Universitat de les Illes Balears

MASTER'S THESIS

Modelling the Mass Mortality Event of *Pinna nobilis*

Alex Giménez Romero

Master's Degree in Physics of Complex Systems

(Specialization/Pathway in Complex Systems)

Centre for Postgraduate Studies

Academic Year 2019/2020

Modelling the Mass Mortality Event of *Pinna nobilis*

Alex Giménez Romero

Master's Thesis

Centre for Postgraduate Studies

University of the Balearic Islands

Academic Year 2019/2020

Keywords:
Ecology, Epidemiology, Biological Modelling

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

Thesis Supervisor's Name: Dr. Iris Eline Hendriks

To all my family, friends and those who I've ever met,
without whom I wouldn't be who I am.

Acknowledgements

This work would not have been possible without the supervision of Dr. Manuel Matías and Dr. Iris Hendriks, who I would like to thank for their guidance, advisement and patience. Moreover, I would like to thank them for the professional and personal mark they have left on me, which will help me throughout my life.

I am grateful to all of those with whom I have had the pleasure to share this master in Physics of Complex Systems during this year, which will be always in my memories. In particular, I wish to show my gratitude to Ana for the insightful conversations about the topics regarding this thesis, besides all the good moments shared within this year.

I also would be pleased to acknowledge my childhood friends, who have really motivated me in my worst moments and brought me always a smile. I would like to mention Alex Jiménez and Marc Gutiérrez in particular, who conform the science crew of my childhood friends. Last but by no means least, I would like to specially thank Lidia Guerrero who, besides being by my side in the last half of my life, brought this thesis to a higher professional level with her grammar corrections.

A loving mention to Carmela, for all the support and affection she has given me this year despite the distance, for pushing me to pursue my goals staying always by my side and to let me be myself.

Finally, I wish to acknowledge the support and great love of my family.

A. Giménez-Romero

Abstract

Since the early Autumn of 2016 a mass mortality event is devastating the populations of fan mussels, *Pinna nobilis*, in the Mediterranean Sea, where it is an endemic species. This ecological disaster is linked to a Haplosporodian endoparasite, *Haplosporodium pinnae*, being the cause of an epidemic that is about to extinguish one of the largest bivalves of the world. The parasite spreads by surface currents and the disease expression is closely related to temperatures above $13.5\,^{\circ}\mathrm{C}$ and to a salinity range between $36.5 - 39.7\,\mathrm{psu}$. The aim of this project is to study the epidemic by building mathematical and computational models. In particular, we'll use classical compartmental models of differential equations as well as stochastic models described by a master equation. Therefore, we propose a first step for the mathematical modelling of the mass mortality event taking place in the Mediterranean Sea as a basis for future developments.

Resumen

Desde principios de otoño de 2016, un evento de mortalidad masiva está devastando las poblaciones de mejillones *Pinna nobilis* en el mar mediterráneo, donde es una especie endémica. Este desastre ecológico está relacionado con un endoparásito Haplosporodian, *Haplosporodium pinnae*, que es la causa de una epidemia que está a punto de extinguir a uno de los bivalvos más grandes del mundo. El parásito se propaga por las corrientes superficiales y la expresión de la enfermedad está estrechamente relacionada con temperaturas superiores a $13.5\,^{\circ}\mathrm{C}$ y a un rango de salinidad entre $36.5 - 39.7\,\mathrm{psu}$. El objetivo de este proyecto es estudiar la epidemia mediante la construcción de modelos matemáticos y computacionales. En particular, utilizaremos modelos compartimentales clásicos de ecuaciones diferenciales, así como modelos estocásticos descritos por una ecuación maestra. Por lo tanto, proponemos un primer paso para el modelado matemático del evento de mortalidad masiva que tiene lugar en el Mar Mediterráneo como base para futuros desarrollos.

Abbreviations

LHS Left Hand Side

MF Mean-Field

MME Mass Mortality Event

NGM Next Generation Matrix

RHS Right Hand Side

MSE Mean Squared Error

Contents

1 Introduction

Coastal ecosystems are of great importance in the biosphere, either for ecological reasons, such as their great biodiversity, ore more anthropological reasons, such as the capture of CO_2 by the marine vegetation that conforms these systems or even economic reasons, like their attractiveness to tourists in the case of the Balearic Islands. These ecosystems are being increasingly affected by the global impact of human activities, i.e. the global change, which determines ecosystem functioning, structure and services they provide to society. Global warming is one of the most common anthropogenic pressures in the Mediterranean Sea [1] . In fact, the natural semi-enclosured nature of the Mediterranean Sea makes it especially vulnerable to the anthropogenic pressure of global change, leading it to be one of the ocean areas with faster warming [2]. Global warming, together with the degradation of natural habitats, eutrophication, hypoxia, acidification and biological invasions has been identified as the biggest cause of change in the biodiversity of the Mediterranean Sea since the last century [2–4]. Thus, coastal ecosystems are being severely affected by global change, understood as the set of human activities that have an impact on the biosphere, for example global warming.

The fan mussel (or pen shell) *Pinna nobilis* is the largest endemic bivalve in the Mediterranean Sea and plays a key ecological role in these ecosystems. It occurs at depths between 0.5 and 60 m, mostly on soft-bottom coastal areas and mainly inhabits seagrass meadows, typically of *Posidonia oceanica* [5, 6]. Its lifespan is up to 50 years in favourable conditions and its size can get up to 1.2 m, placing it among the largest bivalves of the world [7]. These fan mussels play a crucial ecological role in their habitat. *P. nobilis* individuals filter water retaining a large amount of organic matter from suspended detritus, contributing to water clarity. Their shell provide a hard-surface within a soft bottom ecosystem, which can be colonised by different benthic species [7], augmenting biodiversity. They are also an important piece of the trophic web, serving as prey for other species, like the common octopus (*Octopus vulgaris*) [8] and are host to symbionts like the crustaceans *Pontonia pinnophylax* and *Nepinnotheres pinnotheres*.

Nowadays, besides the anthropogenic pressures causing the decline of fan mussel population, like habitat destruction and illegal fishing, a Mass Mortality Event (MME) is causing a devastating reduction of this species, which caused a status change from the "Vulnerable" category to "Critically Endangered" with a serious extinction risk (Orden TEC/1078/2018) [7]. There is evidence that the main cause of this mortality is the protozoan *Haplosporodium pinnae*, a new species that belong to the genus Haplosporidium, one of the four genera of the protist order Haplosporida [6], where it has been found that other Haplosporidian parasites are behind the extensive mortality of several oyster species. Life stages include uninucleate 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, which could cause the extinction of the species. As a consequence, Spanish authorities changed the status of P. nobilis from "endangered" to "endangered with extinction" along the Spanish coasts (Orden TEC/ 596/2019) [9]. The follow-up has helped to better understand the spread of the disease, with surface currents being 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 [7].

The MME of *Pinna nobilis* is basically caused by an infectious disease produced by the parasite *Haplosporodium pinnae*. The modern approach to the mathematical study of infectious diseases came by the hand of W.O. Kermack and A.G. McKendrick in a series of three papers published between 1927 and 1933 [10–12]. The basic compartmental models to describe the transmission of communicable diseases were reported in the first of these papers and what is often called the Kermack-McKendrick epidemic model, e.g. SIR model, is actually a special case of the general model introduced in this paper. Basically the models consist in dividing the population into different compartments or states and allow the individuals to move from one compartment to another following some dynamical rules. A key point of these models is that the disease is transmitted by direct contact between infected individuals. However, in a typical marine infectious disease of sessile individuals, like fan mussels, the infection process can not be described by direct contact of infected individuals, as they can't move.

Instead, the pathogen can either be waterborne or vectorborne, which means, spread by marine currents or mobile organisms that carry the pathogen from one host to another [13]. Thus, there is an explicit need, at least at first, to model the parasite population and not only the host individuals.

In this master's thesis we propose a mathematical modelling of this marine infectious disease to better understand the MME and quantify the future development of the epidemic. The mathematical model will take into account the biological and epidemiological information presented above. The work will be structured as follows: First, a theoretical introduction to mathematical epidemiology will be carried out in Section 2, where the deterministic and stochastic approaches of epidemiological compartmental models are discussed. In Section 3, a model is proposed to study the epidemic in mathematical terms, based on both deterministic and stochastic approaches, and an analytical study is performed. A numerical analysis of the proposed model is presented in Section 4, so that the obtained analytical results can be contrasted and further validated. After that, in Section 5 we try to validate our model by contrast with experimental data, but due to the lack of data only a simplified version of our model can be fitted to the available experimental data. Finally, our conclusions are presented in Section 6 together with the proposed future research line in Section 7.

2 Theoretical background

2.1 Compartmental models in epidemiology

A classical technique to model mathematically infectious diseases are the so-called compartmental models. It consist in dividing the population into different compartments or states and allow the individuals to move from one compartment to another following some dynamical rules. The key assumptions of these models are about the nature and the time rate of transfer from one state to another and the spatial distribution or connectivity of the individuals.

There are different kinds of compartmental models regarding different kind of diseases. Two of the most common ones are the SIS and SIR models. The first one stands for *Susceptible-Infected-Susceptible*, which is the cycle the individuals follow in the development of the disease. It is used to model diseases where the individuals can recover without immunity, so being susceptible again, as for instance the flu (in a large term point of view. The second one is referred as the *Susceptible-Infected-Recovered/Removed* model and it is used for the modelling of infection diseases where the individuals can be recovered, obtaining immunity to the disease, or can die. In general, we can consider the individuals being in the R state as non-infectious non-susceptible individuals. For instance, if we considered that after a given period of time infected individuals can not infect anymore, due to any reason, they can be considered to be in the R state regardless on their actual biological state. In Fig. 1 a graphical representation of the SIR model is provided. Here, β and γ are the infection and recovery rates, so the probability per unit time to change from the susceptible to the infected compartment and from the infected to the recovered compartment, respectively.

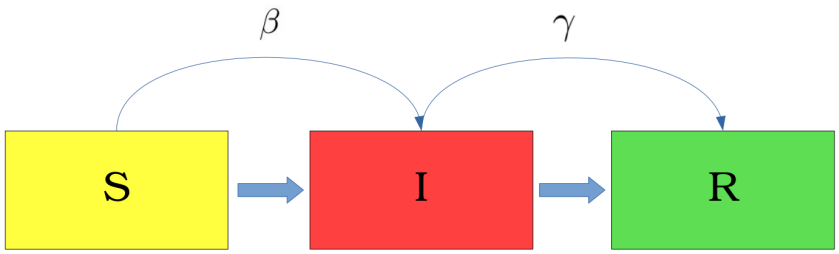


Figure 1: Graphical representation of the SIR model

Other compartmental models take into account more subtle and precise features of some diseases, such that the $SEIR$ or $SEIS$ models, where there is an exposure state between being infected and becoming infective or the $SIRS$ models, with temporary immunity on recovery from infection [14].

These compartmental models can be treated analytically and numerically from two basic points of view, the deterministic and the stochastic approach. The deterministic approach accounts for the mean trend of the process and it is appropriate when we are dealing with large populations and the stochastic effects can be neglected. On the other hand, if the population is small enough or some variable of the model follow some probability distribution, then the stochastic effects become important so that a probabilistic approach is needed.

In the following sections both treatments will be discussed taking as example the SIR model, as the model object of this thesis will keep some similarities with the SIR model.

2.2 Deterministic approach

The model describe the change with time of the populations of the different compartments, that are written in terms of rates of the processes through which the populations change compartments. As a result, our models are formulated initially as systems of differential equations [14]. The hypothesis to construct the model are the following:

1. The various states are uniformly mixed, that is, every pair of individuals has equal probability of coming into contact with one another.
2. An average member of the population makes contact sufficient to transmit infection with βN others per unit time, where N represents total population size.
3. Infectives leave the infective state at rate αI per unit time.
4. The incubation period is short enough to be negligible; that is, a susceptible who contracts the disease is infective right away.
5. There is no entry into or departure from the population, no birth or natural deaths are assumed, so that the population remains constant.

Note that the first two hypothesis basically assume that all individuals are connected between them which is also known as the mean field assumption. According to the mean field approximation, the probability that an infected individual contacts a susceptible one is given by S/N. Then, the number of new infections will be given by the probability that an infected individual contacts a susceptible one multiplied by the actual number of infected individuals and the mean number of infections, $S/N \cdot I \cdot \beta N = \beta SI$. This leads to the following description of the model given by a system of differential equations,

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

Although this system of differential equation is not fully solvable in general, some analytical calculations can be performed with it. The approach followed in the present review is based in reference [15].

First, note that our assumption of population conservation is inherently included in the model,

$$\dot{S} + \dot{I} + \dot{R} = 0 \Longrightarrow S + I + R = \text{const.} = N \ . \tag{2.2}$$

Then, by using the fact that the initial conditions must be $S(0) > 0$, $I(0) > 0$ and $R(0) = 0$ a threshold for the develop of an epidemic can be found. It is straight forward to see that a condition for the develop of an epidemic is nothing but an increase of the number of infected individuals at the beginning of the situation, $\left.\frac{\mathrm{d}I}{\mathrm{d}t}\right|_{t=0} > 0$,

$$\left.\frac{\mathrm{d}I}{\mathrm{d}t}\right|_{t=0} = I_0\left(\beta S_0 - \gamma\right) \Longrightarrow \begin{cases} \left.\frac{\mathrm{d}I}{\mathrm{d}t}\right|_{t=0} < 0 \Longleftrightarrow S_0 < \frac{\gamma}{\beta} \\ \\ \left.\frac{\mathrm{d}I}{\mathrm{d}t}\right|_{t=0} > 0 \Longleftrightarrow S_0 > \frac{\gamma}{\beta} \end{cases} . \tag{2.3}$$

Eq. (2.3) defines a threshold for the developing of an epidemic in the SIR model,

$$S_c = \frac{\gamma}{\beta} \equiv \rho \ . \tag{2.4}$$

The critical parameter $\rho = \gamma/\beta$ is sometimes called the *relative removal rate* and its reciprocal, $\sigma = \beta/\gamma$, the *infection's contact rate*. This threshold behaviour can be further simplified by considering the so-called *basic reproduction number*, R_0, defined as,

$$R_0 = \frac{S_0}{\rho} = S_0\sigma = S_0\frac{\beta}{\gamma} \ , \tag{2.5}$$

which measures the number of secondary infections given by a primary infection in a whole susceptible population.

It can be proved that this quantity defines a threshold for the develop of the epidemic from the fact that $\dot{S}$ is a monotonically decreasing function, which implies $S(t) < S_0$. Then, it yields the following relation,

$$\frac{\mathrm{d}I}{\mathrm{d}t} = I\left(\beta S - \gamma\right) \leq I\left(\beta S_0 - \gamma\right) = \gamma I\left(\frac{S_0}{\rho} - 1\right) = \gamma I\left(R_0 - 1\right) \ , \tag{2.6}$$

so that for $R_0 < 1$, $\dot{I} < 0 \ \forall t$ and thus $I_0 > I(t)$ as $t \to \infty$, which basically means that the epidemic dies out, while for $R_0 > 1$ the epidemic grows.

This threshold behaviour agrees with our intuition given the definition of the basic reproduction number: if a primary infection produces more than 1 secondary infection an epidemic will develop, while if it does not reach to infect at least one individual it will die out. Furthermore, note that because of our mean field approach, the number of secondary infections produced by a primary one refers to the *average* number of secondary infections.

Another important analytical result that can be obtained from this model is the maximum of infected individuals, that gives an idea of how severe the epidemic will be $\dot{I} = 0 \Longrightarrow S(t) = \rho$. So the maximum of infected individuals, given the development of a proper epidemic, will take place when $S = \rho$.

Now, lets derive the phase plane trajectories for S and I. Dividing the differential equations for S and I (considering $I \neq 0$)) we obtain,

$$\frac{\mathrm{d}I}{\mathrm{d}S} = -1 + \frac{\gamma}{\beta S} = -1 + \rho/S \quad (I \neq 0) \ .$$

Integrating the relation,

$$\int_{I_0}^{I} \mathrm{d}I = \int_{S_0}^{S} \mathrm{d}S\left\{-1 + \rho/S\right\} \Longrightarrow I - I_0 = S_0 - S + \rho\ln\left(\frac{S}{S_0}\right) \ ,$$

we obtain the phase plane trajectories for S and I given by,

$$I + S = I_0 + S_0 + \rho\ln\left(\frac{S}{S_0}\right) = N + \rho\ln\left(\frac{S}{S_0}\right) \ , \tag{2.7}$$

as clearly $I_0 + S_0 = N$ given that $R(0) = 0$.

Finally we find an expression for the maximum of infected individuals by substituting the condition for the maximum in the last expression, Eq. (2.7),

$$I_{max} = N + \rho\left[\ln\left(\frac{\rho}{S_0}\right) - 1\right] \ . \tag{2.8}$$

Now we can find the final number of susceptible individuals, by dividing the differential equations for S and R,

$$\frac{\mathrm{d}S}{\mathrm{d}R} = -\frac{\beta}{\gamma}S = -\frac{S}{\rho} \Longrightarrow \int_{S_0}^{S}\frac{\mathrm{d}S}{S} = -\frac{1}{\rho}\int_{R(0)=0}^{R}\mathrm{d}R \Longrightarrow \ln\left(\frac{S}{S_0}\right) = -\frac{R}{\rho} \ .$$

As $I(\infty) = 0$, necessarily $R(\infty) = N - S(\infty)$ so that we get the following expression for the final number of susceptible individuals,

$$S_\infty - \rho\ln(S_\infty) = N - \rho\ln(S_0) \ . \tag{2.9}$$

S_∞ is nothing but the positive root of the transcendental equation Eq. (2.9). In fact, this transcendental equation can be solved by means of the Lambert's W function (see Appendix B for more information about the Lambert's W function),

$$S_\infty = -\rho \cdot W_0\left[-\frac{S_0}{\rho}\, e^{-N/\rho}\right] = -\rho \cdot W_0\left[-R_0\, e^{-N/\rho}\right] \ . \tag{2.10}$$

The Lambert's $W_l(x)$ function is bivalued for $x \in (-1/e, 0)$, so that it is necessary to choose a branch for such values of x. In this epidemiological context it has been already shown that the branch $l = 0$ has to be chosen when $S(t) < S_c$ [16], and clearly $S_\infty < S_c$. With this analysis we ensure that the final number of susceptible individuals is a strictly positive number, $S_\infty > 0$, given the properties of the Lambert's function.

