

**Modeling Microbial Dynamics: Effects on  
Environmental and Human Health**

by

Jude Dzevela Kong

A thesis submitted in partial fulfillment of the requirements for the degree of

Doctor of Philosophy

in

Applied Mathematics

Department of Mathematical and Statistical Sciences  
University of Alberta

© Jude Dzevela Kong, 2017

## Abstract

This thesis focuses on formulating and analyzing non linear models for microbial dynamics vis-a-vis human and environmental health. Firstly, we develop and investigate a stoichiometric organic matter decomposition model in a chemostat culture that incorporates the dynamics of grazers. This mechanistic biodegradation model lead to reliable and suggestive ecological insights in the preservation and restoration of our fragile ecosystems. Questions we attempt to answer include: (i) What mechanisms allow microbes and resources to persist uniformly or go extinct? (ii) How do grazing and dead microbial residues affect decomposition? (iii) How can the rate of decomposition be maximized or minimized? Secondly, we designed a greenhouse gas biogenesis model, which may be used to (i) predict the volume of greenhouse gasses emitted at any given time in an oil sands tailing pond and an end pit lake, (ii) calculate the time required to produce a given volume of cumulative greenhouse gases from them and (iii) estimate how long it will take for an oil sands tailing pond and an end pit lake to stop emitting greenhouse gases. Lastly, we formulate and analyze directly and indirectly transmitted infectious disease models. The questions aim to answer include: (i) Why are there irregularities in seasonal patterns of outbreaks amongst different countries? (ii) How can we estimate the transmission function of an infectious disease from a given incidence or prevalence data set? (iii) What is the estimated value of the basic reproduction number in affected regions? (iv) How can we control the period and intensity of pathogenic disease outbreaks?

This thesis is dedicated to my hero and loving mom, Anastasia Bongyu Kong. I love you mom.

## Acknowledgements

I would like to express my profound appreciation and gratitude to my advisor, Dr. Hao Wang, for investing a considerable amount of time and energy in guiding me throughout my research. His help, encouragement and patience are all invaluable. I would also like to thank my committee chair, Dr. Mark Lewis, and my committee member, Dr. Tariq Siddique, for their infectious enthusiasm, biological and mathematical insights as well as their availability and willingness to discuss new aspects and ideas of my research; even when their schedules were tight. I am also grateful to Drs. Thomas Hillen and Julia Foght for their friendly guidance and thought-provoking suggestions.

I would like to acknowledge the unquestionable love and support from my mom, Anastasia Bongyu Kong. A special appreciation also goes to my siblings Eunice Lenzemo Mbokam, Theresia Yeveyuvi, Ernest Berinyuy Kong and Kingsly Tardzenyuy Kong for their love and encouragement along the way. I would like to recognize my lab members and friends especially Michelle Michelle and Dr. Cyprain Ngolah for their optimism and support over the years. I am truly fortunate to have met all of you. Finally and most importantly, my deep gratitude goes to the Almighty God for his grace, wisdom and blessings throughout my entire life.

This research was generously supported by Natural Sciences and Engineering Research Council of Canada (NSERC), Pacific Institute of Mathematical Sciences (PIMS), Centre for Mathematical Biology (CMB) and Department of Mathematical and Statistical Sciences at the University of Alberta.

# Table of Contents

<b>1</b>	<b>Introduction</b>	<b>1</b>
1.1	Goal and objectives . . . . .	4
1.2	Outline of the thesis . . . . .	5
<b>2</b>	<b>Microbial per capita growth kinetic models</b>	<b>8</b>
2.1	Introduction . . . . .	8
2.2	Monod model . . . . .	11
2.3	Blackman model . . . . .	12
2.4	Exponential model (Tessier model) . . . . .	14
2.5	Haldane model . . . . .	15
2.6	Moser model . . . . .	16
2.7	Contois model . . . . .	17
2.8	Logarithmic model (Westerhoff model) . . . . .	17
2.9	Droop's cell quota model . . . . .	18
<b>3</b>	<b>A stoichiometric organic matter decomposition model in a chemostat culture</b>	<b>20</b>
3.1	Introduction . . . . .	21

3.2	The model . . . . .	25
3.3	Model analysis . . . . .	30
3.3.1	Positivity and boundedness . . . . .	31
3.3.2	Possible equilibrium points . . . . .	33
3.3.3	Stability of the equilibrium points . . . . .	40
3.3.4	Persistence-extinction criteria . . . . .	48
3.3.5	Maximizing decomposition rate . . . . .	60
3.4	Numerical experiments . . . . .	62
3.4.1	Impact of grazing on decomposition . . . . .	67
3.4.2	Impact of recycled death bacteria on decomposition . . . . .	70
3.4.3	Switching time . . . . .	71
3.4.4	Sensitivity analysis . . . . .	72
3.5	Discussion . . . . .	75
<b>4</b>	<b>Predictive model for methane emissions from oil sands tailings ponds and end pit lakes</b>	<b>79</b>
4.1	Graphical abstract . . . . .	80
4.2	Introduction . . . . .	81
4.3	Material and method . . . . .	84
4.3.1	Model development . . . . .	84
4.3.2	Source of data for parameter estimation and model validation . . . . .	88
4.3.3	Mathematical analysis . . . . .	92
4.3.4	Fitting the model to data . . . . .	93
4.3.5	Maximum theoretical methane yield . . . . .	94

4.3.6	Model validation . . . . .	95
4.4	Results and discussion . . . . .	96
4.4.1	Mathematical analysis . . . . .	96
4.4.2	Fitting the model to data . . . . .	106
4.4.3	Model validation . . . . .	112
4.4.4	Discussion . . . . .	113
4.5	Conclusions . . . . .	117
<b>5</b>	<b>Stability and sensitivity analysis of the iSIR model for in-</b>	
	<b>directly transmitted infectious diseases with immunological</b>	
	<b>threshold</b>	<b>119</b>
5.1	Introduction . . . . .	120
5.2	iSIR model formulation . . . . .	123
5.3	Mathematical results . . . . .	128
5.3.1	Forward invariance . . . . .	128
5.3.2	Equilibria of the system . . . . .	130
5.3.3	Local stability of $E_0, E_1$ and $E^*$ . . . . .	134
5.3.4	Local stability of $E_{1,2}^+$ . . . . .	139
5.3.5	Global stability of $E_1$ and $E^*$ . . . . .	141
5.4	Numerical simulations . . . . .	151
5.5	Local sensitivity analysis . . . . .	153
5.5.1	Sensitivity of the outbreak peak . . . . .	153
5.5.2	Sensitivity of the outbreak peak time . . . . .	155
5.5.3	Sensitivity of the endemic steady state . . . . .	157
5.6	Discussion . . . . .	159

<b>6</b>	<b>Dynamics of a cholera transmission model with immunological threshold and natural phage control in reservoir</b>	<b>163</b>
6.1	Introduction . . . . .	164
6.2	Model formulation . . . . .	168
6.3	Forward invariance . . . . .	171
6.4	Existence and stability of equilibria with no shedding . . . . .	173
6.4.1	Existence of equilibria . . . . .	173
6.4.2	Linearization . . . . .	176
6.4.3	Stability of the disease free, bacteria free, phage free equilibrium $E_0$ . . . . .	176
6.4.4	Stability of the disease free, phage free equilibrium $E_K$	176
6.4.5	Stability of the disease free equilibrium $E_1$ . . . . .	177
6.4.6	Stability of the phage free endemic equilibrium $E_K^*$ . . . . .	179
6.4.7	Stability of the interior endemic equilibrium $E_1^*$ . . . . .	180
6.4.8	Local stability summary, bifurcation diagrams and numerical simulations . . . . .	181
6.5	Existence and stability of equilibria with shedding . . . . .	185
6.5.1	Existence of equilibria . . . . .	185
6.5.2	Linearization . . . . .	189
6.5.3	Stability of disease free, bacteria free, phage free equilibrium $E_0$ . . . . .	190
6.5.4	Stability of the boundary equilibrium $E_K$ . . . . .	190
6.5.5	Stability of the disease free equilibrium $E_1$ . . . . .	191
6.5.6	Stability of endemic equilibria $E^*$ and $E_{1,2}^*$ . . . . .	192

6.5.7	Local stability summary, bifurcation diagrams and numerical simulations . . . . .	193
6.6	Chaos . . . . .	202
6.7	Discussion . . . . .	207
<b>7</b>	<b>The inverse method for a childhood infectious disease model with its application to pre-vaccination and post-vaccination measles data</b>	<b>210</b>
7.1	Introduction . . . . .	212
7.2	The SEIRA model . . . . .	217
7.3	Qualitative analysis . . . . .	220
7.4	Sensitivity analysis . . . . .	226
7.4.1	Sensitivity analysis of the outbreak peak value . . . . .	227
7.4.2	Sensitivity analysis of the outbreak peak time . . . . .	229
7.4.3	Sensitivity analysis of the endemic steady state . . . . .	230
7.5	Extracting the time-dependent transmission rate $\beta(t)$ from prevalence pre vaccination data . . . . .	231
7.6	Extracting the time dependent transmission rate from pre-vaccination incidence data . . . . .	237
7.6.1	Solution of the inverse problem for the SEIRA model . . . . .	237
7.7	The SEIRA model with vaccination . . . . .	243
7.8	Qualitative analysis . . . . .	244
7.9	Sensitivity analysis . . . . .	246
7.9.1	Sensitivity analysis of the outbreak peak value . . . . .	246
7.9.2	Sensitivity analysis of the endemic steady state . . . . .	248

7.10	Extracting the time dependent transmission rate $\beta(t)$ from prevalence post vaccination data . . . . .	249
7.11	Extracting the time dependent transmission rate $\beta(t)$ from incidence post vaccination data . . . . .	252
7.12	Discussion . . . . .	255
<b>8</b>	<b>Future directions</b>	<b>259</b>
	<b>Bibliography</b>	<b>265</b>

# List of Tables

3.1	Definition and values of parameters in System (4.1) . . . . .	30
3.2	Notations and expressions for the equilibrium points of System (4.1). . . . .	39
3.3	Criteria for existence of the equilibrium points of System (4.1).	39
3.4	Summary of conditions for asymptotic stability of possible equilibrium points of System (4.1). . . . .	48
3.5	Sensitivity of degradation rate . . . . .	74
4.1	Definition and values of some of the bacteria related parameters of System (4.2) . . . . .	94
4.2	Parameter values of the biodegradation model . . . . .	108
4.3	Zero-and first-order model related parameters values for some of the biodegradable hydrocarbons . . . . .	115
4.4	Model comparison . . . . .	116
5.1	Parameter values . . . . .	127
5.2	Parameter values . . . . .	127
5.3	Sensitivity of parameters to outbreak peak . . . . .	154
5.4	Sensitivity of parameters to outbreak time . . . . .	155

5.5	Sensitivity of parameters to endemic equilibrium . . . . .	157
6.1	Parameter values . . . . .	170
6.2	Ordering of the roots to the endemic equation . . . . .	188
7.1	Parameter values for measles . . . . .	219
7.2	Sensitivity indices of the outbreak peak value . . . . .	227
7.3	Sensitivity of the outbreak peak time . . . . .	229
7.4	Sensitivity of the endemic steady state . . . . .	230
7.5	Sensitivity of the outbreak peak value with vaccination . . . . .	247
7.6	Sensitivity of the outbreak peak time with vaccination . . . . .	247
7.7	Sensitivity of the endemic steady state with vaccination . . . . .	248
8.1	Definition of the parameters of System (8.1) . . . . .	262

# List of Figures

3.1	Phase portrait of System (3.17) . . . . .	53
3.2	one-parameter bifurcation diagram for System (4.1) . . . . .	64
3.3	Bacteria-grazers dynamics when $C_{in} = 0.3mgC/dm^3$ . . . . .	65
3.4	Bacteria-grazers dynamics when $C_{in} = 1.5 mgC/dm^3$ . . . . .	65
3.5	Bacteria-grazers dynamics when $C_{in} = 5.5 mgC/dm^3$ . . . . .	66
3.6	Bacteria-grazers dynamics when $C_{in} = 12.08 mgC/dm^3$ . . . . .	66
3.7	Two-parameter bifurcation for System (4.1) . . . . .	67
3.8	Impact of grazing on decomposition . . . . .	69
3.9	Critical value of $C(0) : N(0)$ that ensures complete degradation when $G = 0$ . . . . .	69
3.10	mpact of recycled death bacteria on decomposition . . . . .	70
3.11	Switching time. $N_{in} = 15 mgN/dm^3$ and $C_{in} = 3 mgC/dm^3$ .	72
3.12	Switching time. $N_{in} = 8 mgN/dm^3$ and $C_{in} = 25 mgC/dm^3$ .	72
4.1	Nullclines for Case 1.1 . . . . .	98
4.2	Nullclines for Case 1.2 . . . . .	99
4.3	Nullclines for Case 1.3 . . . . .	100
4.4	Nullclines for Case 2.1 . . . . .	101

4.5	Nullclines for Case 2.2 . . . . .	102
4.6	Nullclines for Case 2.3 . . . . .	103
4.7	System (4.2) fit to measured alkane biodegradation values . .	109
4.8	System (4.2) fit to measured biodegradable BTEX compounds data. . . . .	110
4.9	System (4.2) fit to measured biodegradable isoalkenes' data. .	111
4.10	Simulated vs measured methane data . . . . .	112
4.11	Comparison between simulated (by all our three models) and measured methane data . . . . .	117
5.1	Indirect incidence term . . . . .	124
5.2	Flow diagram . . . . .	125
5.3	Bacteria upper bound . . . . .	129
5.4	Invariant region . . . . .	130
5.5	Endemic equilibria equation . . . . .	132
5.6	Trajectories in the phase space for $C > 1$ and $B^* \geq C$ . . . . .	140
5.7	Monotonicity . . . . .	141
5.8	Disease free phase diagram . . . . .	151
5.9	Endemic phase diagram . . . . .	152
6.1	Bifurcation diagrams . . . . .	183
6.2	Disease causing cycles . . . . .	185
6.3	Bifurcation diagram 1 . . . . .	195
6.4	Disease free cycles with shedding . . . . .	197
6.5	Disease causing cycles with shedding . . . . .	198
6.6	Bifurcation diagram 2 . . . . .	199

6.7	Disease causing cycles with shedding 2 . . . . .	201
6.8	Effects of $\xi$ on the period of outbreaks and peak value of I . . .	202
6.9	Monthly cholera outbreaks . . . . .	204
6.10	Chaotic bifurcation diagram 1 . . . . .	205
6.11	Lyapunov exponents . . . . .	207
7.1	Transmission rate $\beta(t)$ extracted from fake prevalence data . . .	236
7.2	Transmission rate $\beta(t)$ extracted from fake incidence data . . .	240
7.3	Measles weekly notification data in Liverpool and London . . .	241
7.4	Transmission rate of Liverpool from 1944 to 1966 . . . . .	242
7.5	Transmission rate of London from 1944 to 1966 . . . . .	242
7.6	Transmission rate of Liverpool from 1974 to 1986 . . . . .	253
7.7	Transmission rate of London from 1974 to 1985 . . . . .	254
7.8	1950-1966 weekly notification data for England and Wales . . .	257
7.9	Transmission rate of England and Wales from 1950 to 1952 . . .	258
8.1	General schematic biodegradation pathway of organic matter . . .	260

# Chapter 1

## Introduction

Microorganisms are microscopic or submicroscopic organisms with undifferentiated unicells [14]. They include bacteria, archaea, and fungus [23]. Microorganisms are ubiquitous. Despite their small sizes, they have a huge impact on environmental and human health. Some have profound beneficial effects and others are seriously harmful. However, many microorganisms produce mixed effects.

One of the most significant positive effect of the microorganisms on human and environmental health is their ability to degrade and detoxify pollutants. The contamination of the environment with hazardous and toxic chemicals is one of the main problems faced by today's world. These chemicals are either released on purpose or by accident. Microbes, mostly bacteria, degrade the organic chemicals to innocuous compounds and eventually to methane and carbon dioxide [2]. This process is known as biodegradation. With inorganic pollutants like heavy metals, microbes help reduce their spread in the environment by changing their oxidation state, which affects their mobility [23].

The activity of microbes on the pollutants is usually natural and thus does not require any external energy. Due to a rising level of industrial activities, the rate at which we pollute our environment with organic substances is far beyond the rate of biodegradation. There is a need to accelerate biodegradation in any environment with pollutants. This is the subject of Chapter 3 of this thesis.

Microbes are also essential in supporting life in the lakes and oceans, and subsequently in sustaining the fish population. Algae, one of the microbes, are the primary produces in open water. They grow in open water by taking up nutrients such as phosphorus and nitrogen from the water and capturing energy from sunlight. This energy and nutrients are indirectly transferred to the fish (they provide food for protist which in turn are eaten by zooplankton that are a source of food for the fish [23]).

Albeit biodegradation helps to remove the hazardous and toxic organic chemicals from our environment, the end-products such as greenhouse gases (methane and carbon dioxide) have negative effects on our climate. The greenhouse effect increases the temperature of the Earth by trapping heat in our atmosphere. These gases are powerful infrared absorber gases that cause the earth's temperature to increase by trapping heat in our atmosphere (greenhouse effect) [31]). Because of greenhouse gases, the average temperature of the earth is 37 °C higher than it would have been without them [23]. Global warming hypothesis postulates that increasing the concentration of greenhouse gases leads to more heat being trapped within the atmosphere of the earth leading to increase in temperature. It is predicted in [2] that a 1 to 2 °C increase in the average global temperature might among other consequences

1) cause the ocean level to rise as oceans undergo thermal expansion, 2) alter the species composition of forest and grasslands and this might lead to the potential displacement or extinction of some species and 3) cause the disappearance of some wetlands due to flooding. Thus the ongoing warming of the atmosphere has significant effects and therefore is of particular concern.

Many microorganisms cause diseases in human beings. These diseases are called infectious diseases and include:

- 1) (caused by bacteria) salmonella, tetanus, typhoid, cholera, gangrene, bacterial dysentery, diphtheria, tuberculosis, bubonic plague, meningococcal meningitis, pneumococcal pneumonia,
- 2) (caused by viruses) rabies, influenza (flu), measles, mumps, polio, rubella (german measles), chicken pox, colds, warts, cold sores,
- 3) (caused by protoctists) malaria, amoebic dysentery,
- 4) (caused by fungi) athlete's foot, ringworm.

The disease-causing microbes are called pathogens. Most infectious diseases can be transmitted through the following means:

- i) by consuming food or water containing pathogens or their toxic products (e.g. salmonella, typhoid, cholera),
- ii) by inhaling or ingesting droplets of moisture which have been coughed, sneezed or breathed out by an infected person (e.g. include colds, flu),
- iii) by entry through a wound or sore (e.g. tetanus),
- iv) by direct contact with an infected person (e.g. athlete's foot, ringworm),

v) through vectors (e.g. malaria transmitted by mosquitoes).

## 1.1 Goal and objectives

Drawing on the research requirements stated above, the goal of this study is to develop and analyze nonlinear models for microbial dynamics pertaining to human and environmental health. The objectives of the thesis are fivefold:

Objective 1: To design and analyze a stoichiometric organic matter decomposition model in a chemostat culture. In particular, we 1) design a stoichiometric biodegradation model, 2) determine the mechanisms that allow microbes and resources to persist uniformly or go extinct, 3) study the effects of bacteria grazers on decomposition, 4) perform the sensitivity analysis of the degradation rate to model parameters, 5) carry out a bifurcation analysis of the system, 6) determine the impact of dilution and dead microbes residues on decomposition etc.

Objective 2: To develop a predictive model for methane emissions from oil sands tailings ponds and end pit lakes. Specifically, we 1) develop a greenhouse gas emission model 2) estimate the parameter values of the model using empirical data, and 3) validate the model.

Objective 3: To perform the stability and sensitivity analysis of an indirectly transmitted infectious disease model with an immunological threshold.

Objective 4: To Formulate and analyze a cholera transmission model which explicitly includes the dynamics of bacteriophage and bacteria and contains an

indirect infection term which accounts for a minimum infectious dose of the pathogen *V. cholerae*.

Objective 5: To Design an algorithm for estimating the transmission function of an infectious disease from a given incidence or prevalence data set. In particular, we 1) formulate childhood infectious diseases models with and without vaccination, 2) determine the positive and bounded set for the solutions of the systems, 3) perform the stability analysis of both models, 4) perform the sensitivity analysis of the outbreak peak value, time of outbreak peak, and the steady state value of the infected compartment to the parameters of the systems, 5) construct inverse algorithms for extracting the time-dependent transmission rates from prevalence and incidence pre-and post-vaccination data and 6) test the performance of the incidence algorithms using pre- and post-vaccination measles data from Liverpool and London.

## 1.2 Outline of the thesis

This thesis is organized into eight chapters as follows:

Chapter 2: Chapter 2 provides the literature review of some of the existing models for microbial growth kinetics.

Chapter 3: In this Chapter, we first design stoichiometric organic matter decomposition model in a chemostat culture. Secondly, we prove the positivity and boundedness of the formulated system. Thirdly, we determine the criteria for the uniform persistence and extinction of the species.

Fourthly, we determine the optimal value of grazers that maximizes the per capita decomposition rate of organic carbon. Moreover, we perform a one and two-parameter bifurcation analyses. Furthermore, we numerically compare the effects of grazers on degradation in a microcosm and in a chemostat. In addition to these, we determine the switching time of bacterial growth rate from carbon dependent to nitrogen dependent for different continuous cultures. We also discuss the sensitivity of the degradation rate to model parameters. Lastly, we discuss the results.

Chapter 4: Here, we develop a predictive model for methane emissions from oil sands tailing ponds and end pit lakes. The model is fitted to experimental data and cross validated. We end the chapter with discussion on the model and results as well as a conclusion.

Chapter 5: In Chapter 5, we carry out a stability and sensitivity analysis of the model in [133] and discuss the results.

Chapter 6: In this Chapter, firstly, we propose a cholera transmission model which explicitly includes the dynamics of bacteriophage and bacteria and also contains an indirect infection term which accounts for a minimum infectious dose of the pathogen *V. cholerae*. Next, we provide a forward invariance domain for the model. Thirdly, we perform the stability and bifurcation analysis. Fourthly, we provide an explanation to the different nature of the outbreaks experienced around the world and end the chapter with a discussion of the results.

Chapter 7: In this chapter we first formulate SEIRA models for childhood infectious

diseases with and without vaccination. Secondly, we study the positivity, boundedness and equilibria of the SEIRA models. Thirdly, we calculate, analyze and compare the normalized forward sensitivity indices of the outbreak peak value, time of outbreak peak and steady state value of the infected compartment to the parameters of the systems. Fourthly, we construct algorithms using an inverse method for extracting the time-dependent transmission rate from prevalence and incidence pre- and post-vaccination data. Lastly, we discuss the results.

Chapter 8: Finally, Chapter 8 concludes with an outlook on planned research.

# Chapter 2

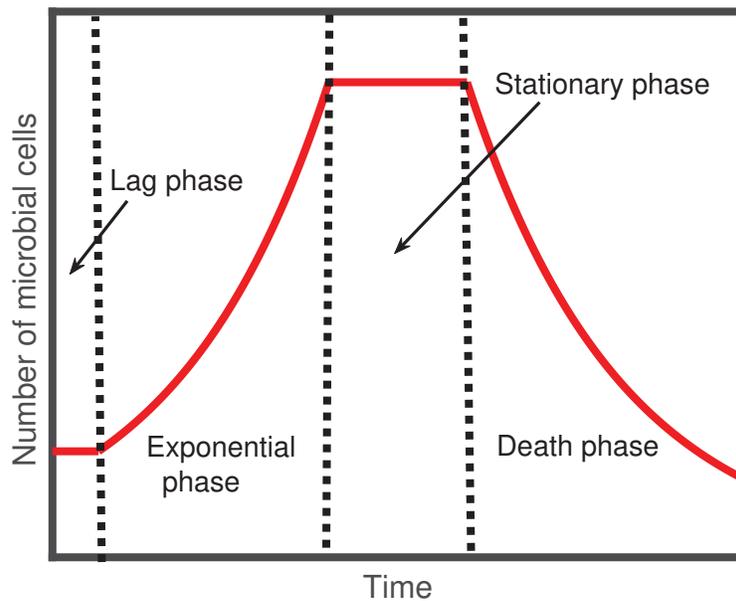
## Microbial per capita growth kinetic models

### 2.1 Introduction

A basic microorganisms growth kinetic model answers the question of how microbial concentration will change in the future, given its current concentration and the concentration of substrate in the ecosystem that the microbe is exposed to. These changes in the concentration of microbes may be changes in the total number of individuals present; i.e., in the number of microbes members, but may also pertain to changes in the composition of the microbes. The elements in the ecosystem affect microbes growth rate as they are all required by microbes cells due to the elements' specific roles in biosynthesis. Following Justus Von Liebig's minimum law, the growth rate of bacteria is proportional to the most limiting elements. The per capita microbial growth rate, gives the gain in microbial biomass per unit time as a function of the limiting elements.

Obviously, if there is no element in the environment, then the microbes cannot eat. One expects intuitively that the more elements that are present, the more the microbes eat, so that the per capita microbial growth rate is an increasing function of its limiting elements. However, at a certain point, these limiting elements may no longer be the limiting elements. Further, huge abundance of an element may even inhibit microbial growth.

Figure (2.1) displays a common time course of microbial growth in a batch culture.



**Figure 2.1:** Microbial kinetic curve

When microbes are introduced in an unfamiliar ecosystem or growth limiting elements are introduced into an ecosystem with microbes, there is a lag phase before the microbes growth kinetics are affected by each of the fresh growth limiting substrate. During this lag phase, individual microbes syn-

thesize RNA, enzymes, and other molecules they need for growth prior to resuming their division. This phase could as well be due to the physiological adjustments of microbes to changes in the concentration of elements or culture conditions. Increasing the growth limiting nutrient concentration will lead to an increase in the growth function. This phase is known as the exponential growth phase (sometimes called the log phase). In the course of this phase, cells grow and divide as an exponential function of time. Eventually, a concentration is reached at which further increases do not lead to an increase in growth. At this point either some other environmental factors are in a limited supply or else the microbes cells themselves have reached their own limits for the present cultural conditions. With time, the growth function will enter a decline phase or death phase. This phase might be as a result of the decrease in the concentration of the given limiting element leading to the microbial death, environmental temperature above or below the tolerance level for microbes, inhibition of microbial growth by the given limiting element or other adverse conditions. This behaviour is atypical of microbes and has an obvious importance in the metabolic activities of microbes in industrial fermentations, biological waste treatment processes, and in other biodegradation processes. Many microbial kinetics models assume that only a single element is usually limiting. The first models for microbial growth rate that tried to capture the dynamics of a microbial growth function and its limiting unique element were proposed by Blackman in 1905 [4] and Monod in 1942 [73]. These models relate the concentration of a given limiting element to the per capita growth rate of microbes using simple empirical equations. These equations provide a foundation that accounts for more specialized cases represented by the com-

monly used empirical models. In this section we review some of the models mostly used in describing microbial specific growth rate.

## 2.2 Monod model [73]

The Monod kinetic model was proposed by Jacques Lucien Monod (1910-1976), based on his empirical findings. It is also one of the earliest microbial growth models. This model describes the per capita growth of microbes with a constant yield on a single growth-limiting substrate ( $S$ ). The Monod model is given by:

$$\mu_M(S) = \mu_{M\max} \frac{S}{K_{SM} + S}, \quad (2.1)$$

where  $\mu_M$  is the specific growth rate,  $\mu_{M\max}$  is the maximum specific growth rate of microbes,  $S$  is the concentration of substrate and  $K_{SM}$  is the substrate concentration which supports the half-maximum specific growth rate. At a low substrate concentration ( $K_{SM} \gg S$ ) the model can be approximated by a first-order kinetics model. As substrate concentration is increased far above the half-saturation constant ( $K_{SM} \ll S$ ), a gradual change to zero-order behaviour occurs. Because of the gradual transition to zero-order, this model may underestimate the per capita growth rate in the neighbourhood of the point where the transition occurs. Also, this model is unable to capture the lag phase as well as the death phase. As a consequence, the model is bound to overestimate microbial growth during the lag phase where there is no growth as well during the death phase. Monod model is therefore more

applicable in the following situations: 1) ecosystems with no adverse conditions (perfect temperature), 2) ecosystems in which at high substrate concentration, microbial growth and death rates are equal 3) ecosystems with no substrate inhibition or toxic substrates 4) ecosystems with mature microbes that are accustomed to them and their substrates.

Next we will look at the Blackman model. This model behaves almost like the Monod model but unlike the Monod model, it allows a rapid transition from first-order to zero-order when substrate concentration exceed half-saturation constant. Thus, the Blackman model may avoid underestimating the per capita growth rate in the neighbourhood of the point where the transition occurs.

## 2.3 Blackman model [4]

This is one of the first per capita microbial growth models. The assumptions that led to the model are:

- at a low substrate concentration the rate of the uptake is proportional to the given substrate concentration, and
- at a high substrate concentration both uptake and growth rates are independent of the given substrate concentration because some other substrates are limiting or microbial growth limiting factors are present.

Just like the Monod model, at a low given substrate concentration, Blackman modelled the microbial per capita growth using the first-order kinetic model and at high substrate concentration, he modelled it using the zero-order kinetic

model. His model is a piecewise function that makes a sharp switch from a first-order to a zero-order when substrate concentration exceeds the concentration that supports half-maximum specific growth rate. This piecewise function is not continuous. The function is given by the following equation:

$$\mu_B(S) = \begin{cases} \frac{\mu_{B\max}S}{K_{SB}} & S < K_S, \\ \mu_{\max} & S \geq K_S, \end{cases} \quad (2.2)$$

where  $\mu_B$  is the specific growth rate,  $\mu_{B\max}$  is the maximum specific growth rate of microbes,  $S$  is concentration of substrate and  $K_{SB}$  is a constant. Also like the Monod model, this model is unable to capture the lag phase as well as the death phase (in Figure (2.1)). Thus as with the Monod model, the Blackman model is bound to overestimate microbial growth during the lag and death phases. Unlike the Monod model, this model saturates rapidly and thus may avoid underestimating the per capita growth rate around the neighbourhood of the point where the transition occurs. The condition under which the model is applicable are similar to those listed under the Monod model and will be omitted here. Because of the sharp transition, the Blackman model may turn to overestimate the per capita growth rate. The next model is a compromise between the Monod model and the Blackman model; it allows a continuous switch from zero-order to first-order, but this transition is sharper than that allowed by the Monod model.

## 2.4 Exponential model (Tessier model)

The exponential specific microbial growth rate model also known as the Tessier model is based on the assumption that the dependence of the per capita growth rate  $\mu_T$  on the limiting substrate is proportional to the difference between the per capita growth rate and the maximum growth rate of microbes. That is

$$\frac{d\mu_T}{dS} = \frac{1}{K_{ST}}(\mu_{T\max} - \mu_T),$$

where  $K_{ST}$  is the proportional constant. Integrating gives the famous Tessier model for specific microbial growth rate, Equation (2.3).

$$\mu_T(S) = \mu_{T\max}(1 - e^{-\frac{S}{K_{ST}}}), \quad (2.3)$$

where  $\mu_T$  is the specific growth rate,  $\mu_{T\max}$  is the maximum specific growth rate of microbes,  $S$  is the concentration of substrate and  $K_{ST}$  is a constant. The Monod model, Blackman model and Tessier model mainly differ in the way the change from first-order to zero-order behaviour occurs. At low and very high substrate concentrations, the three models predictions are approximately the same. At the point where the transition from low substrate concentration dynamics to high substrate concentration dynamics occurs, the Tessier kinetics will lie in between those predicted by the Monod and Blackman. Thus the Tessier growth curve always lies in between the Blackman and Monod curves [44]. The model in the following section is an extension of the Monod model to capture the death phase of microbial growth curve (effects of inhibitory substrates on microbial growth).

## 2.5 Haldane model [18]

Empirical studies reveal that at high concentrations, the specific growth rate of a microbial population may be inhibited by the substrate (death phase) [48]. In this case, the Haldane model given in Equation (2.4) can be used to describe the specific microbial growth rate. This equation can be viewed as an extended version of the Monod equation. It was proposed by Haldane [18] to model enzymes forming an inactive enzyme-substrate complex involving two molecules of substrate per enzyme molecule. This is a famous model for inhibition of microbial growth [19].

$$\mu_H(S) = \mu_{H\max} \frac{S}{K_{SH} + S + \frac{S^2}{K_i}}, \quad (2.4)$$

where  $\mu_H$  is the specific growth rate,  $\mu_{H\max}$  is the maximum specific growth rate of microbes,  $S$  is concentration of substrate and  $K_{SH}$  and  $K_i$  are constants. In addition to the exponential phase and stationary phase of microbial kinetics captured by the models in the previous sections, this model equally captures the death phase. If the inhibition constant  $K_i$  is very large, the model reduces to the Monod model. In situations where the Monod model is applicable, this model is likely to give more realistic results and thus it is advisable to use this as oppose to the Monod model. None of the models covered so far is able to capture the lag phase of microbial kinetics. The model that follows generalizes the Monod model such that for a certain range of its parameter values, the lag phase may be approximately captured.

## 2.6 Moser model [29]

Herman Moser's model [29] is a derivation from the Monod model. Moser modified the Monod model by adding a third parameter,  $n > 1$ , to capture the effects of adoption of microbes to stationary processes by mutation. The model is given by

$$\mu_{Mo}(S, B) = \mu_{M\text{omax}} \frac{S^n}{K_{SMo} + S^n}, \quad (2.5)$$

where  $\mu_{Mo}$  is the specific growth rate,  $\mu_{M\text{omax}}$  is the maximum specific growth rate of microbes,  $S$  is concentration of substrate,  $n$  is an adjustable parameter and  $K_{SMo}$  is a constant. This model provides a degree of flexibility in fitting data and can simulate interesting dynamics in a continuous stirred tank reactor [6]. It is similar to the Hill equation in enzymology. Like the Monod model, this model does not capture the death phase of microbial kinetics. Unlike the Monod model, for large values of  $n$ , the lag phase may be approximately captured by this model. The Moser model is a better model compared to the Monod model as  $n$  can be fine-tuned to capture some complex dynamics that may not be perfectly captured with the Monod model. In comparison with the Haldane model, the best model will be based on the nature of the culture. If the substrate in the culture is toxic, and the microbes were familiar with the culture and ready for cell division, the Haldane model will be preferable. On the other hand, if the death phase were negligible and the lag phase significant, the Moser model will be preferable.

## 2.7 Contois model [11]

Contois microbes growth model [11] is also a derivation from the Monod model. Contois proposed the model following the results of his work after applying the Monod model to *Aerobacter aerogenes* on defined mineral salt media containing ammonium as a nitrogen source and glucose or succinic acid as carbon sources. His work reveals that the specific growth rate of microbes did not only depend on the limiting nutrients but also was inversely proportional to the quantity of microbes,  $B$  in the culture. He then proposed the following model for microbial specific growth rate:

$$\mu_C(S) = \frac{S}{K_{SC}B + S}, \quad (2.6)$$

where  $\mu_C$  is the specific growth rate,  $S$  is concentration of substrate and  $K_{SC}$  is a constant. [15] compared the experimental and simulated data of fed-batch and batch fermentation by various growth kinetic models and the Contois model gave the best results among all the other models that included the Monod, Tessier, Moser and Haldane models. Like the Monod, Blackman, and Tessier models, this model does not capture the lag and death phases.

## 2.8 Logarithmic model (Westerhoff model)

The Westerhoff model [50] also known as the logarithmic model like some of the models seen so far, describes the per capita growth rate as a continuous function of the substrate concentration. The model equation is given by

Equation 2.7:

$$\mu_W = a + b \ln(S), \quad (2.7)$$

where  $\mu_W$  is the specific growth rate,  $a$  and  $b$  are constants. The model predicts a period of rapid increase, followed by a period where the growth slows down, but the growth continues to increase without bound. Thus the model does not capture the stationary phase, the death phase as well as lag phase of microbial growth. This function may overestimate growth in the lag phase, exponential phase, and death phase. According to the model, if the substrate concentration is very low, the growth may be negative. This is unrealistic as the per capita growth rate is always positive. Compared to the model seen so far, this model is not an appropriate model for microbial growth kinetics.

## 2.9 Droop's cell quota model [13]

There is a difference between substrate uptake and substrate-controlled growth. Substrate-controlled growth depends on substrate inside the cell. Michael Droop defined the cell quota,  $Q$ , as the total cell substrate per unit biomass [13]. This definition permits the growth rate to depend on an internal substrate pool. While analyzing observations from vitamin  $B_{12}$  limited chemostat cultures of *Monochrysis lutheri*, Droop discovered that the specific growth rate,  $\mu_D$  is related to the cell quota  $Q$ , by the relationship given in Equation (2.8).

$$\mu_D(Q) = \mu_{Dmax} \left(1 - \frac{q}{Q}\right), \quad (2.8)$$

where  $q$  is the smallest amount of internal nutrient on which the cell can exist and  $\mu_{Dmax}$  is the maximum specific growth rate. This model does not capture the lag phase as well as the death phase of microbial growth.

The above specific growth rate models are valid for a single limiting element. However, when microbes are growing in an environment in which more than one of its required substrates are present at less than saturating levels, the limiting element might switch. In this case, we need to employ a threshold approach and follow Leibig's law of the minimum. In this situation, stoichiometry plays an important role in the decomposition model. This is addressed in the next chapter.

# Chapter 3

## A stoichiometric organic matter decomposition model in a chemostat culture <sup>1</sup>

### Abstract

Biodegradation, the disintegration of organic matter by microorganism, is essential for the cycling of environmental organic matter. Understanding and predicting the dynamics of this biodegradation have increasingly gained attention from the industries and government regulators. Since changes in environmental organic matter are strenuous to measure, mathematical models are essential in understanding and predicting the dynamics of organic matters.

---

<sup>1</sup>This Chapter has been published. Reference: Jude D. Kong, Paul Salceanu, and Hao Wang. “A stoichiometric organic matter decomposition model in a chemostat culture.” *Journal of Mathematical Biology* (2017): 1-36.

Empirical evidence suggests that grazers' preying activity on microorganism helps to facilitate biodegradation. In this paper, we formulate and investigate a stoichiometry-based organic matter decomposition model in a chemostat culture that incorporates the dynamics of grazers. We determine the criteria for the uniform persistence and extinction of the species and chemicals. Our results show that 1) if at the unique internal steady state, the per capita growth rate of bacteria is greater than the sum of the bacteria's death and dilution rates, then the bacteria will persist uniformly; 2) If in addition to this, (a) the grazers' per capita growth rate is greater than the sum of the dilution rate and grazers' death rate, and (b) the death rate of bacteria is less than some threshold, then the grazers will persist uniformly. These conditions can be achieved simultaneously if there are sufficient resources in the feed bottle. As opposed to the microcosm decomposition models' results, in a chemostat culture, chemicals always persist. Besides the transcritical bifurcation observed in microcosm models, our chemostat model exhibits Hopf bifurcation and Rosenzweig's paradox of enrichment phenomenon. Our sensitivity analysis suggests that the most effective way to facilitate degradation is to decrease the dilution rate.

### **3.1 Introduction**

Biodegradation is the process by which organic matters are broken down into smaller compounds by microorganisms [72]. This occurs when microorganisms use organic matters as a source of energy and nutrients. This is an essential process that helps keep our planet clean and healthy as well as return nutrients

into the ecosystem. The pollution of the waste product in our environment has increasingly gained the attention of industrial and government regulators who are interested in preventing the potential negative impacts on human or ecosystem's health. This goal can be efficiently achieved only by mathematical modeling because changes in environmental organic matter are difficult to measure, especially in a long period of time, and the future density of organic matter can only be predicted via modeling.

The most important organisms that break down organic matter in many environments are bacteria. Bacteria are minute organisms which occur everywhere [79]. The growth rate of bacteria depends on resource availability. The population of these vital microorganisms are heavily grazed by their predators [76]. These grazers include protozoans, saprophagous and nematodes. Grazing has several effects on the community. It leads to a decrease in bacterial population [65], stimulates mineralization of nutrients [60, 68, 54] and respiration [68], and speeds up degradation of organic substances [83]. The grazers could produce bacterial growth-limited substances [68] and release density dependent factors that limit bacterial population growth [61]. Thus biodegradation is by no doubt greatly affected by bacterial grazers. There is a need to understand the interactions between bacteria and their grazers, in order to optimize biodegradation.

Organic matter subject to decomposition usually contains a mixture of different elements such as carbon, nitrogen, sulphur and phosphorus. The growth rate of bacteria depends on the various elements that they obtain from decomposing organic matters. These elements are all required by bacterial cells because they fulfill specific roles in biosynthesis. Based on Justus Von

Liebig's minimum law, the growth rate of bacteria is proportional to the most limiting resource. If in an organic matter all other elements except carbon and nitrogen are present in abundance, the growth rate of microbes will be a function of nitrogen and carbon. In this simplified situation, we already need to incorporate the stoichiometry of microbes in modeling decomposition. Ecological stoichiometry is the study of the balance of energy and chemical elements in ecological interactions [82]. It is based on the observation that there are stoichiometric mismatches between different levels in the food chain that greatly affect trophic efficiency and nutrient fluxes [67] .

A majority of existing biodegradation models formulate bacterial per capita growth rate using Monod function [73]. When multiple limiting resources are considered, we need to apply Liebig's minimum law with multiple Monod functions. In this case, stoichiometry plays an important role in the decomposition model. In addition, we include grazing dynamics in our model.

A biodegradation model that incorporates the effect of the grazers on bacteria is developed in [84]. The bacteria and grazers are assumed to be in a batch culture, and nitrogen and carbon are considered to be the only limiting elements. Bacterial growth follows the Liebig's law of minimum and the decomposition rate of organic carbon is proportional to the bacterial growth rate. Their main objective was to provide a quantitative explanation to the reported studies that indicate that the rate or the extent of organic matter decomposition often increases in the presence of grazers. For simplicity, the authors did not include the replenishment of organic matter by dead microbial residues in their model. In addition, they did not determine the criteria for the uniform persistence and extinction of the species and chemicals.

Here, a chemostat version of the model in [84] is constructed and analyzed. A chemostat is an apparatus used to mimic the continuous culture of microorganisms in a laboratory [75]. One could think of it in terms of having three bottles: a supply bottle, culture bottle, and collecting bottle. Bacteria and grazers are considered to be inoculated into the culture bottle which is well stirred. Fresh medium with resources is continuously added from the supply bottle to the culture bottle. Culture liquid containing resources, bacteria, and grazers are continuously removed at the same rate to keep the culture volume constant. This set-up helps us to provide a good idealization of a natural aquatic ecosystem. The input and removal of resources mimic the continuous turnover of resources in a natural environment. The washout of bacteria and grazers analogs non-age specific death and emigration which always occurs in nature [81]. These natural processes can not be captured with a microcosm. The chemostat is employed in the industry for the cost friendly production of microbes [66]. Moreover, it is used in the industry to simulate biological waste decomposition and water purification by microbes [86]. It is important in ecological studies because it helps in generating data that mimic the natural environment. Its mathematics is easy to handle as well [81]. It is argued in [81] that having an on site analysis of an ecosystem is pretty much impossible due to the absence of appropriate controls; hence, we have no choice but to use chemostat models. Chemostat models thus yield more realistic ecological models and interesting mathematical problems. In addition to the model in [84] we allow the replenishment of organic matter from dead microbial residues.

The rest of the chapter is organized as follows. In Section 3.2, we formulate the model. In Section 3.3, we prove the positivity and boundedness of

the formulated system. We determine criteria for the uniform persistence and extinction of the species. We end the section by determining the optimal value of grazers that maximizes the per capita decomposition rate of organic carbon. In Section 3.4, we perform one and two-parameter bifurcation analyses. Furthermore, we numerically compare the effect of grazers on degradation in a microcosm and in a chemostat. Moreover, in this section, we determine the switching time of bacterial growth rate from carbon dependent to nitrogen dependent for different continuous cultures. We end the section by discussing the sensitivity of the degradation rate to model parameters. We end the chapter with a discussion section.

## 3.2 The model

Our model is made up of four coupled nonlinear differential equations, which track the rates of change of four state variables, i.e. the concentration of bacterial biomass measured in carbon ( $B$ ), concentration of carbon in media ( $C$ ), concentration of nitrogen in media ( $N$ ) and that of grazers ( $G$ ) in the culture bottle. It describes the bacteria-grazers interaction in a carbon or nitrogen limiting environment. Microbial biomass in the ecosystem is often estimated by measuring the microbial volume per cell (biovolume) and then converting it to microbial biomass in terms of carbon content using an appropriate conversion factor [55]. Thus we assume that microbial biomass is already in terms of carbon content.

The change in bacterial biomass can result from four processes: growth,

dilution, grazing, and natural mortality. That is

$$[\text{change in bacteria biomass}] = [\text{growth}] - [\text{washout}] - [\text{grazing}] - [\text{death}].$$

The functions  $f(N)$  and  $g(C)$  are the specific bacterial growth rates as a function  $N$  or  $C$  when that resource is limiting. Using the threshold approach, the gross bacterial growth rate is given by  $\min\{f(N), g(C)\}$ . This approach follows directly from Leibig's law of minimum, which states that an organism's growth is limited by the most limiting resource [82]. Fresh medium with constant resource concentrations is continuously added from the feed bottle to the culture bottle at a constant input rate  $D$ . Culture liquid containing resources, bacteria and grazers are continuously removed at the same rate. The bacterial death rate (per capita)  $\varepsilon$  is assumed to be density independent. The per capita grazing efficiency  $h(B)$  and the functions  $f(N)$  and  $g(C)$  are monotonically increasing, saturating functions, which take the value of zero when their arguments are zero. These assumptions lead to the following equation:

$$\frac{dB}{dt} = \underbrace{-DB}_{\text{washout}} + \underbrace{\mu_B B \min\{f(N), g(C)\}}_{\text{growth}} - \underbrace{\mu_G h(B)G}_{\text{grazing}} - \underbrace{\varepsilon B}_{\text{death}} \quad (3.1)$$

The change in the grazers' biomass can be caused by grazing, dilution and death. That is

$$[\text{change in grazers biomass}] = [\text{grazing}] - [\text{dilution}] - [\text{death}].$$

The death rate of grazers (per capita),  $\gamma$  is assumed to be density independent.

Thus we have that

$$\frac{dG}{dt} = - \underbrace{DG}_{\text{washout}} + \underbrace{\alpha\mu_G h(B)G}_{\text{grazing}} - \underbrace{\gamma G}_{\text{death}} \quad (3.2)$$

The concentration of carbon in the culture bottle could be increased by carbon supply from the feed bottle which has a fixed concentration of carbon  $C_{in}$  or carbon recycled from dead bacteria and grazers. I.e.

$$\begin{aligned} \text{[change in organic carbon in the environment]} &= \text{[dilution]} - \text{[decomposition]} \\ &\quad + \text{[recycling of death microbes]}. \end{aligned}$$

The bacterial growth rate is proportional to the rate at which it is decomposing organic substances i.e,

$$\text{[bacteria growth rate]} \propto \text{[its decomposition rate]}.$$

This implies that

$$\text{[bacteria growth rate]} = r \text{[its decomposition rate]}$$

where  $r$  is the proportionality constant. Hence,

$$\text{[decomposition rate of organic substances]} = \frac{1}{r} \text{[bacteria growth rate due to its uptake]}.$$

I.e.,

$$[\text{decomposition of organic substances}] = \frac{1}{r} \min\{f(N), g(C)\}.$$

This constant of proportionality reflects the conversion of organic substances to microbes biomass. Thus we refer to it as the “yield” constant. It can be determined by measuring

$$\frac{\text{mass of biomass formed}}{\text{mass of organic substance consumed}}.$$

These decomposed organic substances are converted into gasses and the rest are taken in by the bacteria for growth. The death microbial are assumed to immediately contribute to the organic carbon pool. In reality there is always a delay and the conversion is usually partial. This partial conversion could be accounted for by the formation of recalcitrant biomass decay product. These lead to the following dynamics for the organic carbon compartment

$$\frac{dC}{dt} = \underbrace{D(C_{in} - C)}_{\text{dilution}} - \underbrace{\frac{1}{r} \mu_B B \min\{f(N), g(C)\}}_{\text{decomposition}} + \underbrace{\gamma G + \varepsilon B}_{\text{recycling of death microbes}} \quad (3.3)$$

We further assume that the change in available nitrogen in the culture bottle could be due to the nitrogen excreted by grazers, nitrogen recycled from dead grazers and bacteria, uptake by bacteria, washout or as a result of

supply from feed bottle which has a constant concentration  $N_{in}$ . I.e.

$$\begin{aligned} & [\text{Change in environmental nitrogen concentration}] = \\ & [\text{dilution}] + [\text{nitrogen excreted by grazers}] + [\text{recycling of death microbes}] \\ & - [\text{nitrogen required for bacteria growth}]. \end{aligned}$$

Empirical evidence indicates that bacteria are nutrient richer than their grazers ( $\theta_B > \theta_G$ ) [82, 57]. Thus due to the stoichiometric mismatch, when grazers consumes bacteria, they exude the extra acquired nutrients back into the ecosystem [56, 53, 64, 82]. Hence if upon grazing, grazers obtained  $\theta_B \mu_G h(B) G mg N / dm^3$  from bacteria, and requires just  $\alpha \theta_G \mu_G h(B) G mg N / dm^3$  for their growth, then the remaining  $(\theta_B \mu_G h(B) G - \alpha \theta_G \mu_G h(B) G) mg N / dm^3$  concentration will be exuded back into the environment. Since the growth of bacteria is  $\mu_B B \min\{f(N), g(C)\}$ , the total amount of nitrogen required by bacteria for this growth is

$$\theta_B \mu_B B \min\{f(N), g(C)\}.$$

Hence the dynamics of nitrogen compartment is given as follows:

$$\begin{aligned} \frac{dN}{dt} = & \underbrace{D(N_{in} - N)}_{\text{dilution}} + \underbrace{\theta_G \gamma G + \theta_B \varepsilon B}_{\text{recycling of death microbes}} + \underbrace{(\theta_B - \theta_G \alpha) \mu_G h(B) G}_{\text{nitrogen excreted by grazers}} \quad (3.4) \\ & - \underbrace{\theta_B \mu_B B \min\{f(N), g(C)\}}_{\text{uptake by bacteria}} \end{aligned}$$

Equations (3.1, 3.2, 3.3, 3.4) lead to the following system of equations:

Parameter	Definition	Value	Unit	References
$\mu_B$	Maximum growth rate of bacteria	0.5	$\text{h}^{-1}$	[70]
$K_h$	H.S.C. for grazing	1	$\text{mgC}/\text{dm}^3$	[70]
$K_f$	Nitrogen-dependent H.S.C. for bacterial growth	1.21	$\text{mgN}/\text{dm}^3$	-
$K_g$	Carbon-dependent H.S.C. for bacterial growth	8	$\text{mgC}/\text{dm}^3$	[70]
$\mu_G$	Maximum grazing rate	0.25	$\text{h}^{-1}$	[70]
$\varepsilon$	Bacteria death rate	0.025	$\text{h}^{-1}$	[70]
$\gamma$	Grazers death rate	0.0075	$\text{h}^{-1}$	[70]
$r$	Yield constant	0.31-0.75	-	[63, 85]
$\theta_B$	N:C of bacteria	0.11 – 0.25	-	[82]
$\theta_G$	N:C of grazers	$< \theta_B$	-	[57]
$D$	Dilution rate	-	$\text{h}^{-1}$	-
$C_{in}$	Concentration of carbon in feed bottle	-	$\text{mgC}/\text{dm}^3$	-
$N_{in}$	Concentration of nitrogen in feed bottle	-	$\text{mgN}/\text{dm}^3$	-
$\alpha$	Conversion efficiency of bacteria to grazers	0-1	-	[84]

**Table 3.1:** Definition and values of parameters in System (4.1)

$$\begin{aligned}
\frac{dB}{dt} &= -DB + \mu_B B \min\{f(N), g(C)\} - \mu_G h(B)G - \varepsilon B, \\
\frac{dG}{dt} &= -DG + \alpha \mu_G h(B)G - \gamma G, \\
\frac{dC}{dt} &= D(C_{in} - C) - \frac{1}{r} \mu_B B \min\{f(N), g(C)\} + \gamma G + \varepsilon B, \\
\frac{dN}{dt} &= D(N_{in} - N) + \theta_G \gamma G + \theta_B \varepsilon B + (\theta_B - \theta_G \alpha) \mu_G h(B)G - \theta_B \mu_B B \min\{f(N), g(C)\}.
\end{aligned} \tag{3.5}$$

In numerical simulations, we use Monod or Michaelis-Menten form for the functions  $f, g, h$ :  $f(N) = \frac{N}{N+K_f}$ ,  $g(C) = \frac{C}{C+K_g}$  and  $h(B) = \frac{B}{B+K_h}$ . All parameters and their definitions, units and values are listed in Table 3.1. H.S.C. represents half-saturation constant. The value of  $K_f$  was estimated from that of  $K_g$  using the Redfield ratio of  $C_{106} : N_{16}$  [82].

### 3.3 Model analysis

Here we present mathematical analysis for the model. We begin by verifying the boundedness and positivity of the solutions, and we then compute the

equilibrium points and determine criteria for their stability. Next, we determine conditions for persistence and extinction of bacteria and grazers. We end the section by determining the optimal grazers concentration that maximizes the degradation rate.

### 3.3.1 Positivity and boundedness

**Theorem 1.** *System (4.1) is point dissipative, i.e. there exists a compact set in  $R_+^4$  that attracts all solutions of (4.1), as  $t \rightarrow \infty$ .*

*Proof. Positivity:*

$$\text{Let } x = \begin{pmatrix} x_1 \\ x_2 \\ x_3 \\ x_4 \end{pmatrix} = \begin{pmatrix} B(t) \\ G(t) \\ C(t) \\ N(t) \end{pmatrix} \text{ and}$$

$$f(t, x) = \begin{pmatrix} f_1(t, x) \\ f_2(t, x) \\ f_3(t, x) \\ f_4(t, x) \end{pmatrix} = \begin{pmatrix} -DB + \mu_B B \min\{f(N), g(C)\} - \mu_G h(B)G - \varepsilon B \\ -DG + \alpha \mu_G h(B)G - \gamma G \\ D(C_{in} - C) - \frac{1}{r} \mu_B B \min\{f(N), g(C)\} + \varepsilon B + \gamma G \\ D(N_{in} - N) + (\theta_B - \theta_G \alpha) \mu_G h(B)G - \theta_B \mu_B B \min\{f(N), g(C)\} + \theta_G \gamma G + \theta_B \varepsilon B \end{pmatrix}$$

Let  $i=1,2,3,4$ . It is easy to see that  $f(t,x)$  exist and is continuous, and that if  $x_i = 0 \implies f_i(t, x) \geq 0 \forall i, t$ . Hence nonnegative initial data can only give rise to nonnegative solutions.

**Eventual boundedness:**

Let  $y(t) = \left[ \left( \frac{1}{r} + \theta_B \right) B + \left( \theta_G + \frac{1}{\alpha r} \right) G + C + N \right]$ . Observe that

$$\begin{aligned} y'(t) &= \left[ \left( \frac{1}{r} + \theta_B \right) B + \left( \theta_G + \frac{1}{\alpha r} \right) G + C + N \right]' \\ &= D \left[ C_{in} + N_{in} - \left( \left( \frac{1}{r} + \theta_B \right) B + \left( \theta_G + \frac{1}{\alpha r} \right) G + C + N \right) \right] \\ &\quad + \varepsilon B \left( 1 - \frac{1}{r} \right) + \gamma G \left( 1 - \frac{1}{\alpha r} \right) \\ &\leq D \left[ C_{in} + N_{in} - \left( \left( \frac{1}{r} + \theta_B \right) B + \left( \theta_G + \frac{1}{\alpha r} \right) G + C + N \right) \right] \\ &= D[C_{in} + N_{in} - y(t)], \end{aligned}$$

which implies that

$$y(t) \leq y(0)e^{-Dt} + (C_{in} + N_{in})(1 - e^{-Dt}) \quad (3.6)$$

and hence

$$\limsup_{t \rightarrow \infty} y(t) \leq C_{in} + N_{in}.$$

Thus the solution is eventually bounded. □

The Corollary below follows from Theorem 7

**Corollary 1.** *The set*

$$\Delta = \left\{ (B, G, C, N) : B, G, C, N, \geq 0, \frac{B}{r} + \frac{G}{\alpha r} + C \leq C_{in} \text{ and } \theta_B B + \theta_G G + N \leq N_{in} \right\} \quad (3.7)$$

is compact and positively invariant for System (4.1).

### 3.3.2 Possible equilibrium points

In this section,  $\lambda_C, \lambda_N, \delta$  are used to denote the break even concentrations for carbon, nitrogen and bacteria respectively. They denote the concentrations of carbon, nitrogen and bacteria for which the following relationship holds:  $\mu_B f(\lambda_N) = D + \varepsilon, \mu_B g(\lambda_C) = D + \varepsilon, \alpha \mu_G h(\delta) = D + \gamma$ . Let  $(B^*, G^*, C^*, N^*)$  be a generic equilibrium point. This implies that

$$\begin{aligned} -DB^* + \mu_B B^* \min\{f(N^*), g(C^*)\} - \mu_G h(B^*)G^* - \varepsilon B^* &= 0 & (3.8) \\ -DG^* + \alpha \mu_G h(B^*)G^* - \gamma G^* &= 0 \\ D(C_{in} - C^*) - \frac{1}{r} \mu_B B^* \min\{f(N^*), g(C^*)\} + \varepsilon B^* + \gamma G^* &= 0 \\ D(N_{in} - N^*) + \theta_G \gamma G^* + \theta_B \varepsilon B^* + (\theta_B - \theta_G \alpha) \mu_G h(B^*)G^* - \theta_B \mu_B B^* \min\{f(N^*), g(C^*)\} &= 0 \end{aligned}$$

a) If  $B_1^* = 0$  and  $G_1^* = 0$ , then we have the equilibrium point

$$E_1 = (0, 0, C_{in}, N_{in}). \text{ This equilibrium point always exists.}$$

b) Suppose  $G_2^* = 0$  and  $C_2^* = \lambda_C$ , then the third equation of System (4.2)

implies

$$D(C_{in} - \lambda_C) - \frac{1}{r}B_2^*(D + \varepsilon) + \varepsilon B_2^* = 0$$

$$\implies B_2^* = \frac{rD(C_{in} - \lambda_C)}{D + \varepsilon(1 - r)}. \text{ From the last equation of System (4.2)}$$

$$\text{we have that } N_2^* = N_{in} - \theta_B B_2^* = N_{in} - \theta_B \left( \frac{rD(C_{in} - \lambda_C)}{D + \varepsilon(1 - r)} \right).$$

Thus we have the equilibrium point

$$E_2 = \left( \frac{rD(C_{in} - \lambda_C)}{D + \varepsilon(1 - r)}, 0, \lambda_C, N_{in} - \theta_B \left( \frac{rD(C_{in} - \lambda_C)}{D + \varepsilon(1 - r)} \right) \right),$$

$$\text{for } \lambda_C < C_{in} \text{ and } T_1 := \frac{N_{in}}{C_{in} - \lambda_C} > \frac{D\theta_B r}{D + \varepsilon(1 - r)} := S.$$

- c) If  $G_3^* = 0$  and  $N_3^* = \lambda_N$ , then we have from the last equation of System (4.2) that

$$D(N_{in} - \lambda_N) - \theta_B B_3^* D = 0.$$

$$\implies B_3^* = \frac{N_{in} - \lambda_N}{\theta_B}.$$

From the third equation of System (4.2) we have that

$$C_3^* = \frac{1}{D} \left[ DC_{in} - \frac{N_{in} - \lambda_N}{r\theta_B} (D + \varepsilon(1 - r)) \right].$$

Hence we have the equilibrium point

$$E_3 = \left( \frac{N_{in} - \lambda_N}{\theta_B}, 0, \frac{1}{D} \left[ DC_{in} - \frac{N_{in} - \lambda_N}{r\theta_B} (D + \varepsilon(1 - r)) \right], \lambda_N \right),$$

for  $\lambda_N < N_{in}$  and  $\frac{(N_{in}-\lambda_N)(D+\varepsilon(1-r))}{DC_{in}} < r\theta_B$ .

