Agent-Based Modelling of Infectious Diseases

Hayman Gosain
(MS15125)

A dissertation submitted for the partial fulfillment of BS-MS dual degree in Science

Indian Institute of Science Education and Research Mohali
June 2020

Certificate of Examination

This is to certify that the dissertation titled "Agent-Based Modelling of Infectious Diseases" submitted by Mr. Hayman Gosain (Reg. No. MS15125) for the partial fulfilment of BS-MS dual degree programme of the institute, has been examined by the thesis committee duly appointed by the institute. The committee finds the work done by the candidate satisfactory and recommends that the report be accepted.

Dr. Abhishek Chaudhuri	Prof. Somdatta Sinha	Dr. Rajeev Kapri
(Supervisor)	(Co-supervisor)	

Date: 15/06/2020

Declaration

The work in this dissertation has been carried out by me under the guidance of Prof. Somdatta Sinha and Dr. Abhishek Chaudhuri at the Indian Institute of Science Education and Research Mohali.

This work has not been submitted in part or in full for a degree, a diploma, or a fellowship to any other university or institute. Whenever contributions of others are involved, every effort is made to indicate this clearly, with due acknowledgement of collaborative research and discussions. This thesis is a bonafide record of original work done by me and all sources listed within have been detailed in the bibliography.

Hayman Gosain
(Candidate)
Date: 15/06/2020

In our capacity as the supervisors of the candidate's project work, we certify that the above statements by the candidate are true to the best of our knowledge.

Dr. Abhishek Chaudhuri
(Supervisor)

Prof. Somdatta Sinha
(Co-supervisor)

Acknowledgements

I wish to record a deep sense of gratitude to Prof. Somdatta Sinha for her valuable guidance and constant support, providing insights and encouragement at all stages of my thesis. Her endeavour for excellence and persistent efforts in this period of time serve as an inspiration which I shall continue to draw from. I would also like to thank Dr. Abhishek Chaudhuri for his generous help and feedback, and Dr. Rajeev Kapri for his comments on the work, which also facilitated the thesis work.

I also take the occasion to acknowledge the discussions and meetings of the Mathematical Modelling and Computational Biology group with the lab members; a special thanks to Inayat, Himanshu, Saurabh, and Sveekruth for the tea breaks and the fun amidst the work as well.

The mentorship and encouragement of parents are indispensable to any work done by a person. My temperament for questioning the workings of nature and critical thinking is due to the support and guidance of my parents, Deepa Gossain and Raman Kumar, for which I'm eternally thankful.

I lastly thank Gaurav, Kabeer, Shridhar, Tinku and my other hostel mates for being by my side and for all the enjoyment to keep me going in this year.

Hayman Gosain

List of Figures

List of Tables

Contents

Abstract

Understanding the spread of infectious disease in a population of susceptible individuals is an active field of theoretical research involving interdisciplinary fields. The most common approach is epidemiological compartment models (commonly known as the SIR models) represented by differential equations, where the host-pathogen interaction is studied assuming the transfer of individuals between the different infection states, such as the Susceptible, Infected and Recovered classes. This is a mean-field approach and does not consider heterogeneity among individuals in a class/compartment. A microscopic approach to study spread of infection is through the Agent Based Models (ABM), where each agent is an individual at a particular infection state, and the rules of interactions describe the change in infection states.

The work presented in this thesis first studies the characteristics of different types of differential equation epidemiological compartment models (SI, SIS, SIR, and SIRS) using mathematical and computational analysis of steady states and their temporal dynamics. As an application of this approach to specific diseases, the basic model of Malaria by Ronald Ross is studied. With an aim to compare and contrast the mean field models with the microscopic approach, a detailed study of the equivalent ABMs is developed and analysed. Analytic solutions and numerical results from differential equation models are compared with individual and averaged dynamics of the Agent-Based Models over 50 runs for many parameter sets and initially infected agent densities.

An in-depth comparison between the two approaches of modelling shows that ABMs can lead to low probability states depending on initial conditions, even with the same parameters - a behaviour that is absent in mean field models. The role of spatial structure in the distribution of the same number of initial infected agents is shown to impact the rate of reaching the steady states along with their dynamics. The presence of oscillations, a feature completely absent in differential equation models, is observed in ABMs. Similar studies are done with the ABM realization of the Ross model as well. These results clearly show that in realistic situations, i.e.,

when a particular infection spreads through the host population, the dynamics and steady state behaviour of the diseased states may exhibit differences due to the inherent randomness in the agents, even in absence of other biological or social factors.

Chapter 1

Introduction

1.1 Background

The study and control of infectious diseases is a critical part of sustainable development in the modern globalised world, where the outbreak of epidemics is responsible for the morbidity of millions of people each year and the cause of 17% of global deaths in addition to high economic costs [GBD 14]. Creation of strategies for management of infectious diseases assumes even more importance for mutating pathogens and new strains of diseases which can cause worldwide pandemics like the 2019 Coronavirus epidemic which has claimed over 300,000 lives as of print.

Infectious diseases refer to communicable diseases, the class of diseases caused by agents or pathogens that can spread to other individuals through a medium or by contact with bodily fluids. The causative agents may be viruses, bacteria, parasites, arthropods, or fungi. Epidemiology, the study of infectious diseases by understanding the profile of infected people such as age, environment, genetics etc., the geographical spread and the temporal evolution of disease, thus aims to examine the spatio-temporal patterns of disease prevalence in order to predict and control it.

1.2 Infectious diseases modelling

The tools of epidemiological analysis of infectious diseases involve modelling them by using mathematical functions to describe the host-pathogen interactions based on biological facts, and using statistical modelling methods on the past and current disease prevalence and incorpo-

rating these as parameters, as well as other environmental or epidemiological parameters such as vaccination coverage. There are also detailed statistical data analysis and prediction methods available that help to determine disease trend, which have important applications in public health to control the spread of disease.

1.3 Types of modelling

A system can be studied and modelled at different scales of detail-

- Macroscopic scale with continuous space and time, such as differential equations
- Mesoscopic scale methods like lattice systems, where space is discrete but time is still continuous
- Microscopic scale having both discrete space and time

Below the macroscopic and microscopic methods are discussed as they have been used in this study.

1.3.1 Differential equation modelling

The use of mathematical modelling for diseases began back in 1766 by Daniel Bernoulli who applied it to smallpox [Bernoulli 66, Blower 04], and since then the tools of differential equations began to be commonly used. The development of compartmental models to examine the spread of diseases occurred dates back to the 1920s with the Kermack-McKendrick model [Kermack 27, Brauer 01].

Compartmental models involve the creation of different classes or compartments in a population which are in different stages of a disease. These models make the simplification that all individuals belonging to a certain state or compartment are identical. Common classes/compartments include-

$$\frac{dS}{dt} = -\beta SI$$
$$\frac{dI}{dt} = \beta SI - \gamma I$$
$$\frac{dR}{dt} = \gamma I$$

(*a*) ODE version

$$\frac{\partial S}{\partial t} = -\beta SI + D_S \frac{\partial^2 S}{\partial x^2}$$
$$\frac{\partial I}{\partial t} = \beta SI - \gamma I + D_I \frac{\partial^2 I}{\partial x^2}$$
$$\frac{\partial R}{\partial t} = \gamma I + D_R \frac{\partial^2 R}{\partial x^2}$$

(*b*) PDE version

Figure 1.1: The Susceptible-Infected-Recovered (SIR) model as (a) ODEs and (b) PDEs with a spatial dimension. Compartment transfer rate constants are β and γ, and diffusion constants D_S, D_I and D_R

- Susceptible (S), referring to individuals prone to contracting the infection and not having immunity against the disease
- Exposed (E), having acquired the infection but present in a latent manner without the capability to spread
- Infected (I), having acquired the infection and able to spread it
- Recovered (R), having no infection and immune to further infection

Ordinary differential equation models (ODEs) describe the change in the population size in each compartment with time, whereas partial differential equation models (PDEs) also account for the dependence of infection in different compartment on space or any other variable such as age structure, environmental variables like temperature, humidity etc.

Differential equation-based modelling, especially using the law of mass action, has been one of the earliest methods of modelling diseases where transition rates define the rate of conversion of one class of individual into another based on interaction in a deterministic fashion. Differential equation modelling may include spatial dimensions and involve diffusion, but all of them are mean field in nature, where interaction is effectively between the population numbers or fractions [Edelstein-Keshet 88]. This is useful for large populations as a good approximation but it exhibits only homogeneous interactions, and heterogeneous interactions and rare states are left out [Agrawal 17, Liu 09].

1.3.2 Agent-based modelling

Agent based modelling (ABM) is a microscopic or bottom-up technique to study interacting systems in discrete space and time which makes use of tracking each agent, i.e., an autonomous individual, to give rise to emergent behaviour [Railsback 19]. For ABM simulations we use Netlogo [Wilensky 99, Wilensky 15], a multi-agent open source programming software which is also an Integrated Development (IDE) for modelling, which may even be integrated with python [Gunaratne 18]. Agent-Based modelling using Netlogo finds many applications across fields as diverse as mathematical finance [Kurahashi 19], immigrant dynamics [Perez 19], life cycle modelling [Wang 19] among others.

Interactions in the model can also incorporate environment, and heterogeneity is accounted for in the system due to knowledge of agent properties. This can result in characteristics both in the mean field and low probability or rare events. An ABM thus has the following features-

- Autonomous agents that have local interaction rules such that no long distance/averaging interactions are present by default
- Interaction rules governing the behaviour of the model instead of equations
- A explicit spatial environment in 2 or 3 dimensions with which the agents can interact and vice-versa

A more detailed description of this methodology is discussed in Chapter 2.

1.4 Malaria modelling

Malaria is one of the oldest infectious diseases plaguing the earth from antiquity. Malaria parasite requires two hosts (human and mosquito) for its life cycle to complete, and the disease is spread through mosquito bites by infected mosquitoes and from infected humans [Anderson 91]. The use of modelling in malaria is of special importance due to the persisting presence of the disease and about 228 million yearly infections and mortality of 405,000 in 2018, despite elimination in certain regions over past decades [WHO 19]. Resistance of the parasite to drugs and

mosquitoes to insecticides as well as warming temperatures that facilitate the survival of malarial parasites present emerging risks that necessitate the understanding of its dynamics and both preventative and public health intervention strategies.

The first malaria model was developed by Sir Ronald Ross in 1910 [Ross 10] to explain the population relationship in humans and mosquitoes which was further developed by incorporating biologically realistic elements like latency of parasite development [Anderson 91, MacDonald 57], age dependent susceptibility [Anderson 91, Aron 82], the time dependent immunity conferred by infection [Aron 82, Aron 88], and spatial and genetic heterogeneity in humans and parasites [Hasibeder 88, Gupta 94]. After the development of Agent-Based Models, such an evolution accounting for the complexities of the disease and accurate reproduction of the dynamics was also attempted. [Gu 05, Smith 06, Smith 18]

1.5 Thesis outline

The work embodied in this thesis is presented in five chapters. Along with a complete review (with mathematical analysis and numerical simulations of equations) of the epidemiological compartmental ODE models and the ABM method, the equivalent ABMs are developed and analysed. Analytic solutions and numerical results from differential equation models are compared with individual and averaged dynamics of the Agent-Based Models. Similar studies are done with the ABM realization of the Ross model also. The aim is to compare and contrast differential equations with the effects of an ABM in compartmental disease models and the Ross malaria model while varying across parameters. The chapter-wise outlines are as under-

1. Chapter 1 gives the introduction to the thesis work.

2. Chapter 2 explains the methods and analysis of compartmental differential equations of SI, SIS, SIR, and SIRS models, and the basic ODE malaria model (Ross model) along with basic mathematical results known. The method of Agent-Based Models is also introduced here.

3. Chapter 3 presents the results of the Agent-Based Models corresponding to the compartmental differential equation models (SI, SIS, SIR, and SIRS).

4. Chapter 4 discusses the ABM realization of the Ross malaria model along with comparisons with the ODE model.

5. Chapter 5 gives a summary of the results presented in the thesis, and proposes further work that may be done.

Chapter 2

Models and methods

2.1 Differential equations methods

The differential equation time series have been produced using numerical integration in C with the Runge-Kutta 4 (RK4) method [Butcher 63]. Calculations for the stability conditions of the different types of fixed points are analytic except for the Ross model. All numerical solutions of the compartmental equations assume the total population, N=1250, which stays conserved across all models; the time step of integration adopted universally is t=0.01 units. We refer to susceptible individuals as 'susceptibles', infected ones as 'infecteds' and likewise hereon. We shall colour compartmental population time series in the following shades- Susceptible, Infected and Recovered.

2.2 Epidemiological models

2.2.1 SI model

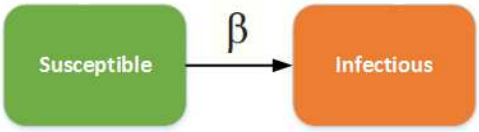


Figure 2.1: Schematic of the SI model

$$\frac{dS}{dt} = -\beta SI \qquad \frac{dI}{dt} = \beta SI$$

Figure 2.2: The SI system of equations

The Susceptible-Infected (SI) model leads to spread of disease from the infected to susceptible class by the infection rate β (Figure 2.1). Example diseases include those with high percentage of mortality like Bovine Spongiform Encephalopathy (BSE), rabbit haemorrhagic disease, Leishmaniasis etc.

The stability analysis of the system leads to a single fixed point of complete infection (I_∞ = N) state, independent of the initial conditions, illustrated by systems with different initial conditions and rates in Figure 2.3. Higher I_0 in system (S1,I1) leads to faster approach to the fixed point despite lower β, but both systems 1 and 2 converge to the fixed point (S_f,I_f)=(0,1250).

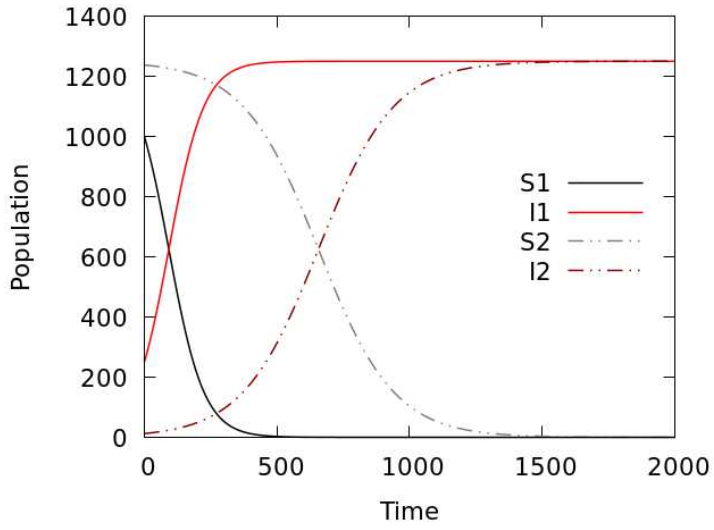


Figure 2.3: Illustrative states (S1,I1) when ($\beta, I_0/N$) = (1.5,0.2) and (S2,B2) when ($\beta, I_0/N$) = (5,0.01); N=1250

2.2.2 SIS model

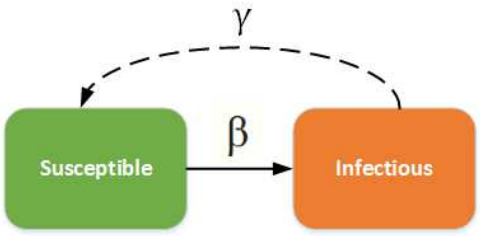


Figure 2.4: Schematic of the SIS model

$$\frac{dS}{dt} = -\beta SI + \gamma I \qquad \frac{dI}{dt} = \beta SI - \gamma I$$

Figure 2.5: The SIS system of equations

In the SIS model, after the spread of disease from the infected to susceptible class by infection rate β, infected individuals get rid of the infection with rate γ without acquiring immunity in the process (Figure 2.4). Diseases which follow this model are retroviruses, sexually transmitted infections and bacterial infections.