The second hypothesis in the model of a rate of contacts per infective which is proportional to population size N, called mass action incidence, was used in all the early epidemic models. However, it is quite unrealistic. It is more realistic to assume a contact rate which is a non-increasing function of total population size. For instance, a situation in which the number of contacts per infective in unit time is constant, called standard incidence, is a more accurate description for contact transmitted diseases. This leads to the following system of differential equations,

$$\begin{aligned} \dot{S} &= -\beta SI/N \\ \dot{I} &= \beta SI/N - \gamma I \\ \dot{R} &= \gamma I \ . \end{aligned} \tag{2.11}$$

Note that with the standard incidence choice the system of differential equations can be written in the same form that Eq. (2.1) by considering the density of individuals instead of their absolute number. Dividing both sides of all the equations by the population size, N, we get,

$$\begin{aligned} \dot{S} &= -\beta SI \\ \dot{I} &= \beta SI - \gamma I \\ \dot{R} &= \gamma I \ , \end{aligned} \tag{2.12}$$

where now S, I and R are the density of susceptible, infected and recovered/removed individuals instead of their absolute numbers.

2.3 Stochastic approach

The underlying assumption of the deterministic model presented above can be simplified by saying that the sizes of the compartments are large enough that the mixing of the different members is homogeneous. This assumptions might be reasonable when the epidemic is well underway, but at the beginning of the disease outbreak the situation may be quite different. In essence, at the beginning of the epidemic the number of infected individuals is low so that the evolution of infected individuals strongly depends on the contact pattern between members of the population. Moreover, transmission of and recovery from an infection are intrinsically stochastic processes, and the deterministic SIR model does not account for fluctuations. These fluctuations are particularly important at the beginning of an epidemic when the number of infecteds is very small. In this regime, a probabilistic description must be used.

In the SIR model the stochastic infection dynamics is given by the following scheme,

$$S + I \xrightarrow{\beta} 2I \qquad I \xrightarrow{\gamma} R \ . \tag{2.13}$$

The reaction equations in Eq. (2.13) indicate that an encounter of a susceptible individual with an infected one results in two infected individuals at a probability rate β while the infected individuals recover (or are removed) at rate γ. Also the recovery (removal) of individuals can be treated as $I \xrightarrow{\gamma} \varnothing$. The quantity of interest is the probability $p(n_S, n_I, n_R; t)$ of finding a number S

of susceptibles, I of infecteds and R of recovered (removed) in a population of size N at time t. However, as the population is conserved one of the terms is redundant, i.e $R = N - S - I$. Then we can simplify our quantity of interest to $p(n_S, n_I; t)$.

Assuming that the process is Markovian on the relevant time scales, the dynamics of this probability is governed by the following master equation [1],

$$\frac{\partial p(n_S, n_I; t)}{\partial t} = \sum_k \left(E^l - 1\right) \left[C_l((n_S, n_I))p(n_S, n_I; t)\right] =$$

$$= \left(E_s {E_I}^{-1} - 1\right) \left[\beta \frac{n_S n_I}{N} p(n_S, n_I : t)\right] + \left(E_I - 1\right) \left[\gamma n_I p(n_S, n_I; t)\right] =$$

$$= \beta \frac{(n_S + 1)\,(n_I - 1)}{N} p(n_S + 1, n_I - 1; t) + \gamma(n_I + 1)p(n_S, n_I + 1; t) - \left(\beta \frac{n_S n_I}{N} + \gamma n_I\right) p(n_S, n_I; t) \ ,$$

where $C_l((n_S, n_I)) \equiv \Omega((n_S, n_I) \to (n_{S'}, n_{I'}))$.

This equation is not exactly solvable analytically but it could be solved numerically, for instance with the so-called Gillespie algorithm [18]. Moreover, one can make use of some approximate methods to find the evolution of the system in the limiting case where the population is large. With the mean field approach it is possible to obtain very easily approximated equations for the first moments of the probability $p(n_S, n_I; t)$. For instance, the evolution of the first moment is given by the following exact equation [17],

$$\frac{\mathrm{d}\left\langle n_i \right\rangle}{\mathrm{d}t} = -\sum_\ell \ell_i \left\langle C_{\ell_i}((n_S, n_I))\right\rangle \ . \tag{2.14}$$

Then, applying Eq. (2.14) to the SIR model we get,

$$\begin{cases} \dfrac{\mathrm{d}\left\langle n_S \right\rangle}{\mathrm{d}t} = -\dfrac{\beta}{N} \left\langle n_S n_I \right\rangle \\[2ex] \dfrac{\mathrm{d}\left\langle n_I \right\rangle}{\mathrm{d}t} = \dfrac{\beta}{N} \left\langle n_S n_I \right\rangle - \gamma \left\langle n_I \right\rangle \ . \end{cases}$$

Now we perform the mean-field approximation by considering that N is large so that the fluctuations are small, $\sigma^2(n_i) = \left\langle {n_i}^2 \right\rangle - \left\langle n_i \right\rangle^2 \approx 0$. Then we can approximate $\left\langle {n_i}^2 \right\rangle \approx \left\langle n_i \right\rangle^2$, or similarly $\left\langle n_i n_j \right\rangle \approx \left\langle n_i \right\rangle \left\langle n_j \right\rangle$. It is convenient to consider the density of individuals, $i = n_i/N$, as it is a fixed quantity. With all together we have,

$$\begin{cases} \dfrac{\mathrm{d}S}{\mathrm{d}t} = -\beta SI \\[2ex] \dfrac{\mathrm{d}I}{\mathrm{d}t} = \beta SI - \gamma I \ , \end{cases}$$

which is the deterministic SIR model where S, I are the density of susceptible and infected individuals, respectively. The number of recovered (removed) individuals then could be computed as $R = N - S - I$.

[1] For more information about the derivation of the master equation see[17]

3 Modelling the fan mussel epidemic

3.1 Model considerations

The phenomenon we will try to model here has a particular difference with the ones described in the previous section. The *Pinna nobilis* individuals that we aim to describe with our model, don't get infected by direct contact with other infected individuals, as fan mussels are sessile organisms, and do not move after an initial larval stadium. Instead, the individuals get infected by filtering water containing the parasite, *Haplosporodium pinnae*. Thus, there is a need to explicitly model the parasite population in the medium. To properly model the parasite and fan mussel populations some key biological aspects must be considered in the model.

The hosts, fan mussels, as invertebrates, lack adaptive immunity [19]. Nevertheless, they have innate immunity, mainly through hemocytes, so that they can combat the infection by parasites to some extent. Thus, infected individuals can return to the susceptible state if they recover from infection by beating the parasite. Infected fan mussels can be understood as "parasite factories", as there is an excretion process of parasites from infected hosts to the medium by their respiratory and excretory system. In general, this excretion process depends on several other processes: the parasite proliferation/reproduction rate, the parasite death rate inside hosts (that could be different to the death rate in the outer medium), some kind of excretion efficiency, etc. Moreover, the excreted parasite population would also depend on the current population inside the hosts, which can change over time. However we consider that all the sub-processes inside the infected hosts are in equilibrium and the excretion process is finally given by an effective rate.

We assume parasites can only reproduce inside fan mussels, while they can die both inside and outside the hosts. Moreover, the parasites can be found in different states or phases, such as uninucleate cells, binucleate cells, multinucleate plasmodia or spores. Uninucleate and binucleate stages are mostly found in the connective tissue while scarcity of plasmodia phases is observed. This suggests that parasite proliferation to reach systemic infection involves karyokinesis in the uninucleate cells producing binucleate cells followed by cytokinesis resulting in two daughter uninucleate cells. A sporulation stage is exclusively found in the digestive gland tubules so that migration of uninucleate and binucleate cells from the connective tissue towards digestive gland tubules can be deduced [6].

These different phases will not be considered in our models, for the sake of simplicity, given the following reasons. First, the distinction between uninucleate and binucleate cells seems unnecessary, as these phases only participates in the parasite proliferation inside infected hosts, a process that we consider in an effective way. Second, the evolution time scale of the disease is much faster than the typical lifetime of fan mussels so no vital dynamics will be considered. Thus, the sporulation phase is irrelevant to our study, as it only affects the growth of new individuals [9].

In the following sections a mathematical model is presented, together with its analytical analysis, taking into account the biological aspects discussed above.

3.2 4D deterministic model

The first step to model the epidemic is to consider a vast zone of substrate with several fan mussels attached in a random fashion. To study the evolution of the fan mussel population, we consider a 4-compartmental model according to the following reaction processes:

$$S + P \xrightarrow{\beta} I \quad I \xrightarrow{\gamma} R \quad I \xrightarrow{\alpha} S \quad I \xrightarrow{\lambda} I + P \quad P \xrightarrow{\mu} \varnothing \,, \tag{3.1}$$

which is graphically summarised in the following figure

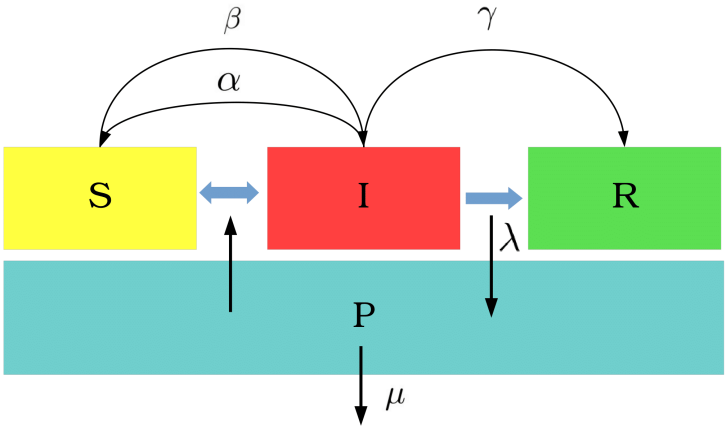


Figure 2: Graphical scheme for the 4D deterministic model

According to the scheme in Fig. 21 we consider the host in 3 possible states: susceptible, S, infected by the parasite, I and removed (death), R. Then we introduce the parasite population in the medium (water), P. Regarding the parameters, β is the infection rate of susceptible hosts, γ the mortality of infected hosts, α the recovery rate of infected hosts, λ the production rate of parasites from infected hosts and μ the death rate of parasites. γ and α represent also the inverse of the typical mean time for the host to die and recover (return to susceptible state, as invertebrates lack an adaptive immunity [19]), respectively, and μ the inverse of the typical life time of the parasite. We assume standard incidence, so that the number of contacts per infected individual in unit time is constant. This means that the number of fan mussels that are infected per unit of parasite is the same as the total size of the population increases. This is to be contrasted with an alternative choice, called mass action incidence, in which doubling, say, the populations also doubles the number of fan mussels that are infected. We consider no vital dynamics for the hosts, i.e. the sum of the 3 host populations is constant, as the evolution time scale of the disease is much faster than the typical life cycle of fan mussels. We consider in this simplified model that the parasites only develop inside the host. The model is based on [20], considering that the recovered hosts can be infected again immediately, as discussed in Section 3.1. As stated in Section 1, the expression of the disease strongly depends on temperature and salinity. Thus, we implicitly assume that the parameters of our model, that are kinetic constants, contain this dependence, so we won't be showing it explicitly.

In order to build the deterministic model we consider that the population is large enough to neglect fluctuations and that it is fully mixed. Then, the probability to find a susceptible individual is S/N so that the average number of contacts between a parasite and susceptible Pen-Shell is given by PS/N. Then, the average number of infected individuals will be given by $\beta PS/N$. With that, the change in the number of susceptible Pen-Shells will be given by $\dot{S} = -\beta PS/N$. We note that it is convenient to express the equations in terms of the density of Pen-Shells individuals, that leads to consider the ratio of the parasite to the Pen-Shell population, $\dot{\rho_S} = -\beta f_P \rho_S$. Notice that with the alternative mass-action formulation, one would have $S = -\beta PS$ instead, where S and P are the number of individuals and not their respective densities.

Following this argumentation one can write the following deterministic $SIRP$ model,

$$\begin{aligned} \dot{S} &= -\beta PS + \alpha I \\ \dot{I} &= \beta PS - \gamma I - \alpha I \\ \dot{R} &= \gamma I \\ \dot{P} &= \lambda I - \beta PS - \mu P \ . \end{aligned} \tag{3.2}$$

Notice that in this expression, written for a standard incidence, S, I, R represent densities and no longer populations as in the discussion above and P is the ratio of the parasite to the Pen-Shell

population.

Initially the R and I population are non-existing, so that $R_0 = I_0 = 0$. Then, we note that initially the I disease compartment increases while the P disease compartment decreases, regardless of the value of the parameters,

$$\frac{\mathrm{d}I}{\mathrm{d}t}\Big|_{t=0} = \beta P_0 S_0 > 0$$

$$\frac{\mathrm{d}P}{\mathrm{d}t}\Big|_{t=0} = -P_0\left(\beta S_0 + \mu\right) < 0 \ .$$

Qualitatively, it is reasonable to think that for a small initial condition in P, $P_0 \ll 1$, it is necessary a change of sign in $\dot{P}$ for the develop of an epidemic. In other words, if the parasite population P is a monotonically decreasing function of time and its initial population is small, although there will be an increase of the infected individuals at first, this will be negligible and no epidemic will take place. On the other hand, for a large enough initial parasite population P_0 a non negligible fraction of the population will be affected even though the parasite population decreases monotonically. In simpler terms, while the parasite population is decreasing form its initial large value the susceptible individuals will get infected anyway.

Then, the threshold can be defined from the study of the fixed point obtained from the intersection of the nullclines $\dot{I} = 0$ and $\dot{P} = 0$, considering $P_0 \ll 1$, as is done similarly in [21]. In this threshold the two forces, infection and recovery, for I, and creation, use and death of parasites, for P, balance. That is,

$$I = \frac{\beta S_0}{\gamma + \alpha} P \qquad\qquad I = \frac{\beta S_0 + \mu}{\lambda} P \ ,$$

for the two nullclines, respectively, from which we obtain the threshold R_0 that will happen when the 2 nullclines meet,

$$R_0 = \frac{\beta \lambda S_0}{(\gamma + \alpha)(\beta S_0 + \mu)} \ . \tag{3.3}$$

We can recover this, in principle non-rigorous, result applying the Next Generation Matrix (NGM) method (see Appendix A).

For our system we have,

$$\mathcal{F} = \begin{pmatrix} \beta P S \\ 0 \end{pmatrix} \qquad \mathcal{V} = \begin{pmatrix} (\gamma + \alpha) I \\ -\lambda I + \beta P S + \mu P \end{pmatrix} \ ,$$

where the first entry corresponds to the compartment I and the second to the compartment P.

Then, the corresponding matrices are,

$$F = \begin{pmatrix} 0 & \beta S_0 \\ 0 & 0 \end{pmatrix} \qquad V = \begin{pmatrix} \gamma + \alpha & 0 \\ -\lambda & \beta S_0 + \mu \end{pmatrix} \qquad V^{-1} = \begin{pmatrix} \dfrac{1}{\gamma + \alpha} & 0 \\ \dfrac{\lambda}{(\alpha + \gamma)\left(\beta S_0 + \mu\right)} & \dfrac{1}{\beta S_0 + \mu} \end{pmatrix} .$$

Note that the element $(V^{-1})_{ij}$ is the expected time that an individual who presently has state j will spend in state i during its entire future "life" (in the epidemiological sense), as stated in [22]. Then, the next generation matrix (the next generation matrix with large domain in [22]) is given by,

$$K = FV^{-1} = \begin{pmatrix} \dfrac{\lambda\beta S_0}{(\alpha+\gamma)(\beta S_0+\mu)} & \dfrac{\beta S_0}{\beta S_0+\mu} \\ 0 & 0 \end{pmatrix} \Longrightarrow R_0 = \frac{\lambda\beta S_0}{(\alpha+\gamma)\,(\beta S_0+\mu)} \ ,$$

from which we directly obtain the dominant eigenvalue as the threshold.
Note that the obtained expression for R_0 can be simplified as follows,

$$R_0 = \frac{\lambda}{\Gamma\,(1+\nu)} \ , \tag{3.4}$$

where we have defined $\nu \equiv \mu S_0/\beta$ and $\Gamma \equiv \alpha + \gamma$.

Intuitively, Γ is the rate at which infected individuals leave the I compartment while ν is the ratio between the death rate and infection rate of the parasite multiplied by the initial susceptible population. With this, we see that the epidemic will depend basically on these three parameters.

Other quantities of interest to be obtained from the model are, for instance, the final density of susceptible individuals or the epidemic peak, given by the maximum density of infected individuals, i.e. the maximum of the $I(t)$ curve. Unfortunately, in this 4-D model with 5 parameters this features are nearly impossible to find analytically. For this reason, a reduction to a 3D model is presented in the next section.

3.3 3D deterministic model

As we have seen in the previous section, the 4D-deterministic model considered is hard to analyse analytically due its phase space dimensionality, and high number of parameters. Thus, here is proposed a reduction to a 3D model considering a separation of time scales. Following [23], we consider that the parasite population evolves in a much faster time scale than the hosts, and adjusts its value almost instantaneously to changes in the dynamics of the different host compartments, what implies that one can consider $\dot{P} = 0$ to a very good approximation. This approximation yields,

$$\dot{P} = 0 \Longrightarrow P = \frac{\lambda}{\beta S + \mu} I \ . \tag{3.5}$$

The substitution of the above relation into Eq. (3.2) leads to the following reduced SIR model with a Michaelis-Menten like (or Monod) like saturation term,

$$\begin{aligned} \dot{S} &= -\frac{\lambda S I}{S+\kappa} + \alpha I \\ \dot{I} &= \frac{\lambda S I}{S+\kappa} - \gamma I - \alpha I \\ \dot{R} &= \gamma I \ , \end{aligned} \tag{3.6}$$

where $\kappa \equiv \mu/\beta$.