**Remark 1.** *It is worth mentioning that if  $\lambda_N \neq f^{-1}(g(\lambda_C))$ , the grazer-extinction equilibria above will never coexist. If  $\lambda_N = f^{-1}(g(\lambda_C))$ , and  $N_{in} \neq \frac{\theta_B r D (C_{in} - \lambda_C)}{D + \varepsilon(1 - r)} + f^{-1}(g(\lambda_C))$  then System (4.1) will have two coexisting grazer-extinction equilibria given by  $\left[ \frac{rD(C_{in} - \lambda_C)}{D + \varepsilon(1 - r)}, 0, \lambda_C, \lambda_N \right]$  and  $\left[ \frac{(N_{in} - \lambda_N)}{\theta_B}, 0, \lambda_C, \lambda_N \right]$ .*

- d) Suppose  $g(C_4^*) < f(N_4^*)$  and  $B_4^* = \delta$ . From the first equation of System (4.2) we have that

$$G_4^* = \frac{\delta(-(D + \varepsilon) + \mu_B g(C_4^*))}{\mu_G h(\delta)} = \frac{\delta \alpha (-(D + \varepsilon) + \mu_B g(C_4^*))}{\gamma + D}. \quad (3.9)$$

The third equation gives

$$\begin{aligned} D(C_{in} - C_4^*) &= \frac{\mu_B \delta g(C_4^*)}{r} - \varepsilon \delta - \gamma G_4^*. \\ \implies DC_4^* &= g(C_4^*) \left[ \frac{\gamma \alpha \delta \mu_B}{\gamma + D} - \frac{\mu_B \delta}{r} \right] + DC_{in} - \frac{\gamma \delta \alpha (D + \varepsilon)}{\gamma + D} + \varepsilon \delta. \end{aligned} \quad (3.10)$$

Equation (3.10) can have at most one solution for  $C_4^*$ . If  $\frac{\gamma \alpha \delta \mu_B}{\gamma + D} - \frac{\mu_B \delta}{r} < 0$  and  $DC_{in} - \frac{\gamma \delta \alpha (D + \varepsilon)}{\gamma + D} + \varepsilon \delta > 0$  or  $\frac{\gamma \alpha \delta \mu_B}{\gamma + D} - \frac{\mu_B \delta}{r} > 0$ , Equation (3.10) will have one unique solution for  $C_4^*$ ; else it will have no solution. Suppose that conditions for Equation (3.10) to have a unique solution hold, then by solving for  $C^*$  we can substitute its expression into Equation (3.9) to

obtain a unique value for  $G_4^*$ . From the last equation, we have

$$N_4^* = \frac{1}{D}(DN_{in} + \theta_G \gamma G_4^* + \theta_B \varepsilon \delta + (\theta_B - \theta_G \alpha) \mu_G h(\delta) G_4^* - \theta_B \mu_B \delta g(C_4^*)).$$

i.e.

$$N_4^* = \frac{1}{D}(DN_{in} + \theta_G \gamma G_4^* + \theta_B \varepsilon \delta + (\theta_B - \theta_G \alpha) \frac{\gamma + D}{\alpha} G_4^* - \theta_B \mu_B \delta g(C_4^*)).$$

To solve this equation for  $N_4^*$ , one has to solve Equations (3.9) and (3.10) for  $C_4^*$  and  $G_4^*$ . In case there exist a unique solution for  $C_4^*$ , then there will exist a unique solution for  $G_4^*$  and  $N_4^*$ . Hence we will have an internal equilibrium point  $E_4 = (\delta, G_4^*, C_4^*, N_4^*)$  where  $\delta$  is the unique solution to  $\alpha \mu_G h(\delta) = \gamma + D$ . In summary, this equilibrium point only exists if all of the following conditions hold:

i)

$$44 \frac{\gamma \alpha \delta \mu_B}{\gamma + D} - \frac{\mu_B \delta}{r} < 0 \text{ and } DC_{in} - \frac{\gamma \delta \alpha (D + \varepsilon)}{\gamma + D} + \varepsilon \delta > 0$$

$$\text{or } \frac{\gamma \alpha \delta \mu_B}{\gamma + D} - \frac{\mu_B \delta}{r} > 0$$

ii)  $\mu_B g(C_4^*) > D + \varepsilon$

iii)  $\theta_B \mu_B \delta g(C_4^*) < DN_{in} + \theta_G \gamma G_4^* + \theta_B \varepsilon \delta + (\theta_B - \theta_G \alpha) \frac{\gamma + D}{\alpha} G_4^* := N_0$

where  $N_0$  is the total nitrogen gained.

e) Suppose  $f(N_5^*) < g(C_5^*)$  and  $B_5^* = \delta$ .

From the second equation of System (4.2), we have that

$\alpha \mu_G h(B_5^*) = \gamma + D$ . That is,  $\alpha \mu_G h(\delta) = \gamma + D$ . The first equation of

(4.2) gives

$$G_5^* = \frac{\delta}{\mu_G h(\delta)} (-(D + \varepsilon) + \mu_B f(N_5^*)). \quad (3.11)$$

Substituting Equation (3.11) into the last equation of System (4.2) leads to

$$\begin{aligned} DN_5^* = DN_{in} - \frac{\theta_G \gamma (D + \varepsilon) \delta}{\mu_G h(\delta)} + \theta_G \alpha \delta [D + \varepsilon] - \theta_B D \delta \\ + \left[ \frac{\theta_G \gamma \mu_B \delta}{\mu_G h(\delta)} - \theta_G \alpha \mu_B \delta \right] f(N_5^*) \end{aligned} \quad (3.12)$$

This equation has a solution for  $N_5^*$  only if

$$DN_{in} - \frac{\theta_G \gamma (D + \varepsilon) \delta}{\mu_G h(\delta)} + \theta_G \alpha \delta [D + \varepsilon] - \theta_B D \delta > 0$$

and

$$\left[ \frac{\theta_G \gamma \mu_B \delta}{\mu_G h(\delta)} - \theta_G \alpha \mu_B \delta \right] < 0 \text{ or } \left[ \frac{\theta_G \gamma \mu_B \delta}{\mu_G h(\delta)} - \theta_G \alpha \mu_B \delta \right] > 0.$$

From the second equation of System (4.2), we have that

$$C_5^* = \frac{1}{D} \left( DC_{in} - \frac{\mu_B \delta f(N_5^*)}{r} + \varepsilon \delta + \gamma \frac{\delta}{\mu_G h(\delta)} (-(D + \varepsilon) + \mu_B f(N_5^*)) \right). \quad (3.13)$$

Thus, to obtain the values of the internal equilibrium point (assuming it exists), one needs to first solve Equation (3.12) for  $N_5^*$ , then substitute the unique value of  $N_5^*$  into Equations (3.11) and (3.13) which

will give a unique  $G_5^*$  and  $C_5^*$ . Thus we have the equilibrium point  $E_5 = (\delta, G_5^*, C_5^*, N_5^*)$  if the following conditions hold:

i)

$$DN_{in} - \frac{\theta_G \gamma (D + \varepsilon) \delta}{\mu_G h(\delta)} + \theta_G \alpha \delta [D + \varepsilon] - \theta_B D \delta > 0$$

and

$$\left[ \frac{\theta_G \gamma \mu_B \delta}{\mu_G h(\delta)} - \theta_G \alpha \mu_B \delta \right] < 0 \text{ or } \left[ \frac{\theta_G \gamma \mu_B \delta}{\mu_G h(\delta)} - \theta_G \alpha \mu_B \delta \right] > 0.$$

ii)  $\mu_B g(N_5^*) > D + \varepsilon$

ii)  $\frac{\mu_B \delta f(N_5^*)}{r} < DC_{in} + \varepsilon \delta.$

**Remark 2.** *If at any of the coexistence equilibria, the concentration of nitrogen  $N^*$  and carbon  $C^*$  are such that  $N^* = f^{-1}(g(C^*))$ , then one can find a parameter set for which two internal equilibria can coexist.*

Table (3.2) contains a summary of all possible critical points and Table (3.3) contains conditions for their existence.

Symbol	Expression
$E_1$	$(0, 0, C_{in}, N_{in})$
$E_2$	$(\frac{rD(C_{in}-\lambda_C)}{D+\varepsilon(1-r)}, 0, \lambda_C, N_{in} - \theta_B(\frac{rD(C_{in}-\lambda_C)}{D+\varepsilon(1-r)}))$
$E_3$	$E_3 = \left( \frac{N_{in}-\lambda_N}{\theta_B}, 0, \frac{1}{D} \left[ DC_{in} - \frac{N_{in}-\lambda_N}{r\theta_B}(D + \varepsilon(1-r)) \right], \lambda_N \right)$
$E_4$	$(\delta, G_4^*, C_4^*, N_4^*)$ . Where $\delta$ solves $\alpha\mu_G h(\delta) = \gamma + D$ , $C_4^*$ solves $DC_4^* = g(C_4^*) \left[ \frac{\gamma\alpha\delta\mu_B}{\gamma+D} - \frac{\mu_B\delta}{r} \right] + DC_{in} - \frac{\gamma\delta\alpha(D+\varepsilon)}{\gamma+D} + \varepsilon\delta$ $G_4^* = \frac{\delta(-(D+\varepsilon)+\mu_B g(C_4^*))}{\mu_G h(\delta)} = \frac{\delta\alpha(-(D+\varepsilon)+\mu_B g(C_4^*))}{\gamma+D}$ $N_4^* = \frac{1}{D}(DN_{in} + \theta_G\gamma G_4^* + \theta_B\varepsilon\delta + (\theta_B - \theta_G\alpha)\frac{\gamma+D}{\alpha}G_4^* - \theta_B\mu_B\delta g(C_4^*))$
$E_5$	$(\delta, G_5^*, C_5^*, N_5^*)$ . Where $\delta$ solves $\alpha\mu_G h(\delta) = \gamma + D$ $N_5^*$ solves $DN_5^* = DN_{in} - \frac{\theta_G\gamma(D+\varepsilon)\delta}{\mu_G h(\delta)} + \theta_G\alpha\delta[D+\varepsilon] - \theta_B D\delta + \left[ \frac{\theta_G\gamma\mu_B\delta}{\mu_G h(\delta)} - \theta_G\alpha\mu_B\delta \right] f(N_5^*)$ $G_5^* = \frac{\delta}{\mu_G h(\delta)} (-(D+\varepsilon) + \mu_B f(N_5^*))$ . and $C_5^* = \frac{1}{D} \left( DC_{in} - \frac{\mu_B\delta f(N_5^*)}{r} + \varepsilon\delta + \gamma\frac{\delta}{\mu_G h(\delta)} (-(D+\varepsilon) + \mu_B f(N_5^*)) \right)$ .

**Table 3.2:** Notations and expressions for the equilibrium points of System (4.1).

Equilibrium point	Conditions for existence
$E_1$	no conditions are needed
$E_2$	$\lambda_C < C_{in}$ and $N_{in} > \frac{\theta_B r D (C_{in} - \lambda_C)}{D + \varepsilon(1-r)}$ .
$E_3$	$\lambda_N < N_{in}$ and $DC_{in} > \frac{(N_{in} - \lambda)(D + \varepsilon(1-r))}{r\theta_B}$
$E_4$	(i) $\frac{\gamma\alpha\delta\mu_B}{\gamma+D} - \frac{\mu_B\delta}{r} < 0$ and $DC_{in} - \frac{\gamma\delta\alpha(D+\varepsilon)}{\gamma+D} + \varepsilon\delta > 0$ or $\frac{\gamma\alpha\delta\mu_B}{\gamma+D} - \frac{\mu_B\delta}{r} > 0$ (ii) $\mu_B g(C_4^*) > D + \varepsilon$ (iii) $\theta_B\mu_B\delta g(C_4^*) < DN_{in} + \theta_G\gamma G_4^* + \theta_B\varepsilon\delta + (\theta_B - \theta_G\alpha)\frac{\gamma+D}{\alpha}G_4^*$
$E_5$	(i) $DN_{in} - \frac{\theta_G\gamma(D+\varepsilon)\delta}{\mu_G h(\delta)} + \theta_G\alpha\delta[D+\varepsilon] - \theta_B D\delta > 0$ and $\left[ \frac{\theta_G\gamma\mu_B\delta}{\mu_G h(\delta)} - \theta_G\alpha\mu_B\delta \right] < 0$ or $\left[ \frac{\theta_G\gamma\mu_B\delta}{\mu_G h(\delta)} - \theta_G\alpha\mu_B\delta \right] > 0$ . (ii) $\mu_B g(N_5^*) > D + \varepsilon$ (iii) $DC_{in} + \varepsilon\delta > \frac{\gamma\delta}{\mu_G h(\delta)}(D + \varepsilon) - \left[ \frac{\gamma}{\mu_G h(\delta)} - \frac{1}{r} \right] \delta\mu_B f(N_5^*)$

**Table 3.3:** Criteria for existence of the equilibrium points of System (4.1).

### 3.3.3 Stability of the equilibrium points

The objective of this section is to determine the stability of all equilibrium points derived above. In deriving the equilibria, we assume that

$$f(N^*) \neq g(C^*).$$

**Stability of  $E_1$  :**

The Jacobian matrix corresponding to this equilibrium point is

$$J(E_1) = \begin{bmatrix} -(D + \varepsilon) + \mu_B \min f(N_{in}), g(C_{in}) & 0 & 0 & 0 \\ 0 & -D - \gamma & 0 & 0 \\ \frac{-\mu_B \min(f(N_{in}), g(C_{in}))}{r} + \varepsilon & \gamma & -D & 0 \\ -\theta_B \mu_B \min f(N_{in}, g(C_{in})) + \theta_B \varepsilon & \theta_G \gamma & 0 & -D \end{bmatrix}.$$

The corresponding eigenvalues are  $\lambda_1 = -D$ ,  $\lambda_2 = -D$ ,  $\lambda_3 = -D - \gamma$ ,  $\lambda_4 = -(D + \varepsilon) + \mu_B \min(f(N_{in}, g(C_{in})))$ . Thus extinction equilibrium  $E_1$  is locally asymptotically stable if

$$\mu_B \min(f(N_{in}), g(C_{in})) < D + \varepsilon.$$

**Biological meaning:** If the resources in the culture bottles are extremely limiting, bacteria will go extinct, which pushes the grazers to go extinct as well.

### Stability of $E_2$ :

Corresponding to  $E_2$  we have the Jacobian matrix

$$J(E_2) = \begin{bmatrix} 0 & -\mu_G h(B_2^*) & \mu_B B_2^* g'(\lambda_C) & 0 \\ 0 & -D + \alpha \mu_G h(B_2^*) - \gamma & 0 & 0 \\ \frac{-(D+\varepsilon)}{r} + \varepsilon & \gamma & -D - \frac{\mu_B B_2^* g'(\lambda_C)}{r} & 0 \\ -\theta_B D & \theta_G \gamma + (\theta_B - \theta_G \alpha) \mu_G h(B_2^*) & -\theta_B \mu_B B_2^* g'(\lambda_C) & -D \end{bmatrix}.$$

The corresponding eigenvalues are

$$\lambda_1 = -D$$

$$\lambda_2 = \frac{-(Dr + \mu_B B_2^* g'(\lambda_C)) + \sqrt{(Dr + \mu_B B_2^* g'(\lambda_C))^2 - 4r(D\mu_B B_2^* g'(\lambda_C) + \mu_B g'(\lambda_C)\varepsilon)}}{2r}$$

$$\lambda_3 = \frac{-(Dr + \mu_B B_2^* g'(\lambda_C)) - \sqrt{(Dr + \mu_B B_2^* g'(\lambda_C))^2 - 4r(D\mu_B B_2^* g'(\lambda_C) + \mu_B g'(\lambda_C)\varepsilon)}}{2r}$$

$$\lambda_4 = -D + \alpha \mu_G h(B_2^*) - \gamma.$$

Thus  $E_2$  is locally asymptotically stable if  $\alpha \mu_G h(B_2^*) < D + \gamma$ .

**Biological meaning:** This grazer-extinction equilibrium point is a sink if the growth rate of grazers is less than its loss rate at this equilibrium.

### Stability of $E_3$ :

The Jacobian matrix for this equilibrium point is

$$J(E_3) = \begin{bmatrix} 0 & -\mu_G h(B_3^*) & 0 & \mu_B B_3^* f'(\lambda_N) \\ 0 & -D + \alpha \mu_G h(B_3^*) - \gamma & 0 & 0 \\ \frac{-(D+\varepsilon)}{r} + \varepsilon & \gamma & -D & -\frac{\mu_B B_3^* f'(\lambda_N)}{r} \\ -\theta_B D & \theta_G \gamma + (\theta_B - \theta_G \alpha) \mu_G h(B_3^*) & 0 & -D - \theta_B \mu_B B_3^* f'(\lambda_N) \end{bmatrix}.$$

The corresponding eigenvalues are  $\lambda_1 = -D$ ,  $\lambda_2 = -D$ ,  $\lambda_3 = -B^* \theta_B \mu_B f'(\lambda_N)$ ,  $\lambda_4 = -D + \alpha \mu_G h(B^*) - \gamma$ . Thus this alternative grazer-extinction  $E_3$  is stable if  $\alpha \mu_G h(B^*) < D + \gamma$ .

**Biological meaning:** This grazer-extinction equilibrium point is a sink if the growth rate of grazers is less than its loss rate at this equilibrium.

**Stability of  $E_4$  :**

$$J(E_4) = \begin{bmatrix} -(D + \varepsilon) + \mu_B g(C_4^*) - \mu_G h'(\delta) G_4^* & -\mu_G h(\delta) & \mu_B \delta g'(C_4^*) & 0 \\ \alpha \mu_G G^* h'(\delta) & 0 & 0 & 0 \\ -\frac{\mu_B g(C_4^*)}{r} + \varepsilon & \gamma & -D - \frac{\mu_B \delta g'(C_4^*)}{r} & 0 \\ \theta_B \varepsilon + (\theta_B - \theta_G \alpha) \mu_G h'(\delta) G_4^* - \mu_B g(C_4^*) \theta_B & \theta_G \gamma + (\theta_B - \theta_G \alpha) \mu_G h(\delta) & -\theta_B \mu_B \delta g'(C_4^*) & -D \end{bmatrix}.$$

**Characteristic Polynomial:** the corresponding characteristic polynomial is

given by

$$\begin{aligned}
& \frac{1}{r} \left[ (x + D) \left( xrD^2 + Dx\epsilon r + 2Dx^2r \right. \right. \\
& + Dx\mu_B\delta g'(C^*) - Dxr\mu_Bg(C^*) + Dxr\mu_Gh'(\delta)G^* + D\alpha\mu_G^2h'(\delta)G^*h(\delta)r \\
& - x^2r\mu_Bg(C^*) + x^2r\mu_Gh'(\delta)G^* + x^3r + x^2\mu_B\delta g'(C^*) + x\mu_B\delta g'(C^*)\mu_Gh'(\delta)G^* \\
& + x\mu_B\delta h'(\delta)\epsilon - x\mu_B\delta g'(C)\epsilon r - \gamma\alpha\mu_Gh'(\delta)G'\mu_B\delta g'(C)r \\
& \left. \left. + x\alpha\mu_G^2h'(\delta)G^*h(\delta)r + \alpha\mu_G^2h'(\delta)G^*h(\delta)\mu_B\delta g'(C^*) \right) \right] = 0.
\end{aligned}$$

This implies that  $x = -D$  or

$$\begin{aligned}
& x^3r + \left( 2Dr + \mu_B\delta g'(C_4^*) - \mu_Bg(C_4^*)r + \mu_Gh'(\delta)G_4^*r \right) x^2 \\
& + \left( -\mu_B\delta g'(C_4^*)\epsilon r + \epsilon rD + \mu_B\delta g'(C_4^*)\epsilon + D^2r - Dr\mu_Bg + Dr\mu_Gh'(\delta)G_4^* \right. \\
& \left. + \mu_B\delta g'(C_4^*)D + \mu_B\delta g'(C_4^*)\mu_Gh'(\delta)G_4^* + \alpha\mu_G^2h'(\delta)G_4^*h(\delta)r \right) x \\
& - \gamma\alpha\mu_Gh'(\delta)G_4^*\mu_B\delta g'r + D\alpha\mu_G^2h'(\delta)G_4^*h(\delta)r + \alpha\mu_G^2h'(\delta)G_4^*h(\delta)\mu_B\delta g'(C_4^*) = 0.
\end{aligned}$$

Let

$$A_3 = r,$$

$$A_2 = 2Dr + \mu_B\delta g'(C_4^*) - \mu_Bg(C_4^*)r + \mu_Gh'(\delta)G_4^*r,$$

$$\begin{aligned}
A_1 = & \left( -\mu_B\delta g'(C_4^*)\epsilon r + \epsilon rD + \mu_B\delta g'(C_4^*)\epsilon + D^2r - Dr\mu_Bg + Dr\mu_Gh'(\delta)G_4^*, \right. \\
& \left. + \mu_B\delta g'(C_4^*)D + \mu_B\delta g'(C_4^*)\mu_Gh'(\delta)G_4^* + \alpha\mu_G^2h'(\delta)G_4^*h(\delta)r, \right)
\end{aligned}$$

$$A_0 = -\gamma\alpha\mu_Gh'(\delta)G_4^*\mu_B\delta g'r + D\alpha\mu_G^2h'(\delta)G_4^*h(\delta)r + \alpha\mu_G^2h'(\delta)G_4^*h(\delta)\mu_B\delta g'(C_4^*).$$

Applying the Routh-Hurwitz stability criterion, we have that the coexistence equilibrium point  $E_4$  is stable if  $A_2$ ,  $A_1$  and  $A_0$  are greater than zero and  $A_2A_1 > rA_0$ .

$$\begin{aligned}
A_1A_2 &= \alpha\mu_G^2h'G_4^*hr\mu_B\delta g' - \mu_B\delta g'\varepsilon r^2\mu_Gh'G_4^* - \alpha\mu_G^2h'G_4^*hr^2\mu_Bg \\
&+ 4Dr\mu_Gh'G_4^*\mu_B\delta g' - \mu_B^2\delta g'\mu_Gh'G_4^*gr + 2\mu_B\delta g'\varepsilon r\mu_Gh'G_4^* + 4\mu_B\delta g'\varepsilon rD \\
&- \mu_B^2\delta g'\varepsilon rg - 2\mu_B\delta g'\varepsilon r^2D + \mu_B^2\delta g'\varepsilon r^2g + 2\varepsilon r^2D\mu_Gh'G_4^* - 2Dr\mu_B^2g\delta g' \\
&+ \mu_B^2\delta^2(g')^2\mu_Gh'G_4^* + \alpha\mu_G^3(h')^2(G_4^*)^2hr^2 - 2\varepsilon r^2D\mu_Bg + Dr^2\mu_G^2(h')^2(G_4^*)^2 \\
&- \mu_B^2\delta^2(g')^2\varepsilon r - \mu_B\delta g'\varepsilon^2r^2 + 3D^2r\mu_B\delta g' + 3D^2r^2\mu_Gh'G_4^* - 3D^2r^2\mu_Bg \\
&+ Dr^2\mu_B^2g^2 + \mu_B^2\delta^2(g')^2D + \mu_B^2\delta^2(g')^2\varepsilon + \mu_B\delta g'\varepsilon^2r + 3D^2r^2\varepsilon + \varepsilon^2r^2D + 2D^3r^2 \\
&- 2Dr^2\mu_Bg\mu_Gh'G_4^* + \mu_B\delta g'\mu_G^2(h')^2(G_4^*)^2r + 2\alpha\mu_G^2h'G_4^*hr^2D + \alpha\mu_G^2h'G_4^*hr^2\varepsilon \\
&= \alpha\mu_G^2h'G_4^*hr\mu_B\delta g' + \alpha\mu_G^2h'G_4^*hr^2D + \mu_B\delta g'\varepsilon r\mu_Gh'G_4^*[2 - r] \\
&+ \alpha\mu_G^2h'G_4^*hr^2[D - \mu_Bg] + r\mu_Gh'G_4^*\mu_B\delta g'[4D - \mu_Bg] + \mu_B\delta g'\varepsilon r[4D - \mu_Bg] \\
&+ Dr\mu_B\delta g'[3D - 2\mu_Bg] + D^2r^2[2D - 3\mu_Bg] + Dr^2\varepsilon[3D - 2\mu_Bg] \\
&+ \mu_B^2\delta^2(g')^2\varepsilon[1 - r] + Dr^2\mu_Gh'G_4^*[3D - 2\mu_Bg] - 2\mu_B\delta g'\varepsilon r^2D \\
&+ \mu_B^2\delta g'\varepsilon r^2g + 2\varepsilon r^2D\mu_Gh'G_4^* + \mu_B^2\delta^2(g')^2\mu_Gh'G_4^* + \alpha\mu_G^3(h')^2(G_4^*)^2hr^2 \\
&+ Dr^2\mu_G^2(h')^2(G_4^*)^2 + Dr^2\mu_B^2g^2 + \mu_B^2\delta^2(g')^2D + \mu_B\delta g'\varepsilon^2[1 - r] \\
&+ \varepsilon^2r^2D + \mu_B\delta g'\mu_G^2(h')^2(G_4^*)^2r + \alpha\mu_G^2h'G_4^*hr^2\varepsilon
\end{aligned}$$

and

$$rA_0 = r[-w\alpha\mu_Gh'(\delta)G_4^*\mu_B\delta g'r + D\alpha\mu_G^2h'(\delta)G_4^*h(\delta)r + \alpha\mu_G^2h'(\delta)G_4^*h(\delta)\mu_B\delta g'(C_4^*)].$$

Observe that for very large values of  $D$ ,  $C_4^* \rightarrow \kappa \leq C_{in}$  and  $g'(C_4^*) \rightarrow 0$ . If  $D$  is chosen such that  $2D \gg \mu_B g(C_4^*)$ ,  $A_2$  and  $A_1$  will be greater than zero and if  $wr < \mu_G h(\delta) = D + \gamma$ ,  $A_0$  will be greater than zero. That is a) if the dilution is greater than the per capita bacterial growth rate  $A_1$  and  $A_2$  will be positive and b) if the per capita grazing rate of bacteria is greater than the bacterial per capita death rate,  $A_0$  will be positive. We also have that if  $\mu_G g(C_4^*) \ll D$   $A_1 A_2$  will be greater than or equal to  $r A_0$ . We could choose  $D$  very large so that  $A_1 A_2 > r A_0$

**Biological meaning:** If the dilution rate is very large,  $E_4$  will be stable.

**Stability of  $E_5$  :**

$$J(E_5) = \begin{bmatrix} -(D + \varepsilon) + \mu_B f(N_5^*) - \mu_G h'(\delta) G_5^* & -\mu_G h(\delta) & 0 & \mu_B \delta f'(N_5^*) \\ \alpha \mu_G h'(\delta) G_5^* & 0 & 0 & 0 \\ -\frac{\mu_B f(N_5^*)}{r} + \varepsilon & w & -D & -\frac{\mu_B \delta f'(N_5^*)}{r} \\ \theta_B \varepsilon + (\theta_B - \theta_G \alpha) \mu_G h'(\delta) G_5^* - \theta_B \mu_B f(N_5^*) & \theta_G w + (\theta_B - \theta_G \alpha) \mu_G h(\delta) & 0 & -D - \theta_B \mu_B \delta f'(N_5^*) \end{bmatrix}$$

**Characteristic polynomial:** the corresponding characteristic polynomial to

this Jacobian matrix is given by

$$\begin{aligned}
& \left( x + D \right) \left( Dx\varepsilon + x^2\varepsilon + xD^2 + Dx\mu_G h'(\delta)G_5^* + D\alpha\mu_G^2 h'(\delta)G_5^* h(\delta) \right. \\
& - Dx\mu_B f(N_5^*) + 2x^2D + Dx\theta_B\mu_B\delta f'(N_5^*) + x^3 + x^2\theta_B\mu_B\delta f'(N_5^*) \\
& - \alpha\mu_G h'(\delta)G_5^*\mu_B\delta f'(N_5^*)\theta_G w - x^2\mu_B f(N_5^*) + x^2\mu_G h'(\delta)G_5^* \\
& + x\alpha\mu_G^2 h'(\delta)G_5^* h(\delta) + x\mu_B\delta f'(N_5^*)\mu_G h'(\delta)G_5^*\theta_G\alpha \\
& \left. + \alpha^2\mu_G^2 h'(\delta)G_5^*\mu_B\delta f'(N_5^*)h(\delta)\theta_G \right) = 0.
\end{aligned}$$

This implies that

$$\begin{aligned}
& \alpha^2\mu_G^2 h'(\delta)G_5^*\mu_B\delta f'(N_5^*)h(\delta)\theta_G + D\alpha\mu_G^2 h'(\delta)G_5^* h(\delta) \\
& - \alpha\mu_G h'(\delta)G_5^*\mu_B\delta f'(N_5^*)\theta_G w + \left( D\mu_G h'(\delta)G_5^* - D\mu_B f(N_5^*) \right. \\
& + \theta_B\mu_B\delta f'(N_5^*)D + \alpha\mu_G^2 h'(\delta)G_5^* h(\delta)\mu_B\delta f'(N_5^*)\mu_G h'(\delta)G_5^*\theta_G\alpha + D^2 \\
& \left. + D\varepsilon \right) x + \left( \mu_G h'(\delta)G_5^* + \theta_B\mu_B\delta f'(N_5^*) + 2D - \mu_B f(N_5^*) + \varepsilon \right) x^2 + x^3 = 0.
\end{aligned}$$

Let

$$B_3 = 1,$$

$$B_2 = \mu_G h'(\delta)G_5^* + \theta_B\mu_B\delta f'(N_5^*) + 2D - \mu_B f(N_5^*) + \varepsilon,$$

$$\begin{aligned}
B_1 &= D\mu_G h'(\delta)G_5^* - D\mu_B f(N_5^*) + \theta_B\mu_B\delta f'(N_5^*)D + \alpha\mu_G^2 h'(\delta)G_5^* h(\delta) \\
&+ \mu_B\delta f'(N_5^*)\mu_G h'(\delta)G_5^*\theta_G\alpha + D^2 + D\varepsilon,
\end{aligned}$$

$$\begin{aligned}
B_0 &= \alpha^2\mu_G^2 h'(\delta)G_5^*\mu_B\delta f'(N_5^*)h(\delta)\theta_G + D\alpha\mu_G^2 h'(\delta)G_5^* h(\delta) \\
&- \alpha\mu_G h'(\delta)G_5^*\mu_B\delta f'(N_5^*)\theta_G w.
\end{aligned}$$

By the Routh-Hurwitz stability criterion, the coexistence equilibrium point  $E_5$  is stable if  $B_2$ ,  $B_1$  and  $B_0$  are greater than zero and  $B_2B_1 > A_0$ .

$$\begin{aligned}
B_2B_1 = & 2\theta_B\mu_B\delta f'D\mu_Gh'G_5^* + 2D\alpha\mu_G^2h'G_5^*h + 3\theta_B\mu_B\delta f'D^2 + 2D\varepsilon\mu_Gh'G_5^* \\
& + D\varepsilon^2 + 3D^2\varepsilon - 2\theta_B\mu_B^2\delta f'Df + 2\theta_B\mu_B\delta f'D\varepsilon - 2D\mu_Bf\mu_Gh'G_5^* \\
& + \theta_B^2\mu_B^2\delta^2(f')^2D + \alpha\mu_G^3(h')^2(G_5^*)^2h + \alpha\mu_G^2h'G_5^*h\varepsilon + D\mu_B^2f^2 \\
& + D\mu_G^2(h')^2(G_5^*)^2 - 2D\mu_Bf\varepsilon + 3D^2\mu_Gh'G_5^* + 2D^3 - 3D^2\mu_Bf \\
& + \mu_B\delta f'\mu_G^2(h')^2G^2\theta_G\alpha - \alpha\mu_G^2h'G_5^*h\mu_Bf + \mu_B\delta f'\mu_Gh'G_5^*\theta_G\alpha\varepsilon \\
& + \alpha\mu_G^2h'G_5^*h\theta_B\mu_B\delta f' - \mu_B^2\delta f'\mu_Gh'G_5^*\theta_G\alpha f + \mu_B^2\delta^2(f')^2\mu_Gh'G_5^*\theta_G\alpha\theta_B \\
& + 2\mu_B\delta f'\mu_Gh'G_5^*\theta_G\alpha D
\end{aligned}$$

$$\begin{aligned}
= & \alpha\mu_G^2h'G_5^*h\theta_B\mu_B\delta f' + D\alpha\mu_G^2h'G_5^*h + \theta_B\mu_B\delta f'D[3D - 2\mu_Bf] \\
& + 3D^2\mu_Gh'G_5^*[3D - 2\mu_Bf] + D\varepsilon[3D - \mu_Bf] + D^2[2D - 3\mu_Bf] \\
& + \alpha\mu_G^2h'G_5^*h[D - \mu_Bf] + \mu_B\delta f'\mu_Gh'G_5^*\theta_G\alpha[2D - \mu_Bf] \\
& + 2\theta_B\mu_B\delta f'D\mu_Gh'G_5^* + 2D\varepsilon\mu_Gh'G_5^* + D\varepsilon^2 + 2\theta_B\mu_B\delta f'D\varepsilon \\
& + \theta_B^2\mu_B^2\delta^2(f')^2D + \alpha\mu_G^3(h')^2(G_5^*)^2h + \alpha\mu_G^2h'G_5^*h\varepsilon + D\mu_B^2f^2 \\
& + D\mu_G^2(h')^2(G_5^*)^2 + \mu_B\delta f'\mu_G^2(h')^2G^2\theta_G\alpha + \mu_B\delta f'\mu_Gh'G_5^*\theta_G\alpha\varepsilon \\
& + \mu_B^2\delta^2(f')^2\mu_Gh'G_5^*\theta_G\alpha\theta_B
\end{aligned}$$

It can be shown that if  $D$  is very large,  $N_5^* \rightarrow \kappa_1 \leq N_{in} + \theta_G\alpha\delta - \theta_B\delta$  and  $f'(N_5^*) \rightarrow 0$ . Thus if  $D \gg \mu_Bf(\kappa_1)$ ,  $B_2$  and  $B_0$  will be greater than zero. Also if  $w < \alpha\mu_Gh(\delta) = D + \gamma$ ,  $B_0$  will be greater than zero. Moreover, if  $D \gg \mu_Bf(\kappa_1)$ ,  $B_2B_1$  will be greater than  $B_0$ .

**Biological meaning:** Suppose the following 2 conditions hold. The first one is that the dilution rate is far greater than the bacterial per capita growth rate. The second is that the natural per capita mortality rate of grazer is less than the per capita grazed bacterial biomass converted to grazers biomass. When these 2 situations hold, the carbon limiting internal equilibrium will be stable.

Table (3.4) contains a summary of the conditions for the asymptotic stability of all the possible equilibrium points.

Equilibrium Point	Conditions for asymptotic stability
$E_1$	$\mu_B \min\{f(N_{in}), g(C_{in})\} < D + \varepsilon.$
$E_2$	$\alpha\mu_G h(B_2^*) < D + w.$
$E_3$	$\alpha\mu_G h(B_3^*) < D + w$
$E_4$	$A_n > 0, n = 0, 1, 2, A_2 A_1 > r A_0$ where $A_n > 0, n = 0, 1, 2$ are as defined above
$E_5$	$B_n > 0, n = 0, 1, 2, B_2 B_1 > B_0$ where $B_n > 0, n = 0, 1, 2$ are as defined above.

**Table 3.4:** Summary of conditions for asymptotic stability of possible equilibrium points of System (4.1).

### 3.3.4 Persistence-extinction criteria

In this section, we obtain sufficient conditions for the uniform persistence as well as extinction of bacteria and grazers and consequently those of System (4.1).  $|\cdot|$  is used to denote a norm on  $\mathbb{R}^9$ . The uniform persistence conditions are important for the food web defined by System (4.1) as they ensure that the bacteria and grazers never go extinct (this is made more precise in Theorem

2 (a) and Theorem 3 (a) below). For the remainder of this section we assume that  $h$  is differentiable at zero. Let  $\lambda_0 = (D, \varepsilon, r, \alpha, \omega, \mu_B, \mu_G, \theta_B, \theta_G) \in \mathbb{R}_+^9$  be a fixed parameter vector and

$$R(\varepsilon) = \mu_B \min\{f(N_{in}), g(C_{in})\} - (D + \varepsilon). \quad (3.14)$$

**Theorem 2** (Bacterial persistence-extinction criteria). *The bacterial persistence-extinction criteria for System (4.1) are as follows:*

(a) *If  $R(\varepsilon) > 0$  then the bacteria are robustly uniformly persistent: there exist  $\eta > 0, \delta > 0$  such that*

$$\liminf_{t \rightarrow \infty} B(t) \geq \eta, \quad (3.15)$$

*for all solutions of (4.1) that satisfy  $B(0) > 0$ , corresponding to all  $\lambda$  satisfying  $|\lambda - \lambda_0| < \delta$ .*

(b) *If  $R(\varepsilon) \leq 0$  then  $(B(t), G(t), C(t), N(t)) \rightarrow E_1$  for all solutions of (4.1).*

*Proof.* (a) Let

$$X_0^B = \{(B, G, C, N) \in \mathbb{R}_+^4 \mid B = 0\}.$$

From Theorem 7, we have that there exists a compact set  $\mathcal{B} \subset \mathbb{R}_+^4$  that attracts all solutions of (4.1). Also, the omega limit set of every solution in  $\mathcal{B} \cap X_0^B$  is the equilibrium point  $E_1 = (0, 0, C_{in}, N_{in})$ . Let

$$\tilde{h}(B) = \begin{cases} \frac{h(B)}{B}, & \text{if } B \neq 0, \\ h'(0), & \text{if } B = 0 \end{cases}$$

Then

$$B' = B(-D - \varepsilon + \mu_B \min\{f(N), g(C)\} - \mu_G \tilde{h}(B)G)$$

Now the conclusion follows from Corollary 4.7 in [77], applied with  $M = \mathcal{B} \cap X_0^B$  and  $T = 1$  (hence  $\Omega(M) = E_1$ , while  $P(1, E_1) = e^{R(\varepsilon)}$ ).

(b) Let  $(B(t), G(t), C(t), N(t))$  be a solution of (4.1) with omega limit set  $\omega$ . Let  $x_0^1 = (B_0^1, G_0^1, C_0^1, N_0^1) \in \omega$  and  $(B^1(t), G^1(t), C^1(t), N^1(t))$  be a solution with  $(B^1(0), G^1(0), C^1(0), N^1(0)) = x_0^1$ . Using Corollary 1 and that  $\omega$  is invariant, we have that  $\omega \neq \emptyset$  and  $C^1(t) \leq C_{in}$ ,  $N^1(t) \leq N_{in}$ , for all  $t \geq 0$ . Thus,

$$B^{1'} \leq B^1(-D - \varepsilon + \mu_B \min\{f(N_{in}), g(C_{in})\}), \forall t \in \mathbb{R}.$$

Since  $R(\varepsilon) \leq 0$ , we have that  $B^{1'}(t) \leq 0$  for all  $t \in \mathbb{R}$ . Suppose  $B_0^1 > 0$ . Then  $B^1(t)$  converges to some  $\bar{B}^1 > 0$ , as  $t \rightarrow -\infty$ . Thus  $\omega(x_0^1) \subset \{(B, G, C, N) \mid B = \bar{B}^1\}$ . Let  $x_0^2 \in \omega(x_0^1)$  and  $(B^2(t), G^2(t), C^2(t), N^2(t))$  be a solution with  $(B^2(0), G^2(0), C^2(0), N^2(0)) = x_0^2$ . If  $G^2(0) > 0$  and  $D + \omega \neq \alpha \mu_G h(\bar{B}^1)$ , it follows that  $G^2(t) \rightarrow \infty$  when either  $t \rightarrow \infty$ , or  $t \rightarrow -\infty$ , which contradicts the boundedness of  $\omega(x_0^1)$ . Hence either  $G^2(0) = 0$ , which implies  $G^2(t) = 0$  for all  $t \in \mathbb{R}$ , or  $D + \omega = \alpha \mu_G h(\bar{B}^1)$ , which implies  $G^2(t) = G^2(0)$  for all  $t \in \mathbb{R}$ . In either case, using the equations for  $C$  and  $D$  in (4.1), we obtain that  $\omega(x_0^1)$  contains an equilibrium point  $(B^*, G^*, C^*, N^*)$  with  $B^* = \bar{B}^1$ . Since  $\bar{B}^1 > 0$ , again, from Corollary 1 we have that  $N^* < N_{in}$  and  $C^* < C_{in}$ . Then

$$B^*(-D - \varepsilon + \mu_B \min\{f(N^*), g(C^*)\}) - \mu_G h(B^*)G^* < B^*(-D - \varepsilon + \mu_B \min\{f(N_{in}), g(C_{in})\}) < 0,$$

which contradicts that  $(B^*, G^*, C^*, N^*)$  is an equilibrium.

Hence  $\omega \subset \{(B, G, C, N) \mid B = 0\}$ . Then, from our discussion above,  $\omega$  cannot

contain points with  $C > 0$ . Hence  $\omega \subset \{(B, G, C, N) \mid B = 0 \text{ and } C = 0\}$ . Let  $x_0^3 = (0, 0, C_0^3, N_0^3) \in \omega$  and  $(B^3(t), G^3(t), C^3(t), N^3(t))$  be a solution with  $(B^3(0), G^3(0), C^3(0), N^3(0)) = x_0^3$ . Suppose  $C_0^3 < C_{in}$ . Then

$$C^3(-t) = (C_0^3 - C_{in})e^{Dt} + C_{in}, \quad \forall t \geq 0,$$

which implies that  $C^3(-t) \rightarrow -\infty$  as  $t \rightarrow \infty$ . This again contradicts the boundedness of  $\omega$ . Similarly, we arrive to a contradiction if we assume that  $N_0^3 < N_{in}$ . This shows that  $\omega = \{(0, 0, C_{in}, N_{in})\}$ .  $\square$

Next we investigate the uniform persistence-extinction scenario for the grazers. For this, it will be useful to understand the grazer-free dynamics, which are given by the system

$$\begin{aligned} B' &= B[-(D + \varepsilon) + \mu_B \min\{f(N), g(C)\}], \\ C' &= D(C_{in} - C) - \frac{\mu_B}{r} B \min\{f(N), g(C)\} + \varepsilon B, \\ N' &= D(N_{in} - N) + \theta_B \varepsilon B - \theta_B \mu_B B \min\{f(N), g(C)\}, \end{aligned} \tag{3.16}$$

**Lemma 1.** *Assume that  $R(\varepsilon) > 0$  and  $\varepsilon(1 - r) < D$ . Then System (3.16) has a unique interior equilibrium that attracts all solutions with  $B(0) > 0$ .*

*Proof.* From the first and third equations in (3.16) we have

$$(N + \theta_B B)' = -D(N + \theta_B B) + DN_{in}.$$

Hence  $N(t) + \theta_B B(t) \rightarrow N_{in}$  as  $t \rightarrow \infty$ . Thus all solutions of (3.16) are attracted to the set  $\{(B, C, N) \in \mathbb{R}_+^3 \mid N = N_{in} - \theta_B B\}$ . The dynamics

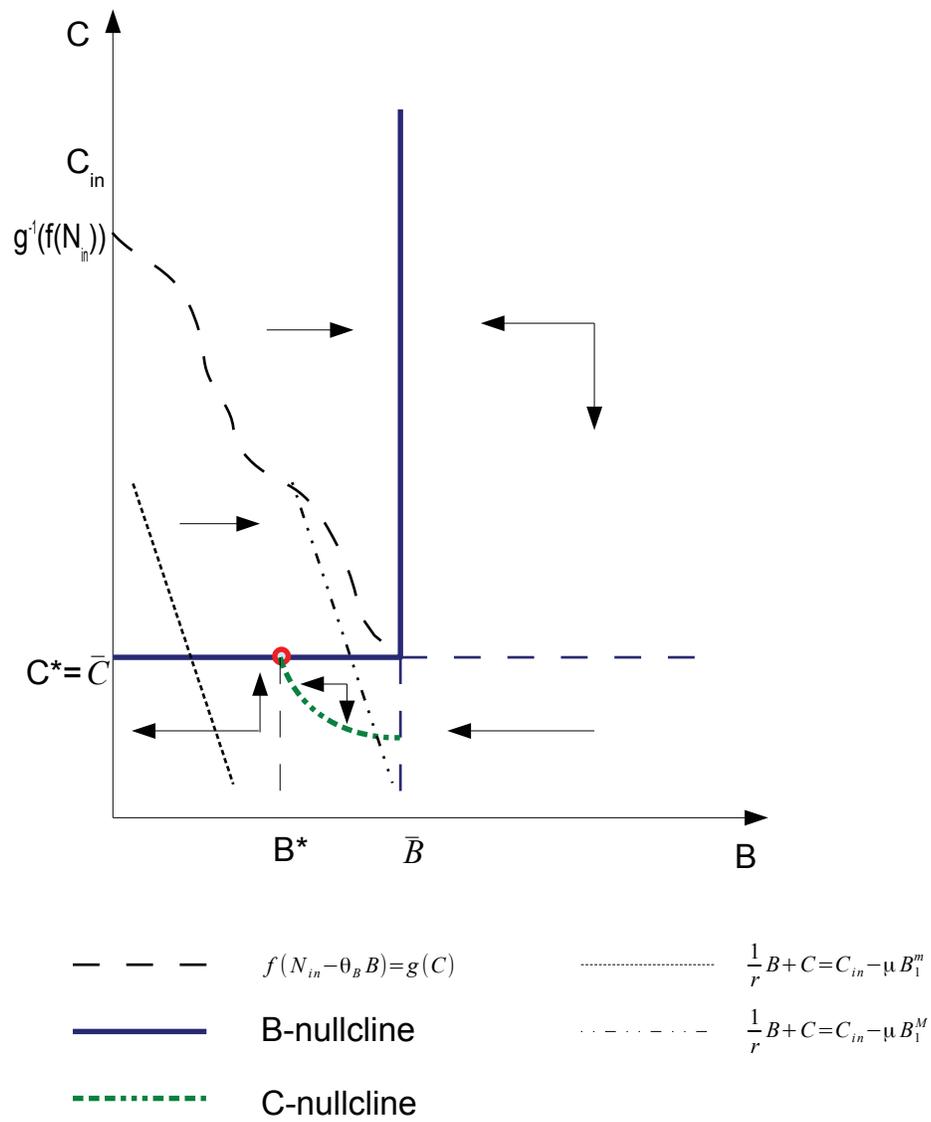
restricted to this set are given by

$$\begin{aligned} B' &= B[-(D + \varepsilon) + \mu_B \min\{f(N_{in} - \theta_B B), g(C)\}] \\ C' &= D(C_{in} - C) - \frac{\mu_B}{r} B \min\{f(N_{in} - \theta_B B), g(C)\} + \varepsilon B \end{aligned} \quad (3.17)$$

Let

$$\bar{B} = \frac{1}{\theta_B} [N_{in} - f^{-1}(\frac{D + \varepsilon}{\mu_B})], \quad \bar{C} = g^{-1}(\frac{D + \varepsilon}{\mu_B}). \quad (3.18)$$

Notice that, since  $R(\varepsilon) > 0$ , both  $\bar{B}$  and  $\bar{C}$  are well defined and positive. The set of points in  $(B, C) \in \mathbb{R}_+^2$  satisfying  $f(N_{in} - \theta_B B) = g(C)$  and  $B \leq \bar{B}$  can be regarded as the graph of the decreasing function  $B \mapsto g^{-1}(f(N_{in} - \theta_B B))$ . Thus, the vector field corresponding to (3.17) looks like in Figure (3.1). Hence the set  $\mathcal{S}_1 = \{(B, C) \in \mathbb{R}_+^2 \mid 0 < B \leq \bar{B}\}$  is positively invariant and attracts all solutions of (3.17).



**Figure 3.1:** Phase portrait of System (3.17)

Next we study the dynamics of (3.17) on this set. It is straightforward to show that (3.17) has a unique equilibrium  $(B^*, C^*)$ , with  $B^* > 0$ . We claim that  $(B^*, C^*)$  attracts all solutions of (3.17) with  $B(0) > 0$ . *Case 1.*  $B^* = \bar{B}$ . Then it is straightforward to argue that the claim holds.

*Case 2.*  $B^* < \bar{B}$ . Then

$$C^* = \bar{C} \text{ and } B^* = \frac{rD(C_{in} - \bar{C})}{\mu_B g(\bar{C}) - r\varepsilon} = \frac{rD(C_{in} - \bar{C})}{D + \varepsilon(1 - r)}. \quad (3.19)$$

Let  $x_0 = (B_0, C_0) \in \mathbb{R}_+^2$  such that  $B_0 > 0$ . Let  $B_1^M = \bar{B}$ . Then

$$\left(\frac{1}{r}B + C\right)' = -D\left(\frac{1}{r}B + C\right) + DC_{in} - \frac{\varepsilon(1 - r)}{r}B \geq -D\left(\frac{1}{r}B + C\right) + DC_{in} - \frac{\varepsilon(1 - r)}{r}B_1^M. \quad (3.20)$$

Hence, all solutions in  $\mathcal{S}_1$  are attracted to the positively invariant set

$$\mathcal{S}'_1 = \{(B, C) \in \mathbb{R}_+^2 \mid \frac{1}{r}B + C \geq C_{in} - \frac{\varepsilon(1 - r)}{rD}\bar{B}\}. \quad (3.21)$$

From this and the direction of the vector field (as depicted in Figure (3.1)) it follows that all solutions in  $\mathcal{S}'_1$  are attracted to the positively invariant set

$$\mathcal{S}''_1 = \{(B, C) \in \mathbb{R}_+^2 \mid B \geq r[C_{in} - \bar{C} - \frac{\varepsilon(1 - r)}{rD}\bar{B}]\} =: B_1^m. \quad (3.22)$$

Then, on  $\mathcal{S}''_1$  we have

$$\left(\frac{1}{r}B + C\right)' \leq -D\left(\frac{1}{r}B + C\right) + DC_{in} - \frac{\varepsilon(1 - r)}{r}B_1^m. \quad (3.23)$$

Hence, all solutions in  $\mathcal{S}_1''$  are attracted to the positively invariant set

$$\mathcal{S}''' = \{(B, C) \in \mathbb{R}_+^2 \mid (\frac{1}{r}B + C) \leq C_{in} - \frac{\varepsilon(1-r)}{rD}B_1^m\}. \quad (3.24)$$

Now consider the line given by

$$(\frac{1}{r}B + C) = C_{in} - \frac{\varepsilon(1-r)}{rD}B_1^m. \quad (3.25)$$

Using (3.22), we have that the point  $(B, C)$  on this line with  $C = \bar{C}$  has

$$\begin{aligned} B = B_2^M &:= r[C_{in} - \bar{C} - \frac{\varepsilon(1-r)}{rD}B_1^m] \\ &= rC_{in}(1 - \frac{\varepsilon(1-r)}{D}) - r\bar{C}(1 - \frac{\varepsilon(1-r)}{D}) + \bar{B}\frac{\varepsilon^2(1-r)^2}{D^2} \end{aligned} \quad (3.26)$$

From this, and using that  $\varepsilon(1-r) < D$ , we obtain that  $B_2^M < B_1^M$  is equivalent to

$$C_{in} - \bar{C} < \frac{\bar{B}}{r}(\frac{\varepsilon(1-r)}{D} + 1). \quad (3.27)$$

On the other hand, using (3.19), we have that

$$\frac{B^*}{r}(\frac{\varepsilon(1-r)}{D} + 1) = C_{in} - \bar{C}. \quad (3.28)$$

Since  $\bar{B} > B^*$ , (3.27) holds, hence  $B_2^M < B_1^M$ . Then all solutions in  $\mathcal{S}_1'''$  are attracted to the positively invariant set

$$\mathcal{S}_2 = \{(B, C) \in \mathbb{R}_+^2 \mid B_1^m \leq B \leq B_2^M\}. \quad (3.29)$$

By repeating (3.20) - (3.26) successively with  $B_1^M$  replaced by  $B_i^M$ ,  $i \geq 2$ , we

obtain two sequences  $(B_i^m)_{i \geq 1}$  and  $(B_i^M)_{i \geq 1}$  that are nondecreasing, respectively nonincreasing, such that  $\omega(x_0)$  is in the set

$$\mathcal{S}_i = \{(B, C) \in \mathbb{R}_+^2 \mid B_i^m \leq B \leq B_{i+1}^M\}, \quad i \geq 1. \quad (3.30)$$

Let  $\bar{B}^M = \lim_{i \rightarrow \infty} B_i^M$ . It follows then that (3.26) holds with both  $B_2^M$  and  $\bar{B}$  replaced by  $\bar{B}^M$ . Hence, we obtain (3.28) with  $B^*$  replaced by  $\bar{B}^M$ , which implies  $B^* = \bar{B}^M$  and, consequently,  $\bar{B}^m = B^*$ . This shows that  $(B^*, C^*) \in \omega(x_0)$ . We have

$$g(C^*) = g(\bar{C}) = f(N_{in} - \theta_B \bar{B}) < f(N_{in} - \theta_B B^*).$$

Hence, in some neighborhood  $V$  on  $(B^*, C^*)$  we have that  $g(C) < f(N_{in} - \theta_B B)$ . The derivative of the vector field in (3.17) restricted to  $V$  is

$$J(B, C) = \begin{pmatrix} -(D + \varepsilon) + \mu_B g(C) & \mu_B B g'(C) \\ \varepsilon - \frac{\mu_B}{r} g(C) & -D - \frac{\mu_B}{r} B g'(C) \end{pmatrix}.$$

We have  $-(D + \varepsilon + \mu_B g(C^*)) = 0$ . So

$$\varepsilon - \frac{\mu_B}{r} g(C^*) = \varepsilon - \frac{\mu_B}{r} \frac{D + \varepsilon}{\mu_B} = \varepsilon \left(1 - \frac{1}{r}\right) - \frac{D}{r} < 0.$$

Thus,  $\text{tr}(J(B^*, C^*)) < 0$  and  $\det(J(B^*, C^*)) > 0$ , which implies that  $(B^*, C^*)$  is asymptotically stable. Hence the claim holds. Now let  $(B(t), C(t), N(t))$  solution of (3.16) with  $B(0) > 0$  and having omega limit set  $\omega_3$ . So  $\omega_3 \subset \{(B, C, N) \in \mathbb{R}_+^3 \mid N = N_{in} - \theta_B B\}$ . Let  $\omega_2 := \{(B, C) \mid (B, C, N_{in} - \theta_B B) \in$

$\omega_3$ . Then  $\omega_2$  is compact and invariant for (3.17). Let  $X := \{(B, C) \in \mathbb{R}_+^2 \mid B > 0\}$ . Then, as shown above,  $\{(B^*, C^*)\}$  is the compact attractor of points in  $X$ , for (3.17). Now, being asymptotically stable,  $\{(B^*, C^*)\}$  attracts a neighborhood of itself. Thus,  $\{(B^*, C^*)\}$  is the compact attractor of compact sets in  $X$  and it contains all compact, invariant ((3.17)) sets in  $X$  (see Theorem ? in [80]). Hence  $\omega_2 = \{(B^*, C^*)\}$  and, consequently,  $\omega_3 = \{(B^*, C^*, N_{in} - \theta_B B^*)\}$ .  $\square$

**Theorem 3** (Grazers persistence-extinction criteria). *The grazers persistence-extinction criteria for System (4.1) are as follows:*

(a) *If  $R(\varepsilon) > 0$ ,  $\varepsilon(1 - r) < D$  and*

$$h(\min\{\frac{1}{\theta_B}[N_{in} - f^{-1}(\frac{D + \varepsilon}{\mu_B})], rD[C_{in} - g^{-1}(\frac{D + \varepsilon}{\mu_B})]/[D + \varepsilon(1 - r)]\}) > \frac{D + \omega}{\alpha\mu_G}. \quad (3.31)$$

*then the grazers are robustly uniformly persistent: there exists  $\eta > 0$ ,  $\delta > 0$  such that*

$$\liminf_{t \rightarrow \infty} G(t) \geq \eta, \quad (3.32)$$

*for all solutions of (4.1) satisfying  $G(0) > 0$ , corresponding to all  $\lambda$  satisfying  $|\lambda - \lambda_0| < \delta$ .*

(b) *If*

$$h(\min\{\frac{1}{\theta_B}[N_{in} - f^{-1}(\frac{D + \varepsilon}{\mu_B})], r[C_{in} - g^{-1}(\frac{D + \varepsilon}{\mu_B})]\}) \leq \frac{D + \omega}{\alpha\mu_G} \quad (3.33)$$

*then  $G(t) \rightarrow 0$  as  $t \rightarrow \infty$ .*

*Proof.* (a) First note that the condition  $R(\varepsilon) > 0$  implies, according to Theorem 2, that the bacteria persists uniformly. Thus, using Theorem 7, we have

that there exists a compact set  $\mathcal{B} \subset \{(B, G, C, N) \in \mathbb{R}_+^4 \mid B > 0\}$  that attracts all solutions of (4.1) with  $B(0) > 0$ . Let

$$X_0^G = \{(B, G, C, N) \in \mathcal{B} \mid G = 0\}.$$

All solutions of (4.1) that originate in  $X_0^G$  satisfy (3.16). From Lemma 1, all solutions of (3.16) with  $B(0) > 0$  converge to an equilibrium  $(B^*, C^*, N^*)$ , where

$$B^* = \min\left\{\frac{1}{\theta_B}[N_{in} - f^{-1}\left(\frac{D + \varepsilon}{\mu_B}\right)], rD[C_{in} - g^{-1}\left(\frac{D + \varepsilon}{\mu_B}\right)]/[D + \varepsilon(1 - r)]\right\}.$$

The conclusion follows now from Corollary 4.7 in [77], applied with  $M = \mathcal{B} \cap X_0^G$  and  $T = 1$  (hence  $\Omega(M) = E_2 = (B^*, 0, C^*, N^*)$ , while  $P(1, E_2) = e^{-(D+\omega)+\alpha\mu_G h(B^*)} > 1$ ).

