S + I) + \mu S \ln(S) + S \bar{\mathcal{C}}(0) - \gamma I \\ \dot{R} &= \gamma I, \end{aligned} \quad (4.6)$$

where $\lambda' = \lambda \bar{\beta}$ and $\bar{\mathcal{C}}(0) = \bar{\beta}(P(0) + \lambda/\gamma(S(0) + I(0) - \mu/\bar{\beta} \ln S(0))) = \bar{\beta} \mathcal{C}(0)$ is a redefinition of the conserved quantity of the SIP model Eq. (4.18), $\mathcal{C}(0)$, a constant, such that it absorbs $\bar{\beta}$ and all initial conditions of the model. The result is that Eq. (4.6) depends on 3 parameters and 1 constant, compared to Eq. (4.5) that depends on 4 parameters, facilitating, thus, the use of the model to fit experimental data.

4.2.3.3 Model reduction through fast-slow separation

The third reduction of the 4-D dynamical model Eq. (4.2) makes the assumption that the time scale in which the parasite population changes are faster than the one corresponding to the host. This means that pathogen deactivation in the medium must be faster than host mortality. In terms of the rates associated to each of these processes, this means $\mu > \gamma$. Taking μ as common factor in \dot{P} one can write,

$$\varepsilon \dot{P} = \lambda I / \mu - \bar{\beta} S P / \mu - P \quad (4.7)$$

where $\varepsilon = 1/\mu$ is small, as μ is large. If furthermore $\mu \gg \bar{\beta}N$ and $\lambda \gg \beta P$ one arrives to,

$$P \approx \frac{\lambda}{\mu} I. \quad (4.8)$$

Under this approximation the slow subsystem can be written,

$$\begin{aligned} \dot{S} &= -\beta' IS \\ \dot{I} &= (\beta' S - \gamma) I \\ \dot{R} &= \gamma I, \end{aligned} \quad (4.9)$$

that is equivalent to the classical SIR model with $\beta' = \bar{\beta}\lambda/\mu$ instead of the infection rate $\bar{\beta}$. The reduced 3-D model Eq. (4.9) from the original 4-D SIRP model Eq. (4.2) depends on 2 parameters instead of 4 as the original model had and 1 initial condition, e.g. $I(0) = N - S(0)$ if $R(0) = 0$, and is much more amenable to be applied to the analysis of experimental data, as shown in Section 4.4. Furthermore, γ could be eliminated through a time rescaling, $t \rightarrow t' = \gamma t$ with a redefinition of $\beta' \rightarrow \beta'' = \beta'/\gamma = \beta\lambda/(\mu\gamma)$, leaving the model as a function of a single effective parameter. However, we will keep both β' and γ for convenience when fitting the model to experimental data in Section 4.4, as we would need to know anyhow γ in order to analyse the experimental data. The validity of this approximation is checked numerically in Section 4.3.

4.3 Numerical analysis of the model

Due to the impossibility of solving the SIRP model analytically, in the present section we perform a numerical characterisation of the model². Moreover, we show the validity range of the performed approximations to reduce the SIRP model to an effective SIR model. We start our numerical analysis by investigating the relative influence of the model parameters on some epidemiological quantities of interest: the basic reproduction number (R_0), related to the existence of an epidemic outbreak, continuing with the final state of the epidemic, given by the final number of dead individuals ($R(\infty)$) and the maximum of infected individuals (I_{\max}) together with the time at which it occurs (t_{\max}).

In order to identify the most influential parameters of our model Sensitivity Analysis (SA) will be performed. SA can be divided into two classes: Local Sensitivity Analysis (LSA) and Global Sensitivity Analysis (GSA). LSA represents the assessment of the local impact of input factors variation on model response by concentrating on the sensitivity in the vicinity of a set of factor values. Such sensitivity is often evaluated through gradients or partial derivatives of the output functions at these factor values, such that other input factors are kept constant. Since epidemic models exhibit a threshold behaviour, controlled by the dimensionless quantity R_0 , it is relevant to study its robustness with respect to small perturbations by means of the LSA explained above, as its analytical expression is known.

On the other hand, GSA will be applied to study the influence of the parameters in the final state of the epidemic and the epidemic peak by exploring a large domain of the parameter space. In turn, GSA is the process of apportioning the uncertainty in outputs to the uncertainty in each input factor over their entire range of interest. A sensitivity analysis is considered to be global when all the input factors are varied simultaneously and the sensitivity is evaluated over the entire range of each input factor, in clear contrast to LSA. Within GSA, first order indices are

²All numerical simulations of the dynamical system Eq. (4.2) have been carried out using a Runge-Kutta 4th order method, with a temporal step $\Delta t = 0.001$. Numerically stable results are obtain with $\Delta t \leq 0.01$.

a measure of the contribution to the output variance given by the variation of the parameter alone averaged over variations in other input parameters while second order indices take into account first order interactions between parameters. While LSA is carried out analytically (if exact expressions are available), GSA is a purely numerical approach. Further mathematical details on Sensitivity Analysis can be found in 4.6.4.

For all the sensitivity analysis performed in the following sections, and in order to avoid ambiguities associated to the definition of $\bar{\beta}$ as a function of N , we assume $N = 1$, so that both possible incidences yield $\bar{\beta} = \beta$ and the numerical results are equivalent.

4.3.1 The basic reproduction number R_0

To study the relevance of parameters involved in an epidemic outbreak a LSA was performed. We analyse the local sensitivity of R_0 through the normalised sensitivity index, so that the function $F(\vec{p})$ of Eq. (4.26) is substituted by the analytical expression of R_0 , Eq. (4.3).

Fig. 4.2(a) shows the sensitivity index for R_0 for specific baseline parameters, where λ , $\bar{\beta}$ and S_0 contribute to increase the basic reproduction number while α , γ and μ contribute to decrease it, as expected. Moreover, we can see that λ and γ are the most influential parameters while μ , $\bar{\beta}$ and S_0 depend on each other. These dependencies cause varying influences on R_0 , which are fully depicted in panels Fig. 4.2(b-d). It can be seen that the influence of $\bar{\beta}$ increases with the increase of μ and the decrease of S_0 . Similarly, the importance of S_0 increases with μ and decreases with $\bar{\beta}$. On the other hand, the impact of μ increases with the decrease of both S_0 or $\bar{\beta}$.

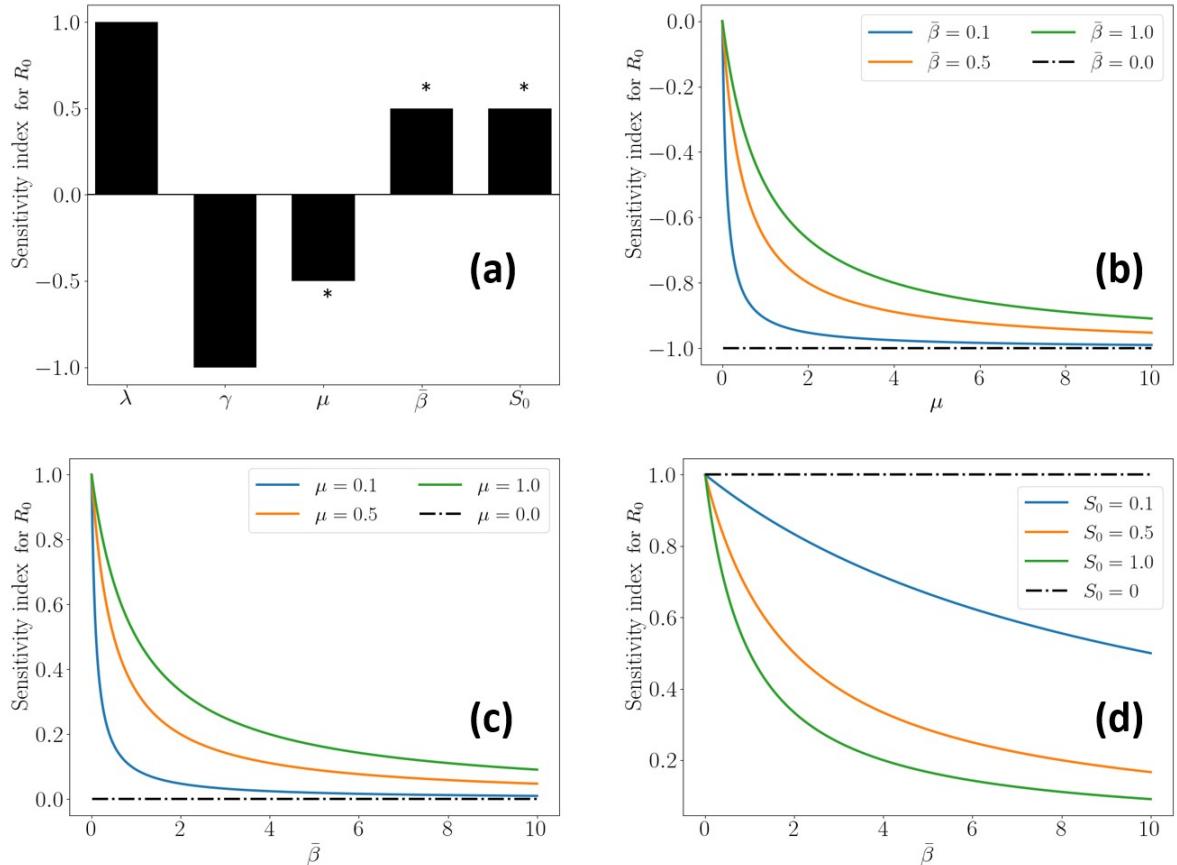


Figure 4.2: Panel (a): Local sensitivity analysis of R_0 for the baseline parameters $\lambda = 1$, $\gamma = 1$ and $\mu = \bar{\beta} = S_0 = 1$. The asterisks mark parameters for which the sensitivity index is not constant, depending on, at least, another parameter. Panels (b-d): Local sensitivity analysis of R_0 with respect to parameters with an asterisk, showing the different dependence with a second parameter and the effect on the varying sensitivity index.

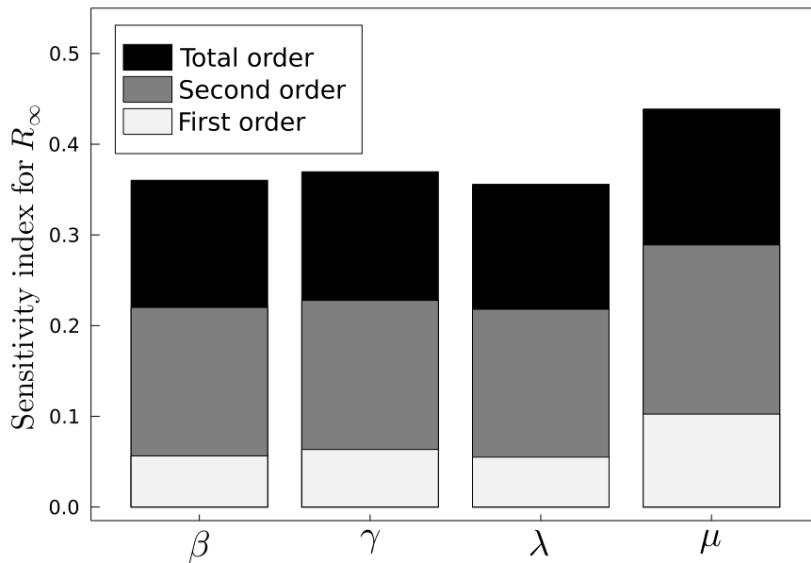


Figure 4.3: Sensitivity indices (LSA) for the final number of dead individuals (R_∞) for each one of the indicated parameters. The black bars represent the total order indices of sensitivity while white (grey) colour represents the contribution of the first (second) order indices.

4.3.2 Final state of the epidemic

Another important quantity in epidemiology is the final state of the epidemic, which can be characterised by the final number of dead individuals, R_∞ . Within our general SIRP model it is not possible to find an analytical expression of $R(t)$ so that we need to tackle the problem numerically. To this end, we perform GSA for the final number of dead individuals in order to determine the most influential parameters for this quantity. In particular, we apply the Sobol method, discussed in 4.6.4. The Confidence Interval, CI, obtained in our study is less than 1% of the index value, indicating a very high accuracy, therefore it is not shown in the figures. The results of the explained procedure are shown in Fig. 4.3, where the total order (black), first order (white) and second order (gray) sensitivity indices for each of the model parameters are detailed. It can be observed that μ has a slightly greater influence than the other parameters with respect to the final number of dead individuals. Note that the second order indices are larger than the first order ones for all the parameters, which indicates a high influence of the nonlinearities in our model, at least for the particular quantity under study.

4.3.3 Maximum of infected individuals

A GSA of the maximum number of infected individuals, I_{\max} and the time it occurs, t_{\max} is performed to study the influence of the model parameters regarding these quantities. In this case, Fig. 4.4, γ has greater influence in the epidemic peak than any of the other parameters, while for the time at which the peak takes place, all the parameters have basically the same influence. Again, the second order indices (the first order interactions between parameters) account for most of parameter sensitivity, in particular in the time of the epidemic peak, indicating the high degree of nonlinearity of this effect.

4.3.4 Numerical verification of the fast-slow approximation

The parasite concentration approximation, based on a timescale separation discussed in Section 4.2.3.3, is now verified by computational means. The verification was performed using both mass action and standard incidence, but for the sake of simplicity we show only the results for the standard incidence case. Worth is to say that, mathematically, changing from standard incidence to mass action involves only a rescaling of the β parameter, so that the numerical results are exactly the same. Fig. 4.5 contains a comparison for 3 different values of the parasite deactivation rate, μ . It can be seen that the approximation is poor when $\mu \sim \gamma, \bar{\beta}N$ Fig. 4.5(a), as it could be expected. On the other hand, the approximation is quite good when μ is one order

of magnitude larger than γ and $\bar{\beta}N$ Fig. 4.5(b), while it is extremely accurate when μ is two orders of magnitude larger than $\gamma, \bar{\beta}N$, Fig. 4.5(c). The figure also shows the numerical value of \dot{P} (pink dashdot), and it can be checked how it becomes smaller as μ increases compared to $\gamma, \bar{\beta}N$, justifying the timescale separation of Section 4.2.3.3. Finally, Fig. 4.5(a-c) also shows (dashed red line) the analytical value for $P(S, I)$ derived in Eq. (4.16), that matches perfectly the result of the numerical integration of Eq. (4.2), as should be the case.

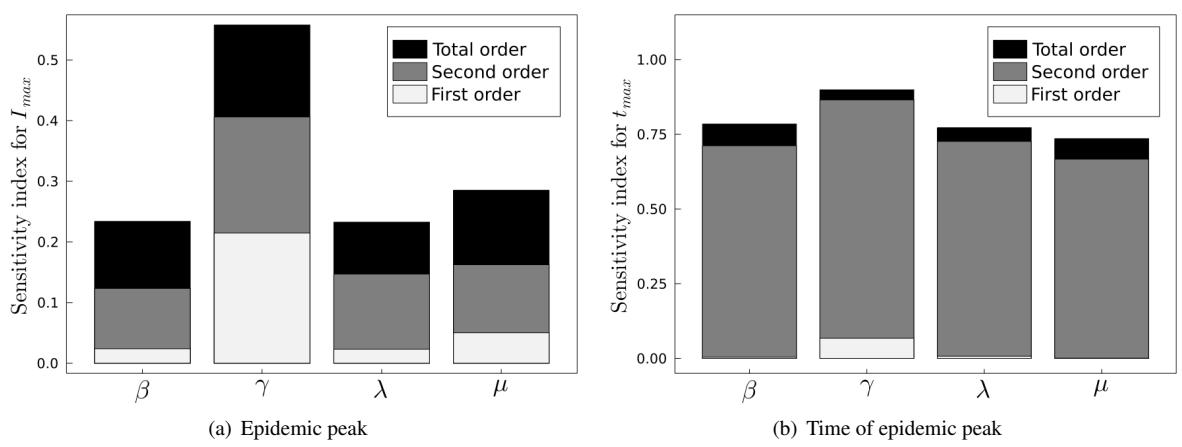


Figure 4.4: Global sensitivity analysis for the maximum of infected individuals I_{\max} (a) and its time occurrence t_{\max} (b). The black bars represent sensitivity at all orders, while white (grey) colour represents the contribution of the first (second) order indices.

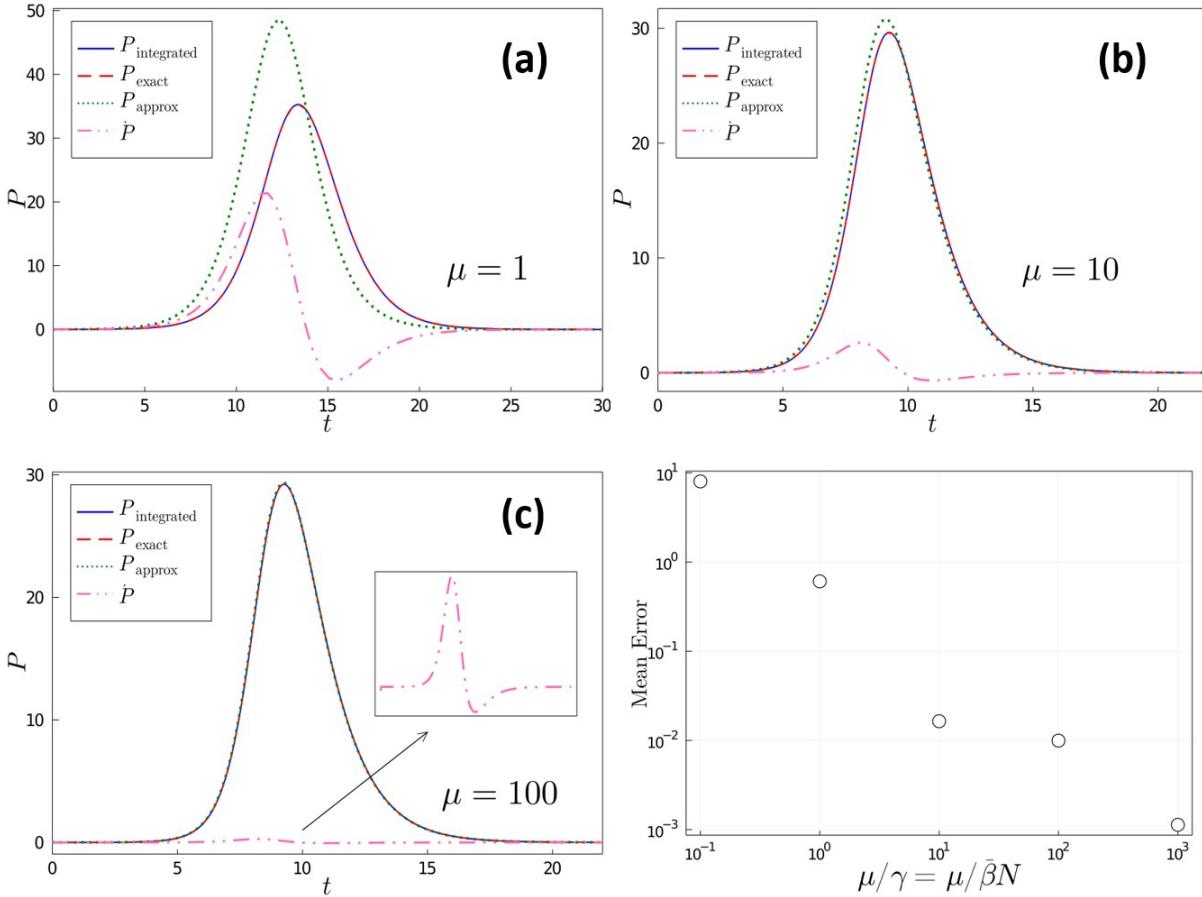


Figure 4.5: Numerical check of the approximate expression for the pathogen concentration, (Eq. (4.8)), $\bar{\beta} = 1/50$ and $\gamma = 1$: (a) $\mu = 1$; (b) $\mu = 10$; (c) $\mu = 100$, while λ is varied to keep $R_0 = 2.5$, defined in (Eq. (4.3)) (with $S_0 = N = 50$), i.e., $\lambda = 5, 27.5, 252.5$ respectively for (a)-(b)-(c), respectively. The blue solid line represents the numerically integrated quantity, the red dashed line (superimposed to the blue solid one as they are identical) is the exact solution for this quantity, (Eq. (4.4)) and the green dotted line accounts for the approximate expression from the timescale separation (Eq. (4.8)). The dash-dotted pink line represents the derivative of P , \dot{P} , in the scaled time frame. Panel (d): Mean error between the approximate and exact solutions for increasing $\mu/\gamma = \mu/\bar{\beta}N$.

4.3.5 Numerical verification of the model approximation from the exact reduction

The numerical verification was performed for both mass action and standard incidence, but for the sake of simplicity in Fig. 4.6 we show only the results for the standard incidence case. First, and as it should be because it is an exact result, the exact reduction of the SIRP model discussed in Section 4.2.3.1 matches perfectly the numerical results obtained from the full model for all possible parameter values, Fig. 4.6(a-c). Regarding the approximation to the exact reduction, one can see how the approximation converges to the exact solution as the parameters fulfil the conditions indicated in Section 4.2.3.2, namely that both $\gamma/\lambda \gg 1$ and $\mu \log(N)/\beta N \gg 1$, becoming very accurate if these ratios are larger than 1 by two orders of magnitude or more (cf. Fig. 4.6(c)). We recall that in this case the SIRP model converges to the SIP model of (18). Conversely, the approximation is poor when any of these two ratios is of order 1 ((cf. Fig. 4.6(a))), while Fig. 4.6(b) presents the result in an intermediate case, in which the approximation is fair.

4.4 Model validation with data of the *Pinna nobilis* Mass Mortality Event

In this section, the general SIRP model is validated against collected data from the *Pinna nobilis* Mass Mortality Event. As explained in Section 4.1, the disease is caused by the parasite *Haplosporidium pinnae* and the hosts, *P. nobilis*, are sessile bivalves endemic of the Mediterranean Sea. Thus, this epidemic is a perfect candidate to be described by the SIRP model. In the model,

parasite production occurs only inside infected hosts, and parasites are released to the medium, either through their respiratory or digestive system. The simultaneous occurrence of the different possible stages of the parasite (uni- and bi-nucleate cells, multinucleate plasmodia, sporocysts and uninucleate spores) in the same host individual is not common among haplosporidans and makes *H. pinnae* different from previously known haplosporidian species (38). The occurrence of uni- and binucleate stages suggest possible direct transmission from infected to healthy fan mussels, as observed in *B. ostreae* and *B. exitiosa* (10, 51, 87). Additionally, the presence of spores (a dormant, resistant stage) could allow long persistence in the environment and the hypothetical involvement of an intermediate host as suggested for *H. nelsoni* and *H. costale* (6, 85, 132). While uninucleate cells are always detected in infected fan mussels, sporulation has been only detected sporadically (38). Thus, we assume that infection occurs mostly through uninucleate (or binucleate) cells by direct transmission (as the experimental observations in captivity point out, see (68)). We do not consider disease transmission through other stages. We do not consider spores, given the infrequent observation of spores and the current lack of experimental information about spore transmission (that could involve another intermediate host species). Regarding plasmodia and sporocyst stages, these stages are too large to be released through the epithelium. The distinction between uninucleate and binucleate cells seems unnecessary at this level of representation, as these phases only participate in parasite proliferation inside infected hosts, a process that we consider in an effective way. Finally, the evidence of the time course of the disease compared to the long life cycle of *P. nobilis* suggests host vital dynamics (i.e. recruitment (reproduction) and natural death) can be neglected.

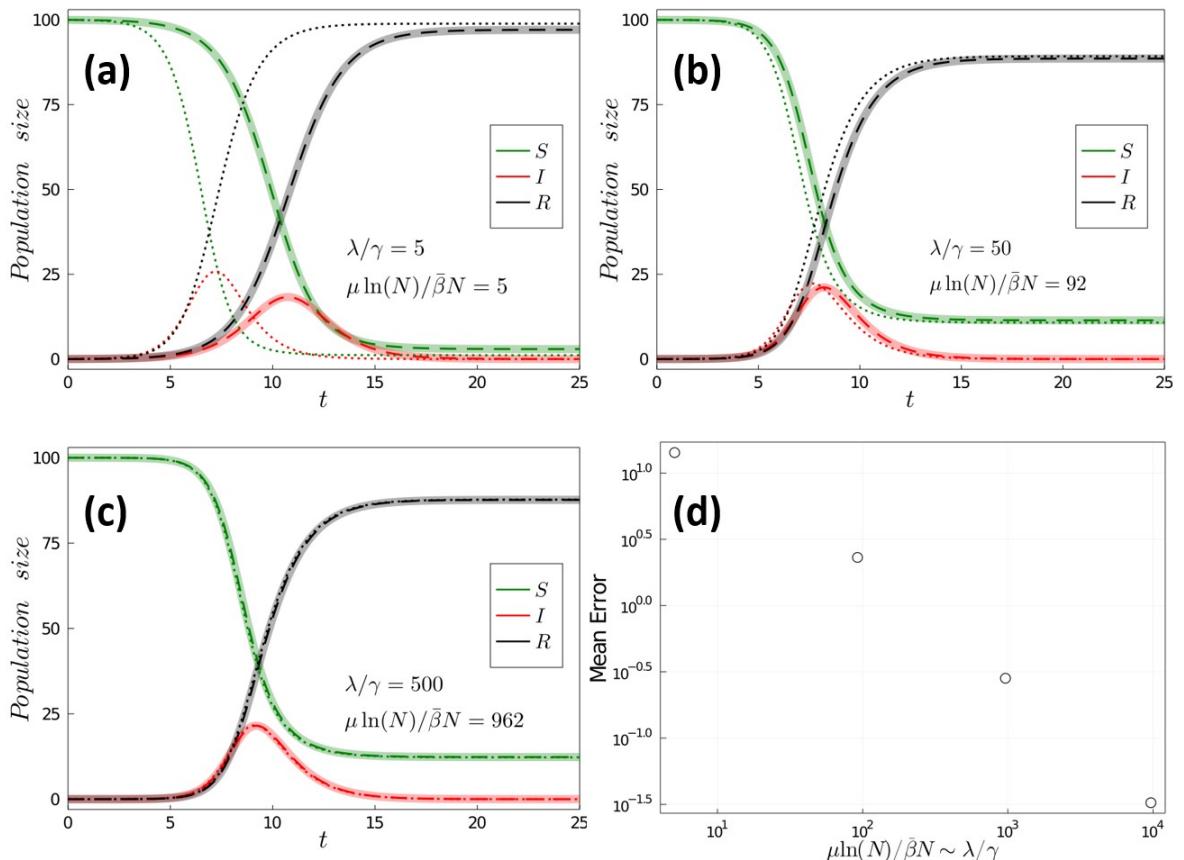


Figure 4.6: Numerical check of the exact model reduction along with the subsequent approximation shown in Section 4.2.3.1 with $N = 100$, $\bar{\beta} = 1/100$, $\gamma = 1$, (a) $\lambda = 5$, $\mu = 1.1$; (b) $\lambda = 50$, $\mu = 20$; (c) $\lambda = 500$, $\mu = 209$. $R_0 = 2.38$ for all the panels. The solid semitransparent lines represent the original 4D model, the dashed lines the exact reduction and the dotted lines the approximate model from the exact reduction. Panel (d): Mean error between the approximate and exact solutions for increasing $\mu \ln(N)/\bar{\beta}N$ and λ/γ while $R_0 = 2.38$ is kept constant.

After an epidemic outbreak that took place in Portlligat, in the north east of Catalonia, 215 *Pinna nobilis* individuals were extracted from their natural medium in order to be preserved as a genetic reserve in several controlled water tanks of different institutions in Spain (68). The institutions that participated in this preservation effort were IFAPA, IEO, IRTA, IMEDMAR-UCV and Oceanogràfic of Valencia. The original idea was to rescue the individuals before infection, however, the subsequent evolution of the rescued *Pinna nobilis* populations indicates that some individuals were already infected at the time of extraction (and/or in contact with some amount of the parasite transferred from sea water). This allowed the opportunity to use the data of the time evolution of the epidemic in the controlled water tanks, reported in (68), to evaluate the described SIRP model³. The empirical data consists of the proportion of survivors as a function of time in the controlled water tanks with a temporal resolution of one month. Despite the fact that the temperature of the water in the tanks was controlled, it was sharply lowered in most of the tanks when mortality started to appear within the population, as a last effort to keep the rest of the population safe and alive, since keeping the temperature below approximately 13.5°C is a known strategy to preserve *Pinna nobilis* individuals as disease expression is minimal (32). Fortunately, two of the tanks kept its temperature approximately constant during the full recorded time. This is the case of the tanks in IFAPA in Huelva and the Oceanogràfic of Valencia (OCE), both Spanish institutes. These water tanks have been selected to validate our model, maintaining constant temperatures of 14°C and 17°C, respectively.

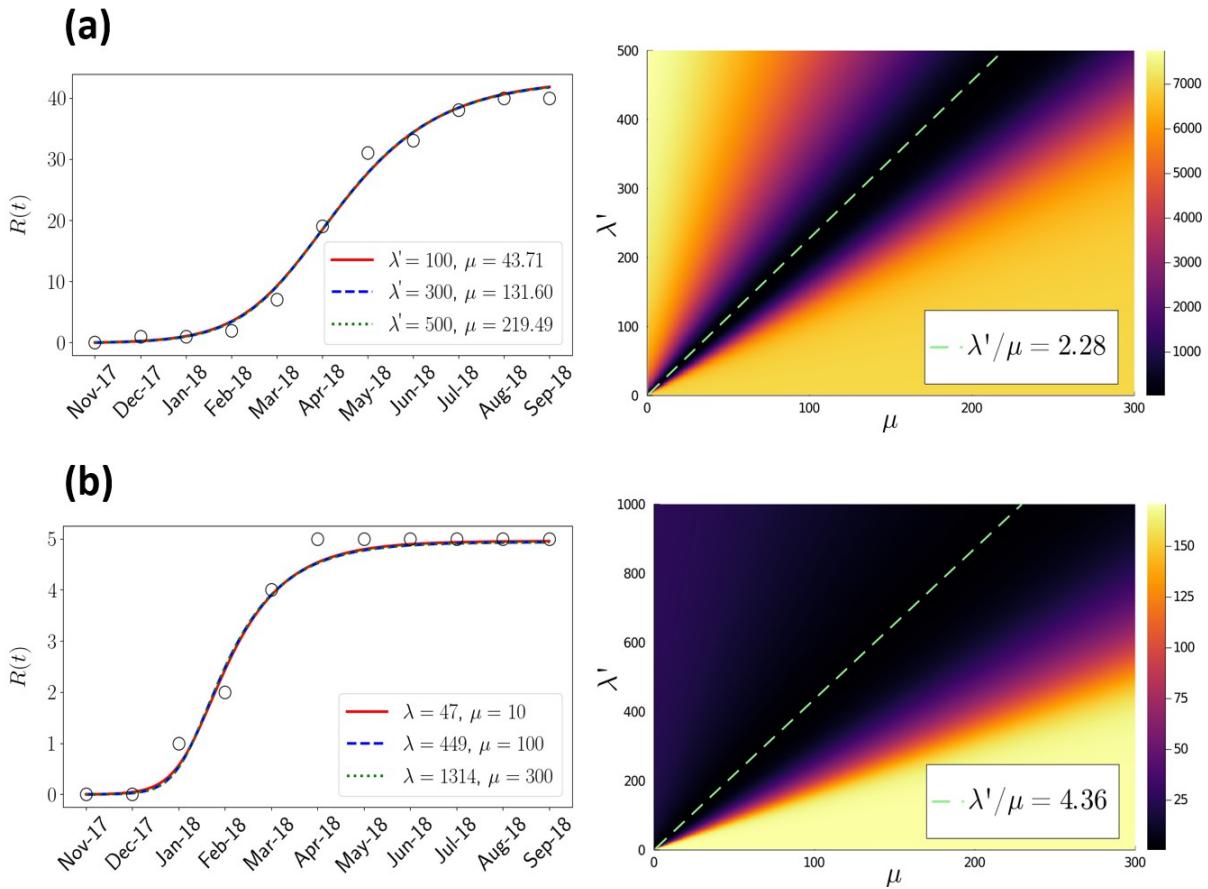


Figure 4.7: Parameter estimation for the approximation from the exact reduction of the SIRP model (Eq. (4.6)) using data from IFAPA (panel (a)) and OCE (panel (b)) water tanks, at 14°C and 17°C respectively. Left figures represent several fits of the model to empirical data of the number of dead hosts ($R(t)$) using different optimal combinations of the parameters. Right figures are the RSS errors as a function of the input parameters, where the green dashed line represents the set of optimal combinations of the parameters with $RSS = 60, 0.8$ for IFAPA and OCE, respectively.

³Data use in this work with the purpose of validating and fitting parameters for the SIRP model have been taken from the Supplementary Information of [68]

First we will fit the exact reduction of the SIRP model, assuming $\mu \log(N) \gg \bar{\beta}N$ and $\lambda/\gamma \gg 1$ as discussed in Section 4.2.3.2, namely Eq. (4.6). This reduced model depends on three parameters (λ', μ, γ) and one constant, $\mathcal{C}(0)$, cf. Section 4.2.3.2, that is related to the initial conditions of the model. The order of magnitude of the mortality rate can be deduced from data, with an estimate value of $\gamma \approx 1 \text{ month}^{-1}$. We fix this parameter in order to give some biological information to our model prior to the computational fit. We focus on the R compartment, as it can be retrieved directly from data in (68)⁴. We use a box-constrained variant⁵ of the well known BFGS optimisation algorithm (66) with a common L2 loss function, also known as Residual Sum of Squares (RSS)⁶. By running this algorithm one observes that the optimal parameters tend to be the ones in the boundary of the box-constrained parameter space. Furthermore, if the box size is increased (or decreased) the optimal parameters continue to be in the boundary of the box-constrained parameter space. This indicates that there exist several parameter combinations that optimally fit the data, and the combination parameters found by the optimisation algorithm are only marginally optimal with respect to other parameter values. The locus (actually a valley) of marginal optimal parameters can be seen in the right hand side panels of Fig. 4.7, where the cost function value of the optimisation algorithm is plotted as heat map.

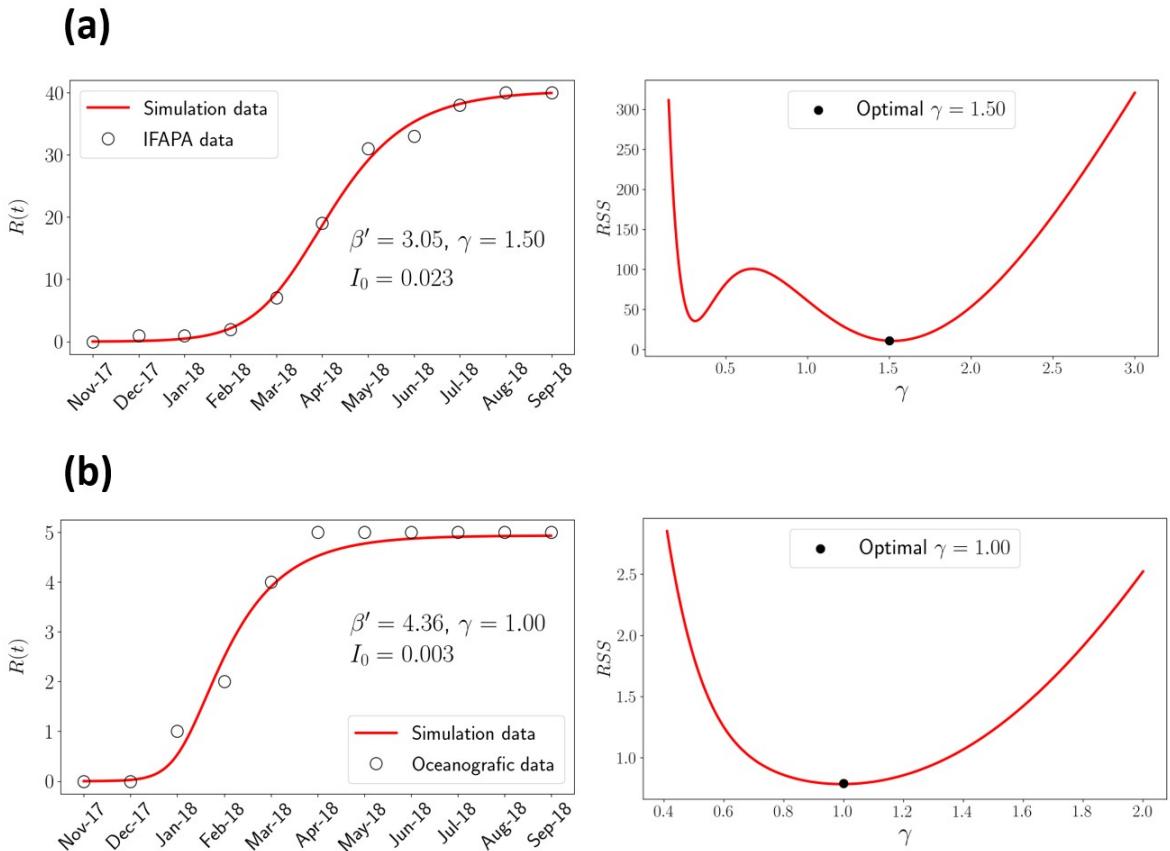


Figure 4.8: Parameter fitting for the R compartment to model (Eq. (4.9)) using data from IFAPA (panel (a)) and Oceanogràfic (panel (b)). The left part of both panels of the figure shows the optimal fit of the model to empirical data with $RSS = 10.9, 0.8$ for IFAPA and OCE, respectively. The right panels show the variation of the RSS error for some values of γ . The β' values have been obtained assuming a standard incidence, as explained in the main text.

Now we reach the point regarding the dilemma between mass action and standard incidence discussed in Section 4.2.1. If one does not correct the $\bar{\beta}$ parameter with the size of the

⁴The number of dead individuals can be obtained as $R = N - S$, where S is the population of survivors and N is the total number of individuals in the tanks, 50 (IFAPA) and 5 (Oceanogràfic), respectively

⁵We constrain the optimisation because the unconstrained optimisation to the full range of the parameters, i.e., from 0 to ∞ is not practical.

⁶The algorithm is implemented within the Julia high-level programming language [15] using the DifferentialEquations.jl package [136].

host population, N , that is equivalent to assuming the mass action incidence $\bar{\beta} = \beta$, the values that one would obtain for $\beta' = \beta\lambda/\mu = \lambda'/\mu$ for both populations take disparate values in both tanks: $\beta' = 0.046$ for the IFAPA data set and $\beta' = 0.87$ for the Oceanogràfic (OCE) data set, a factor of 19 between them while their temperatures differ only by 3°C . These numbers indicate that the standard incidence is more reasonable, what amounts to choosing $\bar{\beta} = \beta/N$, where the final values of the reported parameter β' should be multiplied by $N = 50$ for the IFAPA tank and $N = 5$ for the OCE tank. The final result is then $\beta' = 2.28$ and $\beta' = 4.36$ for IFAPA and OCE tanks, that are the values reported in Fig. 4.7, implying that an almost twofold increase of the β' parameter corresponds to an increase of 3°C . This relation is in good agreement with the typical changes in rates of a wide range of organisms with a 3°C change in temperature, while a 19-fold change in the rate would imply at least a 30°C change in temperature (cf. 4.6.5).