Note that, now, a non-zero initial number of infected individuals is required in order to have the non-trivial solution for the system of differential equations. This initial number of infected individuals is related to the initial population of the parasite of the previous model by Eq. (3.5), which leads to,

$$I_0 = \frac{P_0\,(\beta+\mu)}{\lambda + \beta P_0} \ . \tag{3.7}$$

As in the previous model, using NGM method the basic reproduction number can be found analytically. Here we just have a single disease compartment so that the calculus is even simpler than before,

$$\mathcal{F} = \frac{\lambda S I}{S+\kappa} \ , \qquad \mathcal{V} = (\alpha+\gamma)\,I \quad \Longrightarrow \quad F = \frac{\lambda S_0}{S_0+\kappa} \ , \qquad V = \gamma + \alpha \ .$$

Then, the basic reproduction number is

$$R_0 = FV^{-1} = \frac{\lambda S_0}{(\gamma+\alpha)(S_0+\kappa)} \tag{3.8}$$

Note that substituting $\kappa = \mu/\beta$ we recover the old relation for the basic reproduction number in Eq. (3.3).

However, considering that $S(t)$ is a monotonically decreasing function[2], the threshold value S_c can be obtained as in the original SIR model,

$$\frac{\mathrm{d}I}{\mathrm{d}t} = I(\alpha+\gamma)\left[\frac{\lambda S}{(\alpha+\gamma)(S+\kappa)} - 1\right] < I(\alpha+\gamma)\left[\frac{\lambda S_0}{(\alpha+\gamma)(S_0+\kappa)} - 1\right] = I(\alpha+\gamma)\left[R_0 - 1\right] , \tag{3.9}$$

as $S(t) < S_0 \ \forall t$.

Thus, if $R_0 = \dfrac{\lambda S_0}{(\alpha+\gamma)(S_0+\kappa)} < 1$ then $\dfrac{\mathrm{d}I}{\mathrm{d}t} < 0 \ \forall t$ and the epidemic dies out. This directly defines the basic reproduction number, which coincides with the obtained by the NGM method in Eq. (3.8).

Moreover, we can rearrange the inequality that defines the basic reproduction number to get the threshold value S_c for which if $S_0 > S_c$ the epidemic dies out while for $S_0 < S_c$ there is an outbreak. This gives rise to the following expression,

$$S_c = \frac{\kappa(\alpha+\gamma)}{\lambda-\alpha-\gamma} . \tag{3.10}$$

Note that from the inequalities $S_0 < S_c$ and $S_0 > S_c$, that define the existence or not of an epidemic one can also obtain the "pseudo basic reproduction number",

$$\tilde{R}_0 = \frac{S_0}{S_c} = \frac{S_0(\lambda-\alpha-\gamma)}{\kappa(\alpha+\gamma)} . \tag{3.11}$$

We call this quantity pseudo basic reproduction number as it can not be linked to the number of secondary infections that produces a primary infection, but it is still a valid quantity for defining the threshold. In fact, both expressions, Eq. (3.8) and Eq. (3.11), are equal when both are 1, $R_0 = \tilde{R}_0 = 1$ as they are obtained from the same first inequality, Eq. (3.9). This can be shown easily in a single line,

$$\tilde{R}_0 = \frac{S_0(\lambda-\alpha-\gamma)}{\kappa(\alpha+\gamma)} = 1 \Longrightarrow \lambda S_0 = (S_0+\kappa)(\alpha+\gamma) \Longrightarrow \frac{\lambda S_0}{(\alpha+\gamma)(S_0+\kappa)} = 1 = R_0 .$$

With this simpler 3-D model some quantities of interest can be computed analytically. First consider the maximum of infected individuals, that will occur when the derivative of I with respect to the time vanishes,

$$\frac{\mathrm{d}I}{\mathrm{d}t} = 0 \Longrightarrow S^* = S_c = \frac{\kappa(\alpha+\gamma)}{\lambda-\gamma-\alpha} . \tag{3.12}$$

Note that this relation show that the model is structurally unstable for $\lambda = \gamma + \alpha$ and it requires $\lambda > \gamma + \alpha$ for physical sense, which can also be argued from the pseudo basic reproduction number formula.

Then, the phase plane trajectories $S - I$ can be computed analytically giving rise to an equation for I in terms of S,

$$\frac{\mathrm{d}I}{\mathrm{d}S} = \frac{S(\lambda-\gamma-\alpha)-\kappa(\gamma+\alpha)}{S(\alpha-\lambda)+\alpha\kappa} \Longrightarrow \int_{I_0}^{I} \mathrm{d}I = \int_{S_0}^{S} \frac{S(\lambda-\gamma-\alpha)-\kappa(\gamma+\alpha)}{S(\alpha-\lambda)-\alpha\kappa}\mathrm{d}S ,$$

[2]Note that, in principle, there is the condition $\alpha < \lambda S/(S+\kappa)$ to fulfil S being monotonically decreasing

where the solution of the integrals yield,

$$I = I_0 + \frac{\gamma\kappa\lambda \ln\left(\frac{S\,(\lambda-\alpha)-\alpha\kappa}{S_0\,(\lambda-\alpha)-\alpha\kappa}\right) - (S-S_0)(\alpha-\lambda)(\alpha+\gamma-\lambda)}{(\alpha-\lambda)^2} \ . \tag{3.13}$$

Thus, substituting Eq. (3.12) into Eq. (3.13) the maximum of infected individuals is obtained. Obviously this expression will be only valid for the value of the parameters leading to a proper epidemic, i.e. for the value of the parameters satisfying $R_0 > 1$.

Finally we can also compute the phase plane trajectories $S - R$, from which we will be able to determine an expression for the final number of susceptible individuals in the end of the epidemic. The phase plane trajectories can be obtained as follows,

$$\frac{\mathrm{d}S}{\mathrm{d}R} = \frac{1}{\gamma}\left[-\frac{\lambda S}{S+\kappa} + \alpha\right] \Longrightarrow \int_{S_0}^{S_\infty} \frac{S+\kappa}{S(\alpha-\lambda)+\alpha\kappa}\mathrm{d}S = \frac{1}{\gamma}\int_{R_0=0}^{R_\infty} \mathrm{d}R \ ,$$

where the solution of the integral yields,

$$(S_\infty - S_0)\,(\alpha-\lambda) - \kappa\lambda \ln\left(\frac{\lambda S_\infty - \alpha\,(\kappa+S_\infty)}{\lambda S_0 - \alpha\,(\kappa+S_0)}\right) = \frac{(\alpha-\lambda)^2}{\gamma} R_\infty \ .$$

At the end of the epidemic, by definition, there are not infected individuals left. Then, using $R_\infty = 1 - S_\infty$ we get,

$$S_\infty\left[(\alpha-\lambda) + \frac{(\alpha-\lambda)^2}{\gamma}\right] - \kappa\lambda \ln\left[\lambda S_\infty - \alpha(\kappa+S_\infty)\right] = S_0\,(\alpha-\lambda) - \kappa\lambda\ln(\lambda S_0 - \alpha\,(\kappa+S_0)) + \frac{(\alpha-\lambda)^2}{\gamma} \ .$$

The solution to this transcendental equation can be expressed in terms of the Lambert's W function (see Appendix B for more information about the Lambert's W function),

$$S_\infty = -\frac{\kappa\lambda}{A}\cdot W_0\left[-\frac{A\cdot[\lambda S_0 - \alpha\,(\kappa+S_0)]\exp\left(-\frac{\alpha A}{\lambda(\lambda-\alpha)} - \frac{C}{\kappa\lambda}\right)}{\kappa\lambda(\lambda-\alpha)}\right] + \frac{\alpha\kappa}{\lambda-\alpha} \ , \tag{3.14}$$

where $A = (\alpha-\lambda) + \frac{(\alpha-\lambda)^2}{\gamma}$ and $C = S_0\,(\alpha-\lambda) + \frac{(\alpha-\lambda)^2}{\gamma}$

Note that the A factor can be written in the form $A = (\lambda-\alpha)\cdot\frac{\lambda-\alpha-\gamma}{\gamma}$ and, as $\lambda > \gamma+\alpha$, it will be always positive. Thus, to guarantee that $S_\infty > 0$ we need $\lambda S_0 > \alpha\,(\kappa+S_0)$. This condition restricts the validity range of our expression for S_∞.

3.4 3D deterministic model without recovery

As far as it is known, the number of fan mussels that are able to beat the parasite are insignificant, so it is reasonable to further reduce the model by considering $\alpha = 0$ in Eq. (3.6). This leads to the following system of differential equations,

$$\begin{aligned} \dot{S} &= -\frac{\lambda SI}{S+\kappa} \\ \dot{I} &= \frac{\lambda SI}{S+\kappa} - \gamma I \\ \dot{R} &= \gamma I \ , \end{aligned} \tag{3.15}$$

where $\kappa \equiv \mu/\beta$.

This simplified model is easier to study analytically. Again, the basic reproduction number can be obtained from the NGM method or as in the SIR model, both yielding the same result,

$$R_0 = \frac{\lambda S_0}{\gamma\,(S_0+\kappa)}\ . \tag{3.16}$$

Again, substituting $\kappa = \mu/\beta$ we recover the same R_0 of the 4D model.

The next important analytical result that can be obtained is the maximum of infected individuals,

$$\frac{\mathrm{d}I}{\mathrm{d}t} = 0 \Longrightarrow S_c = \frac{\gamma\kappa}{\lambda-\gamma}\ . \tag{3.17}$$

Note that this relation show that the model is structurally unstable for $\lambda = \gamma$ and it requires $\lambda > \gamma$ for physical sense. This is consistent with the model with recovery setting $\alpha = 0$.

By computing the phase plane trajectories of S and I the actual number of maximum infections can be found,

$$\frac{\mathrm{d}I}{\mathrm{d}S} = -1 + \frac{\gamma\,(S+\kappa)}{\lambda S} \Longrightarrow \int_{I_0}^{I} \mathrm{d}I = \int_{S_0}^{S} \left\{-1 + \frac{\gamma}{\lambda} + \frac{\gamma\kappa}{\lambda}\cdot\frac{1}{S}\right\}\mathrm{d}S\ .$$

The solution of this simple integral gives rise to the following phase plane trajectories,

$$I + S = I_0 + S_0 + \frac{\gamma}{\lambda}\,(S - S_0) + \frac{\gamma\kappa}{\lambda}\ln\!\left(\frac{S}{S_0}\right)\ .$$

From the fact that $S + I + R = 1$ (recall we are working with densities) it is clear that $S_0 + I_0 = 1$ as $R_0 = R(0) = 0$. Then, it has been already found that the maximum of infections will take place when $S = S_c = \lambda\kappa/(\lambda-\gamma)$, so that the maximum density of infected individuals is given by,

$$I_{\max} = 1 - \frac{\gamma\kappa}{\lambda-\gamma} + \frac{\gamma}{\lambda}\left(\frac{\gamma\kappa}{\lambda-\gamma} - S_0\right) + \frac{\gamma\kappa}{\lambda}\ln\!\left(\frac{\gamma\kappa}{S_0\,(\lambda-\gamma)}\right)\ , \tag{3.18}$$

which will be checked numerically afterwards.

Note that the expression can be rewritten in terms of S_c and $\tilde{R}_0$ as,

$$I_{\max} = 1 - \frac{\gamma\kappa}{\lambda}\ln\!\left(\tilde{R}_0\right) + S_c\left[\frac{\gamma}{\lambda}\left(1-\tilde{R}_0\right) - 1\right]\ , \tag{3.19}$$

from which we see that the it isn't defined for $\tilde{R}_0 < 0$, vanishes for $\tilde{R}_0 = 0$ and gets values between 0 and 1 for $\tilde{R}_0 > 1$, as it is expected.

Finally we can determine the density of susceptible individuals at the end of the epidemic. To this end, we first compute the phase space trajectories of R and S,

$$\frac{\mathrm{d}S}{\mathrm{d}R} = \frac{-\lambda S}{\gamma\,(S+\kappa)} \Longrightarrow \int_{S_0}^{S_\infty}\left(1 + \kappa\frac{1}{S}\right)\mathrm{d}S = -\frac{\lambda}{\gamma}\int_{R_0=0}^{R_\infty}\mathrm{d}R\ .$$

The solution of the integrals yield,

$$S_\infty - S_0 + \kappa\ln\!\left(\frac{S_\infty}{S_0}\right) = -\frac{\lambda}{\gamma}R_\infty \Longrightarrow S_\infty + \kappa\ln(S_\infty) = S_0 + \kappa\ln(S_0) - \frac{\lambda}{\gamma}R\ .$$

At the end of the epidemic necessarily $I_\infty = 0$ so that $R_\infty = 1 - S_\infty$. Thus,

$$S_\infty\left(1 - \frac{\lambda}{\gamma}\right) + \kappa\ln(S_\infty) = S_0 + \kappa\ln(S_0) - \frac{\lambda}{\gamma}\ .$$

This is a transcendental equation which solution can be expressed in terms of the Lambert W function,

$$S_\infty = \frac{\kappa}{1-\lambda/\gamma}W_0\left[\frac{1-\lambda/\gamma}{\kappa}S_0\exp(C/\kappa)\right]\ , \tag{3.20}$$

where $C = S_0 - \lambda/\gamma$. Note that this can be rewritten in terms of S_c, Eq. (3.17), and $\tilde{R}_0 = S_0/S_c$ as,

$$S_\infty = -S_c \cdot W_0 \left[-\tilde{R}_0 \exp(C/\kappa)\right] \ . \tag{3.21}$$

Note that, as $S_c > 0$ and $\tilde{R}_0 > 0$, both the prefactor and the argument of the Lambert's W function will be negative, given the positive value of the exponential term, and so W is two-valued in this range the fundamental branch, $l = 0$, is chosen, as $S_\infty < S_c$ [16].

3.5 Stochastic model

As has been previously discussed the deterministic models assume that the population is big and fully mixed, so that fluctuations can be neglected. The fan mussel population is rather small so that fluctuations will become important. In order to take them into account the stochastic approach presented in Section 2.3 is followed. With the scheme in Eq. (3.1) we can write the master equation for the probability $p(n_S, n_I, n_P; t)$. Note that as in the SIR model, the removed population of Pen-Shells can be deduced from the conservation of the total population, i.e. $R = N - S - I$. Thus, the master equation is given by,

$$\frac{\partial p(n_S, n_I, n_P; t)}{\partial t} = \left(E_S E_I{}^{-1} E_P - 1\right) \left[\beta \frac{n_S n_P}{N} p(n_S, n_I, n_P; t)\right] + (E_I - 1)\left[(\gamma + \alpha) n_I p(n_S, n_I, n_P; t)\right] +$$

$$\left(E_I E_S{}^{-1} - 1\right)\left[\alpha n_I p(n_S, n_I, n_P; t)\right] + \left(E_P{}^{-1} - 1\right)\left[\lambda n_I p(n_S, n_I, n_P; t)\right] + (E_P - 1)\left[\mu n_P p(n_S, n_I, n_P; t)\right] =$$

$$= \beta \frac{(n_S + 1)(n_I - 1)(n_P + 1)}{N} p(n_S + 1, n_I - 1, n_P + 1; t) + \gamma (n_I + 1) p(n_S, n_I + 1, n_P; t) +$$

$$+ \alpha (n_I + 1) p(n_S - 1, n_I + 1, n_P; t) + \lambda n_I p(n_S, n_I, n_P - 1; t) + \mu (n_P + 1) p(n_S, n_I, n_P + 1; t) -$$

$$- \left(\beta \frac{n_S n_I n_P}{N} + (\gamma + \alpha + \lambda) n_I + \mu n_P\right) p(n_S, n_I, n_P; t) \ .$$

This complicated equation will be solved numerically with the so-called Gillespie algorithm [18]. Besides the numerical computation of the master equation, the evolution of the first moment of the variables of interest can be studied analytically,

$$\begin{cases} \dfrac{\mathrm{d}n_S}{\mathrm{d}t} = -\dfrac{\beta}{N} \langle n_S n_P \rangle + \alpha \langle n_I \rangle \\ \dfrac{\mathrm{d}n_I}{\mathrm{d}t} = \dfrac{\beta}{N} \langle n_S n_P \rangle - (\gamma + \alpha) \langle n_I \rangle \\ \dfrac{\mathrm{d}n_P}{\mathrm{d}t} = \lambda \langle n_I \rangle - \mu \langle n_P \rangle - \dfrac{\beta}{N} \langle n_S n_P \rangle \end{cases}$$

By means of the mean-field approximation the deterministic model can be recovered. Recall that the basic assumption is that the population is large so that fluctuations can be neglected. Mathematically it leads to $\sigma^2(n_i) = \langle {n_i}^2 \rangle - \langle n_i \rangle^2 \approx 0 \Longrightarrow \langle {n_i}^2 \rangle \approx \langle n_i \rangle^2$. As we are considering the limit of large fan mussel population ($N \to \infty$) we fix a non-vanishing quantity $i = \langle n_i \rangle / N$ (e.g. $S = \langle n_S \rangle / N$). Then, the fixed quantities will be the density of susceptible, infected and recovered (removed) fan mussels (S, I, R) and the ratio of the parasite population to the fan mussels population ($P \equiv \langle n_P \rangle / N$). With this we recover the deterministic system Eq. (3.2),

$$\begin{cases} \dfrac{\mathrm{d}S}{\mathrm{d}t} = -\beta SP + \alpha I \\ \dfrac{\mathrm{d}I}{\mathrm{d}t} = \beta SP - (\gamma + \alpha) I \\ \dfrac{\mathrm{d}P}{\mathrm{d}t} = \lambda I - \mu P - \beta SP \end{cases}$$

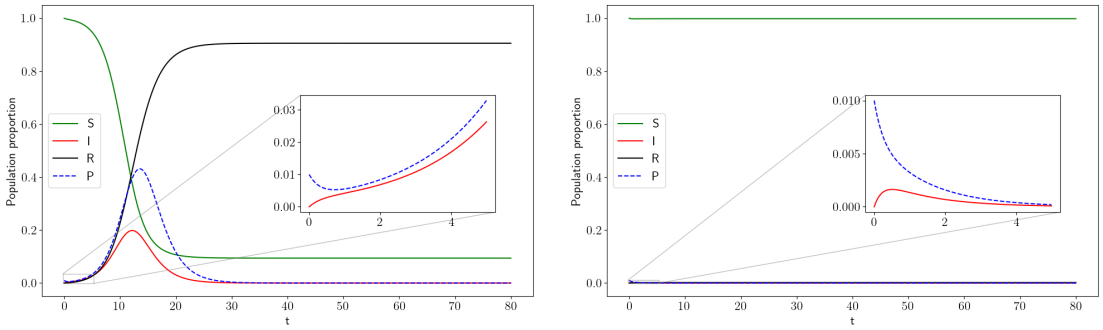
4 Numerical analysis of the developed models

4.1 4D deterministic model

To check the validity of the threshold value found for this model a qualitative and quantitative analysis can be performed by numerically solving the system of differential equations in Eq. (3.2). Qualitatively, we expect an outbreak for the epidemic when $R_0 > 1$ and the nonexistence of epidemic for $R_0 < 1$. In order to observe this phenomenon, the following parameters are fixed

$$S_0 = 1 \qquad I_0 = R(0) = 0 \qquad \gamma = 0.5 \qquad \beta = \mu = 1 \qquad \lambda = 3 \qquad P_0 = 0.01$$

Then, $\alpha = 0.25$, 2.5 are considered so that $R_0 = 2$, 0.5, respectively.