(b) According to Theorem 2, it suffices to consider only the case  $R(\varepsilon) > 0$ . Let  $(B(t), G(t), C(t), N(t))$  be the solution of (4.1) with  $(B(0), G(0), C(0), N(0)) = x_0 \in \Delta$ . From Corollary 1 we have that

$$B' \leq -(D + \varepsilon)B + \mu_B B \min\left\{f(N_{in} - \theta_B B), g\left(C_{in} - \frac{1}{r}B\right)\right\}. \quad (3.34)$$

On the other hand, all solutions of (3.34) with  $B(0) > 0$  and “ $\leq$ ” replaced by “ $=$ ” converge to

$$\hat{B}^* := \min\left\{\frac{1}{\theta_B}[N_{in} - f^{-1}\left(\frac{D + \varepsilon}{\mu_B}\right)], r[C_{in} - g^{-1}\left(\frac{D + \varepsilon}{\mu_B}\right)]\right\}.$$

Hence  $\omega(x_0) \subset \{(B, G, C, N) \in \mathbb{R}_+^4 \mid 0 \leq B \leq \hat{B}^*\}$ .

Suppose  $\omega(x_0) \subset \{(B, G, C, N) \in \mathbb{R}_+^4 \mid B = \hat{B}^*\}$ .

If  $\omega(x_0) \subset \{(B, G, C, N) \in \mathbb{R}_+^4 \mid B = \hat{B}^* \text{ and } G = 0\}$  then there is nothing left to prove. Otherwise, suppose there exists a point  $\hat{x} = (\hat{B}^*, \hat{G}, \hat{C}, \hat{N})$  in  $\omega(x_0)$  with  $\hat{G} > 0$ . Then, for a solution  $(\hat{B}(t), \hat{G}(t), \hat{C}(t), \hat{N}(t))$  of (4.1) originating at  $\hat{x}$ , we have

$$\hat{B}'(0) \leq -\mu_G h(\hat{B}^*) \hat{G} < 0, \quad (3.35)$$

hence this solution would leave  $\omega(x_0)$ , which represents a contradiction, since  $\omega(x_0)$  is invariant. Thus,  $\omega(x_0) \cap \{(B, G, C, N) \in \mathbb{R}_+^4 \mid 0 \leq B < \hat{B}^*\} \neq \emptyset$ . Also, since  $B(t) = \hat{B}^*$  implies  $B'(t) \leq 0$  (see (3.34)), we have that the solution  $(B(t), G(t), C(t), N(t)) \in \{(B, G, C, N) \in \mathbb{R}_+^4 \mid 0 \leq B < \hat{B}^*\}$  for all  $t > T$ , for some  $T \geq 0$ . But then  $G(t)$  converges to some  $\hat{G}^*$  since  $G'(t) \leq 0$  for all  $t > T$ . Suppose  $\hat{G}^* > 0$ . Then, from the equation for  $G$  in (4.1), we have that  $B(t)$  converges to some  $\tilde{B}$  and  $\tilde{B}$  must be in the interval  $(0, \hat{B}^*)$ . From this it follows that  $G(t) \rightarrow 0$  as  $t \rightarrow \infty$ , which is a contradiction. Thus,  $\hat{G}^* = 0$ .  $\square$

**Remark 3.** *Notice that, in the case when  $\varepsilon = 0$  (no natural death rate for bacteria), Theorem 3 gives a sharp threshold that differentiate between persistence and elimination of grazers. We have not been able to prove (either analytically, or through a numerical counterexample) that the condition  $\varepsilon(1 - r) < D$  is necessary in Lemma 1, or that Theorem 3 (b) holds with the left hand side in (3.33) replaced by the left hand side in (3.31) (in which case we would have a sharp threshold).*

### 3.3.5 Maximizing decomposition rate

Mindful of the decomposition-facilitation paradox in [84], we establish the concentration of grazers that optimizes the decomposition rate. We also determine conditions based on parameter values for which this grazer concentration maximizes decomposition when either nitrogen or carbon is limiting. Let

$$\begin{aligned}
 R &:= \frac{\mu_B}{r} B \min\{f(N), g(C)\}. \\
 \frac{dR}{dG} &= \frac{dR}{dt} \times \frac{dt}{dG} \\
 &= \begin{cases} \frac{(-DB + \mu_B Bg(C) - \mu_G h(B)G - \varepsilon B)g(C)\mu_B +}{r(-DG + \alpha\mu_G h(B)G - \gamma G)} + \\ \frac{Bg'(C)\mu_B \left(D(C_{in} - C) - \frac{\mu_B Bg(C)}{r} + wG + \varepsilon B\right)}{r(-DG + \alpha\mu_G h(B)G - \gamma G)} & \text{if } g(C) < f(N); \\ \\ \frac{(-DB + \mu_B Bf(N) - \mu_G h(B)G - \varepsilon B)f(N)\mu_B +}{r(-DG + \alpha\mu_G h(B)G - \gamma G)} + \\ \frac{Bf'(N)\mu_B(D(N_{in} - N) + \theta_G \gamma G + \theta_B \varepsilon B + (\theta_B - \theta_G \alpha)\mu_G h(B)G - \theta_B \mu_B Bf(N))}{r(-DG + \alpha\mu_G h(B)G - \gamma G)} & \text{if } f(N) < g(C). \end{cases}
 \end{aligned}$$

Solving  $\frac{dR}{dG} = 0$  for G, we get

$$G = \begin{cases} \frac{B(g(C)rD - [g(C)]^2 r\mu_B + g(C)r\varepsilon - g'(C)DrC_{in})}{r(-g(C)\mu_G h(B) + Bg'\gamma)} + \\ \frac{B(g'(C)DrC + Bg'(C)\mu_B g(C) - Bg'(C)\varepsilon r)}{r(-g(C)\mu_G h(B) + Bg'(C)\gamma)} & \text{if } g(C) < f(N); \\ \\ \frac{B(f(N)D - [g(C)]^2 \mu_B + f(N)\varepsilon - f'(N)DN_{in})}{-f(N)\mu_G h(B) + Bf'(N)\theta_G \gamma + Bf'(N)\mu_G h(B)\theta_B - Bf'(N)\mu_G h(B)\theta_G \alpha} + \\ \frac{B(f'(N)DN - Bf'(N)\theta_B \varepsilon + Bf'(N)\theta_B \mu_B f(N))}{-f(N)\mu_G h(B) + Bf'(N)\theta_G \gamma + Bf'(N)\mu_G h(B)\theta_B - Bf'(N)\mu_G h(B)\theta_G \alpha} & \text{if } f(N) < g(C). \end{cases}$$

Next we compute  $\frac{d^2R}{dG^2}$ , evaluate it at the value of G calculated above, and determine the conditions based on the concentration of other variables and model parameters for which  $\frac{d^2R}{dG^2}$  is less than zero.

$$\frac{d^2R}{dG^2} = \begin{cases} -\frac{r\mu_B(-g(C)\mu_G h(B) + Bg'\gamma)^2(D - \alpha\mu_G h(B) + \gamma)^{-1}}{B(g(C)rD - [g(C)]^2r\mu_B + g(C)r\varepsilon - g'(C)DrC_{in} + g'(C)DrC + Bg'(C)\mu_B g(C) - Bg'(C)\varepsilon r)} & \text{if } g(C) < f(N); \\ -\frac{(-f(N)\mu_G h(B) + Bf'(N)\theta_G\gamma + Bf'(N)\mu_G h(B)\theta_B - Bf'(N)\mu_G h(B)\theta_G\alpha)^2\mu_B(D - \alpha\mu_G h(B) + \gamma)^{-1}}{B(f(N)D - [g(C)]^2\mu_B + f(N)\varepsilon - f'(N)DN_{in} + f'(N)DN - Bf'(N)\theta_B\varepsilon + Bf'(N)\theta_B\mu_B f(N))} & \text{if } f(N) < g(C). \end{cases}$$

Thus in a carbon limiting environment, given

$$G = \frac{B(g(C)rD - [g(C)]^2r\mu_B + g(C)r\varepsilon - g'(C)DrC_{in} + g'(C)DrC + Bg'(C)\mu_B g(C) - Bg'(C)\varepsilon r)}{r(-g(C)\mu_G h(B) + Bg'(C)\gamma)},$$

if

i)

$$(g(C)rD - [g(C)]^2r\mu_B + g(C)r\varepsilon - g'(C)DrC_{in} + g'(C)DrC + Bg'(C)\mu_B g(C) - Bg'(C)\varepsilon r) > 0$$

and

ii)  $\alpha\mu_G h(B) < \gamma + D$

or

i)

$$(g(C)rD - [g(C)]^2r\mu_B + g(C)r\varepsilon - g'(C)DrC_{in} + g'(C)DrC + Bg'(C)\mu_B g(C) - Bg'(C)\varepsilon r) < 0$$

and

$$\text{ii) } \alpha\mu_G h(B) > \gamma + D,$$

then grazers will maximize bacterial degradation.

Similarly, in a nitrogen limiting environment, given

$$G = \frac{B(f(N)D - [g(C)]^2\mu_B + f(N)\varepsilon - f'(N)DN_{in} + (f'(N)DN - Bf'(N)\theta_B\varepsilon + Bf'(N)\theta_B\mu_B f(N)))}{-f(N)\mu_G h(B) + Bf'(N)\theta_G\gamma + Bf'(N)\mu_G h(B)\theta_B - Bf'(N)\mu_G h(B)\theta_G\alpha},$$

if

$$\text{i) } f(N)D - [g(C)]^2\mu_B + f(N)\varepsilon - f'(N)DN_{in} + f'(N)DN - Bf'(N)\theta_B\varepsilon + Bf'(N)\theta_B\mu_B f(N) > 0 \text{ and}$$

$$\text{ii) } \alpha\mu_G h(B) < \gamma + D,$$

or

$$\text{i) } f(N)D - [g(C)]^2\mu_B + f(N)\varepsilon - f'(N)DN_{in} + f'(N)DN - Bf'(N)\theta_B\varepsilon + Bf'(N)\theta_B\mu_B f(N) < 0 \text{ and}$$

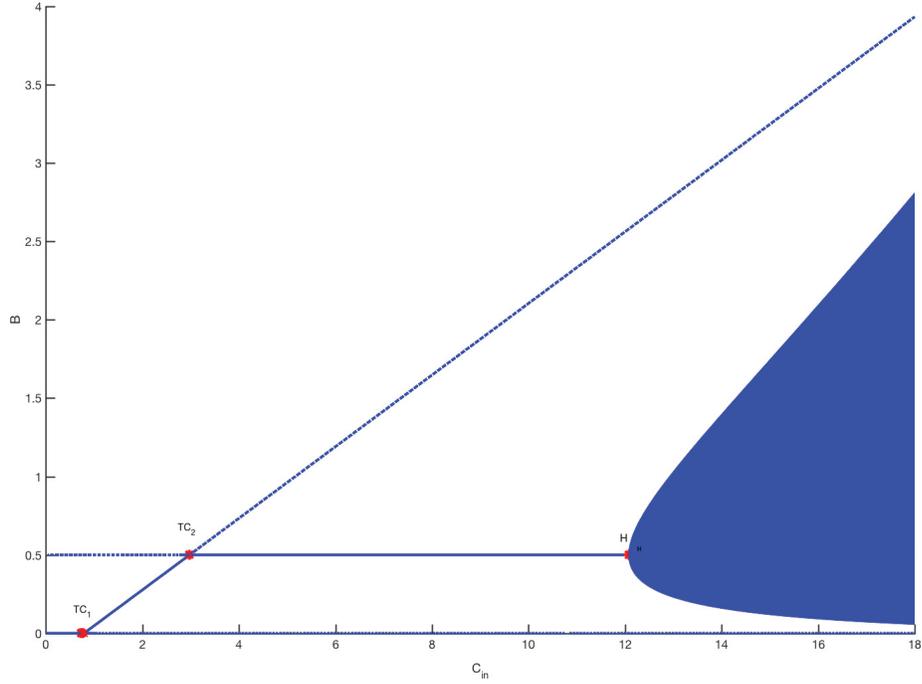
$$\text{ii) } \alpha\mu_G h(B) > \gamma + D,$$

then grazers will maximize bacterial degradation.

### 3.4 Numerical experiments

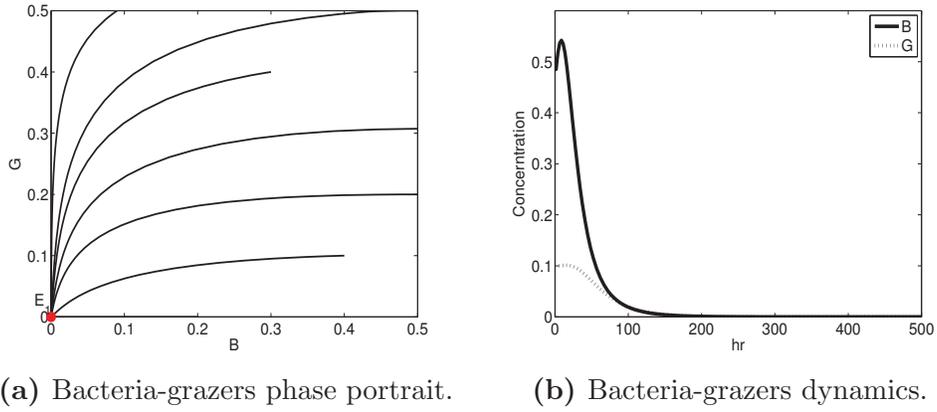
All simulations in this section are performed using Monod or Michaelis-Menten form for the functions  $h$ ,  $f$ , and  $g$ . We begin the section with one- and two-parameters bifurcation diagrams for System (4.1) with dilution rate (per

capita) and concentration of resources in feed bottle as free parameters. The results are obtained using the numerical continuation package MATCONT [59]. Next, we perform numerical simulations using varying values of  $C_{in}$  to illustrate the transfer of stability as the value of this parameter changes. Furthermore, we numerically compare the effect of grazers on degradation in a microcosm and in a chemostat. Moreover, we determine the switching time of bacteria growth rate from carbon dependent to nitrogen dependent for different continuous cultures. We end the section by discussing the sensitivity of the degradation rate to model parameters. Figure (3.2) shows a one-parameter bifurcation diagram obtained by perturbing the concentration of carbon in the supply bottle while holding other parameters fixed. MATCONT detects three important threshold values:  $TC_1$  (labelled BP in Matcont) corresponding to the transcritical bifurcation point that occurs at  $C_{in} = 0.79 \text{ mgC}/dm^3$ ,  $TC_2$  (labelled BP in Matcont) corresponding to the transcritical bifurcation point that occurs at  $C_{in} = 2.98 \text{ mgC}/dm^3$ , H corresponding to the Hopf bifurcation point that occurs at  $C_{in} = 12.08 \text{ mgC}/dm^3$ .



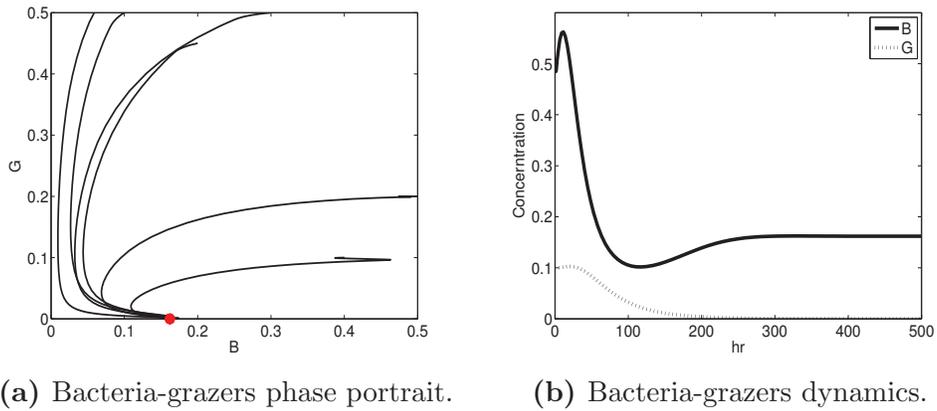
**Figure 3.2:** One-parameter bifurcation diagram for the system (fixed  $N_{in}$  and  $D$ ). The parameters are:  $N_{in} = 7mgN/dm^3$ ,  $D = 0.02$ ,  $\mu_G = 0.5$ ,  $K_g = 8$ ,  $K_f = 1.21$ ,  $\mu_G = 0.25$ ,  $K_h = 1$ ,  $\alpha = 0.36$ ,  $\gamma = 0.01$ ,  $r = 0.4$ ,  $\theta_G = 0.1$ ,  $\theta_B = 0.25$ ,  $\varepsilon = 0.025$ .

For  $C_{in} < TC_1$ , the bacteria and grazers will both go extinct and only the resources will be remaining. For this value of  $C_{in}$ , the insufficient food in the system drives the bacteria population to start going down, when this happens, the grazers population is forced to go extinct. After a while, because of continuous lack of sufficient food for the bacteria, they follow the grazers and go to extinction, see Figures (3.3a, 3.3b).  $R(\varepsilon) := \mu_B \min\{f(N_{in}), g(C_{in})\} - (D + \varepsilon)$ , is less than zero below this threshold value and greater than zero above this threshold value.



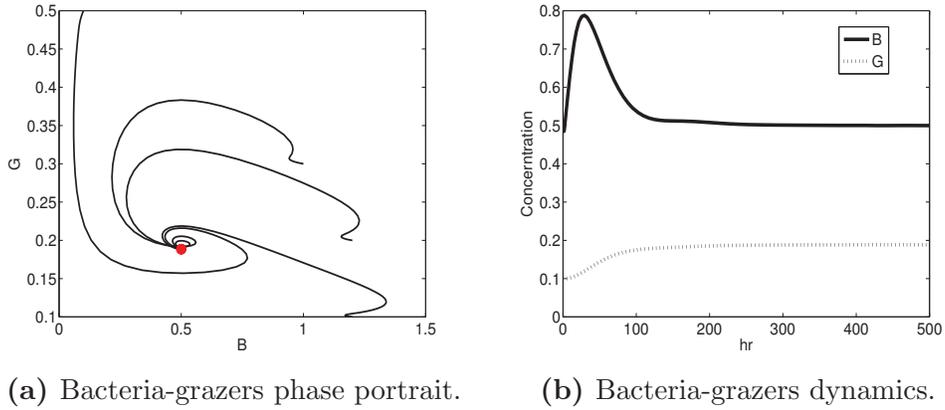
**Figure 3.3:** Bacteria-grazers dynamics.  $C_{in} = 0.3mgC/dm^3$ . Other parameters are same as those in Figure 3.2.

As  $C_{in}$  increases above  $TC_1$  but still less than  $TC_2$ , the food content becomes sufficient to keep the bacteria population from extinction but not the grazers, see Figures(3.4a, 3.4b). Above  $TC_2$ ,  $R(\varepsilon)$  is greater than zero.



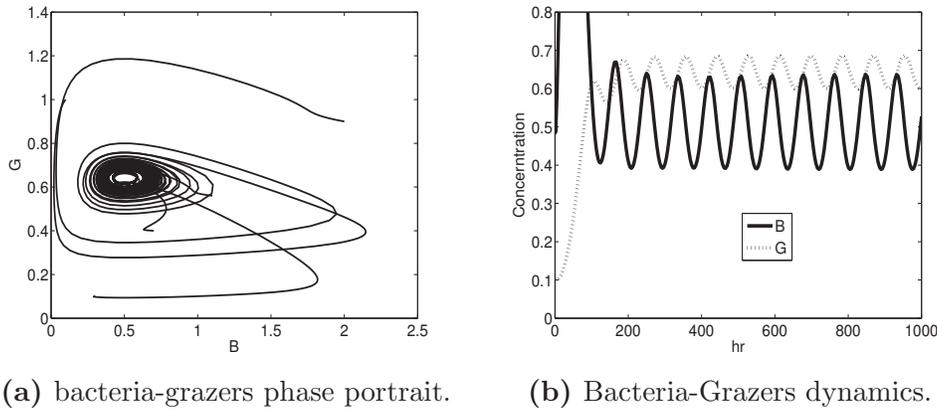
**Figure 3.4:** Bacteria-grazers dynamics. The parameters are:  $C_{in} = 1.5 mgC/dm^3$ . Other parameters are same as those in Figure 3.2.

For  $TC_2 < C_{in} < H$ , both grazers and bacteria coexist with the resources in the culture bottle as shown in Figures (3.5a, 3.5b).



**Figure 3.5:** Bacteria-grazers dynamics. The parameters are:  $C_{in} = 5.5 \text{ mgC}/\text{dm}^3$ . Other parameters are same as those in Figure 3.2.

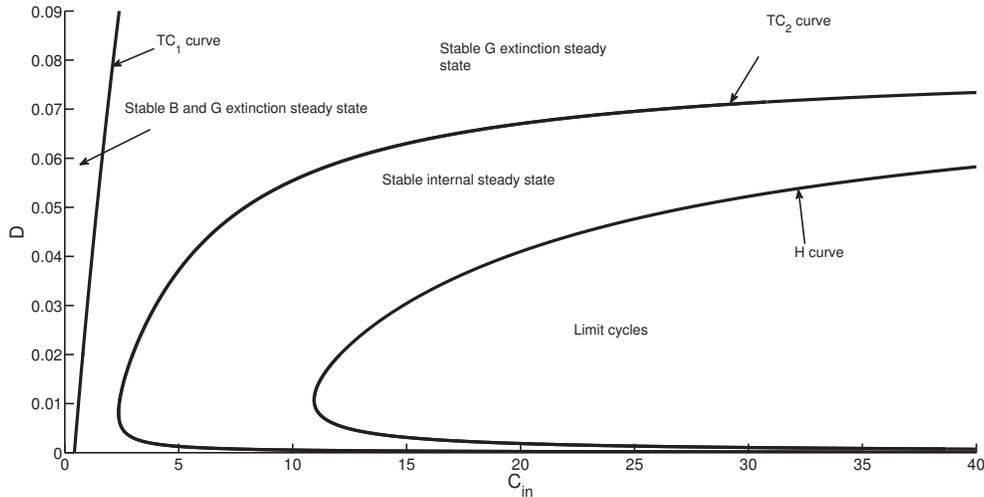
When  $C_{in} = 12.08 \text{ mgC}/\text{dm}^3$ , MATCONT detects a Hopf bifurcation. Figures (3.6a) and (3.6b) respectively illustrate bacteria-grazers phase portrait and the dynamics of bacteria and grazers at this point. Increasing  $C_{in}$  beyond this value causes the internal equilibrium to be unstable and limit cycles appear.



**Figure 3.6:** Bacteria-grazers dynamics. The parameters are:  $C_{in} = 12.08 \text{ mgC}/\text{dm}^3$ . Other parameters are same as those in Figure 3.2.

A similar result is obtained by increasing the concentration of nitrogen in

the influx while leaving other parameters fixed. It is noteworthy that only transcritical bifurcation exists for a microcosm system whereas the chemostat system shows Hopf bifurcations. In Figure (3.7), we show a two-parameter bifurcation diagram. It indicates the values of  $D$  and  $C_{in}$  for which limit cycles occur, for a fixed value of  $N_{in}$ .

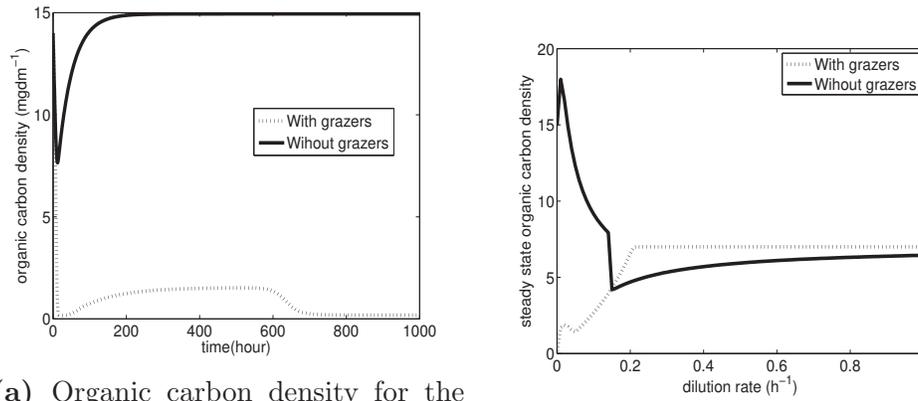


**Figure 3.7:** Two-parameter bifurcation diagram for the system ( $N_{in}$  fixed). The parameters are same as those in Figure 3.2.

### 3.4.1 Impact of grazing on decomposition

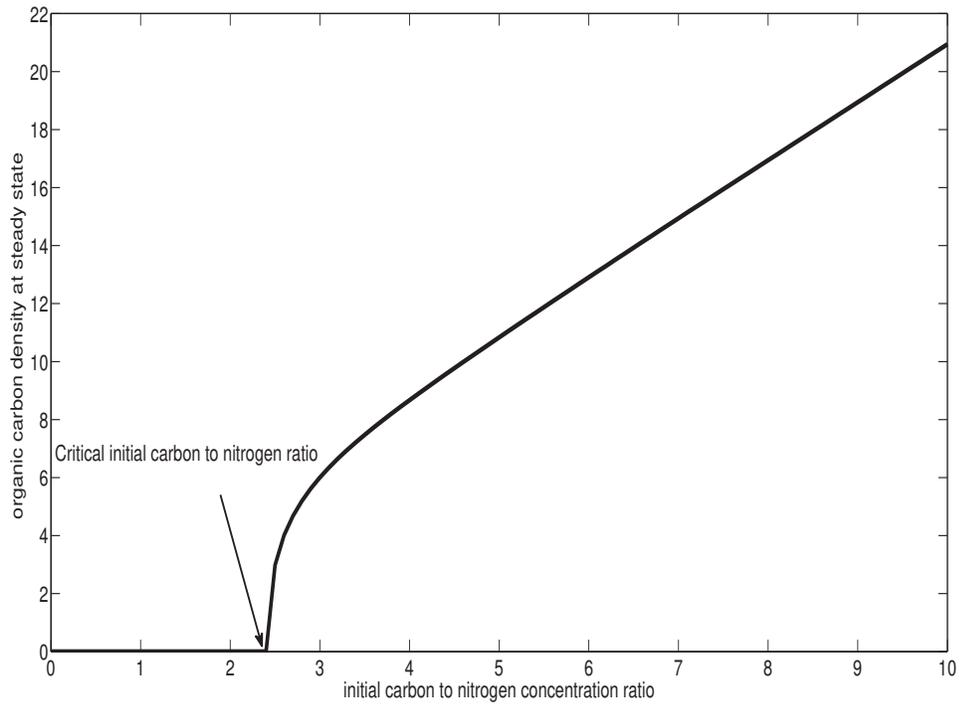
Here we numerically compare the impact of grazers on degradation in a microcosm and in a chemostat by comparing the density of the organic matter left in the system for an ecosystem with and without grazers. Figures (4.8a) and (4.9d) show the dynamics of organic matter in a microcosm and chemostat respectively. Figure (4.8a) confirmed the results in literature that in a microcosm, grazers facilitate the degradation of organic matter. In chemostats, this seems to be the case only if the dilution rate is very small. This is because a

high dilution rate will wash out most of the resources recycled from the microbes. In the microcosm case, these recycled resources end up being used by bacteria for growth. When the environment is initially carbon limiting, grazers have no effect on degradation as the organic carbon will always be completely decomposed. For given parameter values, one can determine a threshold value for the initial ratio of carbon to nitrogen in the environment. Below this value, the organic carbon will always be completely decomposed. In case this ratio is above it, the organic carbon can only be decomposed if at least one of the following conditions is satisfied: a) grazers are present in the environment b) the grazers go extinct only after the organic carbon has been decomposed, and c) the bacteria natural mortality is high (see next subsection for details on this). In Figure (3.9) we use the given parameter values to numerically determine the critical value of the initial carbon to nitrogen concentration ratio for an environment without grazers, above which organic carbon in the environment will never be degraded and below which it would always be completely degraded. Figure (4.9d) predicts that there is a threshold value of the dilution rate for which the grazers will have a positive effect on the organic carbon at steady state. The figure also suggests that there is an optimal value for dilution rate for which the concentration of organic matter in the system at steady state is minimal.



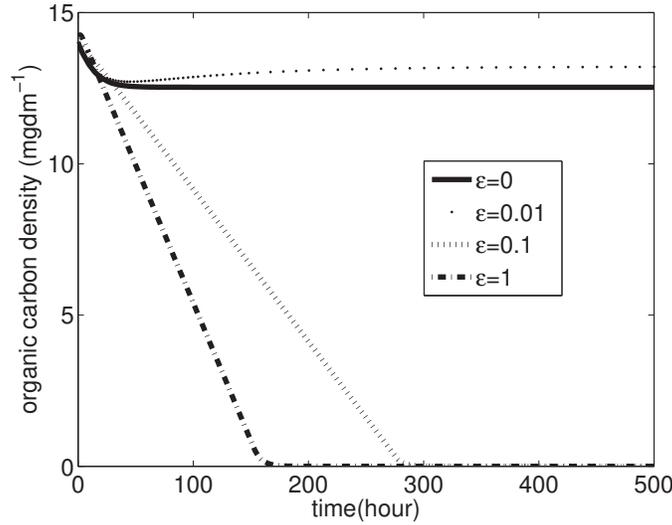
(a) Organic carbon density for the model with or without grazers, and  $D = 0$ . (b) Organic carbon density at steady for the model with or without grazers.

**Figure 3.8:** Organic carbon dynamics. The other parameter values are  $C_{in} = 3 \text{ mgC/dm}^3$ . Other parameters are same as those in Figure 3.2.



**Figure 3.9:** The parameters are:  $C_{in} = 7 \text{ mgCdm}^{-3}$ . Other parameters are same as those in Figure 3.2.

### 3.4.2 Impact of recycled death bacteria on decomposition



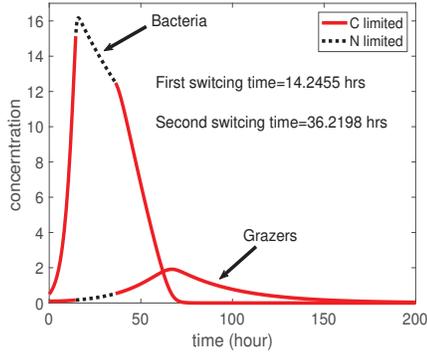
**Figure 3.10:** Impact of recycled death bacteria on decomposition. The parameters are same as those in Figure 3.9.

It is shown in [84] that in the present of grazers, the effect of recycled nutrients from death bacteria is negligible. This is true in a chemostat case as well. In ecosystem with no input and no grazers, the recycled nutrients turn to play a very important rule in degradation as this happen to be the only source of nutrients. Figure 3.10 shows that with no input into a system in the absence of grazers, if there is no natural mortality of bacteria, the organic carbon will not be completely decomposed. For a very small mortality rate, the degradation rate will be a bit low compare to when there is no mortality. This is because, the population of bacteria decreases with mortality and there is no sufficient nutrients to energized the remaining bacteria to increase degradation. With an increase bacteria natural mortality, the degradation rate increases. This

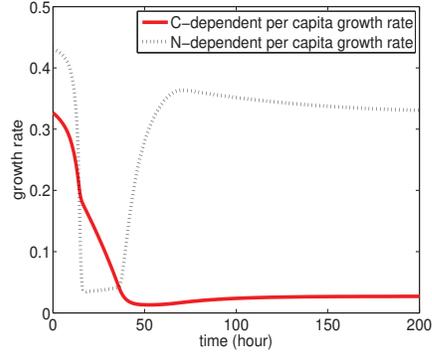
suggest that there is a threshold value of natural mortality, beyond which it has a very positive effect on degradation and below which it has a negative effect. If the time delay that it takes for death bacteria to be converted into nutrients is introduced in the model, the effect will only be noticeable after a long time period.

### 3.4.3 Switching time

Here we determine the switching time of bacteria growth from carbon dependent to nitrogen dependent for different continuous cultures. The objective is to illustrate that bacterial growth could switch from being limited by carbon to being limited by nitrogen and vice versa at different time. We differentiate the continuous cultures by varying the parameters. We do this only for two extreme cases: bacteria and grazers both going extinct and the case where we have limit cycles. The switching time of bacteria growth for a continuous culture in which bacteria and grazers will go extinct (in such an ecosystem, (a)  $R(\epsilon) \leq 0$ ) and one with limit cycles (in such an ecosystem, (b) (i)  $R(\epsilon) > 0$ , (ii)  $\epsilon(1 - r) < D$ , (iii)  $h(\min\{\frac{1}{\theta_B}[N_{in} - f^{-1}(\frac{D+\epsilon}{\mu_B})], rD[C_{in} - g^{-1}(\frac{D+\epsilon}{\mu_B})]/[D + \epsilon(1 - r)]\}) > \frac{D+\gamma}{\alpha\mu_G}$  and resources are sufficient) are respectively shown in Figures (3.11a) and (3.12a). Figures (3.11b) and (3.12b) show the nitrogen and carbon dependent bacterial growth rates for the given continuous cultures.

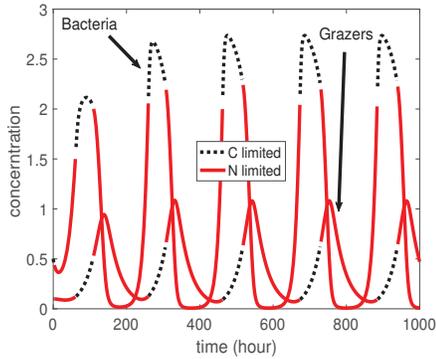


(a) Bacteria-grazer dynamics showing switching time.

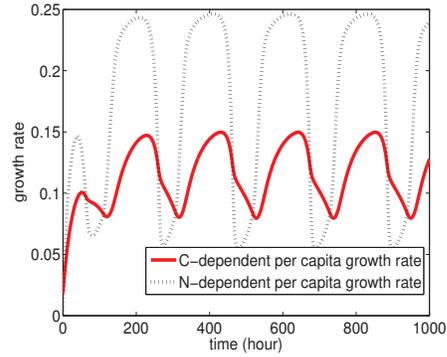


(b) Bacterial per capita growth rate.

**Figure 3.11:** Parameter values are  $N_{in} = 15 \text{ mgN/dm}^3$ ,  $C_{in} = 3 \text{ mgC/dm}^3$ ,  $K_g = 53$ ,  $K_f = 8$ ,  $\theta_G = 1.25$ ,  $\theta_B = 2.5$ . Other parameters are same as those in Figure 3.2.



(a) Bacteria-grazer dynamics showing switching time.



(b) Bacterial per capita growth rate.

**Figure 3.12:** Switching time: 2.5092 hr, 61.9026 hr, 110.0538 hr, 259.7272 hr, 307.0865 hr, 462.4208 hr, 518.0037 hr, 673.3283 hr, 728.8563 hr, 884.1833 hr, 939.5180 hr. Parameter values are:  $N_{in} = 8 \text{ mgN/dm}^3$ ,  $C_{in} = 25 \text{ mgC/dm}^3$ . Other parameters are same as those in Figure 3.11.

### 3.4.4 Sensitivity analysis

The objective of this subsection is to discuss the sensitivity of the degradation rate to model parameters. To do this, we use the normalized forward sensitivity

index [58]:

$$\text{sensitivity index(S.I.)} = \left( \frac{\partial(\text{degradation rate})}{\partial(\text{parameter})} \right) \left( \frac{\text{parameter}}{\text{degradation rate}} \right) \quad (3.36)$$

where the degradation rate is given by  $\frac{1}{r}\mu_B B \min \{f(N), g(C)\} := R$ . Since the variables B, C and N are functions of some parameters that are not explicitly included in the degradation rate formula, we use numerical methods to evaluate the derivatives. Using central difference approximation we have that

$$\frac{\partial R}{\partial \text{parameter}} = \frac{R(\text{parameter}+h) - R(\text{parameter}-h)}{2h} + O(h^2).$$

Letting  $h = 1\%$  of the parameter value (P), Equation (5.6) becomes

$$S.I. = \frac{R(1.01P) - R(0.99P)}{0.02(R(P))}.$$

To evaluate this, we start with a high initial value of resources and ensure that the chosen ecosystem is such that the organic carbon density is decreasing, and then compute the rate of degradation after a day. The sign of the S.I indicates the nature of the relationship between the degradation rate and the parameter in question whereas its magnitude ranks the strength of the relationship in comparison to the other parameters. Table (3.5) shows the computed sensitivity index of the degradation rate with respect to model parameters. The degradation rate is more sensitive to the maximum bacteria growth rate. The positive value of the sensitivity index of the maximum bacteria growth rate indicates that an increase in bacteria growth rate will lead to an increase in the degradation rate. This is as expected, since we assumed that the growth rate

is proportion to the degradation rate. Unfortunately, an experimenter can not make good use of this to regulate the degradation rate as this parameter is not under his/her control. Among the parameters that can be controlled by the experimenter, the dilution rate has the greatest effect on the degradation rate whereas the density of carbon in the feed bottle has the least effect. Thus, an experimenter could increase the degradation of organic carbon in a system by simply decreasing the dilution rate.

Parameter	Sensitivity index
$\mu_B$	57.1471
$K_h$	0.0294
$K_f$	-0.0518
$K_g$	-0.0471
$\mu_G$	-0.0778
$\varepsilon$	-0.6940
$\gamma$	0.0021
$r$	-1.3677
$\theta_B$	-0.1113
$\theta_G$	$-6.9408 \times 10^{-5}$
$D$	-0.4992
$C_{in}$	$3.1612 \times 10^{-5}$
$N_{in}$	0.0024
$\alpha$	-0.0112

**Table 3.5:** The sensitivity of degradation rate to the parameters.

## 3.5 Discussion

In this chapter, a stoichiometry-based model for bacteria-grazer interaction in a continuous culture is proposed and analyzed. Nitrogen and carbon are considered to be the only growth limiting elements in the culture. We use a general class of monotonically increasing, saturating functions, which take the value zero when their arguments are zero to represent the per capita grazing efficiency and per capita bacterial growth rate. Criteria for the uniform persistence and extinction of bacteria and grazers are obtained. Moreover, we determine the optimal value of grazers that maximizes degradation of organic matter. Based on the parameters under the control of the experimenter (elements in feed bottle and dilution rate), we perform one- and two-parameter bifurcation analyses. Furthermore, we determine the switching time of bacterial growth rate from carbon dependent to nitrogen dependent and vice versa for different continuous cultures. The sensitivity of the degradation rate with respect to the model parameters is also shown. We also discuss numerically how the dilution rate affects the decomposition percentage as well as decomposition speed.

In general, we can summarize the uniform persistence and extinction results as follows:

- (a) The resource in the chemostat will never go extinct so long as the supply bottle has positive resource concentrations.
- (b) If at the unique internal steady state, the per capita growth rate of bacteria is greater than the sum of the bacteria's death rate and dilution rates, then the bacteria will persist uniformly otherwise, it will go extinct. It is

thus possible for the experimenter to have the bacteria persist uniformly by either using (i) a high concentration of resources in the supply bottles, (ii) by decreasing the dilution rate, or (iii) carrying out both. Our results equally show that if the bacteria ever go extinct, the grazers will go extinct as well.

- (c) If at the unique internal steady state, (i) the per capita growth rate of bacteria and grazers are respectively greater than the sum of their corresponding dilution and death rates, and (ii) the death rate of bacteria is less than a certain threshold value,  $\frac{D}{1-r} := \varepsilon_2$ , then the grazers will persist uniformly along side bacteria. Should the per capita growth rate of grazers happen to be less than the sum of its dilution rate and death rate at the unique internal steady state, the grazers will go to extinction independent of the concentration of bacteria. An experimenter can prevent grazers from extinction by using high resource concentrations in the supply bottle.

It has been widely reported in microbiology papers that the rate of organic matter degradation often increases in the presence of bacterivorous protists that substantially reduce bacterial abundance [62, 74, 78, 69]. We established conditions based on parameter values for this paradox to hold in a continuous culture. Since some of these parameters like the dilution rate and the resources in the supply bottle could be controlled by the experimenter, one could hold the other parameters fixed and control these to maximize the degradation rate of organic carbon.

It is noteworthy that if there is no dilution and the total nitrogen in the

entire system is high enough, the dynamics of the system will be the same as that of the classical predator-prey models for bacteria-grazer interactions.

It is also worth mentioning that the model in this paper have been modified to one that models the decomposition of tailing hydrocarbons in a manuscript under preparation by Kong et al.[71]. The modified model has been fitted to tailing hydrocarbons' data to determine the optimal parameter values. Using these estimated parameter values, the modified model has been extended to a methane generation model.

Bifurcation analysis shows the rich dynamics of the system. In addition to the transcritical bifurcation observed in microcosm models, our chemostat model exhibits Hopf bifurcation and Rosenzweig's paradox of enrichment phenomenon. Even though the threshold value for which periodic orbits exist is numerically established, we fell short of coming up with a mathematical condition for it in terms of parameter values. This is due to the fact that there are limited tools to establish the existence of limit cycles for non-differentiable but Lipschitz continuous dynamical systems as our model.

In deriving the model, we assume that dead bacteria and grazers are immediately recycled back into the ecosystem. In reality, they are recycled back with a delay. It may be interesting to study the dynamics of the system incorporating a delay in recycling dead bacteria and grazers.

In the process of degrading organic substances different groups of bacteria breaks down different organic substances. Hydrolytic bacteria decomposes complex organic substances, whereas homoacetogens and fermentative bacteria break down simple organic substances. The model in this paper groups all the bacteria into one class and the organic substances into another class.

In a future work, we intend to extend our current model to one that incorporates the dynamics of hydrolytic bacteria, homoacetogens, fermentative bacteria, grazers, nutrients, and complex and simple organic substances. We will study among other things the effect of grazers on the competition between homoacetogens and fermentative bacteria and the effect of grazers on bacterial diversity.

# Chapter 4

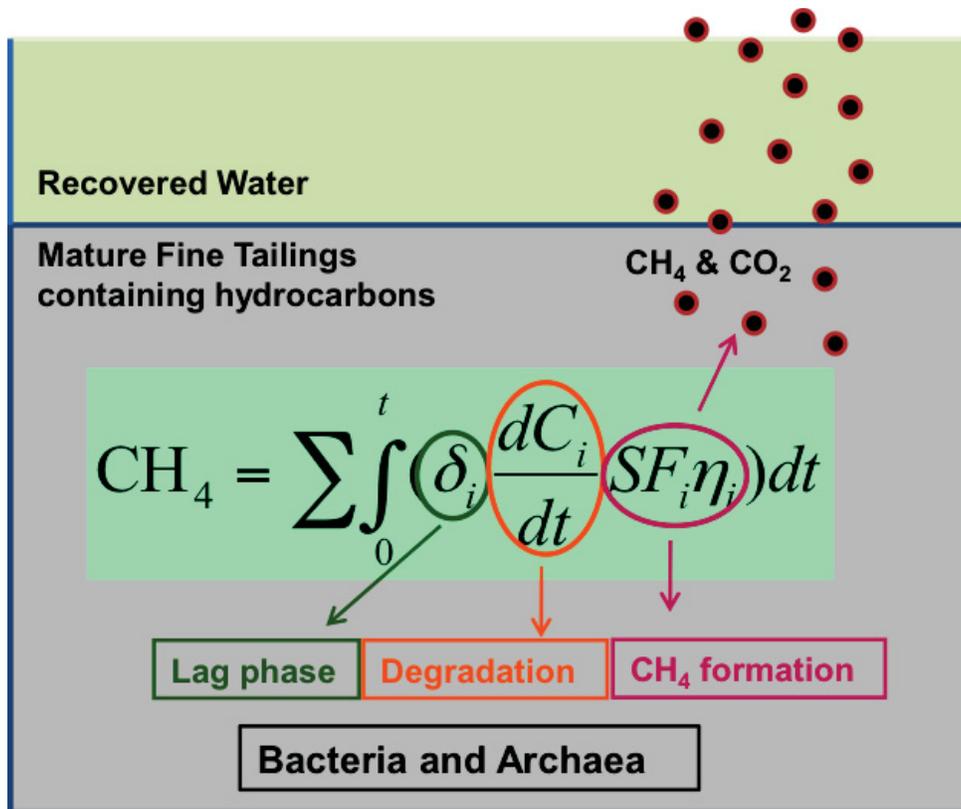
## Predictive model for methane emissions from oil sands tailings ponds and end pit lakes

### Abstract

Microbial metabolism of unrecovered hydrocarbons leads to methane ( $\text{CH}_4$ ) generation in oil sands tailings ponds and End Pit Lakes (EPL). Predicting greenhouse gas emissions is important for both industry and government for mitigation. We developed a biodegradation model which takes into account inflow of carbon (petroleum hydrocarbons), microbial growth and death rates, and kinetics of hydrocarbon biodegradation to predict  $\text{CH}_4$  production. Mathematical analysis reveals that if carbon is the limiting nutrient, the bacteria in the pond will go to extinction in the absence of any inflow of carbon. Model

simulations were performed to determine optimal model parameter values using laboratory experimental data revealing methanogenic biodegradation of hydrocarbons in oil sands tailings. We also extended the biodegradation model to a CH<sub>4</sub> biogenesis model. Using the estimated parameter values, the validity of the CH<sub>4</sub> biogenesis model was determined by comparing the model's output to the measured CH<sub>4</sub> obtained in hydrocarbon biodegradation experiments. Goodness of fit analysis shows that our model's predictions are comparable to testing data.

#### 4.1 Graphical abstract



## 4.2 Introduction

Alberta's oil sands industry is a major economic driver in Canada and is currently producing  $\sim 2.3$  million barrels oil day<sup>-1</sup> which is expected to reach 4 million barrels day<sup>-1</sup> in 2024 (<http://www.energy.alberta.ca/oilsands/oilsands.asp>). However, the oil sands sector has come under international scrutiny concerning industry standards on greenhouse gas (GHG) emissions and other environmental issues. Oil sands operations (mining and upgradation) are responsible for  $\sim 23\%$  of Alberta's overall GHG emissions, with methane (CH<sub>4</sub>) accounting for  $\sim 6\%$  of the oil sands emissions, which does not include emissions from oil sands tailings ponds (OSTP) (<http://environment.gov.ab.ca/info/library/8849.pdf>). CH<sub>4</sub> and carbon dioxide (CO<sub>2</sub>) emissions from tailings ponds are currently measured using floating flux chambers. The total fugitive GHG emissions measured in 2011 from OSTP of major oil sands operators (Suncor, Syncrude and Shell Albian) were 2.8 million tonnes CO<sub>2</sub> equivalent, calculated using 25 as global warming factor (GWP) for CH<sub>4</sub> [5]. The GHG emissions from OSTP result primarily from the biodegradation of residual hydrocarbons which are components of extraction solvent (diluent) entrained in tailings [37, 38, 41, 40, 34]. During bitumen extractions from oil sands ore, naphtha or paraffinic solvent is added as a diluent to assist both in separating the bitumen from solid inorganic matrix and water, as well as in diluting the bitumen so that it may be easily transported to processing and upgrading facilities. Resulting tailings (slurries of alkaline water, sand, silt, clays, unrecovered bitumen and organic diluent) from the extraction process are deposited in OSTP. Indigenous microorganisms in OSTP biodegrade

diluent hydrocarbons syntrophically to produce  $\text{CH}_4$  and  $\text{CO}_2$  [37, 1, 47, 34]. Tailings in OSTP settle slowly to form mature fine tailings (MFT; > 30% solids) and release pore water for recycling. End Pit Lakes (EPLs) are being developed as a new reclamation approach to reintegrate the accumulated tailings into the natural environment. Once the tailings have settled in OSTP, MFT and small amounts of other by-products (e.g. tailings sands, petroleum coke and composite tailings) are transported to mine-out pits and capped with fresh water and/or process-affected water to create a sustainable aquatic system (EPL) over time capable of supporting other economical, ecological and societal uses [22]. Methane biogenesis differs in OSTP and EPLs in that OSTP have a continuous input of diluent containing fresh tailings providing a constant influx of labile hydrocarbons for microbial metabolism, whereas EPLs receive no fresh input of hydrocarbons, rather MFT transferred from OSTP to EPL are generally depleted in labile hydrocarbons due to microbial biodegradation but still contain hydrocarbons that are biodegradable. In reality, the first full-scale demonstration EPL (BML; Base Mine Lake) at Syncrude Canada Ltd. (SCL) produces biogenic gases ( $\text{CH}_4$  and  $\text{CO}_2$ ) that not only contribute to greenhouse gas emissions but also perturb the mudline (MFT-cap water interface) causing turbidity in EPL. Therefore, it is important for both industry and government to accurately predict the  $\text{CH}_4$  emissions from OSTP and EPL for effective tailings management and GHG mitigation. This can be efficiently achieved by mathematical modelling.

Microbial metabolism requires other elements such as nitrogen (N), phosphorus (P) and sulphur (S) as nutrients for growth while biodegrading hydrocarbons; therefore, the growth rate of bacteria and kinetics of hydrocarbon

biodegradation depend on the availability of such nutrients to sustain methanogenesis in OSTP/EPL. Though the tailings microbes have the ability to fix  $N_2$  to meet N requirement for  $CH_4$  production [9], the sustenance of methanogenesis in MFT covered under a body of water that precludes diffusion of atmospheric  $N_2$  still needs to be investigated.

In the first attempt to predict methane biogenesis and flux from MLSB, zero- and first-order kinetic models were developed [39] using the rate constants derived from metabolism of labile hydrocarbon to  $CH_4$  in MFT collected from MSLB [37, 38]. A couple of factors were assumed while designing those models such as (1) certain hydrocarbons such as n-alkane and some monoaromatics are biodegradable under methanogenic conditions, (2) constant bacterial biomass during hydrocarbon metabolism assuming stable microbial population size in tailings pond that receive continuous and consistent input of hydrocarbons, (3) stoichiometric conversion of hydrocarbons to  $CH_4$  using 80% conversion efficiency, and (4) organic carbon concentration ( $C$ ) being very large compared to half the maximum substrate utilization rate ( $C_m$ ) (zero-order; ( $C_m \ll C$ ) or first order under conditions of very low concentrations ( $C \ll C_m$ ). These approaches assume that the constant microbial biomass is implicitly included in the rate coefficient. In this case, the estimated biodegradation rate for each hydrocarbon depends on the concentration of the hydrocarbon in OSTP/EPL in addition to a combination of at least three independent factors:

- a) chemical structure and biodegradability of the hydrocarbon [27],
- b) amount of active microbial biomass associated with and degrading the given OSTP/EPL ([16, 51]), and

- c) the physical probability that the microbes will interact with the given hydrocarbon [43, 26].

In reality, all these factors are not constant in time, and this makes it difficult to find a unique parameter value applicable for various OSTP/EPL and to link a fitted parameter with measured chemical and physical OSTP/EPL properties or with biological features of microbes associated with biodegradation of hydrocarbons.

The present study is designed to: 1) modify our previous model to incorporate microbial dynamics, 2) estimate the parameters of the OSTP/EPL hydrocarbon biodegradation model/methane biogenesis model, 3) validate the model using data that was not used in estimating the parameters of the model (ensuring that the estimated parameters are valid for any given OSTP/EPL).

## **4.3 Material and method**

### **4.3.1 Model development**

#### **Biodegradation model**

The model describes the dynamics of bacteria competing for resources in a resource limited environment. The change in bacterial biomass can result from two processes (growth and death). The per capita bacterial growth rate is assumed to follow the Liebig's law of minimum [45]. According to this law, the growth rate of bacteria is proportional to the most limiting resource in the tailing. We assume that all elements except nitrogen and carbon are present in abundance. Thus, the growth rate of bacteria is a function of only nitrogen

and the hydrocarbons. The total nitrogen in the system (T) is considered to be constant. With this assumption, the total nitrogen available for bacterial growth is given by  $T_N = T - \theta B$ , where  $\theta$  is the nitrogen to carbon ratio of bacteria which is assumed to be constant. The Monod functions  $f(T_N) = \frac{T_N}{T_N + K_f}$  and  $g(C_i) = \frac{C_i}{C_i + K_{g_i}}$  (where  $i = 1 \dots, n$ ,  $n$ , being the total number of hydrocarbons in the medium,  $C_i$  the  $i$ -th hydrocarbon), are used to model the nitrogen and  $i$ th- hydrocarbon dependent per capita bacterial growth rates respectively, where  $K_f = N$  dependent half saturation constant and  $K_{g_i} = C_i$  dependent half saturation constant. Thus, the per capita bacterial growth rate is given by  $\mu_i \min\{f(T_N), g(C_i)\}$ , where  $\mu_i$  is the maximum growth rate of bacteria when only the hydrocarbon  $C_i$  is present in the environment.  $\mu_i \min\{f(T_N), g(C_i)\}$  and  $\mu_j \min\{f(T_N), g(C_j)\}$  for  $i \neq j$ , are assumed to be independent processes. Thus the total per capita growth rate of bacteria is  $\sum_{i=1}^n \mu_i \min\{f(T_N), g(C_i)\}$

The biodegradation rate of each hydrocarbon is assumed to be proportional to the bacterial growth rate due to its uptake, this means that a fixed amount of growth results from the metabolism of a unit quantity of a hydrocarbon i.e.,

$$[\text{bacterial growth rate due to each hydrocarbon}] \propto [\text{biodegradation rate of hydrocarbon}].$$

This implies that

$$[\text{per capita bacterial growth rate due to } C_i] = r_i [\text{per capita biodegradation rate of } C_i]$$

where  $r_i$  is the proportionality constant which we refer to as the yield constant.

It measures the efficiency of conversion of  $C_i$  into bacteria biomass. Hence,

[per capita biodegradation rate of  $C_i$ ] =  $\frac{1}{r_i}$  [per capita bacterial growth rate due to  $C_i$ ],  
i.e., [per capita biodegradation rate of  $C_i$ ] =  $\frac{1}{r_i} \mu_i \min\{f(T_N), g(C_i)\}$

Under an active tailings pond scenario, a constant inflow (concentration) of  $C_i$ ,  $C_i^{in}$  per unit time is expected while no hydrocarbon impacted tailings are released to end pit lake after its establishment. The per capita death rate of bacteria ( $\varepsilon$ ) is assume to be constant. Upon death, bacteria are assumed to be recycled back into useable hydrocarbons immediately. The fraction of  $C_i$  recycled from dead bacteria is assume to be a constant  $\beta_i$ , Where  $0 < \beta_i < 1$ . In line with the observations in[37, 38, 34, 35, 36]c, we assume that each hydrocarbon observed a constant lag period,  $\tau_i$  before the onset of biodegradation. The above assumptions lead to the following systems of equations:

$$\begin{aligned}
g(C_i) &= \begin{cases} 0 & t < \tau_i \\ \frac{C_i}{K_{gi} + C_i} & t \geq \tau_i \end{cases} \\
\dot{B} &= B \sum_{i=1}^n \mu_i \min \left\{ \frac{T_N}{K_f + T_N}, g(C_i) \right\} - \varepsilon B \\
\dot{C}_i &= \frac{-1}{r_i} \mu_i B \min \left\{ \frac{T_N}{K_f + T_N}, g(C_i) \right\} + \beta_i \varepsilon B + C_i^{in} \\
T_N &= T - \theta B \\
B(0) &> 0, C_i(0) \geq 0.
\end{aligned} \tag{4.1}$$

For simplicity, we assume that  $\mu_i$ ,  $r_i$ , and  $\beta_i$  are the same for all  $C_i$ ,

i.e.  $\mu_i = \mu$ ,  $r_i = r$ , and  $\beta_i = \beta$ . With this assumption, System (4.1) becomes

$$\begin{aligned}
g(C_i) &= \begin{cases} 0 & t < \tau_i \\ \frac{C_i}{K_{gi} + C_i} & t \geq \tau_i \end{cases} \\
\dot{B} &= B \sum_{i=1}^n \mu \min \left\{ \frac{T_N}{K_f + T_N}, g(C_i) \right\} - \varepsilon B \\
\dot{C}_i &= \frac{-1}{r_i} \mu B \min \left\{ \frac{T_N}{K_f + T_N}, g(C_i) \right\} + \beta \varepsilon B + C_i^{in} \\
T_N &= T - \theta B \\
B(0) &> 0, C_i(0) \geq 0.
\end{aligned} \tag{4.2}$$

### Methane biogenesis model

Immediate conversion of biodegraded hydrocarbons to bacteria biomass,  $\text{CH}_4$  and  $\text{CO}_2$  by indigenous microbes is a key underlying assumption in the  $\text{CH}_4$  generation model. Thus, in converting the hydrocarbon biodegradation model (System (4.2)) to a  $\text{CH}_4$  generation model, we simply add  $n + 1$  equations to System (4.2);  $n$  of them are ordinary differential equations that dynamically track the increase in the gaseous output ( $\text{CH}_4$  and  $\text{CO}_2$ ) of each of the  $n$ -biodegraded hydrocarbons, denoted by  $G_i$ ,  $i = 1, 2, \dots, n$ . The last additional equation is an algebraic equation that computes the total generated  $\text{CH}_4$ . Our

CH<sub>4</sub> generation model is thus given by the following systems of equations:

$$\begin{aligned}
g(C_i) &= \begin{cases} 0 & t < \tau_i \\ \frac{C_i}{K_{gi} + C_i} & t \geq \tau_i \end{cases} \\
\dot{B} &= B \sum_{i=1}^n \mu \min \left\{ \frac{T_N}{K_f + T_N}, g(C_i) \right\} - \varepsilon B \\
\dot{C}_i &= \frac{-1}{r_i} \mu B \min \left\{ \frac{T_N}{K_f + T_N}, g(C_i) \right\} + \beta \varepsilon B + C_i^{in}, \quad i = 1 \dots, n \\
\dot{G}_i &= \frac{1}{r_i} \mu B \min \left\{ \frac{T_N}{K_f + T_N}, g(C_i) \right\}, \quad i = 1 \dots, n, \\
\text{CH}_4 &= \sum_{i=1}^n \eta_i \Gamma_i G_i \\
T_N &= T - \theta B \\
B(0) &> 0, C_i(0) \geq 0, G_i(0) = 0,
\end{aligned} \tag{4.3}$$

where  $\Gamma_i$  is the expected CH<sub>4</sub> yield when one mole of  $C_i$  biodegrades.  $\eta_i$  is the maximum theoretical yield of CH<sub>4</sub> and CO<sub>2</sub> from  $C_i$ . The product,  $\eta_i \Gamma_i$ , gives the maximum potential CH<sub>4</sub> yield when  $C_i$  biodegrades. We assume that  $\eta_i$  is the same for all  $C_i$ , i.e.  $\eta_i = \eta$ . [37, 38, 34, 35, 36]

### 4.3.2 Source of data for parameter estimation and model validation

Fugitive diluent in tailings is the main source of methanogenesis in tailings ponds [38]. The most commonly used diluents are naphtha and paraffinic solvent. Syncrude Canada Ltd. (Syncrude), Suncor Energy (Suncor), and Canadian Natural Resources Ltd. (CNRL) use naphtha as a solvent. Naphtha

comprises aliphatic and aromatic hydrocarbons (n-, iso- and cycloalkanes and BTEX (benzene, toluene, ethylbenzene and xylenes)) primarily in  $C_6-C_{10}$  range. Shell Albian Sands (Albian) and Imperial Oil use paraffinic solvent (primarily  $C_5-C_6$ ; n- and iso-alkanes) in converting crude oil to bitumen. The data used to fit the parameters and test the model are obtained from [37, 38, 42, 33, 34, 35, 36]. All these experiments were carried out with tailings containing one of the solvents above. The results from these studies reveal that incubating tailings with either naphtha or paraffinic solvents by-products, under methanogenic conditions, for up to  $\sim 1700$  days, only pentane, hexane, heptane, octane, nonane, decane, toluene, o-Xylene, m-plus-p-Xylene, 2-methylpentane, 3-methylhexane, 2-methylheptane, 4-methylheptane, and 2-methyloctane will undergo significant biodegradation. Thus with the goal of predicting methane biogenesis, we estimate only the parameters of the model related to these hydrocarbons. In estimating parameters pertaining to certain hydrocarbons, only the dynamics of those hydrocarbons in the experiment are considered.