The fact that there is an infinite number of combinations of the parameters that optimally fit the real data suggests that, as two parameters are slaved one to each other, that the model admits a further reduction. This reduction corresponds exactly to the approximate *SIR* model derived in Eq. (4.9), with the relationship $\beta' = \lambda'/\mu$, as anticipated. So, this gives further corroboration to the use of the *SIR* model Eq. (4.9) to fit β' as the free parameter (fixing the value of γ and with I_0 as the initial condition determined by the fit). For consistency with the previous fitting we expect to obtain $\beta' = 2.28$ and 4.36 as the optimal parameters for the IFAPA and OCE water tanks, respectively, and this is the case.

Interestingly, as reduced model Eq. (4.9) has fewer parameters to fit we can relax our initial assumption of $\gamma = 1 \text{ month}^{-1}$ and check how the fit improves or worsens when varying γ . In Fig. 4.8 a fit of the reduced SIR model Eq. (4.9) is shown for the IFAPA (top) and Oceanogràfic (bottom) controlled water tanks⁷. Fig. 4.8(c-d) shows the *RSS* error as γ is varied. It can be seen that for the IFAPA water tanks $\gamma = 1.5 \text{ month}^{-1}$ yields more accurate results, while for the Oceanogràfic water tanks $\gamma = 1 \text{ month}^{-1}$ remains optimum. This shows a decrease in the mean removal time $1/\gamma$ for lower water temperatures, with the finite size errors inherent to the OCE tank (as $N = 5$). In the left panels the simulated curve of dead individuals, R compartment, as a function of time for the optimal fitted parameters is confronted to the experimental data, showing a remarkable agreement. With the optimum values of γ , in the IFAPA tank (now with $\gamma = 1.5 \text{ month}^{-1}$) a new value of $\beta' = 3.05$ is obtained, implying a probably more reasonable ratio of 1.43 for β' in both tanks (it was 1.91 in the original fit). From the optimal parameters we obtain the basic reproduction number, since $R_0 = \beta'/\gamma$ we have that $R_0^{\text{IFAPA}} \simeq 2$ and $R_0^{\text{OCE}} \simeq 4$, clearly above the epidemic threshold.

Summarising, the SIRP model is able to fit two sets of experimental data, agreeing with a standard incidence, according to which the infection rate depends on the amount of parasites per pen shell individual. *Pinna nobilis* individuals in the IFAPA experiment were actually distributed in 4 tanks, and the standard incidence is compatible with this experimental aspect. The temperature dependence of the fitted parameters in this range ($14 - 17^\circ\text{C}$, appears to be compatible (although experiments at different temperatures would be needed) with an Arrhenius dependence of the infection parameters, also known as Boltzmann-Arrhenius (27, 120), that can be extended to account for the expect unimodal dependence on temperature, with a maximum infectivity at a characteristic temperature for the parasite (120). Therefore, we can assume that global change (or temperature shifts) is expected to have complex effects on infectious diseases, causing some to increase, others to decrease, and many to shift their distributions (140). In the particular case of pen shell mortality, our model results suggest the proposed mechanism of lower disease expression at lower temperatures. This might have direct consequences for the development of the mortality event and offers a bleak perspective for the future and specifically in the eastern Mediterranean basin, where the mortality was observed later due to current patterns but average temperatures tend to be higher than in the western part of the Mediterranean.

4.5 Conclusions

In this work we have analysed a compartmental model to study marine epizootics for sessile hosts assuming infection by direct transmission through waterborne parasites. Moreover, we have used data from the recent mass mortality event of *Pinna Nobilis* in the Mediterranean

⁷The N correction corresponding to standard incidence has already been applied to these values.

Sea as a case study to validate our model. Compartmental models are routinely used in the study of disease infection and propagation in terrestrial ecosystems, including the study of the current Covid-19 pandemic (see, e.g., (37)). However, these models are starting to be used only recently in the study of marine epizootics (18), while proliferation models have been the most popular in the field (133). A reason for the low popularity of compartment models in the study of marine epizootics is that there are some aspects in its modelling that differ from the now standard application to terrestrial ecosystems (118). An important difference is that, in principle, (micro)parasites need to be modelled explicitly in marine ecosystems, while often they are not included in the description in terrestrial ecosystems (117).

The SIRP model has 4 compartments and depends on 4 parameters, so that it is not quite amenable to theoretical analysis. At the same time, due to the large number of parameters of the model, using it to analyse experimental observations can be cumbersome in practice if the parameter values are unknown. Nevertheless, we have shown three reductions of the model, one exact and two approximate ones, that can be useful to overcome these limitations that are typically present at the first stages of emergent epidemics. Indeed, the timescale approximation is able to fit the collected data of our case study for some optimal parameters, as shown in Section 4.4. This approximation is particularly useful as it only depends on 2 parameters, the death rate of infected hosts, γ and an effective infection rate, β' . Although this approximation simplifies the fitting procedure, there is a price to be paid in this analysis. The infection parameter, β , and the parameters regulating proliferation, λ , and deactivation/dilution of the parasite, μ , become entrained into a single effective parameter, β' . Thus, the full understanding of the different effects at play in the system requires further work. Furthermore, we have shown that an epidemic model for immobile hosts can be reduced to the standard SIR model, which assumes direct contact among the hosts, i.e. that the hosts are mobile. This reduction is only valid when the time scale of the parasites is much faster than that of the hosts, i.e. $\mu \gg \beta N, \gamma$. Thus, our work provides a ground to apply the SIR model in marine epidemics of sessile hosts that fulfil the required conditions.

In a world with many possible new epizootics, we believe that our reduced model can be specifically useful to understand key features of those emerging diseases characterised by the spreading of waterborne parasites in a relatively fast way, provided that the temporal evolution of the disease can be determined for, at least, some set of individuals. Thus, some of the key parameters can be fitted to the available experimental data as shown in Section 4.4. Still, the fitted relevant parameters may need to be supplemented with further information or targeted experiments. We hope that this approach can be useful in understanding emerging diseases in shellfish species of economic not only ecological value, and also, with suitable modifications, in aquaculture. It is noteworthy that our case study is a haplosporidian waterborne parasite. In fact, waterborne haplosporidans have been responsible for some of the most significant and consequential marine disease epizootics on record and are considered the major pathogens of concern for aquatic animal health shellfish industries around the world (8). The SIRP model is the simplest model that one could think of having in mind its practical application, but could be extended to incorporate further effects that are so far described in an effective way.

4.6 Appendix

4.6.1 Finding a conserved quantity for the SIRP model

Starting with the SIRP model,

$$\begin{aligned}\dot{S} &= -\bar{\beta}PS \\ \dot{I} &= \bar{\beta}PS - \gamma I \\ \dot{R} &= \gamma I \\ \dot{P} &= \lambda I - \bar{\beta}PS - \mu P,\end{aligned}\tag{4.10}$$

from the \dot{S} equation, P can be written as follows,

$$P = -\frac{1}{\bar{\beta}} \frac{\dot{S}}{S},\tag{4.11}$$

and summing up the equations for \dot{S} and \dot{I} the following relation for I is obtained

$$I = -(\dot{S} + \dot{I})/\gamma. \quad (4.12)$$

Replacing Eq. (4.11), Eq. (4.12) and the differential equation for \dot{S} in the 4th differential equation in Eq. (4.10) one obtains,

$$\dot{P} = -\frac{\lambda}{\gamma}(\dot{S} + \dot{I}) + \dot{S} + \frac{\mu}{\bar{\beta}} \cdot \frac{\dot{S}}{S} \quad (4.13)$$

As $\dot{S}/S = d(\ln S)/dt$, all terms in the previous equation are exact differentials with respect to time, and the equation can be integrated yielding,

$$P + \frac{\lambda}{\gamma}(S + I) - S - \frac{\mu}{\bar{\beta}} \ln S = C \quad (4.14)$$

with the integration constant C , that is a conserved quantity, i.e., it takes the same value at one time of the dynamical evolution of the system. C is related to the initial conditions by,

$$C = P(0) + \frac{\lambda}{\gamma}(S(0) + I(0)) - \frac{\mu}{\bar{\beta}} \ln S(0) - S(0) = P(0) + \frac{\lambda}{\gamma}(N - R(0)) - \frac{\mu}{\bar{\beta}} \ln S(0) - S(0) \quad (4.15)$$

It is possible to use Eq. (5.13)-Eq. (4.15) to express one of variables as a function of the others, for example the parasite concentration P as,

$$P(S, I) = P(0) - \frac{\lambda}{\gamma}(S + I - N + R(0)) + \frac{\mu}{\bar{\beta}} \ln \frac{S}{S(0)} + S - S(0), \quad (4.16)$$

or equivalently as,

$$P(S, R) = P(0) + \frac{\lambda}{\gamma} \left[R - R(0) + \frac{\mu\gamma}{\bar{\beta}\lambda} \ln \frac{S}{S(0)} \right] + S - S(0) \quad (4.17)$$

From Eq. (5.13), it is easy to show that the SIP model of Ref. (18), that differs from the SIRP model in that the fourth equation is simplified to $\dot{P} = \lambda I - \mu P$, has as exact conserved quantity,

$$P + \frac{\lambda}{\gamma}(S + I) - \frac{\mu}{\bar{\beta}} \ln S = \mathcal{C} \quad (4.18)$$

as the extra term in the SIRP model $-\bar{\beta}SP$ is equal to \dot{S} from the first equation Eq. (4.10).

The SIR model has a conserved quantity (126), that in the case of Eq. (4.9) takes the form,

$$I + S - \frac{\gamma}{\beta'} \ln S = C. \quad (4.19)$$

Rewriting Eq. (5.13) in the alternative form,

$$\frac{\gamma}{\lambda}P + \left(1 - \frac{\gamma}{\lambda}\right)S + I - \frac{\mu\gamma}{\lambda\bar{\beta}} \ln S = C' \quad (4.20)$$

it can be seen that if $\lambda \gg \gamma$ Eq. (4.20) reduces to Eq. (4.19), remembering that in Eq. (4.9) $\beta' = \lambda\bar{\beta}/\mu$. The assumptions used to arrive to Eq. (4.19) in Section 4.2.3.3 where $\mu \gg (\gamma, \bar{\beta})$, and taking into account the expression for R_0 Eq. (4.3), that $\lambda \gtrsim \mu$ is most plausible to keep R_0 above the epidemic threshold ($R_0 > 1$).

4.6.2 Stability analysis of the fixed points of the SIRP model

Here we will assume the initial fixed point of our SIRP model, with $I(0) = P(0) = 0$ right before the introduction of the infection, either through I or P . We will assume that $R(0) = 0$, so that $S(0) = N$. To study the linear stability of the model we need to write the Jacobian, that takes the form,

$$J = \begin{pmatrix} -\bar{\beta}P & 0 & 0 & \bar{\beta}S \\ \bar{\beta}P & -\gamma & 0 & \bar{\beta}S \\ 0 & \gamma & 0 & 0 \\ -\bar{\beta}P & \lambda & 0 & (\bar{\beta}S - \mu) \end{pmatrix} \quad (4.21)$$

and obtain the eigenvalues for both fixed points, where we have already used the standard incidence, $\bar{\beta} = \beta/N$, from the evidence of the validation with experiments. For the pre-epidemic fixed point, the Jacobian becomes,

$$\begin{pmatrix} 0 & 0 & 0 & \bar{\beta}S(0) \\ 0 & -\gamma & 0 & \bar{\beta}S(0) \\ 0 & \gamma & 0 & 0 \\ 0 & \lambda & 0 & (\bar{\beta}S(0) - \mu) \end{pmatrix} \quad (4.22)$$

Matrix Eq. (4.22) has two null (0) eigenvalues and a pair of eigenvalues given by,

$$\Lambda_{1,2} = -\frac{1}{2} \left(\gamma + \mu + \bar{\beta}S(0) \pm \sqrt{\gamma^2 + \mu^2 + (\bar{\beta}S(0))^2 + 2\mu\bar{\beta}S(0) - 2\gamma\mu - 2\gamma\bar{\beta}S(0) + 4\lambda\bar{\beta}S(0)} \right) \quad (4.23)$$

from which one can determine that the fixed point is unstable whenever

$$\lambda\bar{\beta}S(0) > \gamma(\mu + \bar{\beta}S(0)) \quad (4.24)$$

and stable if the inequality is reversed. It can be easily shown that Eq. (4.24) is equivalent to $R_0 > 1$, with R_0 given by Eq. (4.3).

The final point of the epidemic, $S(\infty)$, can be found by solving the transcendental equation,

$$\left(\frac{\lambda}{\gamma} - 1 \right) S(\infty) - \frac{\mu}{\bar{\beta}} \ln(S(\infty)) = C \quad (4.25)$$

where C is determined from the initial conditions (Eq. (4.15)) and $I(\infty) = P(\infty) = 0$. (Eq. (4.25)) has two roots, where $S(\infty)$ represents the smallest one.

4.6.3 Calculation of R_0 using the Next Generation Matrix method

The so called Next Generation Method (NGM) is a method to obtain R_0 , the basic epidemiological quantity that measures the number of secondary cases produced by a typical infected individual during its entire period of infectiousness in a completely susceptible population. It was discussed in 4.6.2 that R_0 is related to the largest non-zero eigenvalue, say Λ , of the fixed point corresponding to the infection-free equilibrium. An outbreak occurs when $\Lambda > 0$ (or equivalently when $R_0 > 1$) and the NGM is an ingenious method to obtain directly R_0 in a reduced linear system. In more concrete terms, within the NGM method R_0 is the dominant eigenvalue of a suitably defined linear operator (a linear matrix in a suitable basis). This operator is obtained from a decomposition of the Jacobian, J of the infected/infecting compartments (i.e. excluding susceptible and removed compartments) in the form $J = T + \Sigma$, where T is the *transmission part*, that describes the production of new infections, and Σ the *transition part*, that describes changes of state (including death). Then, it can be proved (57) that the *basic reproduction number* R_0 is given by the spectral radius (i.e. the largest eigenvalue) of the (next generation) matrix $K = -T\Sigma^{-1}$.

In the case of the SIRP model the decomposition is applied to the 2×2 Jacobian corresponding to the dynamical evolution of the (I, P) infectious compartments, being the decomposition,

$$J = \begin{pmatrix} -\gamma & \bar{\beta}S_0 \\ \lambda & -(\bar{\beta}S_0 + \mu) \end{pmatrix} \quad T = \begin{pmatrix} 0 & \bar{\beta}S_0 \\ 0 & 0 \end{pmatrix} \quad \Sigma = \begin{pmatrix} -\gamma & 0 \\ \lambda & -(\bar{\beta}S_0 + \mu) \end{pmatrix}$$

where the $\bar{\beta}PS$ term in J is the only one that contributes to the transmission matrix, as it is the only process involving infection, while all the other terms in the dynamical equations of I and P imply transitions (to another compartment, like $I \rightarrow R$ or birth and death of P).

Then, the next generation matrix is given by,

$$K = -T\Sigma^{-1} = \begin{pmatrix} \frac{\lambda\beta S_0}{\gamma(\beta S_0 + \mu)} & \frac{\beta S_0}{\beta S_0 + \mu} \\ 0 & 0 \end{pmatrix} \implies R_0 = \frac{\lambda\beta S_0}{\gamma(\beta S_0 + \mu)},$$

This result coincides with the expectation that R_0 should correspond to the number of hosts infected in a single generation by the appearance of an infected host in a completely susceptible population. This can be obtained from the number of parasites produced by an infected

host, λ , times the time in which the infected host is alive producing parasites, $1/\gamma$, multiplied by the number of infected hosts produced per parasite, βS_0 , times the time the parasite is alive available to infect, $1/(\mu + \beta S_0)$, taking into account that parasites are inactivated at a rate μ and also die when infecting at a rate βS_0 , where this result assumes that the susceptible population does not change from its initial value S_0 .

4.6.4 Sensitivity Analysis

One particular way to analyse the local sensitivity (LSA) of a given model function, $F(\vec{p})$, for each of the parameters that conform it, p_i , is through the normalised sensitivity indexes (35),

$$\Omega_{p_i}^F = \frac{\partial F}{\partial p_i} \frac{p_i}{F} \Big|_{p_i=p^0}. \quad (4.26)$$

where the partial derivatives in Eq. (4.26) are determined analytically in our case.

GSA works by studying the influence of a large domain of parameter space in the final state of the epidemic and in the epidemic peak. In our case this will be achieved by means of a variance based analysis, known as Sobol method (154). This particular method provides information no only on how a particular parameter alone influences the model outputs (as happens with LSA), but also on the influence of its interactions with other parameters. This information is organised in what are known as Sobol indices, that have been implemented within the Julia high-level programming language (15) using the DifferentialEquations.jl package (136), and in particular through its subpackage DiffEqSensitivity.jl. This implementation allows the user to sample the parameter space using QuasiMonteCarlo methods and thus obtain confidence intervals (CI) for the sensitivity indices, which are directly related to the committed statistical error.

The total order indices are a measure of the total variance of the output quantity caused by variations of the input parameter and its interactions. First order (or “main effect”) indices are a measure of the contribution to the output variance given by the variation of the parameter alone, but averaged over variations in other input parameters. Second order indices take into account first order interactions between parameters. Further indices can be obtained, describing the influence of higher-order interactions between parameters, but these are not going to be considered. More detailed information about sensitivity analysis can be found in (144).

4.6.5 General rate change with temperature

In (73) the metabolic rate of a wide variety of organisms was studied, showing that the change in the metabolic rate with temperature was similar among them. In particular, the natural logarithm of the metabolic rate linearly depends on the inverse of absolute temperature,

$$\log(R(T)) = a \cdot \left(\frac{100}{T} \right) + b \quad (4.27)$$

and for all the analysed organisms they found that a lies between -5 and -10 and b between 14 and 30 . From their analysis, we can compute the change in the rate for a given increase of temperature,

$$\frac{R(T + \Delta T)}{R(T)} = \exp\left(a \cdot \left(\frac{100}{T + \Delta T} \right) + b\right) / \exp\left(a \cdot \left(\frac{100}{T} \right) + b\right) = \exp\left(a \cdot \frac{-1000}{T + \Delta T} \cdot \frac{\Delta T}{T}\right). \quad (4.28)$$

Substituting $T = 287K$ and $\Delta T = 3K$, that correspond to our available data (cf. Section 4.4) in Eq. (4.28), using both the upper and lower limit of a , we obtain that the expected increase in the effective transmission rate is between 1.2 to 1.4 . This is far from the 19-fold increase that we obtained with the mass action hypothesis in Section 4.4 while it is in good agreement with either the 1.92 ratio we obtained for $\bar{\beta}$ with the reduction of Section 4.2.3.2 or the 1.43 ratio obtained with the fast-slow approximation of Section 4.2.3.3, both obtained using the standard incidence choice.

Fig. 4.9(a) shows the change in the rate with an increase of 3°C for different base temperatures and for all the organisms analysed in (73), and using their fit. Note that for all temperatures

between 0°C and 30°C the rate change lies between 1.2 and 1.45. Fig. 4.9(b) shows the change in the rate for different temperature increases, with a base temperature of $T = 287\text{K}$. Note that in order to obtain a 19-fold increase the temperature change should be at least of 30°C ⁸. The temperature dependence of metabolic rates has been reported in the context of epidemic parameters (44, 153)

The behavior of the metabolic rates re-analysed here has been also found experimentally in epidemic contexts such as (44, 153), i.e. the increase of the rates with temperature fulfill the ranges shown here.

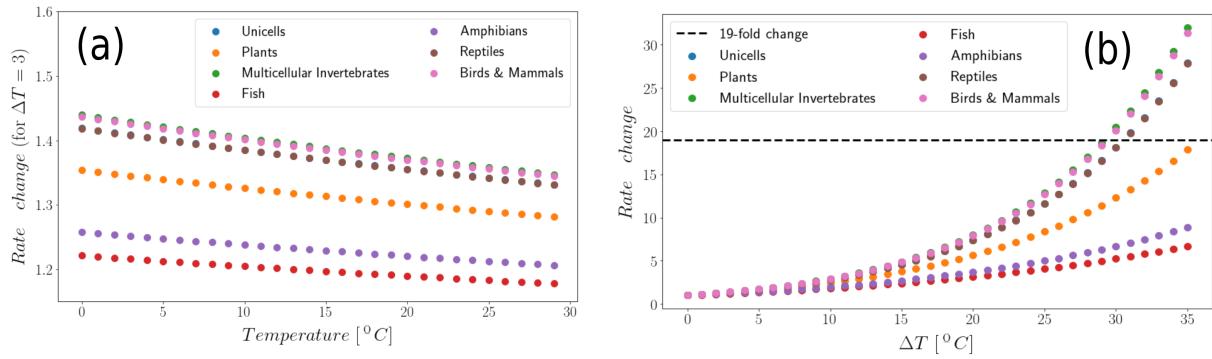
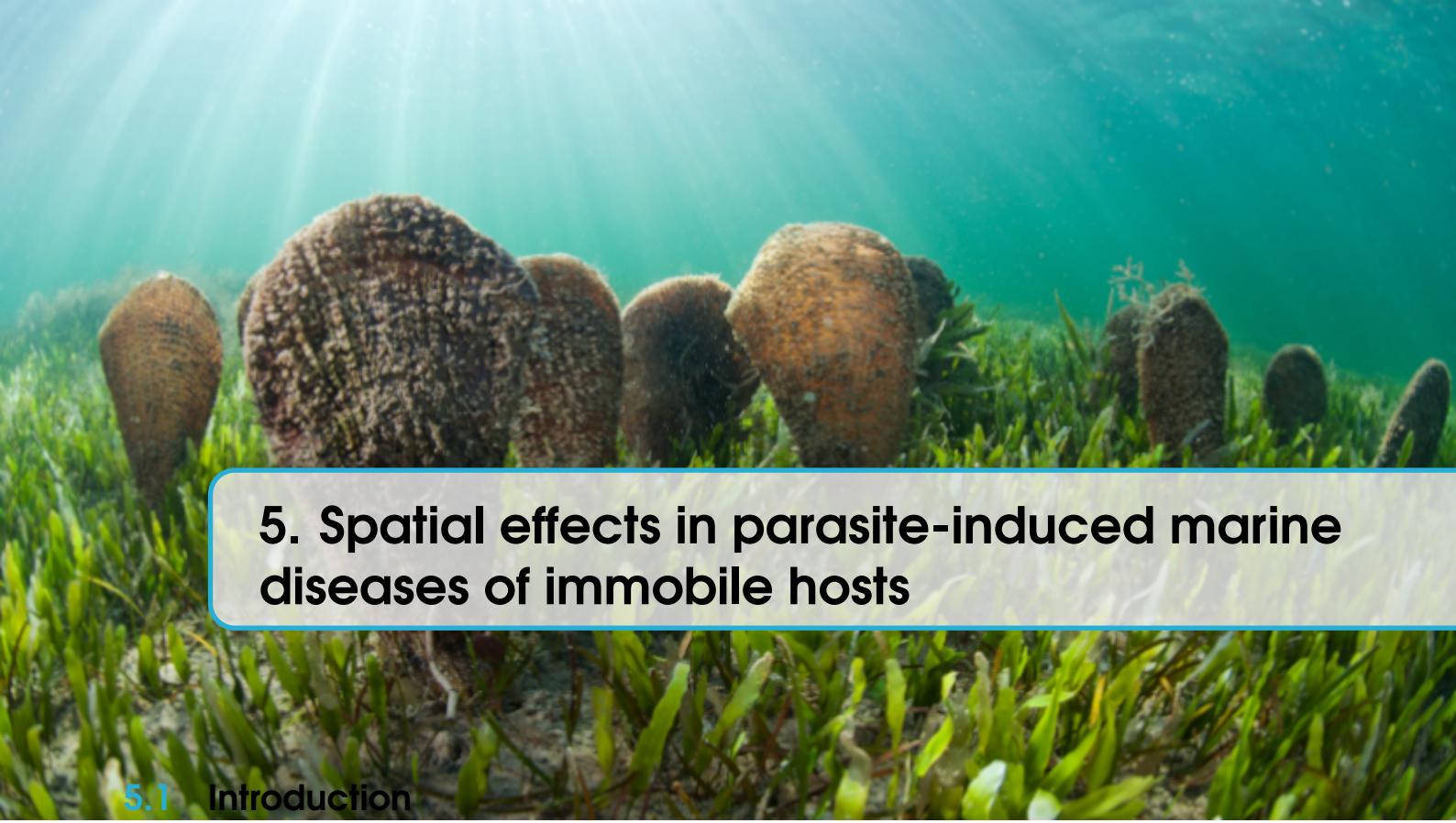


Figure 4.9: Graphical representation of change in the rate (in ordinates) for different reference temperatures (in abscissae) for: (a) a temperature increase of 3°C ; (b) a temperature increase of 14°C . The black dotted line in (b) corresponds to a 19-fold increase in the rate.

⁸A temperature change of 30°C could fall outside the range in which the study of [73] is valid. We just stress that a 19-fold rate change is unlikely for the case of a 3°C that correspond to the 2 data sets that we compare in this section.



5. Spatial effects in parasite-induced marine diseases of immobile hosts

5.1 Introduction

Wildlife emergent infectious diseases represent a substantial threat to ecosystems and the conservation of their biodiversity (55). Their effects can be devastating at the ecological level, causing local extinctions (55) and in some cases pushing endemic species to the verge of extinction, as is the case of *Pinna nobilis* (32); at the economic level, producing losses in agriculture, livestock and aquaculture (129, 160, 168), and impact human health, as is the case of the COVID-19 pandemic (143). For the past decades, parasites have been continuously emerging (53, 123), while globalisation and climate change have contributed to their evolution. This has allowed these parasites to enter in new ecological niches and spread further the diseases they produce (2). In particular, marine infectious diseases are recently increasing due to these and other anthropogenic pressures, like pollution and overfishing (104), inducing widespread mass mortalities in several species (60, 96, 166).

An important subset of marine organisms affected by infectious diseases are sessile (i.e. they cannot move), like bivalves, sponges or corals. An increasing number of outbreaks affecting marine mollusks have been reported, some of them causing mass mortalities in commercially important bivalves (83). Mainly due to the economic importance of some species (e.g. oysters), infectious diseases in bivalve populations have been deeply studied (119, 130, 131, 132). Recently, deterministic compartmental models have been used to describe parasite transmitted diseases in marine sessile bivalves (16, 17, 79), showing to be able to accurately predict disease transmission in some circumstances. The main limitation of these compartmental models is the assumption of a non-spatial description of the system under study. This underlying hypothesis assumes that any pair parasite-host of the system can interact at any time, which is unrealistic in general. A non-spatial description assumes well mixed populations, which implies that the mean distance among hosts is smaller than the typical distance explored by parasites in their lifetime. This assumption can be quite realistic in some situations, as it is in (79) where the hosts were kept in tanks with water renovation. However, a non-spatial model is not expected to yield a good description of spatially extended hosts in a natural setting.

The key quantity in mathematical epidemiology is the basic reproductive number, R_0 , that represents the number of infected individuals generated in one generation by the appearance of a single infected individual in a fully susceptible population. Thus, $R_0 > 1$ ensures the onset of an epidemic, as the number of infected individuals will grow exponentially producing a widespread disease (5). If we first disregard spatial effects and assume a non-spatial description, R_0 can be obtained from standard methods, like the Next Generation Matrix method (57), and will only depend on *intrinsic* characteristics of the pathosystem (host-pathogen system) under study. However, this basic reproduction number is unable to characterise the threshold behaviour in many situations, including spatially extended systems (50, 111, 138). In these systems, the propagation of an epidemic to the entire system needs that a certain spatial threshold

is exceeded (72). Otherwise the disease will only take place in suitable localised parts of the system, not being able to propagate to the total system. Thus, disease spread will be strongly affected by the host spatial distribution and pathogen mobility, which are not accounted for in non-spatial models.

In this work we will try to unravel the transmission mechanisms of a parasite-induced disease affecting immobile hosts in a spatially extended system. We will approach the problem both theoretically and through numerical simulation. The numerical study is based on Individual-Based Modelling (IBM), a method widely used to study ecological systems (82), so that individuals are treated as discrete entities, space is introduced explicitly and the dynamics are stochastic. Representative average behaviours can be obtained by averaging over a sufficient number of realisations, and the accuracy of the approach can be calibrated by deriving the corresponding non-spatial limit, that can be confronted with the suitable compartmental model on which a particular IBM is based. The IBM approach to our problem will allow to study in depth the relation between pathogen mobility and immobile host infection. As parasites move randomly over the space, tracking the position of each parasite at different times turns to be of fundamental importance to properly capture the stochastic dynamics of infections from parasites to hosts. Modelling parasites and hosts as individual entities allows to take into account the spatial and temporal heterogeneity of interactions between them. This heterogeneity and the level of control in microscopic interactions cannot be captured by other mathematical approaches such as partial differential equations. On the other hand, IBMs are mathematically involved, and analytical treatments are normally cumbersome, while their numerical implementation is computationally expensive (26).

Here we introduce a spatially-explicit individual-based model to study parasite-induced marine diseases of immobile hosts. The model is applied to the case of diffusing parasites and uniformly distributed hosts. The system under study is an extension of the compartmental model presented in (79). As a main result, we find that the occurrence of an outbreak will depend on the balance between the intrinsic characteristics of the pathosystem, well represented by the above described non-spatial basic reproductive number, R_0 , and features that characterise parasite mobility. We generalise the basic reproductive number, that we will refer to as \tilde{R}_0 , such that it accounts for the number of hosts that get infected by the appearance of a single infected individual in a fully susceptible population in a spatially extended system. \tilde{R}_0 characterises the global epidemic and can be written as a product between R_0 and a factor describing parasite mobility. The latter factor is smaller and at most equal to 1, which implies that, as it could be expected, it is more difficult to induce a global outbreak in a spatially extended system (a two dimensional lattice in our case) than in a well mixed (non-spatial) population.

The paper is organised as follows: in Section 6.2, we introduce some biological considerations for bivalve epidemics, discussed in more detailed in Ref. (79), and build the spatially-explicit model. In Section 6.3, we present analytical results that are discussed and supported by numerical simulations. Specifically, the high mobility limit is discussed and connected to the the compartmental model. An approximation for the parasite population is discussed. Then, the effect of parasite mobility to the epidemic threshold is characterised, deriving an analytical expression for the basic reproduction number. Furthermore, the spreading speed of the disease and the time-scale to extinction is investigated. Finally, Section 5.4 contains some concluding remarks.

5.2 The SIRP spatial model

The most important biological features of the system under study are as follows. First, hosts are immobile, while the disease is transmitted by parasites produced by infected hosts. There are two mechanisms by which parasites are cleared from the medium: i) they have a finite life time after which they die; ii) they get absorbed after they infect a host and thus are no longer in the medium and cannot infect other hosts. Recruiting (birth) of hosts occur at a very slow rate compared to other timescales in the system, and accordingly it will be considered negligible in the model. Moreover, hosts do not show long-term immunity, as is typical of invertebrates, like mollusks (133). We also assume that recovery (healing) of infected hosts, if it occurs, can be neglected. Furthermore, we consider that dead hosts are not a source of parasites in the medium. See (79) for a detailed presentation of the non-spatial SIRP model, including these

biological modelling considerations.

Under these considerations, we introduce an individual-based model with explicit space characterisation to study the effect of parasite mobility in disease transmission. We consider a square grid of length L with periodic boundary conditions and place a single host per site, so that there are $N = L^2$ hosts. The hosts can be in three discrete states: susceptible, S ; infected, I and dead (or removed), R . Then, we introduce the parasite population as a new individual with a single state, P . Hosts are sessile (i.e. immobile), while parasites are allowed to move between the lattice sites. As initial condition we assume that the entire host population is susceptible, $S(0) = N = L^2$, and that a small initial number of parasites, $P(0)$ is introduced in the system.

Infection occurs when susceptible hosts filter parasites in their close proximity. Accordingly, the infection process is implemented between parasites and susceptible hosts sharing the same lattice site. In particular, susceptible hosts in contact with a parasite become infected at rate β . As the infection event implies the filtering of a parasite by a susceptible host, when a new infection occurs a parasite of that particular site is removed. Infected individuals die at rate γ and produce parasites at rate λ , while parasites die at rate μ . Parasites move randomly between the four neighbouring lattice sites at rate κ , which corresponds to a diffusive motion. Table 5.1 contains the definitions of the variables and parameters of the model , Fig. 5.1 shows a schematic representation of the dynamics and Eq. (5.1) summarises the reactive events



Formally, the model is mathematically described by a system of N master equations for the probabilities of the states in each lattice site i . This is very difficult to manage analytically, so the time evolution of the model is numerically solved using Gillespie's algorithm (71) (the code can be found in (74)).

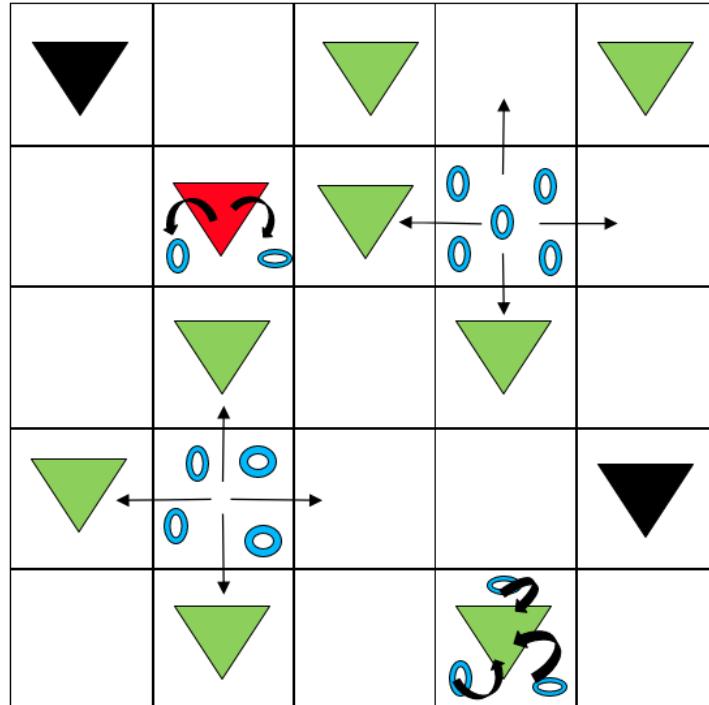


Figure 5.1: Scheme of the individual based model. Green, red and black triangles represent susceptible, infected and dead hosts, respectively. Blue rings represent parasites, which move randomly between cells. Susceptible hosts get infected by filtering parasites while infected hosts produce them. Dead hosts do not participate in the dynamics of the system.

Table 5.1: Variables and parameters of the model

Variable/Parameter	Definition
S	Susceptible host
I	Infected host
R	Dead host
P	Parasite
β	Parasite-host transmission rate
γ	Host mortality rate
λ	Production rate of parasites by infected hosts
μ	Parasite natural death rate
κ	Parasite dispersal rate (mobility)
R_0	Non-spatial basic reproductive number
\tilde{R}_0	Spatial basic reproductive number

5.3 Results

In this section several features of the model are studied, both numerically (from IBM simulations) and analytically. All numerical results were obtained for a square lattice of length $L = 100$, with $N = S(0) = 10^4$ hosts and using a small initial condition of $P(0) = 50$ parasites in the centre site.

5.3.1 Non-spatial limit

An important test of the IBM implementation is to show that, under suitable circumstances, it converges to the non-spatial model on which the IBM is based. This occurs in the limit when the parasites move many times before dying or infecting a susceptible host. In this situation, each parasite visits typically all the hosts of the system and may infect any of them. This is equivalent to infecting a random host of the system, which happens with probability $\beta S/N$, being S the total number of susceptible hosts in the system. An equivalent picture is that parasites will end up uniformly distributed in the lattice, so that there will be P/N parasites in each lattice site at any time. One expects to reach these conditions when $\kappa \gg \mu, \beta$, and, thus the system as a whole can be described by the following system of ordinary differential equations (ODE's),

$$\begin{aligned}\dot{S} &= -\beta PS/N, \\ \dot{I} &= \beta PS/N - \gamma I, \\ \dot{R} &= \gamma I, \\ \dot{P} &= \lambda I - \beta PS/N - \mu P,\end{aligned}\tag{5.2}$$

that is precisely the SIRP non-spatial model (79), where S, I, R are the total number of susceptible, infected and recovered hosts in the system, P the total number of parasites and N is the number of hosts.

The basic reproduction number, R_0 , of this non-spatial model is the dimensionless quantity that yields the number of secondary infections generated by the appearance of a single infected individual in a completely susceptible population, also indicating whether the system will exhibit an epidemic outbreak, $R_0 > 1$, or not, $R_0 < 1$. In our case it can be directly computed as the mean number of parasites produced by an infected host during its mean lifetime, λ/γ , times the mean number of susceptible hosts that get infected by parasites during their mean lifetime, $\beta/(\mu + \beta)$.

$$R_0 = \frac{\lambda}{\gamma} \frac{\beta}{\mu + \beta}.\tag{5.3}$$

This result can be corroborated with standard methods such as the Next Generation Matrix method (57) (see (79)), where $S(0) = N$ has been considered.

Moreover, the model has a conserved quantity \mathcal{C} (79) that allows to find an analytical expression for the final number of dead individuals (cf. Section 5.5.1),

$$R(\infty) = N + \frac{S(0)}{\xi} W_0 \left(-\xi \exp \left(-\frac{\beta}{\mu} C \right) \right), \quad (5.4)$$

with $\xi = S(0) \frac{\beta(\lambda - \gamma)}{\mu\gamma}$ and $C = P_0 + \frac{\lambda}{\gamma} (S(0) + I(0)) - S(0)$.

The non-spatial limit of the model has been evaluated by comparing realisations of the stochastic model (in the limit $\kappa \gg \mu, \beta$) with numerical solutions of the non-spatial ODE system of Eq. (5.2). Furthermore, the analytical expression for $R(\infty)$ using the non-spatial model, Eq. (5.4), is also compared to the numerical results of the individual based model. As shown in Fig. 5.2 (a)-(c), as κ is increased compared to μ the individual based model approaches the non-spatial one. Fig. 5.2(d) shows how the numerical results for $R(\infty)$ for different R_0 values approach the analytical solution in the non-spatial limit.