The dynamics of the SIS system can exhibit an endemic state, which is a fixed point with non-zero values of (S_f, I_f), but also a disease free state (DF), which is a fixed point with (S_f, I_f) = (N,0). The incidence of a DF state occurs when $\beta/\gamma \leq 1$, else the fixed point achieved is $\left(\frac{N\gamma}{\beta}, N\left(1-\frac{\gamma}{\beta}\right)\right)$, with the final state independent of initial conditions. This is illustrated by systems with different initial conditions and rates in Figure 2.6, where system (a) having β/γ = 1.6 approaches the endemic fixed point of (S_f, I_f) = (781.25,468.75), while (b) where β/γ = 0.875 approaches the disease free state (S_f, I_f) = (1250,0)

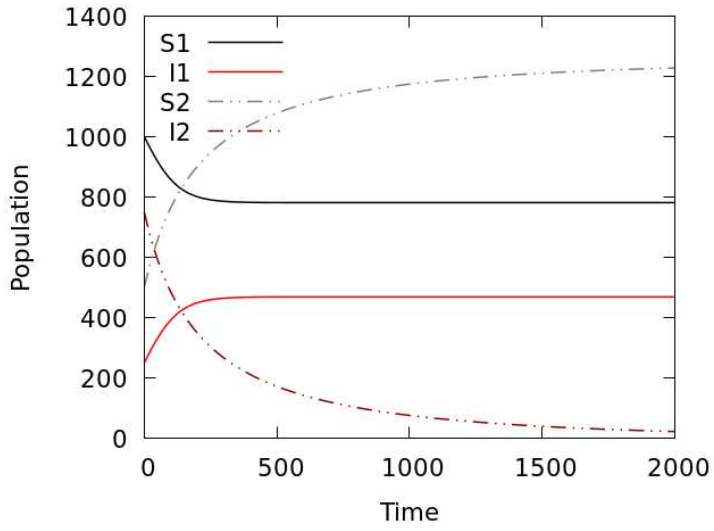


Figure 2.6: Illustrative endemic state (S1,I1) when (β, γ) = (4,2.5) and disease free state (S2,I2) when (β, γ) = (0.7,0.8); N=1250

2.2.3 SIR model

Figure 2.7: Schematic of the SIR model

$$\frac{dS}{dt} = -\beta SI \qquad \frac{dI}{dt} = \beta SI - \gamma I \qquad \frac{dR}{dt} = \gamma I$$

Figure 2.8: The SIR system of equations

In the SIR model, individuals after getting infected with rate β recover from the infection with rate γ, acquiring immunity in the process (Figure 2.7). Diseases such as measles, mumps and rubella can be modelled using the SIR model where immunity is lifelong.

The SIR system can only approach the fixed point of a disease free state since all individuals ultimately become immune, but the intermediate stages of the disease may have an epidemic in which I(t) has a maxima, or I(t) may alternatively decrease monotonously to the fixed point

$(S_\infty, I_\infty = 0, R_\infty)$; S_∞ for general (S_0, I_0, R_0) is given by the transcendental equation

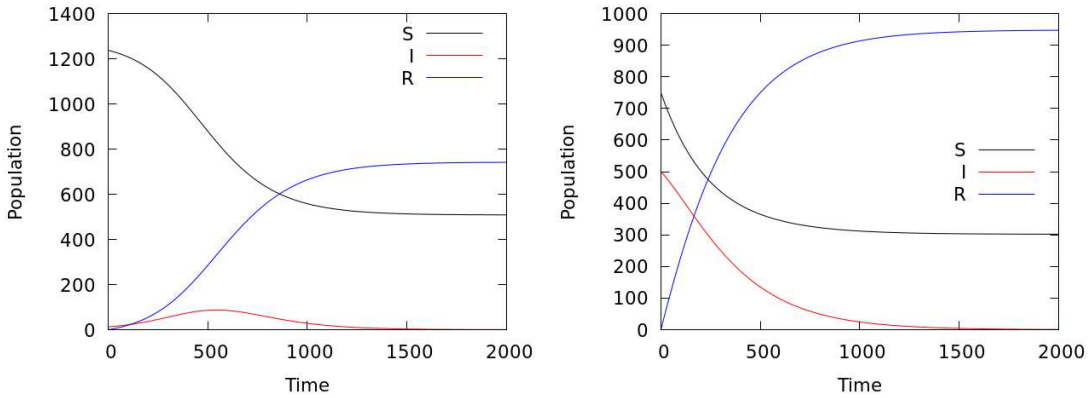
$$1 - \left(\frac{R_0 - S_\infty}{N}\right) + \frac{\gamma}{\beta}\left[\ln\left(\frac{S_\infty}{S_0}\right)\right] = 0$$

and thus $R_\infty = N - S_\infty$. If the epidemic occurs, the maxima is similarly given in general as

$$\frac{I_{\max}}{N} = 1 - \frac{R_0}{N} - \frac{\gamma}{\beta}\left[\ln\left(\frac{\beta S_0}{\gamma N}\right) + 1\right]$$

The epidemic is seen when $\left(\frac{\beta S_0}{\gamma N}\right) \leq 1$, else the disease simply goes extinct; this model notably depends on the initial state as seen in the epidemic condition above.

Figure 2.9 illustrates systems with different initial conditions and rates where (a) satisfies the epidemic condition with $\left(\frac{\beta S_0}{\gamma N}\right) = 1.485$, and (b) where $\left(\frac{\beta S_0}{\gamma N}\right) = 0.72$, and infected population decays monotonously to the fixed point.



(a) $(S_0/N, \beta, \gamma)$ = (0.99,1.8,1.2)

(b) $(S_0/N, \beta, \gamma)$ = (0.6,0.6,0.5)

Figure 2.9: Illustrative time series (a) with epidemic and (b) with no epidemic; N=1250

2.2.4 SIRS model

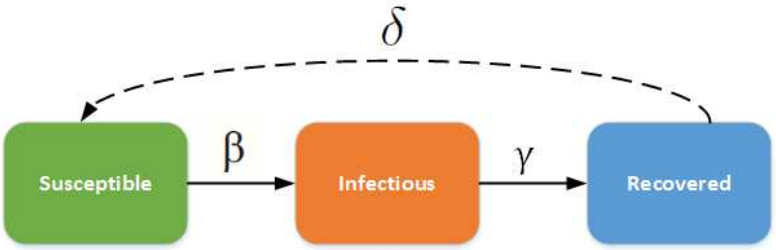


Figure 2.10: Schematic of the SIRS model

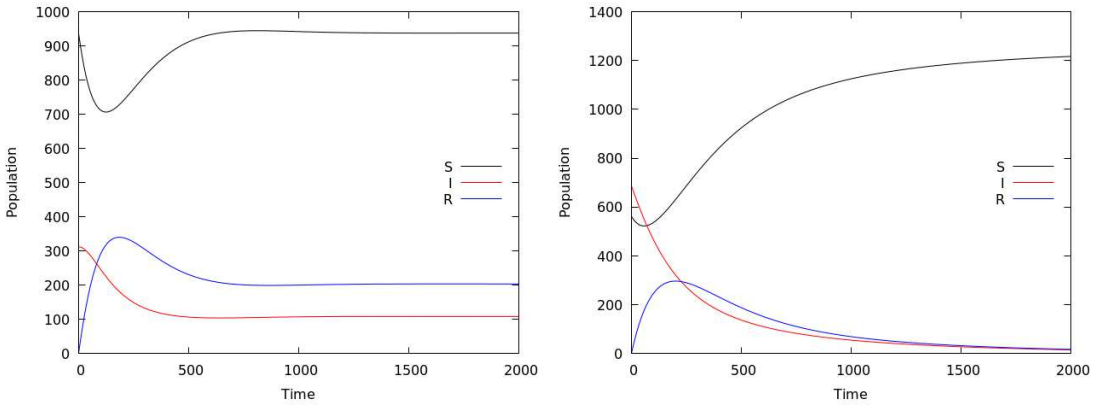
$$\frac{dS}{dt} = \delta R - \beta SI \qquad \frac{dI}{dt} = \beta SI - \gamma I \qquad \frac{dR}{dt} = \gamma I - \delta R$$

Figure 2.11: The SIRS system of equations

In the SIRS model, individuals after getting infected with rate β recover from the infection with rate γ but only acquire transient immunity which is further lost with rate δ. Seasonal influenzas present disease progression of this form with immunity lost over time (Figure 2.10).
The SIRS model can also show the dynamic states of a disease free state or an endemic state, the former when $\beta/\gamma \leq 1$ and the latter otherwise. The fixed point achieved in the case of an endemic state is

$$(S_f, I_f, R_f) = \left(\frac{N\gamma}{\beta}, N\delta\left[\frac{1-\gamma/\beta}{\gamma+\delta}\right], N\gamma\left[\frac{1-\gamma/\beta}{\gamma+\delta}\right]\right)$$

and (N,0,0) in the disease free case, again independent of initial conditions. We illustrate systems with different initial conditions and rates in Figure 2.12, where (a) approaches an endemic fixed point with $(S_f, I_f, R_f) \simeq$(937.5,108.7,203.8) while (b) approaches the disease free state of (S_f, I_f, R_f) = (1250,0,0)



(a) (b)

Figure 2.12: (a) Illustrative endemic state when (β, γ, δ) = (2, 1.5, 0.8) and (b) disease free state at (β, γ, δ)= (0.4,0.5,0.55); N=1250

2.3 Malaria model (Ross model)

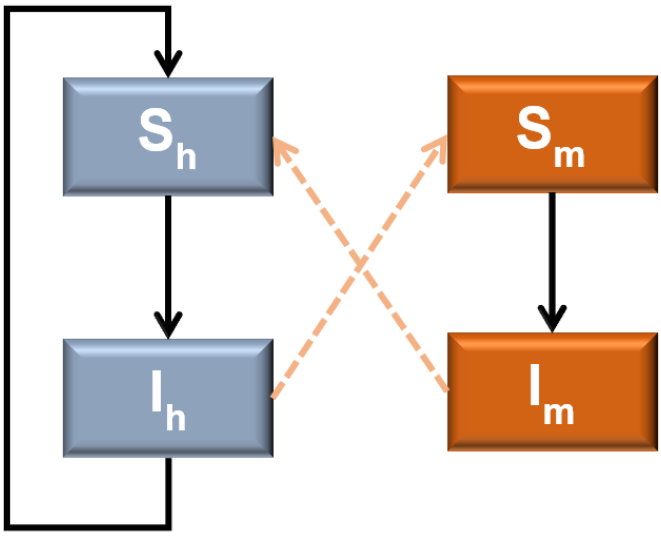


Figure 2.13: Schematic of the Ross model

$$\frac{dI_h}{dt} = a\,b\,m\,I_m\,S_h - \gamma I_h$$
$$\frac{dI_m}{dt} = a\,c\,I_h\,S_m - \mu I_m$$

Figure 2.14: The Ross model system of equations

The Ross model [Ross 10] consists of compartments for both the humans and mosquitoes as below-

1. Susceptible humans (S_h), not having immunity to the disease,
2. Infected humans (I_h), having acquired the infection and are active spreaders,
3. Susceptible mosquitoes (S_m), not having immunity to the disease,
4. Infected mosquitoes (I_m), having acquired the infection and are active vectors.

Infection in both cases does not assume superinfection/reinfection. The variables in the differential equations (Figure 2.14) are the following-

* The man biting rate, a, which is the proportion of mosquitoes that feed on humans in a day [Smith 12],
* The infectious fraction in humans, b, the fraction of infectious bites by mosquitoes leading to contraction of infection in humans,
* The infectious fraction in mosquitoes, c, the fraction of bites by susceptible mosquitoes from an infected human leading to them getting infected,
* The ratio of the density of female mosquitoes to humans, m,
* The death rate of mosquitoes, μ,
* The recovery rate of humans, γ.

We note that the system is effectively a coupled system of SIS model in humans and SI model with death for mosquitoes, having no immunity for either class of individuals (Figure

2.13)

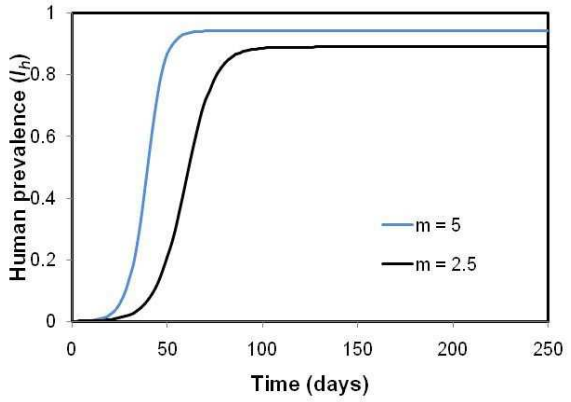
This system has a disease free state of (I_h^*, I_m^*) = (0,0) and an endemic state

$$(I_h^*, I_m^*) = \left(\frac{a^2bcm - \gamma\mu}{a^2bcm + ac\gamma} \ , \ \frac{a^2bcm - \gamma\mu}{a^2bcm + abm\mu} \right)$$

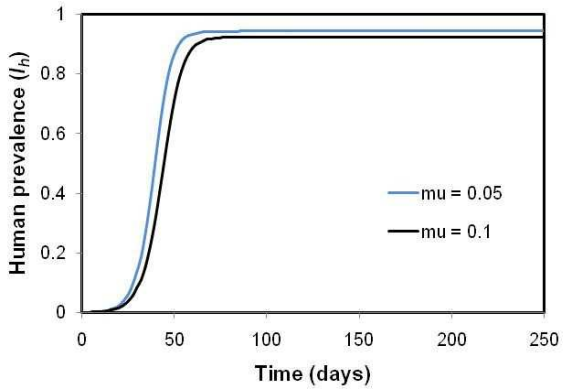
We leave the Jacobi matrix as a function of I_h^* and I_m^* for evaluation of the stability of the states due to the 6 dimensional parameter space,

$$\begin{pmatrix} abmI_m^* - \gamma & abm(1 - I_h^*) \\ ac(1 - I_m^*) & acI_h^* - \mu \end{pmatrix}$$

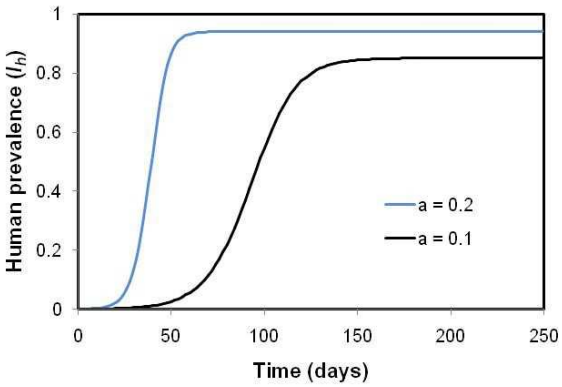
The effect of parameter changes in the Ross model on the fraction of infected humans is depicted in Figure 2.15-



(a)



(b)



(c)

Figure 2.15: Effect of changing (a) m, (b) μ, and (c) a on the prevalence curves (as a fraction of total humans). The reference curve in blue is a=0.2, b=0.5, m=5, γ=0.02, c=0.5, μ=0.05

2.4 Agent-Based Model methods

The agent-based models are produced in Netlogo, making use of the parameter sweeps that can be made to understand the impact of varying parameters as well as heterogeneity from the microscopic level detail. The simulations consist of different compartments as before where individuals have the same state of a disease –

1. Susceptible (S), not having immunity to the disease,
2. Infected (I), having acquired the infection but not liable to reinfection/ superinfection,
3. Recovered (R), having no infection and immune to further infection

The initialisation of the models is in an environment of a grid of 65×65 (4225 in total) patches closed on all sides such that no grid edges are connected, with an initial I_0 number of individuals infected and the rest (S_0) susceptible to the disease. The initial spatial distribution of the agents is randomly done in the grid such that no more than one agent occupies a site. For each ABM simulation, a set of 50 runs is performed with different random configurations of initial distribution of agents to span the entire space-grid. We note that the rate constants in differential equations are now replaced by transition timescales which dictate the progression of an agent from one compartment to the next.