The biodegradation of naphtha components in Syncrude MLSB has been studied in in [37, 38]. In [37], MFT from Syncrude MLSB was amended with some n-alkanes (namely: hexane, heptane, octane, and decane) ( $\sim 0.2\%$ ) in a microcosm and incubated under methanogenic conditions for 322 days. The concentration of each of the n-alkanes left in the microcosm at certain time within the incubation period was quantified. A complete biodegradation of all the n-alkanes was observed during the incubation. Fitting our model to this data, we estimate the initial concentration of microbes in tailings, the nitrogen-dependent H.S.C. for bacterial growth, the total concentration of ni-

trogen in the tailings, C<sub>6</sub>-dependent H.S.C. for bacterial growth, C<sub>7</sub>-dependent H.S.C. for bacterial growth, C<sub>8</sub>-dependent H.S.C. for bacterial growth and C<sub>10</sub>-dependent H.S.C. for bacterial growth as well as as the lag period (the time it takes before biodegradation of the compound starts) of these n-alkanes. In [38] MFT from Syncrude MLSB was amended with  $\sim 0.05\%$  in a micocosm and incubated for 252 days. Same as the above study, the degraded hydrocarbons were quantified. Within the course of the experiment, Toluene was rapidly biodegraded followed by o-Xylene and then m- plus p-xylene. Ethylbenzene was only slightly degraded within the incubation period. Using this data, we estimate Toluene, o-Xylene and m- plus p-xylene -dependent H.S.Cs. for bacterial growth.

[42] contains two set of investigations: one on the metabolism of 5 major iso-alkanes (3-methylhexane, 3-ethylhexane, 2-methylheptane, 4-methylheptane and 2-methylheptane) and 3 major cyclo-alkanes (ethylcycloalkanes, methylcyclohexane and ethylcyclohexane) in MFT collected from MLSB at a depth of 3 meters and the other at a depth of 6 meters. In each set, the MFT was amended with iso- and cyclo-alkanes and incubated for 1700 days. Apart from 3-ethyl hexane, all the iso-alkanes were completely biodegraded within this time period. The cylo-alkanes were only slightly degraded. The data recorded from the biodegradation of hydrocarbons in the MFT collected from a depth of 31 meters was used to estimate 3-methylhexane-, 2-methylheptane-, 4-methylheptane- and 2-methylheptane-dependent H.S.Cs. for bacteria growth and their lag periods.

To estimate C<sub>5</sub> dependent H.S.C. for bacterial growth and its lag period, we use the data in [34]. In this paper, MFT from Albian and CNR are amended

in separate microcosms with either a mixture of two ( $C_5$  and  $C_6$ ) or four ( $C_5$ ,  $C_6$ ,  $C_8$  and  $C_{10}$ ) alkanes and incubated under methanogenic conditions for 600 days. In the course of incubation, the concentration of the alkanes present in the microcosms at certain time were recorded. We use  $C_5$  dynamics data from the microcosm with Albian MFT.

For 2-methylpentane-dependent H.S.C. for bacteria growth we use 2-methylpentane dynamics data in [35]. This paper contains data on the metabolism of paraffinic solvent hydrocarbons to methane in CNR and Albian MFT. To generate the data, MFT from Albian and CNRL were spiked with 0.1% Paraffinic solvent and incubated under methanogenic conditions for 1600 days. Among the iso alkanes in paraffinic solvent, only 2-methyl pentane was completely degraded. The authors also recorded the generated methane values.

Lastly to estimate  $C_9$  related parameters, we use the data from the study in [33], performed to examine the rate of disappearance and methane production by individual compounds in naphtha added to Syncrude Tailings under methanogenic conditions. MFT collected from Syncrude Tailings was spiked with Syncrude naphtha and incubated under methanogenic conditions for 1040 days.  $C_9$  (initially 15 ppm) was completely depleted after 271 days.

We validate the model using methane biogenesis data from Albian MFT amended with paraffinic solvent reported in [35] (briefly summarized above) and methane biogenesis data from naphtha-amended CNRL MFT reported in [36]. Briefly, In [36], Shahimin et al., investigated the biotransformation of naphtha hydrocarbons to methane. To this end, the authors collected MFT from Albian and CNRL, spiked them in two separate microcosms with 0.2% CNRL naphtha and incubated them for 1600 days. The percentage disappear-

ance of individual naphtha compounds and methane production by different batches of tailings was recorded.

### 4.3.3 Mathematical analysis

We provide a basic analysis of System (4.1). First, we group all of the hydrocarbons together and represent them by one variable  $C$ . By doing so, we eliminated the effects of individual hydrocarbons. Also, we assume that the functions  $f(T_N)$  and  $g(C)$  are linear. For simplicity, we consider the total carbon biomass in the system,  $A = \frac{B}{r} + C$ , instead of the total carbon and that death bacteria are not recycled back to usable hydrocarbons. We also assume that  $\tau_i = 0, \forall i$ . Thus System (4.1) becomes:

$$\begin{aligned} \dot{A} &= \frac{r-1}{r}\varepsilon B + C^{\text{in}} = F(B) \\ \dot{B} &= \mu B \min \left\{ f(T - \theta B), g\left(A - \frac{B}{r}\right) \right\} - \varepsilon B = BG(A, B). \end{aligned} \quad (4.4)$$

For a basic mathematical analysis of the System (4.2), we construct the phase plane, i.e. a picture of the solution trajectories mapped out by points  $(A(t), B(t))$  as  $t$  varies over  $(-\infty, +\infty)$ . In particular we identify the steady state solutions i.e. solutions of  $\dot{A} = 0 = \dot{B}$  ( $F(B) = 0$  and  $G(A, B) = 0$ ). We call  $F(B) = 0$  and  $G(A, B) = 0$  (the lines on which trajectories are horizontal or vertical) the nullclines of System (4.4). The steady state solutions are the points where the nullclines (but not different branches of the same nullcline) cross. For the stability of the steady states, we compute the Jacobian matrix corresponding to each equilibrium point  $J(A^*, B^*)$ , where  $(A^*, B^*)$  is a given equilibrium point. We use the sign of the Trace and determinant of  $J(A^*, B^*)$  to deter-

mine the nature of the given equilibrium point. Let  $D = \det J(A^*, B^*)$  and  $T_r = \text{Trace} J(A^*, B^*)$ .

- 1) If  $D < 0$ , the eigenvalues of  $J(A^*, B^*)$  are real and of opposite signs, and the phase portrait is a saddle (which is always unstable).
- 2) If  $0 < D < \frac{T_r^2}{4}$ , the eigenvalues of  $J(A^*, B^*)$  are real, distinct, and of the same sign, and the phase portrait is a node, stable if  $T_r < 0$  and unstable if  $T_r > 0$ .
- 3) If  $0 < \frac{T_r^2}{4} < D$ , the eigenvalues of  $J(A^*, B^*)$  are neither real nor purely imaginary, and the phase portrait is a spiral, stable if  $T_r < 0$  and unstable if  $T_r > 0$ .

#### 4.3.4 Fitting the model to data

We use the non linear regression function *nlinfit(.)* in MATLAB to estimate the values of the initial bacteria in the system, the nitrogen half saturation constant, the total nitrogen available in the system, the half saturation constants of the biodegradable hydrocarbons and their lag periods. *nlinfit(.)* uses the Levenberg-Marquardt algorithm [28] to fit data. We provide the function with our empirical data, the time points at which the data was collected,  $X$ , our simulated results at  $X$ , and a random initial guess of parameter values. The system is integrated by calling a function that takes as input the initial parameter values, the time at which the empirical data was collected, and for a given time uses the MATLAB function *ode15s(.)* to perform the integration. The solution of the system obtained from the function is then evaluated at

$X$ , using the MATLAB function *deval(.)*. We also estimate the 95% confidence intervals of the predicted values. This was done using the MATLAB function *nlparci(.)*. To achieve this, we provided this function with the coefficient estimates, residuals and the estimated coefficient covariance matrix from *nlinfit(.)*. Some of the bacteria related model parameters use in simulation, namely:  $\mu$ ,  $r$ , and  $\theta$  were taken from literature. The units, values and source of these parameters are provided in the Table 4.1

Parameter	Definition	Value	Unit	References
$\mu$	Maximum growth rate of bacteria	1-4	day <sup>-1</sup>	[7, 10]
$r$	Yield constant	0.31-0.75	-	[63, 85]
$\theta$	N:C of bacteria	$\frac{1}{9}-\frac{1}{4}$	-	[82]

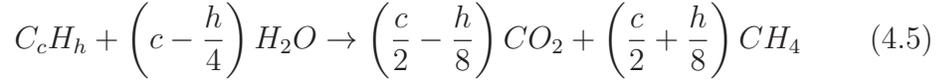
**Table 4.1:** Definition and values of some of the bacteria related parameters of System (4.2)

We assume that no bacteria died in the course of the experiments described above and thus in fitting the data to our model, we take  $d$  to be zero.

### 4.3.5 Maximum theoretical methane yield

In the experiments summarized above, after quantifying the methane and hydrocarbons in the microcosms at given time in the course of the experiment, theoretical calculations were made to compare the actual measured values in the microcosm and theoretical production from the hydrocarbons ( $\eta_i$ ).  $\Gamma_i$ —the expected CH<sub>4</sub> yield when one mole of  $C_i$  biodegrades is calculated using the stoichiometric equation (Equation 4.5 derived from Symons and Buswel

[46]) which describes the complete oxidation of hydrocarbons to  $\text{CO}_2$  and  $\text{CH}_4$  under methanogenic conditions.



A slight difference was observed between the theoretical and measured yield by the authors in [35, 36]. The naphtha-amended CNRL MFT methane production was 55 – 70% of the predicted theoretical maximum methane yield; whereas paraffinic solvent amended Albion MFT methane production was between 68 and 84% of the predicted theoretical maximum methane yield. In comparing our simulated values to measured values, we put into account these percentages. Hence in comparing our simulated methane values to measured values from paraffinic solvent amended Albion MFT methane production we choose  $\eta$  between 68% and 84%. In particular, we choose  $\eta$  to be 0.80. Similarly when comparing the output from our model to the methane produced from the decomposed hydrocarbon in naphtha-amended CNRL MFT, we choose  $\eta$  to lie between 55% and 70%. In particular, we choose  $\eta$  to be 0.65 in this case.

### 4.3.6 Model validation

We evaluate the validity of our model in predicting  $\text{CH}_4$  biogenesis from any given oil sand tailings by putting the estimated parameters in System (4.3) to generate methane data and then compare this generated methane data with measured methane data. The comparison is done using the *goodnessOfFit(.)* function in MATLAB. As inputs, we provided this function with our test data, the simulated data from our model and a cost function. The cost function

determines the goodness of fit. We use the Normalized Mean Square Error (NMSE) for this statistic. It is computed as

$$\text{NMSE} = 1 - \frac{\|[\text{actual}] - [\text{predicted}]\|^2}{\|[\text{actual}] - [\text{mean of actual}]\|^2},$$

where  $\|\cdot\|$  indicates the 2-norm of a vector,  $Y$  the simulated data by our model,  $\bar{Y}$  is the mean of the simulated data, and  $X$  is the test data.  $\text{NMSE} \in [-\infty, 1]$  where  $-\infty$  indicates a bad fit and 1 a perfect fit.

## 4.4 Results and discussion

### 4.4.1 Mathematical Analysis ( $f(T - \theta B)$ and $g(A - \frac{B}{r})$ linear)

Theorem (4) follows from the analysis that follows the theorem.

**Theorem 4.** *System (4.4) may have 0,1 or an infinite number of equilibrium points depending on the concentration of the fresh hydrocarbon input  $C^{in}$ .*

- a) *At OSTP, if  $C^{in}$  is greater than a certain threshold value given by  $C_0 = \left(T - \frac{\varepsilon K_f}{\mu}\right) \frac{\varepsilon(1-r)}{\theta r}$ , System (4.4) will have no equilibrium point. If  $C^{in} = C_0$ , it will have one unique equilibrium point  $E_1 = \left(\frac{\mu C^{in} + \varepsilon^2 K_g(1-r)}{\varepsilon(1-r)\mu}, \frac{r C^{in}}{\varepsilon(1-r)}\right)$  which is unstable. If  $C^{in} < C_0$ , it will have an infinite number of equilibrium points given by*

$$E_2 = \left\{ \left( A, \left( T - \frac{\varepsilon k_f}{\mu} \right) \frac{1}{\theta} \right) : A > \frac{1}{r} \left[ \left( T - \frac{\varepsilon k_f}{\mu} \right) \frac{1}{\theta} + \frac{\varepsilon k_g r}{\mu} \right] \right\}$$

*which are locally asymptotically stable.*

b) At EPLs, System (4.4) have an infinite number of equilibrium points

$$E_3 = (A, 0) : A \geq 0.$$

If  $A < \frac{\varepsilon K_g}{\mu}$  (i.e. if carbon is limiting), these equilibrium points will be locally asymptotically stable and if  $A \geq \frac{\varepsilon K_g}{\mu}$ , they will be unstable.

**Dynamics at OSTP ( $C^{\text{in}} \neq 0$ ):**

**Steady states:**

**A-Nullclines:**

$$\dot{A} = 0 \implies B = \frac{rC^{\text{in}}}{\varepsilon(1-r)}.$$

**B-Nullclines:**

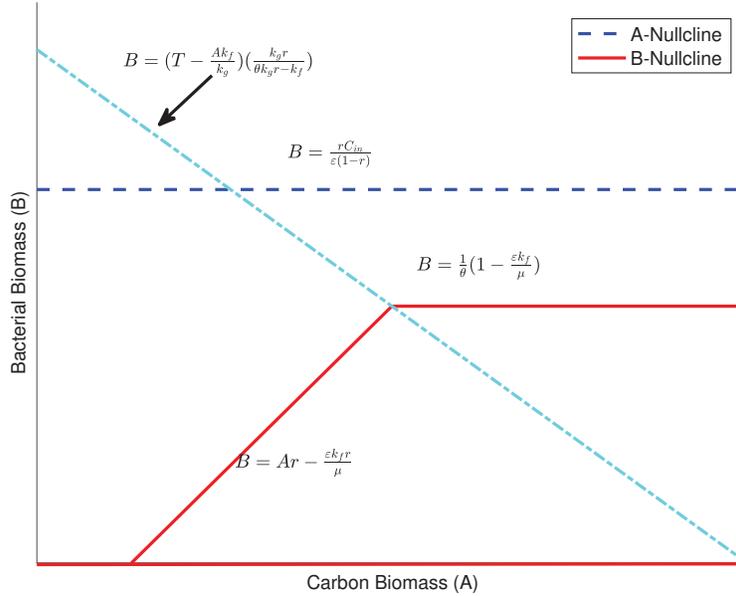
$$\dot{B} = 0 \implies B = 0 \text{ or } G(A, B) = 0.$$

$$G(A, B) = 0 \implies \begin{cases} B = Ar - \frac{\varepsilon k_g r}{\mu} & \text{if } \frac{T - \theta B}{k_f} > \frac{A - \frac{B}{r}}{k_g} \\ B = (T - \frac{\varepsilon k_f}{\mu}) \frac{1}{\theta} & \text{if } \frac{T - \theta B}{k_f} < \frac{A - \frac{B}{r}}{k_g} \end{cases}$$

**Case 1:** Suppose  $\theta - \frac{k_f}{k_g r} > 0$ , then

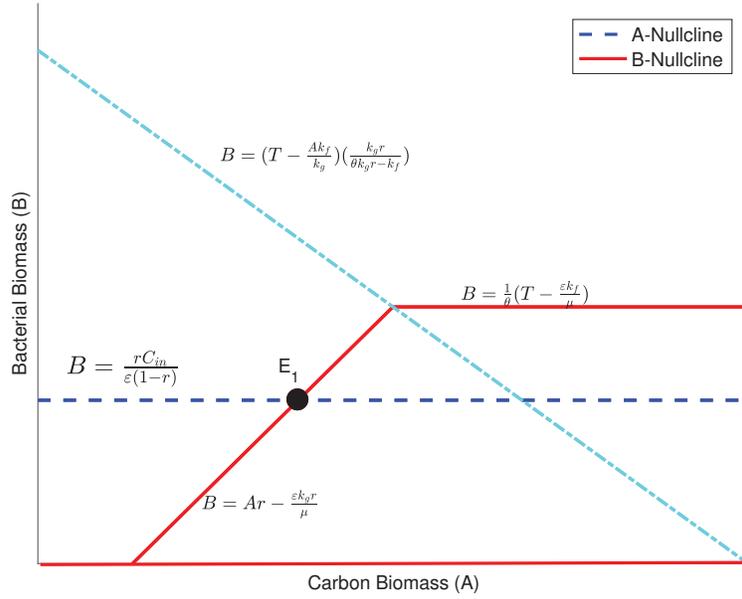
$$G(A, B) = 0 \implies \begin{cases} B = Ar - \frac{\varepsilon k_g r}{\mu} & \text{if } B < \left(T - \frac{Ak_f}{k_g}\right) \left(\frac{k_g r}{\theta k_g r - k_f}\right) \\ B = \left(T - \frac{\varepsilon k_f}{\mu}\right) \frac{1}{\theta} & \text{if } B > \left(T - \frac{Ak_f}{k_g}\right) \left(\frac{k_g r}{\theta k_g r - k_f}\right) \end{cases}$$

**Case 1.1:** If  $\frac{rC^{\text{in}}}{\varepsilon(1-r)} > \left(T - \frac{\varepsilon k_f}{\mu}\right) \frac{1}{\theta}$  i.e  $\theta > \frac{\varepsilon(1-r)}{rC^{\text{in}}} \left(T - \frac{\varepsilon k_f}{\mu}\right)$ , then there will be no intersection between the A and B-nullclines as shown in Figure 4.1. Hence the system will have no equilibrium point.



**Figure 4.1:** The nullclines for  $\theta - \frac{k_f}{k_g r} > 0$  and  $\theta > \frac{\varepsilon(1-r)}{rC^{\text{in}}} \left(T - \frac{\varepsilon k_f}{\mu}\right)$ .

**Case 1.2:** If  $\frac{rC^{\text{in}}}{\varepsilon(1-r)} < \left(T - \frac{\varepsilon k_f}{\mu}\right) \frac{1}{\theta}$  i.e  $\theta < \frac{\varepsilon(1-r)}{rC^{\text{in}}} \left(T - \frac{\varepsilon k_f}{\mu}\right)$ , then the two nullclines will intersect at one unique point  $E_1 = \left(\frac{\mu C^{\text{in}} + \varepsilon^2 k_g(1-r)}{\varepsilon(1-r)\mu}, \frac{rC^{\text{in}}}{\varepsilon(1-r)}\right)$  as shown in Figure 4.2. Hence the system will have one equilibrium point  $E_1$ , which is internal.



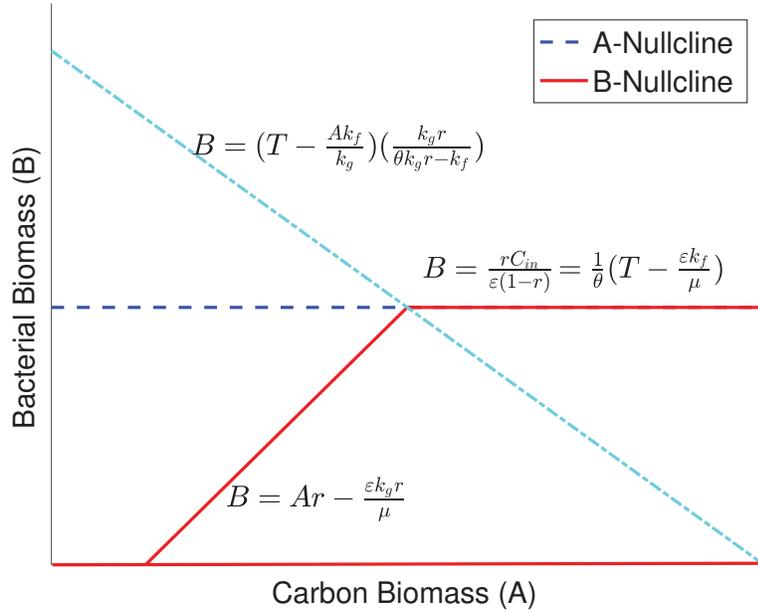
**Figure 4.2:** The nullclines for  $\theta - \frac{k_f}{k_g r} > 0$  and  $\theta < \frac{\varepsilon(1-r)}{rC^{\text{in}}} \left(T - \frac{\varepsilon k_f}{\mu}\right)$ .

**Case 1.3:** If  $\frac{rC^{\text{in}}}{\varepsilon(1-r)} = \left(T - \frac{\varepsilon k_f}{\mu}\right) \frac{1}{\theta}$  i.e  $\theta = \frac{\varepsilon(1-r)}{rC^{\text{in}}} \left(T - \frac{\varepsilon k_f}{\mu}\right)$ , then the two nullclines will intersect on the line

$$\left\{ \left( A, \left( T - \frac{\varepsilon k_f}{\mu} \right) \frac{1}{\theta} \right) : A > \frac{1}{r} \left[ \left( T - \frac{\varepsilon k_f}{\mu} \right) \frac{1}{\theta} + \frac{\varepsilon k_g r}{\mu} \right] \right\}$$

as shown in Figure 4.3. Thus the System (4.4) will have an infinite number of equilibrium points

$$E_2 = \left\{ \left( A, \left( T - \frac{\varepsilon k_f}{\mu} \right) \frac{1}{\theta} \right) : A > \frac{1}{r} \left[ \left( T - \frac{\varepsilon k_f}{\mu} \right) \frac{1}{\theta} + \frac{\varepsilon k_g r}{\mu} \right] \right\}$$



**Figure 4.3:** ]

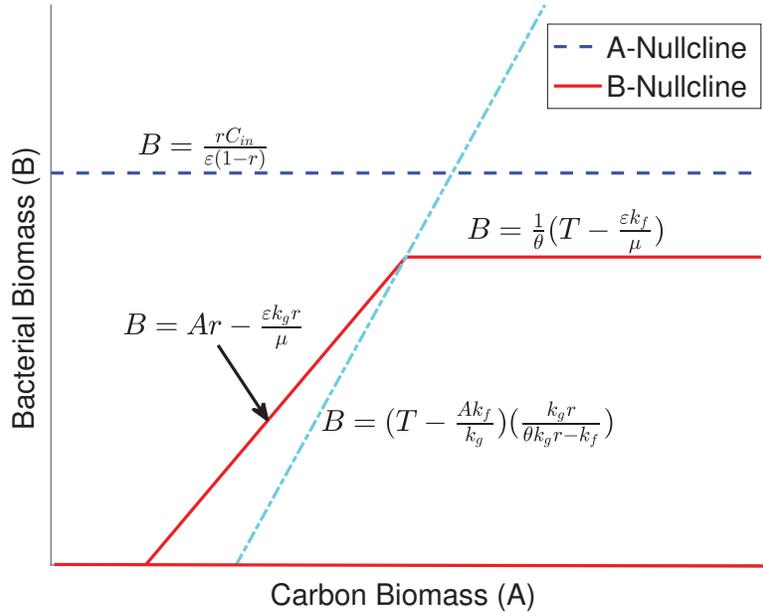
The nullclines for  $\theta - \frac{k_f}{k_g r} > 0$  and  $\theta = \frac{\varepsilon(1-r)}{rC^{\text{in}}} \left( T - \frac{\varepsilon k_f}{\mu} \right)$ .

**Case 2:** Suppose  $\theta - \frac{k_f}{k_g r} < 0$ , then

$$G(A, B) = 0 \implies \begin{cases} B = Ar - \frac{\varepsilon k_g r}{\mu} & \text{if } B > \left( T - \frac{Ak_f}{k_g} \right) \left( \frac{k_g r}{\theta k_g r - k_f} \right) \\ B = \left( T - \frac{\varepsilon k_f}{\mu} \right) \frac{1}{\theta} & \text{if } B < \left( T - \frac{Ak_f}{k_g} \right) \left( \frac{k_g r}{\theta k_g r - k_f} \right) \end{cases}$$

Note that the slope of the line  $B = Ar - \frac{\varepsilon k_g r}{\mu}$  is less than that of the line  $B = \left(T - \frac{Ak_f}{k_g}\right) \left(\frac{k_g r}{\theta k_g r - k_f}\right)$ , since  $\frac{k_f}{k_f - \theta k_g r} > 1$ . Therefore, the point where the line  $B = Ar - \frac{\varepsilon k_g r}{\mu}$  intersects the A-axis,  $\frac{\varepsilon k_g}{\mu}$  must be less than that of the line  $B = \left(T - \frac{Ak_f}{k_g}\right) \left(\frac{k_g r}{\theta k_g r - k_f}\right)$ ,  $\frac{T k_g}{k_f}$ , for the two lines to intersect on the first quadrant.

**Case 2.1:** If  $\frac{rC^{\text{in}}}{d(1-r)} > \left(T - \frac{\varepsilon k_f}{\mu}\right) \frac{1}{\theta}$  i.e  $\theta > \frac{d(1-r)}{rC^{\text{in}}} \left(T - \frac{\varepsilon k_f}{\mu}\right)$  then as before, there will be no intersection between the A and B-nullclines as shown in Figure 4.4. Hence the system will have no equilibrium point.

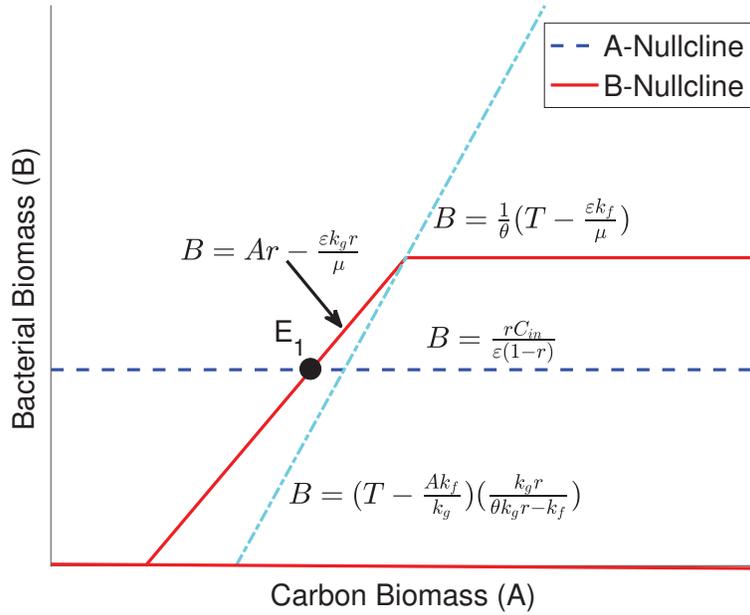


**Figure 4.4:** The nullclines for  $\theta - \frac{k_f}{k_g r} < 0$  and  $\theta > \frac{\varepsilon(1-r)}{rC^{\text{in}}} \left(T - \frac{\varepsilon k_f}{\mu}\right)$ .

**Case 2.2:** If  $\frac{rC^{\text{in}}}{d(1-r)} < \left(T - \frac{\varepsilon k_f}{\mu}\right) \frac{1}{\theta}$  i.e  $\theta < \frac{\varepsilon(1-r)}{rC^{\text{in}}} \left(T - \frac{\varepsilon k_f}{\mu}\right)$ , then two nullclines will intersect at one unique point

$$E_1 = \left( \frac{\mu C^{\text{in}} + \varepsilon^2 k_g (1-r)}{\varepsilon(1-r)\mu}, \frac{rC^{\text{in}}}{\varepsilon(1-r)} \right)$$

as show in Figure 4.5. Hence the system will have one equilibrium point  $E_1$ , which is internal.



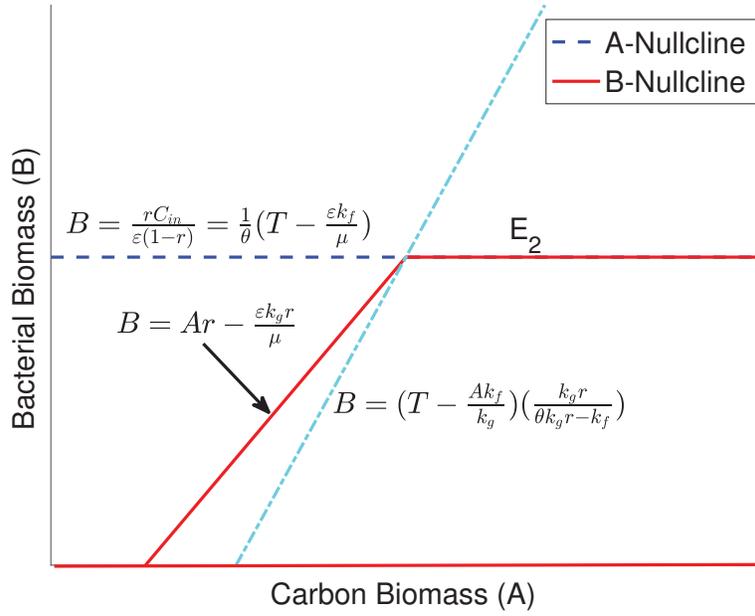
**Figure 4.5:** The nullclines and Equilibrium for  $\theta - \frac{k_f}{k_g r} < 0$  and  $\theta < \frac{\varepsilon(1-r)}{rC^{\text{in}}} \left(T - \frac{\varepsilon k_f}{\mu}\right)$ .

**Case 2.3:** If  $\frac{rC^{in}}{\varepsilon(1-r)} = \left(T - \frac{\varepsilon k_f}{\mu}\right) \frac{1}{\theta}$  i.e  $\theta = \frac{\varepsilon(1-r)}{rC^{in}} \left(T - \frac{\varepsilon k_f}{\mu}\right)$ , then the two nullclines will intersect on the line

$$\left\{ \left( A, \left( T - \frac{\varepsilon k_f}{\mu} \right) \frac{1}{\theta} \right) : A > \frac{1}{r} \left[ \left( T - \frac{\varepsilon k_f}{\mu} \right) \frac{1}{\theta} + \frac{\varepsilon k_g r}{\mu} \right] \right\}$$

as shown in Figure 4.6. Thus the System (4.4) will have an infinite number of equilibrium points

$$E_2 = \left\{ \left( A, \left( T - \frac{\varepsilon k_f}{\mu} \right) \frac{1}{\theta} \right) : A > \frac{1}{r} \left[ \left( T - \frac{\varepsilon k_f}{\mu} \right) \frac{1}{\theta} + \frac{\varepsilon k_g r}{\mu} \right] \right\}$$



**Figure 4.6:** The nullclines and Equilibria for  $\theta - \frac{k_f}{k_g r} < 0$  and  $\theta = \frac{\varepsilon(1-r)}{rC^{in}} \left(T - \frac{\varepsilon k_f}{\mu}\right)$ .

### Stability analysis:

To determine the local stability of the equilibria above, we consider the Jacobian matrix of System (4.4),

$$J(A, B) = \begin{pmatrix} 0 & \frac{(r-1)d}{r} \\ BG_A(A, B) & G(A, B) + BG_B(A, B) \end{pmatrix} \quad (4.6)$$

Where

$$G(A, B) = \begin{cases} \frac{\mu \left( A - \frac{B}{r} \right)}{k_g} - \varepsilon & \text{if } \frac{T - \theta B}{k_f} > \frac{A - \frac{B}{r}}{k_g} \\ \frac{\mu(T - \theta B)}{k_f} - \varepsilon & \text{if } \frac{T - \theta B}{k_f} < \frac{A - \frac{B}{r}}{k_g} \end{cases}$$

$$G_A(A, B) = \begin{cases} \frac{\mu}{k_g} & \text{if } \frac{T - \theta B}{k_f} > \frac{A - \frac{B}{r}}{k_g} \\ 0 & \text{if } \frac{T - \theta B}{k_f} < \frac{A - \frac{B}{r}}{k_g} \end{cases}$$

$$G_B(A, B) = \begin{cases} \frac{-\mu}{k_g r} & \text{if } \frac{T - \theta B}{k_f} > \frac{A - \frac{B}{r}}{k_g} \\ \frac{-\theta \mu}{k_f} & \text{if } \frac{T - \theta B}{k_f} < \frac{A - \frac{B}{r}}{k_g} \end{cases}$$

**Stability of  $E_1$  :**

$$J(E_1) = \begin{pmatrix} 0 & \frac{(r-1)d}{r} \\ \frac{\mu r C^{\text{in}}}{k_g \varepsilon (1-r)} & -\frac{r C^{\text{in}} \mu}{\varepsilon (1-r) r k_g} \end{pmatrix} \quad (4.7)$$

$$\det(J(E_1)) = \left( \frac{\mu r C^{\text{in}}}{k_g \varepsilon (1-r)} \right) \left( \frac{(r-1)\varepsilon}{r} \right) < 0$$

Since  $\det(J(E_1))$  is negative, the eigen values have different signs. Thus  $E_1$  is a saddle and hence unstable.

**Stability of  $E_2$  :**

$$J(E_2) = \begin{pmatrix} 0 & \frac{(r-1)\varepsilon}{r} \\ 0 & -\frac{r C^{\text{in}} \mu \theta}{\varepsilon (1-r) k_f} \end{pmatrix} \quad (4.8)$$

$$\det(J(E_2)) = 0 \text{ and } T_r(J(E_2)) = -\frac{r C^{\text{in}} \mu \theta}{\varepsilon (1-r) k_f} < 0.$$

Since the  $T_r(J(E_2))$  is negative and  $\det(J(E_2))$  is zero, one eigenvalue is zero and the other is negative. Thus  $E_2$  is a line of locally asymptotically stable equilibrium points. Hence the internal equilibrium point  $E_1$  is a unstable and the line of Equilibrium points  $E_2$  is locally asymptotically stable.

**Dynamics at EPLs ( $C^{\text{in}} = 0$ )**

**Steady states:**

**A-Nullclines:**

$$\dot{A} = 0 \implies B = \frac{r C^{\text{in}}}{\varepsilon (1-r)}.$$

**B-Nullclines:**

$$\dot{B} = 0 \implies B = 0 \text{ or } G(A, B) = 0.$$

These nullclines have an infinite number of infinite number of intersections given by

$$E_3 = (A, 0) : A \geq 0$$

Hence at EPLs , System (4.4) have an infinite number of equilibrium points given by  $E_3$ .

**Stability of  $E_3$ :**

$$J(E_3) = \begin{pmatrix} 0 & \frac{(r-1)d}{r} \\ 0 & \frac{\mu A}{K_g} - \varepsilon \end{pmatrix} \quad (4.9)$$

$\det J(E_3) = 0$  and  $T_r J(E_3) = \frac{\mu A}{K_g} - \varepsilon$ .

If  $A < \frac{\varepsilon K_g}{\mu}$ ,  $T_r J(E_3)$  will be less than zero. Hence  $E_3$  will be asymptotically stable. On the other hand, if  $A \geq \frac{\varepsilon K_g}{\mu}$ , then  $T_r J(E_3)$  will be greater than zero and in this case,  $E_3$  will be a line of unstable equilibrium points.

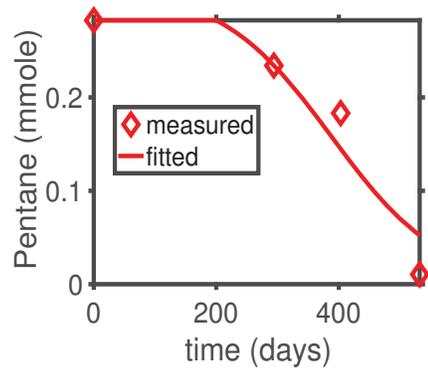
**4.4.2 Fitting the model to data**

Table (4.2) contains the 95 % confidence interval of the parameters obtained from the procedure in Section (4.3.4) to fit the model formulated in Section 4.3.1 to the data discussed in Section (4.3.2). The Table column values contain the values of the parameter used to simulate data in the figures below. Figures (4.7, 4.11b, 4.9) show the simulated concentration of  $C_5$ ,  $C_6$ ,  $C_7$ ,  $C_8$ ,  $C_9$ ,

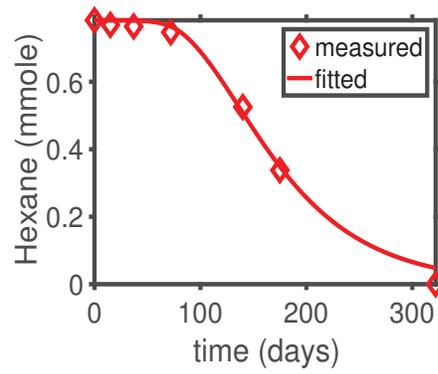
$C_{10}$ , toluene, o-Xylene, m-plus-p-Xylene, 2-methylpentane, 3-methylhexane, 2-methylheptane, 4-methylheptane, and 2-methyloctane along with measured concentration of these hydrocarbons within the given time period. In Table (4.2),  $K_f$  represents the nitrogen-dependent H.S.Cs. for bacterial growth,  $T_N$  the total nitrogen available in the system,  $K_{g_{C_5}}, K_{g_{C_6}}, K_{g_{C_7}}, K_{g_{C_8}}, K_{g_{C_9}}, K_{g_{C_{10}}}, K_{g_{3-MC_6}}, K_{g_{2-MC_7}}, K_{g_{4-MC_7}}, K_{g_{2-MC_8}}, K_{g_{2-MC_5}}$  respectively represents  $C_5$ -,  $C_6$ -,  $C_7$ -,  $C_8$ -,  $C_9$ -,  $C_{10}$ -, 3-MC<sub>6</sub>-, 2-MC<sub>7</sub>-, 4-MC<sub>7</sub>-, 2-MC<sub>8</sub>-, 2-MC<sub>5</sub>- dependent H.S.C. for bacterial growth, and  $Z$ -lag denote the lag period of  $Z$ , where  $Z$  is either  $C_5, C_6, C_7, C_8, C_9, C_{10}, 3-MC_6, 2-MC_7, 4-MC_7, 2-MC_8$  or  $2-MC_5$

Parameter	value	95 % confidence interval	Unit
$B(0)$	0.0004	0.0001-0.0138	mmoleC
$K_f$	0.2622	0.2622	mmole
$T_N$	327.6158	327.6158-327.6158	mmole
$K_{g_{C_5}}$	56.2954	16.1585-96.4324	mmole
$K_{g_{C_6}}$	430.2916	366.0636-494.5195	mmole
$K_{g_{C_7}}$	270.7338	238.9429-302.5247	mmole
$K_{g_{C_8}}$	90.1344	69.3197-110.9490	mmole
$K_{g_{C_9}}$	0.8187	0.6751-0.9623	mmole
$K_{g_{C_{10}}}$	12.0212	10.1740-13.8685	mmole
$K_{g_{\text{toluene}}}$	4.4802	4.1355-4.8248	mmole
$K_{g_{\text{mp-xylenes}}}$	85.0703	76.9272-93.2133	mmole
$K_{g_{\text{o-xylene}}}$	17.4971	14.1566-20.8375	mmole
$K_{g_{3-MC_6}}$	144.6140	102.7117-186.5163	mmole
$K_{g_{2-MC_7}}$	320.4133	183.77548-457.05123	mmole
$K_{g_{4-MC_7}}$	170.3471	121.01706-219.6771	mmole
$K_{g_{2-MC_8}}$	335.8911	179.09790-492.6843	mmole
$K_{g_{2-MC_5}}$	165.94	130.2141-201.6662	mmole
$C_5$ -lag	200	200	days
$C_6$ -lag	26	26	days
$C_7$ -lag	60	40-80	days
$C_8$ -lag	60	60	days
$C_9$ -lag	70	70	days
$C_{10}$ -lag	5	5	days
Touene-lag	30	30	days
m,p-Xylenes-lag	70	70	days
o-Xylene-lag	60	60	days
3-MC <sub>6</sub> -lag	25	24.9820-25.0180	days
2-MC <sub>7</sub> -lag	25	25	days
4-MC <sub>7</sub> -lag	25	25	days
2-MC <sub>8</sub> -lag	25	25	days
2-MC <sub>5</sub> -lag	23	23	days

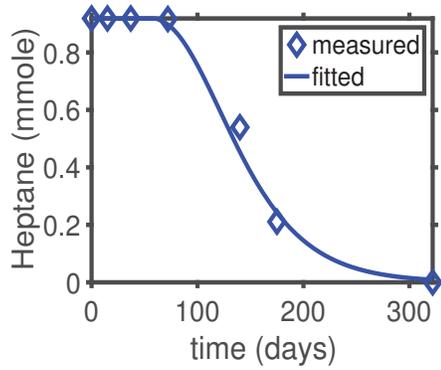
**Table 4.2:** Parameter values of the biodegradation model



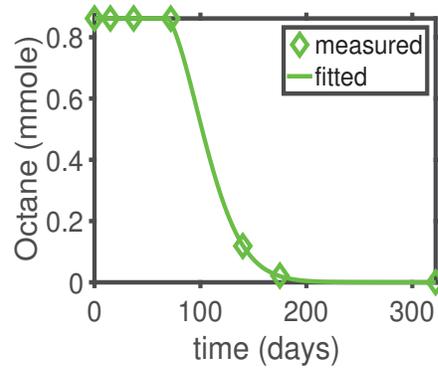
(a)  $C_5$



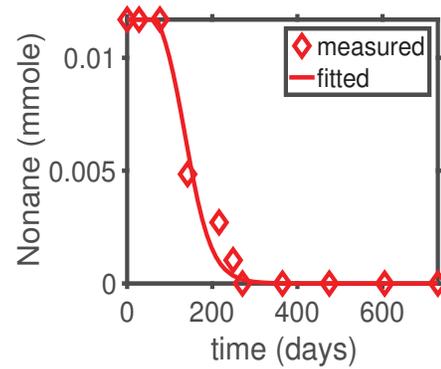
(b)  $C_6$



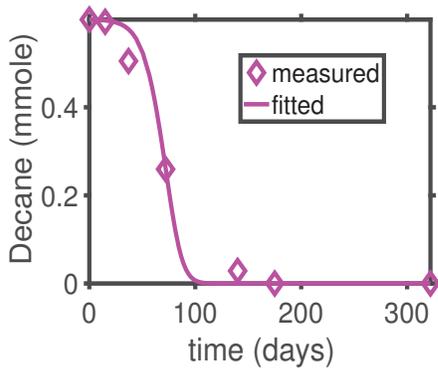
(c)  $C_7$



(d)  $C_8$

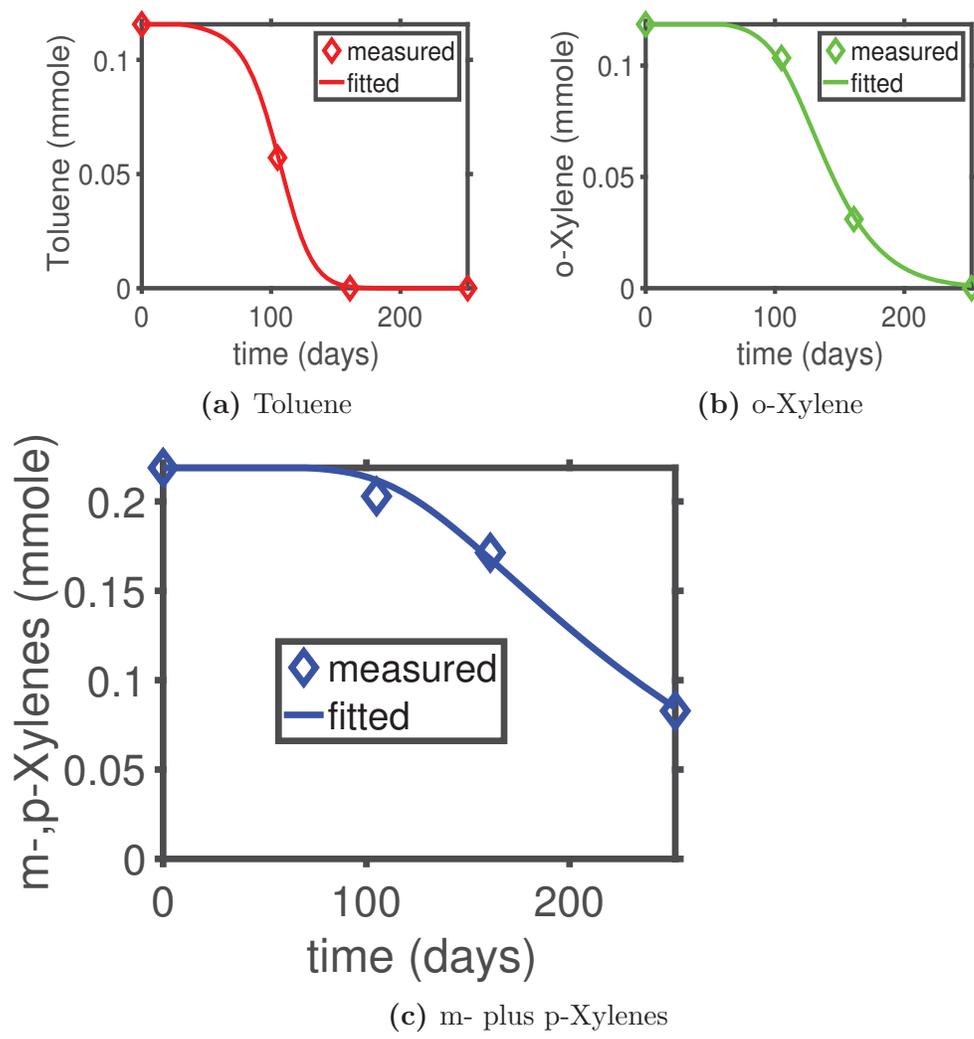


(e)  $C_9$

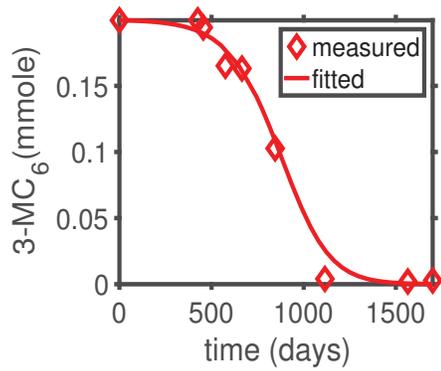


(f)  $C_{10}$

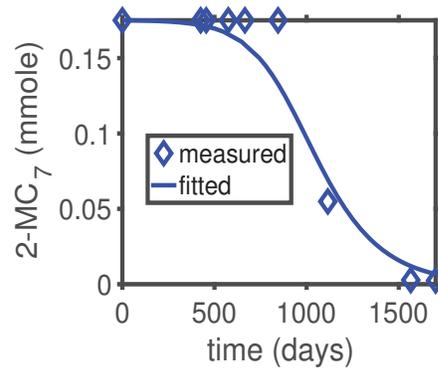
**Figure 4.7:** System(4.2) fit to measured alkane biodegradation values. Solid lines represent model predictions and diamonds denote measured values.



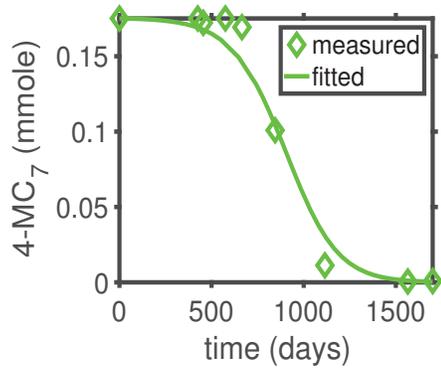
**Figure 4.8:** System(4.2) fit to measured biodegradable BTEX compounds data. Solid lines represent model predictions and diamonds denote measured values.



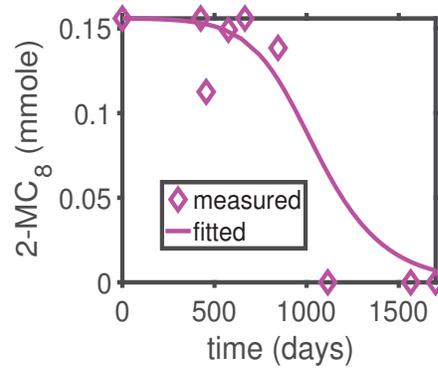
(a) 3-methyl hexane (3-MC<sub>6</sub>)



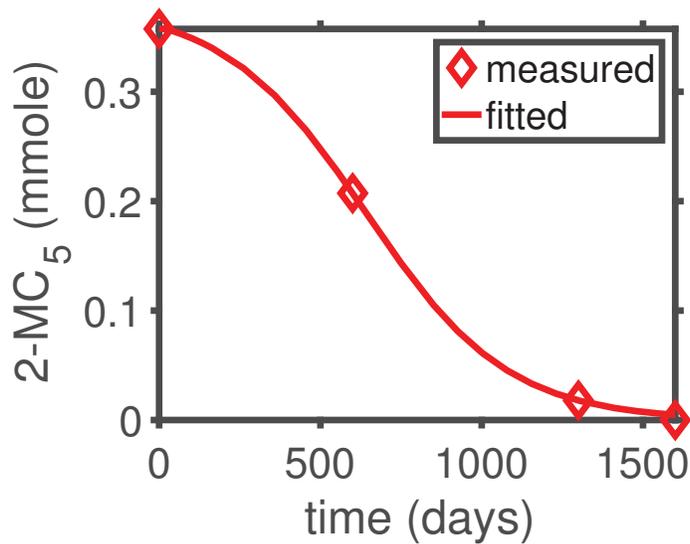
(b) 2-methyl heptane (2-MC<sub>7</sub>)



(c) 4-methyl heptane (4-MC<sub>7</sub>)

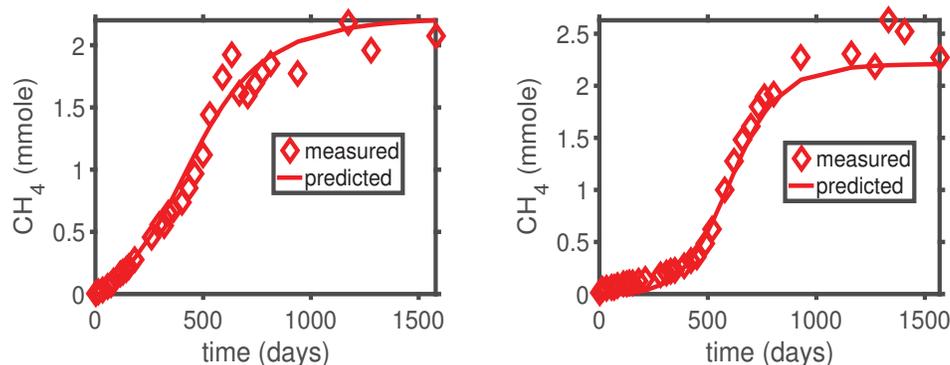


(d) 2-methyl Octane (2-MC<sub>8</sub>)



(e) 2-methyl pentane (2-MC<sub>5</sub>)

**Figure 4.9:** System(4.2) fit to measured biodegradable isoalkenes' data. Solid lines represent fitted values and diamonds denote measured values.



(a) Methane generation from biodegradable naphtha-amended CNRL MFT hydrocarbons: comparison between simulated and measured data.  $\mu = 2$ ,  $\theta = 0.2$ ,  $d = 0$ ,  $r = 0.35$  and  $B(0) = 0.00057$ . The other parameter values are the same as those on the value column of Table (4.2)

(b) Methane generation from biodegradable paraffinic solvent amended Albion MFT hydrocarbons: comparison between simulated and measured data.  $\mu = 2$ ,  $\theta = 0.2$ ,  $d = 0$ ,  $r = 0.35$  and  $B(0) = 0.00329$ . The other parameter values are the same as those on the value column of able (4.2)

**Figure 4.10:** Comparison between simulated methane data and measured methane data. Solid lines represent model predictions and diamonds denote measured values

### 4.4.3 Model validation

Using the parameter values in Table (4.2) with the procedure and methane data described in Sections 4.3.6 and (4.3.2) respectively, we assess the validity of our model in predicting methane biogenesis from any given tailing ponds. Figure (4.10) shows a comparison between our simulated and measured methane data. The goodness-of-fit statistics suggests that System (4.3) with the given parameter values, is a good fit for methane data (NMSE = 0.98 for the comparison with naphtha amended CNRL MFT methane data (Figure (4.11d)) and 0.97 for the comparison with paraffinic solvent amended Albion MFT methane data (Figure (4.10b))). By direct implication, System (4.2) with the given parameter values is a good fit for biodegradation data.

#### 4.4.4 Discussion

We have developed a generic stoichiometry-based microbial metabolism model. This model integrates a key aspect of microbial community (having more than one limiting elements), rarely considered in biodegradation models. This is in agreement with the experimental studies in [32, 3, 49, 52, 17, 30] that show that the degradation rates of hydrocarbons depend also on microbial community among other factors. Though we considered that only two elements are limiting (namely: nitrogen and carbon), the idea can easily be extended to a situation in which more than one element is limiting. Having assumed that the total nitrogen in the tailing pond is constant, to keep track of that available for microbial metabolism, we employed ecological stoichiometry. It postulates a crucial relationship between the balance of elements, typically but not limited to carbon (C), nitrogen (N) and phosphorus (P) in an environment, and their role in determining growth and reproduction of microbes as well as in ecological interactions [45, 20].

While the model provides reasonable goodness of fits to both test data, there are rather obvious areas for model improvement. In deriving the biodegradation model, we assume that only nitrogen and carbon are limiting in the ponds/EPLs. This model also groups all the bacteria into a single group (assumes bacteria population is homogenous). These assumptions seem ecologically plausible, but need not hold in all ponds/EPLs. Different groups of bacteria specializes in biodegrading defined hydrocarbons, based on the complexity of the hydrocarbon. A logical development of the model would be to incorporate other limiting elements and separate the dynamics of the differ-

ent groups of bacteria. This is a topic of our further investigations. A case can also be established against our lag phase. In [24] it is reported that the lag period is a function of the prior exposure of the microbial community to hydrocarbons. This implies that the lag period vary from OSTP to EPL for each hydrocarbon, with the highest values at the OSTP and that of the hydrocarbons at EPL  $\sim 0$ . Since the bacteria in our experiments did not have much prior exposure to the hydrocarbons added to it, our empirical data represent what happens at OSTP. To use our model in EPL, it will be reasonable to take all the lag period to be 0. Our model equally have some room for improvement when it comes to the interaction of bacteria and the hydrocarbons. The population of bacteria in the tailing ponds are heavily grazed by grazers such as protozoans, saprophagous and nematodes [76]. Empirical evidence suggests that the rate or the extent of hydrocarbon decomposition often increases in the presence of grazers that substantially reduce the abundance of bacteria [62, 74, 78, 69]. This could be explained using nutrients recycling hypothesis. Nutrients recycling hypothesis suggests that excretion of mineral nutrients by protists results in stimulated physiological status of bacteria and thus enhanced usage of carbon by bacteria [8]. Thus, incorporating grazers dynamics in the model might lead to a more efficient results.

In [39] we use the hydrocarbon degradation and methane generation data in [37] to estimate  $C_6$ ,  $C_7$ ,  $C_8$  and  $C_{10}$  zero- and first-order methane biogenesis models' related parameters. We equally use the hydrocarbon degradation and methane generation data in [38] to estimate Toluene, o-Xylene and m- plus p-xylene zero- and first-order methane biogenesis models' related parameters. As mentioned above, the same data set is used in this paper to estimate  $C_6$ ,  $C_7$ ,  $C_8$

$C_{10}$ , Toluene, o-Xylene and m- plus p-xylene related parameters of our model. Here, we use these parameter values in their respective models to simulate the cumulative methane expected to be released from the decomposition of these hydrocarbons in the course of the various experiments and compare the results with measured methane values. We equally estimate the zero- and first- order models related parameters for pentane, nonane, 2-methylpentane, 3-methylhexane, 2-methylheptane, 4-methylheptane, and 2-methyloctane (see Table 4.3).