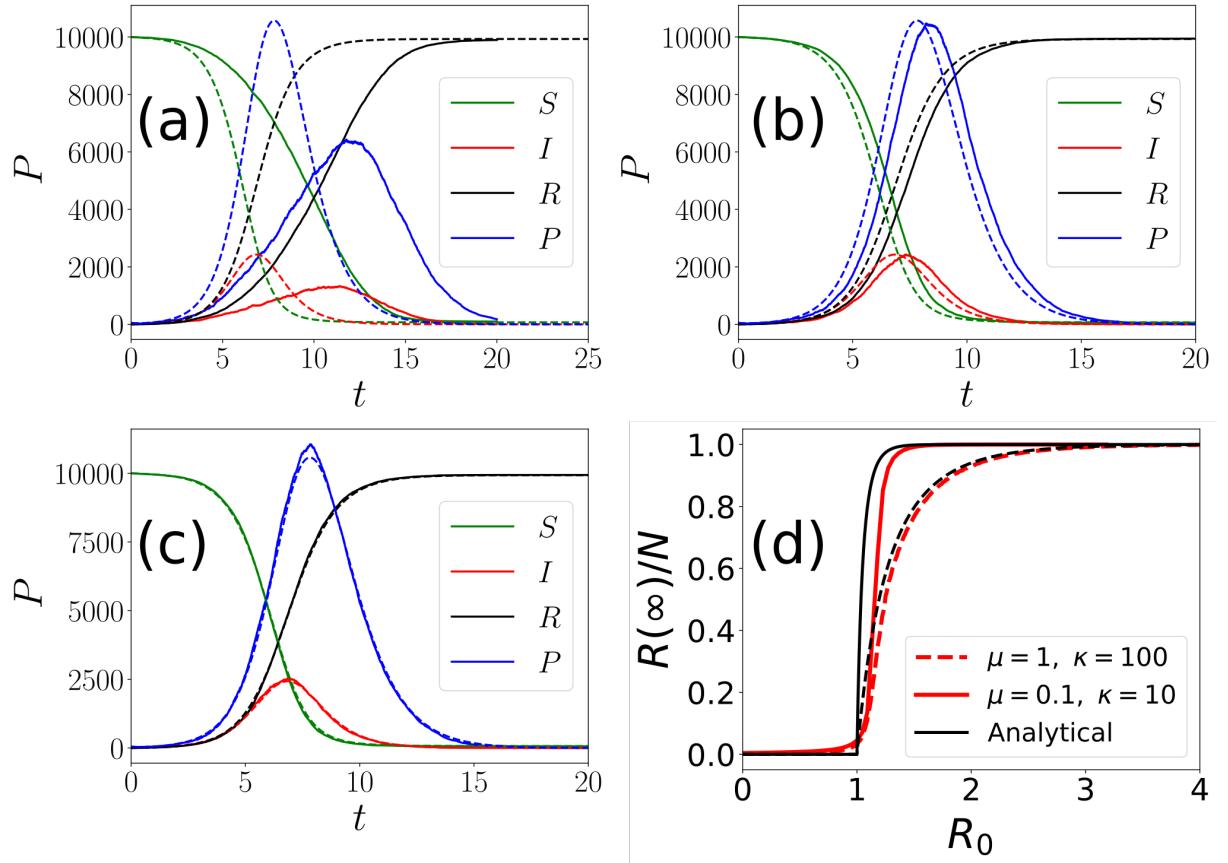


Figure 5.2: Numerical solution of the non-spatial model (Eq. (5.2), dashed lines) compared with numerical solutions of the individual based model (solid lines) approaching the non-spatial limit with fixed $\gamma = \mu = \beta = 1$ and $\lambda = 6$. (a) $\kappa = 10^2$, (b) $\kappa = 10^3$, (c) $\kappa = 10^4$. Panel (d) shows the final fraction of dead hosts, $R(\infty)/N$, as function of R_0 for $\kappa/(\mu + \kappa) = 0.999$ with $\mu = 1, 0.1$ compared to the analytical result.

5.3.2 Approximate relation between parasites and infected hosts

In the limit $\kappa \gg \beta, \mu$ a time-scale approximation can be performed so that the parasite population dynamics directly relates to that of the infected hosts. In the non-spatial limit it was already shown in (79) that, if $\mu \gg \beta, \gamma$ and $\lambda \gg \beta P/N$, the total parasite population of the system can be well described using the approximation (see (79) for a detailed discussion),

$$P(t) \approx \frac{\lambda}{\mu} I(t), \quad (5.5)$$

Here we extend the validity of this approximation to spatial systems far from the non-spatial limit. Consider the local dynamics of the parasite population on a lattice site i . Note that when the host in the site is susceptible, parasites in this site can either infect the host, die, or move to another site. All these processes imply that a parasite will disappear from the current site. Once the host at site i gets infected, infection can no longer occur whereas parasite production is now possible. If κ is small enough compared to λ and μ , the only competing processes in sites with infected individuals will be the production of parasites and their natural death, which can be fairly described by the following rate equation,

$$\frac{dP}{dt} = \lambda - \mu P , \quad (5.6)$$

whose solution is

$$P(t) = \frac{\lambda}{\mu} + \left[P(0) - \frac{\lambda}{\mu} \right] e^{-\mu t} . \quad (5.7)$$

From Eq. (5.7) one may notice that the stationary value of P , λ/μ , is reached in a time proportional to $t_{eq} \propto 1/\mu$. This derivation allows to find a condition for which ?? is valid beyond the non-spatial limit. Basically, if the mean dispersal time, $1/\kappa$, is greater than the equilibrium time, $t_{eq} \propto 1/\mu$, parasites in sites with infected hosts will reach its stationary level before parasites enter or leave the sites. Thus, sites with infected hosts can be considered as a closed system and the approximation holds. In other words, if the dispersal rate of parasites is small compared to the parasite deactivation rate, $\kappa \ll \mu$, the local parasite population of the site will reach its stationary level $P_i = \lambda/\mu$. It is possible to extend the result to the entire system: if there are $I(t)$ infected sites in the system at time t and $\kappa \ll \mu$ is fulfilled, there will be a total parasite population of $P(t) = (\lambda/\mu)I(t)$, which is equivalent to ??.

Thus, for the non-spatial limit ($\kappa \gg \mu$) we have that if $\mu \gg \beta, \gamma$?? is valid, while for $\kappa \ll \mu$ the approximation is also valid regardless of the value of β, γ , as the nature of the approximation is different. Thus, in general, as κ decreases over μ (the lower the parasite mobility becomes) we expect the approximation to work better.

The parasite approximation to infected hosts dynamics, ??, is numerically verified for different mobility conditions. -Fig. 5.3(a)-(b) shows how the approximation improves as μ grows over β, γ (mean errors are 0.18 and 0.0081, respectively) in the non-spatial limit, i.e. $\kappa \gg \mu$, as expected. This result is in perfect agreement with that found in (79). Then, Fig. 5.3(c)-(d) show that the approximation is valid in general when $\kappa \ll \mu$ but improves anyway when $\mu \gg \beta, \gamma$ (mean errors are 0.04 and 0.0026, respectively). Summarising, we see that the lower the value of κ is with respect to μ the more valid ?? is, regardless of the value of β, γ , while in the non-spatial limit, $\kappa \gg \mu$, the condition $\mu \gg \beta, \gamma$ is needed.

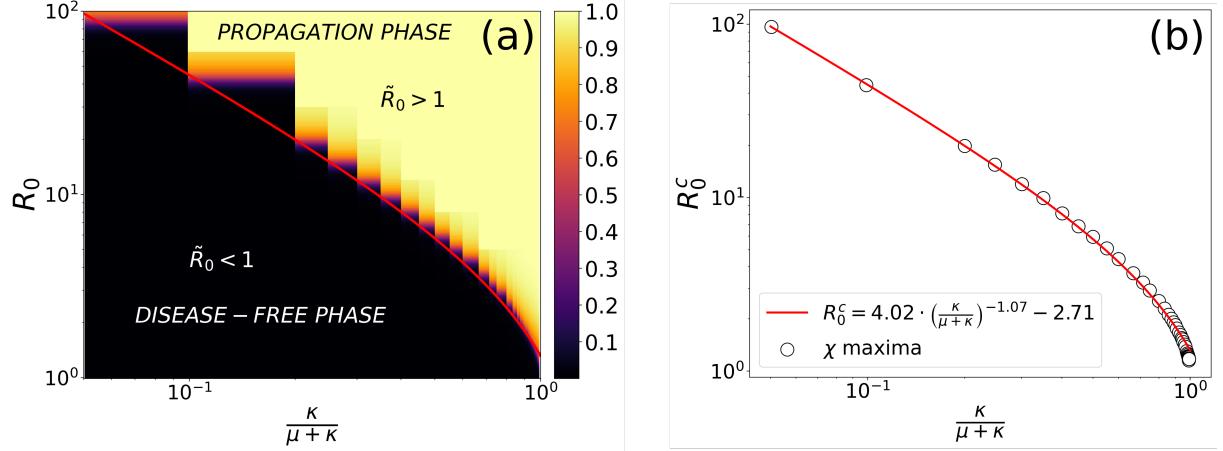


Figure 5.4: (a) Phase diagram showing the transition between the the disease-free phase and the propagation phase for several values of the parasite mobility and R_0 . The colour code represents the fraction of dead individuals (i.e. R/N) in the final state of the epidemic computed by the average over 1000 realisations. (b) Fit for the transition line following Eq. (5.10), where dots are the maximums of the “order parameter” fluctuations, $\chi = \langle R(\infty)^2 \rangle - \langle R(\infty) \rangle^2$

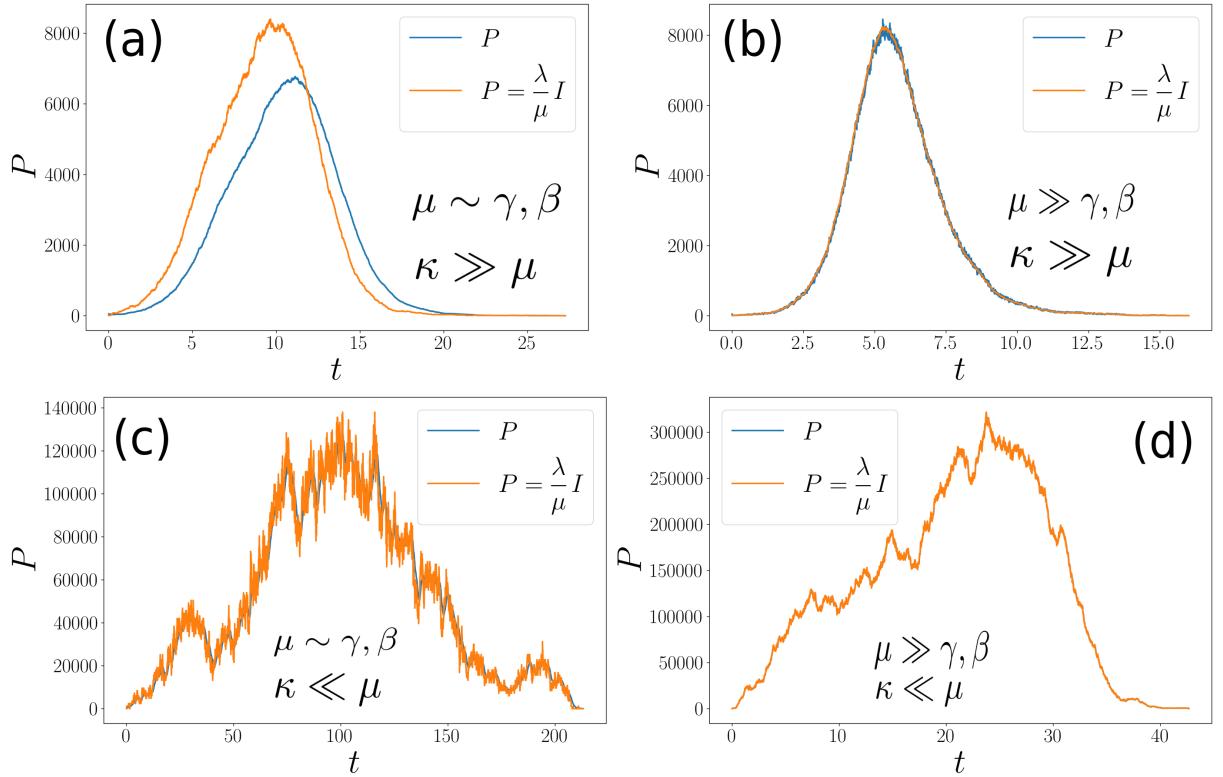


Figure 5.3: Numerical verification of the approximate expression for the parasite population dynamics, ??, for different mobility conditions. The simulations were performed fixing $\beta = \gamma = 1$ for all panels. (a) $\mu = 1$ $\kappa = 10^2$, $\lambda = 6.06$, $\kappa/(\mu + \kappa) = 0.99$; (b) $\mu = 100$, $\kappa = 10^4$, $\lambda = 306$, $\kappa/(\mu + \kappa) = 0.99$; (c) $\mu = 1$, $\kappa = 0.01$, $\lambda = 1200$, $\kappa/(\mu + \kappa) = 0.01$; (d) $\mu = 100$, $\kappa = 1$, $\lambda = 60600$, $\kappa/(\mu + \kappa) = 0.01$

5.3.3 Spatial threshold

One of the main questions in epidemiology is to define the conditions under which an epidemic outbreak occurs, which usually is translated into the existence of a threshold. In a well mixed (non-spatial) system the basic reproduction number (R_0), that characterises this threshold $R_0 = 1$,

can be defined exclusively from *intrinsic* parameters of the pathosystem, as the host-pathogen interaction does not depend on the host spatial structure or pathogen mobility (see Eq. (5.3)). In stochastic spatial models this formulation of R_0 breaks down. First of all, in stochastic models, even above the threshold there is a non-zero probability that the disease is unable to propagate initially, given by $P_{\text{outbreak}} = 1 - (1/R_0)^{I(0)}$ (24). Furthermore, the discrete nature of the populations also modifies the estimates of R_0 (100). On the other hand, the introduction of space changes completely the nature of epidemic outbreaks, modifying the host-pathogen interactions by means of specific host spatial distributions and pathogen mobility patterns. Even if the basic reproduction number of the non-spatial model is above the threshold ($R_0 > 1$), if parasite mobility is not large enough, the epidemic will stay locally confined. Thus, one expects that the threshold at which an epidemic outbreak can propagate to the rest of the system will depend on the balance between the intrinsic pathosystem parameters in R_0 and parasite mobility, defining a spatial basic reproduction number, \tilde{R}_0 .

Having in mind the study in Section 5.3.1, we expect that in the high mobility limit the basic reproduction number is defined by the non-spatial formula, Eq. (5.3). On the other hand, the lower the parasite mobility is, the more difficult will be for a local outbreak to propagate through the system. Thus, it is natural to think of an spatial basic reproduction number of the form $\tilde{R}_0 = R_0 f(\kappa)$, where $f(\kappa)$ is an increasing function of the parasite dispersal rate accounting for parasite mobility fulfilling the limit $\lim_{\kappa \rightarrow \infty} f(\kappa) = 1$.

Indeed, some authors recently showed that the spatial basic reproduction number can be defined as the product between the non-spatial value, R_0 , and a factor accounting for spatially-dependent interactions, $f(r)$, in the form $\tilde{R}_0 = R_0 f(r)$ (63, 157). However, these expressions are not analytical (63, 64) or are not directly related to pathogen mobility (157). Here we propose a simple expression for the spatial basic reproduction number regulating the spatial propagation of the epidemic,

$$\tilde{R}_0 = \frac{\lambda}{\gamma} \frac{\beta}{\mu + \beta} \frac{\kappa}{\mu + \kappa} = R_0 \frac{\kappa}{\mu + \kappa}. \quad (5.8)$$

The derivation of Eq. (5.8) accounts for the number of secondary parasites that are able to produce new infections, or equivalently, the number of secondary infections produced by an initial infected host. If we consider an initial infected individual, on average it will produce λ/γ parasites. Then, these parasites can only move to neighbouring sites or die, so that the dispersal probability is given by $\kappa/(\mu + \kappa)$. Finally, considering that parasites do not affect each other trying to infect the same host, the infection probability is given by $\beta/(\mu + \beta)$. Joining all terms, we finally obtain Eq. (5.8). This expression is valid when parasites move only to sites with susceptible individuals and do not try to infect the same host. Thus, the derived \tilde{R}_0 is only an approximation to the spatial basic reproduction number for the case of an initial introduction of a small quantity of parasites in a fully susceptible population.

Note that, as expected, the spatial basic reproduction number is nothing other than the basic reproduction number of the non-spatial model multiplied by an increasing function of the parasite mobility, $\kappa/(\mu + \kappa)$. Taking the limit $\kappa \gg \mu$ in Eq. (5.8) (non-spatial limit) the basic reproduction number of the non-spatial model is recovered. Conversely, in the limit of very low mobility, the $\kappa/(\mu + \kappa)$ factor is small, and this has to be compensated with a large value of the non-spatial basic reproduction number, R_0 , in order that there is an outbreak, i.e., $\tilde{R}_0 > 1$.

The spatial threshold, $\tilde{R}_0 = 1$, given by Eq. (5.8), has been numerically checked by computing the phase diagram between the absorbing phase $R(\infty) \approx 0$ (no infection, i.e. disease-free state) and the active phase $R(\infty) > 0$ (in which some level of infection has occurred, i.e. propagation phase) for several values of the parasite mobility and the basic reproduction number of the non-spatial model, R_0 Eq. (5.3). The transition is expected to occur at $\tilde{R}_0 = 1$, implying from Eq. (5.8) that the dependence of the critical value of R_0 , say R_0^c , is expected to take the form,

$$R_0^c \sim \left(\frac{\kappa}{\mu + \kappa} \right)^{-1}. \quad (5.9)$$

As discussed above, we expect that $\tilde{R}_0 = 1$ with \tilde{R}_0 given by Eq. (5.8) does not represent exactly the spatial threshold, and for this reason we suggest the more general functional form,

$$R_0^c \sim A \left(\frac{\kappa}{\mu + \kappa} \right)^{-B} - C, \quad (5.10)$$

to be fitted to numerical data, where $A = 1$, $B = 1$ and $C = 0$ would imply a perfect agreement of numerical simulations of the IBM model with Eq. (5.8).

In order to obtain the phase diagram, we compute the absorbing state of the model as an average over 1000 realisations for each value of the mobility and R_0 considered. Then, the critical value R_0^c is computed for each mobility value as the R_0 value for which the fluctuations of the “order parameter” $\chi = \langle R(\infty)^2 \rangle - \langle R(\infty) \rangle^2$ are maximal, as this would be an indication of a transition between the disease-free and the propagation phases.

Fig. 5.4(a) shows the numerical results of the computed transition between the disease-free and propagation phases. The heatmap coding represents the average value of absorbing state $\langle R(\infty) \rangle$ for several values of the mobility factor and R_0 . As expected, the lower the mobility factor is, the higher the value of R_0 is needed for the disease to invade the population. Fig. 5.4(b) shows the fit of Eq. (5.10) with less than a 1% of relative error. Interestingly, we obtain $B = 1.07 \approx 1$ which validates our expression for the spatial threshold as a first approximation. However, the values for $A = 4.02$ and $C = 2.7$ show a significant deviation from Eq. (5.9) and indicate that the expression Eq. (5.8) is an approximation to the spatial basic reproduction number, which however seems to contain the right dependence on $\kappa/(\mu + \kappa)$, and where A could be a geometric factor for a lattice.

5.3.4 Spreading speed of the infected population and time to extinction

Another relevant epidemiological question is how does an infected population spread after the onset of an epidemic. In order to obtain this spreading speed we computed the mean time needed for an infected individual to reach the boundary of the system. More specifically, for each particular choice of the model parameters, 1000 simulations were run for several system sizes ranging from $L = 10$ to $L = 60$. The computed mean time was found to depend linearly with the system size, thus allowing to compute the speed from the slope of this relation. With this procedure, the spreading speed was computed for several values of the parasite mobility and R_0 , large enough to ensure an epidemic outbreak that reached the boundary of the system. In this situation, the spreading speed is expected to depend linearly with the square root of the parasite mobility,

$$v \sim \sqrt{\kappa} . \quad (5.11)$$

Fig. 5.5(a) shows this square root dependence for different values of the fixed R_0 . Similarly, the speed was also computed for several values of the basic reproduction number and a fixed mobility. In this case, it varies with the square root of the distance to the critical value of R_0 , R_0^c , as shown by Fig. 5.5(b). This is in good agreement with other mathematically similar models (14).

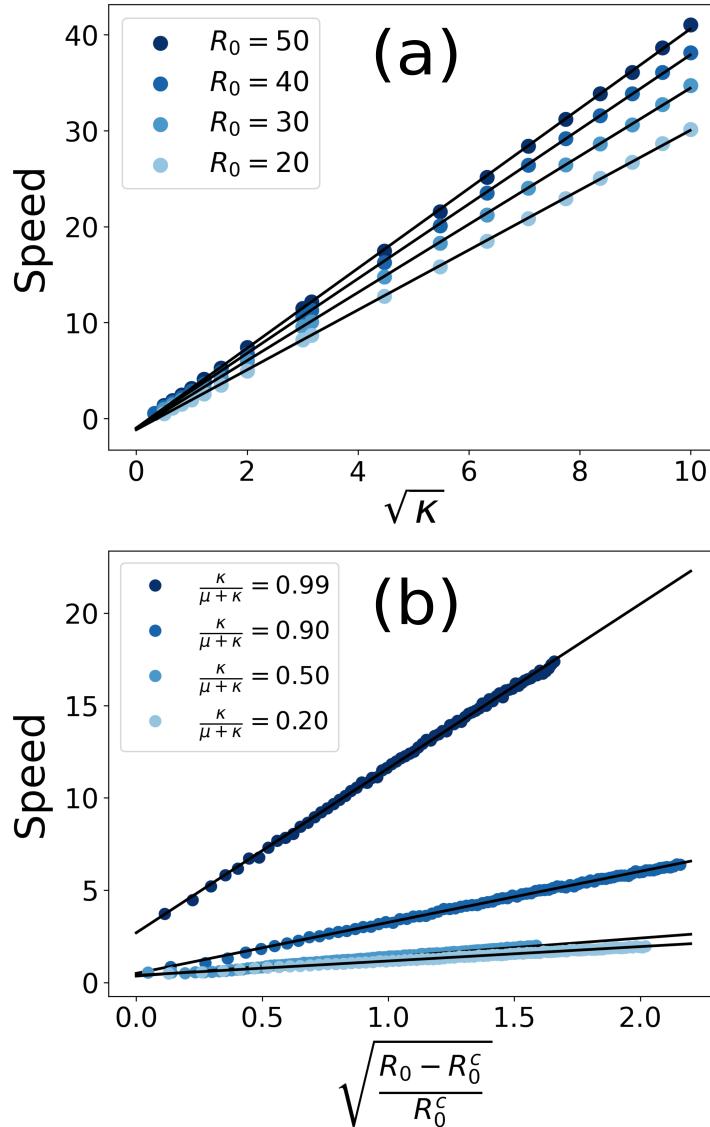


Figure 5.5: (a) Disease spreading speed as function of the square root of the parasite mobility for several values of R_0 . The plot shows a remarkable agreement with Eq. (5.11). (b) Disease spreading speed as function of the square root of the distance to the critical value of R_0 for several values of the parasite mobility.

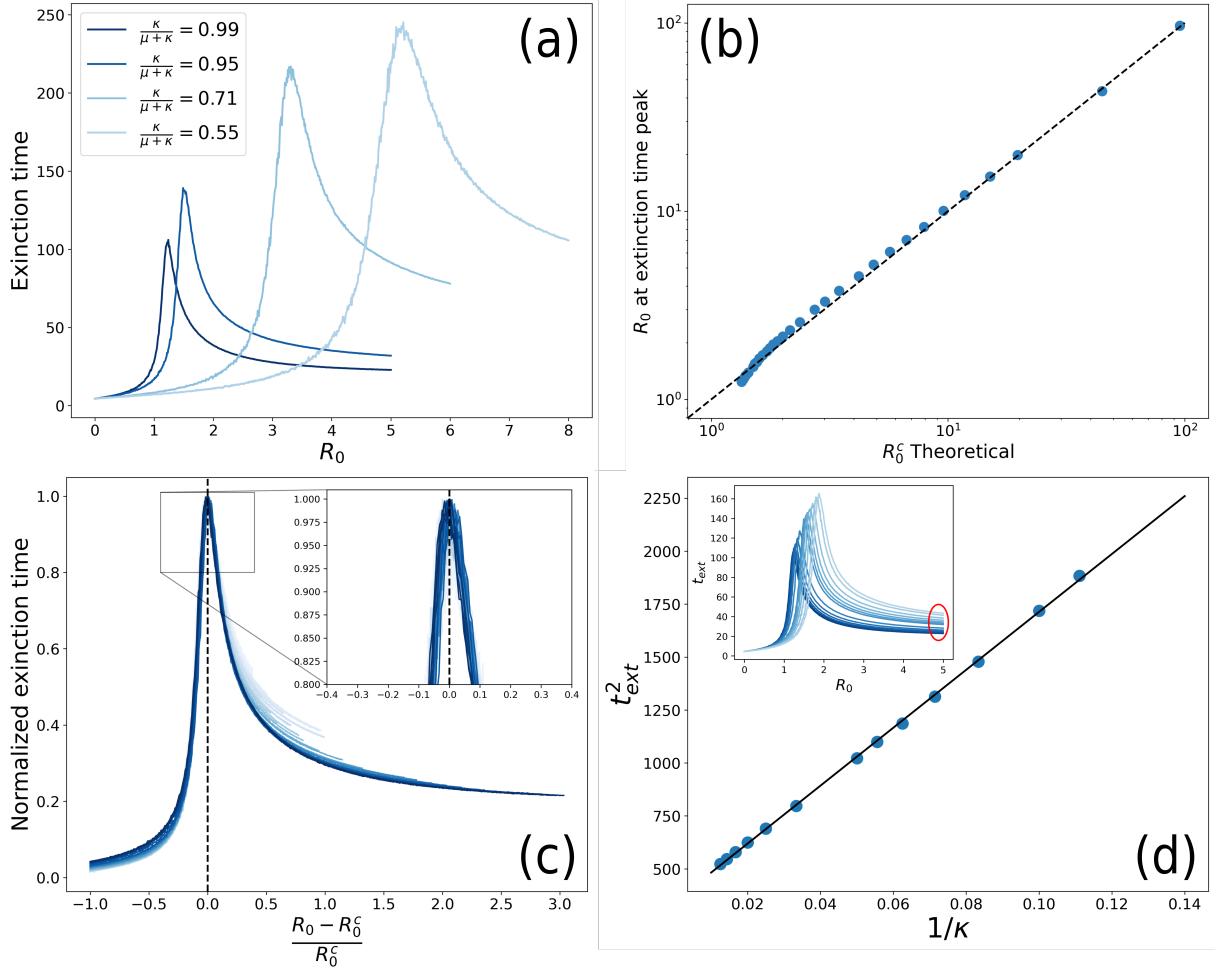


Figure 5.6: (a) Extinction time for some values of the parasite mobility. (b) Comparison of the critical R_0 value computed with Eq. (5.8) compared to the values obtained numerically by computing the maximum of the extinction time. (c) Scaling of the extinction time with several values of the parasite mobility. (d) Representation of the square of the extinction time as function of the inverse of the parasite mobility. The inset shows the zone where this relation has been computed, showing a good agreement with Eq. (5.12).

The extinction time is defined as the time elapsed from the beginning of the epidemic until the system reaches its absorbing state, that is, when no parasites or infected individuals are left. From Eq. (5.11), we expect the time to extinction to increase when the parasite mobility is decreased. Moreover, we expect the extinction time to decrease with the distance to the epidemic threshold, as we expect to reach faster the absorbing state for larger values of the spatial basic reproduction number.

In the limiting case where all (or almost all) hosts die, it is clear that the disease must have spread to the entire system. Thus, in this limit, the extinction time should be proportional to the inverse of the disease spreading speed, $t_{ext} \sim 1/v$. Then, in this limit, we can relate the extinction time with the parasite mobility as follows,

$$t_{ext} \sim \frac{1}{\sqrt{\kappa}} \quad \text{for } \tilde{R}_0 \gg 1. \quad (5.12)$$

However, the absorbing state is not always reached after all hosts becoming infected and Eq. (5.12) is only expected to work far from the epidemic threshold, when the disease is expected to spread to the entire system.

In Fig. 5.6(a) the extinction time is plotted against the basic reproduction number for some values of the parasite mobility. As expected, the extinction time increases for lower values of the parasite mobility. The increasing behaviour before the peak can be understood as the increasing time needed for the initial perturbations to decay to the disease-free phase. After the

peak, the greater the basic reproduction number the faster the epidemic will reach its absorbing state with a non-negligible number of dead individuals. So, with this interpretation, the peaks of the extinction time should coincide with the epidemic threshold for each value of the parasite mobility.

In Fig. 5.6(b) we compare the numerical value of R_0 at which the extinction time peaks with the theoretical value of R_0^c , computed with Eq. (5.10), showing good agreement. Thus, the dependence of the extinction time with R_0 should vanish if plotted against the distance to R_0^c . Furthermore, if the extinction time is normalised (dividing each line by its maximum), all the lines should collapse near the transition point. In Fig. 5.6(c) the normalised extinction time is plotted against the distance to the critical value of R_0 . The scaling is shown to be valid only near the transition point, as expected.

In the limiting case where the epidemic dies by infecting a large part of the host population, i.e. for a large enough R_0 value, the extinction time should follow Eq. (5.12), as previously discussed. In Fig. 5.6(d) we show how the extinction time relates to the parasite mobility in this limit, following the predicted behaviour.

5.4 Conclusions

In this work we have developed a spatially-explicit individual-based model for parasite-produced marine epidemics of immobile hosts. This study has allowed us to tackle important questions in marine epidemiology, as how spatial constraints affect epidemic spreading in filter-feeder populations or how will the infected population of hosts change in space and time. While addressing the aforementioned questions, we have shown that there exists a regime of high parasite mobility where the time progression of both host and parasite populations can be well described by the non-spatial version of the model (i.e. the system of ODE's presented in (79)). We have also shown that a fast-slow approximation for the time progression of the parasite population, already presented in (79), can be extended for spatial systems. Interestingly, the conditions under which this approximation is valid are less restrictive than in the non-spatial case, and regimes in which this approximation is valid for low mobility and comparable time scales are reported in this contribution.

We have derived an approximate analytical expression of the *spatial* basic reproduction number, that allows to predict the onset of a global epidemic in a spatial model. The obtained expression explicitly shows a trade-off between the intrinsic pathosystem dynamics (i.e. R_0) and a factor accounting for parasite mobility. Moreover, the spatial threshold defined by $\tilde{R}_0 = 1$ separates the final state of the system in two different phases, namely a disease-free phase and a propagation phase. In the propagation phase, any initial condition of infected individuals or parasites will propagate throughout the system, causing a proper outbreak. On the other hand, in the disease-free phase the conditions are not sufficient for a local introduction of parasites or infected individuals to spread through the system. The effect of the parasite mobility in the spatial basic reproduction number is clear, the more parasites move the more infections they cause.

The spatiotemporal behaviour of the system has been investigated in the propagation phase. First, we showed that the infected population spreads through the space with a speed directly proportional to the square root of the diffusion coefficient of parasites, showing good agreement between the derived analytical expression and numerical simulations. The time to extinction has been also studied by means of numerical simulations, showing that, if the system is far above enough of the spatial threshold, the time to extinction can be analytically computed, in good agreement with simulations. We obtained that larger values of the parasite mobility yield more severe epidemics in which there are more infections and the extinction is faster.

To summarise, in the present work we have introduced and analysed and Individual-Based approach to epidemic transmission in spatially extended systems of immobile hosts. The infection mechanism is due to mobile parasites, that are in turn produced by infected hosts. The study allows to answer some biologically relevant questions, like predicting the occurrence of a global epidemic outbreak or its velocity of expansion through the system. Thus, the analytical and computational results of the model shed light on the underlying mechanisms underpinning the emergence of a global epidemic outbreak and its spatial progression. This work provides a first step into the spatial-explicit individual-based modelling of marine epidemics of immobile hosts.

Although this work has considered the case of a spatially homogeneous distribution of hosts, we plan to extend the study to more general cases, discussing the effect of inhomogeneous spatial host distributions. Furthermore, other biological relevant effects could be added to the model to enhance the description of different epidemics, e.g. infected individuals could still filter parasites or parasite-load dependent infection process. The model could also describe epidemics on other immobile species such as filter feeders like sponges or other bivalves, corals, intertidal communities or starfishes provided that the necessary modifications in the model are properly included. Stochastic spatially-explicit descriptions like the one presented here could be also extended to the study of epidemics of other immobile hosts, like vector-borne diseases of plants. However, this would imply a quite different model to describe the different epidemic compartments of the vectors and also their ecological features. We hope these studies can be useful in conservation plans or ecosystem management and could serve as a basis for more sophisticated models.

5.5 Appendix

5.5.1 Derivation of the non-spatial equation for R_∞

The model described by the ODE system in Eq. (5.2) has a conserved quantity \mathcal{C} given by (79).

$$\mathcal{C} = P + \frac{\lambda}{\gamma} (S + I) - S - \frac{\mu}{\beta} \ln S \quad (5.13)$$

At $t = \infty$ the system reaches an absorbing state completely determined by $S(\infty)$, as $P(\infty) = I(\infty) = 0$ and $N = S(\infty) + R(\infty)$. Thus, from Eq. (5.13) we have

$$S(\infty) \left(\frac{\lambda}{\gamma} - 1 \right) - \frac{\mu}{\beta} \ln(S(\infty)) = \mathcal{C}_0 \quad (5.14)$$

The transcendental equation Eq. (5.14) can be solved by means of the Lambert's W function,

$$S(\infty) = -\frac{\mu\gamma}{\beta(\lambda-\gamma)} W_0 \left(-\frac{\beta(\lambda-\gamma)}{\mu\gamma} \exp(-\beta\mathcal{C}_0/\mu) \right) \quad (5.15)$$

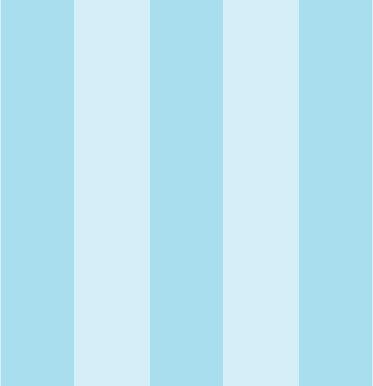
which can be simplified to

$$S(\infty) = -\frac{S(0)}{\xi} W_0 \left(-\xi \exp \left(-\frac{\beta}{\mu} C \right) \right), \quad (5.16)$$

with $\xi = S(0) \frac{\beta(\lambda-\gamma)}{\mu\gamma}$ and $C = P(0) + \frac{\lambda}{\gamma} (S(0) + I(0)) - S(0)$.

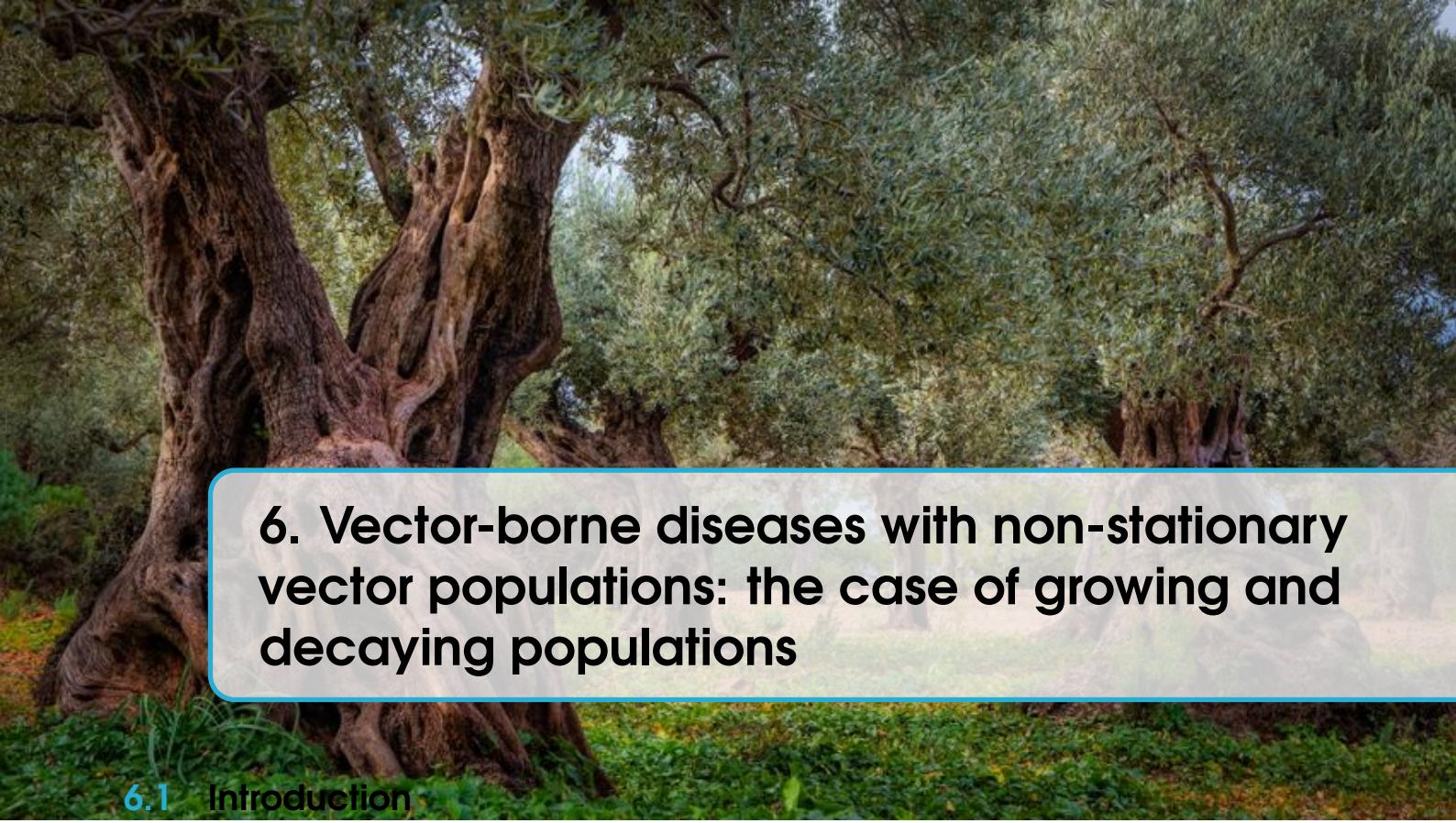
Finally, the absorbing state fulfils the condition $N = S(\infty) + R(\infty)$ so that the final number of dead individuals can be expressed as

$$R(\infty) = N + \frac{S(0)}{\xi} W_0 \left(-\xi \exp \left(-\frac{\beta}{\mu} C \right) \right). \quad (5.17)$$



Realistic models for vector-borne plant diseases

6	Vector-borne diseases with non-stationary vector populations: the case of growing and decaying populations	45
6.1	Introduction	45
6.2	The model	46
6.3	Results	48
6.4	Conclusions	55
6.5	Appendix	56
7	A compartmental model for <i>Xylella fastidiosa</i> diseases with explicit vector seasonal dynamics	59
7.1	Introduction	59
7.2	Materials and Methods	60
7.3	Results	63
7.4	Discussion	69
7.5	Appendix	71



6. Vector-borne diseases with non-stationary vector populations: the case of growing and decaying populations

6.1 Introduction

Vector-borne diseases are caused by infectious agents transmitted by living organisms, called vectors, frequently insects. These diseases represent a significant threat to global human health (9), causing diseases such as malaria, dengue, yellow fever, Zika, trypanosomiasis and leishmaniasis (151). Vector-borne human diseases are responsible of more than 17% of all human infectious diseases, causing millions of cases and more than 700000 deaths annually (174). Moreover, crop production and farm profitability are also affected by bacterial (90) and virus (22) vector-borne diseases. Some examples are the Pierce's Disease of grapevines, that has resulted in an annual cost of approximately \$100 million in California alone (162), the olive quick decline syndrome, which could cause about 5 and 17 billion US\$ of losses in Italy and Spain over the next 50 years in the absence of disease control measures (150) and the multiple diseases caused by viruses (142), with diseases like the tobacco mosaic, tomato spotted wilt, etc., transmitted by aphids and other vectors.