The analysis of oscillatory time series involves the use of Discrete Fourier Transforms (DFTs), defined for a time series x(t) with N data points and sampling frequency f as the following-

$$X(\omega) = \sum_{t=0}^{N-1} x(t) e^{-i\frac{2\pi ft}{N}}$$

We look at the normalised power spectrum or 'probability' ,

$$P(\omega) = \frac{|X(\omega)|^2}{\sum_{\omega=1/N}^{(N-1)/N} |X(\omega)|^2}$$

excluding the constant term at $\omega = 0$, where |z| is the modulus of a complex number z. The implementation of the DFT uses the python library rfft in scipy, which is a Fast Fourier Transform (FFT), speeding calculations by the fact that Discrete Fourier Transforms of a real time series are symmetric about $\omega = N/2$ depending on N even or odd, and that FFTs with N of the form N= $N_1 N_2$ can be broken into independent calculations of N_1 DFTs of size N_2. [Cooley 65] The Full Width at Half Maxima (FWHM) is the width of a peak measured at half the maxima value to quantify its spread. The implementation of FWHM calculation is by finding the difference of frequencies at which $\mathrm{P}(\omega) \geq P_{\max}/2$.

2.4.1 Parameters

The models consist of a total N number of agents (fixed at 1250, ~30% of total patches) and initial compartmental numbers S_0, I_0 and R_0 (R_0=0 always); the models always begin with I_0=1, 5 or 10 with the rest of the initial populations given by S_0=N-I_0. Depending on the model, transition timescales $\tau_{A\to B}$ can also be defined, which represents the time steps taken for an agent in compartment A to transition to compartment B, A and B $\in$[S,I,R] ; say $\tau_{I\to R}$, which represents the time steps after which an infected will become recovered. The timescales used in the compartmental models for simulations were $\tau_{A\to B}$=10,15,20,25 and 30 for all appropriate A and B. Parameters for which time series are presented are chosen to be extreme (minimum and maximum) of the above ranges to highlight differences on account of changing parameters. Error bars when indicated for the ABM results, present mean (μ) +/- one sample standard deviation (σ_s). Simulations are performed for 2500 steps except the SI model, where the simulation dynamically stops when complete infection occurs, or all of the agents are infected. Zoomed time series are presented for the first 1000 steps to study features of the compartmental ABMs. Also, for each ABM simulation, the set of 50 runs performed are shown as light coloured time series in the background superposed with the mean trajectories as solid lines, where the mean time series does not include those cases where the disease has gone extinct.

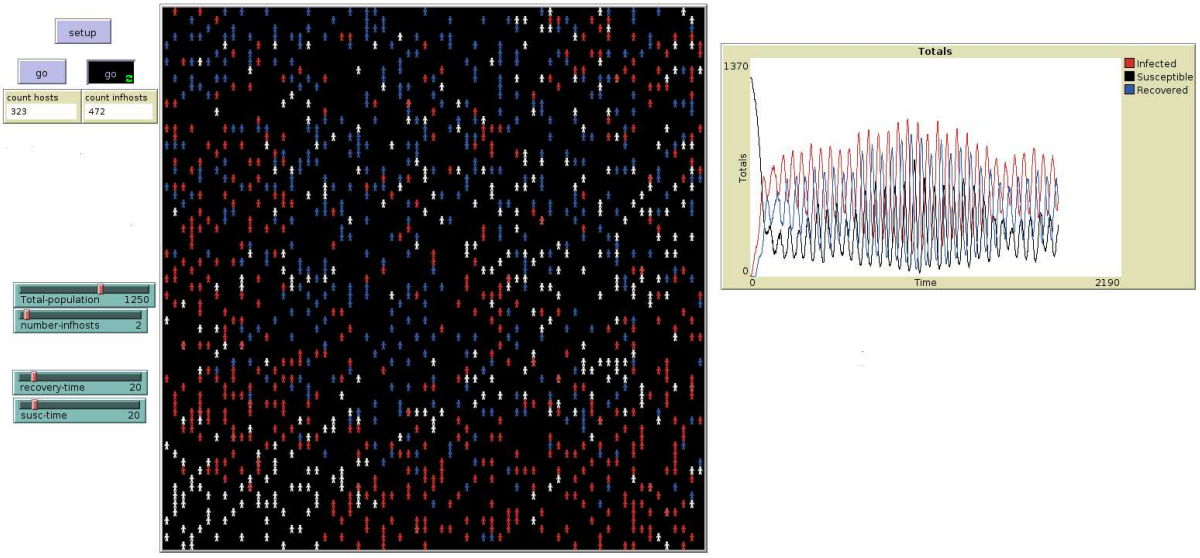


Figure 2.16: An illustration of the Netlogo interface is shown above where we observe the initialisation buttons and plot monitors (top left), parameter sliders (bottom left), the simulation space (centre) and a population graph (top right)

2.4.2 Movement and infection rules

The compartmental model ABM rules are designed to allow only 1-to-1 interactions in a patch such that they are non-ambiguous in nature and thus more than 2 individuals in a patch are not allowed. Infection spreads when an infected is present in a patch with a susceptible, referred to as 'contact'. An agent is allowed to move into its Moore neighbourhood (the set of 8 nearest sites), subject to the constraints below (if the classes exist in a given model)-

- Susceptibles are only allowed to move into empty patches
- Infecteds can move into a patch having no other infecteds
- Recovereds similarly move into patches having no infecteds

Chapter 3

ABM of Epidemiological compartmental models

In this chapter we present the results of the realisation of different epidemiological compartment models to study infection propagation in population using the Agent-based approach. The dynamic behaviour of the compartments in the differential equation approach (mean field description) are given in Chapter 2. Here the results obtained using the microscopic approach of ABM for equivalent compartment models are compared and contrasted. The interaction rules of ABMs are as per section 2.4.2 in Chapter 2. We shall colour compartmental ABM population time series in the following shades- Susceptible, Infected and Recovered.

3.1 Two agent (SI) model

Figure 3.1: Schematic of the SI ABM

The ABM representation of the compartments is modelled by distributing (N-I_0) agents who are named as Susceptible (S) and I_0 agent(s) who are in the Infected (I) state at t=0 for all ABM models. The SI model is one where the susceptible-infected interaction leads to immediate spread of disease from the infected to susceptible class until all agents are infected.

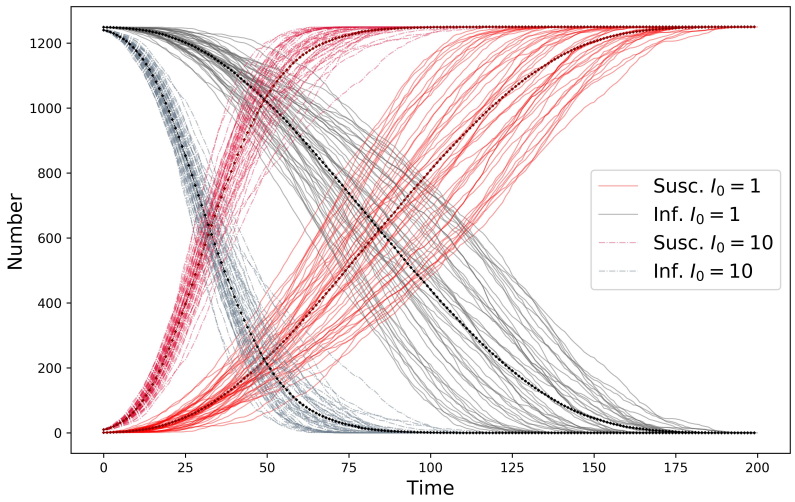
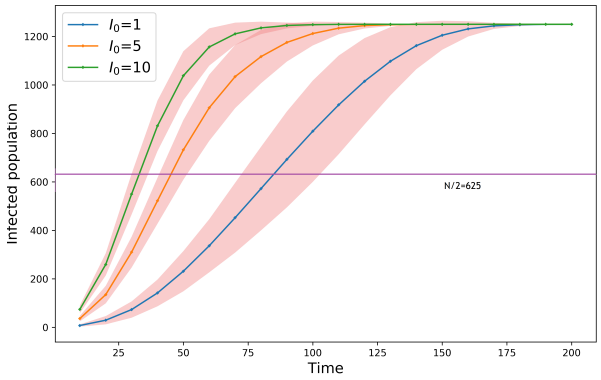
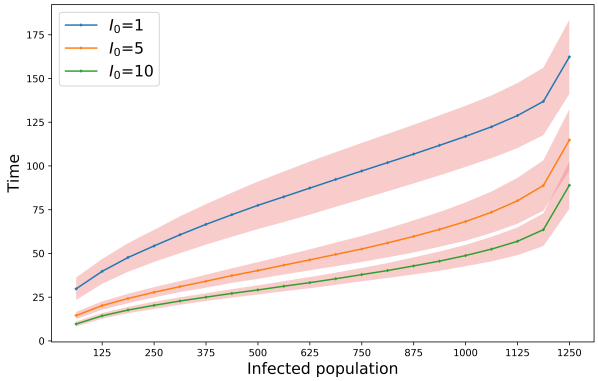


Figure 3.2: 50 superposed runs of time series corresponding to I_0=1 (solid curves) and I_0=10 (dashed curves) with the mean curves as solid lines, coloured Susceptible and Infected



(a) Time course of infected populations taken to reach steady state for I_0 = 1, 5, and 10



(b) Spread (one sample standard deviation) in I across 50 runs with time

Figure 3.3: (a) Average number of infecteds in given time, and (b) Mean time taken to reach given infected individuals. Plots show $\mu \pm \sigma_s$, where the solid coloured lines are the mean values for each I_0

50 simulations are performed each for I_0 = 1,5, and 10. Figure 3.2 shows 50 simulations each for two initial numbers (I_0= 1 and 10) of infected agents with different random initial distributions of S_0 and I_0 in the grid, choosing the minimum and maximum values of I_0 to highlight differences. The increase in infecteds and concomitant decline in susceptibles show that asymptotically a single fixed point state of complete infection (I_∞, S_∞)=(N,0) results. Even though the final steady state is the same for both I_0, it is clear that the steady state is reached faster when I_0 is larger. Also, the time course shows more spread across different runs for lower

I_0. To study these features we computed the time taken for half the total population to get infected ($\tau_{1/2}$) and the error bars/variance (shown in 3.3 (a) and (b)).

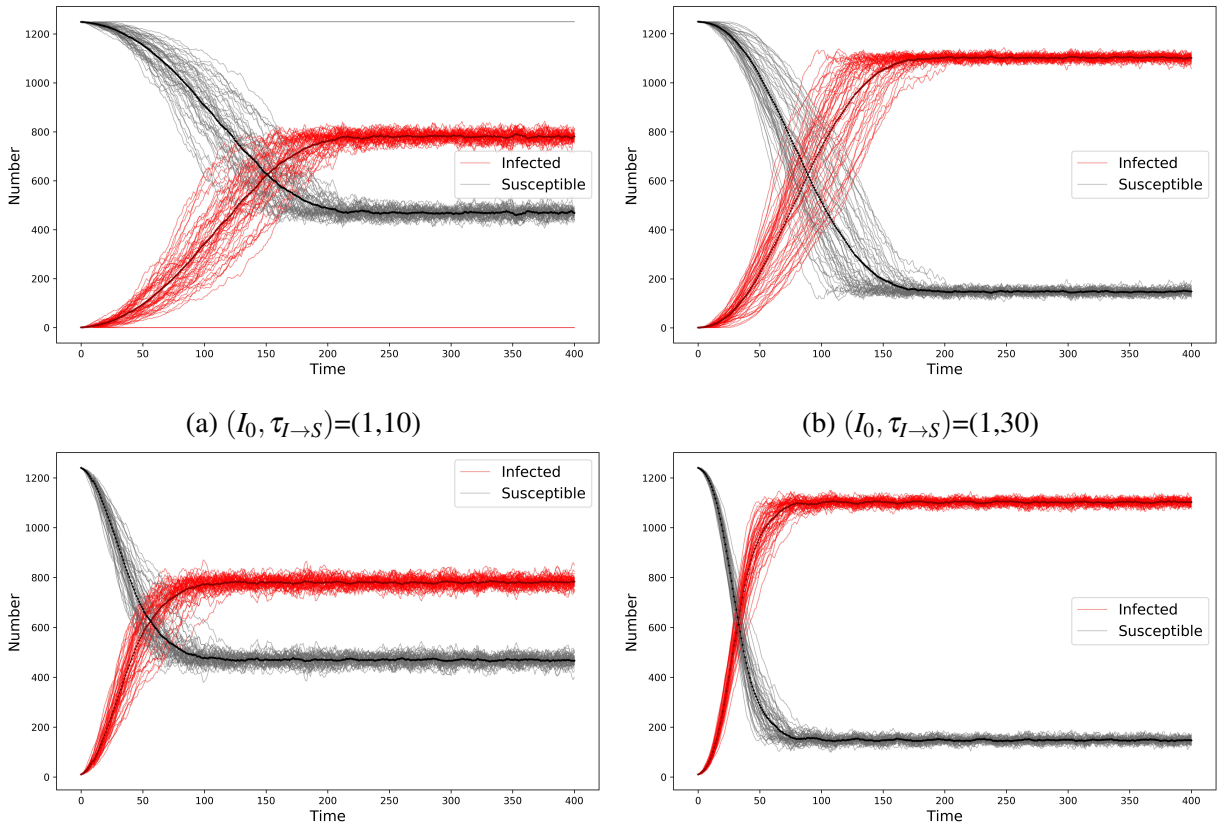
We observe that the time taken for half the population to get infected ($\tau_{1/2}$) is lesser and variance across runs is smaller in the case of higher I_0. Increasing I_0 thus leads to faster spread of infection and lesser heterogeneity in the initial spatial distribution of infected, i.e. lesser variation amongst runs.

In reality this translates to the situations where, irrespective of the distribution of the infecteds in a population at different regions, larger number of initial infected individuals would lead to faster development of disease cases.

3.2 Two agent (SIS) model

Figure 3.4: Schematic of the SIS ABM

This model leads susceptible individuals to infection on contact as before but then has a timescale of eliminating infection, i.e. a timescale from infected to susceptible state, $\tau_{I\to S}$, without any immunity conferred.