Hydrocarbon (mmole)	Lag phase	0th-order parameter (mmole/day)	1st-order parameter (/day)
$C_5$	294	0.0008576	0.01117
3-MC <sub>6</sub>	455	0.0001816	0.003849
2-MC <sub>7</sub>	845	0.00023	0.005258
4-MC <sub>7</sub>	665	0.0001936	0.005663
2-MC <sub>8</sub>	665	0.00017726	0.0006584
2-MC <sub>5</sub>	600	0.00022816	0.0003501
$C_9$	77	2.664e-05	0.01276

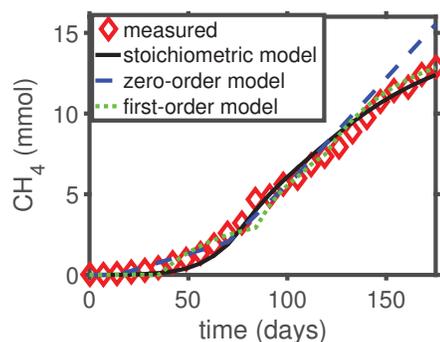
**Table 4.3:** Zero-and first-order model related parameters values for some of the biodegradable hydrocarbons

Using the NMSE as a measure of goodness of fit, we compare the performance of the three models using as test data the measured methane values from alkane amended Syncrude MFT, BTEX amended Syncrude MFT, Naptha amended CNRL MFT and Paraffinic solvent amended Albion MFT reported in [37], [38], [36], [35] respectively. Table (4.4) contains the NMSE obtained from comparing the performance of each of the models against the measured methane values in the experiments.

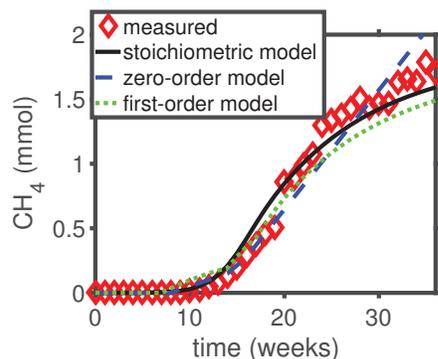
	Zero-order	First-order	Stoichiometric
Alkanes amended methane	0.95	0.99	0.99
Methane data from BTEX	0.96	0.95	0.98
Methane data from paraffinic solvent	-1.10	0.61	0.97
Methane data from naphtha	-1.00	0.82	0.98

**Table 4.4:** Comparing the performance of our three models using measured methane data from alkanes, BTEX, paraffinic solvents and naphtha hydrocarbons

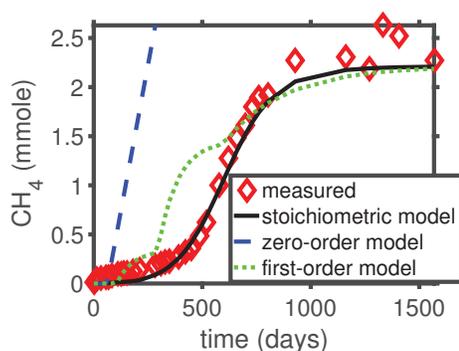
The stoichiometric model has the highest NMSE for all the data. This shows that the simulated results from the stoichiometric model match the training data perfectly well compared to the other models. Figures (4.11a), (4.11b), (4.11c) and (4.11d), respectively contains a plot of the measured and simulated values (by all the three models) for alkane amended Syncrude MFT, BTEX amended Syncrude MFT, Paraffinic solvent amended Albion MFT, and Naptha amended CNRL MFT within the duration of the given experiments.



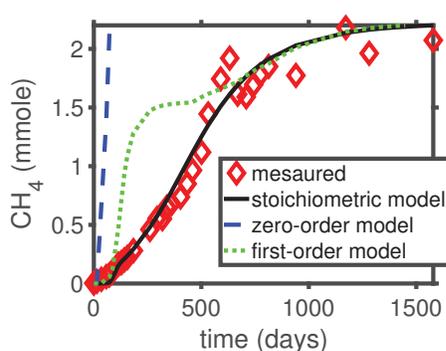
(a) Methane generation from biodegradable alkane amended Syncrude MFT.



(b) Methane generation from BTEX amended Syncrude MFT.



(c) Methane generation from paraffinic solvent amended Albian MFT.



(d) Methane generation from naphthenic solvent amended CNRL MFT.

**Figure 4.11:** Comparison between simulated (by all our three models) and measured methane data. Solid lines represent model predictions and diamonds denote measured values.

## 4.5 Conclusions

A number of goals have been successfully achieved in this chapter. The consideration of the interactions between microbial biomass, limited resources in the OSTP/EPLs and the biodegradation of OSTP/EPLs hydrocarbons in this model allows for the estimation of biodegradation model parameter values (methane biogenesis model parameter values) that are valid for any given

OSTP/EPL. These interactions includes the regulation of OSTP/EPLs hydrocarbons biodegradation by the quantity of microbial biomass and the regulation of microbial growth by the concentration of hydrocarbons and available nitrogen in OSTP/EPLs. After fitting the model to data, we investigated the capabilities of the model in predictions and compared it performance to that of other existing models in the literature. In particular our findings suggest that the present model can be used in 1) predicting the quantity of  $\text{CH}_4$  produced at a certain time in any given OSTP/EPLs, 2) calculating the time required to produce a certain quantity of cumulative  $\text{CH}_4$  from any given OSTP/EPLs, 4) estimating how long it will take before a given OSTP/EPLs will stop emitting GHGs. This can help oil sands operators manage their tailings repositories and evaluate the quality of their EPL system.

# Chapter 5

## Stability and sensitivity analysis of the iSIR model for indirectly transmitted infectious diseases with immunological threshold <sup>2</sup>

### Abstract

Most pathogenic diseases remain epidemic and endemic in the world, causing thousands of deaths annually in less developed countries. Yet, their dynamics are still not fully understood. In this paper, we carry out a thorough stability and sensitivity analysis of an iSIR which incorporates an infection term that

---

<sup>2</sup>This Chapter has been published. Reference: Kong, Jude D., William Davis, Xiong Li, and Hao Wang. “Stability and Sensitivity Analysis of the iSIR Model for Indirectly Transmitted Infectious Diseases with Immunological Threshold.” *SIAM Journal on Applied Mathematics* 74, no. 5 (2014): 1418-1441.

explicitly includes a Minimum Infection Dose (MID), and determine an invariant domain. We discover that if the MID (denoted  $c$ ) is less than bacterial carrying capacity  $K$ , we may have two steady states: the endemic or epidemic steady state and the disease free and bacteria free steady state. The latter is unstable and the previous is globally stable under a certain condition. On the other hand, if  $c \geq K$ , then up to four steady states may exist: an unstable endemic steady state, a locally stable endemic steady state, a conditionally globally stable disease free steady state and an unstable disease free and bacteria free steady state. We find that to control the period and intensity of the outbreaks, it might be better to focus on the bacterial carrying capacity rather than on the shedding rates.

**Keywords:** indirect transmission, immunological threshold, infectious diseases, shedding, global stability analysis, sensitivity analysis.

## 5.1 Introduction

Infectious diseases are diseases caused by pathogenic microorganisms such as bacteria, viruses, parasites and fungi. They can be spread either directly or indirectly. Direct transmission occurs when there is a physical contact between an infected and a susceptible person. Examples of infectious diseases transmitted directly include common cold, scabies and sexually transmitted diseases. Indirect transmission on other hand, occurs when a susceptible individual comes into contact with a contaminated reservoir. Such diseases can be viral in nature, like rotavirus disease or hantavirus pulmonary syndrome

[101, 117]; bacterial, such as cholera or legionellosis [87]; or parasitic, such as schistosomiasis, cryptosporidiosis or giardiasis [104, 99, 92]. The study of diseases spreading through human populations has received attention from mathematicians since the seminal papers of Kermack and McKendrick in the 1920s [119]. However, such attention has mostly been confined to diseases which spread directly. Among the few existing models for indirectly transmitted pathogenic diseases, the main ones that make use of ordinary differential equations are those built upon the Cappasso and Pavari-Fontana model [123] and the Codeço model [125]. While the Capasso and Pavari-Fontana model [123] consisted of two equations, with one for the infected compartment and the other for the aquatic pathogen community, Codeço included the susceptible population and recovered population in the model as well. Denote  $S$ ,  $I$  and  $R$  as the susceptible, infected and recovered compartments from standard SIR models. The recovered compartment is not stated explicitly, as the population is assumed to be of constant size and so the dynamics of the recovered compartment follow directly from the rest of the system noting that  $H = S + I + R$ , where  $H$  is the total population. The model is written

$$\dot{S} = n(H - S) - a\lambda(B)S,$$

$$\dot{I} = a\lambda(B)S - rI,$$

$$\dot{B} = B(n_b - m_b) + eI.$$

The birth and death rate are the same and denoted  $n$ . The parameter  $r$  represents recovery rate, and include natural recovery and death. The pathogen have a net growth rate of proliferation  $n_b$  minus mortality  $m_b$ , and human

contamination increases pathogen levels at a rate  $e$  proportional to the size of the infected class. The infective term consists of the maximum rate of exposure to contaminated water,  $a$ , multiplied by  $\lambda(B) = \frac{B}{K+B}$  which is a Holling II response curve. The use of such a term would overestimate the infectivity of low levels of pathogens, contrary to the idea of a minimum infectious dose, which we think is important.

Key features of the model are that the aquatic reservoir is represented very simply with a linear growth term and linear shedding contribution. This was because the ecological dynamics of the pathogen were not well understood at the time (they are still not completely understood), so Codeço started with the simplest way to model the pathogen population. Unless net growth is naturally zero ( $n_b = m_b$ ), the bacterial population will die out exponentially in the absence of human shedding if  $n_b < m_b$ , or tend to infinity if  $n_b > m_b$ .

Hartley et al.[130] incorporated a hyperinfectious route of transmission to the Codeço model and Joh et al. [133], Tian et al. [140], Jensen et al. [132] and Mukanvire et al. [135] have further built on and branched off from these models.

Joh et al.'s [133] model differs from the other models in that it takes into account the fact that pathogens have to enter the human body in higher concentrations to overwhelm the natural immune response [136] by incorporating a Minimum Infection Dose (MID) into the incidence term. This MID is the rescaled value of the number of pathogens required to override the body's immune response. The infection term is a piecewise continuous function which is zero below the minimum infectious dose and a Holling II response curve above this threshold. Joh et al. analyzed the stability of this model, but

could not carry out all the essential stability and sensitivity analysis. In this paper we carry out a thorough stability and sensitivity analysis on the Joh et al's model. We show that if the bacterial population ( $B$ ) is less than or equal to the ratio of the MID (denoted  $c$ ) to the bacterial carrying capacity  $K$ , then we can only have an unstable disease free and bacteria free steady state. Whereas if  $B > c/K$  and  $c < K$  we could have an unstable disease free and bacteria free equilibrium and a conditionally globally stable endemic or epidemic steady state. Else if  $B > c/K$  and  $c \geq K$  then we may have up to four steady state: an unstable endemic steady state, a locally stable endemic steady state, a conditionally globally stable disease free steady state and an unstable disease free and bacteria free steady state. Further more, we will show using sensitivity analysis that to control the frequency of the outbreaks and the number of person that might be infected, it will be nice to focus on bacterial carrying capacity rather than on the shedding rate.

## 5.2 iSIR model formulation

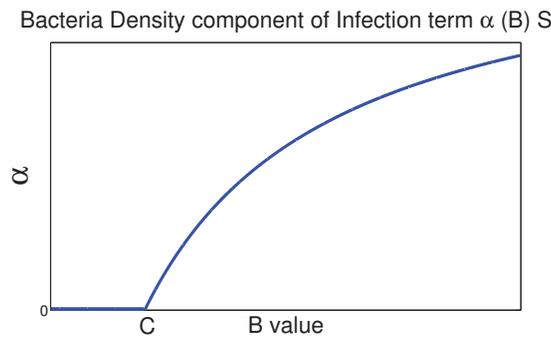
One of the key differences of the iSIR model, proposed in Joh et al. [133], compared to standard SIR models is the incidence term. The rough idea is that humans consume bacteria constantly but do not always get sick. Unlike with viruses where only a small amount of exposure is required, for certain types of bacteria a significant amount of bacterial cells need to be ingested in order to override the body's immune response [136]. This threshold has been measured by the likes of Cash et al. [124] and others [93, 102, 111] to be at least  $10^4$  cells. Simply using Holling I (or mass action) infection terms

or Holling II terms overestimates the infectivity of low levels of bacteria, since the standard Holling I and Holling II functions assume people will be infected with infinitesimally small densities of aquatic pathogens.

The incidence term used in this paper is  $\alpha(B)S$  where  $\alpha(B)$  is the pathogen density dependent component, and the  $S$  term is present for the same reasons as with standard SIR models. The indirect part of the incidence term is defined as

$$\alpha(B) = \begin{cases} 0, & B < c; \\ \frac{\alpha(B-c)}{(B-c)+H}, & B \geq c. \end{cases}$$

When the pathogen density is below a rescaled level corresponding to the MID, there will be no infections even with a nonzero amount of susceptibles, and after the bacterial density is above that threshold, infections will occur via a Holling II response, as shown in Figure 5.1.

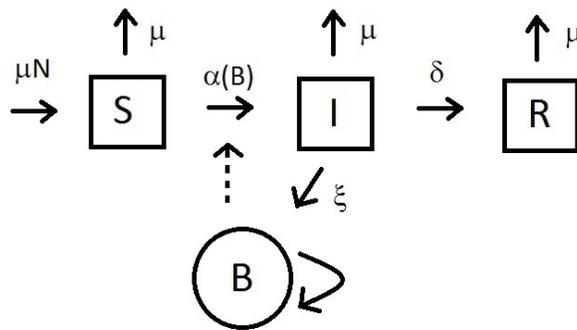


**Figure 5.1:** If bacterial levels are beneath the threshold  $c$ ,  $\alpha(B)$  is zero (no infections). If bacterial levels are above  $c$ , then  $\alpha(B)$  is a Holling II curve.

As pathogens exist naturally in the aquatic environment, the iSIR model uses logistic bacterial growth in the absence of any infected people, in contrast to most other models which have linear terms for the pathogen growth and death [125, 130, 94]. The latter leads to exponential decay in the absence of

infectives, which is consistent as many of those models assume that the aquatic reservoir of pathogen is not relevant to the cause of outbreaks, and so only the short term dynamics of freshly shedded pathogens are considered. In nonendemic areas, where the pathogen does not naturally exist in the environment, the linear form of the pathogen growth makes more sense than logistic growth, as used by Mukandavire et al. [135] in a study on recent outbreaks of Cholera in Zimbabwe. However, most models are intended for endemic areas.

The iSIR model has a positive contribution to the bacterial level when there are sick people shedding pathogens back to the reservoir. This occurs biologically with infected individuals contaminating the water supply through their pathogen laden feces. The dynamics are summarized in Figure 5.2.



**Figure 5.2:** A flow diagram demonstrating the relationship between Susceptibles (S), Infectives (I), Recovered (R) and pathogen (B). Humans have death rate  $\mu$ , contribute to the pathogen reservoir at rate  $\xi$ , recover at rate  $\delta$  and are infected at rate  $\alpha(B)$ .

The variables S, I and R in Figure 5.2, are defined in the usual way as susceptible, infected and recovered categories of the human population. The variable B represents the density of the pathogens in the aquatic reservoir. The first three equations sum to zero, thus the human population is of constant

size. The equations for the model are as follows:

$$\frac{dS}{dt} = -\alpha(B)S - \mu S + \mu N, \quad (5.1a)$$

$$\frac{dI}{dt} = \alpha(B)S - \mu I - \delta I, \quad (5.1b)$$

$$\frac{dR}{dt} = \delta I - \mu R, \quad (5.1c)$$

$$\frac{dB}{dt} = rB \left(1 - \frac{B}{K}\right) + \xi I, \quad (5.1d)$$

$$N = S + I + R. \quad (5.1e)$$

This model was first proposed in Joh et al. [133], though the analysis was preliminary and here we will present a thorough examination of its dynamics. For numerical simulations, the values of the parameters described in Table 7.1 are taken from the literature. The large variations of the key parameter values for certain waterborne diseases are given in Table 5.2. Note that the ranges given in the second column of Table 2 are different from the pathogen shed rate  $\xi$  in Table 7.1. The pathogen shed rate is the number of pathogens shed by an infected per day divided by the total water volume of the reservoir (in the unit of litre).

Parameter	Description	Dimension
$r$	Maximum per capita pathogen growth efficiency	day <sup>-1</sup>
$K$	Pathogen carrying capacity	cell litre <sup>-1</sup>
$H$	Half-saturation pathogen density	cell litre <sup>-1</sup>
$a$	Maximum rate of infection	day <sup>-1</sup>
$\delta$	Recovery rate	day <sup>-1</sup>
$\xi$	Pathogen shed rate	cell litre <sup>-1</sup> day <sup>-1</sup>
$\mu$	Per capita human birth/death rate	day <sup>-1</sup>
$N$	Total Population	persons
$c$	MID	cell litre <sup>-1</sup>

**Table 5.1:** Model parameters.

Disease	Number of pathogens shed by an infected/day (pathogen/day)	MID	Typical concentration (pathogen/liter)
Cholera	$10^{11} - 10^{12}$ [93]	$10^3 - 10^6$ [102, 126]	$10 - 10^3$ [109]
Cryptosporidiosis	$10^8$ [133]	$100 - 300$ [90]	$1 - 5$ [105]
Giardiasis	$10^8 - 10^9$ [112]	$10 - 100$ [112]	$1 - 5$ [105]
Rotavirus disease	$10^{12} - 10^{13}$ [108]	$100$ [133]	$10 - 1000$ [114]

**Table 5.2:** Key parameter values for certain waterborne diseases.

## 5.3 Mathematical results

We can nondimensionalize the system as follows:

$$\mathbf{S} = \frac{S}{N}, \mathbf{I} = \frac{I}{N}, \mathbf{B} = \frac{B}{K},$$

$$\tau = \mu t, \mathbf{A} = \frac{a}{\mu}, \mathbf{C} = \frac{c}{K}, \mathbf{p} = \frac{\mu + \delta}{\mu}, \mathbf{q} = \frac{\xi N}{\mu K}, \mathcal{R} = \frac{r}{\mu}, \boldsymbol{\lambda} = \frac{H}{K}.$$

We redefine the per capita infection rate  $\alpha$  accordingly as

$$\bar{\alpha}(\mathbf{B}) = \begin{cases} 0, & \mathbf{B} < \mathbf{C}, \\ \frac{\mathbf{A}(\mathbf{B}-\mathbf{C})}{(\mathbf{B}-\mathbf{C})+\boldsymbol{\lambda}}, & \mathbf{B} \geq \mathbf{C}. \end{cases}$$

The boldface is now dropped and we arrive at the following nondimensionalized iSIR system:

$$\frac{dS}{d\tau} = -\bar{\alpha}(B)S - S + 1, \quad (5.2a)$$

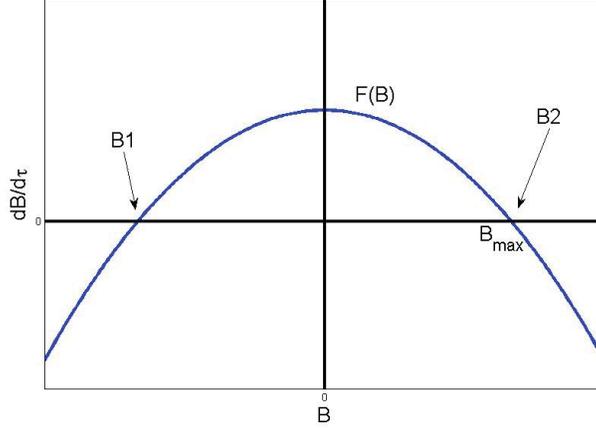
$$\frac{dI}{d\tau} = \bar{\alpha}(B)S - pI, \quad (5.2b)$$

$$\frac{dB}{d\tau} = \mathcal{R}B(1 - B) + qI. \quad (5.2c)$$

### 5.3.1 Forward invariance

First note that in dimensional terms, if  $S = 0$ , then  $\dot{S} = N > 0$  and so  $S(t) > 0$  for  $t > 0$ . If  $I = 0$ , then  $\dot{I} = \bar{\alpha}(B)S$  and because  $\bar{\alpha}(B) \geq 0$  by definition, then  $I \geq 0$  as well. The third equation of (5.1) gives us that  $\dot{R} = \delta I$  when  $R = 0$ , thus  $R(t) \geq 0$ . As  $S + I + R = N$ , we get that  $S, I, R \leq N$  in the usual way.

This transfers over to the nondimensional quantities of  $S$ ,  $I$  and  $R$ , the last of which we typically exclude. We have that  $0 \leq S + I \leq 1$  in particular.



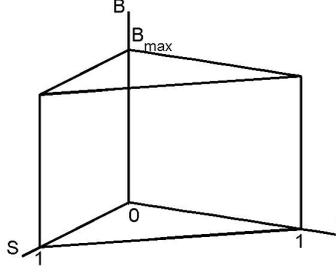
**Figure 5.3:** The derivative of the Bacteria vs. Bacterial Population. When above  $B_{max}$ , the derivative becomes negative. When  $B$  is zero, the derivative is positive.

Once again we drop the boldface for convenience. Looking at the third equation of the nondimensional system we can make note that  $\mathcal{R}B(1 - B) + qI \leq \mathcal{R}B(1 - B) + q$  as  $I \leq 1$ . Define  $F(B) := \mathcal{R}B(1 - B) + q$  which has roots  $B_{1,2} = \frac{\mathcal{R} \pm \sqrt{\mathcal{R}^2 + 4\mathcal{R}q}}{2\mathcal{R}}$  and note the smaller root  $B_1 = \frac{\mathcal{R} - \sqrt{\mathcal{R}^2 + 4\mathcal{R}q}}{2\mathcal{R}} < 0$  because of the positivity of the parameters. The other root  $B_2$  is clearly positive and is denoted as  $B_{max} = \frac{\mathcal{R} + \sqrt{\mathcal{R}^2 + 4\mathcal{R}q}}{2\mathcal{R}} > 1$ . The graph of  $F(B)$  is pictured in Figure 5.3. When  $B = 0$ , we see that  $\dot{B} = qI$  and thus  $B(\tau) \geq 0$  for  $\tau > 0$ . If  $B(0) \in [0, B_{max})$  then  $B(\tau) \in [0, B_{max})$  for any  $\tau > 0$ . The invariant region is pictured in Figure 5.4 and we summarize with a proposition.

**Proposition 1** (Feasible Region). *The set*

$$\Omega = \{(S, I, B) : 0 < S + I \leq 1, 0 \leq B \leq B_{max}, S > 0 \text{ and } I \geq 0\}$$

*defines a forward invariant region of system (5.2).*



**Figure 5.4:** The forward invariant region of system (5.2).

### 5.3.2 Equilibria of the system

Clearly  $E_0 = (1, 0, 0)$  is a steady state of (5.2) and biologically it corresponds to a disease-free and bacteria-free population. When  $C \geq 1$  i.e when the in-reservoir pathogen density is less than or equal to the rescaled MID, this means  $\bar{\alpha}(1) = 0$  and  $E_1 = (1, 0, 1)$  is an equilibrium corresponding to a disease-free state with bacteria at carrying capacity. When  $C < 1$  i.e when the in-reservoir pathogen density is greater than the rescaled MID, we get that  $\bar{\alpha}(1) \neq 0$  and so  $E_1 = (1, 0, 1)$  is not an equilibrium and (5.2) has no equilibrium  $(S^*, I^*, B^*)$  with  $B^* \leq C$  except  $E_0$ . The more complicated steady state  $E^* = (S^*, I^*, B^*)$  arises when  $B^* > C$  which causes  $\bar{\alpha}(B^*) \neq 0$ . Thus, system (5.2) implies that

$$S^* = \frac{B^* - C + \lambda}{(A + 1)(B^* - C) + \lambda},$$

$$I^* = \frac{1}{p} \left( \frac{A(B^* - C)}{(A + 1)(B^* - C) + \lambda} \right) = \frac{\mathcal{R}}{q} B^* (B^* - 1).$$

The expressions for  $I^*$  can be combined to form the equation

$$B^*(B^* - 1) \left( B^* - \left( C - \frac{\lambda}{A + 1} \right) \right) = \frac{q}{p\mathcal{R}} \frac{A}{A + 1} (B^* - C). \quad (5.3)$$

Define  $F_1(B) = f(B) - g(B)$  where  $f(B) = B(B - 1) \left( B - C + \frac{\lambda}{A+1} \right)$  and  $g(B) = \frac{q}{p\mathcal{R}} \frac{A}{A+1} (B - C)$ . Denote  $B_3 = C - \frac{\lambda}{A+1}$  so that if  $C < 1$  we see that

$$F_1(B_3) = \frac{q}{p\mathcal{R}} \frac{A}{A+1} \frac{\lambda}{A+1} > 0,$$

$$F_1(C) = C(C - 1) \left( \frac{\lambda}{A+1} \right) - 0 < 0.$$

Therefore there exists a root  $\bar{B}_1 \in (B_3, C)$ . However, as  $\bar{B}_1 < C$  then  $\bar{\alpha}(\bar{B}_1) = 0$  and equation (5.3) does not apply.

We have that  $f(0) = 0$  and  $g(0) < 0$

$$\implies f(0) > g(0)$$

$$\implies F_1(0) = f(0) - g(0) > 0$$

For  $B = b \ll 0$  we have that  $f(b) < 0$ ,  $g(b) < 0$  and  $f(b) < g(b)$

$$\implies F_1(b) = f(b) - g(b) < 0$$

Since  $F_1(0) > 0$  and  $F_1(b) < 0$  for  $b \ll 0$  and  $F_1$  is continuous on  $(0, b)$ , using the IVT we can find a  $\bar{B}_2 \in (0, b)$  such that  $F_1(\bar{B}_2) = 0$ . Since  $\bar{B}_2 < 0$ , it is not in the feasible region for  $\Omega$ , Lastly,

$$F_1(1) = -\frac{q}{p\mathcal{R}} \frac{A}{A+1} (1 - C) < 0,$$

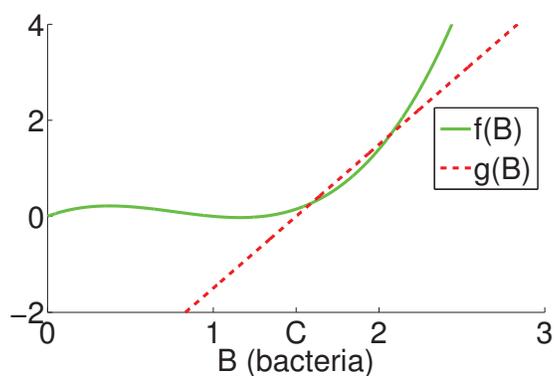
$$F_1(B_{max}) = (B_{max} - C) \left[ \frac{q}{\mathcal{R}} - \frac{q}{\mathcal{R}p} \frac{A}{A+1} \right] + \frac{q}{\mathcal{R}} \frac{\lambda}{A+1} > 0.$$

The latter is true as  $p > 1$ ,  $B_{max} > 1$ , and thus we conclude that there exists  $B^* \in (1, B_{max})$  when  $C < 1$  and it is the unique positive solution to (5.3), giving us a unique interior equilibrium  $E^* = (S^*, I^*, B^*)$ .

Note that  $\lim_{C \rightarrow 1^-} B^*(C) = 1$ , as the  $x$ -intercept of  $g(B)$  is 1 when  $C \rightarrow 1^-$

and  $f(1) = 0$ . As  $C$  increases over the value 1,  $E^*$  becomes  $E_1$  or vice versa if  $C$  is decreased.

We conclude that when  $C \geq 1$ ,  $f(B)$  and  $g(B)$  are as in Figure 5.5 and there are 0, 1 or 2 roots of Equation (5.3) with bacterial values greater than the minimum infectious dose ( $C$ ). For these values  $B_{1,2}^+$ , we find  $S_i^+$  and  $I_i^+$  in the same way as with  $E^*$ , leading us to two distinct equilibria  $E_{1,2}^+$ .



**Figure 5.5:** The left and right hand sides of Equation 5.3 when  $C \geq 1$ . There can be 0, 1 or 2 intersections with bacterial values above the MID.

As  $C$  is greater than 1, it is larger than all of the roots of  $f(B)$ . Thus,  $f(B)$  is concave up on  $[C, \infty)$  and  $f'(B) \geq f'(C) > 0$  for  $B \in [C, \infty)$ . If the slope of  $g(B)$  is less than  $f'(C)$ , there will be no endemic equilibria, as  $g(B)$  will always be below  $f(B)$  and there will be no intersections. Defining  $\zeta = \frac{q}{pR}$  and working backwards, we see that

$$\begin{aligned} \zeta &< f'(C), \\ \implies \zeta \frac{A}{A+1} &< f'(C), \\ \implies g'(B) &< f'(C), \end{aligned}$$

meaning that

$$\zeta < f'(C) \tag{5.4}$$

is a sufficient condition for there being no internal equilibria when  $C \geq 1$ . In dimensional parameters,  $\zeta = \left(\frac{\xi N}{\mu + \delta}\right) \left(\frac{\mu}{rK}\right)$ , and so  $\zeta$  is proportional to the shedding rate  $\xi$ . This motivates the definition of the condition for no internal steady states, as we shall see later. We summarize with a proposition.

**Remark 4** (Biological Interpretation of  $\zeta$ ).  $\zeta = \frac{\frac{\xi N}{\mu + \delta}}{\frac{1}{\mu}(rK)}$  is the ratio of the average number of pathogens shed over the time course of infection if all individuals were infected, to the average number of pathogens reproduced in the reservoir over the time course of an uninfected individual.

**Proposition 2** (Existence of equilibria). *The equilibrium  $E_0 = (1, 0, 0)$  always exists in  $\Omega$ .*

- *When  $C < 1$  (equivalently  $c < K$ ) i.e when the in-reservoir pathogen density is greater than the rescaled MID, there exist two equilibria,  $E_0$  on  $\partial\Omega$  and a unique endemic equilibrium  $E^*$  in  $\overset{\circ}{\Omega}$ .*
- *When  $C \geq 1$  (equivalently  $c \geq K$ ) i.e when the in-reservoir pathogen density is less than or equal to the rescaled MID, then  $E_1 = (1, 0, 1)$  is also an equilibrium and there can be up to two internal equilibria  $E_{1,2}^+$ .*
  - *If  $\zeta < f'(C)$ , there are no internal equilibria, and only  $E_1$  and  $E_0$  exist.*

### 5.3.3 Local stability of $E_0, E_1$ and $E^*$

We calculate the jacobian to analyze the local stability of each of the equilibria.

For the simpler case of  $B \leq C$ ,

$$J_1(S, I, B) = \begin{pmatrix} -1 & 0 & 0 \\ 0 & -p & 0 \\ 0 & q & R - 2\mathcal{R}B \end{pmatrix},$$

and for  $B > C$ ,

$$J_2(S, I, B) = \begin{pmatrix} \frac{-A(B-C)}{(B-C)+\lambda} - 1 & 0 & \frac{-A\lambda}{[(B-C)+\lambda]^2} S \\ \frac{A(B-C)}{(B-C)+\lambda} & -p & \frac{A\lambda}{[(B-C)+\lambda]^2} S \\ 0 & q & \mathcal{R} - 2\mathcal{R}B \end{pmatrix}.$$

When  $C \geq 1$ , the equilibria are  $E_0 = (1, 0, 0), E_1 = (1, 0, 1)$  and up to two  $E_i^+ = (S_i^+, I_i^+, B_i^+)$ . For  $C \geq 1$ , we use  $J_1$  and find that  $E_0$  has eigenvalues  $-1, -p$  and  $\mathcal{R}$ , which indicates that  $E_0$  is a saddle point equilibrium, as all parameter values are assumed positive.  $E_1$  in this case has eigenvalues  $-1, -p$  and  $-\mathcal{R}$  and thus we can conclude that when  $C \geq 1$ , the equilibrium  $(1, 0, 1)$  is locally asymptotically stable: that is, the disease-free equilibrium is locally asymptotically stable. Also, for  $C < 1$ ,  $E_0$  is a saddle point equilibrium for the same reasons.

Now considering  $E^*$  and using the nondimensionalized system (5.2), we

obtain

$$S^* = \frac{B^* - C + \lambda}{(A+1)\left(B^* - C + \frac{\lambda}{A+1}\right)} = \frac{B^* - C + \lambda}{(A+1)\left(\frac{q}{p\mathcal{R}}\frac{A}{A+1}(B^* - C)\frac{1}{B^*(B^*-1)}\right)}$$

$$S^* = \frac{p\mathcal{R}}{Aq} \frac{B^*}{B^* - C} (B^* - 1)(B^* - C + \lambda).$$

We will use  $\gamma$  for eigenvalues as the traditional  $\lambda$  is already used elsewhere.

We can compute

$$\det(\gamma I - J_{E^*}) = \det \begin{pmatrix} \gamma + \frac{A(B^*-C)}{(B^*-C)+\lambda} + 1 & 0 & \frac{A\lambda}{[(B^*-C)+\lambda]^2} S^* \\ \frac{-A(B^*-C)}{(B^*-C)+\lambda} & \gamma + p & \frac{-A\lambda}{[(B^*-C)+\lambda]^2} S^* \\ 0 & -q & \gamma + \mathcal{R}(2B^* - 1) \end{pmatrix}$$

$$= \left( \gamma + \frac{A(B^* - C)}{B^* - C + \lambda} + 1 \right) \left[ (\gamma + p)(\gamma + \mathcal{R}(2B^* - 1)) - \frac{A\lambda q}{(B^* - C + \lambda)^2} S^* \right]$$

$$+ \frac{A^2 \lambda q}{(B^* - C + \lambda)^3} (B^* - C) S^*.$$

Define  $F_2(\gamma) := \det(\gamma I - J_{E^*})$ ,  $h := A \frac{B^* - C}{B^* - C + \lambda} + 1$  and  $m := \frac{A\lambda q}{(B^* - C + \lambda)^2} S^*$ .

Later we will make use of the following alternate forms of these definitions:

$$h = A \left( \frac{B^* - C}{B^* - C + \lambda} + \frac{1}{A} \right) = \frac{(A+1)(B^* - C) + \lambda}{B^* - C + \lambda} > 1$$

and

$$m = \frac{A\lambda q}{(B^* - C + \lambda)^2} \frac{p\mathcal{R}}{Aq} \frac{B^*}{B^* - C} (B^* - 1)(B^* - C + \lambda) = p\mathcal{R} \frac{\lambda}{B^* - C + \lambda} \frac{B^* - 1}{B^* - C} B^*.$$

We can rewrite the characteristic equation with these new expressions taken into account as follows

$$\begin{aligned}
F_2(\gamma) &= (\gamma + h)[(\gamma + p)(\gamma + \mathcal{R}(2B^* - 1)) - m] + \frac{A^2\lambda q}{(B^* - C + \lambda)^3}(B^* - C)S^* \\
&= (\gamma + h)[\gamma^2 + (\mathcal{R}(2B^* - 1) + p)\gamma + p\mathcal{R}(2B^* - 1) - m] + \frac{A^2\lambda q}{(B^* - C + \lambda)^3}(B^* - C)S^* \\
&= \{\gamma^3 + \mathcal{R}(2B^* - 1 + p)\gamma^2 + [p\mathcal{R}(2B^* - 1) - m]\gamma + h\gamma^2 + h(\mathcal{R}(2B^* - 1) + p)\gamma \\
&\quad + [p\mathcal{R}(2B^* - 1) - m]h\} + \frac{A^2\lambda q}{(B^* - C + \lambda)^3}(B^* - C)S^*.
\end{aligned}$$

The Routh-Hurwitz coefficients of the above expression are

$$\begin{aligned}
b_3 &= 1, \\
b_2 &= \mathcal{R}(2B^* - 1) + p + h, \\
b_1 &= p\mathcal{R}(2B^* - 1) - b + h(\mathcal{R}(2B^* - 1) + p), \\
b_0 &= [p\mathcal{R}(2B^* - 1) - b]h + \frac{A^2\lambda q}{(B^* - C + \lambda)^3}(B^* - C)S^*,
\end{aligned}$$

and note that the Routh-Hurwitz stability criterion requires

$$b_1, b_2, b_3 > 0 \quad \text{and} \quad b_2b_1 > b_3b_0$$

as a sufficient condition for stability of the equilibrium. Clearly  $b_2$  and  $b_3$  are positive, and if  $p\mathcal{R}(2B^* - 1) - m > 0$  then  $b_1, b_0 > 0$ . As  $C < 1$  for the internal equilibrium  $E^*$  to exist,  $B^* - 1 < B^* - C$  so that  $\frac{B^* - 1}{B^* - C} < 1$ .

Thus

$$m = p\mathcal{R} \frac{\lambda}{B^* - C + \lambda} \frac{B^* - 1}{B^* - C} B^* < p\mathcal{R}B^*$$

and so

$$p\mathcal{R}(2B^* - 1) - m > p\mathcal{R}(2B^* - 1) - p\mathcal{R}B^* = p\mathcal{R}(B^* - 1) > 0,$$

which means  $b_1, b_0 > 0$ .

As for the second condition  $b_2b_1 > b_3b_0$ , we have the following expression

$$\begin{aligned} b_1b_2 &= [(h+p)\mathcal{R}(2B^* - 1) + hp - m][\mathcal{R}(2B^* - 1) + (p+h)] \\ &= (h+p)\mathcal{R}^2(2B^* - 1)^2 + (h+p)^2\mathcal{R}(2B^* - 1) + (hp-m)\mathcal{R}(2B^* - 1) \\ &\quad + (h+p)(hp-m). \end{aligned}$$

We can define

$$\begin{aligned} B_1 &= 2hp\mathcal{R}(2B^* - 1) - hm, \\ B_2 &= p\mathcal{R}^2(2B^* - 1)^2 - m\mathcal{R}(2B^* - 1), \\ B_3 &= p^2\mathcal{R}(2B^* - 1) - pm, \\ B_4 &= h\mathcal{R}^2(2B^* - 1)^2 + h^2\mathcal{R}(2B^* - 1) + h^2p + hp^2 + hp\mathcal{R}(2B^* - 1). \end{aligned}$$

Using the definition of  $S^*$ , we can express

$$\begin{aligned} b_3b_0 &= b_0 \\ &= [p\mathcal{R}(2B^* - 1) - m]h + \frac{A^2\lambda q}{(B^* - C + \lambda)^3}(B^* - C)\frac{p\mathcal{R}}{Aq}\frac{B^*}{B^* - C}(B^* - 1)(B^* - C + \lambda) \\ &= [p\mathcal{R}(2B^* - 1) - m]h + A\lambda p\mathcal{R}\frac{B^*}{(B^* - C + \lambda)^2}(B^* - 1). \end{aligned}$$

Now we check to see if the inequality  $b_1b_2 > b_3b_0$  is satisfied by noting that

$$\begin{aligned}
& (B^* - C)(2B^* - 1) > (B^* - 1)B^* \\
\Rightarrow & \frac{B^* - C}{B^* - C + \lambda}(2B^* - 1) + \frac{1}{A}(2B^* - 1) > (B^* - 1)\frac{B^*}{B^* - C + \lambda} \\
\Rightarrow & Ap\mathcal{R}\left(\frac{B^* - C}{B^* - C + \lambda} + \frac{1}{A}\right)(2B^* - 1) > Ap\mathcal{R}(B^* - 1)\frac{B^*}{B^* - C + \lambda} \\
& \Rightarrow hp\mathcal{R}(2B^* - 1) > Ap\mathcal{R}\lambda(B^* - 1)\frac{B^*}{(B^* - C + \lambda)^2} \\
\Rightarrow & 2hp\mathcal{R}(2B^* - 1) - hm > hp\mathcal{R}(2B^* - 1) + Ap\mathcal{R}\lambda\frac{B^*}{(B^* - C + \lambda)^2}(B^* - 1) - hm.
\end{aligned}$$

The left-hand side of the above inequality is precisely  $B_1$  and the right-hand side is  $b_0$ . Recall that  $m < p\mathcal{R}B^*$ , and so

$$p\mathcal{R}^2(2B^* - 1)^2 - m\mathcal{R}(2B^* - C) > p\mathcal{R}^2(2B^* - 1)^2 - p\mathcal{R}^2B^*(2B^* - 1) > 0$$

and

$$p^2\mathcal{R}(2B^* - 1) - pm > p^2\mathcal{R}(2B^* - 1) - p^2\mathcal{R}B^* = p^2\mathcal{R}(B^* - 1) > 0.$$

Thus  $B_2, B_3 > 0$  and clearly  $B_4 > 0$ . Lastly,

$$b_1b_2 = \sum_1^4 B_i > B_1 > b_0.$$

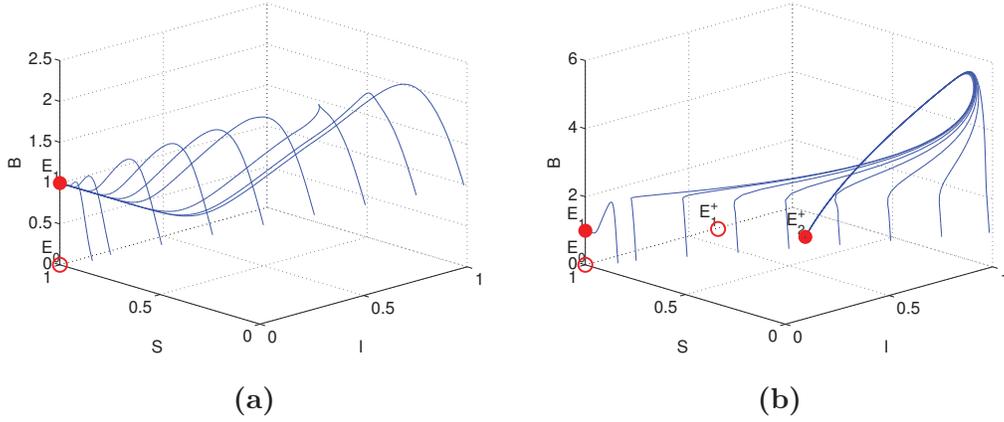
Thus the Routh-Hurwitz conditions are satisfied and  $E^*$  is locally asymptotically stable.

### 5.3.4 Local stability of $E_{1,2}^+$

When  $E_i^+$  exist things are more complicated as we lack exact expressions for the equilibrium quantities, and so the local stability is difficult to find analytically. Numerically it can be demonstrated that  $E_1^+$  (with  $B_1^+ < B_2^+$ ) is a saddle, and  $E_2^+$  is attracting. For example, this can be seen with parameters  $A = 1e3, C = 2, p = 10, q = 1e3, \mathcal{R} = 30$  and  $\lambda = 1, .$  With these parameters  $E_1^+$  and  $E_2^+$  both exist and are given as  $E_1^+ = (0.3986, 0.0601, 2.0015)$  and  $E_2^+ = (0.0036, 0.0996, 2.3898)$ . The eigenvalues corresponding to these steady states are  $\gamma_1 = (-0.9969, 580.8379, -682.4401)$ ,  $\gamma_2 = (-292.1109, -10.6368, -102.1349)$ . The equilibria  $E_1$  and  $E_2^+$  are both locally stable and a situation of bistability occurs, as observed in Figure 5.6a. Almost every solution approaches either the endemic equilibrium or the disease-free equilibrium, depending on initial conditions. Numerically we observe that the basin of attraction is much larger for the endemic equilibrium  $E_2^+$ , meaning that a greater range of initial conditions will lead to an endemic steady state rather than a disease free one. So,  $E_0$  is locally unstable,  $E_1$  and  $E^*$  are locally stable when they exist, and numerically we see that  $E_1^+$  is unstable and  $E_2^+$  is attracting. We summarize the preceding local stability results with a theorem.

**Theorem 5** (Local Stability). *System (5.2) has between two and four equilibria.*

- *When  $C < 1$  (equivalently  $c < K$ ),  $E_0 = (1, 0, 0)$  is unstable and a unique endemic equilibrium  $E^*$  exists and is locally asymptotically stable.*
- *When  $C \geq 1$  (equivalently  $c \geq K$ ), then  $E_0 = (1, 0, 0)$  is unstable and*



**Figure 5.6:** Trajectories in the phase space for  $C > 1$  and  $B^* \geq C$ . Stable equilibria are marked as  $\bullet$  while unstable equilibria are marked as  $\circ$ . a)  $\zeta < f'(C)$  ( $\zeta = 0.5$ ,  $f'(C) = 2.0297$ ). Only two equilibrium points exist: a saddle  $E_0$  and an attracting equilibrium  $E_1$ , the number of infected persons goes to zero as  $t \rightarrow \infty$ . The values used for the parameters are as follows:  $\mathcal{A} = 100$ ,  $C = 2$ ,  $p = 10$ ,  $q = 100$ ,  $\mathcal{R} = 20$ ,  $\lambda = 1$  (original parameter values are  $\delta = 9 \times 10^{-4}$ ,  $K = 1 \times 10^6$ ,  $a = 0.01$ ,  $H = 1 \times 10^6$ ,  $c = 2 \times 10^6$ ,  $N = 1 \times 10^6$ ,  $r = 0.002$ ,  $\mu = 1 \times 10^{-4}$ ,  $\xi = 0.01$ ). b)  $\zeta$  is sufficiently larger than  $f'(C)$  ( $\zeta = 3.3333$ ,  $f'(C) = 2.0030$ ). Four equilibrium points exist: two saddle equilibrium points  $E_1^+$  and  $E_0$  and two attracting equilibrium points  $E_1$  and  $E_2^+$ . The values used for the parameters are as follows:  $\mathcal{A} = 10^3$ ,  $C = 2$ ,  $p = 10$ ,  $q = 10^3$ ,  $\mathcal{R} = 30$ ,  $\lambda = 1$  (original parameter values are  $\delta = 9 \times 10^{-4}$ ,  $K = 1 \times 10^6$ ,  $a = 0.1$ ,  $H = 1 \times 10^6$ ,  $c = 2 \times 10^6$ ,  $N = 1 \times 10^6$ ,  $r = 0.003$ ,  $\mu = 1 \times 10^{-4}$ ,  $\xi = 0.1$ ).

$E_1 = (1, 0, 1)$  is an equilibrium and is locally asymptotically stable. Up to two internal equilibria,  $E_{1,2}^+$ , can also exist.

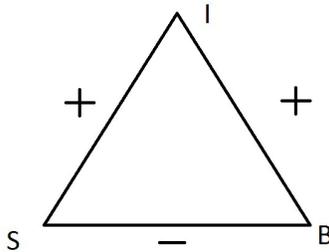
### 5.3.5 Global stability of $E_1$ and $E^*$

We wish to invoke a theorem of Hal Smith in regards to monotone dynamical systems and global stability. Because of the threshold parameter, the Jacobian of (5.2) will have two different forms with  $\alpha(B) = 0$  or not. Either way, the Jacobian is of the form

$$J(S, I, B) = \begin{pmatrix} * & + & - \\ + & * & + \\ - & + & * \end{pmatrix},$$

which is sign stable and sign symmetric in the off-diagonal entries. As demonstrated in Figure 5.7, every closed loop has an even number of edges with + signs and so the system is *monotone* as defined on page 49 in [91] in  $\Omega$  with respect to the partial ordering

$$K_m = \{(S, I, B) : S \geq 0, I \leq 0, B \geq 0\}. \quad (5.5)$$



**Figure 5.7:** The relationship between the three main compartments in the model.

Our argument is as follows: an application of monotone dynamical system theory states that if system (5.2) has a positive periodic orbit in domain  $\Omega$ , then there exists an unstable equilibrium in  $\Omega$  (Proposition 4.3 , p. 44, in [91]). When  $C \geq 1$  and condition (5.4) is satisfied ( $\zeta < f'(C)$ ), then there are only  $E_0$  and  $E_1$ , neither of which is an interior equilibrium. Hence system (5.2) will not have any periodic orbits in  $\Omega$ . As (5.2) is competitive, it reduces to a two-dimensional system [91]. Because of the absence of limit cycles and by the Poincaré-Bendixson theory, the local stability of  $E_1$  implies that  $E_1$  is globally asymptotically stable.

Define

$$H_1 = \{(S, I, B) : B \leq C, 0 < S + I \leq 1\},$$

$$H_2 = \{(S, I, B) : C < B < B_{max}, 0 < S + I \leq 1\},$$

and note that  $H_1 \subset \Omega, H_2 \subset \Omega$  with  $\Omega = H_1 \cup H_2$ . We will show that when  $C \geq 1$  and  $\zeta < f'(C)$ , after some  $\tau_0$  all solutions will stay entirely in  $H_1$  and we can apply our argument about the global asymptotic stability of  $E_1$ .

First we require a result from Hal Smith [91] about competitive systems noting first that  $\ll_m$  and  $\leq_m$  are order relations with respect to  $K_m$  defined in (5.5).

**Lemma 2** (Proposition 4.3, p. 44, in [91]). *Let  $\gamma$  be a non-trivial periodic orbit of a competitive system in  $D \subset \mathbb{R}^3$  and suppose there exists  $p, q \in D$  such that  $p \ll_m q$  and  $[p, q] = \{y \in D : p \leq_m y \leq_m q\} \subset D$ . Then  $W$  is an open subset of  $\mathbb{R}^3$  consisting of two connected components, one bounded and one unbounded. The bounded component,  $W(\gamma)$ , is homeomorphic to the open*

ball in  $\mathbb{R}^3$ .  $W(\gamma) \subset [p, q]$ , is positively invariant and its closure contains an equilibrium.

Now we require some results about the behaviour of solutions of (5.2) with respect to  $H_1$  and  $H_2$ .

**Lemma 3.** *If  $C \geq 1$  and  $\zeta < f'(C)$ , then for all solutions  $x(\tau) = (S(\tau), I(\tau), B(\tau))$  of (5.2), if there exists some  $\tau_0$  such that  $x(\tau_0) \in H_2$ , then there exists some  $\tau_1 > \tau_0$  such that  $x(\tau_1) \in H_1$ .*

*Proof.* Assume  $x(\tau) \in H_2$  for some  $\tau = \tau_1$ . Assume for contradiction that  $x(\tau) \in H_2$  for all  $\tau > \tau_1$ . Then by Monotone Dynamical Systems (MDS) Theory we can reduce this 3d system to a 2d system, as it is competitive, and by the Poincaré-Bendixson Theorem we can conclude all omega limit sets are limit cycles or equilibria.

As there is not an interior equilibria in  $H_2$ , we can conclude by Lemma 2 that there are not any limit cycles in  $H_2$ . As there are also no equilibria of any type in  $H_2$ , we conclude that  $x(\tau)$  exits  $H_2$  at some  $\tau_2 > \tau_1$ . This contradicts our assumption that  $x(\tau) \in H_2$  for all  $\tau > \tau_1$  and our Lemma is proven.  $\square$

**Lemma 4.** *If  $C \geq 1$  and  $\zeta < f'(C)$ , then for all solutions  $x(\tau)$  of (5.2), if there exists  $s_0, s_1$  such that  $s_1 > s_0$  where  $x(s_0) \in H_2$  and  $x(s_1) \in H_1$ , then  $x(\tau) \in H_1$  for  $\tau > s_1$ .*

*Proof.* Suppose there exists  $s_0$  and  $s_1, 0 < s_0 < s_1$  such that  $x(s_0) \in H_2$  and  $x(s_1) \in H_1$ . There exists  $\tau_0 \in (s_0, s_1)$  such that  $B(\tau_0) = C$  and  $\dot{B}(\tau_0) < 0$ . Suppose there exists  $\tau_1 > \tau_0$  where  $B(\tau_1) = C, \dot{B}(\tau_1) > 0$ , meaning that  $x(t)$

is re-entering  $H_2$ . Choose the first such time  $\tau_1$  and note

$$\dot{B}(\tau_1) = \mathcal{R}C(1 - C) + qI(\tau_1) > 0,$$

$$\dot{B}(\tau_0) = \mathcal{R}C(1 - C) + qI(\tau_0) < 0.$$

This means that  $I(\tau_1) > I(\tau_0)$  but  $B(\tau) \leq C$  on  $\tau_0 < \tau < \tau_1$  and so  $\dot{I} = -pI < 0$ . This is a contradiction, so there can be no such  $\tau_1$  as supposed and the Lemma is proven.  $\square$

Thus no solutions can stay in  $H_2$  as  $\tau \rightarrow \infty$  and once  $H_1$  is entered from  $H_2$ ,  $H_1$  is forward invariant. This captures the behaviour of all solutions  $x(\tau)$ .

We can now conclude that  $E_1$  is globally asymptotically stable.

**Proposition 3** (Global Stability of  $E_1$ ). *When  $C \geq 1$  and  $\zeta < f'(C)$ ,  $E_1 = (1, 0, 1)$  is an equilibrium of (5.2) and it is globally asymptotically stable.*

*Proof.* When  $C \geq 1$  and  $\zeta < f'(C)$ , by Lemmas 3 and 4, all solutions eventually exist entirely in  $H_1$  and as there are no interior equilibria (because  $\zeta < f'(C)$ ), by Lemma 2 there are no limit cycles in  $H_1$ . Monotone Dynamical Systems theory says that (5.2) reduces to a two-dimensional system, and so by the Poincaré-Bendixson theorem all omega-limit sets are limit cycles or equilibria. As there are no limit cycles in  $H_1$  and no interior equilibria, by the local stability of  $E_1$ , we conclude that it is globally asymptotically stable.  $\square$

Now we consider the global stability of  $E^*$ . As (5.2) is monotone, it verifies the Poincaré-Bendixson property: every compact omega-limit set without equilibria is a closed orbit. For systems with this property, a criterion on

global stability has been developed by Li, Wang and Muldowney [106, 107]. Note that the *second additive compound* of a  $3 \times 3$  matrix,  $A = [a_{ij}]$ , is denoted  $A^{[2]}$  and defined

$$A^{[2]} = \begin{bmatrix} a_{11} + a_{22} & a_{23} & -a_{13} \\ a_{32} & a_{11} + a_{33} & a_{12} \\ -a_{31} & a_{21} & a_{22} + a_{33} \end{bmatrix}.$$

**Lemma 5** (Theorem 2.5 in [107]). *Let  $\dot{x} = F(x)$  ( $F \in C^1$ ) be a system defined on an open convex subset  $G \subset \mathbb{R}^3$  having a compact global attractor in  $G$ . Assume that*

- 1) *The Poincaré-Bendixson property holds.*
- 2) *There is a unique equilibrium in  $G$  which is locally asymptotically stable.*
- 3) *For each periodic orbit  $p(t)$  in  $G$ , the linear system*

$$\dot{Y} = \frac{\partial F^{[2]}}{\partial x}(p(t))Y$$

*is asymptotically stable.*

*Then the equilibrium is globally asymptotically stable in  $G$ .*

In order to apply this result, we have to study the asymptotic stability of the linear equation

$$\dot{Y} = J^{[2]}(p(t))Y \tag{2.4}$$

where  $p(t)$  is any periodic solution of (5.2) in  $\Omega$ . Given our definition of  $J_{E^*}$ , the second additive compound of  $J_{E^*}$  is

$$J^{[2]} = \begin{pmatrix} -f_0(B) - 1 - p & f'_0(B)S & f'_0(B)S \\ q & -f_0(B) - f'_1(B) - 1 & 0 \\ 0 & f_0(B) & -p - f'_1(B) \end{pmatrix},$$

where  $f_0(B) = \frac{A(B-C)}{B-C+\lambda}$  and  $f'_1(B) = 2\mathcal{R}B - \mathcal{R}$ .

Typically verifying the stability of such a system is nontrivial, but for (2.4) we have a linear, periodic, cooperative, irreducible system with respect to the cone

$$K_1 = \{(S, I, B) : S \geq 0, I \geq 0, B \geq 0\}$$

which suggests we use a comparison result.

**Lemma 6.** (*Proposition 3 in [98]*) *Let  $\dot{Y} = A_i(t)Y$  for  $i = 1, 2$  be two linear, periodic, cooperative and irreducible systems (with the same period) such that  $A_2(t) - A_1(t)$  has nonnegative coefficients. If  $\dot{Y} = A_2(t)Y$  is asymptotically stable, then  $\dot{Y} = A_1(t)Y$  is too.*

For our case, obviously  $A_1 = J^{[2]}(p(t))$  and for  $A_2$  we choose a constant matrix whose entries bound those of  $A_1$  independently of the periodic orbit and denote the matrix  $\bar{J}$  where

$$\bar{J} = \begin{pmatrix} -1 - p - f_0(1) & f'_0(1) & f'_0(1) \\ q & -f_0(1) - f'_1(1) - 1 & 0 \\ 0 & f_0(B_{max}) & -p - f'_1(1) \end{pmatrix}.$$

The characteristic equation,  $P(\gamma)$ , of  $\bar{J}$  is

$$P(\gamma) = [\gamma+1+p+f_0(1)][\gamma+1+f_0(1)+f'_1(1)][\gamma+p+f'_1(1)]-qf'_0(1)[\gamma+p+f'_1(1)+f_0(B_{max})].$$

Expanding this out we can write

$$P(\gamma) = a_3\gamma^2 + a_2\gamma + a_1\gamma + a_0,$$

with coefficients

$$a_3 = 1,$$

$$a_2 = 2(1 + p + f_0(1) + f'_1(1)),$$

$$a_1 = [2 + 2f_0(1) + p + f'_1(1)](p + f'_1(1)) + [1 + p + f_0(1)](1 + f_0(1) + f'_1(1)) - qf'_0(1),$$

$$a_0 = [1 + p + f_0(1)](1 + f_0(1) + f'_1(1))[p + f'_1(1)] - qf'_0(1)[p + f'_1(1) + f_0(B_{max})].$$

To use the Routh-Hurwitz conditions, we require that  $a_i > 0$  and that  $a_1a_2 > a_3a_0$ . First, we consider the positivity of the coefficients. Only the positivity of  $a_1$  and  $a_0$  require checking.

$$\begin{aligned} a_0 &> p + pf'_1(1) + p^2f'_1(1) + pf'_1(1) + pf'_1(1)f'_1(1) + f'_1(1)f'_1(1) \\ &\quad - qA\lambda^{-1}(p + f'_1(1) + A) \\ &= (p + 1)^2f'_1(1) + p^2 + p + pR^2 + R^2 - qA\lambda^{-1}(p + f'_1(1) + A). \end{aligned}$$

Noting that  $f'_1(1) = \mathcal{R} = \frac{r}{\mu}$  it is reasonable to assume that  $\mathcal{R} > 1$  as  $\mu$  is the human birth/date rate and will be very small. Also  $r$ , the maximum bacterial

growth rate, is often greater than 1. Finally, by definition  $p > 1$  thus

$$\begin{aligned} a_0 &> (p+1)^2 \mathcal{R} + p^2 + p + p\mathcal{R}^2 + \mathcal{R}^2 - qA\lambda^{-1}(p + \mathcal{R} + A) \\ &> (p+1)^2 \mathcal{R} + (p+1)^2 - qA\lambda^{-1}(p + A) - qA\lambda^{-1}\mathcal{R} \end{aligned}$$

Suppose the parameters  $p$ ,  $q$ ,  $A$ , and  $\lambda$  are such that

$$(p+1)^2 > qA\lambda^{-1}(p + A). \quad (2.5)$$

Then  $a_0 > 0$ . The positivity of  $a_1$  follows similarly,

$$a_1 > (2+p)p + (1+p) - qA\lambda^{-1} > (1+p)^2 - qA\lambda^{-1},$$

and if parameters satisfy (2.5), then  $a_1 > 0$  too. Now we can consider

$$\Delta = a_1 a_2 - a_3 a_0,$$

$$\begin{aligned} \Delta &= 2(1+p+f_0(1)+f_1'(1))\{(1+f_0(1)+p)(p+f_1'(1)) \\ &\quad + [1+p+f_0(1)][1+f_0(1)+f_1'(1)]\} + [1+p+f_0(1)](1+f_0(1)+f_1'(1))[p+f_1'(1)] \\ &\quad + 2f_1'(1)[1+f_0(1)+f_1'(1)](p+f_1'(1)) + qf_0'(1)[-2-p-2f_0(1)-f_1'(1)+f_0(B_{max})], \\ \Delta &> 2(1+p+f_0(1)+f_1'(1))(1+p)p + (1+p) - qf_0'(1)[2+2p+f_0(1)+2f_1'(1)] \\ &= 2(1+p+f_0(1)+f_1'(1))\left[(1+p)^2 - qf_0'(1)\right]. \end{aligned}$$

Note that  $(1+p)^2 - qf_0'(1) > (1+p)^2 - qA\lambda^{-1}$  and if parameters satisfy (2.5), we have that  $a_2 a_1 > a_3 a_0$ . Lastly considering  $H_1$  and  $H_2$  as before, note that if  $C < 1$  then  $\dot{B}(B=C) = \mathcal{R}C(1-C) + qI > 0$ , so eventually all trajectories exist entirely in  $H_2$ . In particular, any attracting limit cycles are

contained in  $H_2$ . We restate the previous results in a proposition.