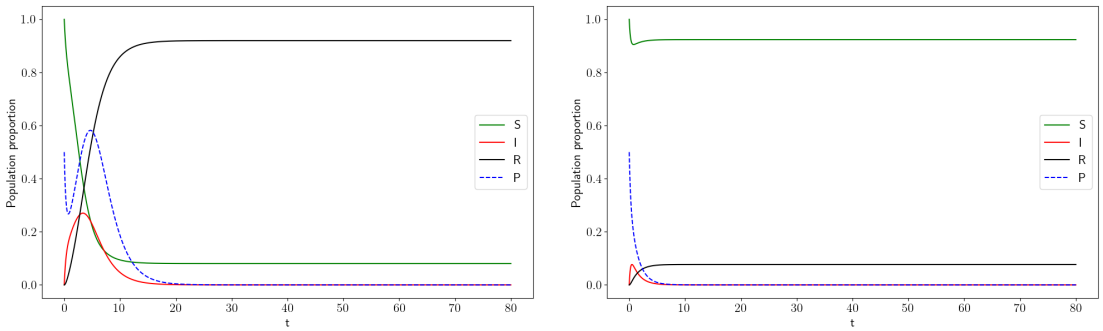


(a) $R_0 = 2$ (b) $R_0 = 0.5$

Figure 3: Different realisations of the deterministic SIRP model: (a) above the threshold, $R_0 = 2$; (b) below the threshold $R_0 = 0.5$. The insets focus on the initial evolution of the I and P compartments.

As can be observed in Fig. 3 the numerical results are in good agreement with the threshold value found by means of theoretical methods. As shown in the insets, the threshold is characterised for the change of the P evolution from monotonically decreasing to "first decreasing and then increasing", which has also been discussed from a theoretical point of view.

In order to analyse the effect of the initial population of parasites, we have now solved the system considering a non negligible initial number of parasites, $P_0 = 0.5$



(a) $R_0 = 2$ (b) $R_0 = 0.5$

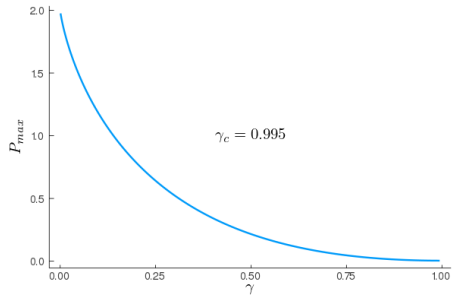
Figure 4: Different realisations of the deterministic SIRP model: (a) above the threshold, $R_0 = 2$; (b) below the threshold $R_0 = 0.5$. The insets focus on the initial evolution of the I and P compartments.

In Fig. 4 we see that for a large enough value of the initial parasite population a considerable part of the population will be affected although $R_0 < 1$. Quantitatively, there is difference with the previous case, where $P_0 \ll 1$, between the susceptible population at the end of the simulations.

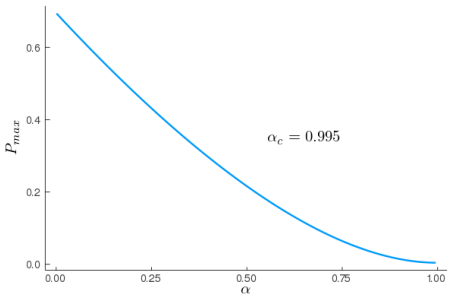
This magnitude is always large when $R_0 < 1$, being in one case almost 100% for P_0 very small and around 90% in the case of P_0 large, while this population is around 10% or less for $R_0 > 1$ in both cases. Moreover, we can see that the qualitative behaviour is different in both cases, below and above the threshold, in the same manner as before. In essence, above the threshold there is a proper epidemic outbreak while there is not below it. Moreover, below the threshold the evolution of P is monotonically decreasing while above the threshold it first decrease and then increase. From this fact, we can numerically determine the threshold by computing the maximum of the parasite population, P. As we have five different parameters that determine the basic reproduction number (i.e. α, β, γ, μ and λ) we should study the threshold behaviour for each of this parameters independently. To this end, we first compute analytically this critical parameters and then we check the results numerically. By fixing three of the parameters and isolating one of them from Eq. (3.3) we obtain

$$\gamma_c = \frac{\beta\lambda}{\beta+\mu} - \alpha \qquad \alpha_c = \frac{\beta\lambda}{\beta+\mu} - \gamma \qquad \lambda_c = \frac{(\gamma+\alpha)(\beta+\mu)}{\beta} \qquad \mu_c = \frac{\beta(\lambda-\gamma-\alpha)}{\gamma+\alpha} \qquad \beta_c = \frac{\mu(\alpha+\gamma)}{\lambda-\gamma-\alpha}$$

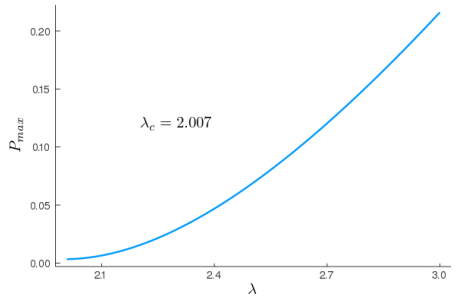
Then, in order to check numerically the validity of the expressions above we set $\lambda = 3$, $\gamma = \alpha = 0.5$ and $\beta = \mu = 1$, so that the critical parameters become $\gamma_c = 1$, $\alpha_c = 1$, $\lambda_c = 2$, $\mu_c = 2$ and $\beta_c = 0.5$.



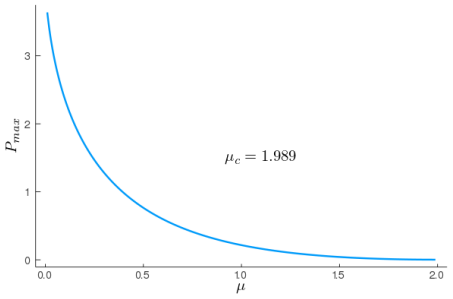
(a) $\gamma_c = 0.995$



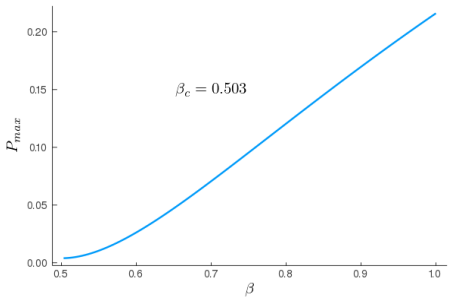
(b) $\alpha_c = 0.995$



(c) $\lambda_c = 2.007$



(d) $\mu_c = 1.989$



(e) $\beta_c = 0.503$

Figure 5: Numerical determination of the critical parameters for the 4D deterministic model, Eq. (3.2)

In Fig. 5 the value of the maximum of P is plotted against each of the parameters up to a point were it vanishes, which is considered the critical value of the parameter. The results for the critical parameters are in perfect agreement with the theoretical results with a maximum of a 1% of relative error.

Now that we have checked the validity of the computed expression for R_0 we perform a sensitivity analysis to study the relevance of the parameters. We analyse the local sensitivity of R_0 for each of the parameters that conform it, p_i, through the normalised sensitivity index [24], given by

$$\Omega_{p_i}^{R_0} = \frac{\partial R_0}{\partial p_i} \frac{p_i}{R_0}\Big|_{p_i=p^0} \tag{4.1}$$

Eq. (4.1) gives rise to the following expressions

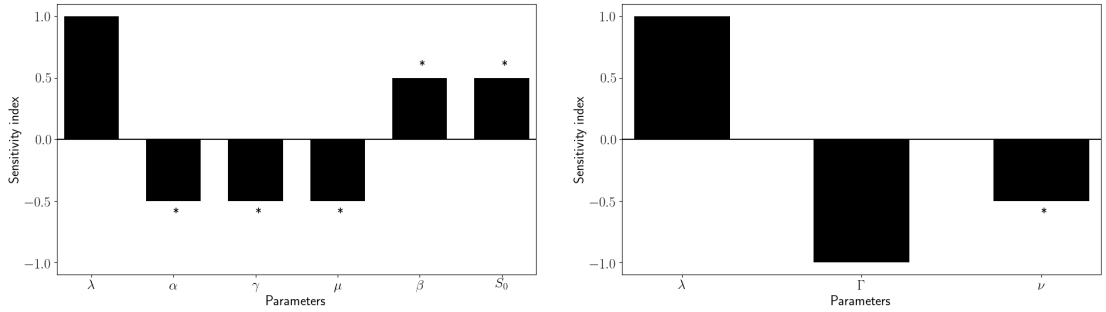
$$\Omega_{\lambda}^{R_0} = 1, \quad \Omega_{\alpha}^{R_0} = -\frac{\alpha}{\alpha+\gamma}, \quad \Omega_{\gamma}^{R_0} = -\frac{\gamma}{\alpha+\gamma}, \quad \Omega_{\mu}^{R_0} = -\frac{1}{1+\frac{\beta S_0}{\mu}}, \quad \Omega_{\beta}^{R_0} = \Omega_{S_0}^{R_0} = \frac{1}{1+\frac{\beta S_0}{\mu}}$$

and for the simplified expression of R_0, Eq. (3.4),

$$\Omega_{\lambda}^{R_0} = 1, \qquad \Omega_{\Gamma}^{R_0} = -1 \qquad \Omega_{\nu}^{R_0} = -\frac{\nu}{1+\nu}$$

Note that the results for the simplified version of R_0 are consistent with the results for the full parameter version in their limiting cases.

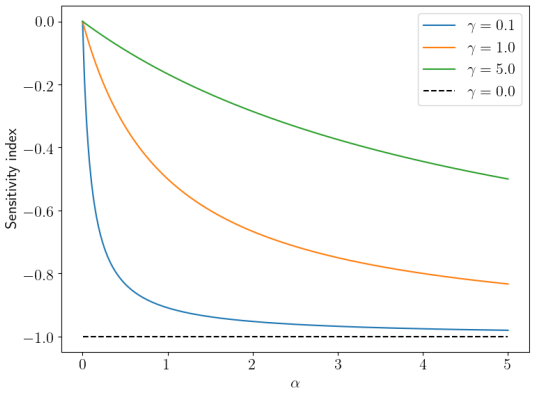
In Fig. 6 the sensitivity index for the full and simplified versions of R_0 is shown for specific baseline parameters while Fig. 7 show the sensitivity analysis for the parameters that, in turn, depend on other parameters.



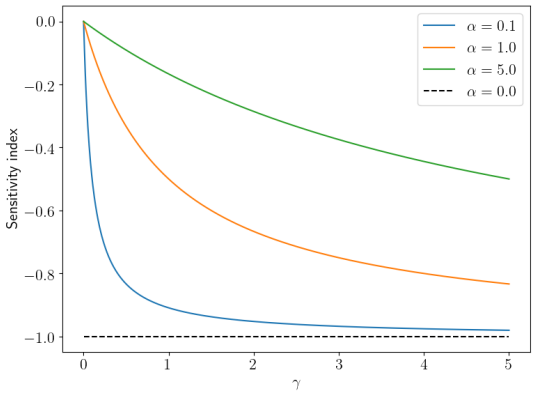
(a) Full R_0 (b) Simplified R_0

Figure 6: Sensitivity analysis of R_0 for the baseline parameters $\lambda = 1$, $\alpha = \gamma = 0.5$ and $\mu = \beta = S_0 = 1$. The asterisks mark parameters for which the sensitivity index is not constant.

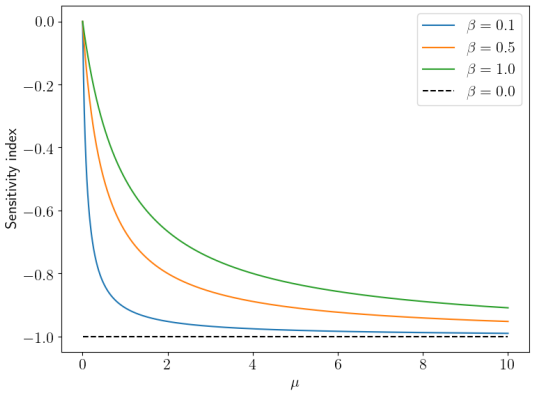
In Fig. 6 we observe that λ, β and S_0 contribute to increase the basic reproduction number while α, γ and μ contribute to decrease it, as expected. Moreover, comparing Fig. 6(a) and Fig. 6(b) we see that the simplified version of R_0 captures the information of its full version. The parameters marked with an asterisk have varying influences on R_0, which is fully depicted in Fig. 7. With this we observe that the influence of α decreases with the increase of γ and the other way round. Regarding β, its relevance increases with the increase of μ and the decrease of S_0. The same happen to S_0, its importance increases with the increase of μ and the decrease of β. On the other hand, the impact of μ increases with the increase of both S_0 or β. Finally, we see in the last plot of the figure that the impact of the summarised parameter ν increases with its own value.



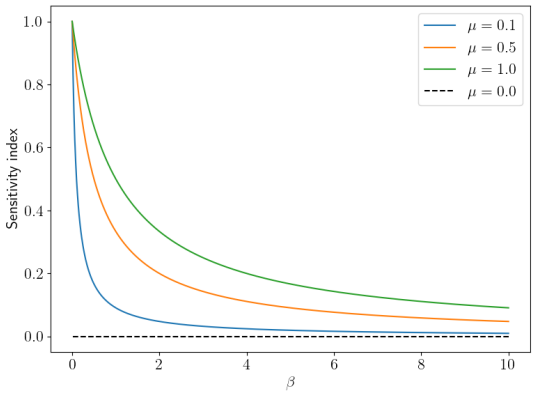
(a) Fan mussel recovery rate, α



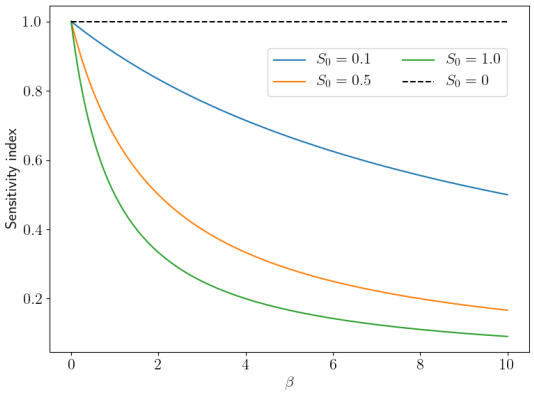
(b) Fan mussel death rate γ



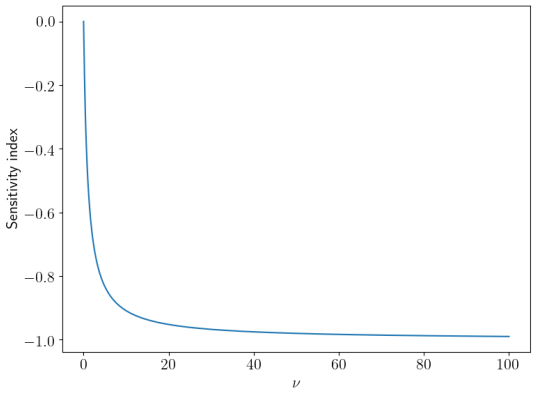
(c) Parasite death rate μ



(d) Infection rate $\beta(\mu)$



(e) Infection rate $\beta(S_0)$



(f) Summarised parameter ν

Figure 7: Sensitivity analysis of R_0 for parameters with varying sensitivity index.