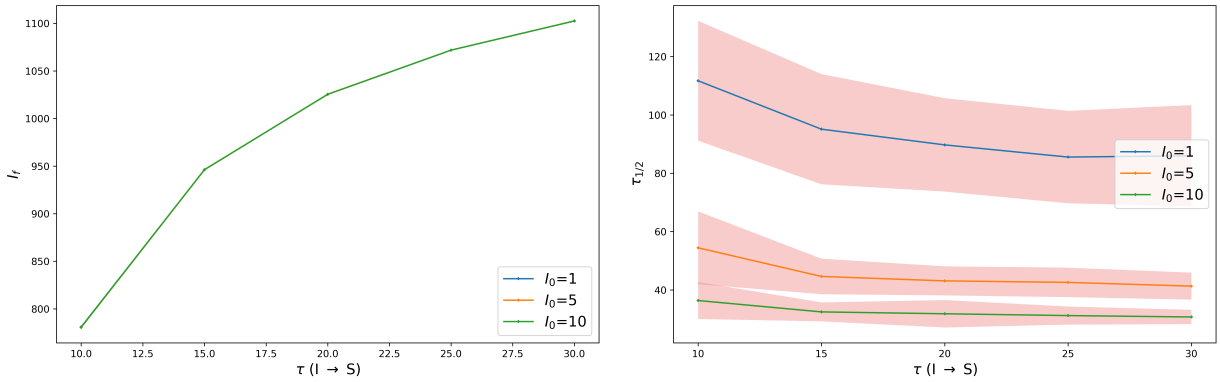
(a) $(I_0, \tau_{I\to S})$=(1,10)

(b) $(I_0, \tau_{I\to S})$=(1,30)

(c) $(I_0, \tau_{I\to S})$=(10,10)

(d) $(I_0, \tau_{I\to S})$=(10,30)

Figure 3.5: 50 superposed runs of time series corresponding to different $(I_0, \tau_{I\to S})$ pairs with the mean curves as solid lines, coloured Susceptible and Infected



(a) I_f vs $\tau_{I\to S}$

(b) $\tau_{1/2}$ vs $\tau_{I\to S}$

Figure 3.6: Mean steady infected population (a) I_f vs $\tau_{I\to S}$ and (b) $\tau_{1/2}$ vs $\tau_{I\to S}$ curves for different I_0

Fixed points for the system can now be both endemic ($0 < I_f \leq$N) and disease free (I_f=0) states with the latter seen at low $\tau_{I\to S}$ when all infecteds transition to susceptibles without

infecting much numbers of other individuals. This incidence of extinction where $I_f = 0$ is thus seen only for I_0=1, seen as flat lines in Figure 3.5 (a), but not I_0=5 or I_0=10. 3 and 2 extinctions were seen for $\tau_{I\to S}$=10 and 15 respectively while no extinctions were found beyond.

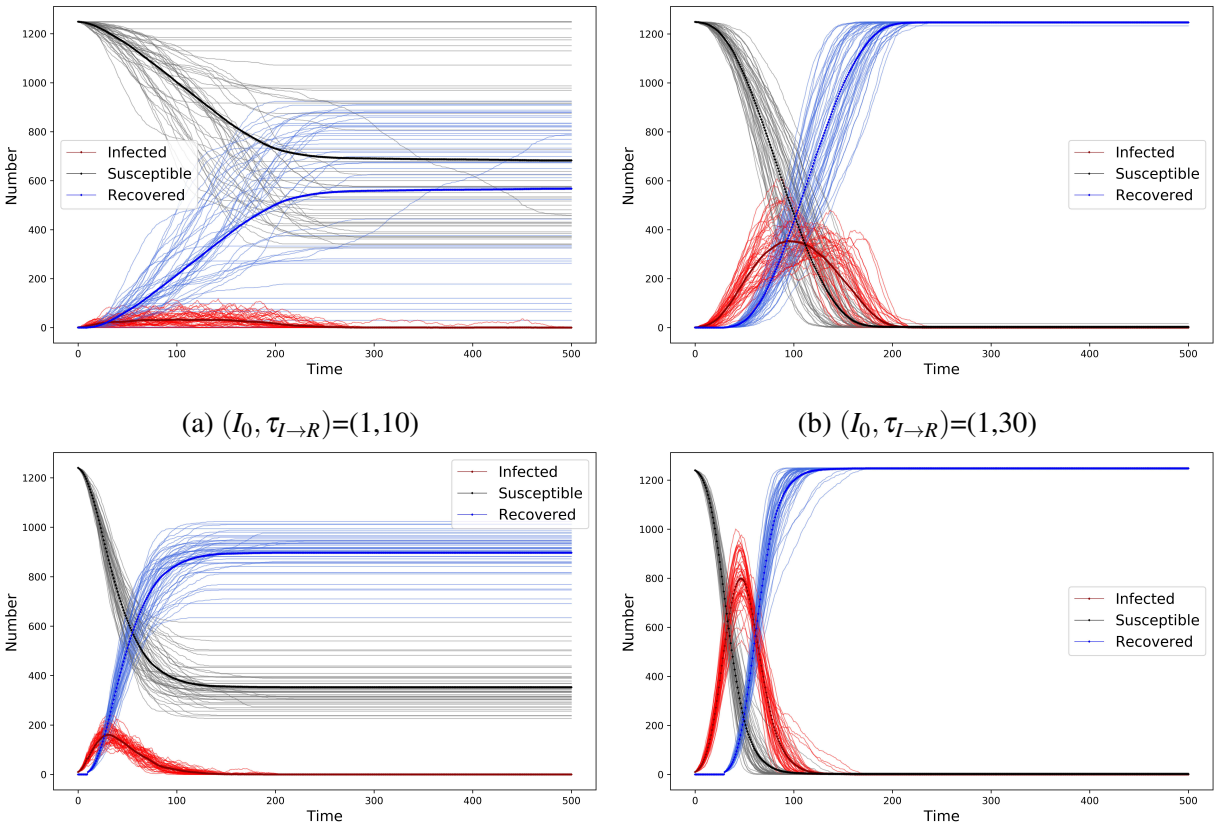
50 simulations each were performed for the 15 pairs (I_0,$\tau_{I\to S}$) where I_0 can take the values 1,5, and 10 while $\tau_{I\to S}$ =10,15,20,25, and 30. Figure 3.5 shows simulations for two different initial infecteds (I_0=1 and 10) and $\tau_{I\to S}$ ($\tau_{I\to S}$=10 and 30) where we again choose parameter minimum and maximum values to highlight differences; we note that the steady infected population increases with $\tau_{I\to S}$ but does not correlate with I_0. This is shown by Figure 3.6a where we plot the steady state infected value I_f, calculated by averaging the mean infected time series (solid red curves in Figure 3.5),<I(t)>, for the last 500 time steps of the simulation.
We again find that steady states are reached faster with both increasing I_0 and $\tau_{I\to S}$ such that $\tau_{1/2}$ falls with increasing I_0 and $\tau_{I\to S}$. Also, σ_s, the sample standard deviation given by the shaded red region in 3.6b, again decreases with increasing I_0 and slightly reduces with higher $\tau_{I\to S}$.

The interpretation of the above results is that in the SIS model, in the case the disease does not get eliminated (likely when I_0 is higher) and achieves endemic status or an infectious equilibrium, the expected number of infected in the population is dependent on the time an individual takes to recover ($\tau_{I\to S}$, infectious period) but not how many diseased individuals were present in the population when epidemic began. Here too, the disease develops faster with higher initial infecteds and $\tau_{I\to S}$.

3.3 Three agent (SIR) model

Figure 3.7: Schematic of the SIR ABM

The SIR model has three classes of agents, Susceptible, Infected, and Recovered. In this model, susceptible-infected interaction leads to infection but infected individuals recover with permanent immunity, the timescale of acquiring immunity being $\tau_{I\to R}$.



(a) $(I_0, \tau_{I\to R})$=(1,10)

(b) $(I_0, \tau_{I\to R})$=(1,30)

(c) $(I_0, \tau_{I\to R})$=(10,10)

(d) $(I_0, \tau_{I\to R})$=(10,30)

Figure 3.8: 50 superposed runs of time series corresponding to different $(I_0, \tau_{I\to R})$ pairs with the mean curves as solid lines, coloured Susceptible, Infected and Recovered

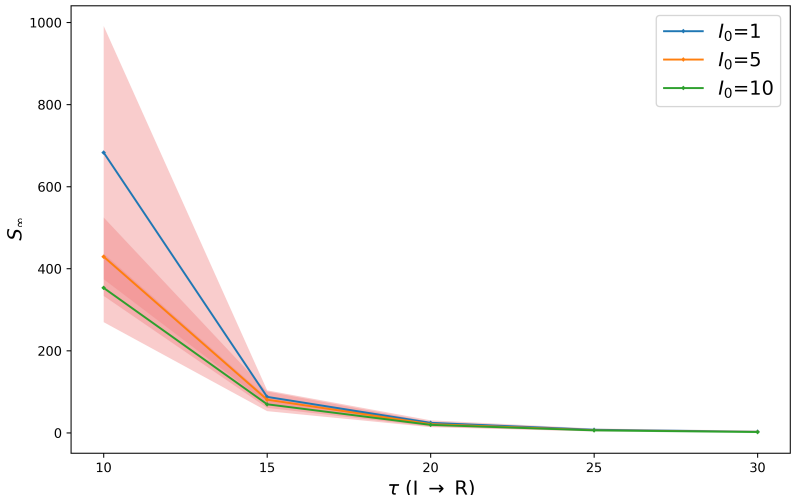
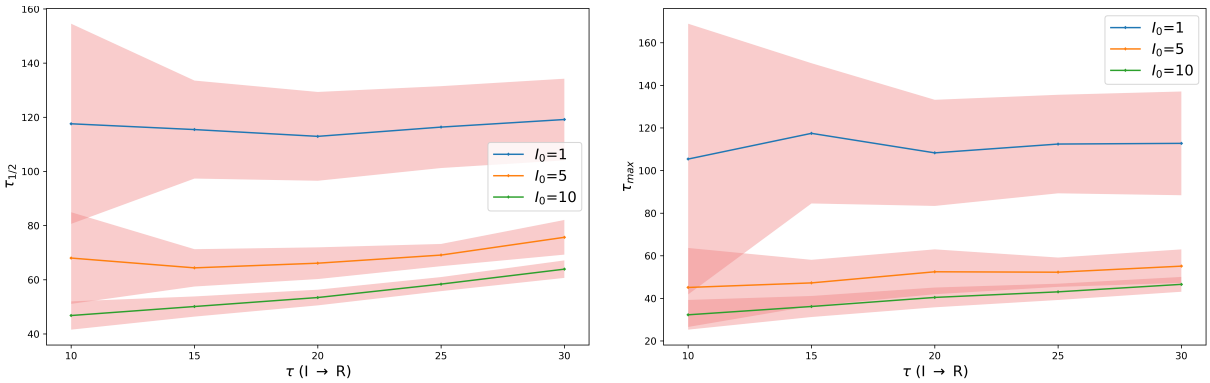


Figure 3.9: Mean S_∞ vs $\tau_{I\to R}$ for different I_0 values



(a) (b)

Figure 3.10: Mean (a) $\tau_{1/2}$ vs $\tau_{I\to R}$ and (b) $\tau_{\max}$ vs $\tau_{I\to R}$ for different values of I_0

The only fixed point for the system is the disease free state (I_∞=0) where uninfected susceptibles and recovereds coexist, but some runs exhibit an infection maxima (epidemic) before approaching the fixed point while others do not.

50 simulations each are performed for the 15 pairs (I_0,$\tau_{I\to R}$) where I_0=1,5, and 10 while $\tau_{I\to R}$ =10,15,20,25, and 30. Figure 3.8 shows simulations for two different initial infecteds (I_0=1 and 10) and $\tau_{I\to R}$ ($\tau_{I\to R}$=10 and 30). In the SIR model, the steady number of recovered increases with both I_0 and $\tau_{I\to R}$, and since $S_\infty = \mathrm{N} - R_\infty$ (N=1250), S_∞ correspondingly falls to 0 at large $\tau_{I\to R}$ (Figure 3.9), and the variation of S_∞ with I_0 also goes away at large $\tau_{I\to R}$; the infected maxima similarly increases with both I_0 and $\tau_{I\to R}$.

This model exhibits a potential bottleneck of spreading infection since immune individuals prevent contact of susceptibles with infecteds, but since longer $\tau_{I\to R}$ also leads to a delay of infection maxima from longer sustained spread of the disease, the resultant trend of $\tau_{1/2}$ with $\tau_{I\to R}$ (Figure 3.10a) is an increase but at higher $\tau_{I\to R}$; similar effects are seen for the time at which infected maxima, $\tau_{\max}$ (Figure 3.10b), occurs. Variance decreases with higher I_0 and $\tau_{I\to R}$ as in earlier ABMs. Extinctions (no epidemic) are found by calculating whether the number of infected reaches 0 within t=$\tau_{I\to R}$. Extinctions are again found for I_0=1 only, with 2 and 1 extinctions for $\tau_{I\to R}$=10 and 15 respectively, and none thereafter.

The results signify that for diseases following the SIR pattern, the number of people infected before an epidemic ends (i.e. R_∞) is highly dependent on the infectious period $\tau_{I\to R}$ but also I_0 if $\tau_{I\to R}$ is small, and that predicting critical times like $\tau_{\max}$ and $\tau_{1/2}$, representing peak periods of disease, can prove difficult, especially at lower I_0 where variance/errors are high, so

one must account for this in planning intervention strategies.

3.4 Three agent (SIRS) model

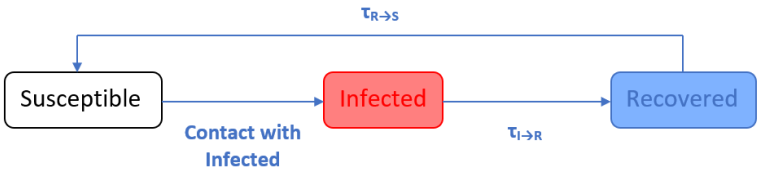
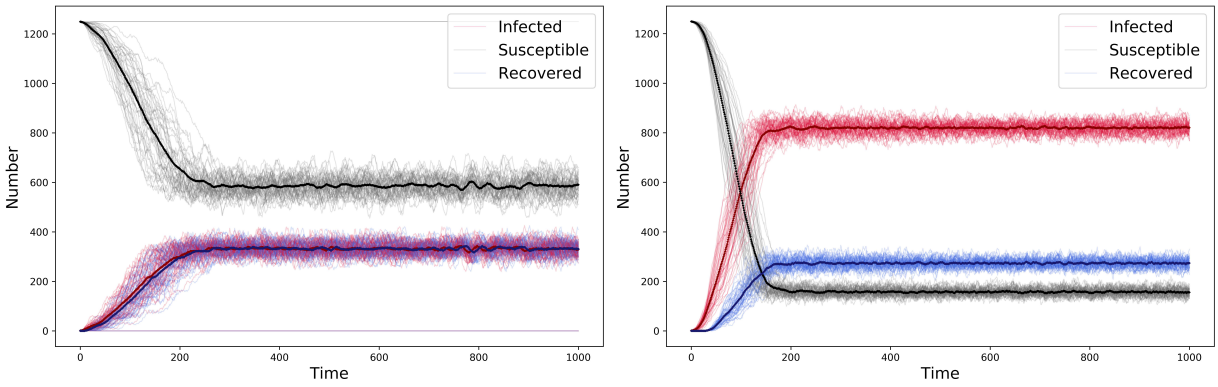


Figure 3.11: Schematic of the SIRS ABM

The SIRS model again causes susceptible individuals to develop infection but has both a timescale of acquiring immunity, $\tau_{I\to R}$, and subsequently losing it at the timescale $\tau_{R\to S}$.