Compartmental deterministic models, e.g. the well known SIR model (102), have been widely used in the modeling of vector-borne diseases after the seminal work of Ross and Macdonald (114), that opened the way to controlling malaria outbreaks by acting on the vectors of the disease (the *Anopheles* mosquito). These models consider that both host and vector populations can be divided into different compartments describing different states of the individuals, such as susceptible, infected or removed (recovered or dead) (24), and the time-evolution of these compartments is expressed as a system of ordinary differential equations, defining a dynamical system. Compartmental models provide a mean-field description, that imply well-mixed (in practice spatially homogeneous) populations. The well mixed approximation will be valid whenever the mean distance among hosts is smaller than the mixing length of vectors before they die. In the case of vector-borne diseases it is also equivalent to every vector effectively interacting with all the hosts and every host with all the vectors. A mean-field description is not always valid in spatially extended systems, but still it is often the first step before writing a spatially explicit description.

The most relevant piece of information about a disease is whether an epidemic outbreak will take place or not. The *basic reproduction number*, R_0 , measures the number of secondary infections caused by an initial infected individual in a fully susceptible population, defining the epidemic threshold (5, 164), that determines the emergence (or not) of an outbreak. If $R_0 > 1$ an epidemic outbreak will occur, while there will be no outbreak otherwise. The standard way of computing R_0 in deterministic compartmental models is based on the existence of an initial disease-free (pre-pandemic) equilibrium, represented by the absence of infected hosts and vectors (97, 109). Some standard methods based on the linear stability condition of this equilibrium have been developed to allow the direct computation of R_0 , such as the Next-Generation Matrix (NGM) method (57).

In the case of vector-borne diseases, some models assume that populations (both hosts and vectors) do not change with time (see e.g. (25, 114, 163)), thus assuming equal birth and death rates. This guarantees the existence of a disease-free equilibrium and the proper use of standard methods to determine R_0 . However, this assumption could be far from reality in several pathosystems. For example, the interaction between temperature, precipitation variations and other factors may lead to strong variations in the vector population (69, 139), implying that the pre-pandemic state may not be an equilibrium state and that standard methods cannot be applied.

Compartmental models of vector-borne diseases have another feature that may hinder their practical applicability. Namely the fact that these models have many compartments, that describe the different states of both hosts and vectors, and as a consequence a relatively large number of parameters. This may lead to an issue known as *parameter identifiability and uncertainty* (43), depending on the available data, that is more likely to be found in models with many compartments and parameters (141). Usually, parameter estimation procedures are needed to connect the models with disease data, mainly using incidence or prevalence over time in the host population. Unfortunately, under many circumstances the underlying model parameters are unidentifiable, so that many different sets of parameter values produce the same model fit (98). Moreover, these parameters can be really difficult to determine from the available experimental data. Nevertheless, in some cases, mathematical manipulations can be performed to reduce the model complexity using exact or approximate relations (79). In such cases, the number of parameters of the models can be usually reduced in terms of new parameters defined as combinations of some of the original parameters.

The plan of this paper is as follows. In Section 6.2 we develop a compartmental model of vector-borne transmitted diseases with constant, but different, birth and death rates for the vectors, that we will use to describe the case of growing and decaying vector populations. For simplicity, the model assumes that there is no host to host direct transmission and that the development of the disease is faster than host recruitment, which is also a realistic assumption in many cases, like plant diseases. Section 6.3 contains the main results of the study. In particular, we show that the asymptotic approach fails to estimate the R_0 of the model, overlooking outbreaks if some conditions are fulfilled. Here, we provide an alternative method to compute R_0 based on the average number of secondary host infections produced by a primary infected host in one generation. It turns out that the validity of the asymptotic approach depends, among other things, on some time-scales of the model. Furthermore, we discuss and apply some approximations that allow to reduce the model in favor of simpler ones, with both fewer compartments and fewer parameters. In particular, we show that if some of the parameters fulfill certain conditions, it is possible to reduce the original model with five compartments and four parameters to a SIR model, with three compartments and two parameters. It is expected that model reductions like this one significantly help in solving possible problems of parameter unidentifiability that plague these models. It is interesting to note that a model in which hosts do not interact directly, but only through vectors, in a certain limit becomes described as if hosts would infect directly one to each other, what is assumed in some studies without suitable confirmation. Finally, the main concluding remarks of the study are presented in Section 6.4.

6.2 The model

The compartment model for vector-borne diseases that we will use to illustrate the points to be discussed in this study consists of 5 compartments, 3 of which describe the host population (susceptible, S_H , infected, I_H , and removed, R_H), while the other 2 describe the vector population (susceptible S_V and infected vectors, I_V). Thus, we consider that the pathogen affects only the hosts and do not consider exposed compartments. In addition, no direct host to host or vertical (or mother to offspring for vectors) transmission is assumed. The model could be also generalized to include an exposed host compartment and the above mentioned transmission modes, which would hinder the theoretical analysis without altering the qualitative conclusions of the study. Finally, for the host population we consider neither recruitment nor natural death and then the total host population, N_H , is constant, $N_H = S_H + I_H + R_H$. Finally, we assume that infected hosts do not have a mechanisms to combat the disease and become susceptible again. These assumptions are reasonable in the case of many phytopathologies.

The model is defined according to the following processes,



which are graphically described in Fig. 7.1, being the birth of new susceptible vectors described as a source term. Thus, the host-vector compartmental model is written as,

$$\begin{aligned} \dot{S}_H &= -\beta S_H I_V / N_H \\ \dot{I}_H &= \beta S_H I_V / N_H - \gamma I_H \\ \dot{R}_H &= \gamma I_H \\ \dot{S}_V &= \delta C - \alpha S_V I_H / N_H - \mu S_V \\ \dot{I}_V &= \alpha S_V I_H / N_H - \mu I_V, \end{aligned} \tag{6.2}$$

where the crossed nonlinear terms, $S_x I_y$, are written divided by the total host population, N_H , what corresponds to the so called standard incidence, which differs of the purely bilinear form known as mass action incidence (116).

The model describes infection of susceptible hosts, S_H , at a rate β through their interaction with infected vectors, I_V , while susceptible vectors, S_V , are infected at a rate α through their interaction with infected hosts I_H . Infected hosts exit the infected compartment at rate γ , while infected vectors stay infected the rest of their lifetime, as we consider that the pathogen does not affect them, as it is customary. Vectors die naturally (or disappear from the population by some mechanism) at rate μ and are born (appear) at a constant rate δ being susceptible. The constant term C sets the scale of the stationary value of the vector population. Fig. 7.1 shows an schematic representation of the model and we refer to (25) for a similar model of vector-borne diseases. However, the model in (25) includes exposed compartments and direct host to host transmission, but assumes that the birth and death rate of vectors are identical, and thus, the population does not change with time and stays as fixed by the initial condition.

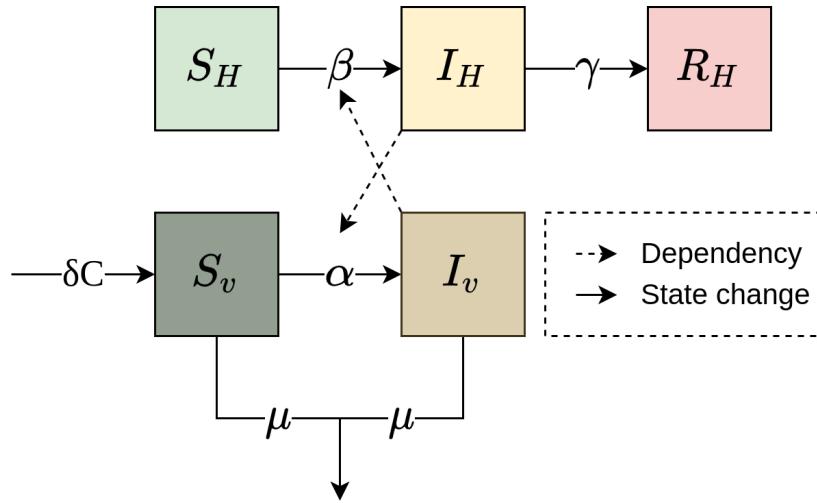


Figure 6.1: Schematic representation of the model n Eq. (6.2). Boxes are the compartments in which the population is divided, solid arrows represent changes in state (so transitions between compartments), and dashed arrows depict the crossed interaction between hosts and vectors.

6.2.1 Preliminary analysis of the model

From Eq. (6.2) it is straightforward to verify that the population of hosts remains constant over time, $N_H = S_H + I_H + R_H$, while the vector population fulfills,

$$\dot{N}_V = \dot{S}_V + \dot{I}_V = -\mu (S_V + I_V) + \delta C = -\mu N_V + \delta C, \tag{6.3}$$

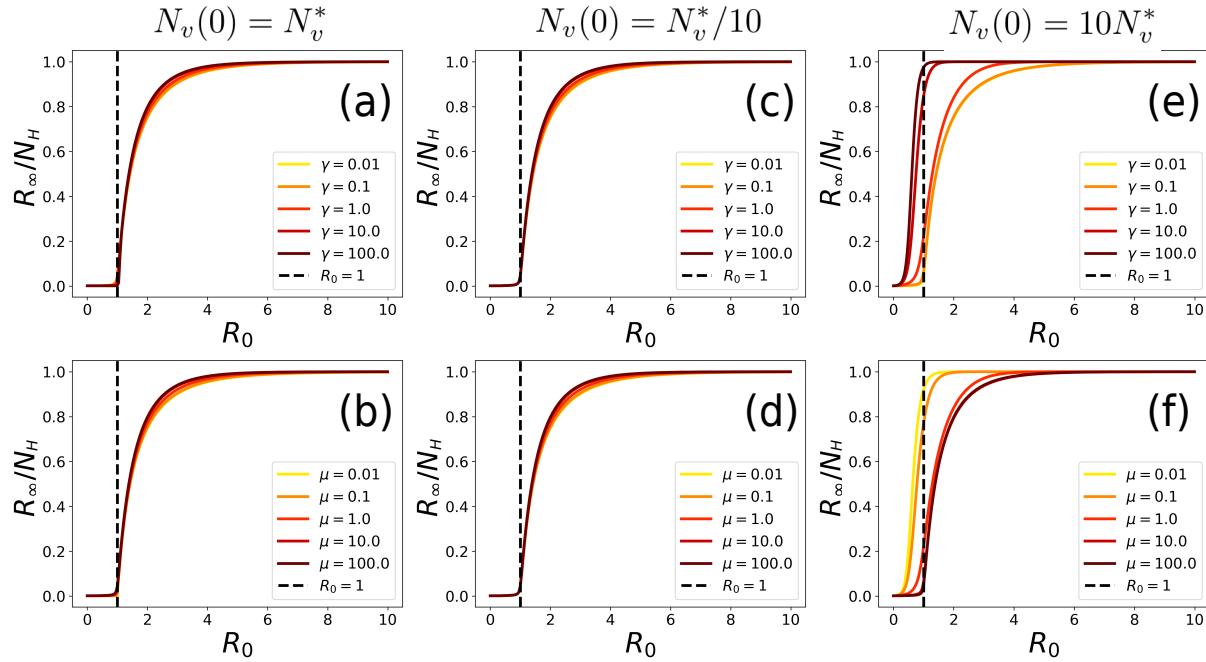


Figure 6.2: Numerical verification of the predictive power of the basic reproduction number relation Eq. (6.7), by plotting the final size of the epidemic, R_∞/N_H as function of R_0 . In panels (a),(b) the initial vector population is in the stationary value, in panels (c),(d) is below, $N_v^*/10$, and in panels (e),(f) above, $10N_v^*$. Panels (a),(c),(e) show realisations for different γ values with a fixed $\mu = 1$ baseline value. Panels (b),(d),(f) show realisations for different μ values with a fixed $\gamma = 1$ baseline value.

with solution,

$$N_v(t) = \frac{\delta}{\mu} C + \left(N_v(0) - \frac{\delta}{\mu} C \right) e^{-\mu t}. \quad (6.4)$$

From Eq. (6.4) one can write the stationary value for the vector population, N_v^* ,

$$N_v^* = \lim_{t \rightarrow \infty} N_v(t) = \frac{\delta}{\mu} C. \quad (6.5)$$

Thus, if the initial population of vectors is below (above) the stationary value, the vector population will grow (decrease) until it reaches the stationary value. On the other hand, if $N_v(0) = N_v^* = \delta C / \mu$ the initial population of vectors is already at the stationary state. The initial condition for the vector population can be written in terms of its stationary value Eq. (6.5), $N_v(0) = f N_v^*$, where both $f < 1$ and $f > 1$ are possible, so that one gets,

$$N_v(t) = N_v^* [1 + (f - 1) e^{-\mu t}]. \quad (6.6)$$

We note that vector-borne disease models that assume constant vector populations (e.g.(25)) can be recovered by setting $\delta = \mu$ and $C = N_v(0)$, so that any initial condition for the vector population is stationary, i.e. $\dot{N}_v = 0$ in Eq. (6.3) and $N_v(t) = N_v(0)$. We note that our model describes populations with an asymptotic stationary vector population and cannot describe periodic vector populations.

6.3 Results

6.3.1 The effect of non-stationary vector populations into the epidemic threshold and disease dynamics

Let us start with the cases in which any initial condition for the vector population is stationary and the total vector population remains unchanged. This will happen when the birth δ and death μ

vector rates are identical, independently on the initial condition of the vector population, or the case in which the initial condition of the vector population is already at its stationary value, $N_v(0) = N_v^*$, independently of the values of δ and μ . In such a case, the initial disease free state of the model, given by $I_H(0) = I_v(0) = 0$, is a fixed point (equilibrium state) of the dynamical system Eq. (6.2) independently of the other initial conditions for the host and vector populations. This allows the definition of the basic reproduction number, R_0 , using standard methods such as linear stability analysis or the Next-Generation Matrix (NGM) method (57) (see 6.5.1).

In other cases, the total vector population will vary with time provided that the initial condition, $N_v(0)$, is not identical to the asymptotic value at large times, N_v^* . In these cases, an initial disease-free state is not an equilibrium (fixed point) of the model. However, in the literature it is customary to apply the standard techniques, i.e. NGM, to compute R_0 using the vector population in the asymptotic state, that is the post-pandemic disease-free equilibrium (61, 108, 152, 170, 175). The use of these methods is supported by the fact that the asymptotic dynamics of the model converges to the dynamics of the subsystem where the vector population is stationary (36, 159). In both cases the basic reproduction number is given by,

$$R_0 = \frac{\beta\alpha}{\mu\gamma} \frac{S_H(0)}{N_H^2} N_v^*. \quad (6.7)$$

As usual, R_0 accounts for the number of secondary infections produced by an infected individual in one generation and controls the threshold behavior of the model: for $R_0 < 1$ the epidemic dies out and for $R_0 > 1$ an outbreak occurs. By one generation we refer to the typical time in which new infections can be produced, being the generation time in our model,

$$t_g = 1/\gamma + 1/\mu. \quad (6.8)$$

Now we will show that Eq. (6.7) is not always predictive about the onset of the epidemic. In Fig. 6.2 the final size of the epidemic, R_∞/N_H , is plotted as a function of R_0 , where R_∞ is the number of dead individuals at the end of the epidemic. Fig. 6.2(a)-(d) show that Eq. (6.7) does indeed regulate the onset of an epidemic when the initial vector population is in its stationary value or below it. This result is general and does not depend on the time-scales of the system, $1/\gamma$ and $1/\mu$, and so all curves in these panels behave similarly. In contrast, Fig. 6.2(e)-(f) shows that Eq. (6.7) does not predict the onset of epidemic outbreak when the initial vector population is larger than the stationary value. Thus, for $R_0 < 1$ (computed using Eq. (6.7)) severe outbreaks appear, yielding mortalities even above 80% of the total population. However, one can observe that as μ is increased, or γ decreased, the predictive power of Eq. (6.7) is progressively recovered.

Thus, only if the vector population reaches its stationary value before infected hosts have produced new infections can the onset of an epidemic be characterized by Eq. (6.7). Let us discuss separately the cases $f > 1$ and $f < 1$, with $N_v(0) = fN_v^*$, namely when the initial vector population is above and below its stationary value, this is, decaying and growing vector populations towards the asymptotic state.

If $f > 1$ Eq. (6.6), the time to approach the stationary value, t^* , is,

$$(1+\varepsilon)N_v^* = N_v^* \left[1 + (f-1)e^{-\mu t^*} \right], \quad (6.9)$$

where $\varepsilon \rightarrow 0$ is a small parameter controlling the amount by which the vector population differs from its asymptotic value at time t^* . Thus, the time to approach the stationary value, with precision ε , is given by

$$t^* = -\frac{1}{\mu} \ln \left(\frac{\varepsilon}{f-1} \right) = \frac{1}{\mu} \left| \ln \frac{\varepsilon}{f-1} \right|, \quad (6.10)$$

where the last equality assumes that the small parameter ε satisfies $\varepsilon < (f-1) > 0$.

If the vector population reaches its stationary value before infected hosts have had time to generate new infections then R_0 as determined from Eq. (6.7) is a good prediction of the onset for an epidemic, what is equivalent to the condition that t^* is much smaller than the hosts infectious period, $t^* \ll 1/\gamma$,

$$\frac{1}{\gamma} \gg \frac{1}{\mu} \left| \ln \frac{\varepsilon}{f-1} \right| \quad \text{or} \quad \frac{\mu}{\gamma} \gg \left| \ln \frac{\varepsilon}{f-1} \right|. \quad (6.11)$$

Otherwise, Eq. (6.7) will not be predictive of the epidemic onset, and as shown in Fig. 6.2(e-f) one may have outbreaks with a substantial final size with $R_0 < 1$.

In the case of growing vector populations, $f < 1$, if $R_0 < 1$ an outbreak cannot occur at all, because R_0 is calculated with the asymptotic population, N_v^* , that is larger than the vector population at any finite time, $N_v(t) < N_v^* \forall t$, and so the threshold condition is never attained. In the $R_0 > 1$ case the behaviour will be richer, and it will depend on the initial condition, $N_v(0)$. One can define an instantaneous basic reproductive number,

$$R_0^{(i)}(t) = \frac{\beta\alpha S_H(0)}{\mu\gamma N_H^2} N_v(t) = R_0 \frac{N_v(t)}{N_v^*}, \quad (6.12)$$

using $N_v(t)$ instead of N_v^* , with $R_0^{(i)}(t) < R_0 \forall t$ because the vector population grows. In particular, if $R_0^{(i)}(0) > 1$ there will be an outbreak occurring for short times, and the population of infected hosts will start growing. If instead, $R_0^{(i)}(0) < 1$, and as $R_0 > 1$ with R_0 being calculated with the asymptotic state, there must be an intermediate time, say t_D , for which $R_0^{(i)}(t_D) = 1$. Thus, from $t > t_D$ an outbreak will occur, not initially but after a finite time, that induces a delay in the outbreak, and the infected host population will start growing.

The difference between the original and the delayed dynamics stems from the waiting time to reach $R_0^{(i)} = 1$, t_D , plus the non-linear effect associated to a new initial condition for the epidemic outbreak at t_D . Thus, in the case that $R_0 > 1$ and $R_0^{(i)}(0) < 1$, from Eq. (7.5) and Eq. (6.6) we can analytically approximate the delay as the time needed to reach $R_0^{(i)}(t_D) = 1$,

$$1 + (f - 1)e^{-\mu t_D} = \frac{1}{R_0}, \quad (6.13)$$

which yields the relation,

$$t_D = -\frac{1}{\mu} \ln \left[\frac{1 - R_0}{(f - 1)R_0} \right], \quad (6.14)$$

where the argument of the logarithm is always positive because $R_0 > 1$ and $f < 1$. Eq. (6.14) is only valid if $f < 1/R_0$, for $R_0^{(i)}(0) = fR_0 < 1$, as if otherwise $R_0^{(i)} > 1$ the outbreak would already occur initially.

From Eq. (6.14) one can see that when the initial vector population is far enough from its stationary value, $f \rightarrow 0$, the delay saturates to a constant value, instead of increasing. This is,

$$\lim_{f \rightarrow 0} t_D = \frac{1}{\mu} \ln \left(\frac{R_0}{R_0 - 1} \right). \quad (6.15)$$

In addition, for increasing values of the basic reproduction number, R_0 , the delay tends to vanish, and from Eq. (6.15). This is,

$$\lim_{R_0 \rightarrow \infty} t_D = \frac{1}{\mu} \ln(1) = 0, \quad (6.16)$$

where the limit $f \rightarrow 0$ is taken simultaneously to guarantee that $R_0^{(i)}(0) = fR_0 < 1$. On the other hand the delay, t_D , scales with the vectors lifetime,

$$t_D \sim \frac{1}{\mu} = \tau_v. \quad (6.17)$$

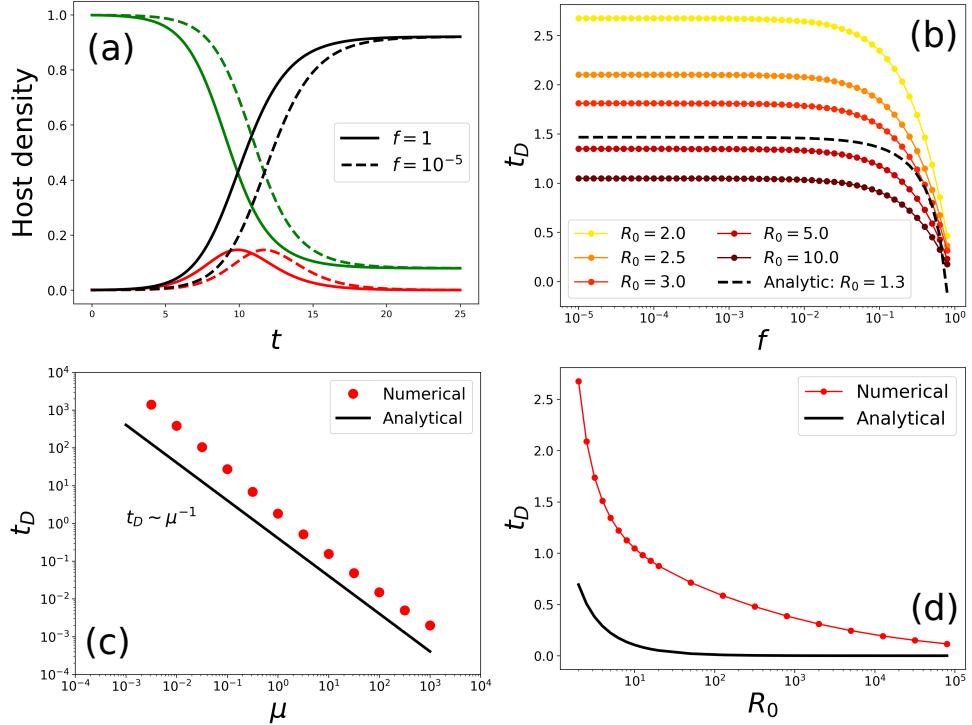


Figure 6.3: Numerical study of the delay induced by growing vector populations. (a) Comparison of hosts dynamics for a stationary vector population ($f = 1$) and a growing vector population ($f = 10^{-5}$). (b) Time delay as function of f for different values of the basic reproduction number R_0 . (c) Time delay as function of the vector natural death rate. (d) Time delay as function of the basic reproduction number, R_0 , with $f = 10^{-5}$.

Fig. 6.3(a) shows an example of the time delay caused in the hosts dynamics when the vector population grows from an initial condition far from the stationary value. In Fig. 6.3(b) we can qualitatively observe that all the predicted properties of the delay are fulfilled, namely, the time delay saturates for low f values and decreases with increasing R_0 . Although the analytical expression (black dashed line) is clearly not exact due to nonlinear effects, Eq. (6.14) captures the basic trends of the time delay, t_D . This is clear from Fig. 6.3(c), that shows that the delay scales with $1/\mu$ and in Fig. 6.3(d) that shows that the delay tends to 0 in the limit $R_0 \rightarrow \infty$, in agreement with the prediction of Eq. (6.16).

6.3.2 The basic reproduction number for non-stationary vector populations

As shown in the previous section, traditional methods to compute the basic reproduction number fail in the case of epidemic models with decaying vector populations, $f > 1$, unless the time scale of vector population fulfills the strong inequality condition Eq. (6.11), as illustrated in Section 6.3.1. Here we introduce an effective, average definition of R_0 , useful to predict the epidemic onset for vector-borne diseases with decaying vector populations, i.e. the case where traditional methods fail. It is defined as the average number of infections produced by an infected individual in one generation Eq. (6.8),

$$\overline{R}_0 = \langle R_0^i(t) \rangle \Big|_0^{t_g} = R_0 \left[1 - \frac{1}{\tau} (f - 1) (e^{-\tau} - 1) \right] = R_0 \cdot \mathcal{F} \quad (6.18)$$

where $\tau = 1 + \mu/\gamma$ and \mathcal{F} accounts for the effect of the decaying vector population on the stationary R_0 (see Section 6.5.2 for the full derivation of Eq. (6.18)).

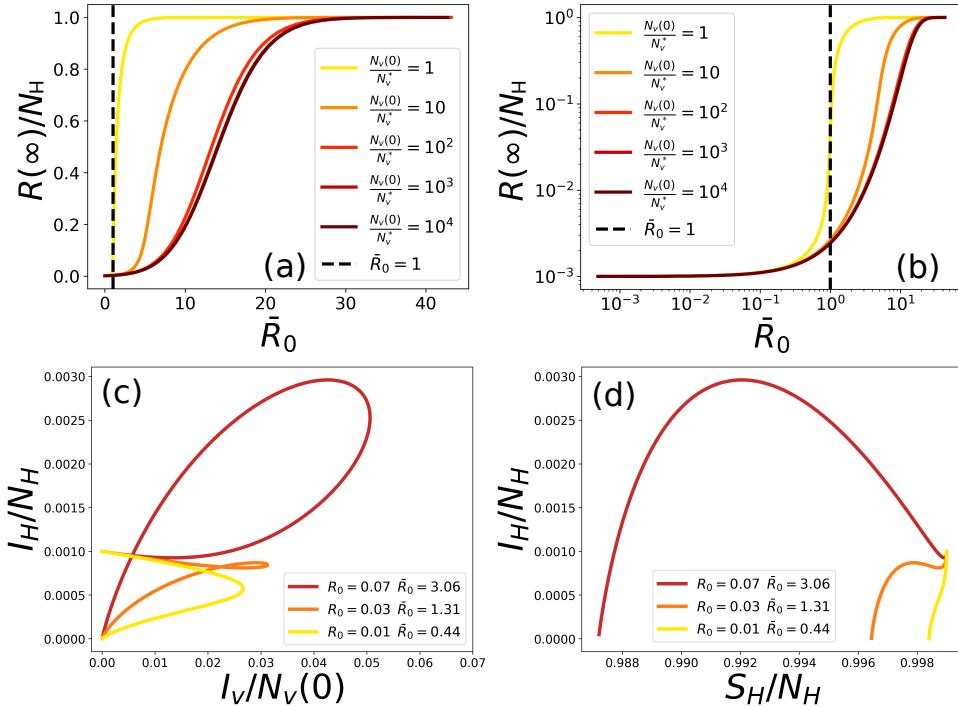


Figure 6.4: Numerical verification of the expression for the basic reproduction number for vector-borne diseases with decaying vector populations Eq. (6.18). Final size of the epidemic as a function of the basic reproduction number in panels: (a) linear scale; (b) logarithmic scale. Phase space trajectories in panels: (c) I_H/N_H vs $I_v/N_v(0)$ and (d) I_H/N_H vs S_H/N_H , where an initial condition $I_H(0)/N_H = 0.01, S_H(0)/N_H = 0.99$ and $I_v(0)/N_v(0) = 0$ has been used for the 3 cases. $\mu = \gamma$ has been used in all the simulations.

A first observation is that $\bar{R}_0 > R_0$ always (for $f > 1$). This stems from the fact that $\tau = 1 + \mu/\gamma > 1$, so that $e^{-\tau} - 1 < 0$, and $f - 1 > 0$, which yields $\mathcal{F} > 1$. This discussion unravels why standard methods fail to predict the onset of an epidemic under decaying vector populations. Another important point is that if $\mu/\gamma \gg 1$, which implies $\tau \gg 1$,

$$\lim_{\tau \gg 1} \mathcal{F} = \lim_{\tau \gg 1} \left[1 - \frac{1}{\tau} (f - 1) (e^{-\tau} - 1) \right] = 1 + \frac{f - 1}{\tau}, \quad (6.19)$$

and if furthermore $\tau \sim \frac{\mu}{\gamma} \gg (f - 1)$ then $\mathcal{F} \rightarrow 1$ and $\bar{R}_0 \rightarrow R_0$. This is in agreement with the discussion in Section 6.3.1 showing that the R_0 computed from standard methods works if $\mu \gg \gamma$.

Fig. 6.4(a-b) contrasts numerically the validity of Eq. (6.18) to predict the final size of the epidemic as a function of the general basic reproduction number, \bar{R}_0 , in linear and logarithmic scale, respectively. We observe that, independently of the initial condition of vectors, the outbreak occurs for $\bar{R}_0 > 1$. However, we may notice that for large values of the initial condition of vectors the final size of the epidemic grows more slowly, so that larger values of \bar{R}_0 are needed to produce a proper outbreak. This can be explained by the fact that for \bar{R}_0 slightly above the threshold, $\bar{R}_0 = 1$, and large values of $f = N_v(0)/N_v^*$, infections are produced only in the transient period of the dynamics, as $R_0 < 1$. This is, while the vector population is decaying to its stationary value, the vectors are able to produce new infections, but once the vector population reaches the stationary value, the epidemics stops. This transmission mechanism is radically different to that of vector-borne diseases with stationary vector populations in which the pre-pandemic disease-free state is an equilibrium of the system. The phase-space plots in Fig. 6.4(c-d) show that the time-averaged basic reproduction number \bar{R}_0 is able to accurately predict the conditions under which the infected host population will grow, in contrast with R_0 computed in the post-pandemic fixed point. In essence, for $\bar{R}_0 > 1$ the infected host population, I_H , grows before reaching the absorbing state, $I_H = I_v = 0$, while for $\bar{R}_0 < 1$ the infected host population is monotonically decreasing. We note that Eq. (6.18) is similar to the time-averaged basic reproduction number presented in (171) for the periodic case, which is a first-order approximation to the true basic reproductive number (12).

6.3.3 Fast-slow approximation

The original 5-D Eq. (6.2) model is certainly not amenable to mathematical analyses due to its high phase-space dimensionality and the fact that it depends on 4 parameters. Moreover, in a real-case application, if the parameters conforming the model are not known, the model could suffer from parameter unidentifiability. However, some approximations can be performed to reduce the mathematical complexity of the model, as for instance a fast-slow (or adiabatic) approximation.

If the time-scale of the vector population evolution is much faster than that of the infected hosts, what is expected to be a good approximation in many practical cases, the vector population will almost instantaneously adapt to its stationary value. Thus, if $1/\mu \ll 1/\gamma$, or equivalently if $\gamma/\mu \ll 1$, we can rewrite the time derivative of the vector infected population as

$$\varepsilon \dot{I}_v = \frac{\alpha}{\mu} S_v \frac{I_H}{N_H} - I_v , \quad (6.20)$$

where time has been re-scaled to $t' \rightarrow \gamma t$ and $\varepsilon = \gamma/\mu$ is a small parameter. Then, \dot{I}_v can be neglected and the infected vector population can be obtained from the relationship,

$$I_v \approx \frac{\alpha}{\mu} \frac{S_v I_H}{N_H} . \quad (6.21)$$

Substituting Eq. (6.21) into the original system Eq. (6.2) and the identity $N_v(t) = S_v(t) + I_v(t)$, while considering that the conditions for which the time-scale approximation is valid, $\mu \gg \gamma$, imply that the vector population will reach its stationary value almost instantaneously, so that $N_v(t) \approx N_v^*$, we obtain the following reduced system,

$$\begin{aligned} \dot{S}_H &= -\beta' \frac{S_H I_H}{\lambda N_H + I_H} \\ \dot{I}_H &= \beta' \frac{S_H I_H}{\lambda N_H + I_H} - \gamma I_H \\ \dot{R}_H &= \gamma I_H , \end{aligned} \quad (6.22)$$

where $\beta' = \beta N_v^*/N_H$ and $\lambda = \mu/\alpha$.

Moreover, if $f \neq 1$ the above mentioned timescales relationship must fulfil $\frac{\mu}{\gamma} \gg \left| \ln \frac{\varepsilon}{f-1} \right|$ (cf. Eq. (6.11)) and not only $\frac{\mu}{\gamma} \gg 1$. It is important to notice that the presence of direct host to host transmission would simply re-scale the coefficient β' , and the SIR reduction Eq. (6.22) would keep its validity.

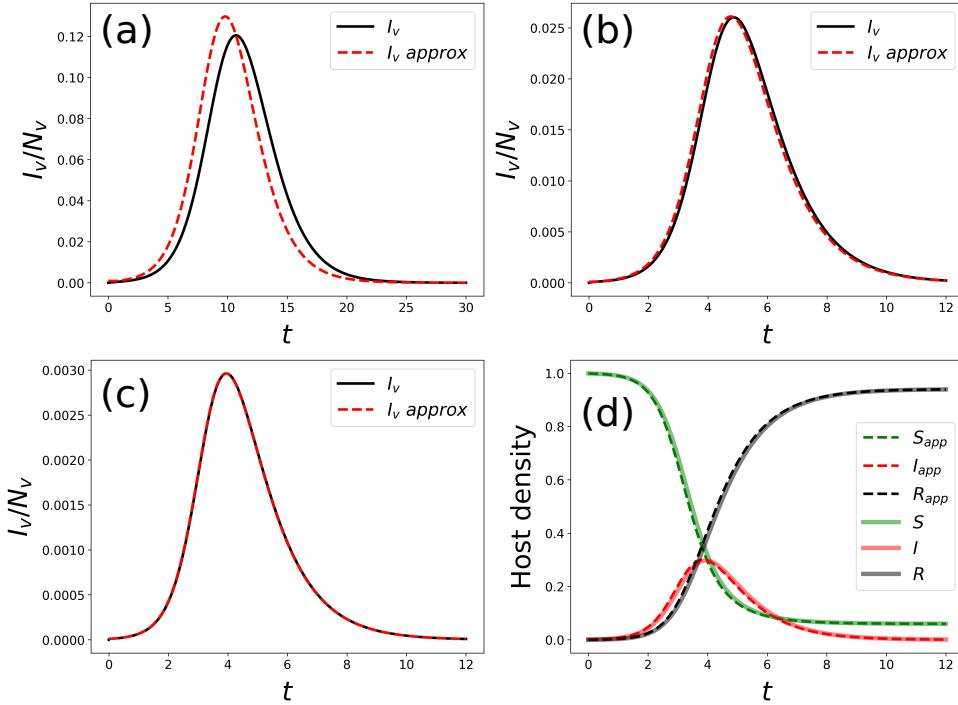


Figure 6.5: Numerical verification of the time-scale approximation (Eq. (6.21)) with $N_H = 100$, $\alpha = \gamma = 1$. β is chosen such that $R_0 = 3$. (a) $\mu = 1$, (b) $\mu = 10$, (c) $\mu = 100$. Panel (d) shows a comparison between the approximate and original models for the parameters used in (c), where the approximated models is expected to represent well the original one.

In Fig. 6.5 we numerically verify the validity of the presented fast-slow approximation. As expected, we observe that the approximation breaks down for $\mu \sim \gamma$ (Fig. 6.5(a)), while as μ becomes larger than γ the approximation improves Fig. 6.5(b) and it becomes quantitative when $\mu \gg \gamma$, Fig. 6.5(c). Finally, we show in Fig. 6.5(d) a comparison between the dynamics of the hosts using both the original and the approximated model using the same parameters than in Fig. 6.5(c), where the results of both models are expected to converge.

6.3.4 Reduction to a SIR model

In addition to the previous condition, $\gamma/\mu \ll 1$, if one has that $\lambda N_H \gg I_H$ also holds (which is indeed plausible in this limit) Eq. (6.22), then the model can be written as a standard SIR model,

$$\begin{aligned} \dot{S}_H &= -\beta_{eff} \frac{S_H I_H}{N_H} \\ \dot{I}_H &= \beta_{eff} \frac{S_H I_H}{N_H} - \gamma I_H \\ \dot{R}_H &= \gamma I_H , \end{aligned} \quad (6.23)$$

$$\text{where } \beta_{eff} = \frac{\beta'}{\lambda} = \frac{\beta \alpha N_v^*}{\mu N_H} .$$

In Fig. 6.6 we show the validity of the reduced models Eq. (6.22) and Eq. (6.23). Fig. 6.6(a) shows that the SIR-like model (Eq. (6.22)) works when the time-scale approximation can be performed (as $\mu/\gamma \gg 1$) but the SIR model fails when the condition $\lambda N_H \gg I_H$ is not fulfilled. Conversely, in Fig. 6.6(b) we show that as $\lambda N_H \gg I_H$ is fulfilled, then the SIR model perfectly matches the original model. Finally, Fig. 6.6(c) shows the decrease in the mean squared error of the approximation as the condition Eq. (6.11) is fulfilled for different values of f .

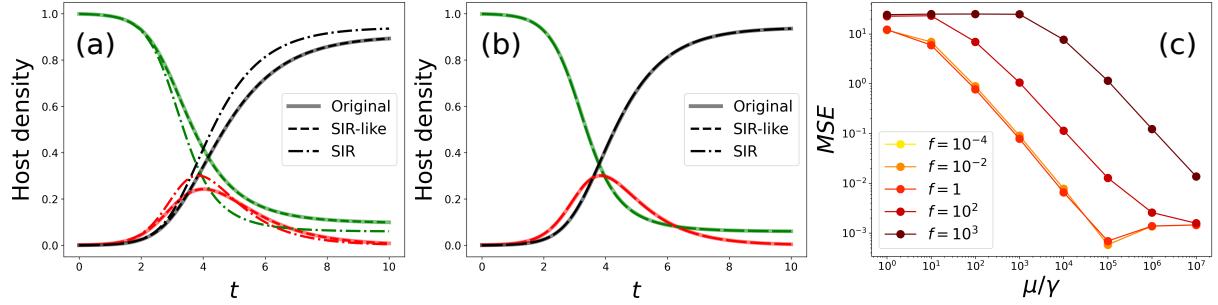


Figure 6.6: Comparison between the original model and the reductions, Eq. (6.22) (SIR-like) and Eq. (6.23) (SIR) with $N = 100$, $\mu/\gamma = 10^3$ and $f = 1$. β was chosen such that $R_0 = 3$. (a) $\lambda = 1$, (b) $\lambda = 10^3$, (c) Mean Squared Error between the original model and the SIR approximations as function of the ratio μ/γ and f .

6.4 Conclusions

In the present work we have analyzed several features of a compartmental deterministic model for vector-borne diseases with 3 compartments for hosts and 2 for vectors, that does not consider neither direct host to host nor vertical transmission. The goal is to study the behavior of the model in the case that the vector population is not stationary. In this case, the pre-pandemic disease-free state is not a fixed point (equilibrium state) of the dynamical system, and, in principle, the methods that are customarily used to determine the basic reproduction number, R_0 do not work. This is so because these methods determine the onset of an outbreak by performing a linear stability analysis of the disease-free state, assuming that it is a fixed point of the model. A common assumption made in the literature is to determine R_0 from the asymptotic state for the vectors (if it is not an extinction state), a fixed point of the model.