Finally we qualitatively analyse the time scale approximation performed to reduce this 4D model to a three dimensional one, i.e. Eq. (3.5). To this end, we solve numerically the system of differential equations of the 4D model, for the previous choice of the parameters giving $R_0 = 2$ and we compare the actual parasite population evolution to that of the effective parasite population given by the approximation,

$$P(S, I) = \frac{\lambda}{\beta S + \mu} I \ .$$

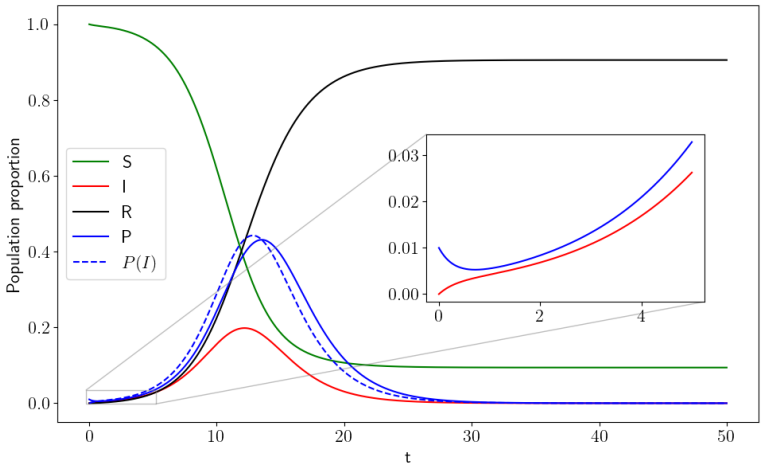


Figure 8: Comparison of $P(t)$ of the 4D model and the effective $P(S, I)$ from the approximation for $R_0 = 2$.

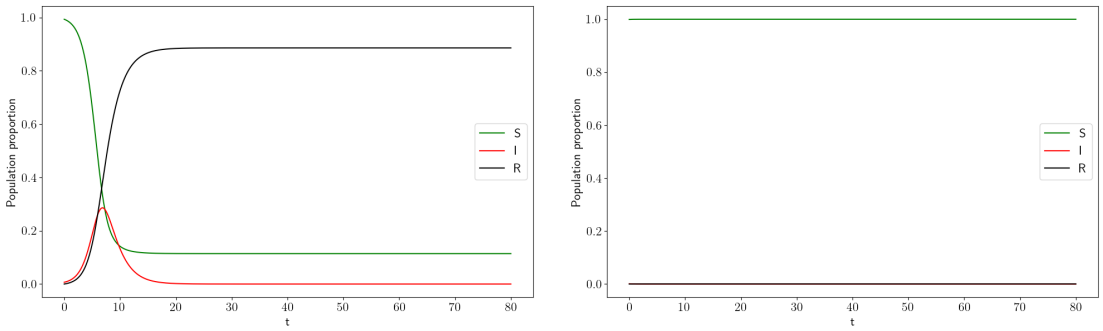
As it can be observed in Fig. 8, the time evolution of the effective parasite population, given by the approximation $P(S, I)$, is quite close to that of the actual parasite population, $P(t)$, of the 4D model. Thus, the analysis suggest the validity of the approximation.

4.2 3D deterministic model

In order to check that this reduced 3D model corresponds to the initial 4D one we make a first preliminary study for a particular choice of the parameters that correspond to the ones used before. In essence we use the following parameters

$$I_0 = 0.01 \qquad S_0 = 0.99 \qquad R_0 = 0 \qquad \beta = \lambda = \mu = 1 \Longrightarrow \kappa = 1 \qquad \gamma = 0.5$$

and as before $\alpha = 0.5$, 2.5 are considered so that $R_0 \approx 2$, 0.5, respectively [3].



(a) $R_0 = 2$ (b) $R_0 = 0.5$

Figure 9: Different realisations of the deterministic SIR model, one above the threshold and one below it for $P_0 \ll 1$. The insets focus on the initial evolution of the I and P compartments.

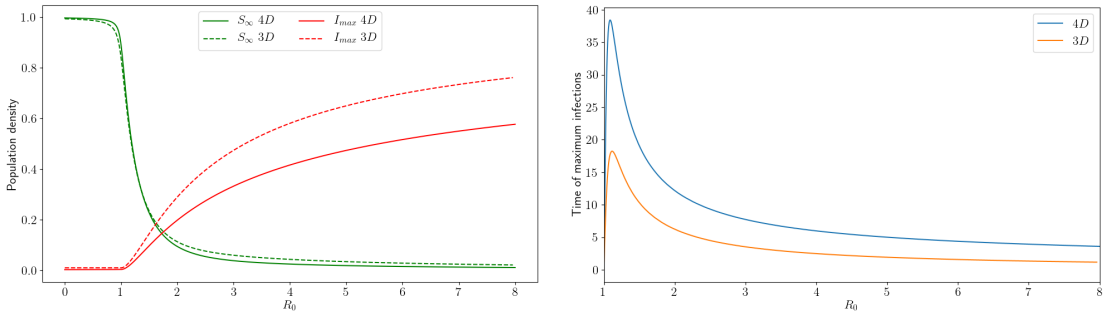
Comparing Fig. 9 with Fig. 3 it seems that, qualitatively, the reduced model successfully reproduces the 4D model. Moreover, the numerical results are in agreement with the threshold value found analytically, at least qualitatively. However, if we compare some quantitative results from both models some differences appear, which are summarised in Table 1. It can be appreciated that the major difference appears in the maximum of infected individuals, so that the approximation works better in reproducing the final state of the solution rather than its time evolution.

[3] Now the result is approximate as S_0 can not be 1, but $S_0 \approx 1$ instead

Table 1: Quantitative comparison between the 4D and 3D models

	4D	3D	Absolute difference	Relative error (%)
S_∞	0.094	0.114	0.02	21
R_∞	0.906	0.886	0.02	2
$I_{\max}$	0.20	0.28	0.08	40

This differences between the 4D and the 3D model for the final state of the epidemic and the maximum density of infected individuals can be studied qualitatively for several values of the basic reproduction number.



(a) S_∞ and I_{max}

(b) Time of maximum infections

Figure 10: Numerical comparison between the 4D and 3D models for the final density of infected individuals and the maximum density of infected individuals.

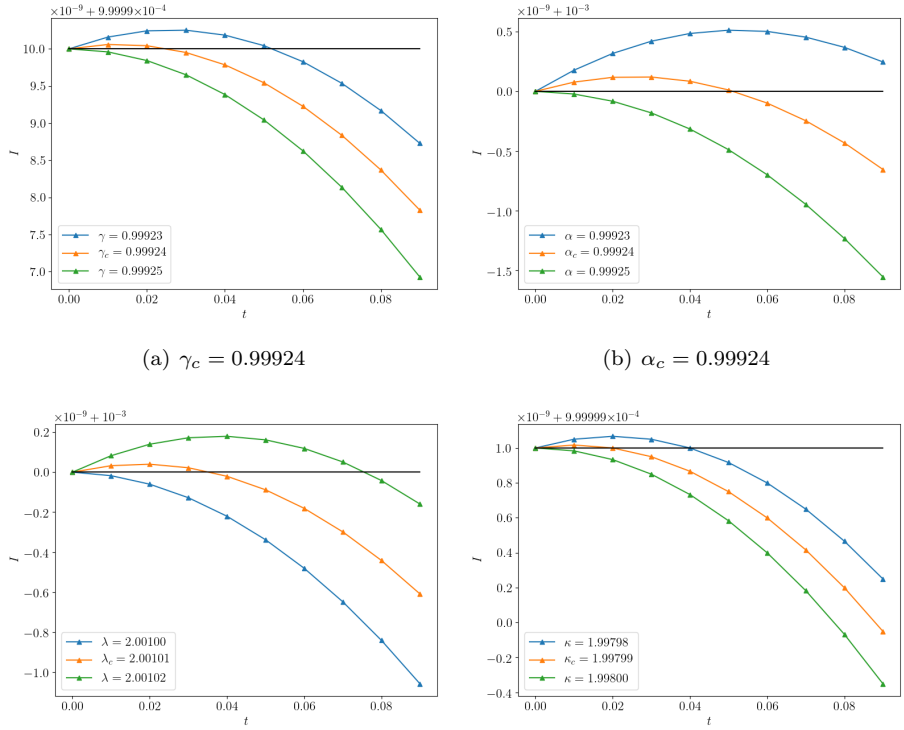
In Fig. 10(a) one can appreciate that, as previously discussed, the 3D model can be considered a good approximation to the initial 4D model for the final states of the solutions rather than its time evolution. However, the 3D model time evolution, which gives rise to quantities such as the maximum density of infected individuals, can be used as an upper limit for the solution. In Fig. 10(b) we also observe a difference in the time the maximum density of infected individuals occur. For this quantity, the 3D model can be useful to set a lower limit for the occurrence of this maximum.

Now the analytical expression found for the basic reproduction number can be checked quantitatively. Assuming that S is a monotonic decreasing function, the basic reproduction number can be deduced as in the original SIR model. This implies that an initial increase on the number of infected individuals is needed to have a proper outbreak, thus the threshold is defined as the critical value to go from $\dot{I} < 0$ to $\dot{I} > 0$. From this fact the critical parameters can be found numerically. From Eq. (3.8) we have

$$\gamma_c = \frac{\lambda S_0}{S_0 + \kappa} - \alpha \qquad \alpha_c = \frac{\lambda S_0}{S_0 + \kappa} - \gamma \qquad \lambda_c = \frac{(\gamma + \alpha)(S_0 + \kappa)}{S_0} \qquad \kappa_c = \frac{S_0(\lambda - \gamma - \alpha)}{\gamma + \alpha}$$

Using $S_0 = 0.999$, $\lambda = 3$, $\gamma = \alpha = 0.5$ and $\kappa = 1$ the critical parameters become

$$\gamma_c = 0.99924... \qquad \alpha_c = 0.99924... \qquad \lambda_c = 2.001001... \qquad \kappa_c = 1.998$$



(a) $\gamma_c = 0.99924$ (b) $\alpha_c = 0.99924$

(c) $\lambda_c = 2.00101$ (d) $\kappa_c = 1.99799$

Figure 11: Numerical determination of the critical parameters for the 3D deterministic model, Eq. (3.6)

In Fig. 11 it can be observed that the numerical determination of the critical parameters is in good agreement with the theoretical ones.

Then the analytical solution for the maximum density of infected individuals can be checked numerically. Fig. 12 shows a comparison between the analytical solution for the maximum of infected individuals, Eq. (3.13) and the numerical results obtained for the following values of the parameters

$$I_0 = 0.01 \qquad S_0 = 0.99 \qquad \kappa = 1 \qquad \lambda = 3 \qquad \gamma = 0.5 \qquad \alpha \in [0, 1] \Longrightarrow R_0 \in [1, 3]$$

A good agreement is observed between the analytical and numerical results.

Finally we check numerically the analytical expression for the final density of susceptible individuals. To this end we fix the following values for the parameters

$$I_0 = 0.01 \qquad S_0 = 0.99 \qquad \kappa = 1 \qquad \lambda = 3 \qquad \alpha = 0.5 \qquad \gamma \in [0, 10] \Longrightarrow R_0 \in [0.07, 3]$$

As can be observed in Fig. 13 there is a perfect agreement between the analytical solution and the numerical results.

After the analysis performed to the 3D model we can conclude with some key points. First, the 3D model keeps exactly the same threshold behaviour than the 4D model for the outbreak of an epidemic, i.e. they have the same R_0. Then, we have checked that the approximation is not perfect, as there are some differences between the time evolution of the fan mussel states of both models. This difference is particularly important for the maximum density of infected individuals. However, the approximation is valid for the final state of the epidemic and it can be used as an upper limit for the maximum density of infected individuals and a lower limit for the occurrence of this maximum.

Moreover, we are able to derive analytical solutions for this approximated model which are in good agreement with numerical results.

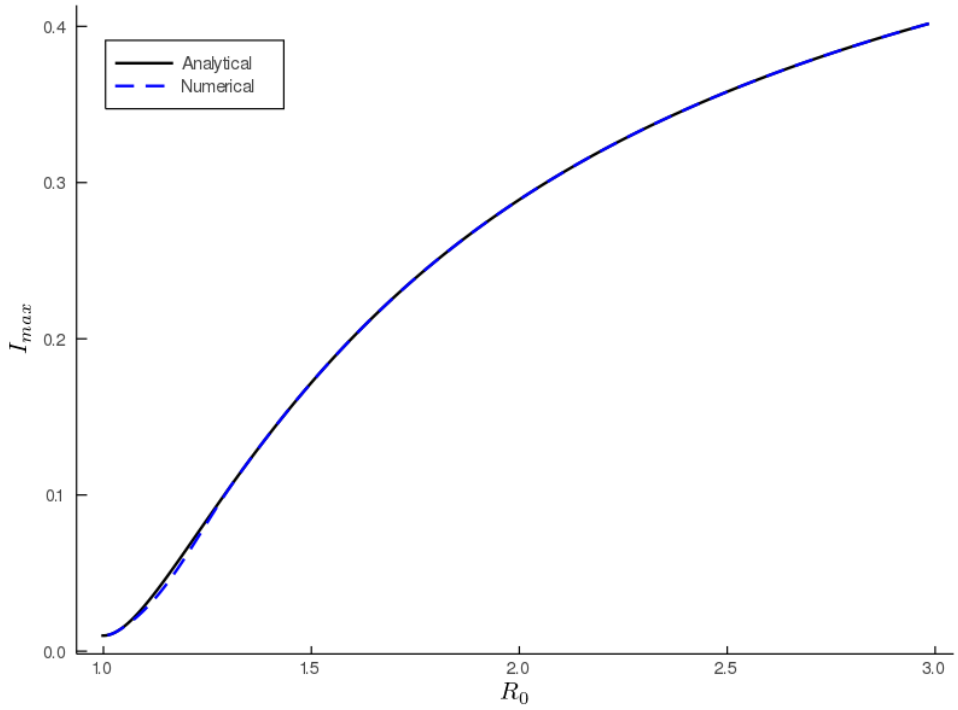


Figure 12: Maximum density of infected individuals as a function of the basic reproduction number. The dashed blue line corresponds to Eq. (3.13) and the solid black line to the numerical results.

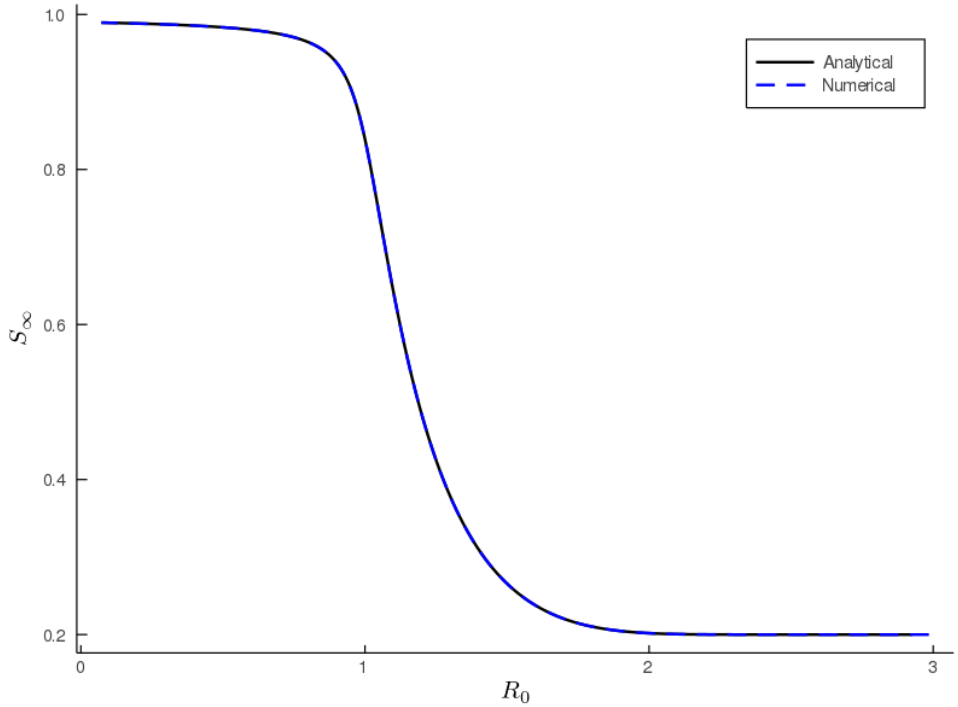


Figure 13: Final density of susceptible individuals as a function of the basic reproduction number. The dashed blue line corresponds to Eq. (3.14) and the solid black line to the numerical results.

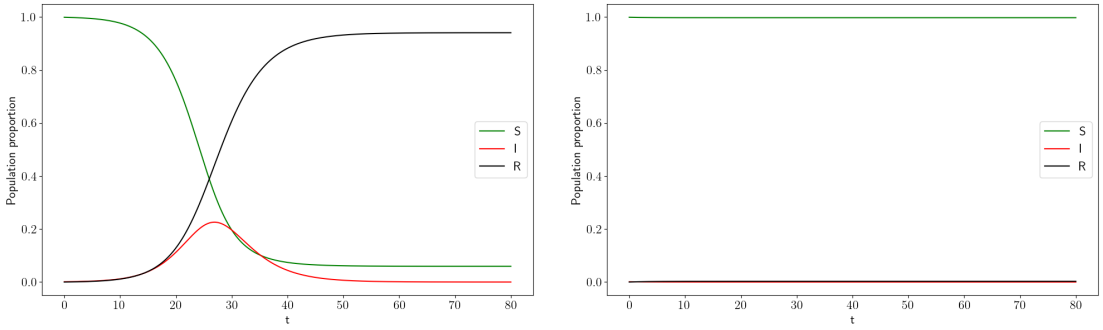
4.3 3D deterministic model without recovery

Now we analyse the 3D deterministic model with $\alpha = 0$. First of all let's check that this simplified model corresponds to the complicated 4D initial model with $\alpha = 0$. A qualitative comparison can be made by taking the same parameters used in both models and substituting $\kappa = \mu/\beta$. However, note that now a non-negative value for the initial density of infected individuals is needed (from Eq. (3.15)). Then we use the following parameters[4]

$$I_0 = 0.002 \qquad S_0 = 0.998 \qquad R_0 = 0 \qquad \beta = \lambda = \mu = 1 \Longrightarrow \kappa = 1 \qquad \gamma = 0.25,\ 0.8$$

Recall that with this choice of the parameters we have $R_0 = 2,\ 0.625$ for the 4D model. Now, substituting in Eq. (3.16) we have $R_0 = 1.9989..., 0.6247....$ This little difference is given by the fact that in the 4D model we had $S_0 = 1$ while now $S_0 = 0.998$ as a non-negative I_0 is needed.