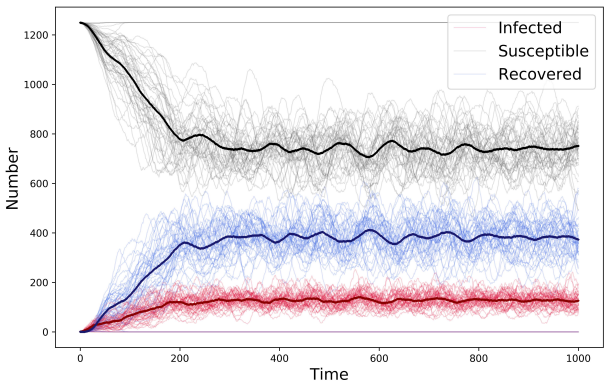
The system possesses fixed points and oscillatory solutions when distinguishing them by the strength of oscillations using the Discrete Fourier Transform (DFT), the fixed points either disease free (I_f=0) or endemic ($0 < I_f \leq$N) and the oscillatory solutions endemic in nature.

50 simulations each are performed for the 75 sets (I_0,$\tau_{I\to S}$,$\tau_{I\to R}$) where I_0 =1,5, and 10 while $\tau_{I\to S}$ or $\tau_{I\to R}$ can both be 10,15,20,25, and 30 each. Figure 3.12 shows simulations for two different initial infecteds (I_0=1 and 10) and ($\tau_{I\to R}$, $\tau_{R\to S}$=10 and 30 for both), coloured Susceptible, Infected and Recovered. The fixed point/oscillation range values are seen to be dependent on $\tau_{I\to R}$ and $\tau_{R\to S}$ but not I_0, with longer $\tau_{I\to R}$ leading to higher steady value of infecteds, and $\tau_{R\to S}$ to higher recovereds and slightly higher susceptibles due to lesser infections from immunity bottleneck (Section 3.3). This system is sensitive to both timescales, and a change in either timescale can lead to a different range of oscillation/steady state for all compartments.

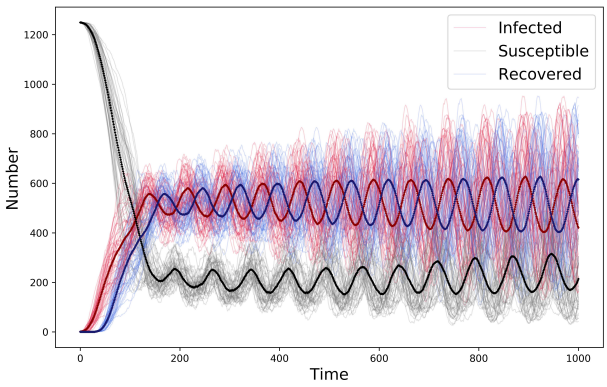


(a) (I_0 , $\tau_{R\to S}$, $\tau_{I\to R}$)=(1,10,10)

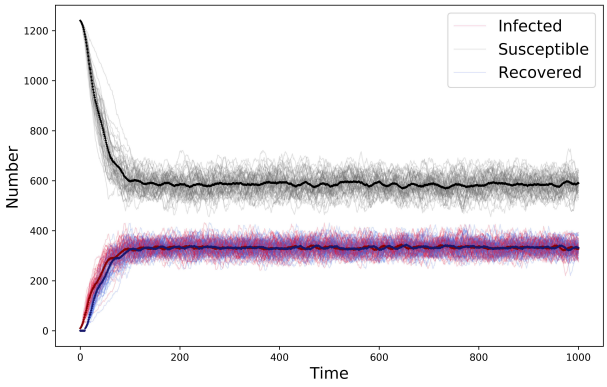
(b) (I_0 , $\tau_{R\to S}$, $\tau_{I\to R}$)=(1,10,30)



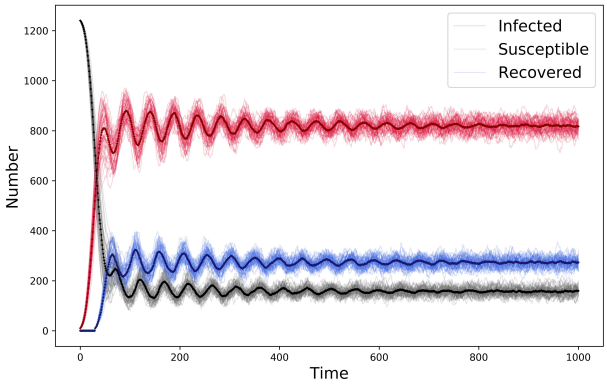
(c) (I_0 , $\tau_{R\to S}$, $\tau_{I\to R}$)=(1,30,10)



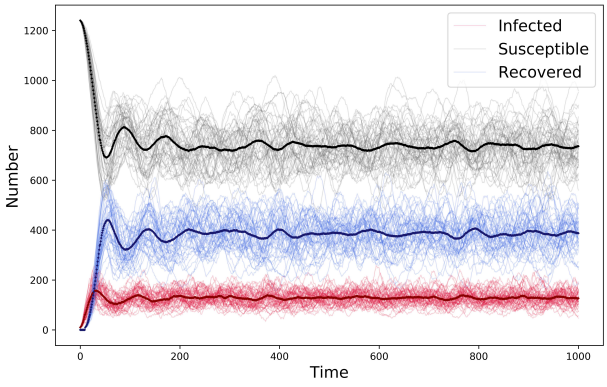
(d) (I_0 , $\tau_{R\to S}$, $\tau_{I\to R}$)=(1,30,30)



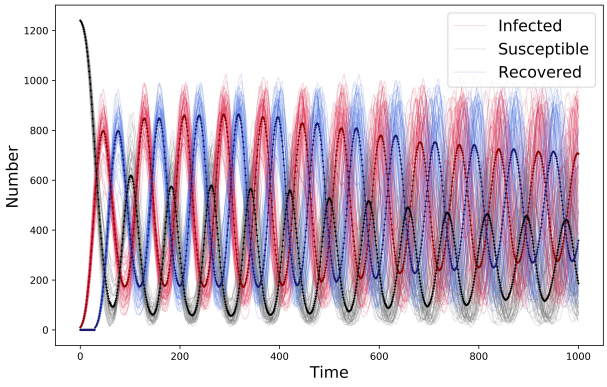
(e) (I_0 , $\tau_{R\to S}$, $\tau_{I\to R}$)=(10,10,10)



(f) (I_0 , $\tau_{R\to S}$, $\tau_{I\to R}$)=(10,10,30)

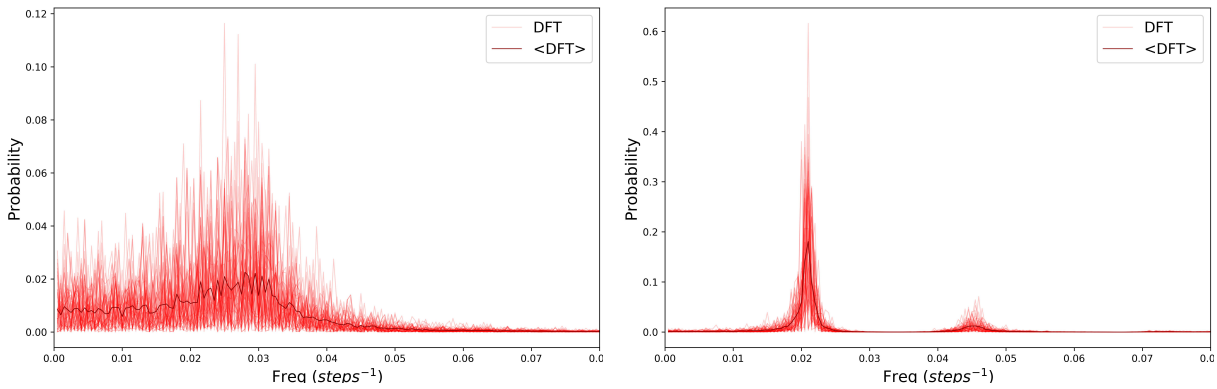


(g) (I_0 , $\tau_{R\to S}$, $\tau_{I\to R}$)=(10,30,10)



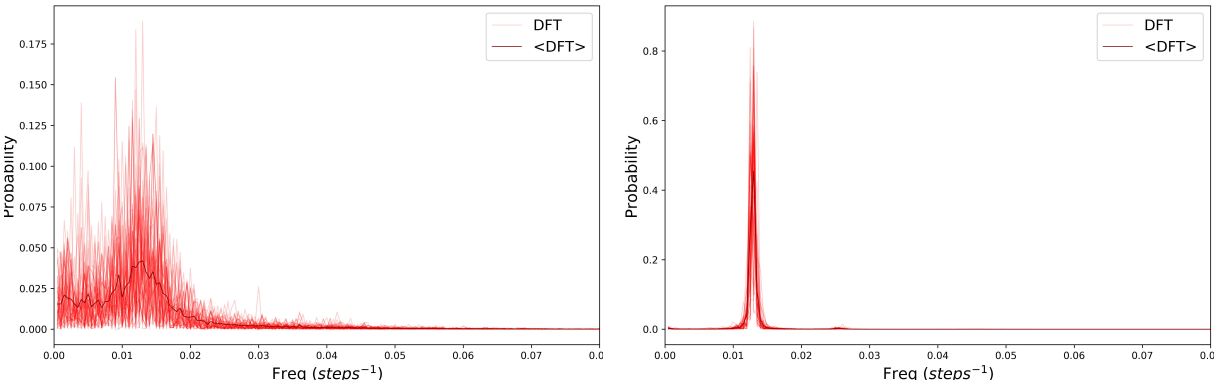
(h) (I_0 , $\tau_{R\to S}$, $\tau_{I\to R}$)=(10,30,30)

Figure 3.12: 50 superposed runs of time series corresponding to different (I_0 , $\tau_{R\to S}$, $\tau_{I\to R}$) combinations with the mean curves as solid lines



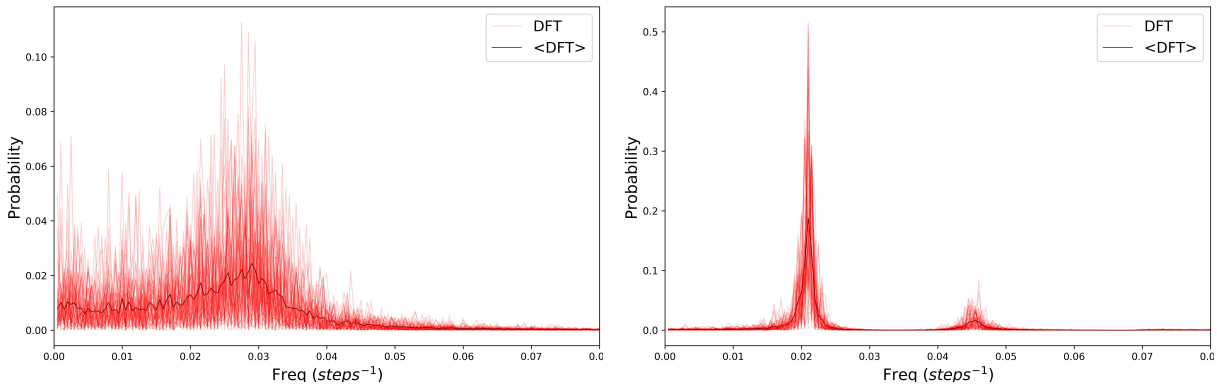
(a) (I_0 , $\tau_{R\to S}$, $\tau_{I\to R}$)=(1,10,10)

(b) (I_0 , $\tau_{R\to S}$, $\tau_{I\to R}$)=(1,10,30)



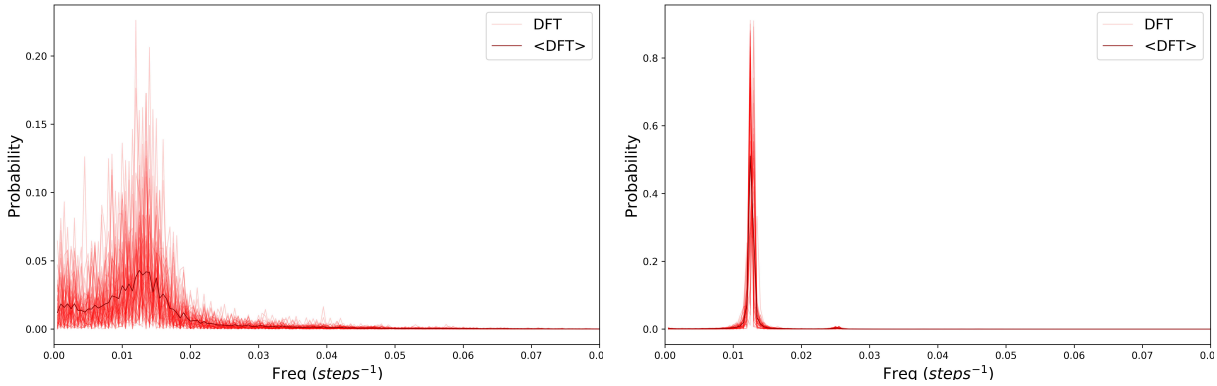
(c) (I_0 , $\tau_{R\to S}$, $\tau_{I\to R}$)=(1,30,10)

(d) (I_0 , $\tau_{R\to S}$, $\tau_{I\to R}$)=(1,30,30)



(e) (I_0 , $\tau_{R\to S}$, $\tau_{I\to R}$)=(10,10,10)