We have analyzed several initial conditions of the vector population, characterizing different regimes. In the case that the initial condition for the number of vectors is below the asymptotic state, implying that the vector population overall grows, then R_0 as determined from the asymptotic state correctly predicts the existence (or not) of an epidemic outbreak, but with a temporal delay in its appearance. This result contrasts with the situation in which the initial state is above the asymptotic state, with an overall decrease in the vector population. In this case R_0 determined from the asymptotic state may fail badly, predicting no outbreak while a large fraction of the population might get infected. We present a simple, albeit useful, generalization of R_0 that is able to give a reasonable prediction of the epidemic threshold for decaying populations, including the case in which vectors become extinct, a case in which the asymptotic estimation to determine R_0 cannot be applied.

Compartmental models of vector-borne diseases usually have many compartments and parameters, which can lead to a problem of parameter unidentifiability. The model analyzed here is not an exception, and when applied to real-world cases many different combinations of the parameters could be able to reproduce the available data. Thus, in order to facilitate the application of the model to experimental data, we have studied a useful fast-slow (or adiabatic) approximation that allows to reduce the model if the parameters fulfill certain conditions. In particular, our study shows that under quite realistic assumptions (the typical timescale of hosts infection and death is much slower than vector timescales) it is possible to obtain a reduced SIR model. We recall that this reduction implies that, under these assumptions, the process by which hosts (that could be immobile) get infected through the action of vectors is equivalent to a direct interaction among hosts.

The deterministic compartmental model analyzed here, with some modifications, is a clear candidate to study many vector-borne diseases, in particular phytopathologies. Furthermore, in case of parameter unidentifiability the model reductions performed in this work could be useful to solve this issue. In any case, this description is still idealized, as compartmental models imply a well-mixed assumption in which space is not explicitly described. This kind of representations are not always applicable to real-world scenarios although are useful as a first approximation. Thus, future research should focus on the integration of space and vector mobility in the model to account for more realistic situations.

6.5 Appendix

6.5.1 Calculation of R_0 from standard methods

The standard methods of calculation of R_0 are based in the linear stability analysis of the disease-free equilibrium, either directly, through the linear analysis of the fixed point, that yields the stability condition from which R_0 can be obtained, or using the Next Generation Method (NGM) (57) that provides directly R_0 by solving a suitable linear problem. Customarily these methods are applied to a pre-pandemic disease-free equilibrium, but as there is no such state in the case of non-stationary populations, here a similar approach is applied to a post-pandemic or asymptotic disease-free equilibrium.

Linear stability analysis

In order to perform the linear stability analysis of the fixed point ($I_H = I_v = 0$) we first need to compute the Jacobian matrix, J ,

$$J = \begin{pmatrix} -\beta \frac{I_v}{N_H} & 0 & 0 & -\beta \frac{S_H}{N_H} \\ \beta \frac{I_v}{N_H} & -\gamma & 0 & \beta \frac{S_H}{N_H} \\ 0 & -\alpha \frac{S_v}{N_H} & -\alpha \frac{I_H}{N_H} - \mu & 0 \\ 0 & \alpha \frac{S_v}{N_H} & \alpha \frac{I_H}{N_H} & -\mu \end{pmatrix} \quad (6.24)$$

Then, we evaluate the Jacobian at the fixed point (or disease free equilibrium, DFE), yielding

$$J|_{DFE} = \begin{pmatrix} 0 & 0 & 0 & -\beta \\ 0 & -\gamma & 0 & \beta \\ 0 & -\alpha \frac{C}{N_H} \frac{\delta}{\mu} & -\mu & 0 \\ 0 & \alpha \frac{C}{N_H} \frac{\delta}{\mu} & 0 & -\mu \end{pmatrix} \quad (6.25)$$

where $S_H = N_H$ has been considered.

The eigenvalues of Eq. (6.25) are,

$$\begin{aligned} \lambda_0 &= 0 \\ \lambda_\mu &= -\mu \\ \lambda_\pm &= -\frac{(\gamma + \mu)}{2} \pm \frac{1}{2} \sqrt{(\gamma - \mu)^2 + 4\beta\alpha \frac{C}{N_H} \frac{\delta}{\mu}} \end{aligned} \quad (6.26)$$

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

$$\lambda_+ = 0 \Rightarrow \frac{\beta\alpha}{\gamma\mu} \frac{C}{N_H} \frac{\delta}{\mu} = 1 \quad (6.27)$$

So, for $\frac{\beta\alpha}{\gamma\mu} \frac{C}{N_H} \frac{\delta}{\mu} < 1 \Rightarrow \lambda_+ < 0$ the fixed point is stable and for $\frac{\beta\alpha}{\gamma\mu} \frac{C}{N_H} \frac{\delta}{\mu} > 1 \Rightarrow \lambda_+ > 0$ a perturbation will grow in the direction of the eigenvector associated to λ_+ . Thus, this threshold defines the basic reproduction number,

$$R_0 = \frac{\beta\alpha}{\gamma\mu} \frac{C}{N_H} \frac{\delta}{\mu} \quad (6.28)$$

If instead of $S_H = N_H$ one considers any initial condition of hosts, $S_H(0)$, the basic reproduction number is given by,

$$R_0 = \frac{\beta\alpha}{\gamma\mu} \frac{C}{N_H} \frac{\delta}{\mu} \frac{S_H(0)}{N_H} \quad (6.29)$$

Next Generation Matrix method

The previous result can also be obtained by means of the NGM method, which is explained in detail in (57). Basically the method is based in decomposing the Jacobian in the form $J = T + \Sigma$, where T is the *transmission part*, that describes the production of new infections, and Σ the *transition part*, that describes changes of state (including death). Then, it can be proved (57) that the *basic reproduction number* R_0 is given by the spectral radius (i.e. the largest eigenvalue) of the (next generation) matrix $K = -T\Sigma^{-1}$.

$$K = -T\Sigma^{-1} = \begin{pmatrix} \frac{\beta\alpha}{\gamma\mu} \frac{C}{N_H} \frac{\delta}{\mu} & \frac{\beta}{\mu} \\ 0 & 0 \end{pmatrix} \quad (6.30)$$

with,

$$T = \begin{pmatrix} 0 & \beta \frac{N_H}{N_H} \\ 0 & 0 \end{pmatrix}, \quad \Sigma = \begin{pmatrix} -\gamma & 0 \\ \alpha \frac{C}{N_H} \frac{\delta}{\mu} & -\mu \end{pmatrix}$$

and $-\Sigma^{-1} = \begin{pmatrix} \frac{1}{\gamma} & 0 \\ \frac{\alpha}{\gamma\mu} \frac{C}{N_H} \frac{\delta}{\mu} & \frac{1}{\mu} \end{pmatrix}$

The basic reproduction number is the spectral radius of this matrix so:

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

Solving for σ one obtains the solutions,

$$\sigma_1 = \frac{\beta\alpha}{\gamma\mu} \frac{C}{N_H}; \quad \sigma_2 = 0 \quad (6.31)$$

Therefore, the basic reproduction number is

$$R_0 = \frac{\beta\alpha}{\gamma\mu} \frac{C}{N_H} \frac{\delta}{\mu} \quad (6.32)$$

If instead of $S_H = N_H$ one considers any initial condition of hosts, $S_H(0)$, the basic reproduction number is given by,

$$R_0 = \frac{\beta\alpha}{\gamma\mu} \frac{C}{N_H} \frac{\delta}{\mu} \frac{S_H(0)}{N_H} \quad (6.33)$$

6.5.2 Calculation of R_0 for non-stationary vector populations

We extend the computation of R_0 in the case of non-stationary and non-periodic vector populations by following the natural definition of *basic reproductive number*. Thus, R_0 is computed by averaging the number of secondary infections produced by an infected individual along one generation, that is equivalent to averaging the instantaneous definition of R_0 , namely R_0^i , over one generation,

$$\overline{R_0} = \langle R_0^i(t) \rangle \Big|_0^{t_g} = \frac{R_0}{N_v^*} \langle N_v(t) \rangle \Big|_0^{t_g} = \frac{R_0}{N_v^*} \frac{1}{t_g} \int_0^{t_g} N_v(t) dt, \quad (6.34)$$

where the integral in Eq. (7.6) is solved as

$$\begin{aligned} \int_0^{t_g} N_v(t) dt &= \left[N_v^* t - \frac{1}{\mu} (N_v(0) - N_v^*) e^{-\mu t} \right]_0^{t_g} = \\ &= N_v^* t_g - \frac{1}{\mu} (N_v(0) - N_v^*) [e^{-\mu t_g} - 1]. \end{aligned} \quad (6.35)$$

Thus, the basic reproduction number for non-stationary vector populations is given by

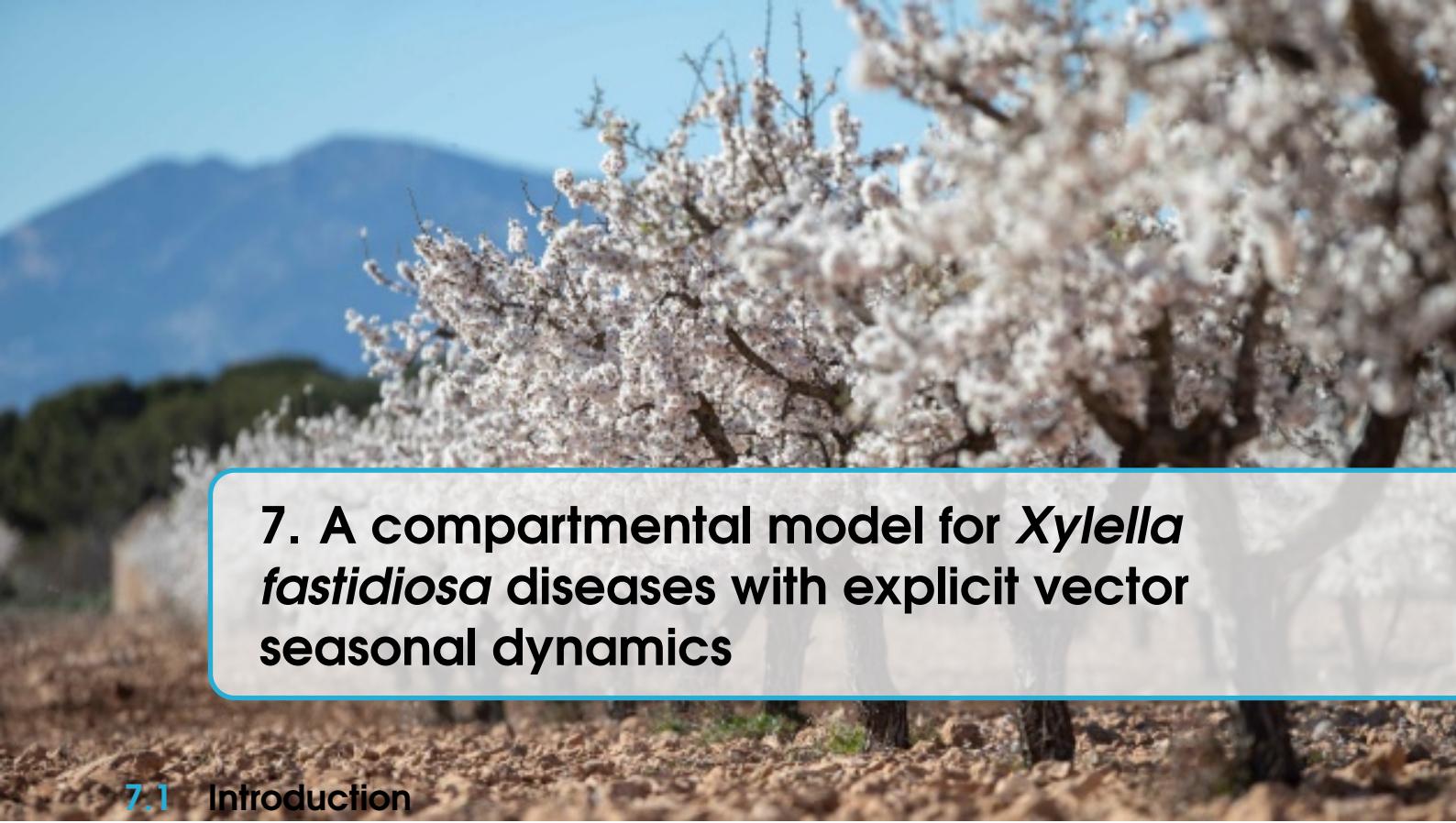
$$\overline{R}_0 = \frac{R_0}{N_v^*} \left\{ N_v^* - \frac{1}{\mu t_g} [N_v(0) - N_v^*] [e^{-\mu t_g} - 1] \right\}, \quad (6.36)$$

where the generation time, t_g , is Eq. (6.8). Eq. (7.8) can be rewritten as,

$$\overline{R}_0 = \langle R_0^i(t) \rangle \Big|_0^{t_g} = R_0 \left[1 - \frac{1}{\tau} (f - 1) (e^{-\tau} - 1) \right] = R_0 \cdot \mathcal{F}, \quad (6.37)$$

where $\tau = 1 + \mu/\gamma$ and \mathcal{F} is the expression in brackets, which accounts for the effect of the decaying vector population on the stationary R_0 .

In our approach, a generation is defined as the time elapsed in the following sequence of processes: 1) A host individual becomes infected; 2) The infected host passes the disease to a susceptible vector; 3) The infected vector dies. Basically, the time elapsed from the first to the last process is the time in which new infections can be produced, i.e. t_g Eq. (6.8).



7. A compartmental model for *Xylella fastidiosa* diseases with explicit vector seasonal dynamics

7.1 Introduction

Mathematical and computational modeling in Ecology and, in particular, Epidemiology have been recently recognized as powerful approaches to guide empirical work and provide a framework for the synthesis, analysis and development of conservation plans and policy-making ((40, 110, 125, 148)). Plant epidemics, mainly plant-virus diseases, have been often described by compartmental models, which deal with the overriding importance of transmission mechanisms in determining epidemic dynamics ((94, 95, 115)). These models have contributed to providing answers to some questions related to the ecology of plant diseases and have led to direct applications in disease control while guiding research directions ((92)).

The emergence of vector-borne plant pathogens in new areas causing huge economic impacts, such as *Xylella fastidiosa* and the *Candidatus Liberibacter* spp. (Huanglongbing or citrus greening), has sparked interest in modeling vector-transmitted plant-disease epidemics ((41, 92)). The vector-borne bacterium *X. fastidiosa* (*Xf*) is a multi-host pathogen endemic to the Americas that causes economically important diseases, mostly in woody crops ((89)). *Xf* is a genetically diverse species with three evolutionary well-defined clades forming the *pauca*, *fastidiosa*, and *multiplex* subspecies, native from South, Central, and North America, respectively ((165)). Within each subspecies, diverse genetic lineages with different host ranges are found. *Xf* is transmitted non-specifically by xylem-sap-feeding insects belonging to the sharpshooter leafhoppers (Hemiptera: Cicadellidae, Cicadellinae) and spittlebugs (Hemiptera: Cercopoidae) ((137)).

Recently, *Xf* has gained renewed interest due to the massive mortality of olive trees in Apulia, Italy ((146)). The first focus of the olive quick decline syndrome (OQDS) was detected in 2013 around Gallipoli (Apulia, Italy)((145)) and since then has spread throughout the region by the meadow spittlebug, *Philaenus spumarius*. Although this was the first official detection of *Xf* in Europe, it has recently been demonstrated that the pathogen arrived much earlier in Corsica ((155)) and in the Balearic islands ((122)). Around 1993, two strains of the subspecies *fastidiosa* (ST1) and *multiplex* (ST81) were introduced from California to Mallorca (Spain) with infected almond plants ((122)). To date, over 80% of the almond trees in Mallorca show leaf scorch symptoms and the outbreak has changed the iconic rural landscape of this Mediterranean island ((127)).

The meadow spittlebug, *P. spumarius* (Hemiptera: Aphrophoridae), has recently been shown to be the main vector of *Xf* in Europe both in transmission experiments and in field studies ((46, 48, 112, 121, 146)). *P. spumarius* is a polyphagous species from the Palearctic region, presenting one generation per year (univoltine) and overwintering as eggs. Foam-forming nymphs emerge at the end of winter, feeding on herbaceous plants. The time required for their development to the adult stage depends mainly on temperature and humidity ((19, 42, 45)). In Mediterranean climates, *P. spumarius* adults generally move from the herbaceous cover to the crop canopy

as evapotranspiration increases in late spring (May–June). In mid-summer, the populations of *P. spumarius* tend to decrease in the crop canopy, while the insects are captured more frequently in trees and shrubs interspersed in crops. Summer dispersal of spittlebugs to wild hosts as refugee seems a common general pattern in Mediterranean crops in Italy ((19, 49)) and Spain ((124)). Because the bacterium has not been detected in spring on insects feeding on the herbaceous cover or in weeds in Europe ((19, 45, 127)), it is assumed that all spittlebug adults acquire the bacteria from the main crop (olive, almond, vine, etc.). Once infected, Xf colonizes the insect foregut in a persistent and non-circulatory manner without transovarial (parent to offspring) or transstadial (inter-stage) transmission ((4, 67, 135)) and without a period latency after vector acquisition ((3, 67)).

Several epidemic models have been already developed for Xf-diseases, but they lack a realistic description of some relevant processes ((92)). Some of these models assume a simple general form for infected host dynamics ((1, 56, 172)) or use a simplified S-I compartmental scheme for hosts, disregarding important features such as the latent period or the host mortality rate ((155)). Models that do take these features into account, however, do not explicitly model the population of vectors responsible for disease transmission ((173)). Other more recent models have taken a step further in explicitly modeling the vector population ((28, 75)), but the characterization of its dynamics is still relatively simple, as it overlooks the known seasonal patterns of vector abundance. Several recent studies have provided new insights into the ecology and temporal dynamics of the transmission of Xf by *P. spumarius* in olive plants ((19, 20)). However, these experimental data of the pathosystem have not been yet integrated at the population level. Thus, there is a need to continue advancing in the modeling of Xf diseases by developing more realistic models that can elucidate the fundamental processes involved in vector-host-pathogen interactions and help to design effective control strategies.

In this work, we develop a deterministic continuous-time compartmental model to describe the general epidemiological dynamics of diseases produced by Xf in Europe. We explicitly account for key biological aspects of the disease, including the seasonal dynamics of its main vector, *P. spumarius*. Our model is able to describe field data from the two major European outbreaks: the olive quick disease syndrome (OQDS) in Apulia, Italy, caused by the subspecies *pauca*, and the almond leaf scorch disease (ALSD) in Mallorca, Spain, caused by subspecies *multiplex* and *fastidiosa*. We aimed to find the most influential parameters in the model with respect to incidence and mortality in both diseases by performing a global sensibility analysis. With this information, the next goal was to explore control strategies acting especially on the vector population.

7.2 Materials and Methods

7.2.1 Epidemic model: the SEIR-V model

We developed a deterministic continuous-time compartmental model that incorporates the specific biological features of Xf diseases in Europe, including the dynamics of the main relevant vector *P. spumarius* ((39)). To build the model we took the following considerations: (i) we assume there is no winter recovery of infected hosts and thus they die sometime after infection; (ii) hosts show an asymptomatic period in which they are non-infectious in practice (exposed compartment) because the bacteria are not yet systemically extended ((156, 158)), while vectors are infectious immediately after acquiring the bacterium ((62)); (iii) vectors have an annual life cycle without mother-to-offspring disease transmission ((67, 135)), so we consider the annual emergence of susceptible newborn vectors and a constant death rate for both susceptible and infected vectors; (iv) infected vectors carry the bacterium during their entire lifespan without affecting their fitness ; and finally, (v) we do not consider host recruitment or natural death given that the typical development time of Xf-epidemics is faster than the typical host's life cycle.

Altogether, our deterministic continuous-time compartmental model consists of six compartments, four describing the host population (susceptible, S_H , exposed, E_H , infectious, I_H , and removed, R_H), and two describing the vector population (susceptible, S_V , and infected, I_V). The

model is defined according to the following processes,



which are illustrated in Fig. 7.1, being the birth of new susceptible vectors described as a source term. Thus, the host-vector compartmental model is written as,

$$\begin{aligned} \dot{S}_H &= -\beta S_H I_V / N_H \\ \dot{E}_H &= \beta S_H I_V / N_H - \kappa E_H \\ \dot{I}_H &= \kappa E_H - \gamma I_H \\ \dot{R}_H &= \gamma I_H \\ \dot{S}_V &= N_V(0) \sum_{n=1}^{\infty} \delta(t - nT) - \alpha S_V I_H / N_H - \mu S_V \\ \dot{I}_V &= \alpha S_V I_H / N_H - \mu I_V. \end{aligned} \quad (7.2)$$

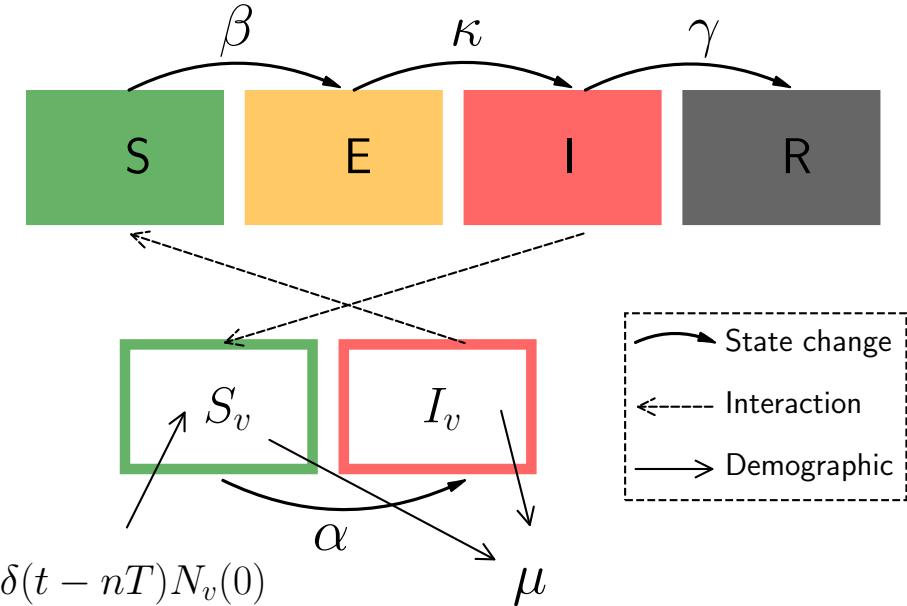


Figure 7.1: Schematic representation of the model Eq. (7.2). Boxes are the compartments in which the population is divided, solid curved arrows represent changes in state, i.e. transitions between compartments, dashed arrows depict the crossed interaction between hosts and vectors and solid straight arrows represent demographic changes in vector population.

The model describes the exposure of susceptible hosts, S_H , at a rate β through their interaction with infected vectors, I_V , while susceptible vectors, S_V , get infected immediately at a rate α through their interaction with infectious hosts I_H . Exposed hosts get infectious at rate κ , being the mean latent period $\tau_E = 1/\kappa$, while infectious hosts die at rate γ , having a mean infectious period of $\tau_I = 1/\gamma$. Infected vectors stay infected and infectious for the rest of their lifetime. Regarding the seasonal dynamics of vectors, we assume that new adults emerge synchronously each year in fields being all susceptible. This is represented by the term $N_V(0) \sum_{n=1}^{\infty} \delta(t - nT)$ in Eq. (7.2), where $T = 1 \text{ yr}$ is the period and $\delta(t - nT)$ is Dirac delta function, and basically implements a yearly pulse of new vectors at a certain moment in the year. Vectors are removed (die, move to herbaceous vegetation and other non-host trees, exit the field, etc) at a given rate μ , which we consider

identical for susceptible and infected vectors. For simplicity, we consider that the quantity of annual newborn adults, $N_v(0)$, is constant. This outburst of new adults followed by an exponential decay resembles the temporal patterns on the abundance of *P. spumarius* observed in crop fields ((7, 13, 46, 113)) (see Fig. 7.8).

In our model (Eq. (7.2)) the crossed nonlinear terms in \dot{S}_H and \dot{S}_v , $S_H I_v$ and $S_v I_H$, are divided by the total host population, N_H . Thus, the vector-to-plant infection process is modeled using mass action incidence, which is density dependent, while the plant-to-vector infection process is modeled using standard incidence, which is frequency dependent ((116)). This implies that doubling the number of vectors in the crop field would double the number of resulting exposed (or infected) hosts, as this process is population-dependent (mass action incidence), while doubling the number of hosts would not result in more vectors per unit area being infected, as this process only depends on the contact probability, being frequency dependent (standard incidence). We think this is the most reasonable assumption because, for a given plantation framework, increasing the number of hosts is expected to also increase the area of the field, while the number of vectors is an independent quantity.

7.2.2 Basic reproductive number

The basic reproductive number, R_0 , of the model cannot be trivially computed using standard methods such as the Next Generation Matrix (NGM) ((57)), as there is no pre-pandemic fixed point in the system of differential equations Eq. (7.2). For periodically varying vector populations, rigorous methods have been developed ((11)), but not for the case of growing or decaying vector populations. Here we use the simple method developed in the work of (76) (see Section 7.5.2), which effectively computes the average number of secondary infections produced by an initially infectious individual in one generation. Thus, the effective basic reproductive number is given by

$$R_0 = \frac{\beta\alpha}{\mu\gamma} \frac{S_H(0)}{N_H^2} \frac{N_v(0)}{\mu\tau} (1 - e^{-\mu\tau}) , \quad (7.3)$$

where τ corresponds to the time length of one generation, in our case one year. This R_0 is calculated using the initial susceptible host population, $S_H(0)$. Below we will also use a time-dependent $R_0(t)$ using $S_H(t)$.

7.2.3 Epidemiological data

Epidemiological data from an ALSD outbreak in the island of Mallorca, Balearic Islands, Spain were taken from (122). Dated phylogenetic analysis and estimates of disease incidence showed that the introduction of both subspecies occurred around 1993 and $\sim 79\%$ almond trees were infected in 2017 ((122)). The annual proportion of infected individuals in the almond tree population between 1993 and 2017 was estimated by analyzing through qPCR the presence of Xf-DNA in the growth rings of 34 sampled trees (cf. Fig. 3 in ((122))). The disease progression curve was estimated without distinguishing whether infections were caused by *multiplex* or *fastidiosa* subspecies. In addition, a two-sided bootstrap confidence interval for each data point was set using the SciPy bootstrap function in Python ((167)). On the other hand, epidemic data for OQDS were retrieved from ((173)). The data consisted of 2 to 3 yearly censuses of symptom prevalence in 17 olive groves infected with Xf subsp. *pauca* in Apulia, Italy, which were aggregated to fit our model as shown in Fig. 4 in ((173)). Because the compartments of our model are not in one-to-one correspondence with those shown in the work of White et al. ((173)), we used the sum of the symptomatic and desiccated infected trees in the dataset ($I_S + I_D$) to fit the sum of the infected and dead trees ($I + R$) and the sum of susceptible and asymptomatic hosts ($S + I_A$) to fit the sum of susceptible and exposed hosts ($S + E$). The processed data used to fit the model can be found in ((74)), while the raw data can be found in the supplementary data accessible online of the cited articles ((122, 173)).

7.2.4 Model fitting through Bayesian Inference

We employed an informative normal $\mathcal{N}(\hat{\mu}, \hat{\sigma}^2)$ prior distribution, with $\hat{\mu}$ and σ , the mean and standard deviation, respectively, for previously measured parameters in the literature, such as

the infectious and latent periods for ALSD, $\tau_I \sim \mathcal{N}(14, 4)$, $\tau_E \sim \mathcal{N}(4, 1)$ ((122, 158)) and OQDS, $\tau_I \sim \mathcal{N}(3.5, 1)$, $\tau_E \sim \mathcal{N}(1.75, 0.5)$ ((62)). The corresponding rates are given by $\gamma = 1/\tau_I$ and $\kappa = 1/\tau_E$, respectively. Similarly, a prior normal distribution was used for the removal rate of vectors, $\mu \sim \mathcal{N}(0.02, 0.0075)$, as the mean value $\mu = 0.02$ already captures the vector dynamics observed in field-data (Fig. 7.8). Regarding the prior distribution for the transmission rates a very wide and uninformative uniform prior distribution, $\beta \sim \mathcal{U}(0.001, 1)$ and $\alpha \sim \mathcal{U}(0.001, 1)$, was used for each parameter. The number of hosts, N_H , was already provided in the datasets, while, given the lack of information about the vector population, we assumed $N_v(0) = N_H/2$ for the initial vector population of each year. However, we tested the robustness of our results by changing $N_v(0)$.

The posterior distributions of the parameters were approximated using the Markov Chain Monte Carlo algorithm No U-Turn Sampler (NUTS) with the recommended target acceptance rate of 65% ((88)). To ensure a proper convergence, we constructed three independent Markov Chains with 10^5 iterations each after a burn-in of 10^4 iterations and checked that the results were statistically equivalent. For each chain, we started at the maximum-likelihood parameters yielded by the Nelder-Mead algorithm with 1000 iterations.

The parameters of our compartmental model were determined by fitting the model to data by means of a Bayesian Inference framework using the Turing.jl package ((70)) in Julia ((15)). The scripts used to fit the model can be found in ((74)).

7.2.5 Sensitivity Analysis

We performed a Global Sensitivity Analysis (GSA) ((144)) of the model to assess the relative contribution of its parameters and their interactions with different features of the epidemic. In contrast to the Local Sensitivity Analysis (LSA), the GSA assesses the influence of a large domain of the parameter space in the desired outputs of the model. We performed GSA by means of a variance-based analysis, the Sobol method ((154)). This particular method provides information not only on how a particular parameter alone influences the model outputs (as happens with LSA), but also due to the nonlinear interactions among two or more parameters. Briefly, the method considers the model output, Y , as a general function of the inputs, $f(x_1, \dots, x_n)$, so that the variance of the output, $Var(Y)$ is decomposed as the sum of the variances given by the variations of the parameters alone and its interactions, $Var(Y) = \sum_{i=1}^n Var(f(x_i)) + \sum_{i < j} Var(f(x_i, x_j)) + \dots$. This information is organized in what are known as Sobol indices. The total order indices are a measure of the total variance of the output quantity caused by variations of the input parameter and its interactions, $S_T = Var(f(x_1, \dots, x_n))/Var(Y)$. First order (or “main effect”) indices are a measure of the contribution to the output variance given by the variation of the parameter alone, but averaged over the variations in other input parameters, $S_i = Var(f(x_i))/Var(Y)$. Second-order indices take into account first-order interactions between parameters, $S_{ij} = Var(f(x_i, x_j))/Var(Y)$. Further indices can be obtained, describing the influence of higher-order interactions between parameters, but these are not going to be considered.

Following the Sobol method, we analyzed the variation of the time at which the infectious population peaks, t_{peak} , the magnitude of this peak, I_{peak} and the final number of dead hosts, R_∞ , relative to variations of the model parameters. The method was implemented within the Julia high-level programming language ((15)) using the sub-package DiffEqSensitivity.jl in DifferentialEquations.jl package ((136)).

7.3 Results

7.3.1 Model fit and parameter estimates

The posterior distributions of the fitted parameters including their estimated mean and median for ALSD and OQDS are shown in Figs. 7.2 and 7.3, respectively, together with the assumed prior distributions. We observe that the literature-driven priors for the latent and infectious period, τ_E and τ_I , were already very good guesses and changed slightly converging to the appropriate distribution that better fitted the epidemic data for both ALSD and OQDS (Fig. 7.2(A-B) and Fig. 7.3(A-B)). Similarly, the prior for the vector removal rate, μ , obtained from field data, was good enough so that little changes were needed for convergence (Fig. 7.2(C) and Fig. 7.3(C)). On the other hand, we also observe that the completely uninformative priors for the transmission rates successfully converged to the posterior distributions (Fig. 7.2(D-E) and Fig. 7.3(D-E)).

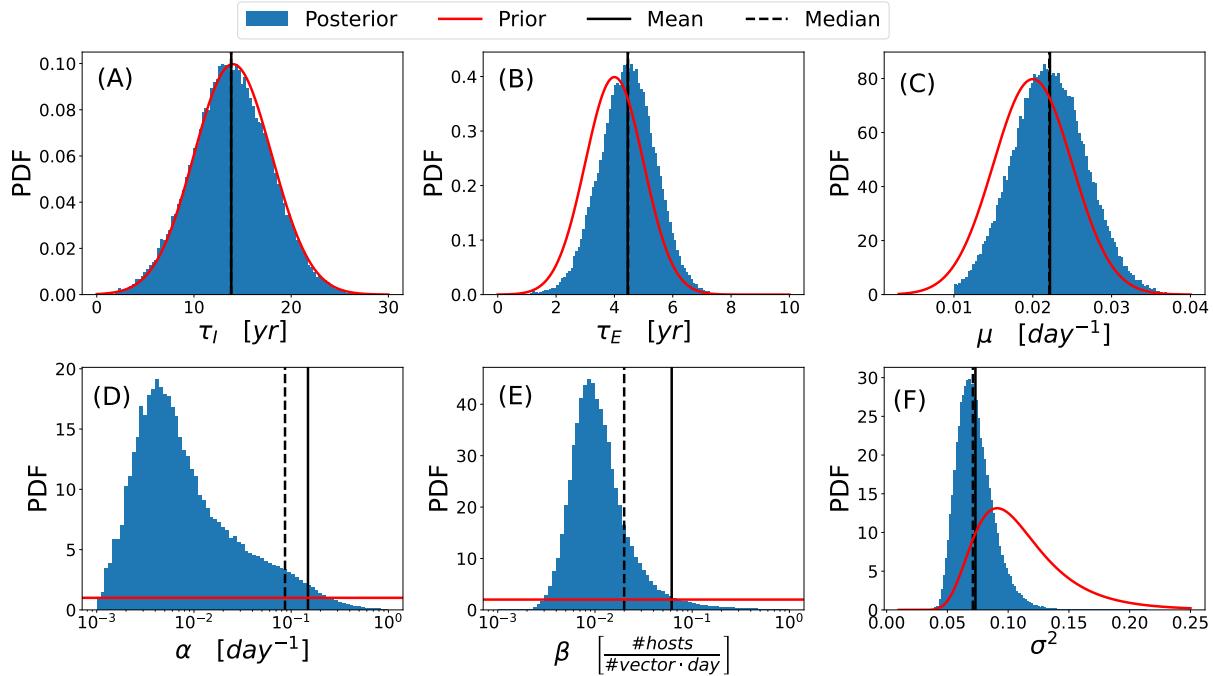


Figure 7.2: Posterior (blue histograms) and prior (red line) distributions of the model parameters for ALSD. Solid and dashed black lines correspond to the mean and median of the posterior distributions. (A) Host infectious period $\tau_I = 1/\gamma$. (B) Host latent period $\tau_E = 1/\kappa$. (C) Vector removal rate μ . (D) Vector infection rate α . (E) Host infection rate β . (F) The variance of the field data σ^2 .

The latter distributions are far from a Gaussian-like shape (note that the x-axis is log-scaled), being heavy-tailed. This kind of distribution highly distorts the statistical measures of mean, median and standard error, indicating that the estimates for transmission rates are not as robust as the estimates for the other parameters. These rather uninformative distributions are most probably arising because of the lack of data about the vector, i.e. $S_v(t)$ and $I_v(t)$, to constrain the fits. In essence, many combinations of α and β can similarly fit the host data while yielding quite different time series for $S_v(t)$ and $I_v(t)$, which cannot be contrasted due to the lack of field data. Nevertheless, the obtained best-fit mean and median parameters, although quite different, are able to perfectly fit the data (Fig. 7.4). Finally, we also observe that the variance for the field data also converged to a bell-shaped distribution.

Mean and median parameter estimates, i.e. the best-fit parameter values for ALSD and OQDS, are summarized in Tables 7.1 and 7.2, respectively. As already seen from the posterior distributions, the best-fit values for τ_E , τ_I and μ are close to the ones given by literature and field data for both diseases. Conversely, α and β are rather uninformative, as their 95% confidence intervals cover almost two orders of magnitude. This again indicates that without some data about the evolution of the vector states in time, $S_v(t)$ and $I_v(t)$, it is nearly impossible to derive the proper values for these parameters.

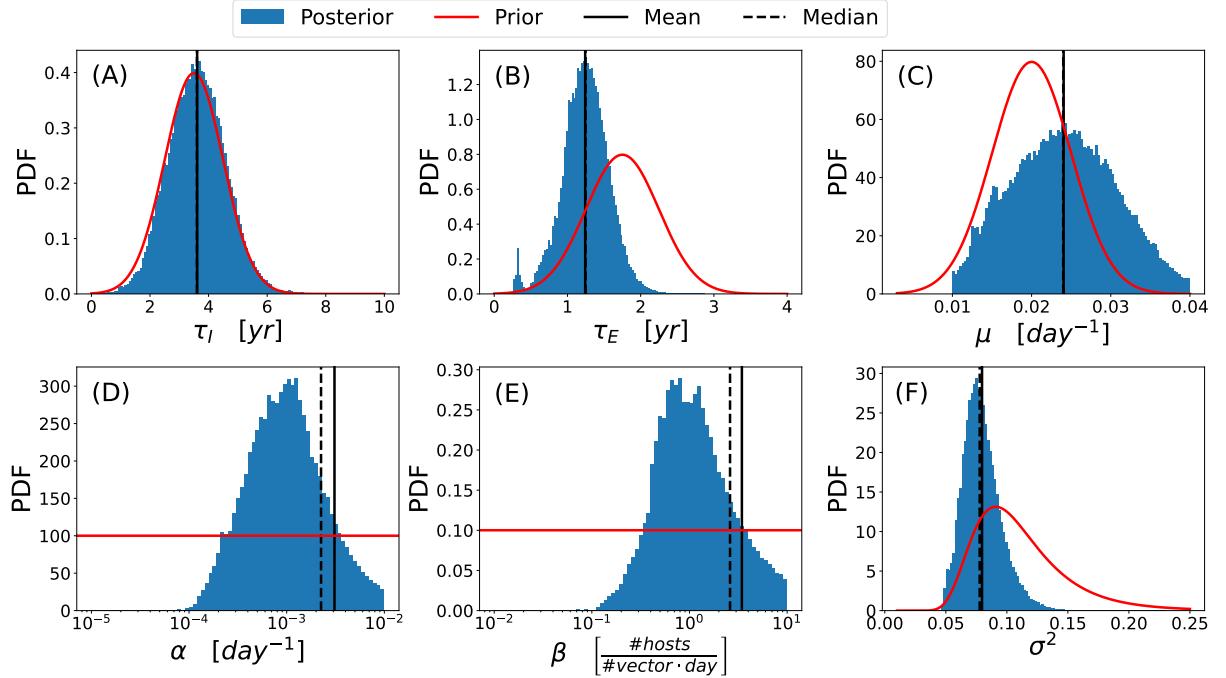


Figure 7.3: Posterior (blue histograms) and prior (red line) distributions of the model parameters for OQDS. Solid and dashed black lines correspond to the mean and median of the posterior distributions. (A) Host infectious period $\tau_I = 1/\gamma$. (B) Host latent period $\tau_E = 1/\kappa$. (C) Vector removal rate μ . (D) Vector infection rate α . (E) Host infection rate β . (F) Variance of the field data σ^2 .

Table 7.1: Estimated epidemiological parameters from Bayesian model fitting to the disease progression curve of ALSD in Mallorca.