**Proposition 4** (Behaviour of limit cycles and Global Stability of  $E^*$ ). *When  $C < 1$ ,  $E^*$  is an equilibrium of (5.2). Any limit cycle, if it exists, should be entirely in  $H_2$ , and if  $(p+1)^2 > qA\lambda^{-1}(p+A)$ , then  $E^*$  is globally asymptotically stable.*

Recalling that  $C = c/K$ , we can summarize our results about the equilibria in this section with the following theorem.

**Theorem 6** (Global Stability). *System (5.2) always has at least two equilibria.*

- *If  $C < 1$  (equivalently  $c < K$ ), the equilibria are  $E_0 = (1, 0, 0)$  which is unstable and  $E^* = (S^*, I^*, B^*)$  which is locally asymptotically stable. Furthermore, if  $(p+1)^2 > qA\lambda^{-1}(p+A)$ , then  $E^*$  is globally asymptotically stable.*
- *If  $C \geq 1$  (equivalently  $c \geq K$ ), the equilibria are  $E_0 = (1, 0, 0)$  which is unstable,  $E_1 = (1, 0, 1)$  which is locally asymptotically stable and up to two internal equilibria  $E_{1,2}^+$ . Further, if  $\zeta < f'(C)$  only  $E_0$  and  $E_1$  exist, and  $E_1$  is globally asymptotically stable.*

Note that if the MID is greater than  $K$  and  $\zeta$  is low enough to satisfy condition (5.4), the disease-free equilibrium  $E_1$  is globally asymptotically stable. As nondimensional  $\zeta$  and the shedding parameter  $\xi$  are proportional, this means that with a nonzero but sufficiently small shedding rate, the disease-free equilibrium is inevitable. This is in contrast to the case where the MID is less than the carrying capacity of bacteria, and the bacteria exist at levels

which naturally cause new infections. In this case, if other parameters agree, the endemic steady state  $E^*$  is globally asymptotically stable for any nonzero shedding rate  $\xi$ . Thus if efforts are taken to decrease  $K$  and  $\xi$  in conjunction, a disease-free globally stable steady state can be attained with a shedding rate that could otherwise lead to an endemic steady state.

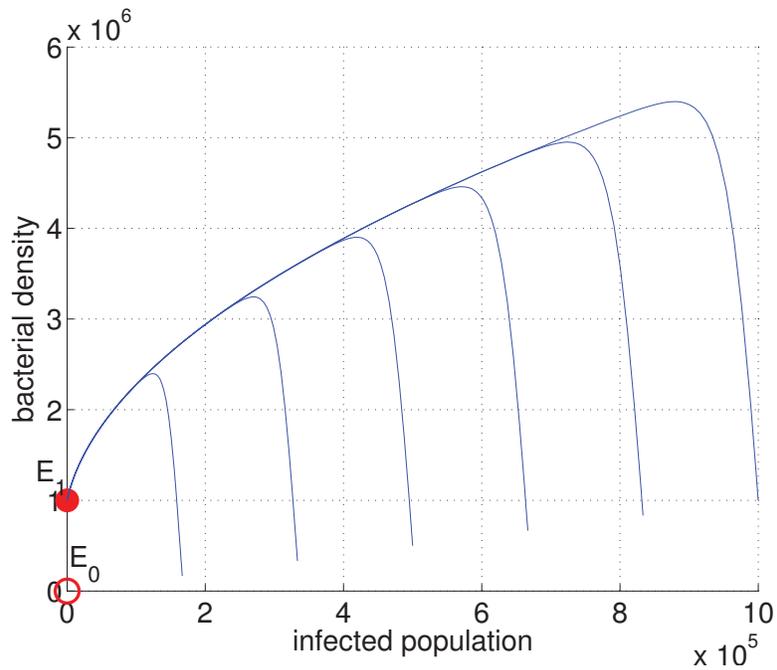
**Remark 5. Biological interpretation of the sufficient condition for globally asymptotic stability of  $E^*$ :** *In terms of the original parameters,*

$$(p + 1)^2 > qA\lambda^{-1}(p + A) \text{ could be written as } N < \frac{\mu H(\delta + 2\mu)^2}{\xi a(\mu + \delta + a)} := N_c.$$

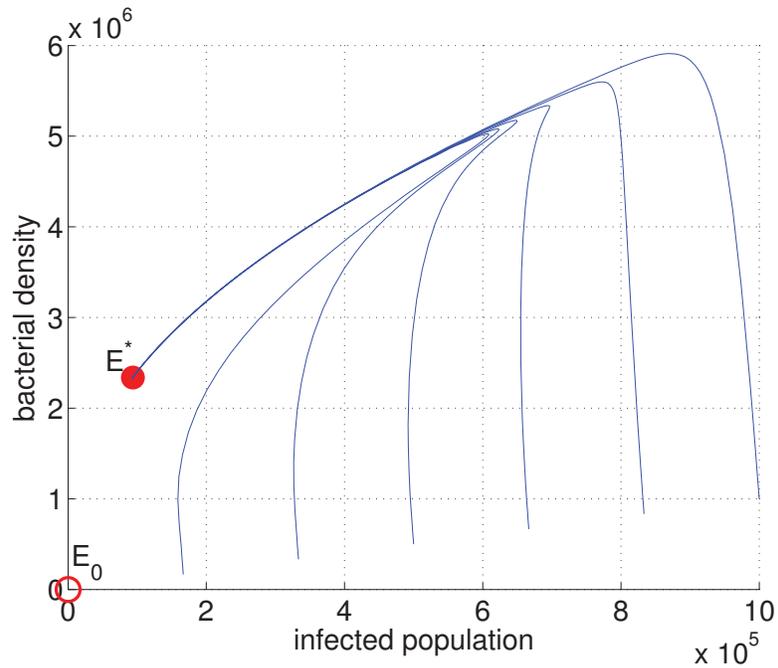
*$N_c$  is a critical population level below which the endemic equilibrium  $E^*$  is globally attracting. If we consider towns with similar population densities and sanitation infrastructures, and with their pathogen carrying capacity in the reservoir over their rescaled MID, a small population size is associated with a small town, which results in a strong connection between individuals and pathogens. Hence, a small population size leads to the robust and persistent disease prevalence in the case that the pathogen carrying capacity is greater than the MID. Note that the given inequality is a sufficient but not necessary condition, thus a large city ( $N \geq N_c$ ) with some restrictions may also lead to the globally attracting  $E^*$ . With the limitation of existing mathematical techniques, the sufficient condition is the best condition we can obtain for the global stability of the internal equilibrium. Using the reasonable parameter values,  $N_c$  can reach up to  $2.5439 \times 10^6$ .*

## 5.4 Numerical simulations

Stability was discussed nondimensionally previously but numerical examples are presented in the following diagrams in dimensional parameters. In Figure 5.8, we see that if  $c \geq K$  ( $C > 1$ ) with a small  $\xi$  value, meaning the minimum infectious dose is larger than the carrying capacity, then there is no endemic equilibrium. Thus the system moves towards the disease free steady state. Intuitively, this means that it takes more than the ‘natural level’ of bacterial density in the water supply to make anyone sick and shedding is low, so no one becomes sick.



**Figure 5.8:** Phase portrait for  $c > K$ , showing trajectories with different initial conditions converging to the disease free steady state  $E_1$ . Parameter values used are  $\delta = 0.1$ ,  $K = 1 \times 10^6$ ,  $a = 0.1$ ,  $H = 1 \times 10^8$ ,  $c = 2 \times 10^6$ ,  $N = 1 \times 10^6$ ,  $r = 0.3704$ ,  $\mu = 5 \times 10^{-5}$ ,  $\xi = 10$ .



**Figure 5.9:** Phase portrait for  $c < K$ , showing many different trajectories approaching the endemic steady state, marked with a solid circle. Parameter values used are  $\delta = 9 \times 10^{-4}$ ,  $K = 1 \times 10^6$ ,  $a = 0.1$ ,  $H = 1 \times 10^8$ ,  $c = 8 \times 10^5$ ,  $N = 1 \times 10^6$ ,  $r = 0.003$ ,  $\mu = 5 \times 10^{-4}$ ,  $\xi = 0.1$ .

Figure 5.9 demonstrates that when the minimum infectious dose is less than the carrying capacity,  $c < K$ , then the internal endemic steady state is attracting. In it, a wide range of initial conditions all follow a similar path towards the endemic steady state after each trajectory first experiences an outbreak. This means that if the MID is small enough that a normal bacterial density can make any individual sick, then the disease will persist in the community if only at a low level. As mentioned, a strong epidemic always occurs with a high outbreak peak.

## 5.5 Local sensitivity analysis

In this section, we compute and analyze the normalized forward sensitivity indices of different quantities to the parameters of the system by computing

$$S.I. = \frac{\partial x^*}{\partial p} \frac{p}{x^*} \quad (5.6)$$

where  $x^*$  is the quantity being considered, and  $p$  is some parameter which  $x^*$  depends upon. Sensitivity indices can be positive or negative which indicates the nature of the relationship, and it is the magnitude that ranks the strength of the relationship as compared to the other parameters. Since we don't have an explicit formula for the quantities we are interested in (outbreak peak, outbreak peak time and the endemic steady state), we estimate  $\frac{\partial x^*}{\partial p}$  using the central difference approximation:

$$\frac{\partial x^*}{\partial p} = \frac{x^*(p + \Delta p) - x^*(p - \Delta p)}{2\Delta p} + \mathcal{O}(\Delta p^2).$$

We choose  $\Delta p = 1\%$  of  $p$ .

plugging all these in (5.6) we get

$$S.I. = \frac{x^*(1.01p) - x^*(0.99p)}{0.02x^*}$$

### 5.5.1 Sensitivity of the outbreak peak

The sensitivity indices of the amplitude of the outbreak peak show how the first epidemic depends on the parameters as seen in Table 7.2. This table has three columns because there is a noticeable difference in the sensitivity indices,

when the bacteria started out above or below the carrying capacity.

Parameter	Sensitivity B(0)< K	Sensitivity B(0)>K	
$\delta$	-1.2024	-0.5296	Recovery rate
K	1.8773	1.1334	Bacterial carrying capacity
a	1.1980	0.9822	Contact rate
H	-1.1905	-0.9623	Half Saturation constant
c	-0.9324	-0.4196	Minimum Infectious Dose
r	-0.2305	-0.5267	Logistic bacterial growth
$\mu$	-5.5e-004	-2.5830 e-004	Human birth/death
$\xi$	0.2352	0.0636	Shedding rate

**Table 5.3:** The sensitivity of the magnitude of the peak outbreak to the parameters. Two columns for the initial density of bacteria below or above its carrying capacity  $K$ .

The carrying capacity  $K$  has the strongest relationship to the magnitude of the outbreak peak. The positive value tells us that a higher carrying capacity would lead to a more severe epidemic. In contrast to the shedding rate  $\xi$  which has among the lowest of sensitivity indices,  $K$  would thus be an important parameter to control in order to reduce the harm of an outbreak.

A negative relationship between  $r$  and the peak magnitude might seem counter intuitive, but the per capita growth rate of bacteria at any given time is  $r(1 - \frac{B}{K})$  and during the peak the bacteria exist over their natural carrying capacity, so the growth rate would be negative and thus there is a negative relationship between  $r$  and the peak amplitude.

The sensitivity index with respect to the human birth/death rate  $\mu$  is very low in comparison to all the others. This makes sense, because the initial

peak of an epidemic occurs relatively quickly after the introduction of sick people or introduction of high levels of bacteria and the birth and death of new susceptibles would not be on the same time scale.

A negative relationship between the minimum infectious dose (MID)  $c$  and peak amplitude is consistent with our understanding of the disease dynamics, because a larger MID means it would take a higher bacterial density to cause any infections at all. Thus a higher MID would mean less infections and a smaller outbreak peak.

The recovery rate  $\delta$  has a strong negative relationship to the peak outbreak level as a higher  $\delta$  leads to few infectives by definition.

### 5.5.2 Sensitivity of the outbreak peak time

Parameter	Sensitivity $B(0) < K$	Sensitivity $B(0) > K$	
$\delta$	0.0772	-6.7542	Recovery Rate
$K$	0.4392	4.8433	Bacterial carrying capacity
$a$	-0.2177	-0.3361	Contact Rate
$H$	0.2159	0.3264	Half Saturation constant
$c$	0.9319	0.0928	Minimum Infectious Dose
$r$	-0.6327	3.9216	Logistic Bacterial growth
$\mu$	3.5491e-005	4.3194e-005	Human birth/death
$\xi$	-0.2269	-0.1425	Shedding rate

**Table 5.4:** The sensitivity of the time of the outbreak maximum to the parameters.

Once again we see from Table 7.3 that the carrying capacity  $K$  has a large influence on the dynamics of the system. It has one of the largest sensitivity

indices, being many times greater than that of the shedding rate  $\xi$ . This suggests that  $K$  is a more important quantity to control to prevent outbreaks. The positive relationship means a smaller carrying capacity would lead to a quicker outbreak as well as a smaller one as we saw in the last section. Noticeable is the lack of effect of  $\mu$ , as with the amplitude of the peak. It has such a negligible effect for the same reasons as outlined previously.

The relationship between contact rate  $a$  and the time of the maximum outbreak is a negative relationship, because a higher contact rate causes more new infections and so the timing of the maximum would be attained earlier than otherwise.

The recovery rate  $\delta$  is interesting because its effect changes sign as well as magnitude considerably with different values of  $B(0)$  in relation to carrying capacity. When  $B(0) > K$  the effect of  $\delta$  is greatest and negative. A higher value of  $\delta$  would mean individuals would be infected, and thus infectious, for less time, so the outbreak should not be as severe and would occur earlier than otherwise. As the magnitude of  $\delta$  is so small when it is positive, the positive relationship does not yield insight into the relationship of outbreak time and recovery rate.

The per capita growth rate is  $r \left(1 - \frac{B}{K}\right)$  and so when  $B > K$  this growth rate is negative. If  $B(0) < K$  then the growth rate will be positive in the beginning of the outbreak, so a larger  $r$  would mean a higher growth rate, and thus the epidemic would peak earlier. This is supported by the negative relationship with  $r$  and peak time when  $B(0) < K$ . If however  $B(0) > K$ , the per capita growth rate will be negative from the start, and as  $B$  will remain above  $K$  for all time, the growth rate will always be negative. So a larger  $r$

value would mean slower growth, and the epidemic wave would take longer to reach a maximum. This is supported by the strong positive relationship of  $r$  and peak time when  $B(0) > K$  as can be seen in Table 7.3.

### 5.5.3 Sensitivity of the endemic steady state

Parameter	Sensitivity of $S^*$	Sensitivity of $I^*$	Sensitivity of $B^*$	
$\delta$	0.0321	-0.9453	-0.0780	Recovery Rate
$K$	-1.9877	-0.1036	1.0666	Bacterial carrying capacity
$a$	-1.0314	-0.0611	0.0402	Contact Rate
$H$	1.0260	0.0606	-0.0399	Half Saturation constant
$c$	0.9932	0.0225	0.0014	Minimum Infectious Dose
$r$	0.0323	-0.0329	0.0474	Logistic Bacterial growth
$\mu$	0.8472	0.8177	0.0811	Human birth/death
$\xi$	-0.0323	0.0123	-0.0179	Shedding rate

**Table 5.5:** The sensitivity of the components of the endemic equilibrium.

We can look at the sensitivity of one of the interior equilibria  $E^*$  with respect to the parameters when  $E^*$  exists. Here we only look at  $B(0) < K$  as the other

case has similar sensitivity results, and we assume the other endemic steady states would yield similar results.

The final size of the susceptible population is most sensitive to the carrying capacity  $K$  and contact rate  $a$ , with a negative relationship in both cases. This is because a higher contact rate causes more infections and a higher carrying capacity causes more bacteria which indirectly leads to more infections. A higher shedding rate would cause more infections which is confirmed with the negative relationship between  $S^*$  and  $\xi$ . But  $S^*$  is many times less sensitive to  $\xi$  than to  $K$  which again points to  $K$  as the more important parameter to focus on in disease control. The minimal infectious dose (MID)  $c$  is nearly as sensitive as the contact rate, but has a positive relationship as a higher MID would lead to fewer infections and a higher  $S^*$ . There is a weak relationship with  $\delta$  but as our model does not allow for reinfection, this accounts for the small magnitude of the sensitivity.

The endemic level of infective individuals is most sensitive to the recovery rate  $\delta$  and the birth/death rate  $\mu$ . The strong negative relationship with  $\delta$  is because recovery is the main way that infectives leave the infected component of our model. The relationship with  $\mu$  is complicated in that the birth rate and death rate are the same in our model. So a larger  $\mu$  means more deaths and thus more infectives leaving the infected component, but also more newly born susceptibles to possibly enter the infected class. The positive relationship means that the positive effect of births is more important to  $I^*$  than the negative effect of deaths. The shedding rate  $\xi$  has a weak relationship with  $I^*$  but the positive relationship is as expected because a larger  $\xi$  leads to more infectives and a higher  $I^*$  value.

The endemic level of the bacterial population is most sensitive to the bacterial carrying capacity  $K$  and has a positive relationship to it as expected. A higher  $K$  means more bacteria and as  $B^* > K$  (equivalent to  $B^* > 1$  in nondimensional form) the relationship is positive. The shedding rate has a small sensitivity which suggests that the logistic part of  $\frac{dB}{dt}$  is more important to the endemic level of  $B^*$ . As such the relationship with the MID is also minimal. The strong relationship with  $K$  and weak one with  $\xi$  also again suggests the important of  $K$  instead of  $\xi$  as a control measure. This could mean, for example, that monitoring the bacterial levels in water reservoirs is more important than simply controlling or restricting access to the water supply to avoid contamination.

## 5.6 Discussion

Cholera has the potential to quickly spread over large areas and can cause many deaths. Thus a full understanding of the dynamics is essential to effectively respond to outbreaks. With the continuing outbreaks there is the opportunity for mathematical modeling to help decipher these dynamics and provide suggestions for governments and health care bodies in effective intervention. An estimate for the basic reproductive number in regions affected by cholera would give important information for controlling future outbreaks and for creating surveillance programs. The potential for amplification in environmental reservoirs and the indirect transmission of the disease make this a nontrivial task. Here we have shown that with  $\mathbf{C} = c/K \geq 1$ , the disease free equilibrium can be globally asymptotically stable. However as bacteria

are existing at a nonzero level, if environmental factors change and alter the carrying capacity enough to make  $\mathbf{C} < 1$ , then there can be outbreaks. If other parameters are in agreement, an endemic equilibrium is globally asymptotically stable. This change to carrying capacity could be seasonally caused as with different amounts of rain in areas like Bangladesh, or it could be a more permanent change due to natural disasters as in Haiti.

An important thing to note about the relationship between  $c$  and  $K$  is that if  $c < K$ , i.e. the minimum infectious dose is less than the carrying capacity, then the unique endemic equilibrium can be globally stable. It is globally stable for any nonzero value of the shedding parameter  $\xi$ . If however, the minimum infectious dose is greater than the natural carrying capacity, if  $\xi$  is low enough, causing the nondimensional  $\zeta$  to be sufficiently small, then the disease free equilibria becomes globally stable. This highlights the importance of being aware of the value of the natural carrying capacity, because decreasing the shedding rate can eliminate the possibility of an endemic steady state, if the MID is larger than the carrying capacity. If the MID was less than the carrying capacity, the unique endemic steady state could be globally asymptotically stable for the same shedding rate. So ideally, efforts need to be taken to reduce both shedding, and in conjunction with this, the bacterial levels in the reservoir.

Our sensitivity analysis suggests that control measures influencing the carrying capacity  $K$  will be more effective in minimizing the epidemic than those concentrating on influencing the shedding rate. While improving the sanitation infrastructure of an area is the obvious step to take to control outbreaks, monitoring and controlling the bacterial levels in the water itself is more im-

portant. Improving the infrastructure would surely help control the bacterial levels in the water by decreasing the amount of human contamination, but *V.cholerae* exist independently of humans and so other factors that influence the natural levels of bacteria in the water need to be considered as well in intervention strategies. As mentioned above, controlling both parameters is important and likely to be the most effective, but the carrying capacity  $K$  is the more influential of the two on its own.

Our analysis accounts for all the situations experienced all over the world. Should the average health and immune system capabilities of the population be sufficient to tolerate bacterially contaminated water, and the shedding rate be sufficiently small, the human and bacteria populations will exist independently of each other. This case reflects both interepidemic periods where cholera outbreaks are common, and also the situation in regions where cholera outbreaks are not experienced. Should however the carrying capacity be sufficiently high (enough to overwhelm the average immune response of the human population), an endemic steady state will exist that can be globally stable. If this situation is only temporary, the iSIR model can thus account for isolated outbreaks of cholera, and if it persists, the model is suitable for regions where cholera cases are constant occurrences.

The original paper of the iSIR model [133] provided some preliminary mathematical results. This analysis adds on to that work, and demonstrates the local stability for most equilibria analytically. In addition, we present the results of dissipativity and determine conditions for global stability.

Further steps to take with this model would be to refine the condition on the global stability of the endemic equilibrium. The condition imposed might not

be required, and a biological explanation is in order. Also, a seasonal carrying capacity could be included to simulate the cycles of cholera which occur in regions like Bangladesh. Further altering the model to include bacteriophage is another possibility, with the idea being that the cycles observed in the human population are caused by cycles in the micro scale of bacteria and bacteriophage as has been suggested by Faruque et al. [129] and others.

# Chapter 6

## Dynamics of a cholera transmission model with immunological threshold and natural phage control in reservoir<sup>3</sup>

### Abstract

Cholera remains epidemic and endemic in the world, causing thousands of deaths annually in locations lacking adequate sanitation and water infrastruc-

---

<sup>3</sup>This Chapter has been published. Reference: Jude D. Kong, William Davis, and Hao Wang. “Dynamics of a cholera transmission model with immunological threshold and natural phage control in reservoir.” *Bulletin of Mathematical Biology*, Vol. 76: 2025-2051 (2014).

ture. Yet, its dynamics are still not fully understood. In this paper, we simplify and improve Jensen et al.'s model (PNAS, 2006) by incorporating a Minimum Infection Dose (MID) into the incidence term. We perform local stability analysis and provide bifurcation diagrams of the bacterial carrying capacity with or without shedding. Choosing parameters such that the endemic or epidemic equilibrium is unstable (as it is the case in reality), we observe numerically that for the bacterial carrying capacity ( $K$ ) less than the MID ( $c$ ), oscillating trajectories exist only in the microbial scale whereas for  $K > c$ , they exist in both the microbial and population scales. In both cases, increasing pathogen shed rate  $\xi$  increases the amplitude of the trajectories and the period of the trajectories for those that are periodic. Our findings highlight the importance of the relationship among the shedding rates,  $K$ , MID, the maximum bacterial growth rate ( $r$ ), and the features of the disease outbreak. In addition, we identified a region in the parameter space of our model that leads to chaotic behaviour. This could be used to explain the irregularity in the seasonal patterns of outbreaks amongst different countries, especially if the positive relationship between bacterial proliferation and temperature is considered. **Keywords:** cholera; indirect transmission; immunological threshold; phage; stability analysis.

## 6.1 Introduction

Cholera is a disease of the intestinal tract that causes severe diarrhoea, leading to dehydration which if left untreated can cause death. It is caused by the bacteria *Vibrio cholerae* and is treatable if caught within 1-2 days of symptoms

first appearing. In places with adequate health care and access to antibiotics, cholera is not much of a problem. However, in countries where such health services are lacking in a permanent sense or because of natural disasters reducing their availability, cholera outbreaks are still a concern. Dhaka, the capital of Bangladesh, for example has two outbreaks of cholera per year [131] that occur with the changes in seasons and the amount of rainfall, both of which affect the quality of the water supply.

Despite being studied for more than 100 years by the likes of English physician John Snow in the mid 1800's [139] and many others, the transmission dynamics of the disease are not fully understood. The role of blue-green algae, which is present in the water supplies of countries like Bangladesh, has been suggested by some [131] as a reservoir that the bacteria can exist in during inter-epidemic times. However this does not explain its natural cycles completely. One problem in understanding the dynamics is that *V. cholerae* is always found in lower levels than predicted, leading to possible explanations that *V. cholerae* can revert from a culturable form to a viable but not culturable form within the human body [126]. Experiments show that *V. cholerae* becomes many times more infectious for a short period of time once it has passed through the human digestive tract [130]. A more complete understanding of how exactly *V. cholerae* is transmitted through populations will certainly help to improve the control of outbreaks in the future for regions where it is an issue.

The contributions from mathematical modelling indicate that mathematical modelling is a promising way to look into the nature of the cholera dynamics. Many mathematical differential equation models have been proposed.

Among these, the main ones that make use of ordinary differential equations are those built upon the Cappasso-Fontona model [123] and the Codeço model [125]. Codeço model [125] is considered to be the first modern cholera model. The author divided the human population into three compartments; Susceptible, Infected and Recovered compartments. The recovered compartment was not started explicitly, as the population was assumed to be constant. The author equally assumed that there is no disease induced mortality. He represented the aquatic reservoir very simply with a linear growth term and linear shedding contribution. This was because the ecological dynamics of *V.cholerae* were not well understood at the time (they are still not completely understood). The oscillations in this system die out over time, so in order to simulate the periodic behaviour of outbreaks observed in some endemic areas, periodic contact rates, shedding rates and net growth rates of bacteria were also included. Hartley et al. [130] incorporated a hyperinfectious route of transmission to the Codeço model and Joh et al. [133], Tian et al. [140], Jensen et al. [132] and Mukanvire et al. [135] have further built on and branched off from these models.

Aimed at taking into account the role of bacteriophage which has been suggested by experimentalists as important, Jensen et al. [132] modified the Codeço model to include a phage compartment P, and dividing infectives into bacteria and phage infected and phage infected individuals. Bacteria are assumed to experience logistic growth with carrying capacity, and predation by phage occurs via a Holling I response. The infection term is a Holling III functional response. The sigmoidal shape is intended to capture the low infectivity of low levels of bacteria, but there are still infections at very small

levels which could overestimate the number of infections in the long run. The focus of this model was to determine the ability of phage to end outbreaks or indirectly cause outbreaks by being reduced in number. Both of these abilities were demonstrated in the analysis. However, they did not examine the role, or existence, of limit cycles caused by the predator-prey like relationship of phage and bacteria, which is something we will pay attention to in the present and absence of human shedding. In particular, we will show that cycles observed in the human population (macro scale) are caused by cycles in the micro scale of bacteria and bacteriophage as has been suggested by Faruque et al. [129] and others. Jensen et al. [132] also failed to take into account the fact that bacteria have to enter the human body in higher concentrations to overwhelm the natural immune response [136] (existence of a Minimum Infection Dose for bacteria), something that was first considered in modelling by Joh et al. [133]. We improve this model by incorporating a Minimum Infection Dose (MID) into the incidence term. This infection term is a piecewise continuous function which is zero below the minimum infectious dose (MID) threshold and a Holling II response curve above the threshold. Similar to Jensen et al. [132] we also allow bacteria to exist naturally under logistic growth. With these adjustments and a few others made to the model, we determine the invariant domain, carry out local stability analysis and locate limit cycles, with or without human shedding, and show numerically that by choosing parameters such that the endemic or epidemic equilibrium is unstable (as it is the case in reality), for the bacteria carrying capacity ( $K$ ) less than the MID ( $c$ ), oscillating trajectories exist only in the microbial scale whereas for  $K > c$ , they exist in both the microbial and population scales, and that in both cases, increasing

pathogen shed rate  $\xi$  increases the amplitude of the trajectories and the period of the trajectories for those that are periodic. Further, in this paper, we demonstrate the existence of a chaotic region in the parameter space, which could account for the different nature of outbreaks observed around the world.

## 6.2 Model formulation

Bacteria and bacteriophage exist in a predator-prey relationship. We capture this interaction by using a Holling II predation term  $\gamma \frac{B}{K_1+B} P$ , where  $\gamma$  is the maximum predation rate,  $B$  and  $P$  represent bacteria and phage densities respectively, and  $K_1$  is the half saturation constant of predation (the bacterial level at which predation occurs at half of the maximum rate). We assume that the bacterial population experiences logistic growth in the absence of predation and human influence, with carrying capacity  $K$  and maximum growth rate  $r$ . As in Jensen et al.'s model [132] and Codeço's model [125], the human population is assumed to be constant and there is assumed to be no disease-induced mortality. The lack of death due to infection as was the case in Dhaka during the 2004 outbreak, where all severe cases were treated (see [132] and reference therein), almost all endemic/epidemic regions nowadays have hospitals that provide the right treatment. No death is inevitable with cholera if the right treatment is provided. This motivates the following model,

which assumes no infection derived immunity for simplicity:

$$\begin{aligned}
 \frac{dS}{dt} &= -\alpha(B)S + \mu I, \\
 \frac{dI}{dt} &= \alpha(B)S - \mu I, \\
 \frac{dB}{dt} &= rB \left(1 - \frac{B}{K}\right) - \gamma \frac{B}{K_1 + B} P + \xi I, \\
 \frac{dP}{dt} &= \beta \gamma \frac{B}{K_1 + B} P - \delta P + \phi \xi I.
 \end{aligned}
 \tag{1.1}$$

The incidence term we use is  $\alpha(B)S$  where  $\alpha(B)$  is the bacterial density dependent component. The ‘indirect’ part of the incidence term  $\alpha(B)$  is defined by

$$\alpha(B) = \begin{cases} 0, & B < c; \\ \frac{\alpha(B-c)}{(B-c)+H}, & B \geq c. \end{cases}$$

Parameter	Values	Description	Units
r	0.3-14.3	Maximum per capita pathogen growth efficiency	day <sup>-1</sup>
K	10 <sup>5</sup> – 10 <sup>7</sup>	Pathogen carrying capacity	cell liter <sup>-1</sup>
H	10 <sup>6</sup> – 10 <sup>8</sup>	Half-saturation pathogen density	cell liter <sup>-1</sup>
a	0.1	Maximum rate of infection	day <sup>-1</sup>
ξ	0- 100	Pathogen shed rate	cell liter <sup>-1</sup> day <sup>-1</sup>
μ	0.1	Human recovery rate	day <sup>-1</sup>
N	10 <sup>6</sup>	Total Population	persons
c	10 <sup>5</sup> – 10 <sup>7</sup>	MID	cell liter <sup>-1</sup>
β	80-100	Phage burst size	virions day <sup>-1</sup>
γ	0 – 0.025	Maximum per capita phage absorption rate	cell virion <sup>-1</sup> day <sup>-1</sup>
δ	0.5-7.9	Phage death rate	virions day <sup>-1</sup>
φ	10 <sup>-6</sup> – 1	Mean phage shed rate	virions cell <sup>-1</sup>
K <sub>1</sub>	< K	Half saturation bacteria predation density	cell liter <sup>-1</sup>

**Table 6.1:** Parameter values expanded from Jensen et al. [132] and Cash et al. [124].

Unlike in larger scale predator prey dynamics, where  $\beta$  would be a measure of the conversion rate of prey into predators, often less than unity,  $\beta$  here represents a ‘burst size’, as each infected bacterial cell will give rise to many new phage cells. Human contamination of the water supply through infected

feces contributes to both bacteria and phage levels and is called ‘shedding’. Bacteria and phage shedding rates need not be the same so the rate for bacteria is  $\xi$  and for phage it is  $\phi\xi$  where  $\phi$  is some constant. In the absence of predators and humans, bacteria will exist at their carrying capacity  $K$ . We assume that phage and bacteria can live naturally without human interference, as in interepidemic times, and so it is assumed that  $\beta\gamma > \delta$ . If this is not so, phage would die out in the absence of human shedding. This maximum predation rate  $\gamma$  is difficult to measure, and for numerical solutions is chosen to satisfy this inequality. The half saturation constant for the predation term,  $K_1$ , was also estimated, and was assumed to be less than the natural carrying capacity  $K$  so that predation does not always occur near the maximal rate. For numerical simulations, parameters are taken from the literature and the ranges are given in Table 6.1.

### 6.3 Forward invariance

We would like to define a forwardly invariant set in which solutions of (1.1) will be bounded. From the first two equations of (1.1) we see that  $\dot{S}(S = 0) = \mu I$  but as  $S + I = N$ , we can write  $\dot{S}(S = 0) = \mu N > 0$ . Thus  $S(t) > 0$  for  $t > 0$ . Even though there is no birth or death in this system, if the entire population were to be infected then there would be people recovering and moving back into the susceptible category.

Similarly,  $\dot{I}(I = 0) = \alpha(B)S \geq 0$  as we just saw that  $S(t) > 0$  for  $t > 0$  and  $\alpha(B) \geq 0$  by definition. As  $S > 0$  and  $I \geq 0$  then as there are only two compartments for humans,  $S \leq N$  and  $I < N$ .

The BP system is more complicated as for upper bounds, but note that  $\dot{B}(B = 0) = \xi I$  thus  $B(t) \geq 0$ . We have that  $\dot{B} < rB(1 - \frac{B}{K}) + \xi N$  and so we can define

$$B_{max} = \frac{rK + K\sqrt{r^2 + \frac{4r}{K}\xi N}}{2r},$$

where if  $B(0) \in [0, B_{max})$  then  $B(t) \in [0, B_{max})$  for  $t \geq 0$ . Lastly, consider  $\dot{P}(P = 0) = \phi\xi I \geq 0$  and so  $P(t) \geq 0$  for all  $t > 0$ . The upper bound of  $P(t)$  requires the following lemma.

**Lemma 7.** *Define positive constants  $u$  and  $v$  such that  $\frac{((r+u)\beta)^2 K}{4r\beta} < v$ . Then for all values of  $B$  the following is true*

$$0 < \frac{r}{K}\beta B^2 - ((r+u)\beta)B + v.$$

We can now show that  $B$  and  $P$  are bounded above, although it was already demonstrated that  $B$  is bounded. Consider

$$\frac{d}{dt}(\beta B + P) < r\beta B - \frac{r}{K}B^2\beta - \delta P + (\beta + \phi)\xi N.$$

By invoking Lemma 7 we see that

$$\frac{d}{dt}(\beta B + P) < -U(\beta B + P) + (\beta + \phi)\xi N + v,$$

where  $U := \min\{u, \delta\}$ , which implies that  $\beta B + P$  is bounded. Defining

$V := (\beta + \phi)\xi N + v$  we can write

$$\limsup_{t \rightarrow \infty} \beta B(t) + P(t) \leq \frac{U}{V} \quad \text{or} \quad \beta B(t) + P(t) \leq \max \left\{ \beta B(0) + P(0), \frac{U}{V} \right\}.$$

We summarize the above results with a proposition

**Proposition 5** (Feasible Region). *The set*

$$\Gamma = \{S, I, B, P \geq 0 : S + I = N, \beta B(t) + P(t) \leq \frac{U}{V}, B < B_{max}\}$$

*defines a forwardly invariant region of system (1.1), where  $V := (\beta + \phi)\xi N + v$  and  $U := \min\{u, \delta\}$ , with  $u, v > 0$  satisfying  $\frac{((r+u)\beta)^2 K}{4r\beta} < v$ .*

## 6.4 Existence and stability of equilibria with no shedding

### 6.4.1 Existence of equilibria

In countries with modern sanitational infrastructure human contamination of the water supply (shedding) is very low; in the ideal case shedding is completely absent. We can determine the number and stability of steady states of (1.1) without shedding by substituting  $\xi = 0$  and further noting that as  $S = N - I$

the first equation is not necessary, leaving us with the following:

$$\begin{aligned}\frac{dI}{dt} &= \alpha(B)(N - I) - \mu I, \\ \frac{dB}{dt} &= rB \left(1 - \frac{B}{K}\right) - \gamma \frac{B}{K_1 + B} P, \\ \frac{dP}{dt} &= \beta \gamma \frac{B}{K_1 + B} P - \delta P.\end{aligned}\tag{1.2}$$

If the bacteria level is below the minimum infectious dose, then  $\alpha(B) = 0$ . The first equation of (1.2) implies that  $I^* = 0$  in this case, and so  $S^* = N$  as well. The second and third equations of (1.2) at steady state become

$$0 = rB \left(1 - \frac{B}{K}\right) - \gamma \frac{B}{K_1 + B} P \quad \text{and} \quad 0 = \left(\beta \gamma \frac{B}{K_1 + B} - \delta\right) P.$$

The second equation of (1.2) at steady state can be solved for  $P$ ;  $P = \frac{r}{\gamma K}(K - B)(K_1 + B) =: F_1(B)$ , having roots  $B = K$  and  $B = -K_1$ . The solution  $B = -K_1$  is not biologically relevant and also is not within our invariant region  $\Gamma$ , as defined in the previous section. The other root,  $B = K$ , of  $F_1(B)$  satisfies  $\alpha(K) = 0$  only if  $K \leq c$ . The third equation at steady state can be solved as well; either  $P = 0$  or  $B = B_1 := \frac{\delta K_1}{\beta \gamma - \delta}$ ; with the latter being relevant only if  $B_1 \leq c$ , i.e.  $\alpha(B_1) = 0$ . To satisfy both equations at once, either  $(B, P) = (0, 0), (K, 0)$  or  $(B_1, P_1)$ , where  $P_1 = F_1(B_1)$ . Thus when  $\alpha(B) = 0$  there are three possible steady states, all of which are disease free. The simplest equilibrium point occurs when  $S = N, I = 0, B = 0$  and  $P = 0$ . The disease free, bacteria free and phage free equilibrium  $E_0 = (N, 0, 0, 0)$  is always an equilibrium of (1.2) for all parameter values. The disease free, phage free equilibrium denoted  $E_K = (N, 0, K, 0)$  is an equilibrium if  $K \leq c$ . Similarly,

if  $B_1 \leq c$  then  $\alpha(B_1) = 0$  and the disease free equilibrium  $E_1 = (N, 0, B_1, P_1)$  exists. However, for the positivity of  $P_1$ , we require that  $B_1 < K$ . Note that if  $B_1 = K$ , then  $P_1 = 0$  and  $E_1$  is simply  $E_K$ . The case of equilibria when  $\alpha(B) \neq 0$  is more complicated, but it can be shown that there are up to two additional endemic equilibria denoted  $E_{1,K}^* = (S_{1,K}^*, I_{1,K}^*, B_{1,K}^*, P_{1,K}^*)$  where all of the entries are strictly positive, making  $E_{1,K}^*$  the only interior equilibria if they exist. If  $\alpha(B^*) \neq 0$ , first note that  $B^* > c$  by definition. Dropping the asterisk on  $B$ , the first equation of (1.2) at equilibrium implies

$$I^* = G_1(B) := Na \frac{(B - c)}{(a + \mu)(B - c) + \mu H}.$$

So for each equilibrium value  $B^*$  such that  $\alpha(B^*) \neq 0$ , there exists a unique value  $I^* = G_1(B^*)$ . To find  $B^*$  and  $P^*$ , it is of no consequence that  $I^* \neq 0$  because with  $\xi = 0$ , the second and third equations of (1.2) do not contain terms including  $I$ . Thus the nontrivial values of  $(B, P)$  that satisfy the second and third equations at steady state are the same as before:  $(B, P) = (K, 0)$  and  $(B_1, P_1)$ , but now  $K > c$  is required so that  $\alpha(K) \neq 0$  and  $B_1 > c$  so that  $\alpha(B_1) \neq 0$ . For the positivity of  $P_1$  it is still necessary that  $B_1 < K$ . Summarizing, there are up to two endemic equilibria of (1.2), denoted  $E_K^* = (S_K^*, I_K^*, K, 0)$  and  $E_1^* = (S_1^*, I_1^*, B_1, P_1)$ , where  $I_K^* = G_1(K)$ ,  $I_1^* = G_1(B_1)$  and  $S_i^* = N - I_i^*$ ,  $i = 1, k$ , with the condition that  $K > c$  and  $c < B_1 < K$  for  $E_K^*$  and  $E_1^*$  to exist respectively. As seen above, if  $B_1 = K$  this would mean  $P_1 = 0$  and  $E_K^* = E_1^*$ , leaving only one endemic equilibrium.

### 6.4.2 Linearization

Due to the threshold in the infection term, linearization yields two cases, one for  $B^* \leq c$ , denoted  $Jac_1$ , and one for  $B^* > c$ , denoted  $Jac_2$ .

$$Jac_1(I, B, P) = \begin{pmatrix} -\mu & 0 & 0 \\ 0 & r - 2\frac{r}{K}B - \frac{\gamma K_1}{(K_1+B)^2}P & -\gamma\frac{B}{K_1+B} \\ 0 & \beta\gamma\frac{aH}{(B-c+H)^2}P & \beta\gamma\frac{B}{K_1+B} - \delta \end{pmatrix},$$

and

$$Jac_2(I, B, P) = \begin{pmatrix} -\alpha(B) - \mu & (N - I)\frac{aH}{(B-c+H)^2} & 0 \\ 0 & r - 2\frac{r}{K}B - \frac{\gamma K_1}{(K_1+B)^2}P & -\gamma\frac{B}{K_1+B} \\ 0 & \beta\gamma\frac{aH}{(B-c+H)^2}P & \beta\gamma\frac{B}{K_1+B} - \delta \end{pmatrix}.$$

### 6.4.3 Stability of the disease free, bacteria free, phage free equilibrium $E_0$

Corresponding to  $E_0$ , we have the following eigenvalues  $\lambda = -\mu, -\delta < 0$  and  $r > 0$ . This means that the disease free, bacteria free and phage free equilibrium  $E_0$ , is a saddle node equilibrium with a one dimensional unstable manifold.

### 6.4.4 Stability of the disease free, phage free equilibrium $E_K$

Corresponding to  $E_K$  we have the following eigenvalues,  $\lambda = -\mu, -r, \beta\gamma\frac{K}{K_1+K} - \delta$ . If  $K < \frac{\delta K_1}{\beta\gamma - \delta} = B_1$  then all eigenvalues are negative and  $E_K$  is a stable

equilibrium, but if  $K > \frac{\delta K_1}{\beta\gamma - \delta} = B_1$  then  $E_K$  is a saddle node equilibrium with a one dimensional unstable manifold. Note that  $B_1 < K$  is required for  $E_1$  to exist, so the existence of  $E_1$  and the stability of  $E_K$  are contrary notions.

### 6.4.5 Stability of the disease free equilibrium $E_1$

For the Equilibrium point  $E_1$  we have the following Jacobian matrix

$$Jac_1(0, B_1, P_1) = \begin{pmatrix} -\mu & 0 & 0 \\ 0 & r - 2\frac{r}{K}B_1 - \frac{\gamma K_1}{(K_1+B_1)^2}P_1 & -\gamma\frac{B_1}{B_1+K_1} \\ 0 & \beta\gamma\frac{K_1}{(K_1+B_1)^2}P_1 & 0 \end{pmatrix}.$$

The matrix is slightly more complicated, so we will use a lemma from McCluskey and van den Driessche [134] with regard to three dimensional matrices.

**Lemma 8** (Lemma 3, McCluskey and van den Driessche (2003)). *Let  $A$  be a  $3 \times 3$  matrix with real entries. If  $tr(A)$ ,  $\det A$  and the determinant of the second additive compound matrix of  $A$ ,  $\det A^{[2]}$  are all negative, then all of the eigenvalues of  $A$  have negative real part.*

The converse of Lemma 8 is also true, which is apparent if you note that the eigenvalues of the second additive compound matrix  $A^{[2]}$ , for a  $3 \times 3$  matrix  $A$ , are just  $\sum \lambda_i + \lambda_j$  for  $i < j$  with  $\lambda_i$  being the eigenvalues of  $A$ .

Defining  $J(2, 2) = r - 2\frac{r}{K}B_1 - \gamma\frac{K_1}{(K_1+B_1)^2}P_1$ , the second additive compound of

$Jac_1$  at  $E_1$  is

$$Jac_1^{[2]}(0, B_1, P_1) = \begin{pmatrix} -\mu + J(2, 2) & -\gamma \frac{B_1}{B_1 + K_1} & 0 \\ \beta \gamma \frac{K_1}{(K_1 + B_1)^2} P_1 & -\mu & 0 \\ 0 & 0 & J(2, 2) \end{pmatrix}.$$

The determinant of  $Jac_1(E_1)$  is

$$\det Jac_1(E_1) = -\mu \gamma \frac{B_1}{B_1 + K_1} \beta \gamma \frac{K_1}{(K_1 + B_1)^2} P_1 < 0,$$

so it will always satisfy its role in the antecedent of Lemma 8. The trace is given by

$$tr(Jac_1(0, B_1, P_1)) = -\mu + J(2, 2).$$

If

$$\begin{aligned} J(2, 2) < 0 & \Rightarrow tr(Jac_1) < 0, \\ J(2, 2) > 0, -\mu + J(2, 2) < 0 & \Rightarrow tr(Jac_1) < 0, \\ J(2, 2) > 0, -\mu + J(2, 2) > 0 & \Rightarrow tr(Jac_1) > 0. \end{aligned}$$

Lastly we need to consider the sign of  $\det Jac_1^{[2]}(E_1)$ .

$$\det Jac_1^{[2]}(E_1) = J(2, 2) \left\{ [-\mu + J(2, 2)][-\mu] + \left( \beta \gamma \frac{K_1}{(K_1 + B_1)^2} P_1 \right) \left( \gamma \frac{B_1}{B_1 + K_1} \right) \right\}.$$

We can see that if  $J(2, 2) < 0$ , then  $\det Jac_1^{[2]} < 0$  and if  $J(2, 2) > 0$  but  $J(2, 2) - \mu < 0$  then  $\det Jac_1^{[2]} > 0$ . If  $J(2, 2) > 0$  and  $J(2, 2) - \mu > 0$  then the

antecedent of Lemma 8 will not be satisfied as  $tr(Jac_1(0, B_1, P_1)) > 0$ . Thus, because  $\det Jac_1(0, B_1, P_1) < 0$  all the time, by Lemma 8 we have that

$$\begin{aligned} E_1 \text{ is stable} &\iff J(2, 2) < 0 \\ &\iff K < \frac{\beta\gamma + \delta}{\beta\gamma - \delta} K_1 = B_1 + \beta\gamma \frac{K_1}{\beta\gamma - \delta} \end{aligned}$$

Define  $B_3 = B_1 + \beta\gamma \frac{K_1}{\beta\gamma - \delta}$  and note that as we assume  $\beta\gamma > \delta$  it follows that  $B_3 > 2B_1$ .

### 6.4.6 Stability of the phage free endemic equilibrium

$$E_K^*$$

Considering the equilibrium point  $E_K^*$ , we have that

$$Jac_2^{[2]}(E_K^*) = \begin{pmatrix} -\alpha(K) - \mu - r & -\gamma \frac{K}{K_1 + K} & 0 \\ 0 & -\alpha(K) - \mu + \left( \beta\gamma \frac{K}{K + K_1} - \delta \right) & (N - I^*) \frac{aH}{(K - c + H)^2} \\ 0 & 0 & -r + \left( \beta\gamma \frac{K}{K + K_1} - \delta \right) \end{pmatrix}.$$

$$tr(Jac_2(E_K^*)) = -\alpha(K) - \mu + \left( \beta\gamma \frac{K}{K + K_1} - \delta \right)$$

with

$$\det Jac_2(E_K^*) = [-\alpha(K) - \mu] [-r] \left[ \beta\gamma \frac{K}{K_1 + K} - \delta \right]$$

and

$$\det Jac_2^{[2]}(E_K^*) = [-\alpha(K) - \mu - r] \left[ -\alpha(K) - \mu + \beta\gamma \frac{K}{K + K_1} - \delta \right] \left[ -r + \beta\gamma \frac{K}{K_1 + K} - \delta \right].$$

Common to all three expressions is that if  $K < B_1$ , then they are each negative. And if  $K > B_1$  then  $\det Jac_2$  at  $E_K^*$  is positive. Hence  $E_K^*$  is stable  $\iff K < B_1$ . Note that the stability condition of  $E_K^*$  is contrary to the existence condition for  $E_1^*$  and so there can only ever be at most one locally stable endemic equilibrium point at a time.

### 6.4.7 Stability of the interior endemic equilibrium $E_1^*$

Lastly considering the local stability of interior endemic equilibrium  $E_1^*$ , we have that

$$Jac_2^{[2]}(I_1^*, B_1, P_1) = \begin{pmatrix} -\alpha(B_1) - \mu + J(2, 2) & -\gamma \frac{B_1}{K_1 + B_1} & 0 \\ \beta \gamma \frac{K_1}{(K_1 + B_1)^2} P_1 & -\alpha(B_1) - \mu & (N - I^*) \frac{aH}{(B_1 - c + H)^2} \\ 0 & 0 & J(2, 2) \end{pmatrix}.$$

$$tr(Jac_2(I_1^*, B_1, P_1)) = -\alpha(B_1) - \mu + J(2, 2)$$

and

$$\det Jac_2(I_1^*, B_1, P_1) = [-\alpha(B_1) - \mu] \gamma \frac{B_1}{B_1 + K_1} \beta \gamma \frac{K_1}{(K_1 + B_1)^2} P_1 < 0.$$

So the determinant is always negative, and the trace can be negative if  $J(2, 2) < 0$ . We now consider the sign of the determinant of  $Jac_2^{[2]}$  evaluate at  $(I_1^*, B_1, P_1)$ .

$$\det Jac_2^{[2]}(I_1^*, B_1, P_1) = J(2, 2) \left\{ [-\alpha(B_1) - \mu + J(2, 2)] [-\alpha(B_1) - \mu] + \beta \gamma \frac{K_1}{(K_1 + B_1)^2} P_1 \gamma \frac{B_1}{B_1 + K_1} \right\}$$

which appears complicated but the important part is that if  $J(2, 2) < 0$ , then

it will be negative. If  $0 < J(2, 2) < \alpha(B_1) + \mu$  implies that  $\det Jac_2^{[2]}(E_1^*) > 0$ , and if  $J(2, 2) > \alpha(B_1) + \mu$  implies that both  $\det Jac_2^{[2]}(E_1^*) > 0$  and  $tr(E_1) > 0$ . As seen previously,  $J(2, 2) < 0$  is equivalent to  $B_1 < B_3$ , thus

$$E_1^* \text{ is stable} \quad \iff \quad B_1 < B_3 = B_1 + \frac{\beta\gamma}{\beta\gamma - \delta}.$$

### 6.4.8 Local stability summary, bifurcation diagrams and numerical simulations

Noting that there are at most 2 equilibria that exist at any one time other than  $E_0$ , and writing ‘un’ for locally unstable and ‘s’ for locally asymptotically stable, we can summarize the preceding local stability results with a proposition.

**Proposition 6** (Local stability of the non-shedding case).  *$E_0$  always exists and is locally stable for all parameter values.*

---

If  $c < K$  and  $c \geq B_1$

$B_3 \leq K$  implies  $E_1(un)$  and  $E_K^*(un)$  exist,

$K < B_3$  implies  $E_1(s)$  and  $E_K^*(un)$  exist

and  $c < B_1$

$K < B_1$  implies  $E_K^*(s)$  exists

$B_1 < K < B_3$  implies  $E_K^*(un)$  and  $E_1^*(s)$  exist

$B_3 \leq K$  implies  $E_K^*(un)$  and  $E_1^*(un)$  exist.

If  $c \geq K$  and  $c \geq B_1$

$K < B_1$  implies  $E_K(s)$  exists

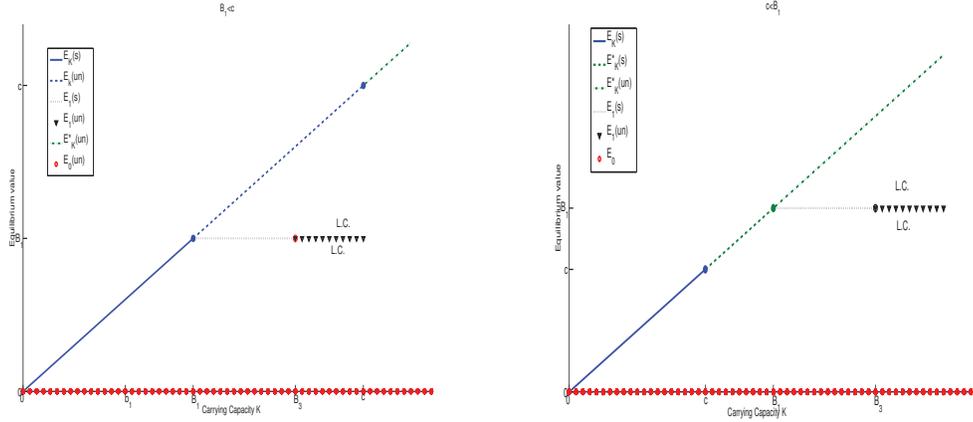
$B_1 < K < B_3$  implies  $E_K(un)$  and  $E_1(s)$  exist

$B_3 \leq K$  implies  $E_K(un)$  and  $E_1(un)$  exist

and  $c < B_1$

as  $B_1 < B_3$  then  $E_K(s)$  exists.

The results of Proposition 6 are perhaps better understood as a bifurcation diagram. Figure 6.1 demonstrates the changes in stability as the carrying capacity  $K$  is varied. The first diagram is for the case when the minimum infectious dose  $c$  is greater than  $B_1$ , which means that only  $E_1$  can exist, and not  $E_1^*$ . The lower figure has  $c < B_1$ , which reverses the situation.



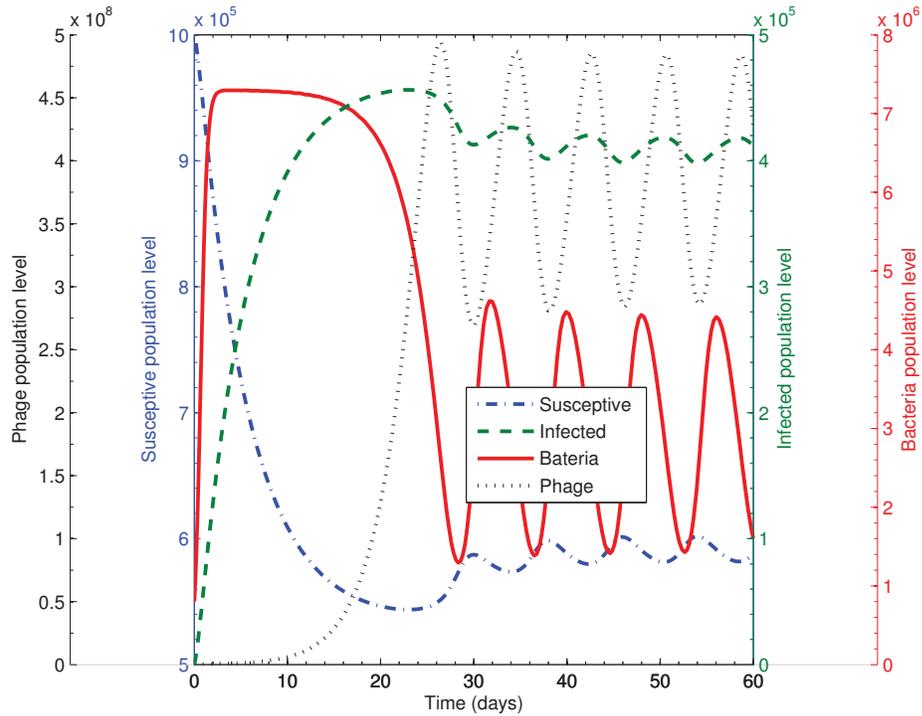
**Figure 6.1:** Bifurcation diagrams when  $\xi = 0$  and there is no shedding. Limit cycles exist when  $E_1$  and  $E_1^*$  undergo Hopf bifurcations, and are denoted L.C.