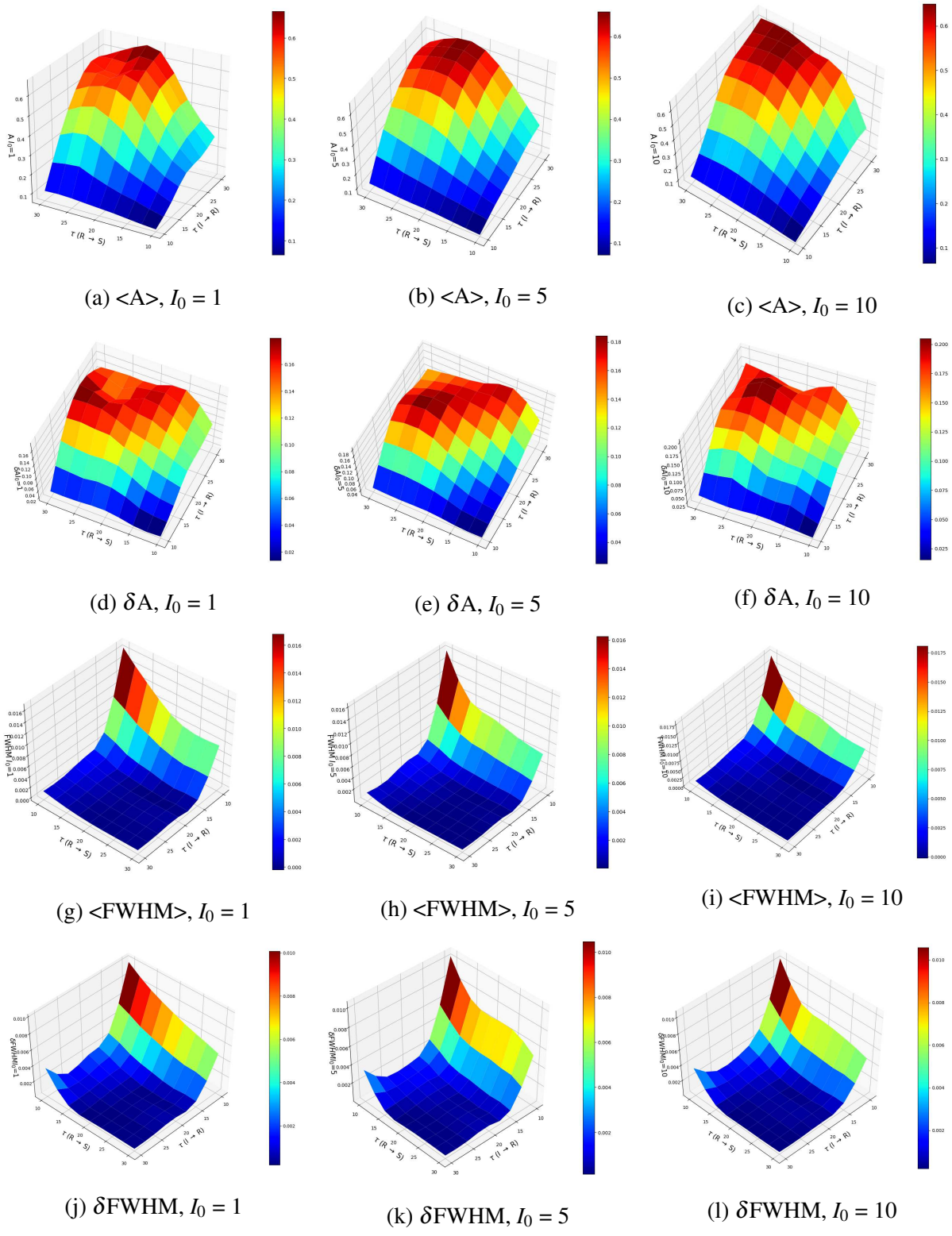
(f) (I_0 , $\tau_{R\to S}$, $\tau_{I\to R}$)=(10,10,30)



(g) (I_0 , $\tau_{R\to S}$, $\tau_{I\to R}$)=(10,30,10)

(h) (I_0 , $\tau_{R\to S}$, $\tau_{I\to R}$)=(10,30,30)

Figure 3.13: Superposed DFTs of the infected time series corresponding to the 50 runs of time series shown previously. Mean curves (<DFT>), taken point wise, are solid lines



(a) <A>, I_0 = 1 (b) <A>, I_0 = 5 (c) <A>, I_0 = 10

(d) δA, I_0 = 1 (e) δA, I_0 = 5 (f) δA, I_0 = 10

(g) <FWHM>, I_0 = 1 (h) <FWHM>, I_0 = 5 (i) <FWHM>, I_0 = 10

(j) δFWHM, I_0 = 1 (k) δFWHM, I_0 = 5 (l) δFWHM, I_0 = 10

Figure 3.14: The mean DFT peak amplitude surface with interpolation (a-c) and errors (d-f), and interpolated mean Full Width at Half Maxima (FWHM) surface (g-i) with errors (j-l)

	$\tau_{R\to S}=10$	$\tau_{R\to S}=15$	$\tau_{R\to S}=20$	$\tau_{R\to S}=25$	$\tau_{R\to S}=30$
$\tau_{I\to R}=10$	3	4	6	5	6
$\tau_{I\to R}=15$	2	0	1	2	1
$\tau_{I\to R}=20$	0	1	0	0	0

Table 3.1: Extinction table for the SIRS model, I_0=1

We also perform Discrete Fourier Transforms (DFTs) to quantify the strength of oscillations, and looking at Figure 3.13 conclude that oscillatory behaviour is present when $\tau_{I\to R}$ and $\tau_{R\to S}$ are higher, while we can consider the final state to be endemic non-oscillatory when either is lower. This behaviour is further quantified by calculating the peak amplitudes (A) and Full Width at Half Maxima (FWHM) in the DFTs in Figure 3.14 averaged over the 50 runs, with cubic interpolation performed in between simulation points for smoother surfaces. Higher timescale values lead to higher peak amplitude, and also a smaller Full Width at Half Maxima (FWHM), the width of the peak measured at half of the peak amplitude, showing strong, regular oscillations at higher timescales.

Variance of the runs largely follows the same pattern as that of A and FWHM (Figure 3.14 d-f and j-l). Extinction is only observed for I_0=1, where correlation with $\tau_{I\to R}$ is strong and negative while that on $\tau_{R\to S}$ is weaker and positive, (Table 3.1) showing the instance of persistence of infection spread with higher $\tau_{I\to R}$ and suppression at higher $\tau_{R\to S}$.

The SIRS model assumes that people become susceptible to the disease after having transient immunity, but since we know that immune individuals prevent the spread of disease, longer transient immunity ($\tau_{R\to S}$) can prevent epidemics. Also, the infectious period ($\tau_{I\to R}$) influences the final proportion of people infected should the disease not die out and become endemic. Just like the SIS ABM, the final states are independent of the initial infecteds present in the population.

Chapter 4

Agent-Based Model for Malaria

The implementation of ABM for a particular disease requires the biological knowledge of the host-pathogen interaction specific for that infection. Here we show the implementation and results of the simplest agent-based model for malaria, which is similar to the Ross Model, which is a SIS type differential equation model (Section 2.2.2) for humans and SI type for mosquitoes (Section 2.2.1). Since the malaria parasite needs two hosts - human and mosquito - to complete its life cycle and disease to spread, the ABM representation also has two sets of agents for human and mosquitoes, each having their own susceptible and infected states. As in Chaper 3, rate constants are replaced by timescales of transition of the different compartments of both humans and mosquitoes, while the other parameters are stochastically implemented. We first develop the Ross ABM model, and study the role of different parameters for malaria spread in the population of agents and compare the results with the mean field approach of the original differential equation model by Sir Ronald Ross. [Ross 10]

4.1 Epidemiological model

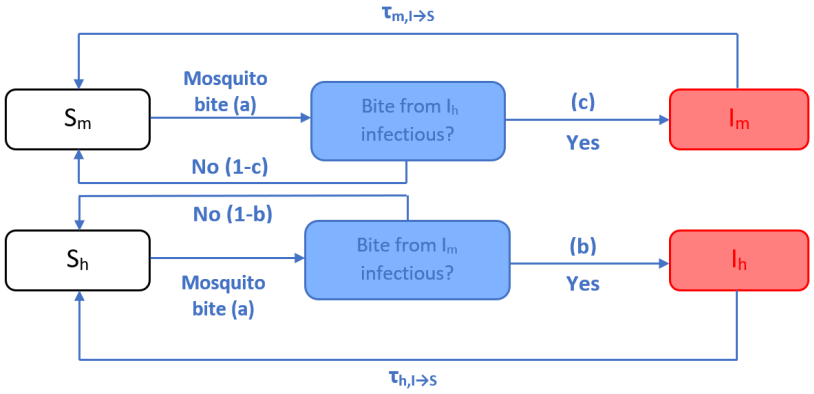


Figure 4.1: Schematic of the general Ross ABM with non-trivial a,b and c, where brackets show the probability of transition of states in terms of the parameters

As in the differential equation model, the Ross ABM also consists of two classes of agents for both humans and mosquitoes–

1. Susceptible human agents (S_h), and infected human agents (I_h), having acquired the infection but not liable to reinfection/ superinfection. The I_h agents can recover from malaria, and become S_h again without having any immunity to the disease.

2. Susceptible mosquito agents (S_m), and infected mosquito agents (I_m), having acquired the infection but not liable to reinfection/ superinfection. I_m agents die on infection. To keep the total population constant, every dead I_m is replaced with a susceptible mosquito agent S_m in the same patch.

The interactions between the human and mosquito agents are defined by timescales that are similar to the rate constants of interactions in the differential equation Ross model. These are the timescale of the transition of infected humans to susceptible ones, $\tau_{h,I\to S}$, and $\tau_{m,I\to S}$ is the infected mosquitoes replaced with susceptibles (after death) - to keep the mosquito population constant.

There are other rates of interactions that are implemented as non- timescale parameters. In the differential equation model both susceptible humans and mosquitoes are infected by biting

of I_m to S_h and when I_h is bitten by S_m (Figure 4.1). These bites by mosquitoes can lead to probabilistic disease transmission based on infectivity and biting rate parameters.

4.2 Parameters

4.2.1 Population parameters

We initialise the ABM simulations with the total number of agents as 1250 (~30% of total space sites). The length of the simulation was for 3200 time steps, but zoomed time series are presented to study features of the ABM. The total population sizes of the two classes of agents are - Humans: $N_h = S_h + I_h = 625$ and Mosquitoes: $N_m = S_m + I_m = 625$. Here we consider that a few infected humans, but no infected mosquitoes are present at the start of simulations, i.e., $S_{m,0} = 625$ and $I_{m,0} = 0$. The disease starts to spread in the populations of agents when the susceptible mosquito takes a blood-meal from an infected human and becomes infected. Then onward whenever I_h and I_m interact with S_m and S_h in the grid respectively, the disease spreads. This study used infected human (I_h) agent population sizes at t=0 as $I_{h,0}$=1, 5 or 10, referred to as simply I_0 here on.

4.2.2 Timescale parameters

To remain close to known parameter values on interaction rates, the human and mosquito infected-to-susceptible transition timescales are taken as $\tau_{h,I\to S} \in$ [90,95,100,105,110] and $\tau_{m,I\to S} \in$ [7,8,9,10,11] respectively. The inverse of the central values, $1/\tau_{h,I\to S} = 0.01$ and $1/\tau_{m,I\to S}$ ~0.11 are kept to be of the same order as the observed values of 0.005-0.05 and 0.05-0.5 respectively [Mandal 11]. Also, the female mosquito to human population ratio, $m=N_m/N_h=1$ in this ABM, whereas the observed range is 0.5-40 ,where it may be noted that all mosquito agents are taken to be potentially infectious/female.

4.2.3 Other parameters

The differential equation Ross model (Section 2.3) also has the following non-timescale parameters –

- The man biting rate, a, which is the proportion of mosquitoes that feed on humans in a day [Smith 12]
- The infectious fraction in humans, b, the fraction of infectious bites by mosquitoes leading to contraction of infection in humans.
- The infectious fraction in mosquitoes, c, the fraction of bites by susceptible mosquitoes from an infected human leading to them getting infected.

The implementation of these parameters in the general ABM is probabilistic, where a random number in [0,1] is generated and if found greater than the parameter value, the condition (bite/infection) is switched on; for instance if a=0.7 and the random number generated when a mosquito and human are in the same patch is 0.6, the mosquito will not bite. We take a=b=c=1 in the simulations for adequate strength of infection spread and simplicity.

4.3 Movement and infection rules

The Ross model ABM, like the compartmental ABMs in Chapter 3, has interactions in a patch which are 1-to-1 for humans and mosquitoes, and no more than 2 agents in total of any type are allowed at the same patch/site. The following description is for general values of a, b and c in the ABM set-up:

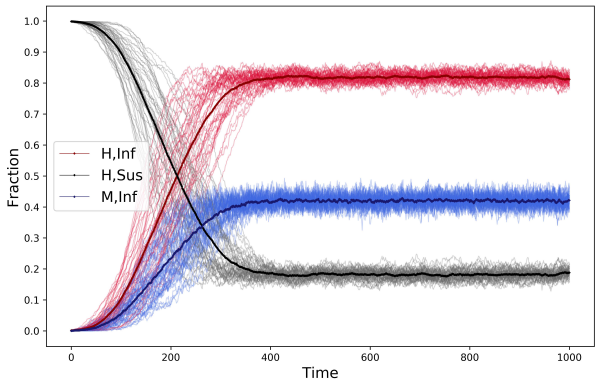
- Infection spreads when an infected mosquito is present in a patch with a susceptible human and bites it, probabilistically triggering infection, or when a susceptible mosquito similarly bites an infected human
- An agent is allowed to move into its Moore neighbourhood (i.e., nearest 8 neighbours) subject to the constraints below-

- Susceptible or Infected humans can only move into patches unoccupied by any other susceptible or infected humans
- Susceptible or Infected mosquitoes can similarly only move into patches with no other susceptible or infected mosquitoes

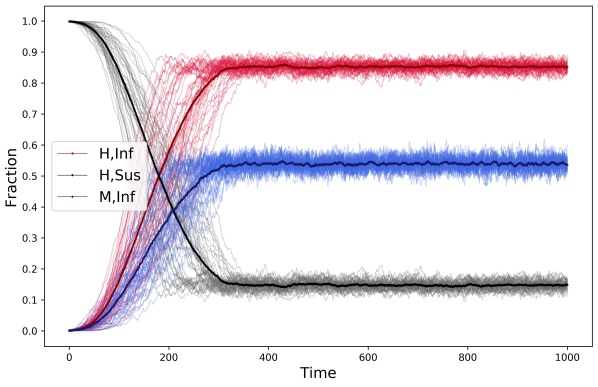
4.4 Results

4.4.1 Parameter set results

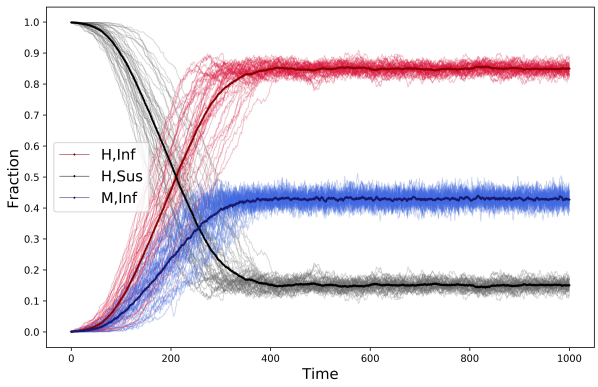
When using the parameters in Section 4.2, the asymptotic dynamics of the Ross ABM is a fixed point similar to the differential equation Ross model. Below we discuss the results of the effect of variations in time scales and I_0 on the collective dynamics of the agents in Ross ABM. The sets of time series in the following section are coloured as susceptible humans (H,Sus), infected humans (H,Inf), and infected mosquitoes (M,Inf).