Parameter	Definition	Units	Posterior Mean	Posterior Median	95% C.I.
τ_I	Host infectious period	yr	13.84	13.82	[7.12, 20.47]
τ_E	Host latent period	yr	4.46	4.47	[2.88, 5.99]
β	Host infection rate	#host/#vector-day	0.062	0.02	[0.0061, 0.3013]
α	Vector infection rate	day ⁻¹	0.15	0.086	[0.0047, 0.54]
μ	Vector removal rate	day ⁻¹	0.0222	0.0221	[0.015, 0.030]
R_0	Basic reproductive number	-	133	25	-

Table 7.2: Estimated epidemiological parameters from Bayesian model fitting to the disease progression curve of OQDS in Apulia.

Parameter	Definition	Units	Posterior Mean	Posterior Median	95% C.I.
τ_I	Host infectious period	yr	3.61	3.60	[2.06, 5.20]
τ_E	Host latent period	yr	1.24	1.25	[0.70, 1.75]
β	Host infection rate	#host/#vector-day	3.44	2.60	[0.55, 8.79]
α	Vector infection rate	day ⁻¹	0.0031	0.0022	[0.0005, 0.0084]
μ	Vector removal rate	day ⁻¹	0.0240	0.0240	[0.014, 0.035]
R_0	Basic reproductive number	-	33	21	-

Overall, the data falls within the 99% confidence limits of the fitted model for both the ALSD and OQDS outbreaks (Fig. 7.4(B,D)). We also computed the instantaneous reproductive number, $R_0(t)$, by using Eq. (7.3) with $S_H(t)$ instead of only $S_H(0)$ along the simulation. Noteworthy, $R_0(t) = 1$ coincides with the stopping of new infections being produced, i.e. the number of exposed hosts does not increase (Fig. 7.4(A,C)). This supports our approximate method for computing the

reproductive number for Xf diseases (Section 7.5.2, Eq. (7.3)). Due to the different time scales of both epidemics ($\tau_I^{ALSD} + \tau_E^{ALSD} > \tau_I^{OQDS} + \tau_E^{OQDS}$), the OQDS outbreak dies out earlier than the one for ALSD.

We notice that for ALSD a large proportion of the vector population gets infected every year (Fig. 7.4(A)), while a very small proportion is needed in OQDS to produce a lethal outbreak (Fig. 7.4(C)). However, this last statement is rather unrealistic, as around 50% of the vectors that are captured in Apulia are indeed infected by Xf ((39, 47)). Thus, the evolution of the infected vector population should be qualitatively similar to that obtained for ALSD (Fig. 7.4(C)). As previously explained, different suitable combinations of α and β parameters should give rise to similar progression curves for the hosts while different ones for the vectors, but the realistic values for these parameters cannot be obtained from the Bayesian fit due to the lack of data of the vector states, $S_v(t)$, $I_v(t)$.

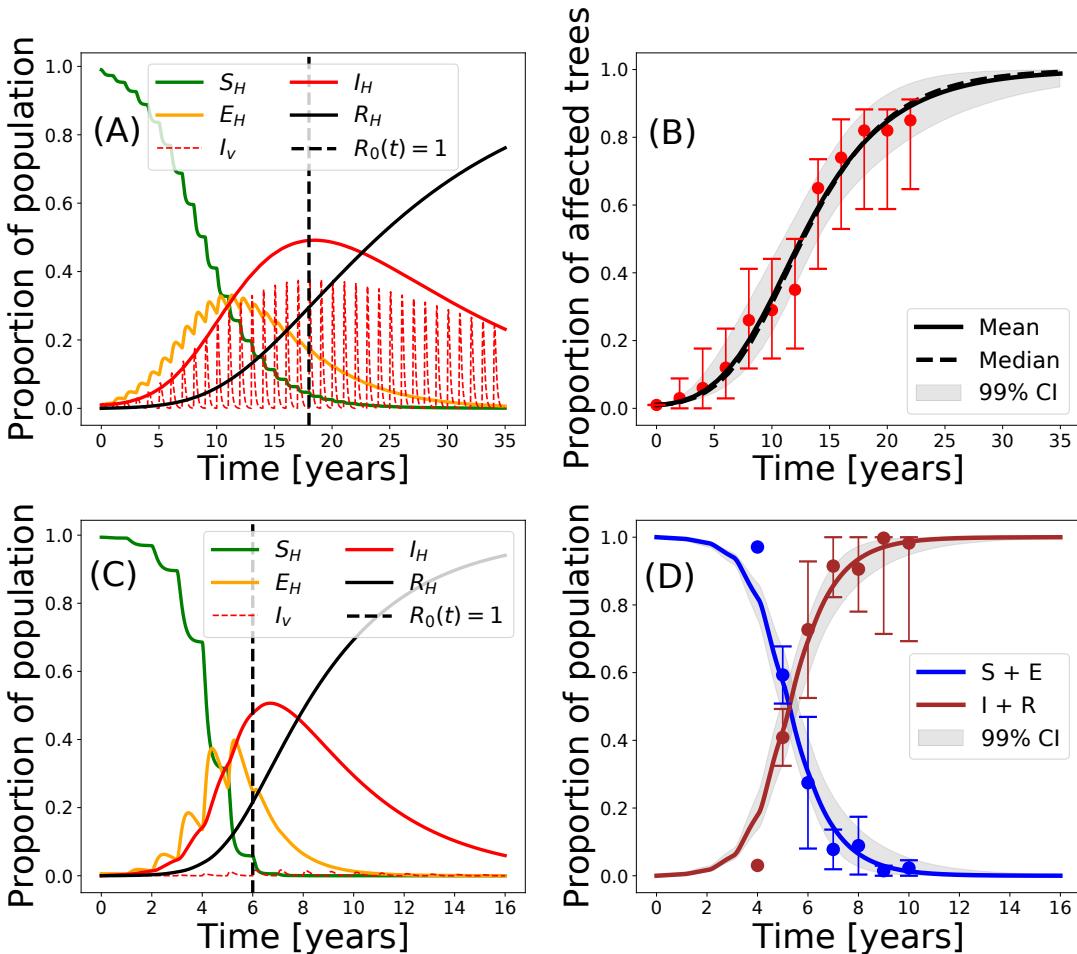


Figure 7.4: (A) Simulation of the model with the best-fit parameters for ALSD. (B) Model fit to field data by means of the mean and median values of the posterior distributions of the parameters for ALSD. (C) Simulation of the model with the best-fit parameters for OQDS. (D) Model fit to field data by means of the mean and median values of the posterior distributions of the parameters for OQDS. The gray-shaded area corresponds to the 99% confidence interval. The error bars for the field data correspond to their 95% confidence interval obtained with a bootstrapping technique.

Nevertheless, by manually exploring other values for α and β parameters, we can obtain a more biologically plausible scenario for the OQDS that is still able to fit the available data for the hosts. Fig. 7.5(A) shows a simulation of the model with previously inferred best-fit median parameters for OQDS. By changing the values of α and β , we obtain a more realistic scenario, i.e. around a 50% of the vector population getting infected during the outbreak (Fig. 7.5(B)) ((39, 47)). Noteworthy, the β value obtained in this way is almost identical to the transmission

rate recently reported by (20) for OQDS. This change in the transmission parameters only affects the progression curve of the infected vector population, being the progression of the host compartments practically unchanged (Fig. 7.5(C)). Anyway, both sets of parameter values for α and β can properly fit the field data, corresponding exclusively to the host population (Fig. 7.5(D)).

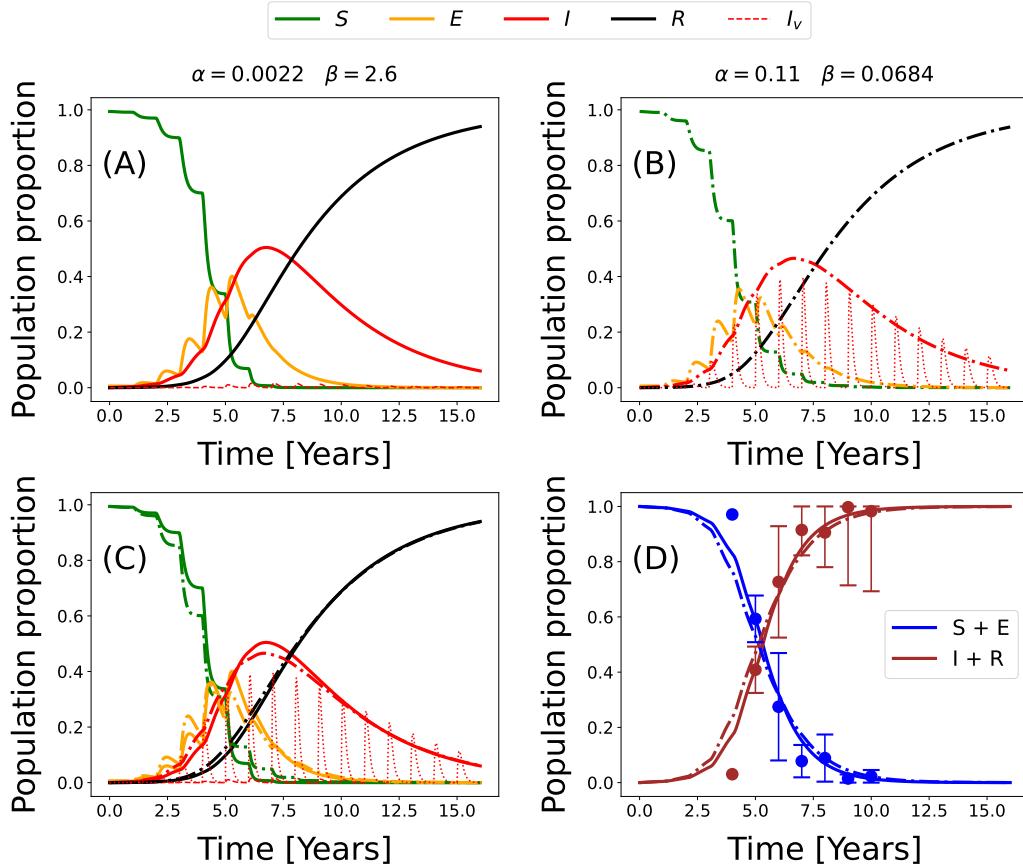


Figure 7.5: (A) Simulation of the model with the original best-fit parameters for OQDS. (B) Simulation of the model with the original best-fit parameters for OQDS but with different α , β values. (C) Comparison of the progression curves. Note that the curves for the hosts are very similar while the curve for the infected vector population is very different. (D) Comparison of the model fit to the data with both simulations. Solid lines correspond to results with the original best-fit parameters while dash-dot lines correspond to the results of the more realistic scenario with different α and β .

The model adjusted to the progression curves of both diseases indicates that the transmission rate α must be greater than β when the proportion of infected vectors is relatively high ($> 30\%$). We checked if the relation between α and β held when changing the assumed $N_v(0) = N_H/2$, obtaining that it kept approximately the same for very different values of the initial vector population.

7.3.2 Global Sensitivity Analysis

We computed the sensitivity indices for the model parameters with respect to the more relevant quantities of interest, namely, the time at which the number of infectious hosts is maximal, t_{peak} , the maximum number of infectious hosts, I_{peak} and the final number of dead hosts, R_∞ . The results were obtained exploring the parameter space constrained to the intervals $\{\beta \in (0.001, 0.1), \tau_E \in (3, 7), \tau_I \in (5, 25), \alpha \in (0.001, 1), \mu \in (0.01, 0.04)\}$ using 10^4 Quasi-Monte Carlo samples and are summarized in Fig. 7.6.

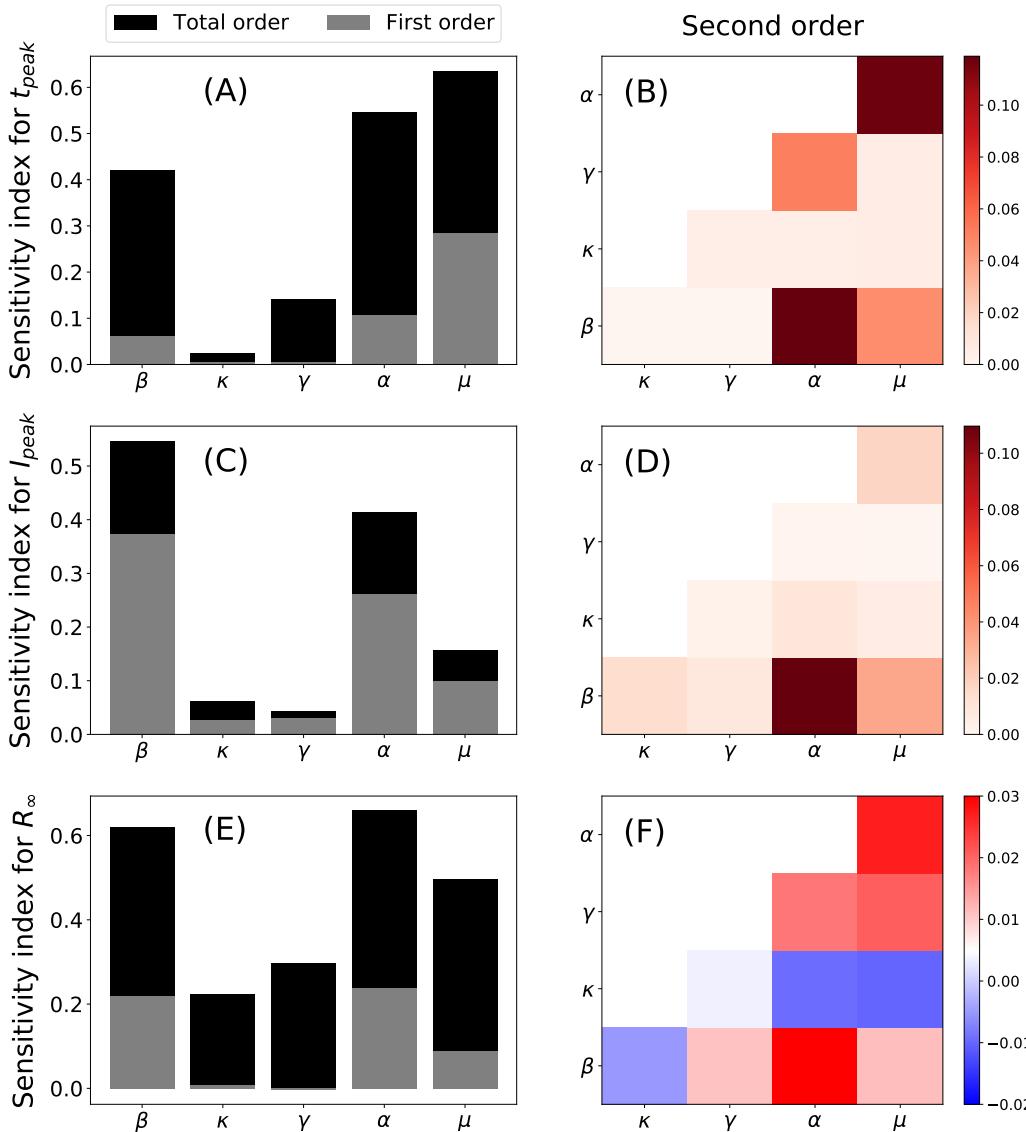


Figure 7.6: Global Sensitivity Analysis of the model parameters performed with the Sobol method with respect to the time at which the infectious population peaks, t_{peak} (A-B), the magnitude of this peak, I_{peak} (C-D) and the final number of dead hosts, R_∞ (E-F). The left column (A,C,E) shows the total and first-order indices and the right column (B,D,F) shows the second-order indices.

Parameters α , β and μ are the most influential with regard to the time at which the infectious host population peaks, t_{peak} , the magnitude of the peak, I_{peak} , and the final number of dead hosts, R_∞ . The total output variance (total order indices) cannot be explained by the variances of the parameters alone (first order indices) (Fig. 7.6). Therefore, higher-order interactions among the parameters importantly affect the sensitivity of the quantities under study. Indeed, the contribution to the total output variance of γ and κ for t_{peak} and R_∞ come notably from higher-order interactions. This can be checked in panels (B,C,F) of Fig. 7.6, in which the contribution to the output variance from interactions between pairs of parameters (second order indices) is represented. Interactions among the parameters contribute to increasing the output variance

with respect to t_{peak} and I_{peak} , while the effect is more heterogeneous in the case of R_∞ . In particular, the interactions between $\alpha - \beta$ and $\alpha - \mu$ produce the main contributions to the increase of output variance in all cases, while $\kappa - \beta$, $\kappa - \alpha$ and $\kappa - \mu$ interactions decrease the output variance.

7.3.3 Epidemic control through vector management

The sensitivity analysis clearly indicates that acting on α, β and μ is the best strategy to lower disease incidence and mortality. However, controlling transmission rates is cumbersome so a different control strategy based only on vector control is considered in this section. In our model, there are two ways of implementing vector-population control: (i) decreasing the typical time, $1/\mu$, that vectors spend between crops each year by some mechanism (thus increasing μ) and (ii) reducing the initial number of vectors that invade crops each year (e.g. lowering $N_v(0)$ via egg or nymph control ((106)).

We analyzed the effect of vector management by simulating epidemic outbreaks using different values of μ and $N_v(0)$, and keeping the rest of parameters as fitted for both ALSD and OQDS outbreaks (Fig. 7.7). In both epidemics, decreasing the presence time as well as the number of vectors contribute to controlling the epidemic by lowering R_0 and, consequently, the final size of the epidemic, R_∞ . Furthermore, we observe that decreasing vector presence is more efficient than decreasing its annual initial population, i.e. we further reduce R_∞ , the final size of the epidemic, by applying a similar reduction in the residence time $1/\mu$. This could also be anticipated as R_0 depends quadratically on $1/\mu$ while only linearly on $N_v(0)$ (Eq. (7.3)). However, the minimal intervention strategy, starting from the current situation in the $(1/\tau, N_v(0))$ parameter space that yields an absolute control of the epidemic, $R_0 < 1$, involves a mixed strategy of lowering both $1/\mu$ and $N_v(0)$.

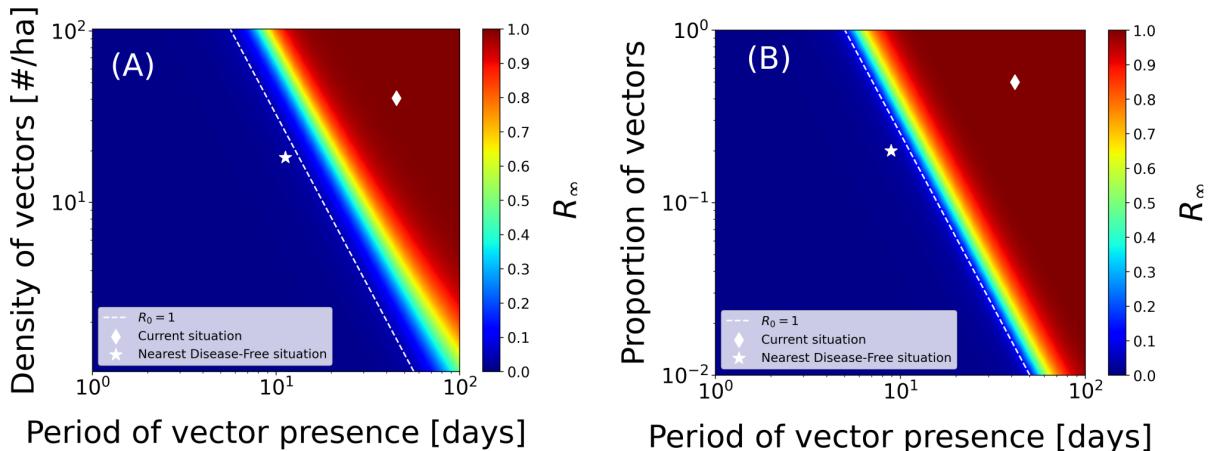


Figure 7.7: Epidemic control through vector management for ALSD in Mallorca (A) and OQDS in Apulia (B). The white shaded line denotes $R_0 = 1$, the white diamond corresponds to the parameter values of the fitted model. The white star is the closest disease-free state to the current situation in this representation.

7.4 Discussion

In this work, we have developed a deterministic continuous-time compartmental model for *Xylella fastidiosa* vector-borne diseases in Europe. The model attempts to characterize the main biotic processes that lead to the development of epidemics, including the seasonal dynamics of the main vector, *P. spumarius*. We show how the model is sufficiently general to represent with some accuracy the parameters that determine the ALSD in Mallorca (Spain) and the OQDS in Apulia (Italy), both transmitted by *P. spumarius*. To our best knowledge, this is the first mathematical model describing Xf epidemics that considers the temporal pattern of vector abundance observed in field data, faithfully representing the known biological information about the pathosystem. It includes a dynamic approximation of the non-stationary populations of *P. spumarius*, mathematically represented by a sporadic source term through which vectors

are born every year, and an exponential decay term. Due to the non-stationarity of the vector dynamics, R_0 in the model cannot be computed with standard methods such as the Next Generation Matrix ((57)). To circumvent this problem, we applied an approximate method to compute it as previously proposed by ((76)). We show that this approximate R_0 correctly characterizes the epidemic, further validating the method proposed by ((76)).

Nonlinear mathematical models of disease transmission enhance our understanding of the different mechanisms operating in an epidemic, especially compared with correlative or machine learning methods, often very useful in practice but offering very little understanding. A key aspect to render these models useful is the determination of the parameters from available data. If this step can be properly performed, these models become very predictive and especially helpful to design disease control strategies. However, an appropriate calibration of the model relies on access to good-quality field data, which is often the bottleneck for the application of this kind of models. In the present study, the parameters have been obtained using a Bayesian inference framework, which relies on probability distributions rather than point-like measures. This way, mean or median values can be considered together with their confidence intervals able to characterize the robustness of the obtained parameters. In general, we obtained different values of the parameters for the ALSD and OQDS outbreaks in Mallorca and Apulia, respectively. The fitted values, however, are in good agreement with previous field-based measures for each disease while the differences observed between both outbreaks may reflect differences between the Xf subspecies and crops involved (deciduous vs. evergreen).

One of the conclusions of the study is that the available data for both diseases is not enough to obtain robust estimates for all of the model parameters. The lack of data about the vector population compartments yields many possible values for the parameters that regulate transmission, α and β , provided that the progression of the host compartments correctly fits the field data. In other words, very infectious vectors (high β) that hardly ever get infected (low α) can produce a similar outbreak within the host population to that produced by very low infectious vectors (low β) that get infected very often (high α). The great difference in these situations would be that, in the former, the infected vector population would be very low, while in the latter, it would be quite high. This is a manifestation of parameter unidentifiability from the fit ((43, 141)), which stresses the importance of transmission and calls for detailed measurements of the vector population, and not just of the hosts. Furthermore, to compare transmission rates between different diseases caused by Xf (e.g. β , α), it is necessary to know the vector-host population ratio of the pathosystem (N_v/N_H), since β is expressed as a number of hosts per vectors per day. Although, in general, populations of *P. spumarius* in the canopy of olive trees are much larger than those found in the almond trees of the Balearic Islands during the months of July and August ((113)), our work is based on data from studies in which information of the vector populations is not provided. Without this information, therefore, conclusive results regarding transmission cannot be obtained.

In any case, our model shows that the vector-to-plant transmission process, mediated by β , is somehow different from that from the plant-to-vector one, mediated by α . In essence, β must be smaller than α in order to reproduce the observed outbreaks and have a sufficiently large vector population getting infectious, being this fact independent of the particular choice of $N_v(0)/N_H$. This heterogeneity can be caused by several factors: differences in the efficiency of plant-to-vector transmission with respect to vector-to-plant transmission, differences in contact rates, i.e. susceptible vectors contact trees at a different rate than infected vectors; vector feeding preferences, i.e. differences in the probability of contacting a susceptible host compared to an infectious host, etc. Indeed, our mathematical model assumes constant contact rates with no preferences over any host state, so that under these assumptions, it indicates that the probability of effectively transmitting the pathogen from plant to vector is greater than from vector to plant. However, this interpretation is subject to this particular assumption, so that to fully disentangle this question experimental work in form of transmission assays should be performed. Furthermore, we found that the timing and magnitude of the infectious host peak and the final number of dead hosts are mostly controlled by the vector-to-plant transmission rate, β , the plant-to-vector transmission rate, α and the vector removal rate μ . Because these parameters are strongly related to the vector, the analysis makes clear that enhancing the knowledge about the vector, as well as obtaining precise data, is crucial to improve the modeling of Xf diseases and pose important questions to be solved in specifically designed experiments.

The fact that the most influential parameters of the model are those related to the vector can be used to design appropriate disease control strategies. Because acting on transmission rates is rather cumbersome, we argue that control strategies should focus on reducing the vector population in crop fields. In our model, this depends on two parameters, μ , the rate at which vectors die (or move to herbaceous vegetation and other non-host trees or exit the field) and $N_v(0)$, the number of newborns susceptible vectors every year (assumed constant in this study). Our results show that a mixed strategy acting on both parameters is optimal to lower disease prevalence and, eventually, eradicate the disease. Interestingly, we also show that acting on the vector removal μ is more effective than controlling the newborn vector population $N_v(0)$. In fact, most control strategies carried out in practice for Xf diseases focus on the latter factor, reducing $N_v(0)$ via egg or nymph control ((48, 106, 112)). However, our results indicate that alternative strategies based on increasing the removal (or dispersal) rate of vectors should be explored. Furthermore, the evolution of the population compartments of the hosts and vectors provides relevant information on the epidemiology of both diseases. In both cases, the newly defined basic reproductive number that accounts for a decaying vector population is very predictive of the moment in which new infections are not produced anymore, coinciding approximately with the peak of infectious hosts. Therefore, any intervention with control measures after this peak would have marginal effects on future disease progression.

Our mathematical model is still rather simple, implementing only a few relevant epidemic processes in contrast to the high complexity of the pathogen-vector-host interactions occurring in plant epidemics. Indeed, the model itself raises some questions about these interactions, for example, whether or not contact rates are homogeneous. Another simplification of the model is the fact that the spatial constraints and the intrinsic stochasticity of the transmission processes are neglected. A straightforward extension of the model would be to include a specific spatial setting and implement the explicit motion of the vector within a stochastic framework, such as Individual Based Models ((82)). With this, the effectiveness of current and further control strategies could be tested and improved controlling for the motion of the vector. For instance, the control strategy based on the removal of symptomatic trees together with their surrounding trees at a given distance could be implemented in the model, evaluate the current effectiveness according to the present protocols and even provide improved parameters to be implemented in the field. Of course, implementing a model in which the spatial degrees of freedom are explicitly represented would require access to further information about vector mobility and spatially resolved data to confront the model, which is not currently available.

Mathematical models tested against experimental data increase our understanding of the system under study. They also help to identify critical parameters that require better prior information to adjust functions relating to different variables and make the model predictions more accurate to suggest and test control strategies ((52, 93)). Our mathematical model suggests a certain lack of knowledge of the transmission processes and reveals that the currently available data is not enough to fit complex models dealing with the explicit dynamics of the vector population.

7.5 Appendix

7.5.1 Vector population dynamics

Fig. 7.8 shows a time series for the population of *Philaenus spumarius* in Mallorca, taken from ((113)) (in blue). Superimposed (in orange) is the assumption used in our model Eq. (7.2), the $\delta(t - nT)$, i.e. every year susceptible vectors appear in the system.

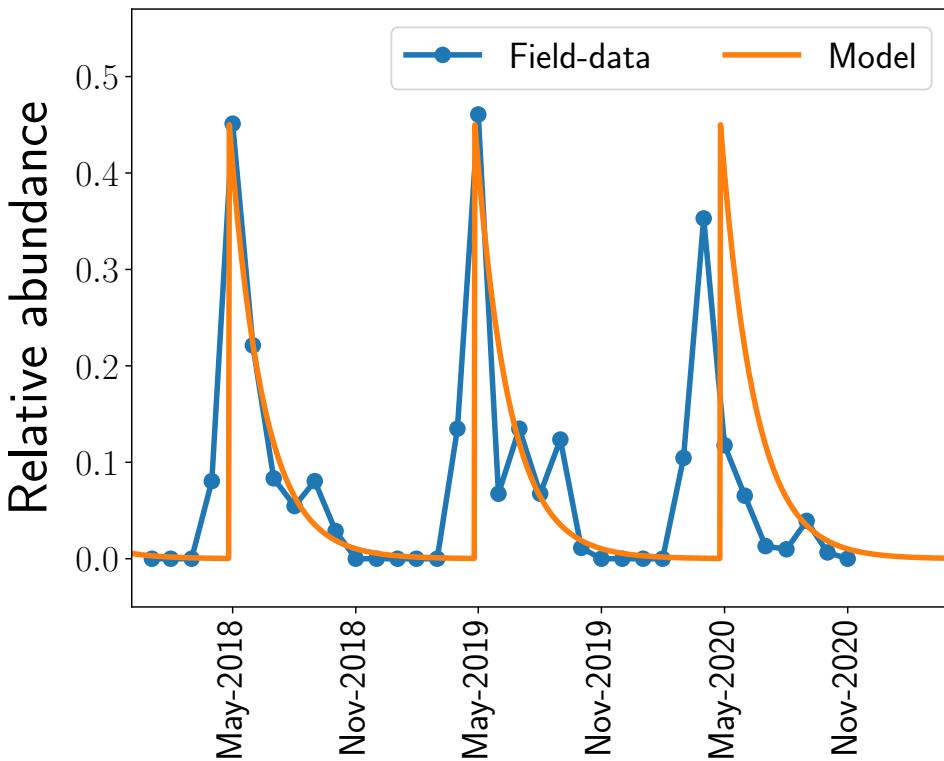


Figure 7.8: Vector dynamics produced by the model compared to field-data from ([113]).

7.5.2 Determination of R_0

The handicap of determining the basic reproductive number of the model Eq. (7.2) is that the pre-pandemic fixed point given by $I_H = I_v = 0$ and $S_H = S_H(0)$ is not a fixed point of the system of differential equations, because vector population decays, so that the standard methods to compute R_0 such as the Next Generation Matrix ((57, 76)) do not apply. In ((76)) a method was suggested to determine the basic reproductive number in the case of compartmental models of vector-borne transmitted diseases in which the vector population grows or decays. It consists in averaging the instantaneous basic reproductive number over the time of a generation.

To proceed we consider that $I_H = I_v = 0$, $S_H = S_H(0)$ is indeed a fixed point of the system. Then, the basic reproductive number could be determined, e.g. as shown in ((25)). First, an infectious host infects vectors at a rate $\beta S_H(0)/N_H$ for a time $1/\gamma$. This produces $\beta S_H(0)/\gamma N_H$ infected vectors. The second stage is that these infectious vectors infect hosts at a rate $\alpha N_v(0)/N_H$ for a time $1/\mu$, producing $\alpha N_v/\mu N_H$ infectious hosts per vector. The net result of these two stages is

$$\tilde{R}_0 = \frac{\alpha \beta}{\mu \gamma} \frac{S_H(0)}{N_H^2} N_v(0) = R_0^* \cdot N_v(0) . \quad (7.4)$$

This result coincides with the value of R_0 obtained using the standard NGM method, that can be applied in this case because we are assuming that we use a nongeneric initial condition that sits at the fixed point of the model.

In practice, our initial condition will never be a fixed point of the model, and, as mentioned above, we will obtain an approximate basic reproductive number, to which we will refer as R_0 using the method suggested in ((76)), that consists in calculating the average number of secondary infections produced by an infectious host in one generation. One first defines an instantaneous basic reproductive number,

$$R_0^{(i)}(t) = \frac{\beta \alpha}{\mu \gamma} \frac{S_H(0)}{N_H^2} N_v(t) = R_0^* N_v(t) , \quad (7.5)$$

from which the average is simply computed as

$$R_0 = \left\langle R_0^{(i)}(t) \right\rangle \Big|_0^\tau = R_0^* \langle N_v(t) \rangle \Big|_0^\tau = R_0^* \frac{1}{\tau} \int_0^\tau N_v(t) dt . \quad (7.6)$$

In our model, the time-dependent vector population can be obtained from Eq. (7.2),

$$\dot{N}_v = \dot{S}_v + \dot{I}_v = -\mu N_v \implies N_v(t) = N_v(0)e^{-\mu t} , \quad (7.7)$$

and introducing this expression for $N_v(t)$ in Eq. (7.6) the integral can be solved

$$R_0 = \frac{\beta \alpha S_H(0)}{\mu \gamma N_H^2} \frac{N_v(0)}{\mu \tau} (1 - e^{-\mu \tau}) = R_0^* \frac{N_v(0)}{\mu \tau} (1 - e^{-\mu \tau}) , \quad (7.8)$$

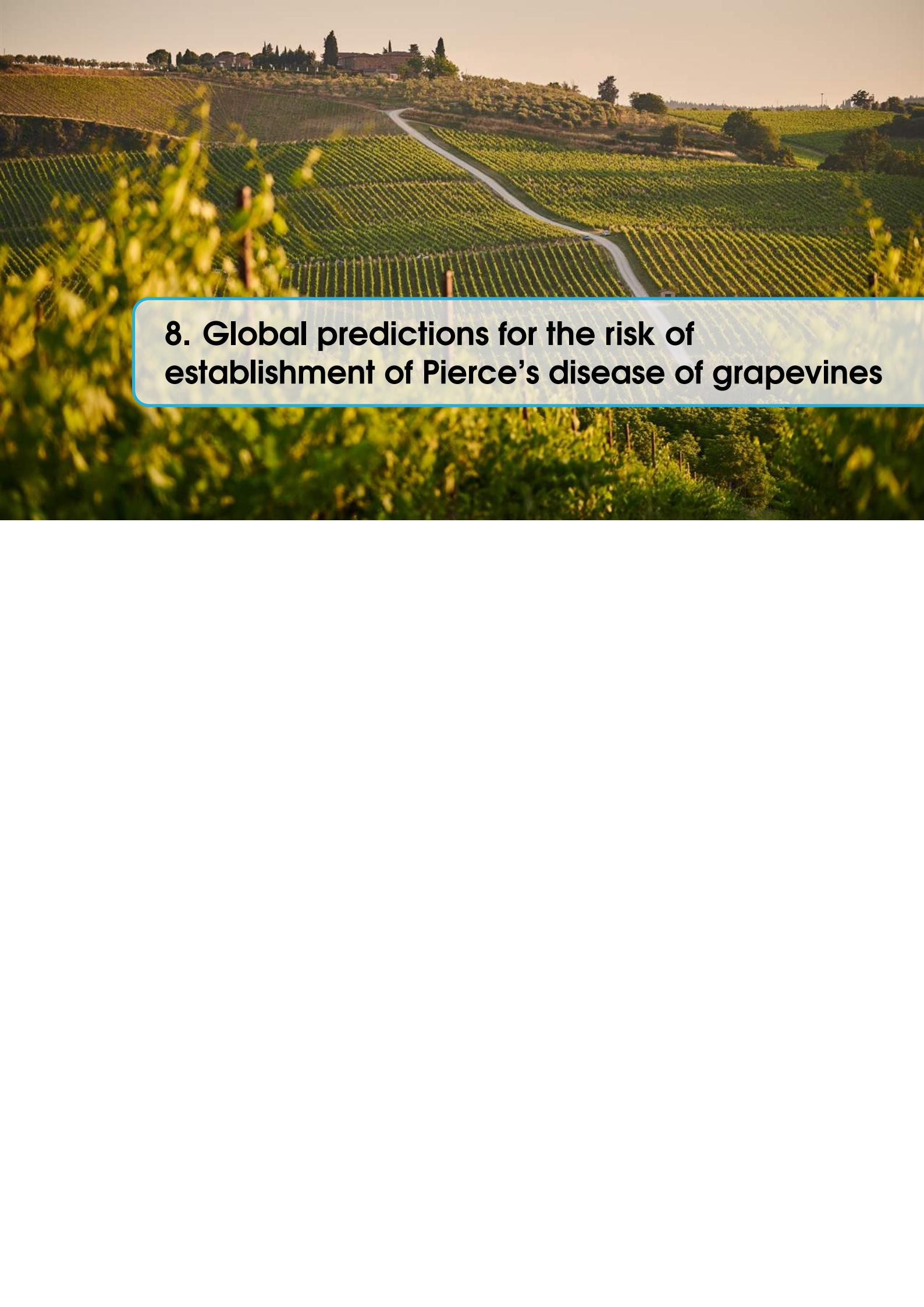
that is an approximated expression to the basic reproductive number for our model, in which the vector population is nonstationary, where, in Eq. (7.5) and Eq. (7.8) it has been defined, $R_0^* = (\beta \alpha S_H(0)) / (\mu \gamma N_H^2)$.

Note that in our model one generation correspond to one year and that $N_v(0)$ is reset every year.