If  $K = B_3$ , implying  $J(2, 2) = 0$ , then computing  $\det[\lambda I - Jac_1(E_1)]$  we find that,

$$\det(\lambda I - Jac_1(E_1)) = (\lambda + \alpha(B_1) + \mu) \left\{ \lambda^2 + \beta\gamma \frac{K_1}{(B_1 + K_1)^2} P_1 \frac{\gamma B_1}{B_1 + K_1} \right\}$$

which has one real negative and two purely imaginary roots. We conclude that  $E_1$  undergoes a Hopf bifurcation as  $K$  passes  $B_1$ , and  $E_1$  changes from locally stable to unstable. In Figure 6.2 we demonstrate the existence of limit cycles occurring as a result of the unstable  $E_1$ . The stability conditions for the endemic equilibrium  $E_1^*$  are the same as for  $E_1$  and it also undergoes a Hopf bifurcation when it exists and  $K$  increases past  $B_3$ . The only difference in the calculation of the eigenvalues of  $Jac_2(E_1^*)$  is that the second entry in  $Jac_2$  is nonzero, but as  $\xi = 0$  the zeros in the first column reduce the calculation of the eigenvalues of  $Jac_2(E_1^*)$  to that shown above. When  $E_1$  or  $E_1^*$  was unstable and the carrying capacity  $K$  was less than the MID ( denoted

*c*), cycles were observed numerically in the bacteria-phage system (BP) but not the susceptible-infected (SI) system. If  $K$  was sufficiently larger than  $c$ , implying that the MID was at a level such that bacteria at carrying capacity would cause infections, the cycles existed in both the SI and BP systems with the infected population peaking 4 days after and the phage population 8 days after in Figure 6.2. While these cycles are far too short to match real world situations, as the infected class peaks occurred after the bacteria class, and because with shedding at zero the BP system influences the SI system unidirectionally, their existence does support the idea that cycles that naturally occur are bottom-up and not top-down in cause. When shedding is included, the cycles lengthen to relevant levels as we shall see in the next section.



**Figure 6.2:** Cycles in Phage, Bacteria and Infected populations. The Bacteria peak ends first, followed by the Infected 4 days later and Phage 8 days later. The parameters are  $r = 3$ ,  $K = 7.3e6$ ,  $\gamma = 0.02$ ,  $K_1 = 1.6e6$ ,  $\beta = 80$ ,  $\delta = 1$ ,  $\xi = 0$ ,  $a = 0.1$ ,  $c = 7.1e6$ ,  $\mu = 0.1$  and  $H = 1e6$ .

## 6.5 Existence and stability of equilibria with shedding

### 6.5.1 Existence of equilibria

The complete absence of human contamination of the water supply is the ideal, but it is certainly not the reality anywhere and particularly not in places where the disease is endemic. Noting that the first equation of (1.1) is not necessary

as  $S + I = N$ , we can rewrite it as follows:

$$\begin{aligned}\frac{dI}{dt} &= \alpha(B)(N - I) - \mu I, \\ \frac{dB}{dt} &= rB \left(1 - \frac{B}{K}\right) - \gamma \frac{B}{K_1 + B} P + \xi I, \\ \frac{dP}{dt} &= \beta \gamma \frac{B}{K_1 + B} P - \delta P + \phi \xi I.\end{aligned}\tag{1.3}$$

If  $\alpha(B) = 0$ , we will have the same endemic equilibria  $E_0 = (N, 0, 0, 0)$ ,  $E_K = (N, 0, K, 0)$  and  $E_1 = (N, 0, B_1, P_1)$  as before, with the same conditions.

If  $\alpha(B) \neq 0$  at steady state, then from the first equation of (1.3), we have that  $I^* = G_1(B^*) = \frac{Na(B^* - c)}{(a + \mu)(B^* - c) + \mu H}$ , and from the third equation we have that  $P^* = \frac{-\phi \xi I^*}{\beta \gamma \frac{B^*}{K_1 + B^*} - \delta} = \frac{-\phi \xi G_1(B^*)(K_1 + B^*)}{\beta \gamma B^* - \delta(K_1 + B^*)}$ . This expression for  $P^*$  provides a condition for  $B^*$  as the denominator must be strictly negative in order to have a well defined and positive value for  $P^*$ . That is  $B^* < \frac{K_1 \delta}{\beta \gamma - \delta} = B_1$ . Solving the second equation of (1.3) at equilibrium for  $B^*$  assuming that  $\alpha(B) \neq 0$ , will then possibly lead to endemic equilibria  $E^*$ . Dropping the asterisks on  $B^*$  for convenience, the second equation at steady state becomes

$$0 = rB \left(1 - \frac{B}{K}\right) + \xi I^* - \gamma \frac{B}{K_1 + B} P^* = F(B) + G(B)\tag{1.4}$$

where

$$F(B) := rB \left(1 - \frac{B}{K}\right) [(a + \mu)(B - c) + \mu H][(\beta \gamma - \delta)B - \delta K_1]$$

$$G(B) := Na \xi [\gamma(\phi + \beta)B - \delta K_1 - \delta B](B - c).$$

Note that  $F(B)$  is a quartic with roots  $0$ ,  $K$ ,  $B_2 := c - \frac{\mu H}{a + \mu}$  and  $B_1$  which was defined previously as  $B_1 = \frac{\delta K_1}{\beta\gamma - \delta}$ , that opens downwards. The three roots of  $0$ ,  $B_1$  and  $K$  are nonnegative, but  $B_2$  could be negative or zero with realistic parameters. To solve  $0 = F(B) + G(B)$  it suffices to find the intersections of  $F(B)$  and  $-G(B)$ . As such note that  $-G(B)$  is a downward opening parabola, with roots  $c$  and  $b_1 := \frac{\delta K_1}{\gamma(\phi + \beta) - \delta}$ .

The roots of the two functions have some obvious relationships which limit the number of possibilities we need to consider when looking for points where the two functions intersect. Consider  $b_1$  and  $B_1$  which are clearly related. We assume  $\beta\gamma - \delta > 0$  and as  $\phi > 0$ , being part of the shedding term for the phage population  $\phi\xi$ , it is clear that  $0 < b_1 < B_1$ . Also  $B_2 < c$  as all parameter values are positive. Previously we found that  $c < B^* < B_1$  to ensure  $P^* > 0$  and  $\alpha(B^*) > 0$ , so this implies  $c < B_1$  is a condition for any  $B^*$  to exist.

Between the third and fourth roots of  $F$ , we see that  $F(B) > 0$ , whatever those roots may be. If the largest root is  $B_1$ , then we find a problem if  $c$  is the next largest of  $\{b_1, c, K, B_1, B_2\}$ . As  $-G(c) = 0$  with  $-G(B) < 0$  for  $B \geq c$ , and  $F(c) > 0$ , then clearly no intersections can occur until  $B > B_1$  when  $F(B)$  is no longer nonnegative. Any such intersection would be inadmissible as  $B^* > B_1$  for that  $B^*$ . Also, any intersections between  $F$  and  $-G$  with  $B < c$  are also inadmissible as we require  $B^* > c$ . Hence, if  $c < B_1$  and  $\{c, B_1\}$  are the largest of  $\{b_1, c, K, B_1, B_2\}$  then there will be no endemic equilibria.

Case	Subcase	Ordering
Ia)		
$0 < B_2 < B_1 < K$	i) $B_2 < b_1 < B_1$	$B_2 < b_1 < c < B_1$ $B_2 < c < b_1 < B_1$
	ii) $0 < b_1 < B_2$	$b_1 < B_2 < c < B_1$
Ib)		
$0 < B_2 < K < B_1$	i) $b_1 > K$	$K < c < b_1 < B_1$ $c < K < b_1 < B_1$
	ii) $B_2 < b_1 < K$	$B_2 < b_1 < c < K$ $B_2 < c < b_1 < K$
	iii) $b_1 < B_2$	$b_1 < B_2 < c < K < B_1$
Ic)		
$K < B_2 < B_1$	i) $B_2 < b_1 < B_1$	$B_2 < c < b_1 < B_1$
II a)		
$B_2 < 0 < B_1 < K$	i) $b_1 < B_1$	$b_1 < c < B_1 < K$ $c < b_1 < B_1 < K$
II b)		
$B_2 < 0 < K < B_1$	i) $K < b_1$	$K < c < b_1 < B_1$ $c < K < b_1 < B_1$
	ii) $b_1 < K$	$b_1 < c < K < B_1$ $c < b_1 < K < B_1$

**Table 6.2:** Possible ordering of  $\{B_2, B_1, K, b_1, c\}$ . The first column determines the order of  $\{B_2, B_1, K\}$ , the second places  $b_1$  in that ordering, and the third column places  $c$  within the ordering.

The possible orderings of  $\{B_2, B_1, K, b_1, c\}$  are outline in Table 6.2. There are only 9 combinations of  $\{B_2, B_1, K, b_1\}$  and 15 orderings of all of the roots

of  $F$  and  $G$  when all of the restrictions are considered. There can be up to 4 equilibria at one time depending on the relationships among the parameters. We can summarize these results with a proposition.

**Proposition 7** (Existence of Equilibria). *The equilibrium  $E_0 = (N, 0, 0, 0)$  always exists.*

---

If  $c \geq K$  and  $B_1 \leq c$ ,

if also  $B_1 > K$ , then only  $E_K$  exists.

if  $B_1 \leq K$  then  $E_K$  and  $E_1$  exist.

and  $B_1 > c$ ,

if  $c < b_1$  then there are up to  $B_{1,2}^* \in (c, b_1)$  and  $E_K$ .

if  $c > b_1$  then there are no internal equilibria but  $E_K$  exists.

If  $c < K$  and  $B_1 > c$

then  $B^*$  exists between the second and third in the ordering of  $\{c, K, b_1, B_1\}$ .

and  $B_1 \leq c$

there are no internal equilibria and  $E_1$  is an equilibrium.

---

## 6.5.2 Linearization

Due to the threshold in  $\alpha(B)$ , we will have two linearizations of (1.3) with the first having  $\alpha(B) = 0$ , denoted  $J_1$ :

$$J_1(I, B, P) = \begin{pmatrix} -\mu & 0 & 0 \\ \xi & r - 2\frac{r}{K}B - \frac{\gamma K_1}{(K_1+B)^2}P & -\gamma\frac{B}{K_1+B} \\ \phi\xi & \beta\gamma\frac{aH}{(B-c+H)^2}P & \beta\gamma\frac{B}{K_1+B} - \delta \end{pmatrix},$$

and the second linearization applies when  $\alpha(B) \neq 0$ , denoted  $J_2$

$$J_2(I, B, P) = \begin{pmatrix} -\alpha(B) - \mu & (N - I) \frac{aH}{(B-c+H)^2} & 0 \\ \xi & r - 2\frac{r}{K}B - \frac{\gamma K_1}{(K_1+B)^2}P & -\gamma \frac{B}{K_1+B} \\ \phi\xi & \beta\gamma \frac{aH}{(B-c+H)^2}P & \beta\gamma \frac{B}{K_1+B} - \delta \end{pmatrix}.$$

### 6.5.3 Stability of disease free, bacteria free, phage free equilibrium $E_0$

Corresponding to  $E_0$ , we have the following eigenvalues  $-\mu$ ,  $-\delta < 0$  and  $r > 0$ . This means that the equilibrium  $E_0$  is a saddle node equilibrium with a one dimensional unstable manifold.

### 6.5.4 Stability of the boundary equilibrium $E_K$

Considering the equilibrium point  $E_K$ , we have that

$$J_1^{[2]}(0, K, 0) = \begin{pmatrix} -\mu - r & \left(-\gamma \frac{K}{K_1+K}\right) & 0 \\ 0 & -\mu + \beta\gamma \frac{K}{K_1+K} - \delta & 0 \\ -\phi\xi & \xi & -r + \left(\beta\gamma \frac{K}{K_1+K} - \delta\right) \end{pmatrix}$$

$$\text{tr}(J_1(0, K, 0)) = -\mu - r + \left(\beta\gamma \frac{K}{K_1+K} - \delta\right), \det J_1(0, K, 0) = \mu r \left(\beta\gamma \frac{K}{K_1+K} - \delta\right)$$

and

$\det J_1^{[2]} = (-\mu - r) \left(-\mu + \beta\gamma \frac{K}{K_1+K} - \delta\right) \left(-r + \beta\gamma \frac{K}{K_1+K} - \delta\right)$ . If  $\beta\gamma \frac{K}{K_1+K} - \delta < 0$  this implies that  $\text{tr}(J_1)$ ,  $\det J_1$  and  $\det J_1^{[2]}$  are all negative, but if  $\beta\gamma \frac{K}{K_1+K} - \delta > 0$  then  $\det J_1 > 0$ . By Lemma 8 this means that  $E_K$  is stable if, and only if  $K < \frac{\delta K_1}{\beta\gamma - \delta} = B_1$ . Note that  $B_1 < K$  is required for  $E_1$  to exist,

so the existence of  $E_1$  and the stability of  $E_K$  are contrary notions.

### 6.5.5 Stability of the disease free equilibrium $E_1$

Defining  $J(2, 2) = r - 2\frac{r}{K}B_1 - \gamma\frac{K_1}{(K_1+B_1)^2}P_1$  again as in the previous non-shedding case, the second additive compound matrix of  $J_1$  is

$$J_1^{[2]}(0, B_1, P_1) = \begin{pmatrix} -\mu + J(2, 2) & -\gamma\frac{B_1}{B_1+K_1} & 0 \\ \beta\gamma\frac{K_1}{(K_1+B_1)^2}P_1 & -\mu & 0 \\ -\phi\xi & \xi & J(2, 2) \end{pmatrix}.$$

The determinant of  $J_1$ ,  $\det J_1(0, B_1, P_1) < 0$ , so it will always satisfy the antecedent of Lemma 8. The trace is given by  $tr(J_1(0, B_1, P_1)) = -\mu + J(2, 2)$ .

As before  $J(2, 2) < 0$  if, and only if,  $K < B_3 = B_1 + \frac{\beta\gamma K_1}{\beta\gamma - \delta}$ . If

$$\begin{aligned} J(2, 2) < 0 & \Rightarrow tr(J_1) < 0 \\ J(2, 2) > 0, -\mu + J(2, 2) < 0 & \Rightarrow tr(J_1) < 0 \\ J(2, 2) > 0, -\mu + J(2, 2) > 0 & \Rightarrow tr(J_1) > 0. \end{aligned}$$

Lastly consider the sign of  $\det J_1^{[2]}$

$$\det J_1^{[2]}(0, B_1, P_1) = J(2, 2) \left\{ [-\mu + J(2, 2)][-\mu] + \left( \beta\gamma\frac{K_1}{(K_1+B_1)^2}P_1 \right) \left( \gamma\frac{B_1}{B_1+K_1} \right) \right\},$$

where it is clear that if  $J(2, 2) < 0$  then  $\det J_1^{[2]} < 0$  and if  $J(2, 2) > 0$  but  $J(2, 2) - \mu < 0$  then  $\det J_1^{[2]} > 0$ . If  $J(2, 2) > 0$  and  $J(2, 2) - \mu > 0$  then the antecedent of Lemma 8 will not be satisfied as  $tr(J_1(0, B_1, P_1)) > 0$ . Thus since  $\det J_1(0, B_1, P_1) < 0$  all the time, by Lemma 8 we have that  $E_1$  is stable  $\iff$

$J(2, 2) < 0$  i.e  $E_1$  is stable if and only if  $K < \frac{\beta\gamma+\delta}{\beta\gamma-\delta}K_1 = B_1 + \beta\gamma\frac{K_1}{\beta\gamma-\delta} = B_3$

### 6.5.6 Stability of endemic equilibria $E^*$ and $E_{1,2}^*$

The stability of the endemic steady states was found numerically. If  $c > b_1$  and  $c < B_1$ , there is a unique endemic equilibrium,  $E^*$ . Using parameter values  $r = 1, \gamma = 0.02, K_1 = 3.6e5, \beta = 80, \delta = 1, \xi = 50, \phi = 1, a = 0.1, H = 1e6, c = 5.9e5$  and  $\mu = 0.1$ , will achieve such a relationship. If  $K = 7e5 > c$  or larger,  $E_K$  does not exist and  $E_0$  and  $E^*$  are the only equilibria. Limit cycles are observed and if  $B^*$  and its eigenvalues are computed numerically, we see that  $E^*$  is a saddle-node equilibrium with a two-dimensional unstable manifold.

If  $K = 4e5 < c = 5e5$ , with all other parameters the same, then  $K < c < b_1$  and two endemic equilibria exist, along with  $E_0$  and  $E_K$ . In this case they are both saddle-node equilibria, where  $E_1^*$  (with the smaller  $B^*$  value) has a one-dimensional unstable manifold, and the other has a two-dimensional unstable manifold. If  $K = 5.9e5 > c$  or greater, there is a unique endemic equilibria and  $E_K$  no longer exists. As before, it is a saddle-node with two dimensional unstable manifold.

Lastly, if  $K_1$  is decreased to  $K_1 = 1.6e5$ , with all other parameters as before, if  $c > b_1$  and the endemic equilibrium is unique, it is again a saddle-node with two-dimensional unstable manifold. If instead  $c < b_1$ , then when the two internal equilibria exist,  $E_1^*$  (with the smaller  $B^*$ ) is a saddle-node with one dimensional unstable manifold as before. The larger however, is stable as its eigenvalues all have negative real part. Finally, if  $K > c$ , for example  $K = 2.2e6$  or higher, there is a unique endemic equilibrium  $E^*$  which is locally

stable as all eigenvalues have negative real part.

### **6.5.7 Local stability summary, bifurcation diagrams and numerical simulations**

We can summarize the local stability results of the previous sections with a proposition, writing ‘un’ for locally unstable, and ‘s’ for local asymptotic stability. The goal of the bifurcation diagrams below is to exhibit all possible cases of the model, thus their associated parameter ranges are wider than the variation of parameter values in simulations.

**Proposition 8** (Local Stability).  *$E_0$  is always locally unstable.*

---

If  $c \geq K$  and  $B_1 \leq c$ ,

if  $K > B_3$ , then  $E_1(un)$  and  $E_K(un)$ .

if  $B_1 < K < B_3$  then  $E_1(s)$  and  $E_K(un)$ .

if  $K < B_1$  then  $E_K(s)$ .

and  $B_1 > c$ ,

if  $c < b_1$  then there are up to  $E_{1,2}^*$  and  $E_K(s)$ ,

with  $B_i^* \in (c, b_1)$ .

if  $c > b_1$  then there are no internal equilibria but  $E_K(s)$ .

If  $c < K$  and  $B_1 > c$ ,

then  $E^*$  exists,

with  $B^*$  between the second and third in the ordering of  $\{c, K, b_1, B_1\}$ .

and  $B_1 \leq c$ ,

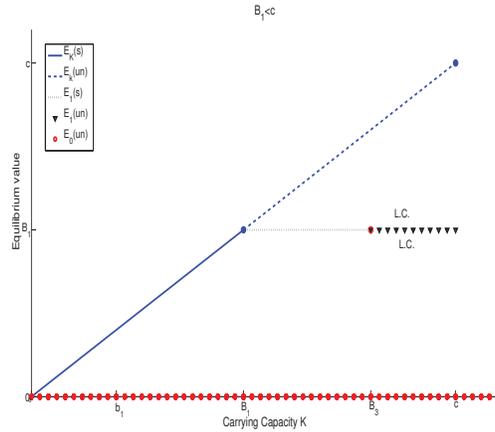
if  $K > B_3$ , then  $E_1(un)$ .

if  $K < B_3$ , then  $E_1(s)$ .

---

In addition to the infection free, bacteria free and phage free equilibrium  $E_0$ , which always exists and is always locally unstable, there are at most three other equilibria for any given set of parameters. Note that  $E_1$  and  $E_K$  are never both stable at the same time, as the condition for the local stability of  $E_K$  implies that  $E_1$  does not exist. Of the two,  $E_1$  is more realistic as it has the phage population existing at nonzero levels, which is certainly the case during inter-epidemic times. The existence of a stable endemic equilibrium, either when  $E^*$  is unique, or when it exists with another, which is unstable, does not match the usual pattern of explosive outbreaks of cholera, but if the  $B^*$  level

is low enough, perhaps it could be biologically relevant for certain areas. Our main interest is on the existence of limit cycles, as will be discussed below. The results of Proposition 8 are perhaps better understood with bifurcation diagrams. Figure 6.3 shows the case when  $B_1 < c$ , and only nonendemic equilibria are possible. In the figure  $B_3 < c$ , which means that both  $E_1$  and  $E_K$  can be unstable at the same time. If this was reversed and  $B_3 > c$ , the difference would be that  $E_1$  would be unstable only when  $E_K$  does not exist, and the two could not be unstable for the same set of parameters.



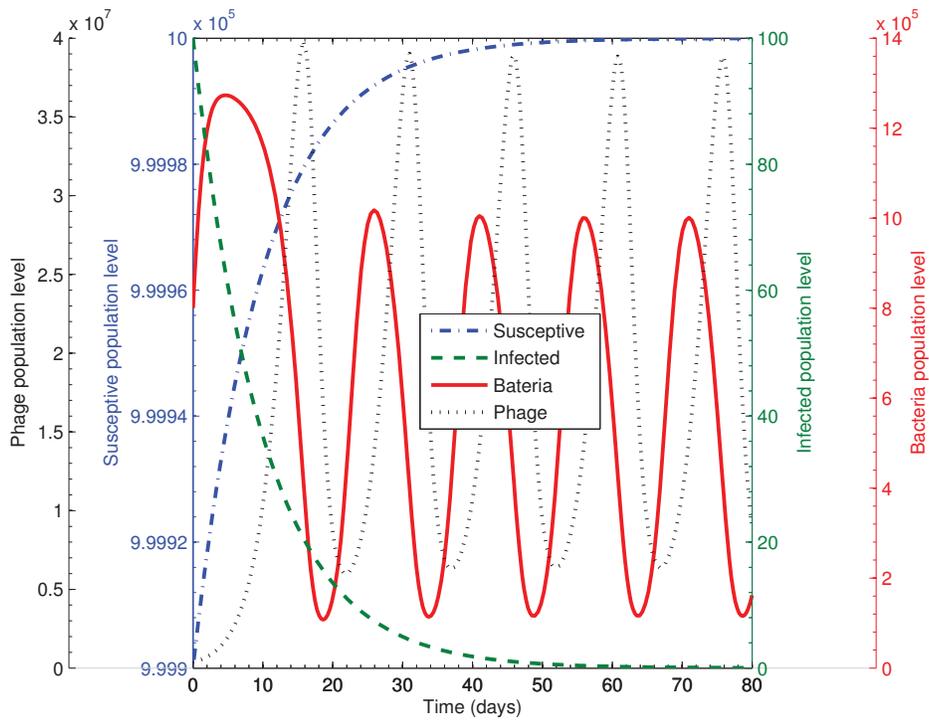
**Figure 6.3:** Bifurcation diagrams with all parameters positive and  $B_1 < c$ , which implies only nonendemic equilibria exist. Equilibrium  $E_1$  undergoes a Hopf bifurcation when carrying capacity  $K$  increases past  $B_3$ , leading to limit cycles denoted L.C.

Equilibrium  $E_1$  is only present in the first diagram, and when the carrying capacity  $K = B_3$ , we can calculate  $\det[\lambda I - J_1(E_1)]$ , noting that  $J(2, 2) = 0$  to see that

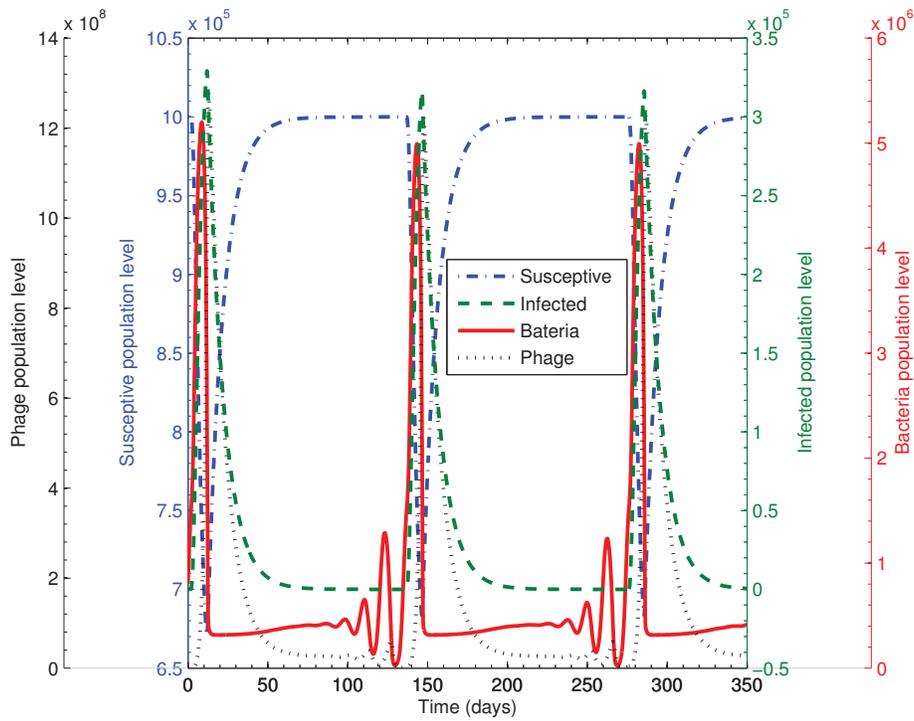
$$\det[\lambda I - J_1(E_1)] = (\lambda + \mu) \left\{ \lambda^2 + \beta\gamma \frac{K_1}{(K_1 + B_1)^2} P_1 \gamma \frac{B_1}{B_1 + K_1} \right\},$$

and observe that  $J_1(E_1)$  would have one negative eigenvalue and two purely

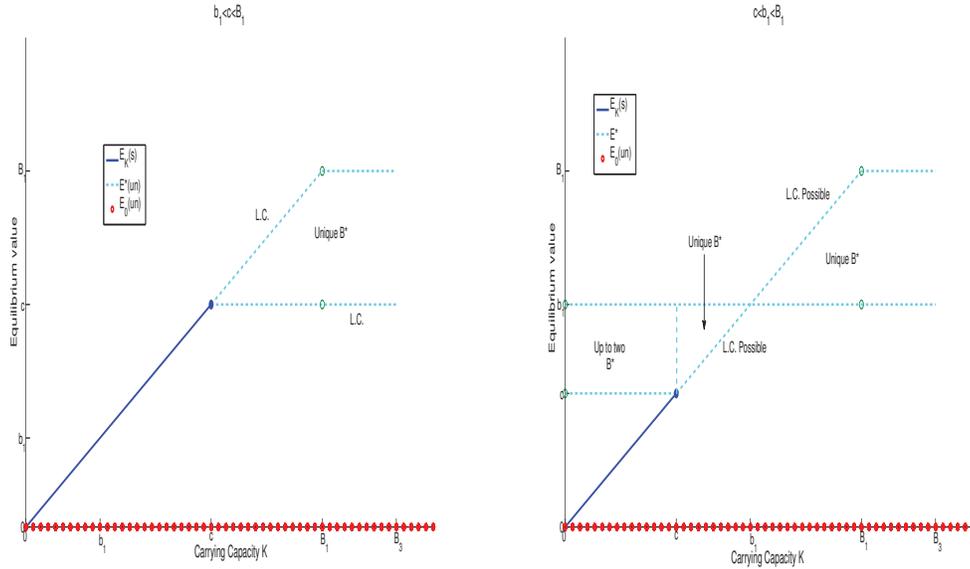
imaginary eigenvalues. Thus  $E_1$  undergoes a Hopf bifurcation as  $K$  increases past  $B_3$  and  $E_1$  switches from locally stable to unstable. With parameters in the region where  $E_1$  is unstable, we found limit cycles to exist. If  $K < c$ , then these cycles existed only in the BP community and did not cause any infections. Figure 6.4 demonstrates such limit cycles, with period of only 14 days, and phage peaking 5 days after the bacteria class does. If  $K > c$  by a large enough amount, meaning that the minimum infectious dose is less than the normal carrying capacity of bacteria, the cycles entered the human population as well and increase greatly in period. Unlike the case with  $\xi = 0$ , the period of these cycles could even be approximately 180 days, which could correspond to the biannual outbreaks observed in some endemic areas. Figure 6.5 is an example of of such cycles with period of 150 days. The bacteria are the first to peak, followed by the human infected population 3 days later, and the phage 1 day after the infected class. As these cycles can exist at low levels and only enter the human population when the bacteria levels increase passed the MID, and because the bacteria peak before the infected human population, we conclude that the BP system is ‘driving’ these limit cycles.



**Figure 6.4:** Dynamics of I, S, B and P when  $E_1$  is unstable and  $K < c$ . When the bacteria levels pass the minimum infectious dose (MID), the cycles spread to the human population as well. The period is approximately 15 days and the Phage peak 4 days the Bacteria. The parameters are  $r = 1, K = 1.3e6, \gamma = 0.02, K_1 = 2.5e5, \beta = 80, \delta = 1, \xi = 50, \phi = 1, a = 0.1, H = 1e6, c = 1.5e6$  and  $\mu = 0.1$ .



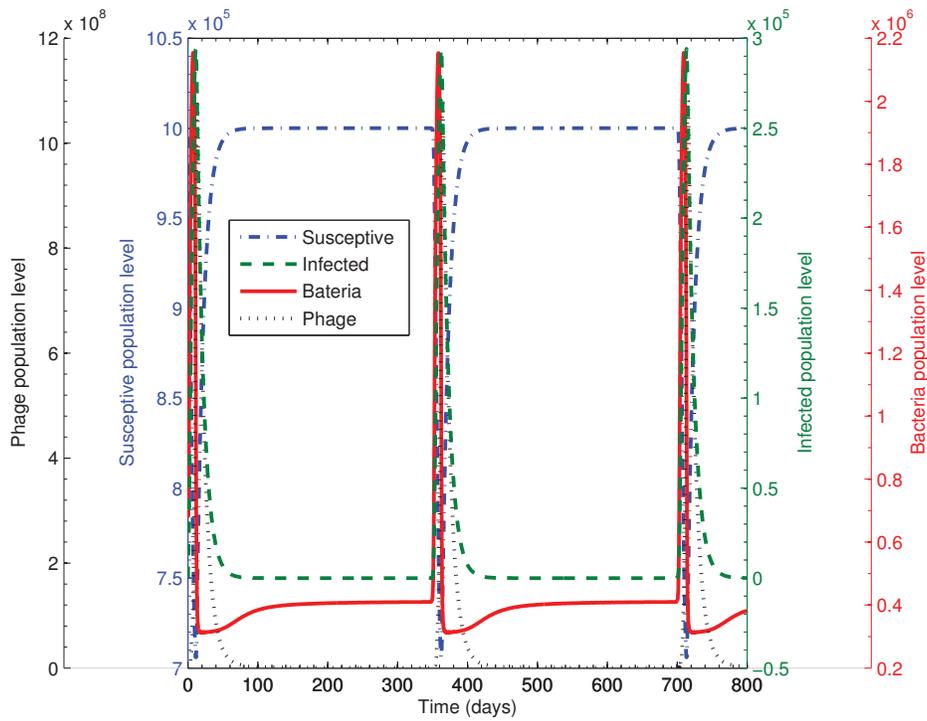
**Figure 6.5:** Dynamics of I, S, B and P when  $E_1$  is unstable and  $K > c$ . When the bacteria levels pass the minimum infectious dose (MID), the cycles spread to the human population as well. The Infected class peaks 3 days after the Bacteria, and the Phage 4 days after. The parameters are  $r = 1, K = 1.8e6, \gamma = 0.02, K_1 = 2.5e5, \beta = 80, \delta = 1, \xi = 50, \phi = 1, a = 0.1, H = 1e6, c = 1.5e6$  and  $\mu = 0.1$ . The period is approximately 150 days which corresponds to biannual outbreaks in endemic areas.



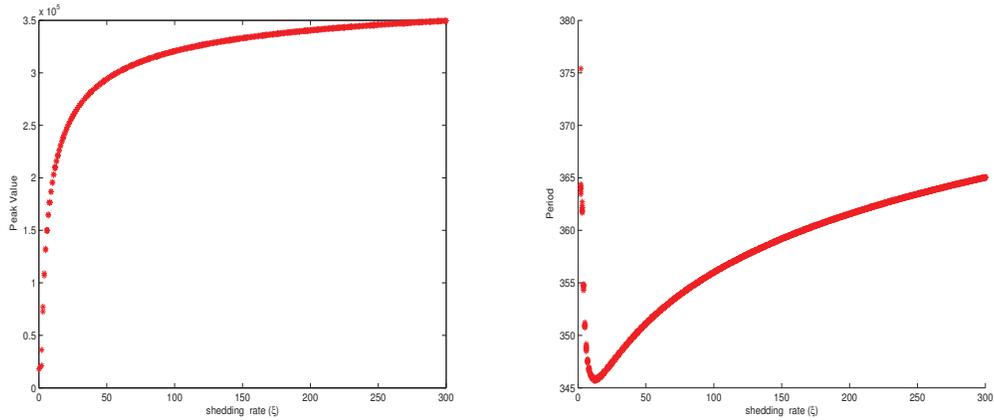
**Figure 6.6:** Bifurcation diagrams with all parameters positive. Limit cycles exist after when  $E^*$  is unique and unstable and are denoted L.C. When  $c > b_1$ ,  $E^*$  was always unstable and there were cycles. When  $c < b_1$ , these cycles existed when  $E^*$  was unstable, but were absent when it was stable. If there were two equilibria  $E_{1,2}^*$ , they were either both unstable, or the one with the smaller  $B_i^*$  was unstable and the larger was stable. Limit cycles were not observed with parameters in this range.

Figure 6.6 contains bifurcation diagrams for the cases of endemic equilibria. The first diagram of Figure 6.6 is for when  $E^*$  is unique, and numerically it was found to always be unstable and causing limit cycles. The second diagram is very similar, except up to two  $E_{1,2}^*$  can exist. These are either both unstable, or the equilibrium with the smaller  $B_i^*$  value was unstable and the larger was stable. The unique  $E^*$  in the second diagram could be either stable or unstable for realistic parameter values, and when it was unstable limit cycles were found to exist. These cycles ranged in period, but could be found with periods of approximately 360 days, as in Figure 6.7, which correspond to the annual outbreaks observed in some endemic areas. The period of outbreaks and maximum number of people infected in an outbreak, differs from one endemic re-

gion to another depending largely on sanitational infrastructures. Figure 6.8a, shows that whenever an outbreak occurs in an endemic/epidemic region with poor sanitational infrastructures, many people are infected compared to when it occurs in an endemic region with better sanitational infrastructures. The number of infected persons increases monotonically as  $\xi$  increases from 0 to 300. Thus improving the sanitational infrastructure of an endemic region could lead to a reduction in the number of infected persons whenever an outbreak occurs. This alone, is unable to eradicate cholera. Figure 6.8b, shows the effect of poor sanitational infrastructure on the period of the outbreaks. By increasing  $\xi$  from 0 to 300, the period of the outbreaks decreases monotonically as  $\xi$  moves from 0 to 15 and attends a minimum value of approximately 346 days and there after, increases monotonically as  $\xi$  goes above 15. The fact that the period of the outbreaks decreases as the shedding rate increases from 0 to 15, is counterintuitive.



**Figure 6.7:** Dynamics of I, S, B and P when when  $E^*$  is unique and unstable with  $K > c$ . When the bacteria levels pass the minimum infectious dose (MID), the cycles spread to the human population as well. The Infected class peaks 3 days after the Bacteria, and the Phage 6 days after. The parameters are  $r = 1, K = 4.12e5, \gamma = 0.02, K_1 = 2.5e5, \beta = 80, \delta = 1, \xi = 50, \phi = 1, a = 0.1, H = 1e6, c = 4.1e5$  and  $\mu = 0.1$ . The period is approximately 351 days, which corresponds to annual outbreaks in endemic areas.



(a) The effect of an increase in  $\xi$  on the peak value of I cycles when  $E^*$  is unique and unstable with  $K > c$ . (b) The effect of an increase in  $\xi$  on the period of outbreaks when  $E^*$  is unique and unstable with  $K > c$ .

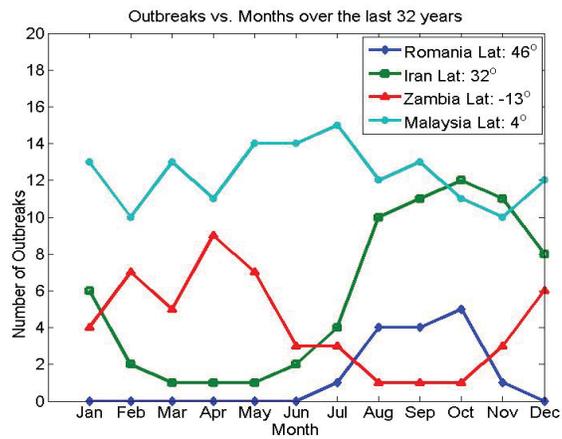
**Figure 6.8:** The parameters are  $r = 1, K = 4.12e5, \gamma = 0.02, K_1 = 2.5e5, \beta = 80, \delta = 1, \phi = 1, a = 0.1, H = 1e6, c = 4.1e5$  and  $\mu = 0.1$ . For these parameter values, each  $\xi$  value leads to an oscillating trajectory with a unique period.

## 6.6 Chaos

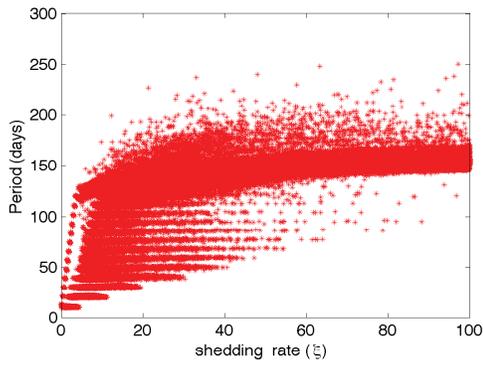
In this section we will attempt an explanation to the different nature of the outbreak experience around the world. We will attempt to explain why you might have countries with the same sanitational infrastructure but the outbreak in one might be sporadic whereas those in the other are periodic. For instance, Malaysia and Zambia have approximately the same sanitational infrastructure but the outbreaks in Malaysia are sporadic whereas those in Zambia are periodic.

In many countries that experience endemic cholera, there are annual cholera outbreaks which appear to be periodic. However, in countries with similar sanitation infrastructure the outbreaks are much more frequent and lack an overwhelmingly periodic structure. The general trend is that countries closer

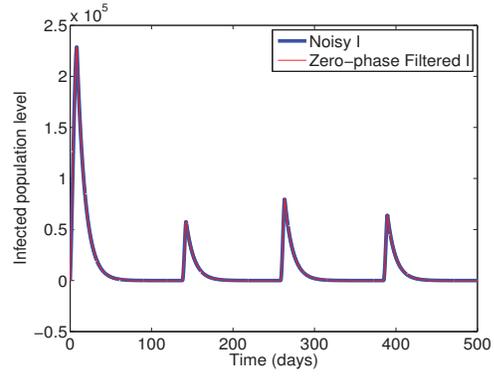
to the equator have higher levels of outbreaks with greater frequency, while countries that are further from the equator typically have seasonal outbreaks [128]. An explanation for this trend may lie in the existence of chaotic behaviour in (1.1) for certain values of the shedding parameters  $\xi$  and  $\phi$ . This window of chaos depends on other parameters in (1.1) as well, and not just  $\xi$  and  $\phi$ . The maximal growth rate of bacteria  $r$  is proportional to the values  $\xi_c$  and  $\phi_c$  where chaos first occurs. If this value of  $r$  is itself proportionate to average temperatures, and thus inversely proportional to the distance from the equator, then warmer countries with a higher  $r$  value could have chaotic behaviour of bacterial levels, and thus outbreaks with the same values of  $\xi$  and  $\phi$ . A positive relationship between bacteria proliferation and average temperature is known to exist, so this explanation is plausible [138]. Figure 6.9 demonstrates different trends in cholera outbreaks for countries at different latitudes. Malaysia for example is the closest to the equator of the four countries shown, at a latitude of  $4^\circ$ , and has a somewhat uniform distribution of monthly outbreaks when summed over 32 years. The other three countries of Romania, Iran and Zambia which are at a distance of at least  $\pm 13^\circ$  from the equator, have much stronger trends in what month cholera outbreaks typically occur. For Romania and Iran, which are both in the Northern Hemisphere, outbreaks typically occur between August and November. In Zambia, which is in the Southern Hemisphere, outbreaks occur most often between February and May. A larger value of the maximal bacterial growth rate  $r$  for countries closer to the equator, which also corresponds to a lower value of  $\xi_c$  and  $\phi_c$ , could explain why the outbreaks in warmer countries occur less seasonally than in countries further away from the equator.



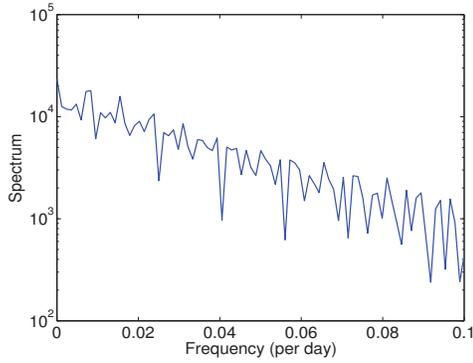
**Figure 6.9:** Sums of monthly cholera outbreaks over the last 32 years in countries at different latitudes, adapted from Emch et al. [128].



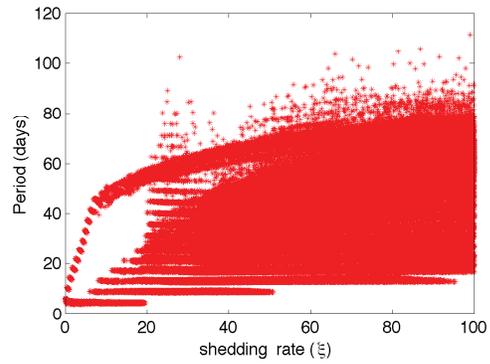
(a) shedding rate vs period with  $r=1$ .



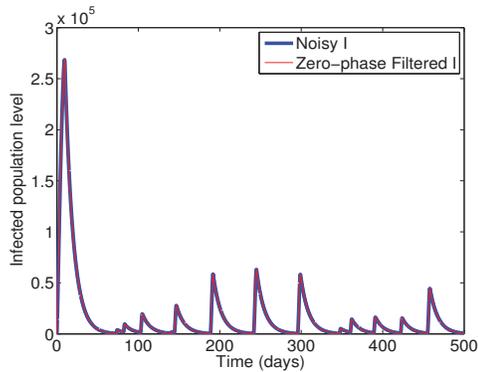
(b) trajectory for  $r=1$  and  $\xi = 11$ .



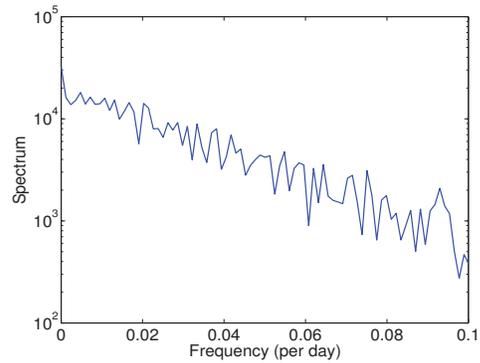
(c) single-sided amplitude spectrum of filtered I with  $r=1$ ,  $\xi = 11$ .



(d) shedding rate vs period with  $r=5$ .



(e) trajectory for  $r=5$  and  $\xi = 53$ .

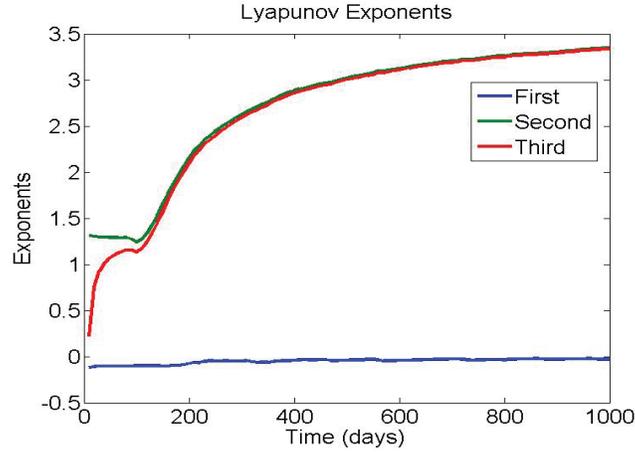


(f) single-sided amplitude spectrum of filtered I with  $r=5$ , and  $\xi = 53$ .

**Figure 6.10:** Chaotic behaviour with two different  $r$  (maximum bacterial growth) values. The remaining parameters are  $K = 1e6$ ,  $\gamma = 0.021$ ,  $K1 = 1/4K$ ,  $\beta = 100$ ,  $\delta = 1$ ,  $\phi = 0$ ,  $a = 0.1$ ,  $c = 5e5$ ,  $\mu = 0.1$ ,  $H = 1e6$ .

In the panels (a) and (d) of Figure 6.10, the shedding rate  $\xi$  is increased by 0.01 after each time step, and for each value, the peak values of the trajectory produced are numerically determined and used to estimate the length of each cycle in it. For the maximal growth rate  $r$  comparatively low, we have trajectories oscillating with varying peak values and periods occurring for lower values of  $\xi$ . In Figure 6.10a,  $r$  is comparatively low and for most values of  $\xi$  approximately less than 30, the corresponding trajectory has varying periods (See Figure 6.10b for the behaviour of the trajectories for a typical  $\xi$  in this interval). On the contrary, for  $r$  comparatively high, oscillating trajectories with varying peak values and periods occur for higher values of  $\xi$  as can be seen in Figure 6.10d where  $r=5$  and for  $\xi$  approximately greater than 20, the trajectory for each  $\xi$  has varying peak values and length of cycles (see Figure 6.10e for the behaviour of the trajectories for a  $\xi$  taken from this interval). Thus the same values of  $\xi$  and  $\phi$  could cause different trajectories for different values of  $r$ . In Figure 6.10, it is the unstable  $E_1$  that causes the cycles, but an unstable  $E^*$  could also be used. Figure 6.10c and Figure 6.10f present the Fast Fourier Transform (FFT) of the filtered infected population trajectories in Figure 6.10b and Figure 6.10e respectively. These FFT show continuous spectrum of frequencies. This behaviour is typical of chaotic motions. However, it is not enough to determine if the behaviour is chaotic: we should consider the Lyapunov Exponents. In Figure 6.11 we see that due to the positivity of the largest Lyapunov exponent, that the behaviour is in fact chaotic. Furthermore the largest two exponents are both positive and almost the same value. They correspond to the bacteria and phage categories from the model, which suggests that it is the bacteria-phage system which drives the cyclic behaviour

of the entire system and that they are the most sensitive to perturbations.



**Figure 6.11:** Lyapunov Exponents for (1.1). The parameters used are the same as those in Figure 10(a) with  $\xi = 11$ . The largest is positive which is enough to determine the trajectory is chaotic.

## 6.7 Discussion

We have presented a model, which is an extension of the one in Jensen et al. [132]. This model explicitly includes the dynamics of bacteriophage and bacteria and also contains a new indirect infection term which accounts for a minimum infectious dose of the pathogen *V.cholerae*. Unlike Jensen et al. [132], we focused on the existence of stable limit cycles, in order to account for the periodicity observed in outbreaks of cholera in endemic areas. As these cycles exist in the absence of human contribution to the bacteria and phage levels, and because the bacteria cycles peak before the human cycles when they exist in both systems, we conclude that it is the bacteria and phage which are driving the cycles, and not the reverse situation. If the minimum infectious dose is less than the carrying capacity of the bacteria, we observe that the bacteria cycles usually fail to surpass the minimum infectious dose,

so there are no new infections and the system is disease free. However, if the natural carrying capacity is sufficiently larger than the minimum infectious dose, these cycles are able to enter the human population, which highlights the importance of understanding the relationship between the two. Additionally, as a control measure if the phage levels could be enhanced in some way to keep the bacteria below this minimum infectious dose, then the cycles would remain in the bacteria and phage system alone. This idea links back to the 1930s when the use of injections of bacteriophage was explored as a treatment of cholera by limiting *V.cholerae* levels within the human host [121, 137].

Additionally, a chaotic region in the parameter space was identified. The existence of chaotic behaviour could explain the lack of clear periodicity in some endemic areas, with seasonal or other factors increasing the height of these chaotic peaks annually or biannually and creating a pseudo-periodic pattern. The exact role of these external factors would be difficult to determine, given the sensitivity of such a system. As the existence of this chaotic parameter region can be positively correlated with the proliferation rate of *V.cholerae* and overall climate, it could also explain the unpredictable nature of outbreaks in countries nearer the equator.

Future work on the model could be to explicitly include the role of infection derived immunity through the use of a recovered class, even though immunity is somewhat accounted for in the value of the minimum infectious dose. Exact conditions for the existence of limit cycles would be valuable as well as a definitive relationship between the amplitude and period of the cycles to other parameters in the system. This would be useful in establishing useful connections between simulations and the data. Furthermore, including

a second disease causing serogroup of *V.cholerae* would increase the realism of the model for use in regions where outbreaks caused by serogroups O1 and O139 occur simultaneously.

## Chapter 7

# The inverse method for a childhood infectious disease model with its application to pre-vaccination and post-vaccination measles data <sup>4</sup>

### Abstract

In this paper, we improve the classic *SEIR* model by separating the juvenile group and the adult group to better describe the dynamics of childhood

---

<sup>4</sup>This Chapter has been published. Reference: Jude D. Kong, Chaochao Jin, and Hao Wang. “The inverse method for a childhood infectious disease model with its application to pre-vaccination and post-vaccination measles data.” *Bulletin of Mathematical Biology*, Vol. 77: 2231-2263 (2015)

infectious diseases. We perform stability analysis to study the asymptotic dynamics of the new model, and perform sensitivity analysis to uncover the relative importance of the parameters on infection.

The transmission rate is a key parameter in controlling the spread of an infectious disease as it directly determines the disease incidence. However, it is essentially impossible to measure the transmission rate for certain infectious diseases. We introduce an inverse method for our new model, which can extract the time-dependent transmission rate from either prevalence data or incidence data in existing open databases. Pre- and post-vaccination measles data sets from Liverpool and London are applied to estimate the time-varying transmission rate.

From the Fourier transform of the transmission rate of Liverpool and London, we observe two spectral peaks with frequencies 1/year and 3/year. These dominant frequencies are robust with respect to different initial values. The dominant 1/year frequency is consistent with common belief that measles is driven by seasonal factors such as environmental changes and immune system changes and the 3/year frequency indicates the superiority of school contacts in driving measles transmission over other seasonal factors. Our results show that in coastal cities the main modulator of the transmission of measles virus, paramyxovirus is school seasons. On the other hand, in landlocked cities, both weather and school seasons have almost the same influence on paramyxovirus transmission.

**Keywords:** childhood infectious disease, time-dependent transmission rate, incidence algorithm, prevalence algorithm, inverse problem, measles, Fourier

transform, sensitivity analysis, vaccination.

## 7.1 Introduction

Many infectious diseases such measles, whooping cough, chickenpox, polio, mumps and rubella take special interest in children. The major reason for this is that they have not yet developed immunity to them. Although vaccination has reduced the infection from these diseases and eradicated some of them like small pox, they continue to kill thousands of children around the world every year. Thus understanding childhood disease transmission and control remains essential. Because there are enough infection data, mathematical models are increasingly useful for these purposes. Most of these models are compartmental models which classify the population with respect to various stages of disease infection.

The classical compartmental model for childhood infectious diseases is the McKendrick-Kermack SIR model with vital dynamics. This model has three compartments:

- **Susceptible  $S(t)$ :** individuals who are currently susceptible to the disease;
- **Infectious  $I(t)$ :** individuals who are currently infected with the disease;
- **Removed  $R(t)$ :** individuals who have recovered from the disease and therefore possess immunity.

It is assumed that the birth rate and the natural death rate are both  $\mu$ . This assumption keeps the population size constant at a normalized value one.

Infected children recover at a constant rate  $\nu$  and then become immune to the disease. Finally, it is assumed that the incidence is proportional to the product of the number of susceptible and infected individuals. These lead to the following system of equations:

$$\begin{aligned}\frac{dS(t)}{dt} &= \mu - \beta(t)S(t)I(t) - \mu S(t), \\ \frac{dI(t)}{dt} &= \beta(t)S(t)I(t) - \nu I(t) - \mu I(t), \\ \frac{dR(t)}{dt} &= \nu I(t) - \mu R(t),\end{aligned}$$

where  $S(t)+I(t)+R(t) = 1$ . Many authors have extended this model to include individuals who are exposed (E) to infection, but are not yet infectious. They assumed that children leave the exposed compartment and enter the infectious (I) compartment at some constant rate  $a$ , the reciprocal of which equals the mean latent period. This model, called the SEIR model, is described by the following system of equations:

$$\begin{aligned}\frac{dS(t)}{dt} &= \mu - \beta(t)S(t)I(t) - \mu S(t), \\ \frac{dE(t)}{dt} &= \beta(t)S(t)I(t) - aE(t) - \mu E(t), \\ \frac{dI(t)}{dt} &= aE(t) - \nu I(t) - \mu I(t), \\ \frac{dR(t)}{dt} &= \nu I(t) - \mu R(t).\end{aligned}$$

The rate at which susceptibles become infected is called the transmission rate  $\beta$ . This transmission rate depends on factors such as the the frequency

and closeness of contacts, the infectivity of the infectious individuals, and the susceptibility of susceptible individuals. Thus the transmission rate will be higher when children are packed together and lower when they are not. Since children turn to be crowded together in school seasons and separated during holidays, the transmission rate will therefore be extremely high when children are in school and extremely low when they are on holidays. This means that the transmission rate for a childhood disease varies dramatically in time.

According to Section 3.4.9 of Anderson and May [141], ‘... the direct measurement of the transmission rate is essentially impossible for most infections. But if we wish to predict the changes wrought by public health programmes, we need to know the transmission rate ...’. Thus unlike other parameters of infectious disease models such as recovery rate, birth rate and death rate, that can be easily measured directly via public health databases, it is essentially impossible to measure the transmission rate directly for certain infectious diseases. Moreover, other parameters are relatively stable, while the transmission rate varies dramatically. Given that the important events of disease transmission are incorporated in the transmission rate, there is urgent need to have an estimate for this parameter. Some researchers model the transmission rate using a step function based on school calendars [144]. Other authors use the sinusoidal function  $\beta(t) = \beta_0(1 + \alpha \cos(2\pi t))$ , where  $\beta_0$  is the mean transmission rate and  $\alpha$  is the amplitude of the seasonal variation [147].

These models do not take into account all the seasonal factors that drive the transmission rate of childhood infectious diseases. Even though the neglected factors are of little importance, including them in the model will paint a complete picture of the nature of the transmission rate of these diseases.

One additional drawback of these models is the lack of transmission data to validate them with. There is therefore a need to extract the time dependent transmission rate through a solution of an inverse problem so as to have a complete picture of the transmission rate of these diseases and and equally validate the assumed transmission rate functions in literature.

Some researchers [143, 145] used the discrete time SI or SIR model to extract the time dependent transmission rate. [145] used the model:

$$I(t + 1) = I(t)S(t)\beta(t)$$

$$S(t + 1) = S(t) - I(t + 1) + B(t) - V(t)$$

Where  $I(t)$  and  $I(t + 1)$  denote the infected population at time periods  $t$  and  $t+1$ ,

$S(t + 1)$  and  $S(t)$  the susceptible population at time periods  $t$  and  $t+1$ ,

$\beta(t)$  the time dependent transmission parameter,

$B(t)$  the number of susceptible introduced or born into the population,

$V(t)$  the vaccinated population. From this model the authors obtained the recursive formula:

$$\beta(t) = \frac{I(t + 1)}{\beta(t)S(t)}.$$

This formula requires  $S(t)$ , which is often difficult to estimate especially for outbreaks in which a reasonable percentage of the population is immune to infection. Also the formula is not explicit. For other drawbacks of this formula, see [152, 153, 148]. In 2010 and 2012, Pollicott, Wang, and Weiss [152, 153] originally introduced the inverse method for estimating a contin-

uous time transmission rate from prevalence data. In 2011, Haderler [148] extended the inverse method for incidence data and other possibilities. This inverse method does not require the knowledge of  $S(t)$  and leads to an explicit formula for  $\beta(t)$ . The inverse method for deriving these algorithms is applicable to a majority of infectious diseases, but the algorithms have only been constructed and applied to pre-vaccination data so far. One can reduce the number of susceptible children with vaccines. Vaccination is the administration of antigenic material to stimulate an individual's immune system to develop adaptive immunity to a pathogen. Vaccines can ameliorate both mortality and morbidity. The effectiveness of vaccination has been widely studied and verified since the first work of Edward Jenner on smallpox [151]. In this paper, we first derive the algorithms with vaccination which is dominantly important in the control of infectious diseases nowadays Assuming that all children grow to adults at some point in the future, in this paper, we extend the SEIR model for childhood diseases to an SEIRA model by adding the adult compartment (A). The SEIRA model is extended to include a compartment for children that have been vaccinated and the effect of vaccination on the dynamics of  $I(t)$  is studied. We carry out stability and sensitivity analysis of these two models and extend both the prevalence and incidence algorithms respectively.

Our sensitivity analysis emphasizes the importance of the transmission rate in controlling outbreaks and thus the need to estimate this key parameter. The applicability of our derived incidence algorithm is illustrated with pre- and post-vaccination measles data from Liverpool and London. The transmission rate estimated using our algorithm has two dominant spectral peaks

of frequencies 1 and 3 times per year. These dominant frequencies are the same for pre- and post-vaccination situations for both cities. The dominant frequency of 1 per year is consistent with common belief that measles is driven by seasonal factors and the 3 times per year frequency indicates the superiority of school contacts in driving measles transmission over other seasonal factors. The peak values of 1 per year and the 3 per year frequencies are comparable for both pre- and post-vaccination data from Liverpool, while the 1 per year peak value is larger than the 3 per year peak value for both pre- and post-vaccination data from London. This is because London is a landlocked city and thus has relatively high temperature variations which strongly affect the seasonality of the measles virus with subsequent influence on its transmission. Liverpool, on the other hand, is a coastal city with relatively stable temperature and thus the main modulator of the transmission of measles virus for this city was school dates. We find that the dominant frequencies of the Fourier transform of the transmission rate in London have less noise than that of the Fourier transform of the transmission rate in Liverpool. This could be attributed to the city size as London is a much larger city than Liverpool.

## 7.2 The SEIRA model

Recall that in the derivation of the *SEIR* model, the total population is divided into four compartments: susceptible, exposed, infective and recovered. This model can be applied to all infectious diseases satisfying its assumptions. However, it is not suitable for childhood infectious diseases since, in the *SEIR* model, adults who have never been infected nor vaccinated are also considered

as susceptibles.

When studying childhood infectious diseases, we first classify the population into the adult ( $A$ ) and the juvenile groups. Then, divide the juvenile group into susceptibles ( $S$ ), exposed ( $E$ ), infective ( $I$ ), and recovered ( $R$ ).