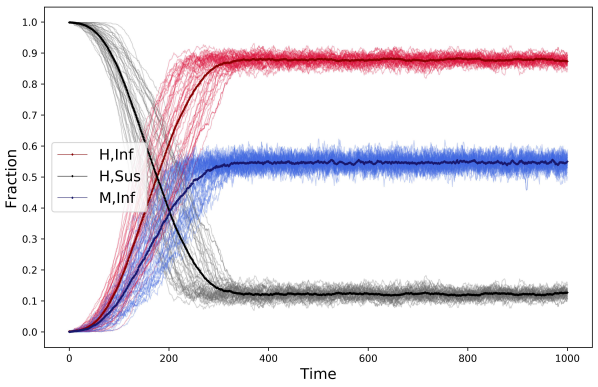
(a) $(I_0, \tau_{h,I\to S}, \tau_{m,I\to S})$ =(1,90,7)



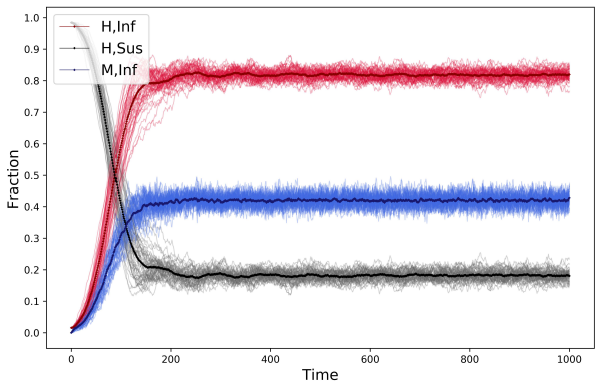
(b) $(I_0, \tau_{h,I\to S}, \tau_{m,I\to S})$ =(1,90,11)



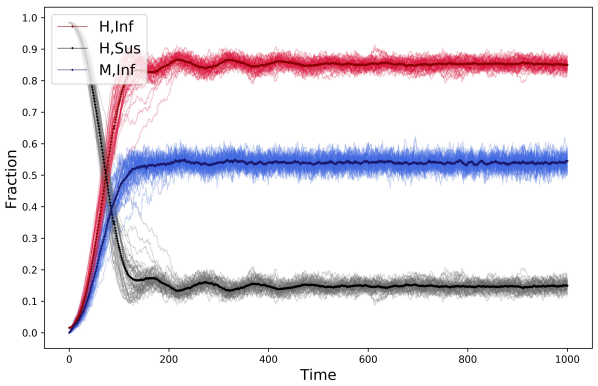
(c) $(I_0, \tau_{h,I\to S}, \tau_{m,I\to S})$ =(1,110,7)



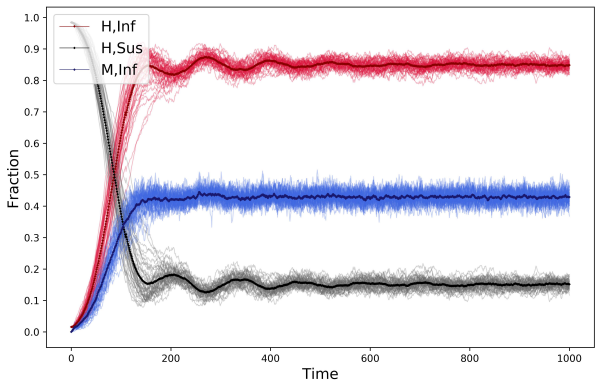
(d) $(I_0, \tau_{h,I\to S}, \tau_{m,I\to S})$ =(1,110,11)



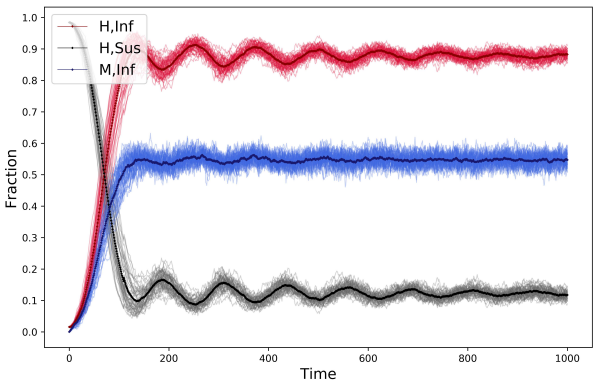
(e) $(I_0, \tau_{h,I\to S}, \tau_{m,I\to S})$ =(10,90,7)



(f) $(I_0, \tau_{h,I\to S}, \tau_{m,I\to S})$ =(10,90,11)



(g) $(I_0, \tau_{h,I\to S}, \tau_{m,I\to S})$ =(10,110,7)



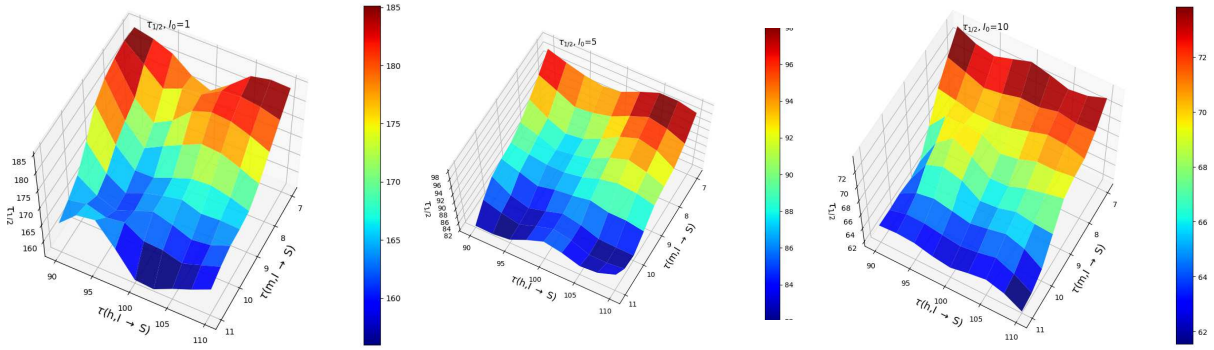
(h) $(I_0, \tau_{h,I\to S}, \tau_{m,I\to S})$ =(10,110,11)

Figure 4.2: 50 superposed runs of time series corresponding to different $(I_0, \tau_{h,I\to S}, \tau_{m,I\to S})$ combinations with the mean curves as solid lines.

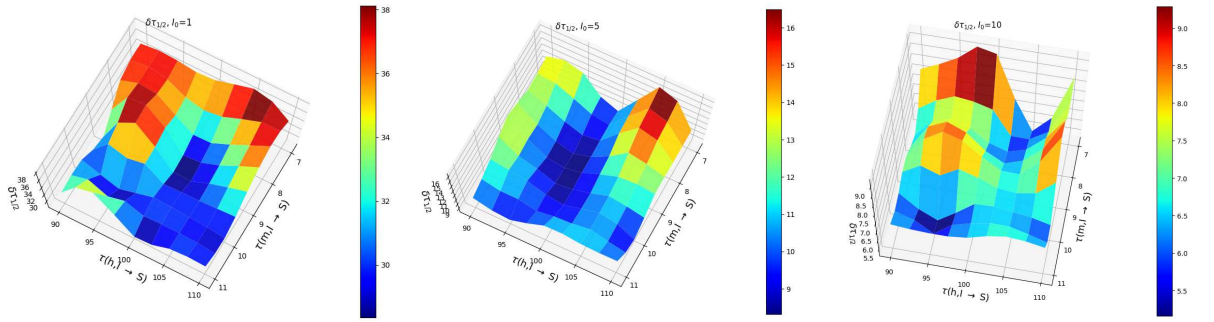
The section 4.2 parameters only lead to fixed points which are again disease free ($I_{h,f}$=0 and $I_{m,f}$=0) or endemic ($0 < I_{h,f} \leq$ N and $0 < I_{m,f} \leq$ N). Extinctions, where $I_{h,f} = I_{m,f} = 0$ are very sporadic with one case each for ($\tau_{h,I\to S}, \tau_{m,I\to S}$) = (90,8) and (100,10) at I_0=1.

From a large range of parameter combinations studied, in Figure 4.2, representative time series are presented with parameter values I_0=1 and 10, $\tau_{m,I\to S}$=7 and 11, and $\tau_{h,I\to S}$ =90 and 110 to highlight differences from parameter variation. The susceptible mosquito population time series is omitted for the sake of plot clarity, but is given as N_m - I_m = 625 - I_m. The large time mosquito population, $I_{m,f}$ is largely dependent on the mosquito death timescale $\tau_{m,I\to S}$ but has a slight dependence on $\tau_{h,I\to S}$ seen by the incline along $\tau_{h,I\to S}$ axis in Figure 4.3 (g-i); note again the averaging over 50 runs and interpolation to give mean values and smoother surfaces respectively. The large time human infected (I_h) and susceptible (S_h) are similarly affected largely by $\tau_{h,I\to S}$; the dependence of both on I_0 is absent. Also, we study the time taken for half the mosquito population to get infected ($\tau_{1/2}$, Figure 4.3 a-c) and see that both $\tau_{h,I\to S}$ and $\tau_{m,I\to S}$ are inversely related to $\tau_{1/2}$ In the endemic cases decaying oscillations can be seen but with longer decay time as $\tau_{h,I\to S}$, $\tau_{h,I\to S}$ or I_0 increase. As in the case of compartmental ABMs, variance in runs still falls with increasing I_0 but there is no specific pattern with $\tau_{m,I\to S}$ or $\tau_{h,I\to S}$ as seen in Figure 4.3 (d-f and j-l).

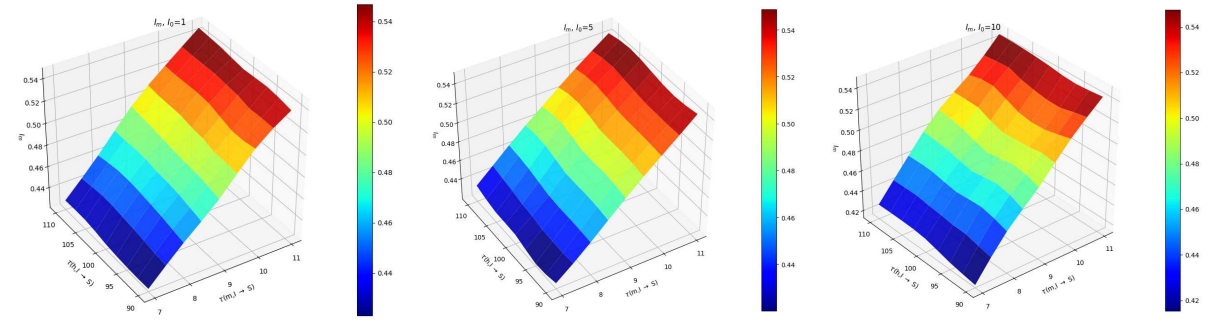
We thus see that the expected number of mosquitoes infected by malaria in endemic states depends on the infectiousness of humans and mosquitoes, but that of humans is largely dependent on their own lifespan when infected. Oscillations may be triggered by the sustained spread of the disease, which is a result of lesser heterogeneity with increasing $\tau_{m,I\to S}, \tau_{h,I\to S}$, and I_0.



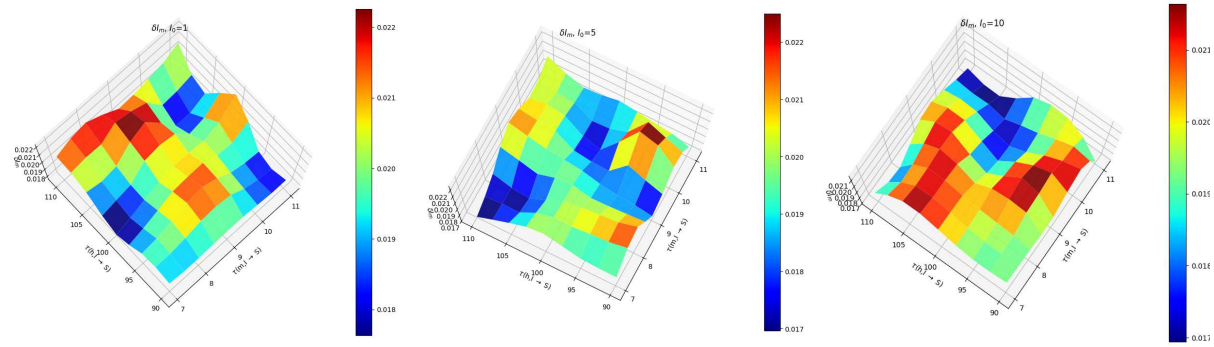
(a) $\tau_{1/2}$, $I_0 = 1$ (b) $\tau_{1/2}$, $I_0 = 5$ (c) $\tau_{1/2}$, $I_0 = 10$



(d) $\delta\tau_{1/2}$, $I_0 = 1$ (e) $\delta\tau_{1/2}$, $I_0 = 5$ (f) $\delta\tau_{1/2}$, $I_0 = 10$



(g) $<I_{m,f}>$, $I_0 = 1$ (h) $<I_{m,f}>$, $I_0 = 5$ (i) $<I_{m,f}>$, $I_0 = 10$



(j) $\delta I_{m,f}$, $I_0 = 1$ (k) $\delta I_{m,f}$, $I_0 = 5$ (l) $\delta I_{m,f}$, $I_0 = 10$

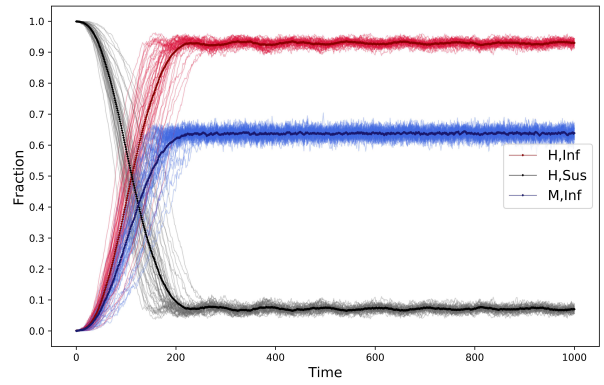
Figure 4.3: The mean $\tau_{1/2}$ surface with interpolation (a-c) and errors (d-f), and interpolated mean $I_{m,f}$ (large time infected mosquito fraction) surface (g-i) with errors (j-l)

4.4.2 Total population size

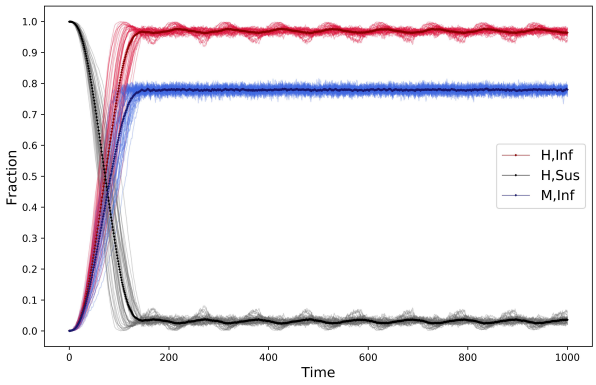
The model dynamics under different population sizes of human and mosquito agents with N_h=N_m and initial infected host agent numbers I_0 is also studied keeping $\tau_{h,I\to S}$ =100, $\tau_{m,I\to S}$=9 constant (Figure 4.4, a-d).

Compared to Figure 4.2 it is seen that damping is reduced and the dynamic behaviour transitions to oscillations at higher populations and I_0, as is also shown in the corresponding DFTs (Figure 4.4, e-h).

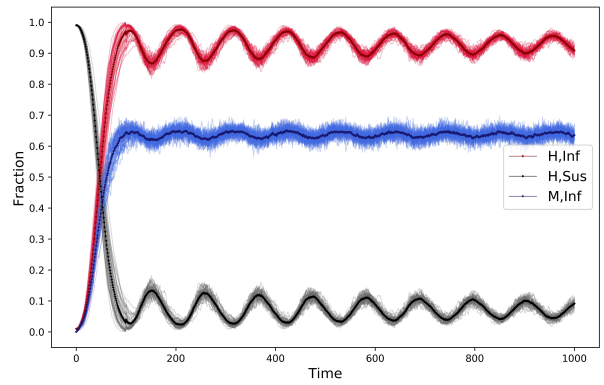
Oscillatory dynamics is not observed in the differential equation Ross model (Figure 2.14), and is one of the interesting outcomes of the ABM realization. This may be a combined effect of high population density and time scales of interactions leading to sustained regular variation in spread distributed in space, and increased I_0 causing less heterogeneity/variance as before. This is purely a spatial effect dependent on initial conditions since the ratio m= N_m / N_h is preserved, and thus all parameters are constant.