Modelling the risk of vector-borne plant diseases

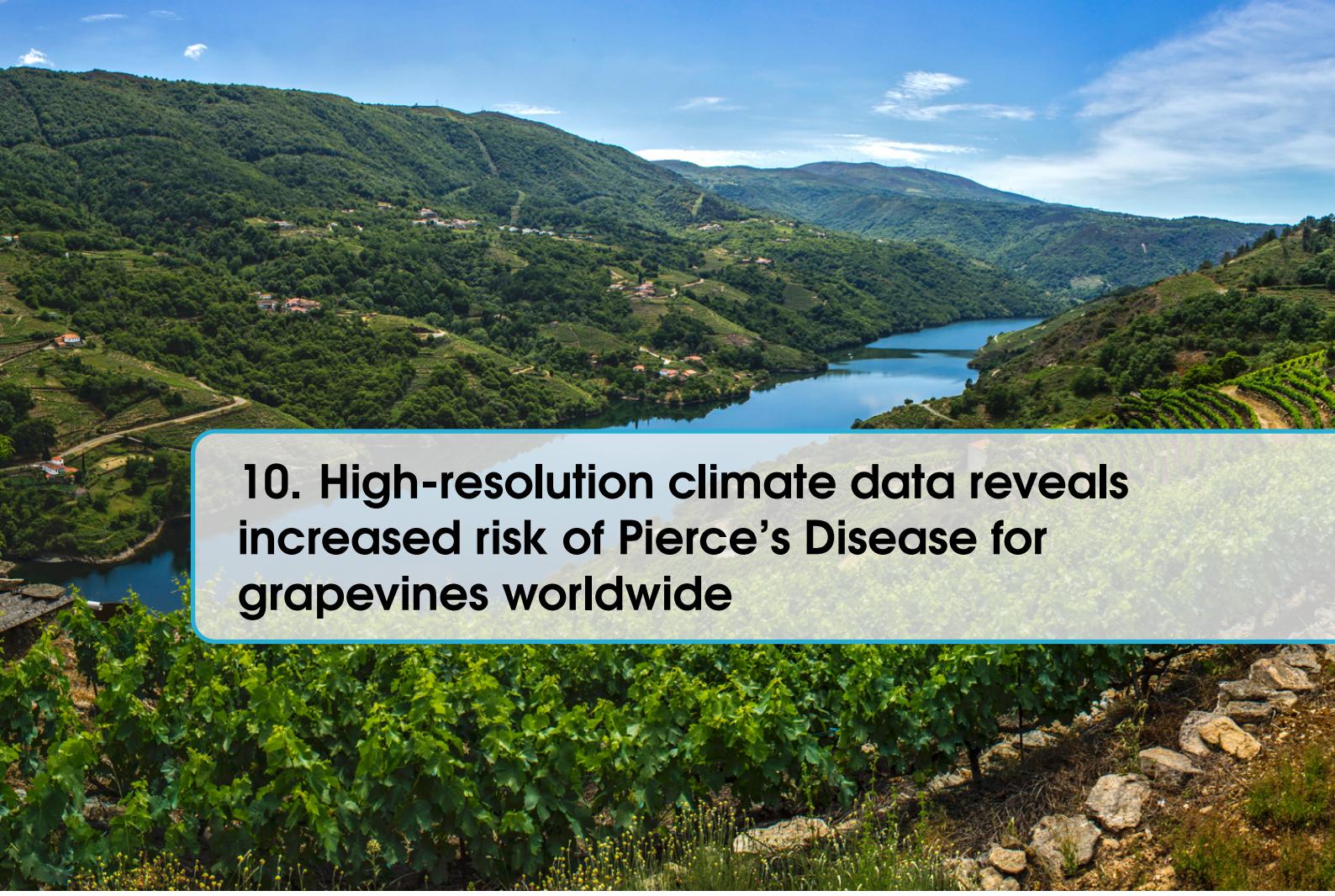
- 8 Global predictions for the risk of establishment of Pierce's disease of grapevines 77
- 9 Global warming significantly increases the risk of Pierce's disease epidemics in European vineyards 79
- 10 High-resolution climate data reveals increased risk of Pierce's Disease for grapevines worldwide .. 81



8. Global predictions for the risk of establishment of Pierce's disease of grapevines



9. Global warming significantly increases the risk of Pierce's disease epidemics in European vineyards

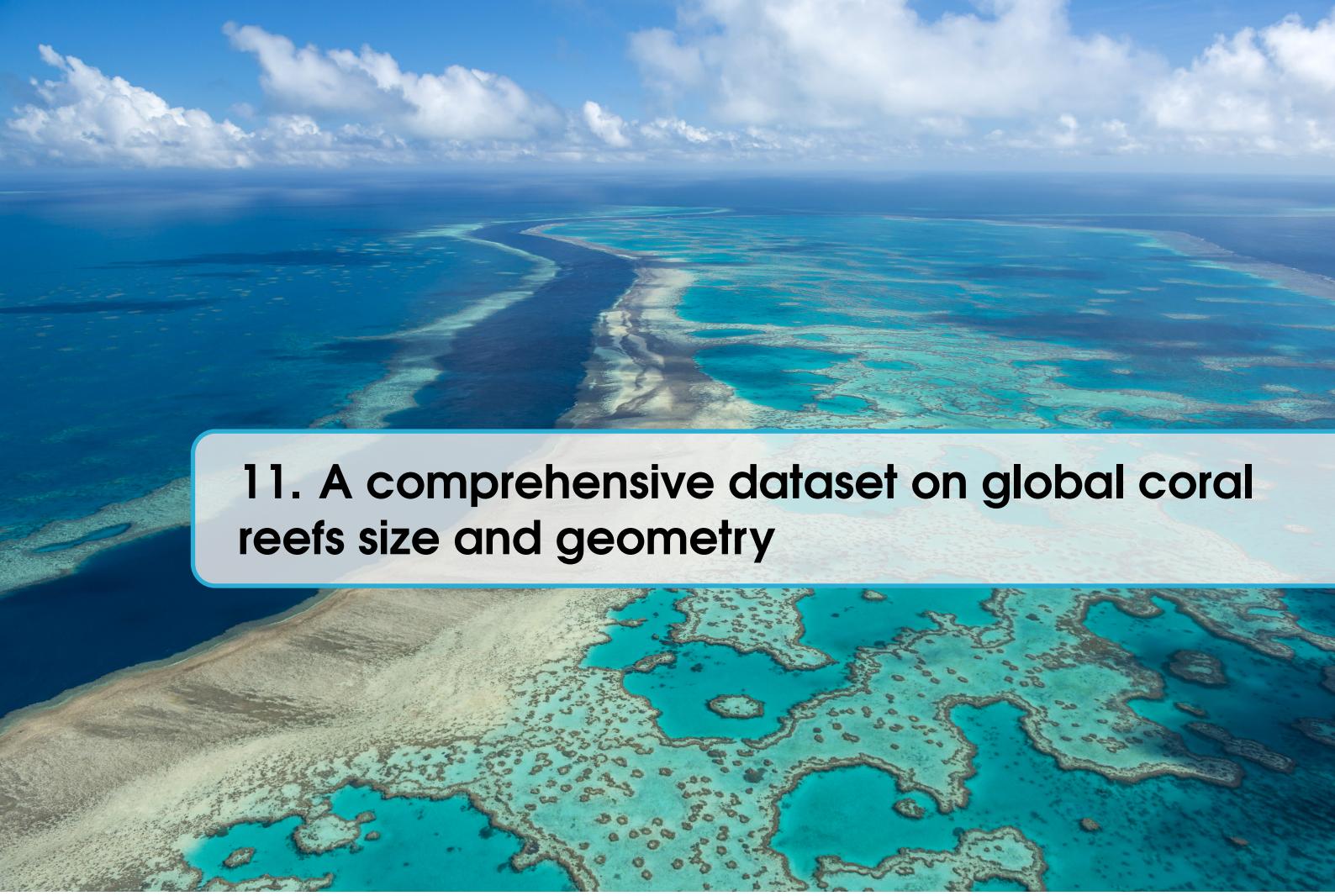


10. High-resolution climate data reveals increased risk of Pierce's Disease for grapevines worldwide

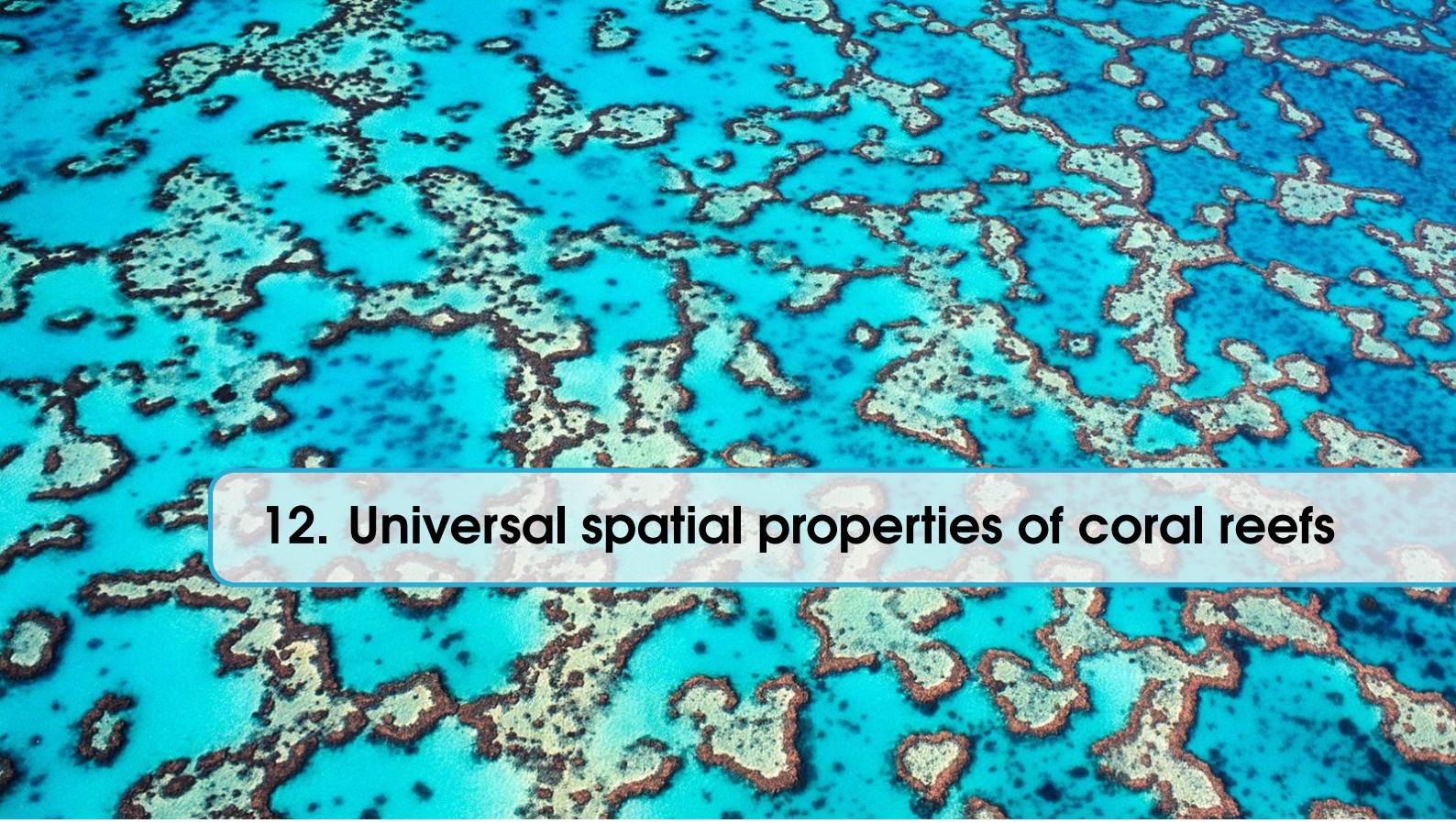


Data-driven methods for global ecological problems

- | | | |
|----|---|----|
| 11 | A comprehensive dataset on global coral reefs size and geometry | 85 |
| 12 | Universal spatial properties of coral reefs | 87 |
| 13 | pH trends and seasonal cycle in the coastal Balearic Sea reconstructed through machine learning | 89 |
| 14 | Mapping the distribution of seagrass meadows from space with deep convolutional neural networks | 91 |



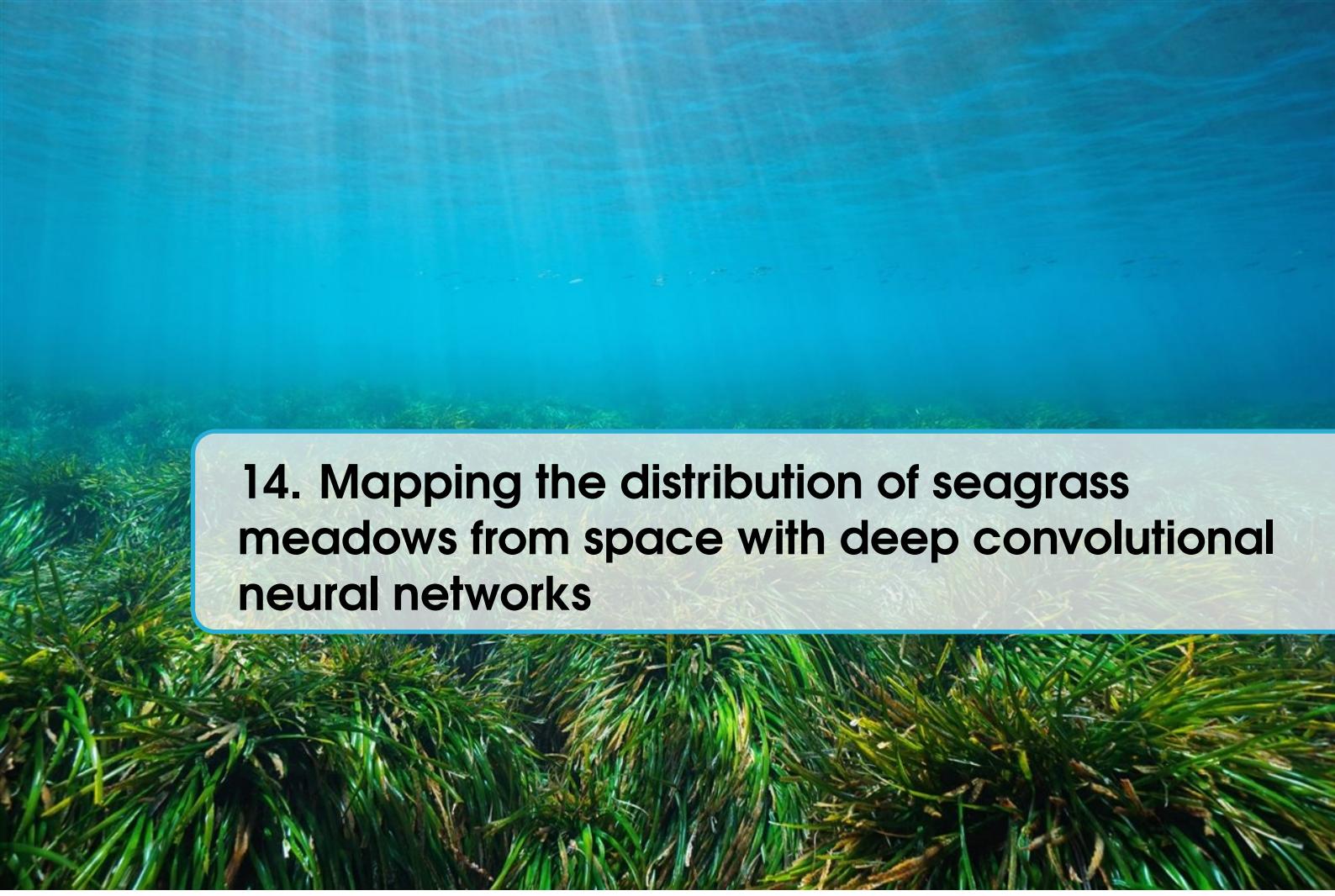
11. A comprehensive dataset on global coral reefs size and geometry



12. Universal spatial properties of coral reefs



**13. pH trends and seasonal cycle in the coastal
Balearic Sea reconstructed through machine
learning**

A photograph of a seagrass meadow underwater. The dense, green, blade-like leaves of the seagrass grow in patches across the sandy ocean floor. The water is clear, allowing light to penetrate and illuminate the plants. The background shows more of the seagrass extending towards the horizon.

14. Mapping the distribution of seagrass meadows from space with deep convolutional neural networks

Bibliography

Bibliography

- [1] C. Abboud et al. “Dating and localizing an invasion from post-introduction data and a coupled reaction-diffusion-absorption model”. In: *Journal of Mathematical Biology* 79.2 (July 2019), pages 765–789. ISSN: 1432-1416. DOI: [10.1007/s00285-019-01376-x](https://doi.org/10.1007/s00285-019-01376-x). URL: <https://doi.org/10.1007/s00285-019-01376-x> (cited on page 60).
- [2] A. Alonso Aguirre and Gary M. Tabor. “Global Factors Driving Emerging Infectious Diseases”. In: *Annals of the New York Academy of Sciences* 1149.1 (2008), pages 1–3. DOI: <https://doi.org/10.1196/annals.1428.052> (cited on page 29).
- [3] R. P. P. Almeida and L. Nunney. “How do plant diseases caused by *Xylella fastidiosa* emerge?” In: *Plant Disease* 99 (11 1987), pages 1457–1467. DOI: [10.1094/PDIS-02-15-0159-FE](https://doi.org/10.1094/PDIS-02-15-0159-FE) (cited on page 60).
- [4] R. P. P. Almeida and A. H. Purcell. “Biological Traits of *Xylella fastidiosa* Strains from Grapes and Almonds”. In: *Applied and Environmental Microbiology* 69.12 (2003), pages 7447–7452. DOI: [10.1128/AEM.69.12.7447-7452.2003](https://doi.org/10.1128/AEM.69.12.7447-7452.2003). eprint: <https://journals.asm.org/doi/pdf/10.1128/AEM.69.12.7447-7452.2003>. URL: <https://journals.asm.org/doi/abs/10.1128/AEM.69.12.7447-7452.2003> (cited on page 60).
- [5] Roy M. Anderson. “Discussion: The Kermack-McKendrick epidemic threshold theorem”. In: *Bulletin of Mathematical Biology* 53 (1991), pages 3–32. URL: <https://www.sciencedirect.com/science/article/pii/S0092824005800394> (cited on pages 7, 29, 45).
- [6] J. Andrews. “Epizootiology of diseases of oysters (*Crassostrea virginica*), and parasites of associated organisms in eastern North America”. In: *Helgolander Meeresuntersuchungen* 37 (Jan. 1984), pages 149–166. URL: <https://scholarworks.wm.edu/vimsarticles/1447> (cited on page 19).
- [7] S. Antonatos et al. “Seasonal Appearance, Abundance, and Host Preference of *Philaenus spumarius* and *Neophilaenus campestris* (Hemiptera: Aphrophoridae) in Olive Groves in Greece”. In: *Environmental Entomology* 50.6 (Sept. 2021), pages 1474–1482. ISSN: 0046-225X. DOI: [10.1093/ee/nvab093](https://doi.org/10.1093/ee/nvab093). eprint: <https://academic.oup.com/ee/article-pdf/50/6/1474/41745778/nvab093.pdf>. URL: <https://doi.org/10.1093/ee/nvab093> (cited on page 62).
- [8] Isabelle Arzul and Ryan B. Carnegie. “New perspective on the haplosporidian parasites of molluscs”. In: *Journal of Invertebrate Pathology* 131 (2015), pages 32–42. DOI: [10.1016/j.jip.2015.07.014](https://doi.org/10.1016/j.jip.2015.07.014) (cited on pages 8, 23).
- [9] Tejas Athni et al. *How vector-borne disease shaped the course of human history*. June 2020. DOI: [10.22541/au.159135318.82566392](https://doi.org/10.22541/au.159135318.82566392) (cited on page 45).
- [10] Corinne Audemard et al. “*Bonamia exitiosa* transmission among, and incidence in, Asian oyster *Crassostrea ariakensis* under warm euhaline conditions”. In: *Diseases of Aquatic Organisms* 110 (July 2014), pages 143–50. DOI: [10.3354/dao02648](https://doi.org/10.3354/dao02648) (cited on page 19).
- [11] Nicolas Bacaër. “Approximation of the Basic Reproduction Number R0 for Vector-Borne Diseases with a Periodic Vector Population”. In: *Bulletin of Mathematical Biology* 69.3 (Apr. 2007), pages 1067–1091. ISSN: 1522-9602. DOI: [10.1007/s11538-006-9166-9](https://doi.org/10.1007/s11538-006-9166-9) (cited on page 62).
- [12] Nicolas Bacaër and Souad Guernaoui. “The epidemic threshold of vector-borne diseases with seasonality”. In: *Journal of Mathematical Biology* 53.3 (Sept. 2006), pages 421–436. ISSN: 1432-1416. DOI: [10.1007/s00285-006-0015-0](https://doi.org/10.1007/s00285-006-0015-0) (cited on page 52).
- [13] D. J. Beal et al. “Seasonal Abundance and Infectivity of *Philaenus spumarius* (Hemiptera: Aphrophoridae), a Vector of *Xylella fastidiosa* in California Vineyards”. In: *Environmental Entomology* 50.2 (Jan. 2021), pages 467–476. ISSN: 0046-225X. DOI: [10.1093/ee/nvaa178](https://doi.org/10.1093/ee/nvaa178). eprint: <https://academic.oup.com/ee/article-pdf/50/2/467/37313684/nvaa178.pdf>. URL: <https://doi.org/10.1093/ee/nvaa178> (cited on page 62).
- [14] E. Bertuzzo et al. “On spatially explicit models of cholera epidemics”. In: *Journal of The Royal Society Interface* 7.43 (2010), pages 321–333. DOI: [10.1098/rsif.2009.0204](https://doi.org/10.1098/rsif.2009.0204) (cited on page 37).

- [15] Jeff Bezanson et al. “Julia: A fresh approach to numerical computing”. In: *SIAM Review* 59.1 (2017), pages 65–98. DOI: [10.1137/141000671](https://doi.org/10.1137/141000671) (cited on pages 21, 26, 63).
- [16] G. Bidegain et al. “Microparasitic disease dynamics in benthic suspension feeders: Infective dose, non-focal hosts, and particle diffusion”. In: *Ecological Modelling* 328 (2016), pages 44–61. ISSN: 0304-3800. DOI: [10.1016/j.ecolmodel.2016.02.008](https://doi.org/10.1016/j.ecolmodel.2016.02.008) (cited on pages 8, 10, 29).
- [17] G. Bidegain et al. “Modeling the transmission of *Perkinsus marinus* in the Eastern oyster *Crassostrea virginica*”. In: *Fisheries Research* 186 (2017), pages 82–93. ISSN: 0165-7836. DOI: [10.1016/j.fishres.2016.08.006](https://doi.org/10.1016/j.fishres.2016.08.006) (cited on pages 8, 29).
- [18] Gorka Bidegain et al. “Marine infectious disease dynamics and outbreak thresholds: contact transmission, pandemic infection, and the potential role of filter feeders”. In: *Ecosphere* 7.4 (2016), e01286. DOI: [10.1002/ecs2.1286](https://doi.org/10.1002/ecs2.1286) (cited on pages 8, 9, 11, 13, 18, 23, 24).
- [19] N. Bodino et al. “Phenology, seasonal abundance and stage-structure of spittlebug (Hemiptera: Aphrophoridae) populations in olive groves in Italy”. In: *Scientific reports* 9.1 (2019), pages 1–17. DOI: [10.1038/s41598-019-54279-8](https://doi.org/10.1038/s41598-019-54279-8) (cited on pages 59, 60).
- [20] N. Bodino et al. “Temporal dynamics of the transmission of *Xylella fastidiosa* subsp. *pauca* by *Philaenus spumarius* to olive plants”. In: *Entomologia Generalis* 41.5 (Oct. 2021), pages 463–480. DOI: [10.1127/entomologia/2021/1294](https://doi.org/10.1127/entomologia/2021/1294). URL: <http://dx.doi.org/10.1127/entomologia/2021/1294> (cited on pages 60, 67).
- [21] Antonio Box et al. “Reduced Antioxidant Response of the Fan Mussel *Pinna nobilis* Related to the Presence of *Haplosporidium pinnae*”. In: *Pathogens* 9.11 (Nov. 2020), page 932. ISSN: 2076-0817. DOI: [10.3390/pathogens9110932](https://doi.org/10.3390/pathogens9110932) (cited on page 8).
- [22] C. Bragard et al. “Status and Prospects of Plant Virus Control Through Interference with Vector Transmission”. In: *Annual Review of Phytopathology* 51.1 (2013), pages 177–201. DOI: [10.1146/annurev-phyto-082712-102346](https://doi.org/10.1146/annurev-phyto-082712-102346) (cited on page 45).
- [23] Fred Brauer. “Models for the spread of universally fatal diseases”. In: *Journal of Mathematical Biology* 28 (1990), pages 451–462. DOI: [10.1007/BF00178328](https://doi.org/10.1007/BF00178328) (cited on page 10).
- [24] Fred Brauer. “Compartmental Models in Epidemiology”. In: *Mathematical Epidemiology*. Edited by Fred Brauer, Pauline van den Driessche, and Jianhong Wu. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, pages 19–79. ISBN: 978-3-540-78911-6. DOI: [10.1007/978-3-540-78911-6_2](https://doi.org/10.1007/978-3-540-78911-6_2) (cited on pages 36, 45).
- [25] Fred Brauer et al. “Some models for epidemics of vector-transmitted diseases”. In: *Infectious Disease Modelling* 1.1 (2016), pages 79–87. ISSN: 2468-0427. DOI: [10.1016/j.idm.2016.08.001](https://doi.org/10.1016/j.idm.2016.08.001) (cited on pages 46–48, 72).
- [26] Broder Breckling. “Individual-Based Modelling Potentials and Limitations”. In: *The Scientific World Journal* 2 (Apr. 2002), page 684985. ISSN: 2356-6140. DOI: [10.1100/tsw.2002.179](https://doi.org/10.1100/tsw.2002.179) (cited on page 30).
- [27] James H. Brown et al. “Toward a Metabolic Theory of Ecology”. In: *Ecology* 85.5 (2004), pages 1771–1789. DOI: [10.1890/03-9000](https://doi.org/10.1890/03-9000) (cited on page 22).
- [28] M. Brunetti et al. “A mathematical model for *Xylella fastidiosa* epidemics in the Mediterranean regions. Promoting good agronomic practices for their effective control.” In: *Ecological Modelling* 432 (2020), page 109204. ISSN: 0304-3800. DOI: [10.1016/j.ecolmodel.2020.109204](https://doi.org/10.1016/j.ecolmodel.2020.109204). URL: <https://www.sciencedirect.com/science/article/pii/S030438002030274X> (cited on page 60).
- [29] Colleen A. Burge et al. “Climate change influences on marine infectious diseases: Implications for management and society”. In: *Annual Review of Marine Science* 6 (2014), pages 249–277. ISSN: 19411405. DOI: [10.1146/annurev-marine-010213-135029](https://doi.org/10.1146/annurev-marine-010213-135029) (cited on page 7).
- [30] Eugene M. Burreson and Susan E. Ford. “A review of recent information on the Haplosporidia, with special reference to *Haplosporidium nelsoni* (MSX disease)”. In: *Aquatic Living Resources* 17.4 (Oct. 2004), pages 499–517. ISSN: 0990-7440. DOI: [10.1051/alr:2004056](https://doi.org/10.1051/alr:2004056) (cited on page 8).
- [31] A Butler, Nardo Vicente, and Béatrice De Gaulejac. “Ecology of the pterioid bivalves *Pinna bicolor* Gmelin and *Pinna nobilis* L”. In: *Marine Life* 3.1-2 (1993), pages 37–45. URL: <https://www.semanticscholar.org/paper/Ecology-of-the-pterioid-bivalves-Pinna-bicolor-and-Butler-Vicente-ead2e17a9578a52c274c058bef27084566e12f33> (cited on page 8).
- [32] Miguel Cabanellas-Reboredo et al. “Tracking a mass mortality outbreak of pen shell *Pinna nobilis* populations: A collaborative effort of scientists and citizens”. In: *Scientific Reports* 9.1 (Sept. 2019), page 13355. ISSN: 2045-2322. DOI: [10.1038/s41598-019-49808-4](https://doi.org/10.1038/s41598-019-49808-4) (cited on pages 8, 20, 29).

- [33] Danielle L. Cantrell et al. “Modeling Pathogen Dispersal in Marine Fish and Shellfish”. In: *Trends in Parasitology* 36.3 (2020), pages 239–249. ISSN: 1471-4922. DOI: [10.1016/j.pt.2019.12.013](https://doi.org/10.1016/j.pt.2019.12.013) (cited on page 7).
- [34] F. Carella et al. “A mycobacterial disease is associated with the silent mass mortality of the pen shell *Pinna nobilis* along the Tyrrhenian coastline of Italy”. In: *Scientific Reports* 9.1 (Feb. 2019), page 2725. ISSN: 2045-2322. DOI: [10.1038/s41598-018-37217-y](https://doi.org/10.1038/s41598-018-37217-y) (cited on page 8).
- [35] Jessica Cariboni et al. “The Role of Sensitivity Analysis in Ecological Modelling”. In: *Ecological Modelling* 203 (Apr. 2007), pages 167–182. DOI: [10.1016/j.ecolmodel.2005.10.045](https://doi.org/10.1016/j.ecolmodel.2005.10.045) (cited on page 26).
- [36] C. Castillo-Chavez and H. R. Thieme. “Asymptotically autonomous epidemic models”. In: *Mathematical Population Dynamics: Analysis of Heterogeneity, Vol. I, Theory of Epidemics*. Edited by D. Arino O. Axelrod, M. Kimmel, and M. Langlais. Wuerz publishing (Winnipeg), 1995, pages 33–50. URL: <https://ecommons.cornell.edu/bitstream/handle/1813/31834/BU-1248-M.pdf> (cited on page 49).
- [37] Mario Castro et al. “The turning point and end of an expanding epidemic cannot be precisely forecast”. In: *Proceedings of the National Academy of Sciences* 117.42 (2020), pages 26190–26196. ISSN: 0027-8424. DOI: [10.1073/pnas.2007868117](https://doi.org/10.1073/pnas.2007868117) (cited on pages 12, 23).
- [38] Gaetano Catanese et al. “*Haplosporidium pinnae* sp. nov., a haplosporidan parasite associated with mass mortalities of the fan mussel, *Pinna nobilis*, in the Western Mediterranean Sea”. In: *Journal of Invertebrate Pathology* 157 (2018), pages 9–24. ISSN: 0022-2011. DOI: [10.1016/j.jip.2018.07.006](https://doi.org/10.1016/j.jip.2018.07.006) (cited on pages 8, 19).
- [39] V. Cavalieri et al. “Transmission of *Xylella fastidiosa* Subspecies *Pauca* Sequence Type 53 by Different Insect Species”. In: *Insects* 10.10 (2019), page 324. ISSN: 2075-4450. DOI: [10.3390/insects10100324](https://doi.org/10.3390/insects10100324) (cited on pages 60, 66).
- [40] Y. H. Chew et al. “Mathematical Models Light Up Plant Signaling”. In: *The Plant Cell* 26.1 (Jan. 2014), pages 5–20. ISSN: 1040-4651. DOI: [10.1105/tpc.113.120006](https://doi.org/10.1105/tpc.113.120006). eprint: https://academic.oup.com/plcell/article-pdf/26/1/5/35892793/plcell_v26_1_5.pdf. URL: <https://doi.org/10.1105/tpc.113.120006> (cited on page 59).
- [41] C. Chiyaka et al. “Modeling huanglongbing transmission within a citrus tree”. In: *Proceedings of the National Academy of Sciences* 109.30 (2012), pages 12213–12218. DOI: [10.1073/pnas.1208326109](https://doi.org/10.1073/pnas.1208326109) (cited on page 59).
- [42] S. M. Chmiel and M. C. Wilson. “Estimation of the Lower and Upper Developmental Threshold Temperatures and Duration of the Nymphal Stages of the Meadow Spittlebug, *Philaenus spumarius*”. In: *Environmental Entomology* 8.4 (Aug. 1979), pages 682–685. ISSN: 0046-225X. DOI: [10.1093/ee/8.4.682](https://doi.org/10.1093/ee/8.4.682). eprint: <https://academic.oup.com/ee/article-pdf/8/4/682/18278470/ee8-0682.pdf>. URL: <https://doi.org/10.1093/ee/8.4.682> (cited on page 59).
- [43] Gerardo Chowell. “Fitting dynamic models to epidemic outbreaks with quantified uncertainty: A primer for parameter uncertainty, identifiability, and forecasts”. In: *Infectious Disease Modelling* 2.3 (2017), pages 379–398. ISSN: 2468-0427. DOI: [10.1016/j.idm.2017.08.001](https://doi.org/10.1016/j.idm.2017.08.001) (cited on pages 46, 70).
- [44] Joao R Coelho and Fernando SM Bezerra. “The effects of temperature change on the infection rate of *Biomphalaria glabrata* with *Schistosoma mansoni*”. en. In: *Memórias do Instituto Oswaldo Cruz* 101 (Mar. 2006), pages 223–224. ISSN: 0074-0276. DOI: [10.1590/S0074-02762006000200016](https://doi.org/10.1590/S0074-02762006000200016) (cited on page 27).
- [45] D. Cornara, D. Bosco, and A. Fereres. “*Philaenus spumarius*: when an old acquaintance becomes a new threat to European agriculture”. In: *Journal of Pest Science* 91.3 (2018), pages 957–972. DOI: [10.1007/s10340-018-0966-0](https://doi.org/10.1007/s10340-018-0966-0) (cited on pages 59, 60).
- [46] D. Cornara et al. “Spittlebugs as vectors of *Xylella fastidiosa* in olive orchards in Italy”. In: *Journal of Pest Science* 90.2 (Mar. 2017), pages 521–530. ISSN: 1612-4766. DOI: [10.1007/s10340-016-0793-0](https://doi.org/10.1007/s10340-016-0793-0). URL: <https://doi.org/10.1007/s10340-016-0793-0> (cited on pages 59, 62).
- [47] D. Cornara et al. “Transmission of *Xylella fastidiosa* by naturally infected *Philaenus spumarius* (Hemiptera, Aphrophoridae) to different host plants”. In: *Journal of Applied Entomology* 141.1-2 (2017), pages 80–87. DOI: [10.1111/jen.12365](https://doi.org/10.1111/jen.12365) (cited on page 66).
- [48] D. Cornara et al. “EPG combined with micro-CT and video recording reveals new insights on the feeding behavior of *Philaenus spumarius*”. In: *PLoS One* 13.7 (July 2018), pages 1–20. DOI: [10.1371/journal.pone.0199154](https://doi.org/10.1371/journal.pone.0199154) (cited on pages 59, 71).

- [49] D. Cornara et al. “Natural areas as reservoir of candidate vectors of *Xylella fastidiosa*”. In: *Bulletin of Insectology* 74 (2021), pages 173–180 (cited on page 60).
- [50] Paul C Cross et al. “Utility of R_0 as a predictor of disease invasion in structured populations”. In: *Journal of The Royal Society Interface* 4.13 (2007), pages 315–324. DOI: [10.1098/rsif.2006.0185](https://doi.org/10.1098/rsif.2006.0185) (cited on page 29).
- [51] S Culloty and M Mulcahy. “*Bonamia ostreae* in the Native Oyster *Ostrea edulis*”. In: *Marine Environment and Health Series* 29.1649 (2007), pages 1–36. ISSN: 1649-0053. URL: <https://oar.marine.ie/bitstream/handle/10793/269/No%202029%20Marine%20Environment%20and%20Health%20Series.pdf?sequence=1&isAllowed=y> (cited on page 19).
- [52] N. J. Cunniffe et al. “Thirteen challenges in modelling plant diseases”. In: *Epidemics* 10 (2015), pages 6–10. DOI: [10.1016/j.epidem.2014.06.002](https://doi.org/10.1016/j.epidem.2014.06.002) (cited on page 71).
- [53] Andrew A. Cunningham, Peter Daszak, and James L. N. Wood. “One Health, emerging infectious diseases and wildlife: two decades of progress?” In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 372.1725 (2017), page 20160167. DOI: [10.1098/rstb.2016.0167](https://doi.org/10.1098/rstb.2016.0167) (cited on page 29).
- [54] Susana Darriba. “First haplosporidan parasite reported infecting a member of the Superfamily Pinoidea (*Pinna nobilis*) during a mortality event in Alicante (Spain, Western Mediterranean)”. In: *Journal of Invertebrate Pathology* 148 (2017), pages 14–19. ISSN: 0022-2011. DOI: [10.1016/j.jip.2017.05.006](https://doi.org/10.1016/j.jip.2017.05.006) (cited on page 8).
- [55] Peter Daszak, Andrew A. Cunningham, and Alex D. Hyatt. “Emerging Infectious Diseases of Wildlife—Threats to Biodiversity and Human Health”. In: *Science* 287.5452 (2000), pages 443–449. ISSN: 0036-8075. DOI: [10.1126/science.287.5452.443](https://doi.org/10.1126/science.287.5452.443) (cited on page 29).
- [56] M. P. Daugherty and R. P. P. Almeida. “Understanding How an Invasive Vector Drives Pierce’s Disease Epidemics: Seasonality and Vine-to-Vine Spread”. In: *Phytopathology* 109.2 (2019), pages 277–285. DOI: [10.1094/PHYTO-07-18-0217-FI](https://doi.org/10.1094/PHYTO-07-18-0217-FI). URL: <https://doi.org/10.1094/PHYTO-07-18-0217-FI> (cited on page 60).
- [57] O. Diekmann, J. A. P. Heesterbeek, and M. G. Roberts. “The construction of next-generation matrices for compartmental epidemic models”. In: *Journal of The Royal Society Interface* 7.47 (2010), pages 873–885. DOI: [10.1098/rsif.2009.0386](https://doi.org/10.1098/rsif.2009.0386) (cited on pages 12, 25, 29, 33, 45, 49, 56, 57, 62, 70, 72).
- [58] Odo Diekmann, Hans Heesterbeek, and Tom Britton. *Mathematical Tools for Understanding Infectious Disease Dynamics*. Princeton: Princeton U.P., 2013. ISBN: 978-0-691-15539-5 (cited on page 9).
- [59] P. van den Driessche and James Watmough. “Further Notes on the Basic Reproduction Number”. In: *Mathematical Epidemiology*. Edited by Fred Brauer, Pauline van den Driessche, and Jianhong Wu. Berlin, Heidelberg: Springer Berlin Heidelberg, 2008, pages 159–178. ISBN: 978-3-540-78911-6. DOI: [10.1007/978-3-540-78911-6_6](https://doi.org/10.1007/978-3-540-78911-6_6) (cited on page 12).
- [60] Morgan E. Eisenlord et al. “Ochre star mortality during the 2014 wasting disease epizootic: role of population size structure and temperature”. In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 371.1689 (2016), page 20150212. DOI: [10.1098/rstb.2015.0212](https://doi.org/10.1098/rstb.2015.0212) (cited on page 29).
- [61] Lourdes Esteva and Cristobal Vargas. “Analysis of a dengue disease transmission model”. In: *Mathematical Biosciences* 150.2 (1998), pages 131–151. ISSN: 0025-5564. DOI: [10.1016/S0025-5564\(98\)10003-2](https://doi.org/10.1016/S0025-5564(98)10003-2) (cited on page 49).
- [62] A. Fierro, A. Liccardo, and F. Porcelli. “A lattice model to manage the vector and the infection of the *Xylella fastidiosa* on olive trees”. In: *Scientific Reports* 9.1 (June 2019), page 8723. ISSN: 2045-2322. DOI: [10.1038/s41598-019-44997-4](https://doi.org/10.1038/s41598-019-44997-4). URL: <https://doi.org/10.1038/s41598-019-44997-4> (cited on pages 60, 63).
- [63] J. A. N. Filipe and M. M. Maule. “Analytical methods for predicting the behaviour of population models with general spatial interactions”. In: *Mathematical Biosciences* 183 (2003), pages 15–35. DOI: [10.1016/S0025-5564\(02\)00224-9](https://doi.org/10.1016/S0025-5564(02)00224-9) (cited on page 36).
- [64] J. A. N. Filipe and M. M. Maule. “Effects of dispersal mechanisms on spatio-temporal development of epidemics”. In: *Journal of Theoretical Biology* 226 (2004), pages 125–141. DOI: [10.1016/S0022-5193\(03\)00278-9](https://doi.org/10.1016/S0022-5193(03)00278-9) (cited on page 36).
- [65] Susana Flecha et al. “pH trends and seasonal cycle in the coastal Balearic Sea reconstructed through machine learning”. In: *Scientific Reports* 12.1 (July 2022), page 12956. ISSN: 2045-2322. DOI: [10.1038/s41598-022-17253-5](https://doi.org/10.1038/s41598-022-17253-5). URL: <https://doi.org/10.1038/s41598-022-17253-5> (cited on page 17).