In Fig. 14 the time evolution of the S, I and R compartments are plotted for cases below and above the threshold, $R_0 = 1$. We can try to check quantitatively the level of similarity between both models, 4D and 3D with $\alpha = 0$, by comparing the values of the maximum density of infected individuals and final density of susceptible and removed individuals. This results are summarised in Table 2, where it can be appreciated that, although the errors committed because of the approximation are very small, the approximated model reproduces better the final state of the epidemic rather than its time evolution.



(a) $R_0 = 2$ (b) $R_0 = 0.625$

Figure 14: Different realisations of the deterministic 3D SIR model, one above the threshold and one below it.

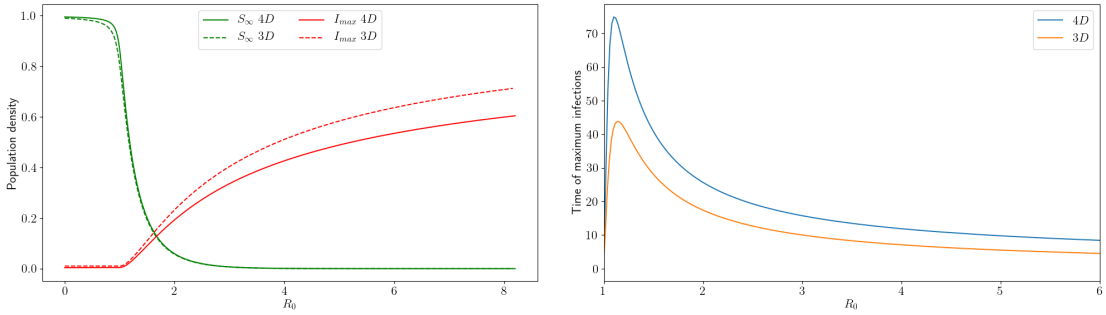
Table 2: Comparison between the 4D and 3D models

	4D	3D	Absolute difference	Relative error (%)
S_∞	0.05955	0.05923	$3.2 \cdot 10^{-4}$	0.54
R_∞	0.9401	0.94075	$6.5 \cdot 10^{-4}$	0.07
$I_{\max}$	0.1930	0.2258	0.033	18.13

The quantitative comparison between the 4D and 3D deterministic models shows good agreement for the final values of the density of susceptible and recovered (removed) individuals. An "a priori" non negligible difference appears in the maximum density of infected individuals. However, although the relative error seems quite large note that the numbers are very similar, so in practice the approximation would be useful. Thus, if we were interested only in the final situation this 3D model would be a perfect approximation and it should be a valid one also if we are interested in the evolution of the epidemic.

[4]Note that the I_0 value is chosen according to Eq. (3.7)

This differences between the 4D and the 3D model for the final state of the epidemic and the maximum density of infected individuals can be studied qualitatively for several values of the basic reproduction number.



(a) S_∞ and I_{max}

(b) Time of maximum infections

Figure 15: Numerical comparison between the 4D and 3D models for the final density of infected individuals and the maximum density of infected individuals.

In Fig. 15(a) one can appreciate that, as previously discussed, the 3D model can be considered a good approximation to the initial 4D model for the final states of the solutions rather than its time evolution. However, the 3D model time evolution, which gives rise to quantities such as the maximum density of infected individuals, can be used as an upper limit for the solution. We can also observe a difference in the time the maximum density of infected individuals occur, represented in Fig. 15(b). For this quantity, the 3D model can be useful to set a lower limit for the occurrence of this maximum. Moreover, in comparison to the 3D model with recovery we see that now the final states predicted are even more accurate than before, being basically equal to the solution of the 4D model and that the difference between the models in the time evolution of the fan mussel states is smaller.

Then we can check the validity of the basic reproduction number, R_0. Now we find it numerically in exactly the same way that analytically,: we look for the value of the selected parameter where the density of infected individuals increases at $t = 0$. From Eq. (3.16) we obtain that the critical parameters are

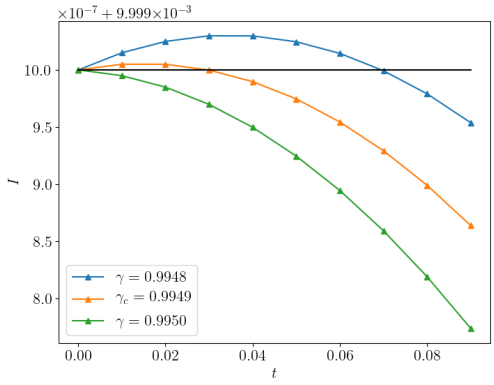
$$\gamma_c = \frac{\lambda S_0}{S_0 + \kappa} \qquad \lambda_c = \frac{\gamma S_0 + \kappa}{S_0} \qquad \kappa_c = S_0 \left(\frac{\lambda}{\gamma} - 1 \right) = 0.99$$

Using $S_0 = 0.999$, $\lambda = 2$, $\gamma = 1$ and $\kappa = 1$ the critical parameters become

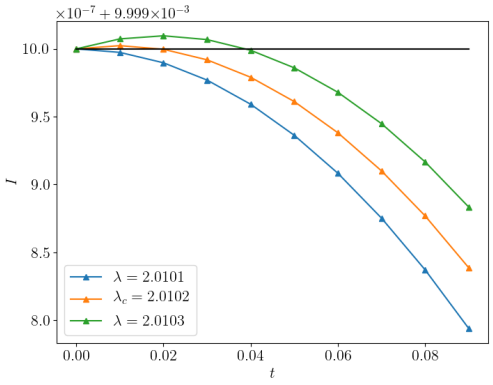
$$\gamma_c \approx 0.99497... \qquad \lambda_c \approx 2.01010... \qquad \kappa_c = 0.999$$

Fig. 16 shows the numerical determination of the critical parameters and their exact agreement with the analytical ones found above.

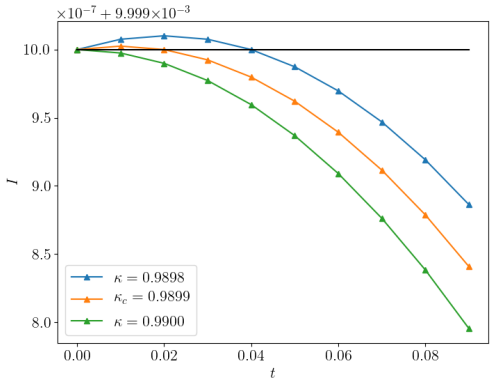
Finally, the final density of susceptible individuals and the maximum of infected individuals have been computed numerically and compared to Eq. (3.20) and Eq. (3.18), respectively. In Fig. 17 and Fig. 18 this quantities are plotted as a function of the basic reproduction number and in both figures show an extremely good agreement between the analytical and numerical solutions.



(a) $\gamma_c = 0.9949$



(b) $\lambda_c = 2.0102$



(c) $\kappa_c = 0.9989$

Figure 16: Numerical determination of the critical parameters for the 3D deterministic model, Eq. (3.15)

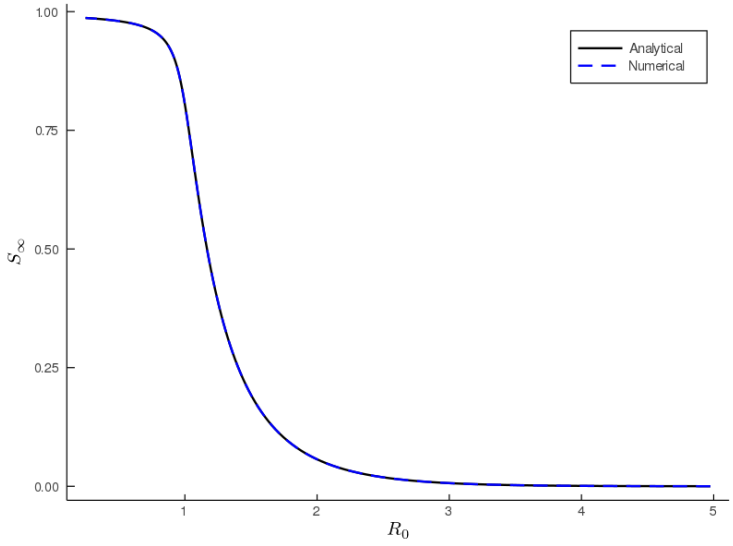


Figure 17: Final density of susceptible individuals as a function of the basic reproduction number. The dashed blue line corresponds to Eq. (3.20) and the solid black line to the numerical results.

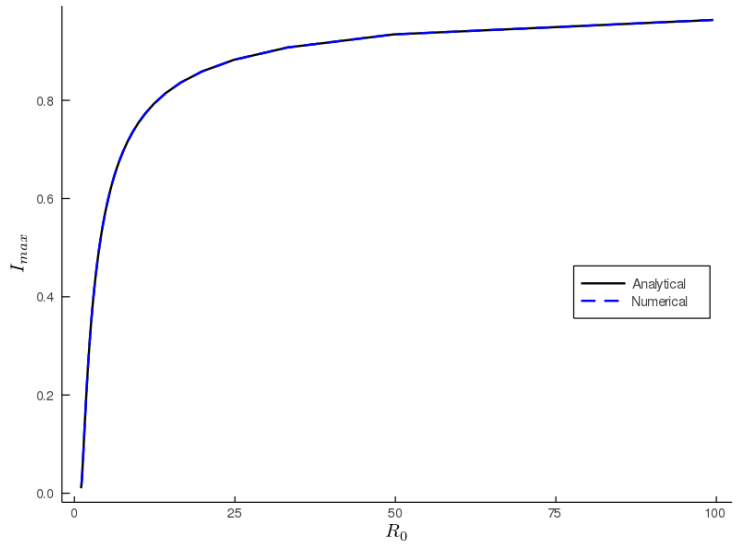


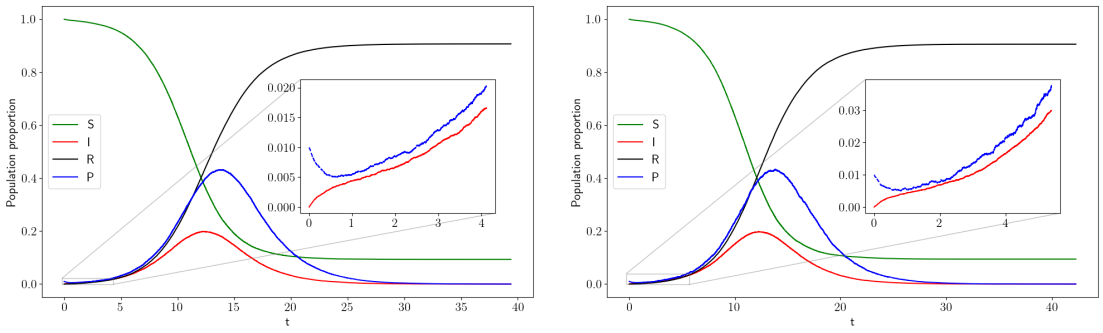
Figure 18: Maximum density of infected individuals as a function of the basic reproduction number. The dashed blue line corresponds to Eq. (3.18) and the solid black line to the numerical results.

4.4 Stochastic model

In this section the study of the stochastic model is presented by means of the Gillespie algorithm. First of all we check that its mean-field (MF) solution correspond to the deterministic 4D model, as was discussed. Moreover, we check the approximation of continuum release of parasite by infected fan mussels, still in the MF limit. This last point is done by comparing the stochastic simulation of continuum release with that of discrete release (expulsion of the generated quantity of parasite at the moment of death), which is feasible to consider numerically.

To consider the MF limit we set a number $N = 10^6$ of hosts. For such a big number of individuals, fluctuations should be small enough so that a single realisation approached the deterministic model. Then we set the same parameters that were used in the 4D deterministic model

$$S_0 = N \qquad I_0 = R(0) = 0 \qquad \gamma = 0.5 \qquad \beta = \mu = 1 \qquad \lambda = 3 \qquad \alpha = 0.25 \qquad P_0 = 0.01 \cdot N$$



(a) Continuous release

(b) Discrete release

Figure 19: Qualitative comparison between continuous and discrete release of parasite in infected hosts for a single realisation with $N = 10^6$ hosts.

As it can be observed in Fig. 19 for a number of $N = 10^6$ hosts the fluctuations in the results are negligible small so that the MF solution is achieved. We also see that both continuous and

discrete release of parasite by infected hosts give rise to the same results in the MF limit, at least qualitatively. Moreover, comparing Fig. 19 with Fig. 3 we see that, in the MF limit, the stochastic model can be described by the corresponding deterministic model of differential equations, as it has been discussed previously.

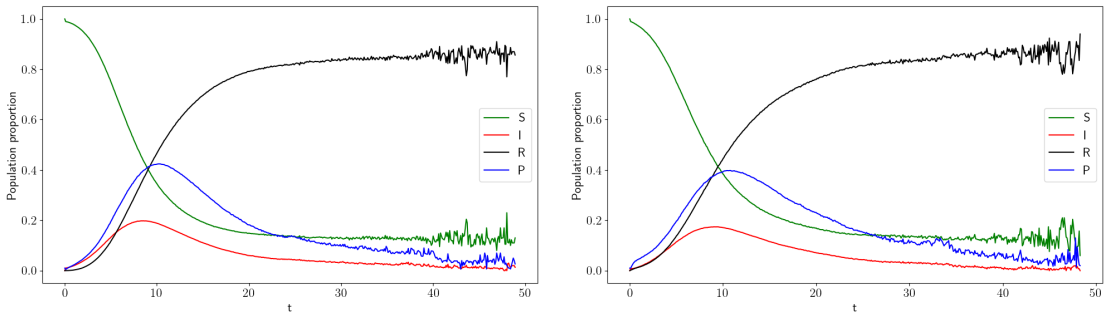
Then, we compute the average over $M = 100$ realisations in order to compare the results quantitatively with the deterministic 4D model and quantify the difference between the continuous and discrete release of parasite.

Table 3: Quantitative comparison between continuous and discrete parasite production in infected hosts. Results are averaged over $M = 100$ realisations for $N = 10^6$ individuals.

	$I_{\max}$	$S\infty$	R_∞
Continuous release	$0.198609 \pm 5.6 \cdot 10^{-5}$	$0.094073 \pm 4.0 \cdot 10^{-5}$	$0.905927 \pm 4.0 \cdot 10^{-5}$
Discrete release	$0.198660 \pm 4.4 \cdot 10^{-5}$	$0.094102 \pm 4.0 \cdot 10^{-5}$	$0.905898 \pm 4.0 \cdot 10^{-5}$
Relative error (%)	0.03	0.03	0.003
Deterministic	0.198338	0.094108	0.905892
Relative error (%)	0.7	0.1	0.01

In Table 3 a extremely agreement between the continuous and discrete release of parasite is observed with a negligible small relative error, so that, basically, the two models can be considered to yield the same results. Moreover, both models are in good agreement with the deterministic 4D model, despite the values do not match perfectly within errors.

Then the MF limit is approached by setting an small number of hosts, $N = 100$, and averaging the results over a huge number of realisations, in particular $M = 10^4$. The qualitative results are shown in Fig. 20 while the quantitative results are summarised in Table 4



(a) Continuous release

(b) Discrete release

Figure 20: Qualitative comparison between continuous and discrete release of parasite in infected hosts for a $M = 10^4$ realisations with $N = 100$ hosts.

As we can see in Fig. 20 the solutions of both continuous and discrete release models give rise to the same qualitative results. Moreover we can observe that at the end of the simulation the averages, despite of being over $M = 10^4$ realisations, begin to fluctuate. Probably the increase of the size of the fluctuations as time increases is due the collapse of an increasing number of realisations into a state with $S_\infty = 0$, so that the final part of the average has poorer statistics.

In Table 4 the maximum density of infected individuals is computed for both continuous and discrete release models, while S_∞ and R_∞ are not computed because of the huge fluctuations in the final part of the simulation. We observe that, with a small number of hosts, there is a non-negligible

difference between the two models. However, the maximum of infected individuals of both models are in good agreement within errors with the deterministic model.

Table 4: Quantitative comparison between continuous and discrete parasite production in infected hosts. Results are averaged over $M = 10^4$ realisations for $N = 100$ individuals.

	Continuous release	Discrete release
$I_{\max}$	0.198 ± 0.068	0.174 ± 0.055

5 Model validation with experimental data

The final aim of this thesis is to validate the presented models with experimental data. Unfortunately, at the present moment there is neither sufficient data nor quality data to clearly validate our models. Nevertheless, 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. The institutions that participated in this preservation effort were IFAPA, IEO, IRTA, IMEDMAR-UCV and Oceanografic of Valencia. The original idea was to rescue the individuals before infection, however, they appear to be have been carrying the pathogen at the time of extraction. This effort to save and preserve part of the fan mussel population of Portlligat also gave the opportunity to record data of the epidemic in the controlled water tanks, which is reported in [9]. The experimental data consists of the density of survivors as a function of time in the controlled water tanks with a temporal resolution of months. 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. Fortunately, two of the tanks kept its temperature approximately constant during all the recorded time. This is the case of IFAPA and Oceanografic water tanks, that are the ones we will take into account to validate our model.

Our initial model, the $4D$ deterministic model, has 6 free parameters: the recovery rate, α, the infection rate, β, the mortality rate of the fan mussels, γ, the production rate of parasites in infected fan mussels λ, the death rate of parasites, μ and the initial number of susceptible individuals, S_0. By visual inspection the mortality rate of fan mussels, γ, can be inferred from the experimental data, being $\gamma \approx 1\,\text{month}^{-1}$. Moreover, this value is supported by biologists working on the subject through their observations in the field. Unfortunately, all the other parameters remain unknown as there is no more experimental data or biological knowledge about them. Thus, the remaining parameters should be numerically fit to the available experimental data, which is basically impossible.