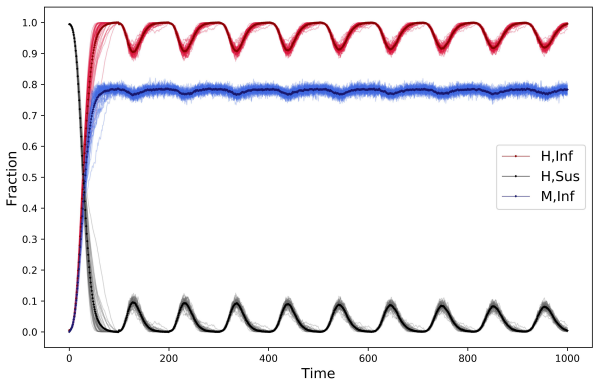
(a) (I_0 , N_h, N_m)=(1,1000,1000)



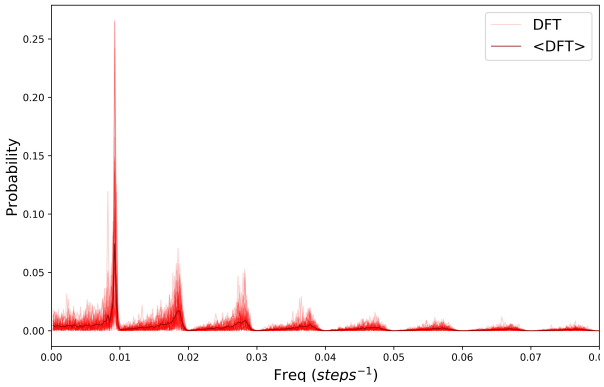
(b) (I_0 , N_h, N_m)=(1,1800,1800)



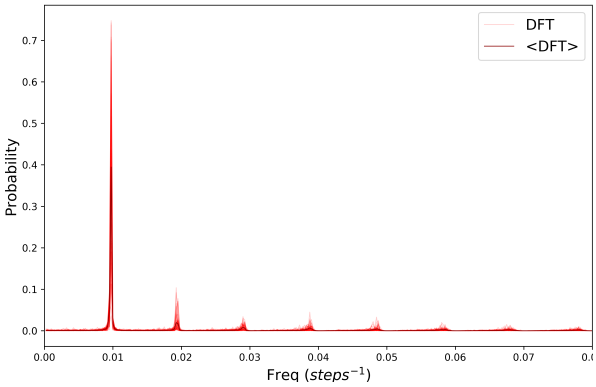
(c) (I_0 , N_h, N_m)=(10,1000,1000)



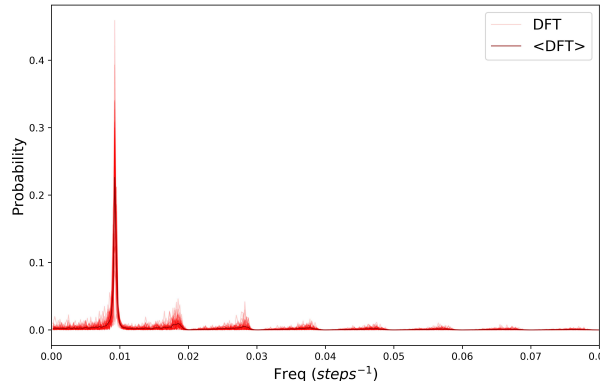
(d) (I_0 , N_h, N_m)=(10,1800,1800)



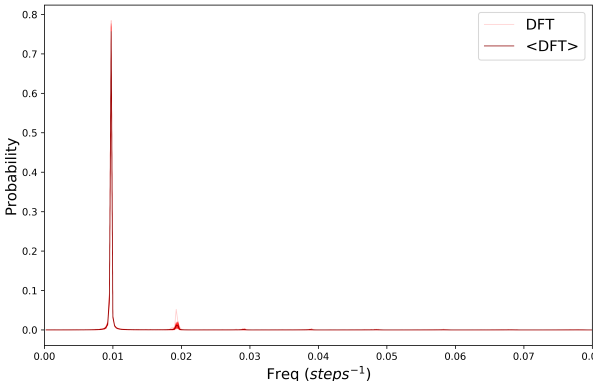
(e) DFT at (I_0 , N_h, N_m)=(1,1000,1000)



(f) DFT at (I_0 , N_h, N_m)=(1,1800,1800)



(g) DFT at (I_0 , N_h, N_m)=(10,1000,1000)



(h) DFT at (I_0 , N_h, N_m)=(10,1800,1800)

Figure 4.4: 50 superposed runs of time series (a-d) corresponding to different (I_0 ,N_h ,N_m) combinations with the mean curves as bold lines, and the DFTs of I_h time series (e-h)

Chapter 5

Summary and future work

5.1 Summary

Understanding the mechanism of the spread of an infection in a population is critical to its management, control, and eradication. Epidemiological modelling has been a commonly used approach to enhance our knowledge and predictive capability in managing infectious diseases. In this thesis, two of the theoretical modelling approaches - the mean-field differential equation models and microscopic Agent-Based Models - have been used. ABMs incorporate heterogeneity in agent properties and have agents distributed in space, which may lead to differing evolution of their time course. The work aimed to compare and contrast the results obtained for epidemiological compartment modelling from the two approaches. First, the differential equation models for two and three compartments, namely the SI, SIS, SIR, and SIRS models, are studied and their steady state and local stability analysis done to understand the long term dynamics of disease spread. All these models exhibit fixed point asymptotic dynamics, and either disease free or endemic infection of population, depending on the interaction rates and initial conditions.

To demonstrate the behaviour of the microscopic approach of agent populations in similar compartment models, the ABMs of equivalent descriptions are developed in a 45×45 spatial grid with relevant time scales for interaction rates. A large number of simulations are done spanning a range for each parameter for 50 different initial population distributions. The major results obtained for each of the compartment models are summarized below-

- SI model: We observe that the time taken for half the population to get infected is lesser and variance across runs is smaller in the case of higher initial infecteds. The rate of

infection in the case of low initial infecteds shows a higher dependence on their location since infecteds at grid edges have less average neighbours, which explains why some outliers may exist and give rise to variations in ABM runs across all model types. Higher infecteds at the beginning thus leads to faster spread of infection and lesser heterogeneity in the initial spatial distribution of infected.

In reality, this translates to the situations where irrespective of the distribution of the infecteds in a population at different regions, a larger number of initial infected individuals would lead to faster development of disease cases.

- SIS model: In the SIS model, extinction or a state where the population is completely free of infecteds can occur at low initial infecteds but if state achieved is endemic status or an infectious equilibrium, the expected number of infected in the population is dependent on the time an individual takes to recover or the infectious period but not how many diseased individuals were present in the population when the epidemic began. Here too, the disease develops faster with higher initial infecteds and infectious period.

- SIR model: The results signify that for diseases following the SIR pattern, the number of people infected before an epidemic ends is highly dependent on the infectious period but also initial infecteds if the infectious period is small, and that predicting critical times like time of infection maximum or $\tau_{1/2}$, representing peak periods of disease can prove difficult especially at low initial infecteds, so one must account for this in planning intervention strategies.

- SIRS model: Higher timescale values lead to higher peak amplitude or strength of oscillations, i.e., oscillatory behaviour, and also a smaller Full Width at Half Maxima (FWHM), the width of the peak measured at half of the peak amplitude. Strong, regular oscillations are seen at higher timescales, while lower timescales correspond to either non-oscillatory endemic or disease free states.

  Extinction is only observed for low infecteds with the dependence on timescales being such that higher infectious period strongly increases the persistence of infection while increasing recovery period/transient immunity stage suppresses epidemics though with relatively moderate intensity.

  The SIRS model allows transient immunity, after which individuals again become susceptible, but since immune individuals prevent contact between susceptibles and infecteds and thus inhibit the spread of disease, longer transient immunity can help prevent

epidemics. Also, the infectious period influences the final proportion of people infected should the disease not die out and become endemic. The SIRS and SIS models show similarities in that final endemic states are independent of the initial infecteds present in the population, both notably being models where a return to susceptibility or 'cyclicity' of transitions is permitted.

Along with these general models, a more complex model specific to the disease Malaria (the Ross model), is also studied using both the approaches. Here also both approaches showed fixed point dynamics, but the ABM realization displayed oscillatory dynamics for certain time scales and population sizes.

The above description depicts both the similarities that result from agent-based models when converted from differential equation models but also highlight the effects of spatial dimensions seen at the microscopic scale. Agent-based models incorporate stochasticity of parameters and implement timescales in individual agents leading to heterogeneity as seen in a spatially explicit grid and also have interactions between agents instead of averaged population interactions, so the above characteristics result in distinctive features. Another important differentiation between the two is the strong dependence of the ABM on initial conditions, leading to potential extinction / disease free states by spatial inhibition of interaction, and also outliers or strong deviations in certain runs from the average ABM behaviour occurring due to particular spatial distributions of the system, all of which are low probability or rare states. These occur for the same parameter sets in contrast to differential equations where parameters and initial conditions completely and deterministically provide the trajectory of the system.

5.2 Future work

While conversion of differential equation models to agent-based models requires appropriate equivalent parameters and behaviour rules, realistic models will additionally need verification of rules, calibration of parameters, and replacement of uniform parameters with distributions as found in populations.

This work may be extended into the analysis and conversion of detailed malaria models [Mandal 11, Sequeira 19] or used to create ABMs for containment strategies [Jin 20]. Agent-

based models may also be created corresponding to differential equation models of malaria or other diseases that have effectively reproduced results in specific geographies [Mandal 13]. ABMs may also be integrated with population data and networks of people realistically situated in a city or region [IIIangakoon 15], even making use of Geospatial Information Systems (GIS) software like ArcGIS and demographic profiles of ages to accurately predict disease outbreaks and plan interventions and public health responses.

References

[Agrawal 17] V Agrawal, P Moitra & S Sinha. *Emergence of Persistent Infection due to Heterogeneity*. Scientific Reports (Nature), 2017.

[Anderson 91] R M Anderson & R M May. Infectious diseases of humans: dynamics and control. Oxford University Press, 1991.

[Aron 82] J L Aron & R M May. The population dynamics of malaria. Chapman and Hall, 1982.

[Aron 88] J L Aron. *Mathematical modeling of immunity to malaria*. Mathematical Biosciences, 1988.

[Bernoulli 66] D Bernoulli. *Essai d'une nouvelle analyse de la mortalité causée par la petite verole et des avantages de l'inoculation pour la prévenir*. Mémoires de Mathématique et de Physique à l'Académie Royale des Sciences à Paris, 1766. *An attempt at a new analysis of the mortality caused by smallpox and of the advantages of inoculation to prevent it*.

[Blower 04] S Blower & D Bernoulli. *An attempt at a new analysis of the mortality caused by smallpox and of the advantages of inoculation to prevent it*. Reviews in Medical Virology, 2004.

[Brauer 01] F Brauer & C Castillo-Chávez. Mathematical models in population biology and epidemiology. Springer International Publishing, 2001.

[Butcher 63] J C Butcher. *Coefficients for the study of Runge-Kutta integration processes*. Journal of the Australian Mathematical Society, 1963.

[Cooley 65] J W Cooley & J W Tukey. *An algorithm for the machine calculation of complex Fourier series*. Mathematics of Computation, 1965.

[Edelstein-Keshet 88] L Edelstein-Keshet. Mathematical models in biology. 1988.

[GBD 14] GBD. *Global, regional, and national age-sex specific all-cause and cause-specific mortality for 240 causes of death, 1990-2013: a systematic analysis for the Global Burden of Disease Study*. Lancet, 2014.

[Gu 05] W Gu & R J Novak. *Habitat-based modeling of impacts of mosquito larval interventions on entomological inoculation rates, incidence, and prevalence of malaria*. American Journal of Tropical Medecine and Hygiene, 2005.

[Gunaratne 18] C Gunaratne & Garibay I. *NL4Py: Agent-Based Modeling in Python with Parallelizable NetLogo Workspaces*. CoRR, 2018.

[Gupta 94] S Gupta, J Swinton & R M Anderson. *Theoretical studies of the effects of heterogeneity in the parasite population on the transmission dynamics of malaria*. Proceedings of the Royal Society of London, 1994.

[Hasibeder 88] G Hasibeder & C Dey. *Population dynamics of mosquito-borne disease: persistence in a completely heterogeneous environment*. Theoretical Population Biology, 1988.

[IIIangakoon 15] C IIIangakoon, R D McLeod & M R Friesen. *Agent-based modelling of malaria*. IEEE Canada IHTC Conference 2014, 2015.

[Jin 20] X Jin, S Jin & D Gao. *Mathematical Analysis of the Ross–Macdonald Model with Quarantine*. Bulletin of Mathematical Biology, 2020.

[Kermack 27] W O Kermack & A G McKendrick. *A Contribution to the Mathematical Theory of Epidemics*. Proceedings of the Royal Society of London, 1927.

[Kurahashi 19] S Kurahashi. *Agent-Based Gaming Approach for Electricity Markets*. 2019.

[Liu 09] Q-X Liu, R-H Wang & Z Jin. *Persistence, extinction and spatio-temporal synchronization of SIRS spatial models*. Journal of Statistical Mechanics: Theory and Experiment, 2009.

[MacDonald 57] G MacDonald. The epidemiology and control of malaria. Oxford University Press, 1957.

[Mandal 11] S Mandal, R R Sarkar & S Sinha. *Mathematical models of malaria - a review*. Malaria journal, 2011.

[Mandal 13] S Mandal, S Sinha & R R Sarkar. *A Realistic Host-Vector Transmission Model for Describing Malaria Prevalence Pattern*. Bulletin of Mathematical Biology, 2013.

[Perez 19] L Perez, S Dragicevic & J Gaudreau. *A geospatial agent-based model of the spatial urban dynamics of immigrant population: A study of the island of Montreal, Canada*. PLoS ONE, 2019.

[Railsback 19] S F Railsback & V Grimm. Agent-based and individual-based modeling: A practical introduction, second edition. Princeton University Press, 2019.

[Ross 10] R Ross. The prevention of malaria. 1910.

[Sequeira 19] J Sequeira, J Louçã, A M Mendes & P G Lind. *Transition from endemic behaviour to eradication of malaria due to combined drug therapies: An agent-model approach*. Journal of Theoretical Biology, 2019.

[Smith 06] T Smith, G F Killeen, N Maire *et al. Mathematical modeling of the impact of malaria vaccines on the clinical epidemiology and natural history of Plasmodium falciparum malaria: Overview*. American Journal of Tropical Medecine and Hygiene, 2006.

[Smith 12] D L Smith, K E Battle, S I Hay, C M Barker, T W Scott & F E McKenzie. *Ross, Macdonald, and a Theory for the Dynamics and Control of Mosquito-Transmitted Pathogens*. PLoS Pathogens, 2012.

[Smith 18] N R Smith, J M Trauer, M Gambhir *et al. Agent-based models of malaria transmission: a systematic review*. Malaria Journal, 2018.

[Wang 19] H-H Wang, W E Grant, N C Elliott, M J Brewer, T E Koralewski, J K Westbrook, T M Alves & G A Sword. *Integrated modelling of the life cycle and aeroecology of wind-borne pests in temporally-variable spatially-heterogeneous environment*. Ecological Modelling, 2019.

[WHO 19] WHO. World malaria report 2019. 2019.

[Wilensky 99] U Wilensky. Netlogo. 1999. http://ccl.northwestern.edu/netlogo/.

[Wilensky 15] U Wilensky & W Rand. An introduction to agent based modelling. The MIT Press, 2015.