- [66] R. Fletcher. “Newton-Like Methods”. In: *Practical Methods of Optimization*. John Wiley & Sons, Ltd, 2013. Chapter 3, pages 44–79. ISBN: 9781118723203. DOI: [10.1002/9781118723203.ch3](https://doi.org/10.1002/9781118723203.ch3) (cited on page 21).
- [67] J. H. Freitag. “Host range of the Pierce’s disease virus of grapes as determined by insect transmission”. In: *Phytopathology* 41.10 (1951), pages 920–934 (cited on page 60).
- [68] J.R. García-March et al. “Can we save a marine species affected by a highly infective, highly lethal, waterborne disease from extinction?” In: *Biological Conservation* 243 (2020), page 108498. ISSN: 0006-3207. DOI: [10.1016/j.biocon.2020.108498](https://doi.org/10.1016/j.biocon.2020.108498) (cited on pages 8, 19–21).
- [69] R Girms, JF Walsh, and JB Davies. “Studies on the reinvansion of the *Onchocerciasis* Control Programme in the Volta River Basin by *Simulium damnosum* sI with emphasis on the south-western areas”. In: *Tropenmedizin und Parasitologie* 30.3 (1979), pages 345–362 (cited on page 46).
- [70] H. Ge, K. Xu, and Z. Ghahramani. “Turing: a language for flexible probabilistic inference”. In: *International Conference on Artificial Intelligence and Statistics, AISTATS 2018, 9–11 April 2018, Playa Blanca, Lanzarote, Canary Islands, Spain*. 2018, pages 1682–1690. URL: <http://proceedings.mlr.press/v84/ge18b.html> (cited on page 63).
- [71] Daniel T. Gillespie. “Exact stochastic simulation of coupled chemical reactions”. In: *The Journal of Physical Chemistry* 81.25 (Dec. 1977), pages 2340–2361. ISSN: 0022-3654. DOI: [10.1021/j100540a008](https://doi.org/10.1021/j100540a008) (cited on page 31).
- [72] Christopher A. Gilligan and Frank van den Bosch. “Epidemiological Models for Invasion and Persistence of Pathogens”. In: *Annual Review of Phytopathology* 46 (2008), pages 385–418. DOI: [10.1146/annurev.phyto.45.062806.094357](https://doi.org/10.1146/annurev.phyto.45.062806.094357) (cited on page 30).
- [73] James F. Gillooly et al. “Effects of Size and Temperature on Metabolic Rate”. In: *Science* 293.5538 (2001), pages 2248–2251. ISSN: 0036-8075. DOI: [10.1126/science.1061967](https://doi.org/10.1126/science.1061967) (cited on pages 26, 27).
- [74] Alex Giménez-Romero. “A compartmental model for *Xylella fastidiosa* related diseases”. In: *GitHub Repo* (2022). URL: <https://github.com/agimenezromero/Spatial-effects-in-parasite-induced-marine-diseases-of-immobile-hosts> (cited on pages 31, 62, 63).
- [75] Alex Giménez-Romero et al. “Global predictions for the risk of establishment of Pierce’s disease of grapevines”. In: *Communications Biology* 5.1 (Dec. 2022), page 1389. ISSN: 2399-3642. DOI: [10.1038/s42003-022-04358-w](https://doi.org/10.1038/s42003-022-04358-w). URL: <https://doi.org/10.1038/s42003-022-04358-w> (cited on pages 17, 60).
- [76] Àlex Giménez-Romero, Rosa Flaquer-Galmés, and Manuel A. Matías. “Vector-borne diseases with nonstationary vector populations: The case of growing and decaying populations”. In: *Phys. Rev. E* 106 (5 Nov. 2022), page 054402. DOI: [10.1103/PhysRevE.106.054402](https://doi.org/10.1103/PhysRevE.106.054402). URL: <https://link.aps.org/doi/10.1103/PhysRevE.106.054402> (cited on pages 17, 62, 70, 72).
- [77] Àlex Giménez-Romero, Eduardo Moralejo, and Manuel A Matias. “High-resolution climate data reveals increased risk of Pierce’s Disease for grapevines worldwide”. In: *bioRxiv* (2024). DOI: [10.1101/2024.03.06.583743](https://doi.org/10.1101/2024.03.06.583743). eprint: <https://www.biorxiv.org/content/early/2024/03/11/2024.03.06.583743.full.pdf>. URL: <https://www.biorxiv.org/content/early/2024/03/11/2024.03.06.583743> (cited on page 18).
- [78] Àlex Giménez-Romero, Eduardo Moralejo, and Manuel A. Matías. “A Compartmental Model for *Xylella fastidiosa* Diseases with Explicit Vector Seasonal Dynamics”. In: *Phytopathology®* 113.9 (2023). PMID: 36774557, pages 1686–1696. DOI: [10.1094/PHYTO-11-22-0428-V](https://doi.org/10.1094/PHYTO-11-22-0428-V). eprint: <https://doi.org/10.1094/PHYTO-11-22-0428-V>. URL: <https://doi.org/10.1094/PHYTO-11-22-0428-V> (cited on page 17).
- [79] Àlex Giménez-Romero et al. “Modelling parasite-produced marine diseases: The case of the mass mortality event of *Pinna nobilis*”. In: *Ecological Modelling* 459 (2021), page 109705. ISSN: 0304-3800. DOI: <https://doi.org/10.1016/j.ecolmodel.2021.109705>. URL: <https://www.sciencedirect.com/science/article/pii/S030438002100260X> (cited on pages 17, 29, 30, 32–34, 40, 41, 46).
- [80] Àlex Giménez-Romero et al. “Spatial effects in parasite-induced marine diseases of immobile hosts”. In: *Royal Society Open Science* 9.8 (2022), page 212023. DOI: [10.1098/rsos.212023](https://doi.org/10.1098/rsos.212023). eprint: <https://royalsocietypublishing.org/doi/pdf/10.1098/rsos.212023>. URL: <https://royalsocietypublishing.org/doi/abs/10.1098/rsos.212023> (cited on page 17).

- [81] Àlex Giménez-Romero et al. “Contrasting Patterns of Pierce’s Disease Risk in European Vineyards Under Global Warming”. In: *bioRxiv* (2023). DOI: [10.1101/2023.07.17.549293](https://doi.org/10.1101/2023.07.17.549293). eprint: <https://www.biorxiv.org/content/early/2023/07/18/2023.07.17.549293.full.pdf>. URL: <https://www.biorxiv.org/content/early/2023/07/18/2023.07.17.549293> (cited on page 17).
- [82] Volker Grimm and Steven F Railsback. *Individual-based Modeling and Ecology*. Princeton (NJ): Princeton University Press, 2005. ISBN: 978-0691096650 (cited on pages 30, 71).
- [83] Ximing Guo and Susan E. Ford. “Infectious diseases of marine molluscs and host responses as revealed by genomic tools”. In: *Philosophical Transactions of the Royal Society B: Biological Sciences* 371.1689 (2016), page 20150206. DOI: [10.1098/rstb.2015.0206](https://doi.org/10.1098/rstb.2015.0206) (cited on page 29).
- [84] Drew Harvell et al. “The rising tide of ocean diseases: Unsolved problems and research priorities”. In: *Frontiers in Ecology and the Environment* 2.7 (2004), pages 375–382. ISSN: 15409309. DOI: [10.1890/1540-9295\(2004\)002\[0375:TRTOOD\]2.0.CO;2](https://doi.org/10.1890/1540-9295(2004)002[0375:TRTOOD]2.0.CO;2) (cited on pages 7, 9).
- [85] Harold H Haskin and Jay D Andrews. “Uncertainties and speculations about the life cycle of the eastern oyster pathogen *Haplosporidium nelsoni* (MSX)”. In: *Special Publication (American Fisheries Society)* 18 (1988). URL: <https://ir.library.oregonstate.edu/downloads/pn89df056?locale=en> (cited on page 19).
- [86] Iris E. Hendriks et al. “Seagrass Meadows Modify Drag Forces on the Shell of the Fan Mussel *Pinnina nobilis*”. In: *Estuaries and Coasts* 34.1 (Jan. 2011), pages 60–67. ISSN: 1559-2731. DOI: [10.1007/s12237-010-9309-y](https://doi.org/10.1007/s12237-010-9309-y) (cited on page 8).
- [87] P.M. Hine. “The ecology of Bonamia and decline of bivalve molluscs”. In: *New Zealand Journal of Ecology* 20.1 (1996), pages 109–116. URL: <https://newzealandecology.org/nzje/1993> (cited on page 19).
- [88] M. D. Homan and A. Gelman. “The No-U-Turn Sampler: Adaptively Setting Path Lengths in Hamiltonian Monte Carlo”. In: *J. Mach. Learn. Res.* 15.1 (Jan. 2014), pages 1593–1623. ISSN: 1532-4435. URL: <https://dl.acm.org/doi/10.5555/2627435.2638586> (cited on page 63).
- [89] D. L. Hopkins and A. H. Purcell. “*Xylella fastidiosa*: Cause of Pierce’s Disease of Grapevine and Other Emergent Diseases”. In: *Plant Disease* 86 (Oct. 2002), pages 1056–1066. DOI: [10.1094/PDIS.2002.86.10.1056](https://doi.org/10.1094/PDIS.2002.86.10.1056) (cited on page 59).
- [90] Weijie Huang et al. “Bacterial Vector-Borne Plant Diseases: Unanswered Questions and Future Directions”. In: *Molecular Plant* 13.10 (2020), pages 1379–1393. ISSN: 1674-2052. DOI: [10.1016/j.molp.2020.08.010](https://doi.org/10.1016/j.molp.2020.08.010) (cited on page 45).
- [91] IUCN. “IUCN. The IUCN Red List of Threatened Species. Version 2019-3”. In: (2019). URL: <http://www.iucnredlist.org> (cited on page 8).
- [92] M. J. Jeger and C. Bragard. “The Epidemiology of *Xylella fastidiosa*; A Perspective on Current Knowledge and Framework to Investigate Plant Host–Vector–Pathogen Interactions”. In: *Phytopathology* 109.2 (2019), pages 200–209. DOI: [10.1094/PHYTO-07-18-0239-FI](https://doi.org/10.1094/PHYTO-07-18-0239-FI) (cited on pages 59, 60).
- [93] M. J. Jeger, L. V. Madden, and F Van Den Bosch. “Plant virus epidemiology: Applications and prospects for mathematical modeling and analysis to improve understanding and disease control”. In: *Plant Disease* 102.5 (2018), pages 837–854. DOI: [10.1094/PDIS-04-17-0612-FE](https://doi.org/10.1094/PDIS-04-17-0612-FE) (cited on page 71).
- [94] M. J. Jeger et al. “A model for analysing plant-virus transmission characteristics and epidemic development”. In: *Mathematical Medicine and Biology: A Journal of the IMA* 15.1 (Mar. 1998), pages 1–18. ISSN: 1477-8599. DOI: [10.1093/imammb/15.1.1](https://doi.org/10.1093/imammb/15.1.1). eprint: <https://academic.oup.com/imammb/article-pdf/15/1/1/1995834/15-1-1.pdf>. URL: <https://doi.org/10.1093/imammb/15.1.1> (cited on page 59).
- [95] M. J. Jeger et al. “Epidemiology of insect-transmitted plant viruses: modelling disease dynamics and control interventions”. In: *Physiological Entomology* 29.3 (2004), pages 291–304. DOI: [10.1111/j.0307-6962.2004.00394.x](https://doi.org/10.1111/j.0307-6962.2004.00394.x). eprint: <https://resjournals.onlinelibrary.wiley.com/doi/pdf/10.1111/j.0307-6962.2004.00394.x>. URL: <https://resjournals.onlinelibrary.wiley.com/doi/abs/10.1111/j.0307-6962.2004.00394.x> (cited on page 59).
- [96] K. Jones et al. “A review of fibropapillomatosis in Green turtles (*Chelonia mydas*)”. In: *The Veterinary Journal* 212 (2016), pages 48–57. ISSN: 1090-0233. DOI: <https://doi.org/10.1016/j.tvjl.2015.10.041> (cited on page 29).

- [97] Jean Claude Kamgang and Gauthier Sallet. “Computation of threshold conditions for epidemiological models and global stability of the disease-free equilibrium (DFE)”. In: *Mathematical Biosciences* 213.1 (2008), pages 1–12. ISSN: 0025-5564. DOI: [10.1016/j.mbs.2008.02.005](https://doi.org/10.1016/j.mbs.2008.02.005) (cited on page 45).
- [98] Yu-Han Kao and Marisa C. Eisenberg. “Practical unidentifiability of a simple vector-borne disease model: Implications for parameter estimation and intervention assessment”. In: *Epidemics* 25 (2018), pages 89–100. ISSN: 1755-4365. DOI: [10.1016/j.epidem.2018.05.010](https://doi.org/10.1016/j.epidem.2018.05.010) (cited on page 46).
- [99] Stelios Katsanevakis. “Transplantation as a conservation action to protect the Mediterranean fan mussel *Pinna nobilis*”. In: *Marine Ecology Progress Series* 546 (2016), pages 113–122. DOI: [10.3354/meps11658](https://doi.org/10.3354/meps11658) (cited on page 8).
- [100] Matt J. Keeling and Bryan T. Grenfell. “Individual-based Perspectives on R₀”. In: *Journal of Theoretical Biology* 203.1 (2000), pages 51–61. ISSN: 0022-5193. DOI: <https://doi.org/10.1006/jtbi.1999.1064> (cited on page 36).
- [101] W. O. Kermack and A. G. McKendrick. “Contributions to the mathematical theory of epidemics-III. Further studies of the problem of endemicity”. In: *Bulletin of Mathematical Biology* 53.1 (1933), pages 89–118. ISSN: 0092-8240. URL: <http://www.sciencedirect.com/science/article/pii/S0092824005800424> (cited on page 7).
- [102] William Ogilvy Kermack and A. G. McKendrick. “A contribution to the mathematical theory of epidemics”. In: *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 115.772 (1927), pages 700–721. DOI: [10.1098/rspa.1927.0118](https://doi.org/10.1098/rspa.1927.0118) (cited on pages 7, 9, 45).
- [103] William Ogilvy Kermack and A. G. McKendrick. “Contributions to the mathematical theory of epidemics-II. The problem of endemicity”. In: *Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character* 138.834 (1932), pages 55–83. DOI: [10.1098/rspa.1932.0171](https://doi.org/10.1098/rspa.1932.0171) (cited on page 7).
- [104] Kevin D. Lafferty, James W. Porter, and Susan E. Ford. “Are Diseases Increasing in the Ocean?” In: *Annual Review of Ecology, Evolution, and Systematics* 35 (2004), pages 31–54. ISSN: 1543592X, 15452069. DOI: [10.1146/annurev.ecolsys.35.021103.105704](https://doi.org/10.1146/annurev.ecolsys.35.021103.105704) (cited on pages 7, 29).
- [105] Kevin D. Lafferty et al. “Infectious diseases affect marine fisheries and aquaculture economics”. In: *Annual Review of Marine Science* 7 (2015), pages 471–496. ISSN: 19410611. DOI: [10.1146/annurev-marine-010814-015646](https://doi.org/10.1146/annurev-marine-010814-015646) (cited on page 7).
- [106] C. Lago et al. “Degree-day-based model to predict egg hatching of *Philaenus spumarius* (Hemiptera: Aphrophoridae), the main vector of *Xylella fastidiosa* in Europe”. In: *bioRxiv* (2022). DOI: [10.1101/2022.11.10.515963](https://doi.org/10.1101/2022.11.10.515963). eprint: <https://www.biorxiv.org/content/early/2022/11/13/2022.11.10.515963.full.pdf>. URL: <https://www.biorxiv.org/content/early/2022/11/13/2022.11.10.515963> (cited on pages 69, 71).
- [107] Clara Lago et al. “Degree-day-based model to predict egg hatching of *Philaenus spumarius* (Hemiptera: Aphrophoridae), the main vector of *Xylella fastidiosa* in Europe”. In: *Environmental Entomology* 52.3 (Apr. 2023), pages 350–359. ISSN: 0046-225X. DOI: [10.1093/ee/nvad013](https://doi.org/10.1093/ee/nvad013). eprint: <https://academic.oup.com/ee/article-pdf/52/3/350/50615564/nvad013.pdf>. URL: <https://doi.org/10.1093/ee/nvad013> (cited on page 19).
- [108] Abid Ali Lashari and Gul Zaman. “Global dynamics of vector-borne diseases with horizontal transmission in host population”. In: *Computers & Mathematics with Applications* 61.4 (2011), pages 745–754. ISSN: 0898-1221. DOI: [10.1016/j.camwa.2010.12.018](https://doi.org/10.1016/j.camwa.2010.12.018) (cited on page 49).
- [109] István G. Laukó. “Stability of disease free sets in epidemic models”. In: *Mathematical and Computer Modelling* 43.11 (2006), pages 1357–1366. ISSN: 0895-7177. DOI: [10.1016/j.mcm.2005.06.011](https://doi.org/10.1016/j.mcm.2005.06.011) (cited on page 45).
- [110] S. A. Levin, editor. *Mathematics and Biology: The interface, challenges and opportunities*. Lawrence Berkeley Lab., CA (USA), 1992. eprint: <https://escholarship.org/uc/item/870966k3> (cited on page 59).
- [111] Jing Li, Daniel Blakeley, and Robert J. Smith. “The failure of R₀”. In: *Computational and Mathematical Methods in Medicine* 2011 (2011), page 527610. ISSN: 1748-6718. DOI: [10.1155/2011/527610](https://doi.org/10.1155/2011/527610) (cited on page 29).

- [112] J. López-Mercadal, P. Mercadal-Frontera, and M. Á. Miranda. “Mechanical management of weeds drops nymphal density of *Xylella fastidiosa* vectors”. In: *bioRxiv* (2022). DOI: [10.1101/2022.10.18.512680](https://doi.org/10.1101/2022.10.18.512680) (cited on pages 59, 71).
- [113] J. López-Mercadal et al. “Collection of data and information in Balearic Islands on biology of vectors and potential vectors of *Xylella fastidiosa* (GP/EFSA/ALPHA/017/01)”. In: *EFSA Supporting Publications* 18.10 (2021), 6925E. DOI: [10.2903/sp.efsa.2021.EN-6925](https://doi.org/10.2903/sp.efsa.2021.EN-6925) (cited on pages 62, 70–72).
- [114] George Macdonald et al. “The epidemiology and control of malaria.” In: *The Epidemiology and Control of Malaria*. (1957) (cited on pages 45, 46).
- [115] L. V. Madden, M. J. Jeger, and F. van den Bosch. “A Theoretical Assessment of the Effects of Vector-Virus Transmission Mechanism on Plant Virus Disease Epidemics”. In: *Phytopathology* 90.6 (2000), pages 576–594. DOI: [10.1094/PHYTO.2000.90.6.576](https://doi.org/10.1094/PHYTO.2000.90.6.576). eprint: <https://doi.org/10.1094/PHYTO.2000.90.6.576> (cited on page 59).
- [116] Maia Martcheva. *An Introduction to Mathematical Epidemiology*. New York: Springer, 2015. ISBN: 978-1-4899-7611-6. DOI: [10.1007/978-1-4899-7612-3](https://doi.org/10.1007/978-1-4899-7612-3) (cited on pages 10, 47, 62).
- [117] Robert M. May and Roy M. Anderson. “Population biology of infectious diseases: Part II”. In: *Nature* 280.5722 (Aug. 1979), pages 455–461. ISSN: 1476-4687. DOI: [10.1038/280455a0](https://doi.org/10.1038/280455a0). URL: <https://doi.org/10.1038/280455a0> (cited on pages 9, 23).
- [118] Hamish I. McCallum et al. “Does terrestrial epidemiology apply to marine systems?” In: *Trends in Ecology & Evolution* 19.11 (2004), pages 585–591. ISSN: 0169-5347. DOI: [10.1016/j.tree.2004.08.009](https://doi.org/10.1016/j.tree.2004.08.009) (cited on pages 7, 9, 23).
- [119] Joseph B. McLaughlin et al. “Outbreak of *Vibrio parahaemolyticus* Gastroenteritis Associated with Alaskan Oysters”. In: *New England Journal of Medicine* 353.14 (2005). PMID: 16207848, pages 1463–1470. DOI: [10.1056/NEJMoa051594](https://doi.org/10.1056/NEJMoa051594) (cited on page 29).
- [120] Péter K. Molnár et al. “Thermal Performance Curves and the Metabolic Theory of Ecology—A Practical Guide to Models and Experiments for Parasitologists”. In: *Journal of Parasitology* 103.5 (2017), pages 423–439. DOI: [10.1645/16-148](https://doi.org/10.1645/16-148) (cited on page 22).
- [121] E. Moralejo et al. “Insights into the epidemiology of Pierce’s disease in vineyards of Mallorca, Spain”. In: *Plant Pathology* 68.8 (2019), pages 1458–1471. DOI: <https://doi.org/10.1111/ppa.13076>. eprint: <https://onlinelibrary.wiley.com/doi/pdf/10.1111/ppa.13076>. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/ppa.13076> (cited on page 59).
- [122] E. Moralejo et al. “Phylogenetic inference enables reconstruction of a long-overlooked outbreak of almond leaf scorch disease (*Xylella fastidiosa*) in Europe”. In: *Communications Biology* 3.1 (Oct. 2020), page 560. ISSN: 2399-3642. DOI: [10.1038/s42003-020-01284-7](https://doi.org/10.1038/s42003-020-01284-7). URL: <https://doi.org/10.1038/s42003-020-01284-7> (cited on pages 59, 62, 63).
- [123] David M. Morens, Gregory K. Folkers, and Anthony S. Fauci. “The challenge of emerging and re-emerging infectious diseases”. In: *Nature* 430.6996 (July 2004), pages 242–249. ISSN: 1476-4687. DOI: [10.1038/nature02759](https://doi.org/10.1038/nature02759) (cited on page 29).
- [124] M. Morente et al. “Distribution and relative abundance of insect vectors of *Xylella fastidiosa* in olive groves of the Iberian Peninsula”. In: *Insects* 9.4 (2018), page 175. DOI: [10.3390/insects9040175](https://doi.org/10.3390/insects9040175) (cited on page 60).
- [125] J. D. Murray. *Mathematical Biology*. Berlin (Germany): Springer, 1989 (cited on page 59).
- [126] J. D. Murray. *Mathematical Biology. I. An Introduction*. 3rd edition. New York, NY: Springer, 2002. ISBN: 978-0-387-22437-4 (cited on pages 12, 24).
- [127] D. Olmo et al. “Landscape Epidemiology of *Xylella fastidiosa* in the Balearic Islands”. In: *Agronomy* 11.3 (2021), page 473. ISSN: 2073-4395. DOI: [10.3390/agronomy11030473](https://doi.org/10.3390/agronomy11030473) (cited on pages 59, 60).
- [128] Ivane Lilian Pairaud et al. “Impacts of climate change on coastal benthic ecosystems: assessing the current risk of mortality outbreaks associated with thermal stress in NW Mediterranean coastal areas”. In: *Ocean Dynamics* 64.1 (Jan. 2014), pages 103–115. ISSN: 1616-7228. DOI: [10.1007/s10236-013-0661-x](https://doi.org/10.1007/s10236-013-0661-x) (cited on page 7).
- [129] Fabrice Pernet et al. “Infectious diseases in oyster aquaculture require a new integrated approach”. In: *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 371.1689 (Mar. 2016), page 20150213. DOI: [10.1098/rstb.2015.0213](https://doi.org/10.1098/rstb.2015.0213) (cited on page 29).

- [130] Fabrice Pernet et al. "Determination of risk factors for herpesvirus outbreak in oysters using a broad-scale spatial epidemiology framework". In: *Scientific Reports* 8.1 (July 2018), page 10869. ISSN: 2045-2322. DOI: [10.1038/s41598-018-29238-4](https://doi.org/10.1038/s41598-018-29238-4) (cited on page 29).
- [131] Bruno Petton et al. "The Pacific Oyster Mortality Syndrome, a Polymicrobial and Multifactorial Disease: State of Knowledge and Future Directions". In: *Frontiers in Immunology* 12 (2021), page 52. ISSN: 1664-3224. DOI: [10.3389/fimmu.2021.630343](https://doi.org/10.3389/fimmu.2021.630343) (cited on page 29).
- [132] Eric N Powell et al. "Modeling the MSX parasite in eastern oyster (*Crassostrea virginica*) populations. III. Regional application and the problem of transmission". In: *Journal of Shellfish Research* 18.2 (1999), pages 517–537. URL: https://digitalcommons.odu.edu/ccpo_pubs/76/ (cited on pages 19, 29).
- [133] Eric N. Powell and Eileen E. Hofmann. "Models of marine molluscan diseases: Trends and challenges". In: *Journal of Invertebrate Pathology* 131 (2015), pages 212–225. ISSN: 0022-2011. DOI: [10.1016/j.jip.2015.07.017](https://doi.org/10.1016/j.jip.2015.07.017) (cited on pages 8–10, 23, 30).
- [134] Patricia Prado et al. "*Pinna nobilis* in suboptimal environments are more tolerant to disease but more vulnerable to severe weather phenomena". In: *Marine Environmental Research* (2020), page 105220. ISSN: 0141-1136. DOI: [10.1016/j.marenres.2020.105220](https://doi.org/10.1016/j.marenres.2020.105220) (cited on page 8).
- [135] A. H. Purcell and A. Finlay. "Evidence for noncirculative transmission of Pierce's disease bacterium by sharpshooter leafhoppers". In: *Phytopathology* 69.4 (1979), pages 393–395. DOI: [10.1094/Phyto-69-393](https://doi.org/10.1094/Phyto-69-393) (cited on page 60).
- [136] Christopher Rackauckas and Qing Nie. "DifferentialEquations.jl – A Performant and Feature-Rich Ecosystem for Solving Differential Equations in Julia". In: *The Journal of Open Research Software* 5.1 (2017), page 15. DOI: [10.5334/jors.151](https://doi.org/10.5334/jors.151) (cited on pages 21, 26, 63).
- [137] R. A. Redak et al. "The biology of xylem fluid-feeding insect vectors of *Xylella fastidiosa* and their relation to disease epidemiology". In: *Annual Review of Entomology* 49.1 (2004), pages 243–270. DOI: [10.1146/annurev.ento.49.061802.123403](https://doi.org/10.1146/annurev.ento.49.061802.123403). eprint: <https://doi.org/10.1146/annurev.ento.49.061802.123403> (cited on page 59).
- [138] Steven Riley et al. "Five challenges for spatial epidemic models". In: *Epidemics* 10 (2015). Challenges in Modelling Infectious Disease Dynamics, pages 68–71. ISSN: 1755-4365. DOI: [10.1016/j.epidem.2014.07.001](https://doi.org/10.1016/j.epidem.2014.07.001) (cited on page 29).
- [139] Joacim Rocklöv and Robert Dubrow. "Climate change: an enduring challenge for vector-borne disease prevention and control". In: *Nature Immunology* 21.5 (May 2020), pages 479–483. ISSN: 1529-2916. DOI: [10.1038/s41590-020-0648-y](https://doi.org/10.1038/s41590-020-0648-y) (cited on page 46).
- [140] Jason R. Rohr and Jeremy M. Cohen. "Understanding how temperature shifts could impact infectious diseases". In: *PLoS Biology* 18.11 (2020), e3000938. DOI: [10.1371/journal.pbio.3000938](https://doi.org/10.1371/journal.pbio.3000938) (cited on page 22).
- [141] Kimberlyn Roosa and Gerardo Chowell. "Assessing parameter identifiability in compartmental dynamic models using a computational approach: application to infectious disease transmission models". In: *Theoretical Biology and Medical Modelling* 16.1 (Jan. 2019), page 1. ISSN: 1742-4682. DOI: [10.1186/s12976-018-0097-6](https://doi.org/10.1186/s12976-018-0097-6) (cited on pages 46, 70).
- [142] Edward P. Rybicki. "A Top Ten list for economically important plant viruses". In: *Archives of Virology* 160 (1 2015), pages 17–20. DOI: [10.1007/s00705-014-2295-9](https://doi.org/10.1007/s00705-014-2295-9) (cited on page 45).
- [143] Cristiano Salata et al. "Coronaviruses: a paradigm of new emerging zoonotic diseases". In: *Pathogens and Disease* 77.9 (Feb. 2020), ftaa006. DOI: [10.1093/femsdp/ftaa006](https://doi.org/10.1093/femsdp/ftaa006) (cited on page 29).
- [144] Andrea Saltelli et al. *Sensitivity Analysis in Practice: A Guide to Assessing Scientific Models*. USA: Halsted Press, 2004. ISBN: 0470870931 (cited on pages 26, 63).
- [145] M. Saponari et al. "Identification of DNA sequences related to *Xylella fastidiosa* in oleander, almond and olive trees exhibiting leaf scorch symptoms in Apulia (Southern Italy)". In: *Journal of Plant Pathology* 95.3 (2013). DOI: [10.4454/JPP.V95I3.035](https://doi.org/10.4454/JPP.V95I3.035) (cited on page 59).
- [146] M. Saponari et al. "*Xylella fastidiosa* in olive in Apulia: Where we stand". In: *Phytopathology* 109.2 (2019), pages 175–186. DOI: [10.1094/PHYTO-08-18-0319-FI](https://doi.org/10.1094/PHYTO-08-18-0319-FI) (cited on page 59).
- [147] Tomislav Šarić et al. "Epidemiology of Noble Pen Shell (*Pinna nobilis* L. 1758) Mass Mortality Events in Adriatic Sea Is Characterised with Rapid Spreading and Acute Disease Progression". In: *Pathogens* 9.10 (Sept. 2020), page 776. ISSN: 2076-0817. DOI: [10.3390/pathogens9100776](https://doi.org/10.3390/pathogens9100776) (cited on page 8).

- [148] S. Sarkar et al. “Biodiversity conservation planning tools”. In: *Annual Review of Environment and Resources* 31 (2006), pages 123–59. DOI: [10.1146/annurev.energy.31.042606.085844](https://doi.org/10.1146/annurev.energy.31.042606.085844) (cited on page 59).
- [149] Fabio Scarpa et al. “Multiple Non-Species-Specific Pathogens Possibly Triggered the Mass Mortality in *Pinna nobilis*”. In: *Life* 10.10 (Oct. 2020), page 238. ISSN: 2075-1729. DOI: [10.3390/life10100238](https://doi.org/10.3390/life10100238) (cited on page 8).
- [150] Kevin Schneider et al. “Impact of *Xylella fastidiosa* subspecies *pauca* in European olives”. In: *Proceedings of the National Academy of Sciences* 117.17 (2020), pages 9250–9259. ISSN: 0027-8424. DOI: [10.1073/pnas.1912206117](https://doi.org/10.1073/pnas.1912206117) (cited on page 45).
- [151] Sandra K. Schumacher and Jeffrey I. Campbell. “Travel Medicine”. In: *Urgent Care Medicine Secrets*. Edited by Robert P. Olympia, Rory M. O’Neill, and Matthew L. Silvis. Elsevier, 2018. Chapter 56, pages 352–357. ISBN: 978-0-323-46215-0. DOI: [10.1016/B978-0-323-46215-0.00056-2](https://doi.org/10.1016/B978-0-323-46215-0.00056-2) (cited on page 45).
- [152] Nita Shah and Gupta Jyoti. “SEIR Model and Simulation for Vector Borne Diseases”. In: *Applied Mathematics* 4 (Jan. 2013), pages 13–17. DOI: [10.4236/am.2013.48A003](https://doi.org/10.4236/am.2013.48A003) (cited on page 49).
- [153] Lillian L. M. Shapiro, Shelley A. Whitehead, and Matthew B. Thomas. “Quantifying the effects of temperature on mosquito and parasite traits that determine the transmission potential of human malaria”. In: *PLOS Biology* 15.10 (10 Oct. 2017), e2003489. DOI: [10.1371/journal.pbio.2003489](https://doi.org/10.1371/journal.pbio.2003489) (cited on page 27).
- [154] I.M Sobol. “Global sensitivity indices for nonlinear mathematical models and their Monte Carlo estimates”. In: *Mathematics and Computers in Simulation* 55.1 (2001), pages 271–280. ISSN: 0378-4754. DOI: [10.1016/S0378-4754\(00\)00270-6](https://doi.org/10.1016/S0378-4754(00)00270-6) (cited on pages 26, 63).
- [155] S. Soubeirand et al. “Inferring pathogen dynamics from temporal count data: the emergence of *Xylella fastidiosa* in France is probably not recent”. In: *New Phytologist* 219.2 (2018), pages 824–836. DOI: <https://doi.org/10.1111/nph.15177>. eprint: <https://nph.onlinelibrary.wiley.com/doi/pdf/10.1111/nph.15177>. URL: <https://nph.onlinelibrary.wiley.com/doi/abs/10.1111/nph.15177> (cited on pages 59, 60).
- [156] J. F. Stevenson et al. “Grapevine Susceptibility to Pierce’s Disease II: Progression of Anatomical Symptoms”. In: *American Journal of Enology and Viticulture* 55.3 (2004), pages 238–245. ISSN: 0002-9254. DOI: [10.5344/ajev.2004.55.3.238](https://doi.org/10.5344/ajev.2004.55.3.238). eprint: <https://www.ajevonline.org/content/55/3/238.full.pdf>. URL: <https://www.ajevonline.org/content/55/3/238> (cited on page 60).
- [157] Yevhen F. Suprunenko, Stephen J. Cornell, and Christopher A. Gilligan. “Analytical approximation for invasion and endemic thresholds, and the optimal control of epidemics in spatially explicit individual-based models”. In: *Journal of The Royal Society Interface* 18.176 (2021), page 20200966. DOI: [10.1098/rsif.2020.0966](https://doi.org/10.1098/rsif.2020.0966) (cited on page 36).
- [158] B. L. Teviotdale and J. H. Connell. “Almond Leaf Scorch”. In: *ANR University of California* (2003), page 8106. DOI: [10.3733/ucanr.8106](https://doi.org/10.3733/ucanr.8106) (cited on pages 60, 63).
- [159] Horst R. Thieme. “Convergence results and a Poincaré-Bendixson trichotomy for asymptotically autonomous differential equations”. In: *Journal of Mathematical Biology* 30.4 (1992), pages 755–763. DOI: [10.1007/BF00173267](https://doi.org/10.1007/BF00173267) (cited on page 49).
- [160] Fiona M. Tomley and Martin W. Shirley. “Livestock infectious diseases and zoonoses”. In: *Philosophical transactions of the Royal Society of London. Series B, Biological sciences* 364.1530 (Sept. 2009), pages 2637–2642. ISSN: 1471-2970. DOI: [10.1098/rstb.2009.0133](https://doi.org/10.1098/rstb.2009.0133) (cited on page 29).
- [161] S. Trigos et al. “Utilization of muddy detritus as organic matter source by the fan mussel *Pinna nobilis*”. In: *Mediterranean Marine Science* 15.3 (2014), pages 667–674. ISSN: 1791-6763. DOI: [10.12681/mms.836](https://doi.org/10.12681/mms.836) (cited on page 8).
- [162] Kabir B. Tumber, Julian M. Alston, and Kate B. Fuller. “Pierce’s disease costs California \$104 million per year”. In: *California Agriculture* 68.1 (2014), pages 20–29. DOI: [10.3733/ca.v068n01p20](https://doi.org/10.3733/ca.v068n01p20). URL: <https://ucanr.edu/repository/fileAccessPublic.cfm?fn=cav6801p20-133590.pdf> (cited on page 45).
- [163] Frank Van den Bosch and Michael J. Jeger. “The basic reproduction number of vector-borne plant virus epidemics”. In: *Virus Research* 241 (2017), pages 196–202. ISSN: 0168-1702. DOI: [10.1016/j.virusres.2017.06.014](https://doi.org/10.1016/j.virusres.2017.06.014) (cited on page 46).
- [164] Pauline van den Driessche. “Reproduction numbers of infectious disease models”. In: *Infectious Disease Modelling* 2.3 (2017), pages 288–303. ISSN: 2468-0427. DOI: [10.1016/j.idm.2017.06.002](https://doi.org/10.1016/j.idm.2017.06.002) (cited on page 45).

- [165] M. Vanhove et al. “Genomic diversity and recombination among *Xylella fastidiosa* subspecies”. In: *Applied and Environmental Microbiology* 85.13 (2019), e02972–18. DOI: [10.1128/AEM.02972-18](https://doi.org/10.1128/AEM.02972-18) (cited on page 59).
- [166] Maite Vázquez-Luis et al. “S.O.S. *Pinna nobilis*: A Mass Mortality Event in Western Mediterranean Sea”. In: *Frontiers in Marine Science* 4 (2017), page 220. ISSN: 2296-7745. DOI: [10.3389/fmars.2017.00220](https://doi.org/10.3389/fmars.2017.00220) (cited on pages 8, 29).
- [167] P. Virtanen et al. “SciPy 1.0: Fundamental Algorithms for Scientific Computing in Python”. In: *Nature Methods* 17 (2020), pages 261–272. DOI: [10.1038/s41592-019-0686-2](https://doi.org/10.1038/s41592-019-0686-2) (cited on page 62).
- [168] Maurizio Vurro, Barbara Bonciani, and Giovanni Vannacci. “Emerging infectious diseases of crop plants in developing countries: impact on agriculture and socio-economic consequences”. In: *Food Security* 2.2 (June 2010), pages 113–132. ISSN: 1876-4525. DOI: [10.1007/s12571-010-0062-7](https://doi.org/10.1007/s12571-010-0062-7) (cited on page 29).
- [169] Jessica R Ward and Kevin D Lafferty. “The Elusive Baseline of Marine Disease: Are Diseases in Ocean Ecosystems Increasing?” In: *PLoS Biology* 2.4 (Apr. 2004). Edited by Larry Crowder, e120. ISSN: 1545-7885. DOI: [10.1371/journal.pbio.0020120](https://doi.org/10.1371/journal.pbio.0020120) (cited on page 7).
- [170] Hui-Ming Wei, Xue-Zhi Li, and Maia Martcheva. “An epidemic model of a vector-borne disease with direct transmission and time delay”. In: *Journal of Mathematical Analysis and Applications* 342.2 (2008), pages 895–908. ISSN: 0022-247X. DOI: [10.1016/j.jmaa.2007.12.058](https://doi.org/10.1016/j.jmaa.2007.12.058) (cited on page 49).
- [171] Curtis L. Wesley and Linda J.S. Allen. “The basic reproduction number in epidemic models with periodic demographics”. In: *Journal of Biological Dynamics* 3.2-3 (2009), pages 116–129. DOI: [10.1080/17513750802304893](https://doi.org/10.1080/17513750802304893) (cited on page 52).
- [172] S. M. White et al. “Modelling the spread and control of *Xylella fastidiosa* in the early stages of invasion in Apulia, Italy”. In: *Biological Invasions* 19.6 (June 2017), pages 1825–1837. ISSN: 1573-1464. DOI: [10.1007/s10530-017-1393-5](https://doi.org/10.1007/s10530-017-1393-5). URL: <https://doi.org/10.1007/s10530-017-1393-5> (cited on page 60).
- [173] S. M. White et al. “Estimating the epidemiology of emerging *Xylella fastidiosa* outbreaks in olives”. In: *Plant Pathology* 69.8 (2020), pages 1403–1413. DOI: <https://doi.org/10.1111/ppa.13238>. URL: <https://onlinelibrary.wiley.com/doi/abs/10.1111/ppa.13238> (cited on pages 60, 62).
- [174] *Fact sheets of vector-borne diseases*. 2018. URL: https://www.hiobs.org/lib/file/manager/Vector-Borne_Disease_General_Fact_Sheet.pdf (cited on page 45).
- [175] Shi Zhao et al. “Modelling the effective reproduction number of vector-borne diseases: the yellow fever outbreak in Luanda, Angola 2015-2016 as an example”. In: *PeerJ* 8 (Feb. 2020), e8601–e8601. ISSN: 2167-8359. DOI: [10.7717/peerj.8601](https://doi.org/10.7717/peerj.8601) (cited on page 49).
- [176] Maria Zouhou et al. “*Pinna nobilis* in the Greek seas (NE Mediterranean): on the brink of extinction?” In: *Mediterranean Marine Science* 21.3 (July 2020), pages 575–591. ISSN: 1791-6763. DOI: [10.12681/mms.23777](https://doi.org/10.12681/mms.23777) (cited on page 8).