Although the scenario of using experimental data to validate our model seems to be difficult, we are able to fit some parameters to the given data by doing a crucial consideration. As reported in [25], the typical values of marine infection disease accomplish that the infection rate, β, is about two orders of magnitude smaller than the parasite death rate, μ, so that $\mu \gg \beta$ can be considered. Moreover, there is evidence that very few individuals manage to beat the parasite, so we also consider $\alpha = 0$. With these considerations we are able to reduce our 3D deterministic model to an original SIR model with an effective infection rate that captures three of the parameters of the previous model.

Recall our 3D deterministic model with $\alpha = 0$

$$\begin{aligned}\dot{S} &= -\frac{\lambda SI}{S+\kappa}\\ \dot{I} &= \frac{\lambda SI}{S+\kappa} - \gamma I\\ \dot{R} &= \gamma I\end{aligned}$$

With our consideration, $\kappa = \mu/\beta \gg 1$, we have

$$\dot{S} = -\frac{\lambda SI}{S+\kappa} \approx \frac{\lambda}{\kappa} SI = -\beta' SI$$

Because $S \in [0,1]$ and $\kappa \gg 1$. Thus we have

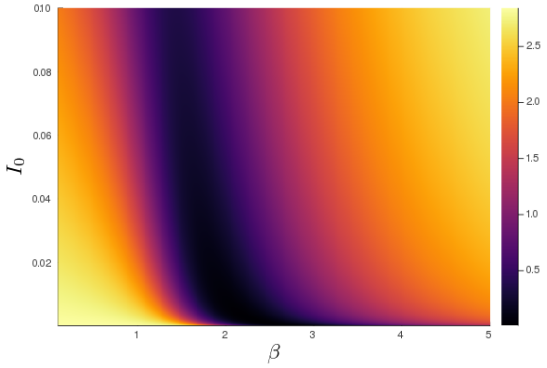
$$\begin{aligned}\dot{S} &= -\beta' SI\\ \dot{I} &= \beta' SI - \gamma I\\ \dot{R} &= \gamma I\end{aligned}$$

where $\beta' = \dfrac{\lambda}{\kappa} = \dfrac{\lambda}{\mu}\beta$.

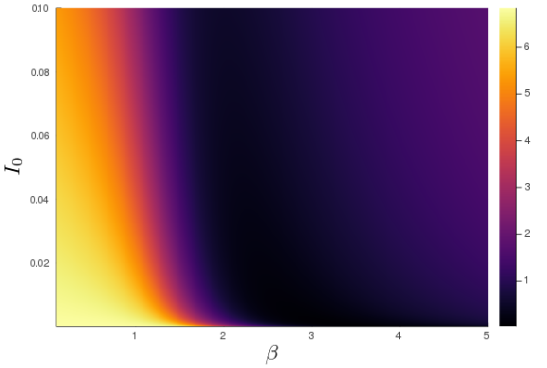
Note that λ/μ is the net production of parasite and β the infection probability so it makes sense to consider the effective infection rate as $\beta' = \lambda\beta/\mu$. Now the threshold value is simply given by $R_0 = \beta' S_0/\gamma$. We recover this result by considering $\nu \gg 1$ in our original expression of R_0

$$R_0 = \frac{\lambda}{\gamma\,(1+\nu)} \approx \frac{\lambda}{\gamma\nu} = \frac{\lambda}{\gamma\kappa} S_0 = \frac{\beta'}{\gamma} S_0$$

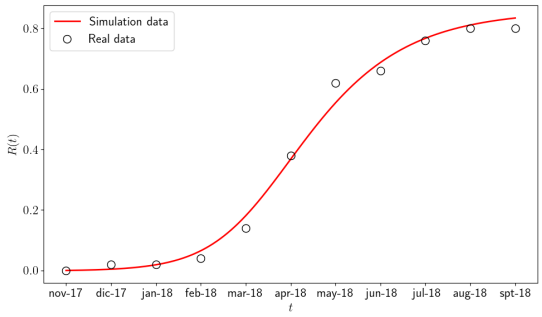
We then can use this super simplified model to fit the β' parameter, as we know that $\gamma \approx 1\,\text{month}^{-1}$. To perform the fitting we focus on the R compartment, as it can be retrieved directly from data[5] and we use a least squares procedure [26]. This give rise to the following results



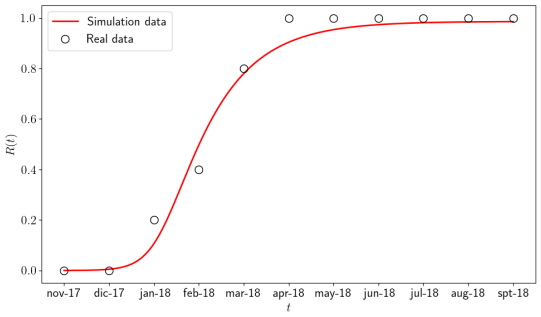
(a) MSE of the fitting procedure.



(b) MSE of the fitting procedure.



(c) $T \approx 14\,^{\circ}\text{C}$, $\beta' = 2.28\,\text{month}^{-1}$, $I_0 = 0.0021$



(d) $T = 17\,^{\circ}\text{C}$, $\beta' = 4.36\,\text{month}^{-1}$, $I_0 = 0.0006$

Figure 21: Parameter fitting for the R compartment using data from IFAPA (panels (a) and (c)) and Oceanografic (panels (b) and (d)).

In Fig. 21 a fit of the previous SIR model is performed for two of the controlled water tanks (IFAPA and Oceanografic, from left to right) reported in [9], which are maintained at constant temperature. In the panels above, Fig. 21(a) and Fig. 21(b), the corresponding mean squared errors (MSE) of all the fitted parameters are plotted. In particular, the MSE committed for the optimal parameters is less than 1%. In the panels below, Fig. 21(c) and Fig. 21(d), the simulated curve of removed individuals for the optimal fitted parameters is confronted to the experimental data, showing a remarkable agreement. Moreover, we check qualitatively the expected relation between R_0 (for the particular case of our parameter, $\gamma = 1$, $R_0 = \beta'$) and temperature, i.e., for a higher temperature we get a higher R_0. Note that if more data could be fitted for different temperatures the dependence between β' and temperature (considering γ independent of temperature) could be inferred.

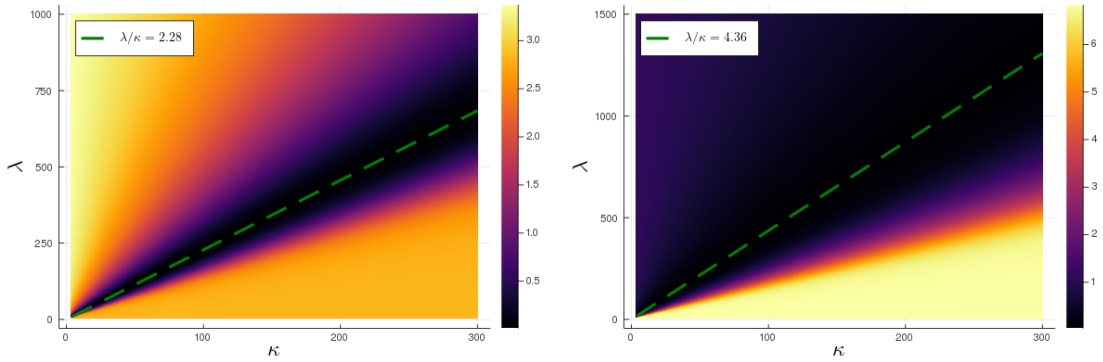
[5]The density of survivors can be the sum of infected and susceptible individuals, so it is clear that subtracting this data to 1 we obtain the density of removed individuals.

Once this fit has been performed, a step back can be done in order to try to fit the 3D model by inferring λ from the relation

$$\beta' = \lambda/\kappa \Longrightarrow \lambda = \beta'\kappa \tag{5.1}$$

As β' is known and it is a constant, this defines a straight line of valid parameters that fit the 3D model to the experimental data. From this fact it is clear than more experimental parameters are needed to continue with the experimental validation of the models presented in this thesis.

In order to show this last point we use the previous fitting procedure with the 3D model. First we consider $\kappa = 100$, as can be inferred from [25], and fit the 3D model for the λ and I_0 parameters. For consistency with the previous fits and with Eq. (5.1), we should obtain the same I_0 and $\lambda = 228,\ 436$ respectively. Indeed, we obtain the same $I_0 = 0.0021,\ 0.0006$, respectively, and a little modification on $\lambda = 229,\ 439$, respectively. Then we keep constant the value of initial density of infected individuals and vary the λ and κ parameters.



(a) IFAPA ($T \approx 14\,^{\circ}$C) (b) Oceanografic ($T = 17\,^{\circ}$C)

Figure 22: MSE for several pairs (κ, λ) parameters. The green dashed line correspond to the parameter values that fulfil Eq. (5.1).

In Fig. 22 we show graphically that, as it has been already discussed, all the parameter values that fulfil Eq. (5.1) minimise the mean squared error of the fitting. Thus, we have dependent parameters that make impossible to validate our 3D model without more experimental data.

To further emphasise the fact that different combinations of the κ and λ parameters are able to fit the experimental data we perform simulations with three different combinations of the parameters and plot the results for the density of removed individuals in front of the experimental data for the IFAPA tank in Fig. 23.

Finally, if we were to fit the 4D model we would encounter the same problem but with higher free dimensionality, this is, we would have an infinite number of combinations of three parameters, fulfilling some conditions, that would fit well the experimental data. For instance, if we consider $\kappa = \mu = \beta = 100$ we have again a straight line relating the parameters, so that all the values of μ and β which ratio is 100 will be valid parameters. Then we should fix the initial condition for the parasite population, P_0. As a final example, the following parameters for the SIR, 3D and 4D model produce simulated curves that fit well with the experimental data of IFAPA tank, as can be observed in Fig. 24.

SIR: $\beta' = 2.28,\ I_0 = 0.0021$

3D : $\kappa = 100,\ \lambda = 228,\ I_0 = 0.0021$

4D: $\beta = 2.4$, $\mu = 240$, $\lambda = 228$, $P_0 = 0.23$, $I_0 = 0$

Note that the parameters are related by $\kappa = \mu/\beta$ and $\lambda/\kappa = \beta'$.

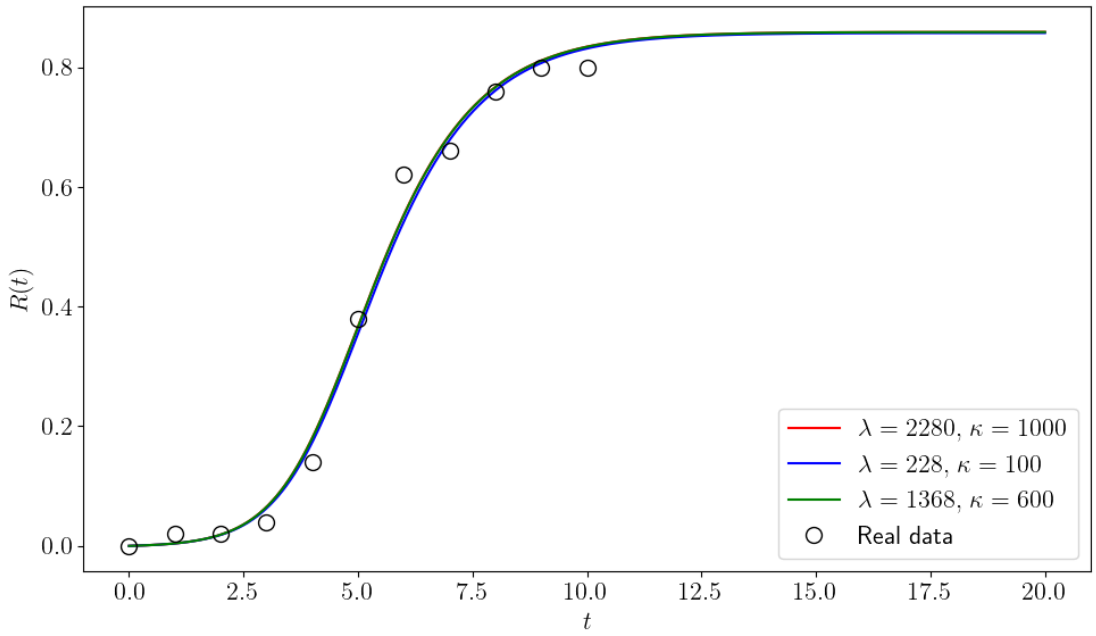


Figure 23: Comparison between experimental data (IFAPA) and simulations for three different valid combinations of λ and κ parameters.

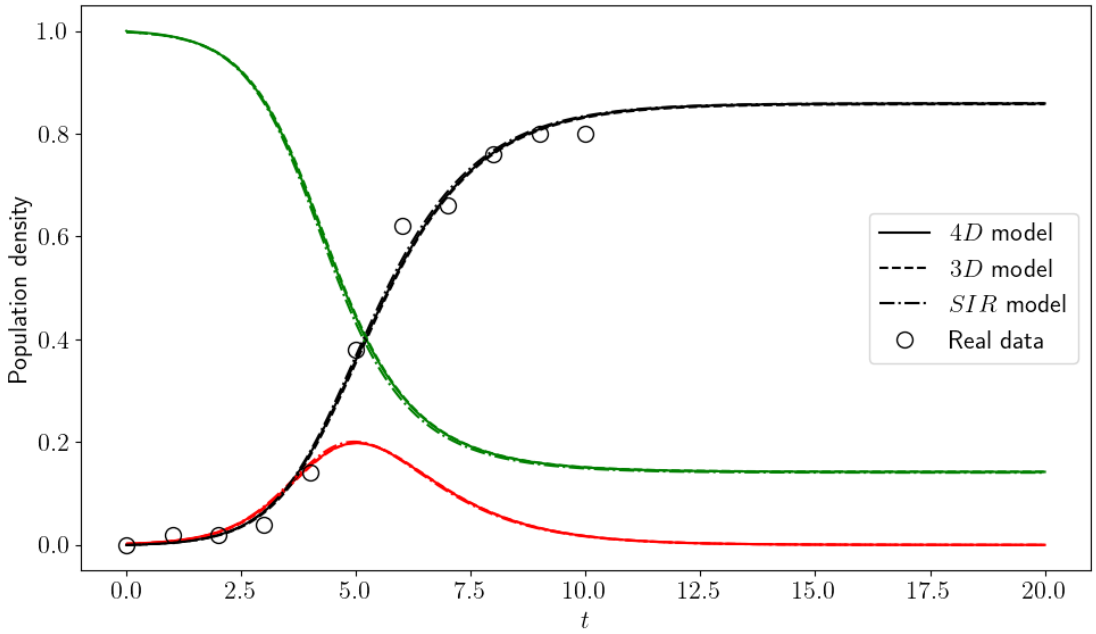


Figure 24: Collapse of the 3 models for a given set of parameters and comparison to experimental data.

6 Conclusions

In this thesis the mass mortality event of *Pinna Nobilis*, an endemic species of the Mediterranean sea, has been modelled by means of mathematical epidemiology. The study carried out in this work includes logical reasoning, mathematical analysis and computer simulations. With this set of techniques we aim to understand the epidemic taking place in the Mediterranean sea and try to perform further predictions.

More specifically, a four dimensional model of differential equations was built with three host states, susceptible, infected and removed and the parasite population. Due the complexity of the model, only the basic reproduction number, R_0, could be studied analytically using the next generation method approach. Thus, it was reduced to a 3D model by considering a separation of time scales, where the dynamics of the parasite population is much faster than the typical evolution time of the hosts. This simpler model admitted some analytical solutions, i.e. an expression for the maximum density of infected individuals and the final number of susceptible individuals. Both 4D and 3D model where studied both analytically and numerically and the numerical simulations were in good agreement with the analytical solutions. Then, the stochastic equivalent to the 4D model was built by means of a master equation for the states probabilities. Despite the high complexity of the master equation a mean field approximation could be performed, showing the equivalence with the deterministic 4D model of differential equations. Then, the stochastic model was shown to approach the deterministic one performing computer simulations by means of the Gillespie algorithm. Moreover, our initial assumption of continuous release of parasite by infected individuals was contrasted to that of a unique release at the end of the infectious period, showing good agreement between them. Finally, a ultimate reduction to our model was performed in order to fit it to experimental data. This simplification yield an original SIR model, showing that a disease fulfilling the biological considerations of the present disease can be cast as a normal SIR model with a specific relation between the general model parameters. Moreover, we were able to fit our simplified model to the experimental data showing the expected qualitative behaviour with temperature, i.e., higher value of R_0 for higher temperatures.

A clear and specific line for future work is emerging. However, due the lack of data and time, this line has not been followed in the present work but is presented in the following section.

7 Ongoing work

First, if more experimental data could be provided, further validations should be performed to our models. For instance, if some experimental values for the transition rates could be fixed, for instance the death rate of parasites, μ, our 4D model could be validated. Moreover, with more experimental data of water tanks at different controlled temperature the dependence of the parameters with temperature could be inferred. With this, the presented models could be finally validated and, in case of successfully reproducing the experimental data, the dependence of the parameters with temperature would be estimated.

Then, if enough data could be collected, the stochastic model should be also fitted. Several techniques are reported in the literature to perform parameter estimation or model fitting for stochastic models, like the maximum likelihood estimation [27]. For our particular model the probability density function for the states of our model can not be solved analytically. Thus, Monte Carlo methods should be used in order to perform the maximum likelihood method numerically [28].

Once the model and the parameters would be known and validated a metapopulation model could be built. The idea is to incorporate the parasite mobility by the oceanic currents in the metapopulation model, as shown in Fig. 25, thus studying the disease evolution in several places of the Mediterranean Sea. This would allow further predictions of the evolution of the epidemic and, perhaps, bring some hope for the fan mussels survivorship.

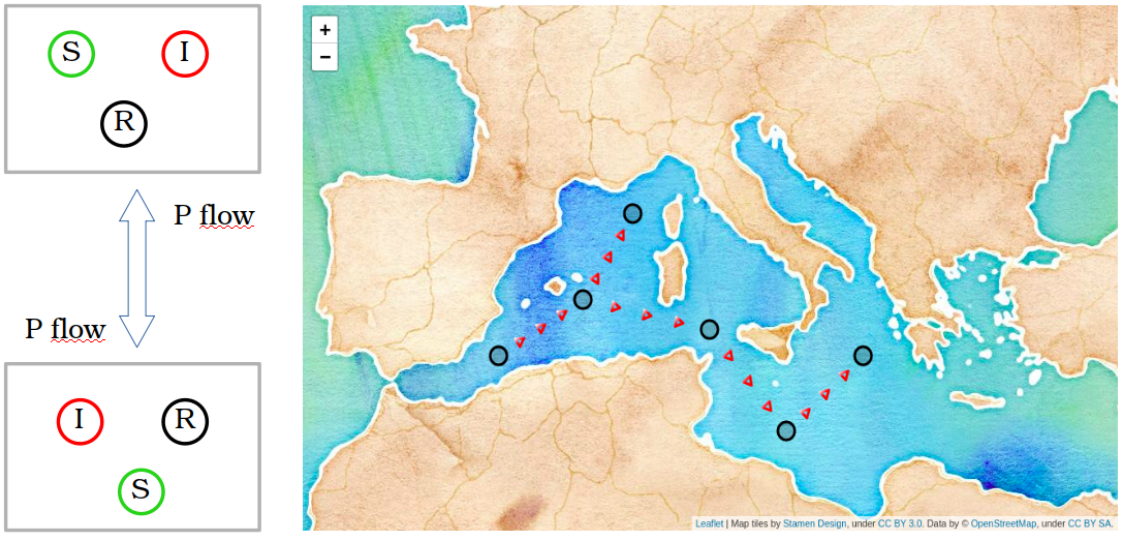


Figure 25: Metapopulation model including parasite mobility by oceanic currents

A Next-generation matrix method

We call a *disease compartment* if if the individuals therein are infected. Generally, the idea is that the disease compartments are the ones related to the production of new infections. The other compartments are then defined as the *nondisease compartments*. Let $x \in \mathbb{R}^n$ and $y \in \mathbb{R}^n$ be the subpopulations in the disease and nondisease compartments, respectively. Further, denote by $\mathcal{F}_i$ the rate secondary infections increase the i_{th} disease compartment and by $\mathcal{V}_i$ the rate disease progression, death and recovery decrease the i_{th} compartment. The compartmental model can then be written in the following form:

$$\begin{aligned} \dot{x}_i' &= \mathcal{F}_i(x,y) - \mathcal{V}_i(x,y), \quad i = 1, \dots n \\ \dot{y}_j &= g_j(x,y), \quad j = 1, \dots, m \end{aligned} \tag{A.1}$$

The derivation of the basic reproduction number is based on the linearization of the ODE model about a disease-free equilibrium. A set of assumptions ensure the existence of this disease-free equilibrium [29]. Then, it can be shown that

$$\frac{\partial \mathcal{F}_i}{\partial y_j}(0, y_o) = \frac{\partial \mathcal{V}_i}{\partial y_j}(0, y_o) = 0 \tag{A.2}$$

is satisfied, which implies the decoupling of the linearized equations of the disease compartments from the other equations. Thus we can write

$$\dot{x} = (F - V)x \tag{A.3}$$

where

$$F = \frac{\partial \mathcal{F}_i}{\partial x_j}(0, y_o) \quad \text{and} \quad V = \frac{\partial \mathcal{V}_i}{\partial x_j}(0, y_o) \tag{A.4}$$

With that, linear stability of the system is completely determined by the linear stability of the jacobian, split as $F - V$.

The number of secondary infections produced by a single infected individual can be expressed as the product of the expected duration of the infectious period and the rate secondary infections occur. The expected time an individual spends in the each compartment is given by $\int_0^\infty \phi(t, x_0)\mathrm{d}t$, where $\phi(t, x_0)$ is the solution of Eq. (A.3) when no secondary infections occur ($F = 0$). Given positive initial conditions we have

$$\dot{x} = -Vx \Longrightarrow \phi(t, x_0) = e^{-Vt}x_0 \tag{A.5}$$

Recalling that V is a matrix, the exponential e^{-Vt} can be computed from its Taylor series

$$e^A = I + A + \frac{A^2}{2} + \frac{A^3}{3!} + \cdots + \frac{A^k}{k!} + \cdots$$

Then, the integral is given by

$$\int_0^\infty \phi(t, x_0) = x_0 \int_0^\infty e^{-Vt}\mathrm{d}t = x_0 \cdot \frac{-1}{V}\left[e^{-Vt}\right]_0^\infty = -x_0 V^{-1}\left[0 - \mathbb{I}\right] = x_0 V^{-1}$$

given that the Taylor series of the exponential converges for all t.

The (i, j) entry of the V^{-1} matrix can be interpreted as the expected time an individual initially introduced into disease compartment j spends in disease compartment i [29] [22]. On the other hand,the (i, j) entry of the F matrix is the rate secondary infections are produced in compartment i by an individual in compartment j. Hence, the expected number of secondary infections produced by an individual is given by

$$K = \int_0^\infty F e^{-Vt} x_0 dt = F V^{-1} x_0 \tag{A.6}$$

Where K is the so called *next-generation matrix* where the (i, j) entry of K is the expected number of secondary infections in compartment i produced by individuals initially in compartment j. The next generation matrix, $K = FV^{-1}$, is nonnegative and therefore has a nonnegative eigenvalue, $\mathcal{R}_o = \rho\left(FV^{-1}\right)$, such that there are no other eigenvalues of K with modulus greater than $\mathcal{R}_o$ and there is a nonnegative eigenvector ω associated with $\mathcal{R}_0$.This eigenvector is in some sense the distribution of infected individuals that produces the greatest number, $\mathcal{R}_0$, of secondary infections per generation [29].

Trivial example: SIR model

For the usual SIR model we only have one disease compartment, i.e. I. Thus $\mathcal{F}$ and $\mathcal{V}$ are one dimensional vectors

$$\mathcal{F} = (\beta SI) \qquad \mathcal{V} = (\gamma I)$$

according tho their definition presented above. Then, F and V will be one dimensional matrices given by

$$F = \frac{\partial \mathcal{F}_i}{\partial x_j}(0, y_o) = (\beta S_0) \quad \text{and} \quad V = \frac{\partial \mathcal{V}_i}{\partial x_j}(0, y_o) = (\gamma)$$

Finally, the next generation matrix is given by

$$K = FV^{-1} = \frac{\beta}{\gamma} S_0$$

So that the basic reproduction number is the trivial eigenvalue, say $R_0 = \dfrac{\beta}{\gamma} S_0$.

Modified SEIR model

Consider the modified $SIER$ model given by [14]

$$\begin{aligned}
\dot{S} &= -\Pi - \mu S - \beta SI \\
\dot{E} &= \beta SI - (\mu + \kappa)E \\
\dot{I} &= \kappa E - (\mu + \alpha)I \\
\dot{R} &= \alpha I - \mu R
\end{aligned}$$

The next-generation matrix method can be applied as follows

$$\mathcal{F} = \begin{pmatrix} \beta SI \\ 0 \end{pmatrix} \qquad \mathcal{V} = \begin{pmatrix} (\mu + \kappa)E \\ -\kappa E + (\mu + \alpha)I \end{pmatrix}$$

So that

$$F = \begin{pmatrix} 0 & \beta S_0 \\ 0 & 0 \end{pmatrix} \qquad V = \begin{pmatrix} \mu + \kappa & 0 \\ -\kappa & \mu + \alpha \end{pmatrix}$$

and

$$K = FV^{-1} = \begin{pmatrix} \dfrac{\kappa \beta S_o}{(\mu + \kappa)(\mu + \alpha)} & \dfrac{\beta S_0}{\mu + \alpha} \\ 1 & 0 \end{pmatrix}$$

Finally R_0 is given by the solution of $R_0 :\ \det|K - \lambda \mathbb{I}| = 0 \Longrightarrow R_0 = \dfrac{\kappa \beta S_o}{(\mu + \kappa)(\mu + \alpha)}$.

B Lambert's W function

The so-called Lambert's W function is the inverse function of $x = f(y) = y\exp(y)$, i.e., $y = W(x)$. It can be seen as a generalised logarithm, a useful analogy, because as the (complex) logarithm function, Lambert's W function is multivalued. It is defined as,

$$x = W_l(x)\exp\left(W_l(x)\right), \quad l \in \mathbb{Z}$$

where, in principle, $x \in \mathbb{C}$ and $l \in \mathbb{Z}$ is the branch index. The principal branch, $W_0(x)$ or simply $W(x)$, has a branch point at $x = -1/e$ and a branch cut along the negative real axis $x \in [-\infty, -1/e]$ $(W_0(-1/e) = -1)$, and is real valued in the interval $x \in [-1/e, \infty]$. Moreover, it is analytic at $x = 0, W_0(0) = 0$, while all the other branches have a branch point at 0. Moreover, $W_{-1}(x)$ is real in the interval $x \in [-1/e, 0]$.

The Lambert function can also be used to find the exact solution of transcendental equations of the type $ax + b\ln x + c = 0$, as

$$\ln x = -(ax + c)/b$$

and, thus

$$\frac{a}{b}x\exp\left(\frac{a}{b}x\right) = \frac{a}{b}\exp\left(-\frac{c}{b}\right)$$

from which ax/b is the Lambert W function of the R.H.S., and one gets

$$x = \frac{b}{a}W\left[\frac{a}{b}\exp\left(-\frac{c}{b}\right)\right]$$

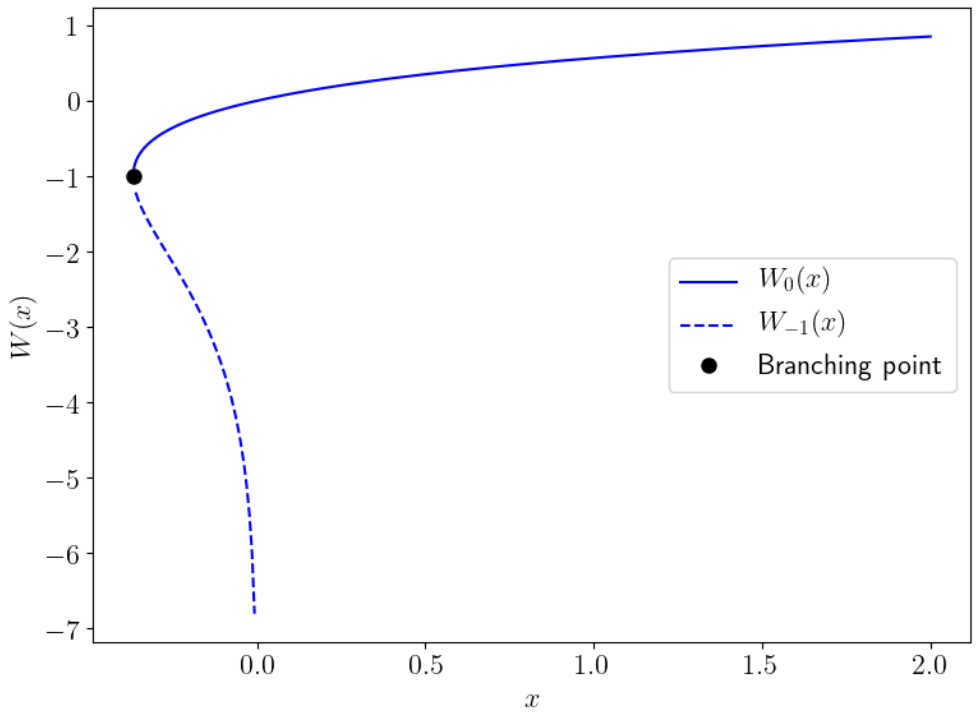


Figure 26: A plot of the two branches of the Lambert's W fucntion

References

[1] B. S. Halpern et al., "A Global Map of Human Impact on Marine Ecosystems", Science **319**, 948–952 (2008).

[2] N. Marbà, G. Jordà, S. Agustí, C. Girard, and C. M. Duarte, "Footprints of climate change on Mediterranean Sea biota", Frontiers in Marine Science **2**, 56 (2015).

[3] M. Coll et al., "The Biodiversity of the Mediterranean Sea: Estimates, Patterns, and Threats", PLOS ONE **5**, 1–36 (2010).

[4] N. Marbà, E. Diaz-Almela, and C. Duarte, "Mediterranean seagrass (Posidonia oceanica) loss between 1842 and 2009", Biological Conservation **176**, 183–190 (2014).

[5] I. E. Hendriks, M. Cabanellas-Reboredo, T. J. Bouma, S. Deudero, and C. M. Duarte, "Seagrass Meadows Modify Drag Forces on the Shell of the Fan Mussel Pinna nobilis", Estuaries and Coasts **34**, 60–67 (2011).

[6] G. Catanese, A. Grau, J. M. Valencia, J. R. Garcia-March, M. Vázquez-Luis, E. Alvarez, S. Deudero, S. Darriba, M. J. Carballal, and A. Villalba, "Haplosporidium pinnae sp. nov., a haplosporidan parasite associated with mass mortalities of the fan mussel, Pinna nobilis, in the Western Mediterranean Sea", Journal of Invertebrate Pathology **157**, 9–24 (2018).

[7] M. Cabanellas-Reboredo et al., "Tracking a mass mortality outbreak of pen shell Pinna nobilis populations: A collaborative effort of scientists and citizens", Scientific Reports **9**, 13355 (2019).

[8] G. Fiorito and F. Gherardi, "Prey-handling behaviour of Octopus vulgaris (Mollusca, Cephalopoda) on Bivalve preys", Behavioural Processes **46**, 75–88 (1999).

[9] J. García-March et al., "Can we save a marine species affected by a highly infective, highly lethal, waterborne disease from extinction?", Biological Conservation **243**, 108498 (2020).

[10] W. O. Kermack, A. G. McKendrick, and G. T. Walker, "A contribution to the mathematical theory of epidemics", Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character **115**, 700–721 (1927).

[11] W. O. Kermack, A. G. McKendrick, and G. T. Walker, "Contributions to the mathematical theory of epidemics-II. The problem of endemicity", Proceedings of the Royal Society of London. Series A, Containing Papers of a Mathematical and Physical Character **138**, 55–83 (1932).

[12] W. Kermack and A. McKendrick, "Contributions to the mathematical theory of epidemics-III. Further studies of the problem of endemicity", Bulletin of Mathematical Biology **53**, 89–118 (1933).

[13] H. I. McCallum, A. Kuris, C. D. Harvell, K. D. Lafferty, G. W. Smith, and J. Porter, "Does terrestrial epidemiology apply to marine systems?", Trends in Ecology & Evolution **19**, 585–591 (2004).

[14] F. Brauer, "Compartmental Models in Epidemiology", in *Mathematical Epidemiology*, edited by F. Brauer, P. van den Driessche, and J. Wu (Springer Berlin Heidelberg, Berlin, Heidelberg, 2008), pages 19–79.

[15] J. D. Murray, "Dynamics of Infectious Diseases: Epidemic Models and AIDS", in *Mathematical Biology: I. An Introduction*, edited by J. D. Murray (Springer New York, New York, NY, 2002), pages 315–394.

[16] R. Martínez-Peña, "A Consumer-Resource Description of Public-Goods Production in Microbes", Master's thesis (UIB-IFISC, 2018).

[17] R. Toral and P. Colet, "Introduction to Master Equations", in *Stochastic Numerical Methods* (John Wiley & Sons, Ltd, 2014) Chap. 8, pages 235–260.

[18] R. Toral and P. Colet, "Numerical Simulations of Master Equations", in *Stochastic Numerical Methods* (John Wiley & Sons, Ltd, 2014) Chap. 9, pages 261–274.

[19] L. Canesi and P. Procházková, "Chapter 7 - The Invertebrate Immune System as a Model for Investigating the Environmental Impact of Nanoparticles", in *Nanoparticles and the Immune System*, edited by D. Boraschi and A. Duschl (Academic Press, San Diego, 2014), pages 91–112.

[20]R. M. May and R. M. Anderson, "Population biology of infectious diseases: Part II", Nature **280**, 455–461 (1979).

[21]M. E. J. Woolhouse, C. Dye, A. Dobson, and J. Foufopoulos, "Emerging infectious pathogens of wildlife", Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences **356**, 1001–1012 (2001).

[22]O. Diekmann, J. A. P. Heesterbeek, and M. G. Roberts, "The construction of next-generation matrices for compartmental epidemic models", Journal of The Royal Society Interface **7**, 873–885 (2010).

[23]R. M. May and R. M. Anderson, "Population biology of infectious diseases: Part II", Nature **280**, 455–461 (1979).

[24]J. Cariboni, D. Gatelli, R. Liska, and A. Saltelli, "The Role of Sensitivity Analysis in Ecological Modelling", Ecological Modelling **203**, 167–182 (2007).

[25]G. Bidegain, E. N. Powell, J. M. Klinck, T. Ben-Horin, and E. E. Hofmann, "Marine infectious disease dynamics and outbreak thresholds: contact transmission, pandemic infection, and the potential role of filter feeders", Ecosphere **7**, e01286 (2016).

[26]H. Banks, S. Hu, and W. Thompson, *Modeling and inverse problems in the presence of uncertainty* (2014).

[27]F. W. Scholz, "Maximum Likelihood Estimation", in *Encyclopedia of Statistical Sciences* (American Cancer Society, 2006).

[28]F. Hartig, J. M. Calabrese, B. Reineking, T. Wiegand, and A. Huth, "Statistical inference for stochastic simulation models – theory and application", Ecology Letters **14**, 816–827 (2011).

[29]P. van den Driessche and J. Watmough, "Further Notes on the Basic Reproduction Number", in *Mathematical Epidemiology*, edited by F. Brauer, P. van den Driessche, and J. Wu (Springer Berlin Heidelberg, Berlin, Heidelberg, 2008), pages 159–178.