With SEIR model we consider natural death rate for every group. But children usually do not die naturally. They die only because of some specific reasons like accidents or diseases and their death rate is much lower than the natural death rate. Therefore, here, we ignore natural death rate for the juvenile group. Instead, we consider growth rate as children will grow up and no longer be susceptible. We assume that children grow to become adults at a rate  $g$ , i.e. people under  $1/g$  years old will be considered as juvenile. Transition terms between  $S$ ,  $E$ ,  $I$  and  $R$  are the same as in  $SEIR$  model.

Our model, called the SEIRA model, is described by the following system of equations:

$$\begin{aligned}
 \frac{dS(t)}{dt} &= \delta A(t) - \beta(t)S(t)I(t) - gS(t), \\
 \frac{dE(t)}{dt} &= \beta(t)S(t)I(t) - aE(t) - gE(t), \\
 \frac{dI(t)}{dt} &= aE(t) - \nu I(t) - gI(t), \\
 \frac{dR(t)}{dt} &= \nu I(t) - gR(t), \\
 \frac{dA(t)}{dt} &= g(S(t) + E(t) + I(t) + R(t)) - \delta A(t).
 \end{aligned}
 \tag{7.1}$$

The parameters are described in Table 7.1, which also contains the default values of these parameters for measles.

Parameter	Value	Description	Units
$p$	0-100%	Vaccinated fraction	no unit
$\delta$	1/64	Natural death rate	year <sup>-1</sup>
$\beta$	$\approx 1000$	Average transmission rate	year <sup>-1</sup>
$g$	1/16	Growth rate	year <sup>-1</sup>
$\nu$	52	Recovery rate	year <sup>-1</sup>
$a$	52	Rate at which exposed individuals become infective	year <sup>-1</sup>

**Table 7.1:** Parameter descriptions and values for measles.

Values of the parameters  $a$  and  $\nu$  for measles are taken from [141, 155]:  $a = 52/year$  and  $\nu = 52/year$ . Thus the infectious period is the same as the exposed period which is  $1 year/52 = 1 week$ . We assume that the average life span is 80 years and only kids under 16 are susceptible to measles. Therefore,  $g = 1/16 \cdot year^{-1}$  and natural death rate  $\delta = 1/64 \cdot year^{-1}$ . Values of  $p$  and  $q$  are mainly determined by a vaccination policy, which can be as low as 0% if there is no vaccination or as high as 100% if every individual is vaccinated.  $p$  will be used in Section 7.7.

In sections 7.3 and 7.4 we assume that  $\beta$  is a constant and in Sections 7.5 and 7.6, it is considered to be a function of time.

### 7.3 Qualitative analysis

In this section, we list qualitative results such as positivity, boundedness, equilibria and their stability of the *SEIRA* model. The proofs of the theorems in this section are presented in Appendix.

**Theorem 7.** *The compact set  $\Omega = \{(S, E, I, R, A): S \geq 0, E \geq 0, I \geq 0, R \geq 0, A \geq 0, S + E + I + R + A = 1\}$  is positively invariant for the semiflow generated by system (7.1).*

*Proof.* Positivity describes the property that for any positive initial values, the solution of a system will stay positive. The solution of an ODE  $\frac{dy_i}{dt} = f_i(y_1, y_2, \dots, y_n)$  ( $n \in \mathbb{N}$ ) is said to be positive for any positive initial values if  $\forall 1 \leq i \leq n$ ,  $f_i \geq 0$  when  $y_i = 0$  and  $y_j \geq 0$ ,  $j \neq i$ . Denoting  $S, E, I, R, A$  as  $f_i$  and the functions on the right hand side of (7.1) as  $y_i$ , for  $1 \leq i \leq 5$ , respectively we have that  $f_1 = \delta y_5 \geq 0, f_2 = \beta y_1 y_3 \geq 0, f_3 = a y_2 \geq 0, f_4 = \nu y_3 \geq 0, f_5 = g(y_1 + y_2 + y_3 + y_4) \geq 0$ . Hence the solution of (7.1) is positive for any positive initial values. Also, we notice that  $S, E, I, R$ , and  $A$  sum to 1 and their derivatives sum to 0. Hence we can conclude that the solution of (7.1) will stay in  $\{(S, E, I, R, A): S \geq 0, E \geq 0, I \geq 0, R \geq 0, A \geq 0, S + E + I + R + A = 1\}$  for any positive initial values.  $\square$

**Theorem 8.** *System (7.1) has two equilibria: the disease-free equilibrium point  $(S_1^*, E_1^*, I_1^*, R_1^*, A_1^*) = \left(\frac{\delta}{g+\delta}, 0, 0, 0, \frac{g}{g+\delta}\right)$  and the endemic equilibrium point*

$$(S_2^*, E_2^*, I_2^*, R_2^*, A_2^*) = \left(\frac{(a+g)(\nu+g)}{a\beta}, \frac{g\delta}{(a+g)(g+\delta)} - \frac{g(\nu+g)}{a\beta}, \frac{ag\delta}{(a+g)(g+\delta)(\nu+g)} - \frac{g}{\beta}, \frac{a\nu\delta}{(a+g)(g+\delta)(\nu+g)} - \frac{\nu}{\beta}, \frac{g}{g+\delta}\right).$$

If  $a\delta\beta < (a + g)(g + \delta)(\nu + g)$ , the disease-free equilibrium will be locally asymptotically stable and the endemic equilibrium will not be feasible. On the other hand, if  $a\delta\beta > (a + g)(g + \delta)(\nu + g)$ , the endemic equilibrium will be locally asymptotically stable and the disease-free equilibrium will be unstable.

*Proof.* Suppose  $(S^*, E^*, I^*, R^*, A^*)$  is an equilibrium point, then it should satisfy

$$0 = \delta A^* - \beta S^* I^* - g S^*, \quad (7.2)$$

$$0 = \beta S^* I^* - (a + g) E^*, \quad (7.3)$$

$$0 = a E^* - (\nu + g) I^*, \quad (7.4)$$

$$0 = \nu I^* - g R^*, \quad (7.5)$$

$$0 = g - (g + \delta) A^*. \quad (7.6)$$

From (7.6), we know that  $A^* = \frac{g}{g+\delta}$ . From (7.4), we have  $E^* = \frac{\nu+g}{a} I^*$ . Plugging it into (7.3) we get

$$\begin{aligned} \beta S^* I^* - (a + g) \frac{\nu + g}{a} I^* &= 0 \\ \Rightarrow (\beta S^* - \frac{(a + g)(\nu + g)}{a}) I^* &= 0 \\ \Rightarrow S^* = \frac{(a + g)(\nu + g)}{a\beta} \text{ or } I^* &= 0 \end{aligned}$$

If  $I^* = 0$ , then  $E^* = \frac{\nu+g}{a} I^* = 0$ ,  $R^* = \frac{\nu}{g} I^* = 0$ ,  $S^* = \frac{\delta}{g} A^* = \frac{\delta}{g+\delta}$ .

If  $S^* = \frac{(a+g)(\nu+g)}{a\beta}$ , then

$$\begin{aligned} I^* &= \frac{\delta A^* - gS^*}{\beta S^*} = \frac{ag\delta}{(a+g)(g+\delta)(\nu+g)} - \frac{g}{\beta}, \\ E^* &= \frac{\nu+g}{a} I^* = \frac{g\delta}{(a+g)(g+\delta)} - \frac{g(\nu+g)}{a\beta}, \\ R^* &= \frac{\nu}{g} I^* = \frac{a\nu\delta}{(a+g)(g+\delta)(\nu+g)} - \frac{\nu}{\beta}. \end{aligned}$$

Hence, there are two equilibria: the disease-free equilibrium point  $(S_1^*, E_1^*, I_1^*, R_1^*, A_1^*) = \left(\frac{\delta}{g+\delta}, 0, 0, 0, \frac{g}{g+\delta}\right)$  and the endemic equilibrium point

$$(S_2^*, E_2^*, I_2^*, R_2^*, A_2^*) = \left(\frac{(a+g)(\nu+g)}{a\beta}, \frac{g\delta}{(a+g)(g+\delta)} - \frac{g(\nu+g)}{a\beta}, \frac{ag\delta}{(a+g)(g+\delta)(\nu+g)} - \frac{g}{\beta}, \frac{a\nu\delta}{(a+g)(g+\delta)(\nu+g)} - \frac{\nu}{\beta}, \frac{g}{g+\delta}\right).$$

To determine the stability of the Equilibria, we first calculate the Jacobian matrix of the *SEIRA* model. Since  $S, E, I, R, A$  sum up to 1, there are only four free variables. To calculate the Jacobian matrix, we only need to consider any four of them. We ignore the fourth equation of system (7.1) and obtain the Jacobian matrix as

$$J(S, E, I, A) = \begin{pmatrix} -\beta I - g & 0 & -\beta S & \delta \\ \beta I & -(a+g) & \beta S & 0 \\ 0 & a & -(\nu+g) & 0 \\ 0 & 0 & 0 & -(g+\delta) \end{pmatrix}.$$

- Notice that  $I_2^*$  can be rewritten as  $I_2^* = \frac{g}{\beta(a+g)(g+\delta)(\nu+g)}(a\delta\beta - (a+g)(g+\delta)(\nu+g))$ , thus for  $a\delta\beta < (a+g)(g+\delta)(\nu+g)$ ,  $I_2^* < 0$ ,  $R_2^* = \frac{\nu}{g}I_2^* < 0$ ,

and  $E_2^* = \frac{\nu+g}{a}I_2^* < 0$ . Thus the endemic equilibrium is not feasible for  $a\delta\beta < (a+g)(g+\delta)(\nu+g)$ .

For the disease-free equilibrium, the Jacobian matrix is

$$J(S_1^*, E_1^*, I_1^*, A_1^*) = \begin{pmatrix} -g & 0 & -\frac{\delta}{g+\delta}\beta & \delta \\ 0 & -(a+g) & \frac{\delta}{g+\delta}\beta & 0 \\ 0 & a & -(\nu+g) & 0 \\ 0 & 0 & 0 & -(g+\delta) \end{pmatrix}$$

and the characteristic equation is  $(\lambda+g)(\lambda+\nu+g)(\lambda^2+(a+2g+\nu)\lambda+(a+g)(\nu+g)-\frac{a\delta}{g+\delta}\beta) = 0$ . Solving this, we obtain the following eigen values  $\lambda_1 = -g < 0$ ,  $\lambda_2 = -(g+\delta) < 0$  and

$$\lambda_{3,4} = \frac{-(a+2g+\nu) \pm \sqrt{(a+2g+\nu)^2 - 4((a+g)(\nu+g) - \frac{a\delta}{g+\delta}\beta)}}{2} = \frac{-(a+2g+\nu) \pm \sqrt{(a-\nu)^2 + 4\frac{a\delta}{g+\delta}\beta}}{2}$$

If  $(a+g)(\nu+g) - \frac{a\delta}{g+\delta}\beta > 0$ ,  $\lambda_3$  and  $\lambda_4$  will both be less than zero.

Hence the disease-free equilibrium point is asymptotically stable  $\iff$

$$(a+g)(\nu+g) - \frac{a\delta}{g+\delta}\beta > 0 \iff a\delta\beta < (a+g)(g+\delta)(\nu+g).$$

- When  $a\delta\beta > (a+g)(g+\delta)(\nu+g)$ , we have that  $S_2^* = \frac{(a+g)(\nu+g)}{a\beta} > 0$ ,  $A_2^* = \frac{g}{g+\delta} > 0$ .

Also,  $I_2^* = \frac{g}{\beta(a+g)(g+\delta)(\nu+g)}(a\delta\beta - (a+g)(g+\delta)(\nu+g)) > 0$ , and  $R_2^* = \frac{\nu}{g}I_2^* > 0$ ,  $E_2^* = \frac{\nu+g}{a}I_2^* > 0$ . Thus the endemic equilibrium is feasible when all the parameters are positive and  $a\delta\beta > (a+g)(g+\delta)(\nu+g)$ . Also, when  $a\delta\beta > (a+g)(g+\delta)(\nu+g)$ , from the previous proof, we know that the disease-free equilibrium is unstable.

For the endemic equilibrium, once again, we calculate the Jacobian matrix:

$$J(S_2^*, E_2^*, I_2^*, A_2^*) = \begin{pmatrix} -\beta I_2^* - g & 0 & -\beta S_2^* & \delta \\ \beta I_2^* & -(a + g) & \beta S_2^* & 0 \\ 0 & a & -(\nu + g) & 0 \\ 0 & 0 & 0 & -(g + \delta) \end{pmatrix}.$$

From the above matrix, we can see that one eigenvalue of the characteristic equation is  $\lambda_1 = -(g + \delta)$ , and the other three satisfy  $\lambda^3 + (-tr(M))\lambda^2 + (\alpha(M))\lambda + (-Det(M)) = 0$ ,

where  $tr(M) = -(\beta I_2^* + a + 3g + \nu)$ ,  $\alpha(M) = (\beta I_2^* + g)(a + g) + (\beta I_2^* + g)(\nu + g) + (a + g)(\nu + g) - a\beta S_2^*$ ,  $Det(M) = -(\beta I_2^* + g)(a + g)(\nu + g) + ag\beta S_2^*$ .

To study stability, we apply the third order Routh-Hurwitz stability criterion. We first review the third order Routh-Hurwitz stability criterion.

**Third-order Routh-Hurwitz stability criterion:** Real parts of all solutions of a third-order polynomial  $P(s) = a_3s^3 + a_2s^2 + a_1s + a_0 = 0$  are negative if the coefficients satisfy  $a_3 > 0, a_2 > 0, a_1 > 0, a_0 > 0$  and  $a_2a_1 > a_3a_0$ .

Thus for our characteristic equation we have that

$$a_3 = 1 > 0,$$

$$a_2 = -tr(M) = \beta I_2^* + 3g + a + \nu > 0,$$

$$\begin{aligned}
a_1 &= \alpha(M) = (\beta I_2^* + g)(a + g) + (\beta I_2^* + g)(\nu + g) + (a + g)(\nu + g) - a\beta S_2^* \\
&= (\beta I_2^* + g)(a + g) + (\beta I_2^* + g)(\nu + g) + (a + g)(\nu + g) - a\beta \frac{(a+g)(\nu+g)}{a\beta} \\
&= (\beta I_2^* + g)(a + 2g + \nu) > 0, \\
a_0 &= -\text{Det}(M) = (\beta I_2^* + g)(a + g)(\nu + g) - ag\beta S_2^* \\
&= \beta I_2^*(a + g)(\nu + g) + g(a + g)(\nu + g) - ag\beta S_2^* \\
&= \beta I_2^*(a + g)(\nu + g) + g(a + g)(\nu + g) - ag\beta \frac{(a+g)(\nu+g)}{a\beta} \\
&= \beta I_2^*(a + g)(\nu + g) > 0, \\
a_2 a_1 &= \alpha(M) * (-\text{tr}(M)) \\
&= (\beta I_2^* + g)(a + 2g + \nu)(\beta I_2^* + 3g + a + \nu) \\
&> \beta I_2^*(a + g)(\nu + g) = -\text{Det}(M) = a_3 a_0.
\end{aligned}$$

Hence the conditions of Routh-Hurwitz stability criterion are satisfied. The endemic equilibrium is thus asymptotically stable if all the parameters are positive and  $a\delta\beta > (a + g)(g + \delta)(\nu + g)$ .

□

**Conjecture 1.** *When*

$$R_0 < 1, \quad \lim_{t \rightarrow +\infty} (S(t), E(t), I(t), R(t), A(t)) \longrightarrow \text{Disease-free equilibrium (DFE)}$$

*and when*

$$R_0 > 1, \quad \lim_{t \rightarrow +\infty} (S(t), E(t), I(t), R(t), A(t)) \longrightarrow \text{Endemic equilibrium (EE)}$$

, where  $R_0 = \frac{a\delta\beta}{(a+g)(g+\delta)(\nu+g)}$  is the basic reproduction number.

Recall that when  $R_0 > 1$ , the disease can spread and when  $R_0 < 1$ , the

disease will finally disappear. We can rewrite  $R_0$  as

$$R_0 = \frac{\beta}{\nu + g} \cdot \frac{a}{a + g} \cdot \frac{\delta}{\delta + g}.$$

We know that  $\beta SI$  is the number of new cases over a unit time. Hence the average number of new cases caused by one infective individual is  $\frac{\beta SI}{I} = \beta S$ . Therefore, the number of susceptibles an infective individual can infect is the average number over a unit time multiply by the average length of duration an infective individual stay infectious which is  $\beta S \cdot \frac{1}{\nu+g}$ . The fraction of the infected individuals who can finally become infective is the probability that an exposed individual will become infective in a unit time multiply by the time duration of an exposed individual stay exposed which is  $a \cdot \frac{1}{a+g}$ . The expected fraction of susceptibles is  $\frac{\delta}{\delta+g}$  which is the value of susceptibles at the disease-free steady state. Therefore,  $\beta S \cdot \frac{1}{\nu+g} \cdot \frac{a}{a+g} = \beta \cdot \frac{\delta}{\delta+g} \cdot \frac{1}{\nu+g} \cdot \frac{a}{a+g}$  is the average number of susceptibles that one infective individual can infect.

## 7.4 Sensitivity analysis

In this section, we calculate, analyze and compare the normalized forward sensitivity indices of the outbreak peak value, time of outbreak peak, and steady state value of  $I(t)$ , to the parameters of the system by computing

$$S.I. = \frac{p}{X^*} \frac{\partial X^*}{\partial p}, \quad (7.7)$$

where  $X^*$  is the quantity being considered, and  $p$  is the parameter which  $X^*$  depends upon. Sensitivity indices can be positive or negative which indicate

Parameter	Sensitivity of peak	Description
$\delta$	0.0099	Death/Birth Rate
$\beta$	1.7200	Average transmission rate
$g$	-0.0195	Growth Rate
$\nu$	-2.8204	Removal Rate
$a$	1.1102	Rate at which exposed individuals become infective

**Table 7.2:** The sensitivity indices of the value of the outbreak peak respect to the parameters values  $\delta = 1/64/12/month$ ,  $\beta = 55/month$ ,  $g = 1/16/12/month$ ,  $\nu = 52/12/month$ ,  $a = 52/12/month$  and initial values  $S(0) = 0.2$ ,  $E(0) = 0.002$ ,  $I(0) = 0.002$ ,  $R(0) = 0.006$ ,  $A(0) = 0.79$ .

the nature of the relationship. The magnitude of S.I. indicates the strength of the relationship.

When studying quantities like the peak value or the peak time which do not have explicit formulas, we compute an approximate value of their S.I. as follows

$$S.I. = \frac{p}{X^*(p)} \frac{X^*(p + \Delta p) - X^*(p - \Delta p)}{2\Delta p}. \quad (7.8)$$

We calculate  $S.I.$  with respect to one specific parameter by perturbing this parameter only and keeping the others unchanged. Here, we take  $\Delta p = 1\%p$ .

### 7.4.1 Sensitivity analysis of the outbreak peak value

The sensitivity indices of the amplitude of the outbreak peak show how the first epidemic depends on the parameters as seen in Table 7.2.

The removal rate  $\nu$  has the strongest relationship to the magnitude of the outbreak peak. The negative value tells us that a lower removal rate would

lead to a more severe epidemic. In contrast to the birth/death rate  $\delta$  which has among the lowest of sensitivity indices,  $\nu$  would thus be an important parameter to control in order to reduce the harm of an outbreak.

Both the average transmission rate  $\beta$  and  $a$  have strong positivity relationship to the peak outbreak as a higher  $\beta$  value would lead to a higher number of people in the exposed compartment and a higher  $a$  would move more exposed to the infectives compartment.

The sensitivity index with respect to the human birth/death rate  $\mu$  is very low in comparison to all the others. This makes sense, because the initial peak of an epidemic occurs relatively quickly after the introduction of sick people, and the birth and death of new susceptibles would take much longer time.

The sensitivity of the growth rate  $g$  to the outbreak peak is negative because a larger growth rate will move people from the infectives compartment to the adult compartment faster, thus reducing the outbreak peak. Similar to birth/death rate  $\delta$ , growth rate  $g$  has a small influence on the peak value.

In fact, parameters related to demography, such as birth/death rate and growth rate, would have small influence on the outbreak level as the initial peak appears relatively quickly. Parameters which are directly related to infection would have important influence on the initial peak. For instance the following parameters:  $a$  which determines how fast exposed individuals will become infectious,  $\nu$  which determines how quickly infectives will move to the recovered compartment and  $\beta$  which determines how many susceptibles will be infected all have a strong relationship with the outbreak as expected because they are directly related to infection.

## 7.4.2 Sensitivity analysis of the outbreak peak time

Sensitivity indices of the outbreak peak time measure how the first epidemics outbreak time depends on different parameters as seen in the Table 7.3.

Parameter	Sensitivity of peak time	Description
$\delta$	-0.0011	Death/Birth Rate
$\beta$	-0.7403	Average transmission rate
$g$	0.0027	Growth Rate
$\nu$	0.3075	Removal Rate
$a$	-0.2908	Rate at which exposed individuals become infective

**Table 7.3:** The sensitivity of the outbreak peak time respect to the parameters with values  $\delta = 1/64/12/month$ ,  $\beta = 55/month$ ,  $g = 1/16/12/month$ ,  $\nu = 52/12/month$ ,  $a = 52/12/month$ . and initial values  $S(0) = 0.2$ ,  $E(0) = 0.002$ ,  $I(0) = 0.002$ ,  $R(0) = 0.006$ ,  $A(0) = 0.79$ .

As outlined previously, we have the same reason that birth/death rate  $\delta$  and growth rate  $g$  have less influence on the outbreak time than the other three parameters.

We can see from Table 7.3 that the average transmission rate  $\beta$  has the strongest influence on the dynamics of the system. This suggests that  $\beta$  is a more important parameter to control to prevent outbreaks. The negative relationship tells us that a larger average transmission rate would lead to a quicker outbreak.

The relationship between rate  $a$  and the time of the maximum outbreak is negative, because a higher contact rate (shorter latent period) will cause more

new infections and the timing of the maximum would be attained earlier.

The removal rate still has an important effect on the outbreak time. The positive relationship between  $\nu$  and the outbreak time is because patients will recover faster with a larger  $\nu$  value thereby postponing the outbreak time.

### 7.4.3 Sensitivity analysis of the endemic steady state

Endemic steady state determines the levels of the different groups of an endemic infectious disease. It represents the expectation of the final size of all the groups. In Table 7.4, we list sensitivity indices of  $I_2^*$  with respects to all the parameters.

Parameter	Sensitivity of $I_2^*$	Description
$\delta$	1.3221	Death/Birth Rate
$\beta$	0.6526	Average transmission rate
$g$	-0.3260	Growth Rate
$\nu$	-1.6508	Removal Rate
$a$	0.0020	Rate at which exposed individuals become infective

**Table 7.4:** Sensitivity of the endemic steady state with respect to the parameters with values  $\delta = 1/64/12/month$ ,  $\beta = 55/month$ ,  $g = 1/16/12/month$ ,  $\nu = 52/12/month$ ,  $a = 52/12/month$ .

The endemic level of infective individuals is most sensitive to the recovery rate  $\nu$  and birth/death rate  $\delta$ . It has a strong negative relationship with  $\nu$  because recovering is the main way that infectives leave the infected compartment. The relationship with  $\delta$  is positive as a larger  $\delta$  means more susceptible

newborns will possibly become infectives. Rate  $a$  has a weak but positive relationship with  $I_2^*$  as expected because a larger  $a$  causes more people in the exposed compartment to become infectives. The average transmission rate  $\beta$  is also of great importance in controlling the endemic level of the infectives. The positive relationship is obvious since larger  $\beta$  means more susceptibles will get infected.

## 7.5 Extracting the time-dependent transmission rate $\beta(t)$ from prevalence pre vaccination data

In this section, we assume that  $\beta(t)$  is a function of time, compute the formula for  $\beta(t)$  based on prevalence pre-vaccination data and then use it to construct an algorithm.

**Theorem 9.** *Suppose the epidemic is observed over the time interval  $[0, T]$ , where  $t = 0$  and  $t = T$  are respectively the start and end of observations, then the time dependent transmission function  $\beta(t)$  for System (7.1) satisfy  $M\beta''\beta^2 + N(\beta')^2\beta + P\beta'\beta^2 - L\beta^4 - Q\beta^3 = 0$  where  $f(t)$  is a smooth positive function which matches the infection data in the interval  $[0, T]$ ,*

$$\begin{aligned}
H(t) &\triangleq f''(t) + (\nu + 2g + a)f'(t) + (a + g)(\nu + g)f(t) \\
M = D &= -Hf^3, \\
N = C &= 2Hf^3, \\
P &= B - (2g + \delta)Hf^3 = 2Hf'f^2 - 2H'f^3 - (2g + \delta)Hf^3, \\
Q &= -(A + (2g + \delta)H'f^3 - (2g + \delta)Hf'f^2 + g(g + \delta)Hf^3) \\
&= -(H''f^3 - Hf''f^2 - 2H'f'f^2 + 2H(f')^2f + (2g + \delta)H'f^3 - (2g + \delta)Hf'f^2 + g(g + \delta)Hf^3), \\
L &= -(H'f^4 + (g + \delta)Hf^4 - ag\delta f^4).
\end{aligned}$$

*Proof of Theorem 9.* Rewriting the third equation of System (7.1) as  $f'(t) + (\nu + g)f(t) = aE(t)$  and differentiating both sides we get:  $f''(t) + (\nu + g)f'(t) = aE'(t)$ . Plugging the second and third equations of (7.1) into the above equation, we have

$$\begin{aligned}
f''(t) + (\nu + g)f'(t) &= aE'(t) \\
&= a\beta(t)S(t)f(t) - (a + g)(aE(t)) \\
&= a\beta(t)S(t)f(t) - (a + g)(f'(t) + (\nu + g)f(t).)
\end{aligned}$$

Rewriting the above equation as

$$aS(t) = \frac{f''(t) + (\nu + 2g + a)f'(t) + (a + g)(\nu + g)f(t)}{\beta(t)f(t)}.$$

$$\text{i.e } aS(t) = \frac{H(t)}{\beta(t)f(t)}.$$

Taking the first and second derivatives of the above equation, we get

$$\begin{aligned}
aS'(t) &= \left( \frac{H(t)}{\beta(t)f(t)} \right)' = \frac{H'(t)\beta(t)f(t) - H(t)(\beta'(t)f(t) + \beta(t)f'(t))}{\beta(t)^2 f(t)^2}, \\
aS''(t) &= \left( \frac{H'(t)\beta(t)f(t) - H(t)(\beta'(t)f(t) + \beta(t)f'(t))}{\beta(t)^2 f(t)^2} \right)' \\
&= \frac{[H'\beta f - H(\beta'f + \beta f')]'\beta^2 f^2 - [H'\beta f - H(\beta'f + \beta f')](\beta^2 f^2)'}{(\beta^2 f^2)^2} \\
&= \frac{H''f^3\beta^3 + H'f^3\beta'\beta^2 + H'f'f^2\beta^3 - H'f^3\beta'\beta^2 - H'f'f^2\beta^3 - Hf^3\beta''\beta^2 - 2Hf'f^2\beta'\beta^2 - Hf''f^2\beta^3}{\beta^4 f^4} \\
&\quad - \frac{2(H'f^3\beta'\beta^2 + H'f'f^2\beta^3 - Hf^3(\beta')^2\beta - Hf'f^2\beta'\beta^2 - Hf'f^2\beta'\beta^2 - H(f')^2 f\beta^3)}{\beta^4 f^4} \\
&= \frac{[H''f^3 - Hf''f^2 - 2H'f'f^2 + 2H(f')^2 f]\beta^3 + [2Hf'f^2 - 2H'f^3]\beta'\beta^2 + 2Hf^3(\beta')^2\beta - Hf^3\beta''\beta^2}{\beta^4 f^4} \\
&= \frac{A\beta^3 + B\beta'\beta^2 + C(\beta')^2\beta + D\beta''\beta^2}{\beta^4 f^4},
\end{aligned}$$

where  $A = H''f^3 - Hf''f^2 - 2H'f'f^2 + 2H(f')^2 f$ ,  $B = 2Hf'f^2 - 2H'f^3$ ,  $C = 2Hf^3$ ,  $D = -Hf^3$ .

Since  $S + E + I + R + A = 1$ , i.e.  $S + E + I + R = 1 - A$ , the fifth equation of (7.1) becomes  $A' = g(S + E + I + R) - \delta A = g(1 - A) - \delta A = g - (g + \delta)A$ .

Taking the derivative of the first equation of (7.1) with respect to  $t$  we have:  $S'' + (\beta SI)' + gS' = \delta A'$ , plugging in the first and fifth equation of (7.1) we get

$$S'' + (\beta SI)' + gS' = \delta A' = \delta(g - (g + \delta)A) = g\delta - (g + \delta)(\delta A) = g\delta - (g + \delta)(S' + \beta SI + gS).$$

Multiplying both sides of the above equation by  $a$  and simplifying gives  $aS'' + (2g + \delta)(aS') + g(g + \delta)(aS) + (a\beta SI)' + (g + \delta)(a\beta SI) = ag\delta$ .

Substituting for  $aS$ ,  $aS'$  and  $aS''$  we have

$$\begin{aligned} & \left( \frac{A\beta^3 + B\beta'\beta^2 + C(\beta')^2\beta + D\beta''\beta^2}{\beta^4 f^4} \right) \\ & + (2g + \delta) \left( \frac{H'(t)\beta(t)f(t) - H(t)(\beta'(t)f(t) + \beta(t)f'(t))}{\beta(t)^2 f(t)^2} \right) \\ & + g(g + \delta) \left( \frac{H(t)}{\beta(t)f(t)} \right) + H' + (g + \delta)H = ag\delta. \end{aligned}$$

Multiplying both sides of the equation by  $\beta^4 f^4$  and expanding we get

$$\begin{aligned} & A\beta^3 + B\beta'\beta^2 + C(\beta')^2\beta + D\beta''\beta^2 + (2g + \delta)H'f^3\beta^3 - (2g + \delta)Hf^3\beta'\beta^2 \\ & (2g + \delta)Hf'f^2\beta^3 + g(g + \delta)Hf^3\beta^3 + H'f^4\beta^4 + (g + \delta)Hf^4\beta^4 = ag\delta\beta^4. \end{aligned}$$

This implies

$$\begin{aligned} & (H'f^4 + (g + \delta)Hf^4 - ag\delta f^4)\beta^4 + (A + (2g + \delta)H'f^3 - (2g + \delta)Hf'f^2 \\ & + g(g + \delta)Hf^3)\beta^3 + (B - (2g + \delta)Hf^3)\beta'\beta^2 + C(\beta')^2\beta + D\beta''\beta^2 = 0. \end{aligned}$$

i.e.

$$-L\beta^4 - Q\beta^3 + P\beta'\beta^2 + N(\beta')^2\beta + M\beta''\beta^2 = 0, \quad (7.9)$$

where

$$\begin{aligned}
M &= D = -Hf^3, \\
N &= C = 2Hf^3, \\
P &= B - (2g + \delta)Hf^3 = 2Hf'f^2 - 2H'f^3 - (2g + \delta)Hf^3, \\
Q &= -(A + (2g + \delta)H'f^3 - (2g + \delta)Hf'f^2 + g(g + \delta)Hf^3) \\
&= -(H''f^3 - Hf''f^2 - 2H'f'f^2 + 2H(f')^2f + (2g + \delta)H'f^3 - (2g + \delta)Hf'f^2 + g(g + \delta)Hf^3), \\
L &= -(H'f^4 + (g + \delta)Hf^4 - ag\delta f^4).
\end{aligned}$$

Notice that  $N = -2M$ , equation(7.9) becomes

$$M\beta''\beta^2 - 2M(\beta')^2\beta + P\beta'\beta^2 - Q\beta^3 - L\beta^4 = 0.$$

Dividing by  $\beta^4$  we get

$$\begin{aligned}
&M\frac{\beta''\beta^2 - 2(\beta')^2\beta}{\beta^4} + P\frac{\beta'\beta^2}{\beta^4} - Q\frac{\beta^3}{\beta^4} - L\frac{\beta^4}{\beta^4} \\
&= -M(2\beta^{-3}(\beta')^2 - \beta^{-2}\beta'') - P(-\beta^{-2}\beta') - Q(\beta^{-1}) - L \\
&= 0.
\end{aligned}$$

Let  $y = \beta^{-1}$ ,  $y' = -\beta^{-2}\beta'$ ,  $y'' = 2\beta^{-3}(\beta')^2 - \beta^{-2}\beta''$ .

(7.9) can be rewritten as  $My'' + Py' + Qy + L = 0$  with  $y = \beta^{-1}$ .  $\square$

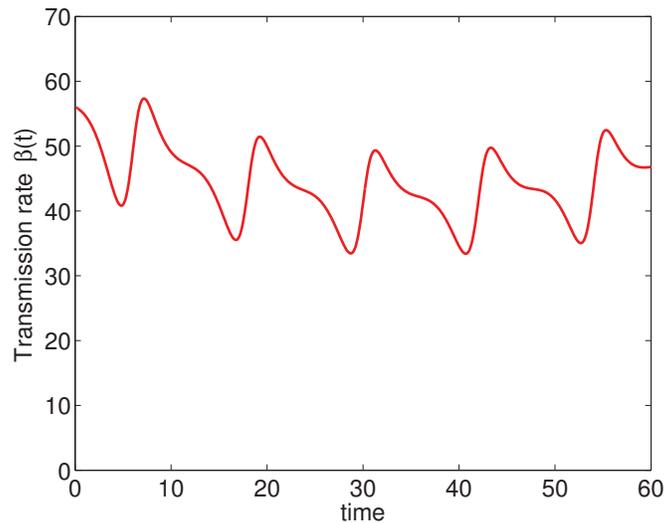
Based on the formula for  $\beta(t)$  derived above, we now construct a prevalence algorithm for extracting the transmission rate from the *SEIRA* model.

**Step 1** Smoothly interpolate the infection data with a spline or trigonometric function to generate a smooth function,  $f(t)$ .

**Step 2** Calculate the function  $H(t) = f''(t) + (\nu + 2g + a)f'(t) + (a + g)(\nu + g)f(t)$ . Compute M, N, P, Q, and L by plugging  $H(t)$  into (7.5).

**Step 3** Choose  $\beta(0)$ ,  $\beta'(0)$ , and interval  $T$ , use an ODE solver to solve equation  $M\beta''\beta^2 + N(\beta')^2\beta + P\beta'\beta^2 - L\beta^4 - Q\beta^3 = 0$  for  $\beta(t)$  on interval the  $[0, T]$ .

In the absence of real prevalence data, we use simulated data to test this algorithm. To this end, consider the following function  $f(t) = 10^{-3}[1.4 + \cos(2\pi t/12)]$ , that approximates fractions of infectives from a typical infectious disease with periodic outbreaks. Figure 7.1 contains the dynamics of  $\beta(t)$  for this data set, using the algorithm above.



**Figure 7.1:**  $\beta(t)$  extracted from fake prevalence data  $f(t) = 10^{-3}[1.4 + \cos(2\pi t/12)]$  with initial value  $\beta(0) = 56$ ,  $\beta'(0) = -1$ , and parameters  $\nu = 52/12$ ,  $a = 52/12$ ,  $g = 1/16/12$ ,  $\delta = 1/64/12$ .

## 7.6 Extracting the time dependent transmission rate from pre-vaccination incidence data

In this section, we compute the formula for  $\beta(t)$  that depends on incidence data and use it to construct the incidence algorithm. As in the previous section, we assume that the transmission rate depends on time.

### 7.6.1 Solution of the inverse problem for the SEIRA model

To construct the algorithm, we first rewrite  $\beta(t)$  in terms of  $\omega(t)$ . With  $\omega(t) = \beta SI$  and  $S + E + I + R + A = 1$ , the *SEIRA* model can be re-written as

$$\frac{dS(t)}{dt} = (\delta A(t) - \omega(t)) - gS(t) \quad (7.10)$$

$$\frac{dE(t)}{dt} = \omega(t) - (a + g)E(t), \quad (7.11)$$

$$\frac{dI(t)}{dt} = aE(t) - (\nu + g)I(t), \quad (7.12)$$

$$\frac{dR(t)}{dt} = \nu I(t) - gR(t), \quad (7.13)$$

$$\frac{dA(t)}{dt} = g - (g + \delta)A(t). \quad (7.14)$$

We have the following theorem:

**Theorem 10.** *For the SEIRA, given a continuous function  $w(t)$  generated from the incidence data,  $\beta(t)$  can be estimated by  $\frac{\omega(t)}{S(t)I(t)}$  with  $S(t)$  and  $I(t)$*

given by (7.15) and (7.16) respectively.

$$S(t) = S(0)e^{-gt} + \int_0^t (\delta(A(0)e^{-(g+\delta)s} + \int_0^s ge^{(g+\delta)(\sigma-s)}d\sigma - \omega(s))e^{g(s-t)}ds \quad (7.15)$$

$$I(t) = I(0)e^{-(\nu+g)t} + \int_0^t a(E(0)e^{-(a+g)s} + \int_0^s \omega(\sigma)e^{(a+g)(\sigma-s)}d\sigma)e^{(\nu+g)(s-t)}ds \quad (7.16)$$

*Proof.* Solving the equations in the system using the method of variation of parameters, we obtain

$$A(t) = C(t)e^{-(g+\delta)t} = (A(0) + \int_0^t ge^{(g+\delta)s}ds)e^{-(g+\delta)t}$$

$$\Rightarrow A(t) = A(0)e^{-(g+\delta)t} + \int_0^t ge^{(g+\delta)(s-t)}ds \quad (7.17)$$

$$S(t) = S(0)e^{-gt} + \int_0^t (\delta A(s) - \omega(s))e^{g(s-t)}ds \quad (7.18)$$

$$I(t) = I(0)e^{-(\nu+g)t} + \int_0^t aE(s)e^{(\nu+g)(s-t)}ds \quad (7.19)$$

$$E(t) = E(0)e^{-(a+g)t} + \int_0^t \omega(s)e^{(a+g)(s-t)}ds \quad (7.20)$$

Plug (7.20) into (7.19) and plug (7.17) into (7.18):

$$S(t) = S(0)e^{-gt} + \int_0^t (\delta(A(0)e^{-(g+\delta)s} + \int_0^s ge^{(g+\delta)(\sigma-s)}d\sigma - \omega(s))e^{g(s-t)}ds \quad (7.21)$$

$$I(t) = I(0)e^{-(\nu+g)t} + \int_0^t a(E(0)e^{-(a+g)s} + \int_0^s \omega(\sigma)e^{(a+g)(\sigma-s)}d\sigma)e^{(\nu+g)(s-t)}ds \quad (7.22)$$

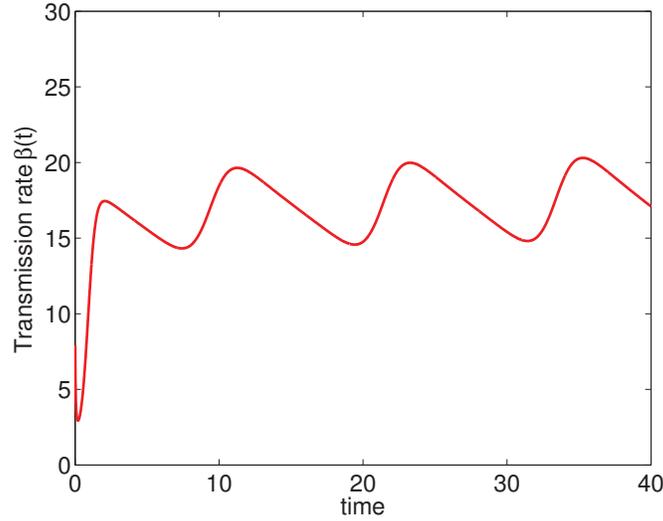
Thus  $\beta(t) = \frac{\omega(t)}{S(t)I(t)}$  with  $S(t)$  and  $I(t)$  given in (7.15) and (7.16).  $\square$

We now turn the above theorem into an algorithm to extract time-dependent transmission rate  $\beta(t)$  numerically, using incidence data:

**Step 1** Smoothly interpolate incidence data with a spline or trigonometric function to generate a smooth  $\omega(t)$  ( In fact, we only need  $\omega(t)$  to be continuous, not necessarily smooth).

**Step 2** Let  $T$  be the whole period of data. Compute  $\beta(t) = \frac{\omega(t)}{S(t)I(t)}$ , for  $t \in [0, T]$ .

Because we have real incidence data, we test the latter algorithm using both fake and real data. Firstly, we test the performance of the incidence algorithm using fake data. As before, we use the function  $f(t) = 10^{-3}[1.4 + \cos(2\pi t/12)]$  to generate the data. Figure 7.2 shows  $\beta(t)$  plotted against time using simulated data. Our algorithm estimates the transmission rate perfectly well, using this data.

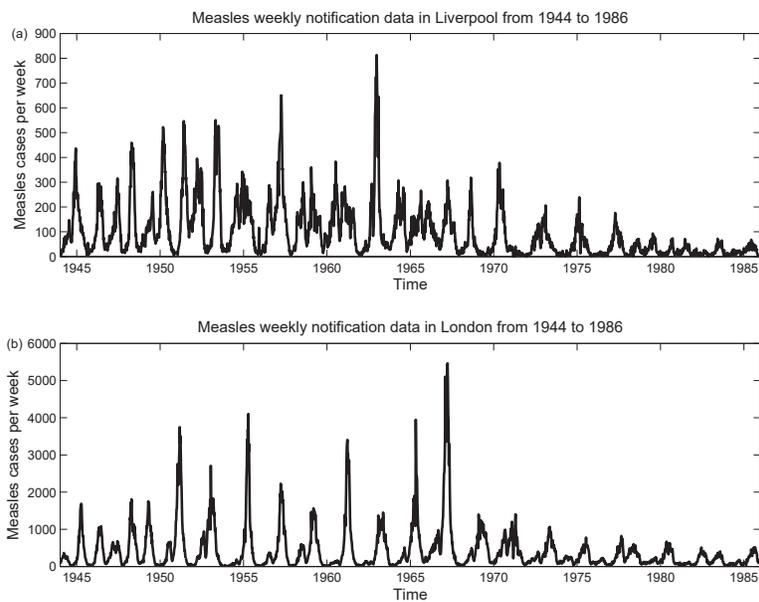


**Figure 7.2:**  $\beta(t)$  extracted from fake incidence data  $f(t) = 10^{-4}[2.7 + 1.5\sin(2\pi t/12)]$  with initial value  $S(0) = 0.25, E(0) = 0.0009, I(0) = 0.0001, A(0) = 0.7$ , and parameters  $\nu = 52/12, a = 52/12, g = 1/16/12, \delta = 1/64/12$ . (a)  $\beta(t)$  from time 0 to 40.(b)  $\beta(t)$  from time 0 to 144.

Secondly, we use real incidence data from Liverpool and London given in Figure 7.3 to test the efficiency of our prevalence algorithm.

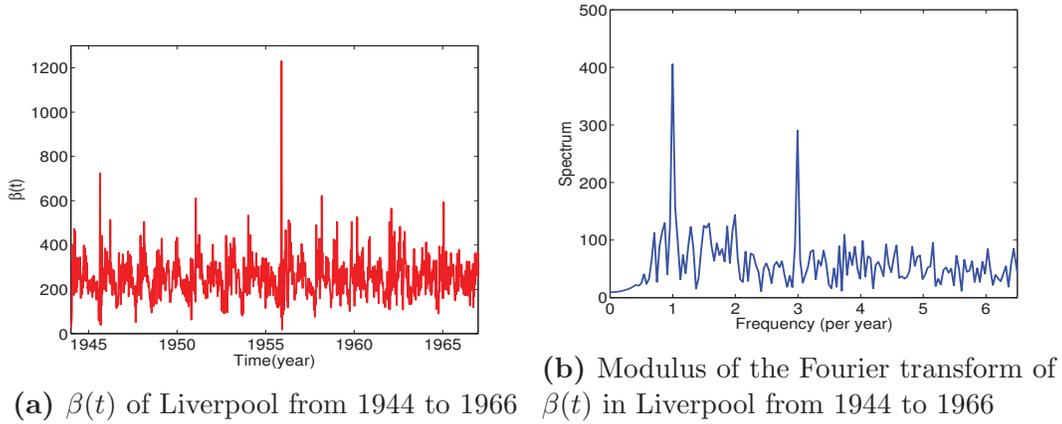
Figure 7.4a shows  $\beta(t)$  extracted from pre-vaccination measles weekly notification data of Liverpool from 1944 to 1966 using the prevalence algorithm. There are more noises with the weekly data. This is because Liverpool is a relatively smaller city compare to London. The population of Liverpool is less than 1/10 that of London. Figure 7.4b plots modulus of Fourier transform of  $\beta(t)$  in Liverpool. We observe two dominant peaks with frequencies 1/year and 3/year.

Figure 7.5a illustrates  $\beta(t)$  extracted from post-vaccination measles weekly notification data of London from 1944 to 1966 by the prevalence algorithm. Figure 7.5b plots modulus of Fourier transform of  $\beta(t)$  in this city. As with

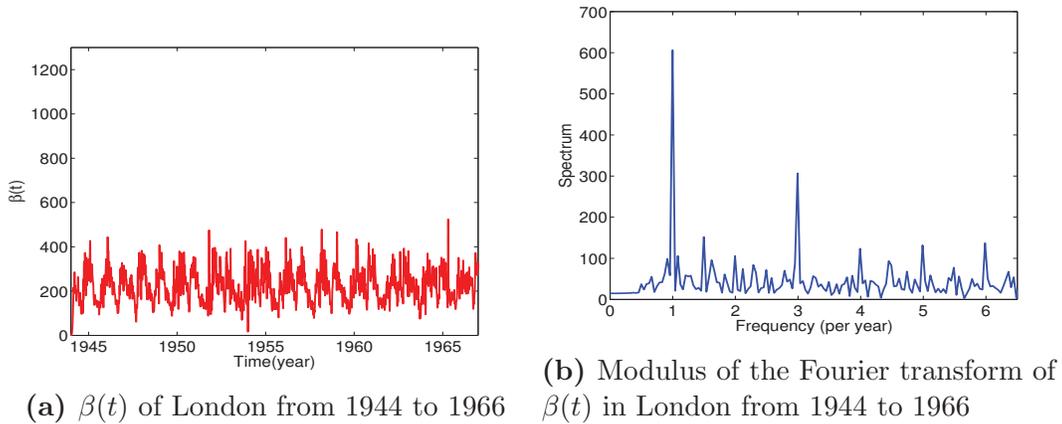


**Figure 7.3:** Measles weekly notification data in Liverpool and London from 1944-86.

that in Liverpool, two dominant peaks with frequencies 1/year and 3/year are observed.



**Figure 7.4:** Time-dependent transmission rate  $\beta(t)$  of Liverpool from year 1944 to 1966 and the modulus of its Fourier transform. The parameters are  $\delta = 1/64/52/week$ ,  $a = 52/52/week$ ,  $\nu = 52/52/week$ ,  $g = 1/16/52/week$  and initial values  $S(0) = 0.2$ ,  $E(0) = 0.001$ ,  $I(0) = 0.001$ ,  $A(0) = 0.78$ .



**Figure 7.5:** Time-dependent transmission rate  $\beta(t)$  of London from year 1944 to 1966 and the modulus of its Fourier transform. The parameters used are  $\delta = 1/64/52/week$ ,  $a = 52/52/week$ ,  $\nu = 52/52/week$ ,  $g = 1/16/52/week$  and initial values  $S(0) = 0.2$ ,  $E(0) = 0.001$ ,  $I(0) = 0.001$ ,  $A(0) = 0.78$ .

The 1/year peak is consistent with common belief that measles is driven by seasonal factors such as environmental changes and immune system changes

and the 3/year peak indicates the superiority of school contacts in driving measles transmission over other seasonal factors and thus support authors in [146, 149] that ignore other seasonal factors in determining the transmission rate of measles. The 1/year frequency peak value is about twice the value of the 3/year frequency for the Fourier transform of the transmission rate in London, whereas there is no great difference in the peak values of the dominant frequencies for that in Liverpool. This might be linked to the geographical location of the two cities. Also, the dominant frequencies are robust with respect to initial values. Moreover, the dominant frequencies in London have less noise than those in Liverpool because the population of London is greater than that of Liverpool and it is less sensitive to unexpected factors.

## 7.7 The SEIRA model with vaccination

In this section, we investigate the effect of vaccination on the SEIRA epidemic model. Vaccination is the process by which a vaccine stimulates the immune system of an individual to build immunity against a pathogen. Vaccination can ameliorate both mortality and morbidity. The effectiveness of vaccination has been widely studied and verified since the first work of Edward Jenner on smallpox [151].

Different vaccination strategies are used to deal with different situations. Pediatric vaccination is an efficient way in preventing dangerous human infectious diseases. Much work has focused on the vaccination of newborn babies or infants to reduce the prevalence of diseases like measles, mumps, and rubella. Mathematical treatment of vaccination is straight forward and only needs a

single addition to the *SEIRA* model. Using  $p$  to denote the fraction of the newborns that are successfully vaccinated, we obtain the following model:

$$\begin{aligned}
\frac{dS(t)}{dt} &= \delta(1-p)A(t) - \beta(t)S(t)I(t) - gS(t), \\
\frac{dE(t)}{dt} &= \beta(t)S(t)I(t) - aE(t) - gE(t), \\
\frac{dI(t)}{dt} &= aE(t) - \nu I(t) - gI(t), \\
\frac{dR(t)}{dt} &= \nu I(t) - gR(t) + \delta p A(t), \\
\frac{dA(t)}{dt} &= g(S(t) + E(t) + I(t) + R(t)) - \delta A(t).
\end{aligned} \tag{7.23}$$

However, it is not cost-effective to control rare infectious diseases by pediatric vaccination. To this end, another vaccination policy, random vaccination are conducted for rare infectious diseases or any potential outbreak. With this policy, all unvaccinated susceptibles and not just newborns are vaccinated.

It is difficult for a disease to spread as long as the fraction of susceptibles is kept low. Therefore, it is more reasonable that we should vaccinate less if the fraction of susceptibles is lower, and vice-versa.

In sections 7.8 and 7.9 we assume that  $\beta$  is a constant and in Sections 7.10 and 7.11, it is considered to be a function of time.

## 7.8 Qualitative analysis

In this section, we list qualitative results such as positivity, boundedness, equilibria and their stability of system (7.23). The proofs of the theorems in this

section are presented in Appendix.

**Theorem 11.** *The compact set  $\Gamma = \{(S, E, I, R, A): S \geq 0, E \geq 0, I \geq 0, R \geq 0, A \geq 0, S + E + I + R + A = 1\}$  is positively invariant for the semiflow generated by system (7.23).*

*Proof of Theorem 11.* Using the same symbols as in the qualitative analysis section before, we have that  $f_1 = \delta(1 - p)y_5 \geq 0$ ,  $f_2 = \beta y_1 y_3 \geq 0$ ,  $f_3 = ay_2 \geq 0$ ,  $f_4 = \nu y_3 + \delta p y_5 \geq 0$ ,  $f_5 = g(y_1 + y_2 + y_3 + y_4) \geq 0$ . Therefore the solution to system (7.23) is positive for any positive initial values. Also, boundedness property is the same as with the SEIRA model without vaccination analyze above. Combining positivity and boundedness, we conclude that the solution of (7.23) will stay in  $\{(S, E, I, R, A): S \geq 0, E \geq 0, I \geq 0, R \geq 0, A \geq 0, S + E + I + R + A = 1\}$  for any positive initial values.  $\square$

**Theorem 12.** *System (7.23) has two equilibria: the disease-free equilibrium  $(S_1^*, E_1^*, I_1^*, R_1^*, A_1^*) = (\frac{\delta(1-p)}{g+\delta}, 0, 0, \frac{\delta p}{g+\delta}, \frac{g}{g+\delta})$  and the endemic equilibrium*

$$(S_2^*, E_2^*, I_2^*, R_2^*, A_2^*) = \left( \frac{(a+g)(\nu+g)}{a\beta}, \frac{g\delta(1-p)}{(a+g)(g+\delta)} - \frac{g(\nu+g)}{a\beta}, \frac{ag\delta(1-p)}{(a+g)(g+\delta)(\nu+g)} - \frac{g}{\beta}, \frac{a\nu\delta(1-p)}{(a+g)(g+\delta)(\nu+g)} - \frac{\nu}{\beta} + \frac{\delta p}{g+\delta}, \frac{g}{g+\delta} \right)$$

*When  $a\delta(1-p)\beta < (a+g)(g+\delta)(\nu+g)$ , the disease-free equilibrium is locally asymptotically stable and the endemic equilibrium is not feasible, and when  $a\delta(1-p)\beta > (a+g)(g+\delta)(\nu+g)$ , the endemic equilibrium is locally asymptotically stable and the disease-free equilibrium is unstable.*

*Proof.* Same as the proof of Theorem 8.  $\square$

The basic reproduction number is

$$R_0 = \frac{a\beta\delta(1-p)}{(a+g)(g+\delta)(\nu+g)}.$$

It can be rewritten as

$$R_0 = \frac{\beta}{\nu+g} \cdot \frac{a}{a+g} \cdot \frac{\delta}{\delta+g} \cdot (1-p).$$

The first three terms have the same meanings as before. When considering vaccination, the level of susceptibles will be ‘discounted’ by  $p \times 100\%$  because vaccinated newborns are not susceptible.

## 7.9 Sensitivity analysis

In this section, we focus only on the sensitivity of the outbreak peak value, time of outbreak peak, and steady state value of  $I(t)$ , to the Vaccinated fraction  $p$ , since we have discussed the sensitive indices of these quantities to the other parameters of the model in one of the previous sections.

### 7.9.1 Sensitivity analysis of the outbreak peak value

Table 7.5 shows the sensitivity of the outbreak peak value to all the parameters of our model. Comparing sensitivity analysis of this model with that of the previous model, we can see that the absolute value of sensitivity indices respect to all parameters are smaller, but vaccination does not change ranks of their importance. This is because, it is applied only to newborn ba-

Parameter	Sensitivity of peak	
$p$	-0.0137	Vaccinated fraction
$\delta$	0.0137	Human birth Rate
$\beta$	0.8629	Transmission Rate
$g$	-0.0152	Growth Rate
$\nu$	-1.3635	Removal Rate
$a$	0.5124	Rate at which exposed individuals become infective

**Table 7.5:** Sensitivity of the value of the outbreak peak to the parameters with the parameter values  $p = 0.5, \delta = 1/64/12/month, \beta = 150/month, g = 1/16/12/month, \nu = 52/12/month, a = 52/12/month$  and initial values  $S(0) = 0.0998, E(0) = 0.0001, I(0) = 0.0001, R(0) = 0.11, A(0) = 0.79$ .

Parameter	Sensitivity of the peak time	
$p$	0.0056	Vaccinated fraction
$\delta$	-0.0056	Human birth Rate
$\beta$	-1.1359	Transmission Rate
$g$	0.0071	Growth Rate
$\nu$	0.5821	Removal Rate
$a$	-0.4484	Rate at which exposed individuals become infective

**Table 7.6:** Sensitivity of the outbreak peak time to the parameters with parameter values  $p = 0.5, \delta = 1/64/12/month, \beta = 150/month, g = 1/16/12/month, \nu = 52/12/month, a = 52/12/month$  and initial values  $S(0) = 0.0998, E(0) = 0.0001, I(0) = 0.0001, R(0) = 0.11, A(0) = 0.79$ .

bies, and newborns constitute only a very small proportion of the total child population. The Vaccinated fraction  $p$  has a negative relationship with the outbreak peak since more pediatric vaccination will result in less infectives. subsectionSensitivity analysis of the outbreak peak time

From Table 7.6, we can see that the Vaccinated fraction  $p$  is the least important to control in preventing outbreaks. This is because, the vaccination here is pediatric vaccination and there are fewer newborns compare to all

Parameter	Sensitivity of $I_2^*$	Description
$p$	-1.4076	Vaccinated fraction
$\delta$	1.1261	Death/Birth Rate
$\beta$	0.4077	Transmission Rate
$g$	-0.1295	Growth Rate
$\nu$	-1.4061	Removal Rate
$a$	0.0017	Rate at which exposed individuals become infective

**Table 7.7:** Sensitivity of the endemic steady state to the parameters with parameter values  $p = 0.5$ ,  $\delta = 1/64/12/month$ ,  $\beta = 150/month$ ,  $g = 1/16/12/month$ ,  $\nu = 52/12/month$ ,  $a = 52/12/month$ .

children.

### 7.9.2 Sensitivity analysis of the endemic steady state

Table 7.7 shows that the Vaccinated fraction  $p$  has the greatest importance in determining the endemic level of infectives. Long-term vaccination to newborn babies will give immunity to most kids after many years, thus the Vaccinated fraction  $p$  is critical in endemics. The negative relationship is because more vaccination will reduce the fraction of susceptibles. Infectives will be less with less susceptibles.

## 7.10 Extracting the time dependent transmission rate $\beta(t)$ from prevalence post vaccination data

In this section, we derive the formula for  $\beta(t)$  from the SEIRA model with vaccination based on prevalence data that can be used to construct an algorithm for extracting transmission rate from post-vaccination prevalence data as with  $\beta(t)$  that we constructed for pre-vaccination data. We will omit the algorithm here as the steps of the algorithm are the same as when extracting  $\beta(t)$  from prevalence pre-vaccination data.

**Theorem 13.** *Suppose the epidemic is observed over the time interval  $[0, T]$ , where  $t = 0$  and  $t = T$  are respectively the start and end of observations, then the time dependent transmission function  $\beta(t)$  for System (7.23) satisfying (7.24),*

$$My'' + Py''Qy + L = 0 \tag{7.24}$$

where  $f(t)$  is a smooth positive function which matches the infection data in the interval  $[0, T]$ ,

$$y = \beta^{-1}$$

$$H = f''(t) + (a + 2g + \nu)f'(t) + (a + g)(\nu + g)f(t),$$

$$M = -(1 - p)Hf^3,$$

$$N = 2(1 - p)Hf^3,$$

$$P = (1 - p)(2Hf'f^2 - 2H'f^3 - (2g + \delta)Hf^3) - p'Hf^3,$$

$$Q = -(1 - p)(H''f^3 - Hf''f^2 - 2H'f'f^2 + 2H(f')^2f + (2g + \delta)H'f^3 - (2g + \delta)Hf'f^2 \\ + g(g + \delta)Hf^3) - p'H'f^3 + p'Hf'f^2 - gp'Hf^3,$$

$$L = -(1 - p)(H'f^4 + (g + \delta)Hf^4 - ag\delta(1 - p)f^4) - p'Hf^4.$$

The formula for the coefficients seem quiet different from the previous ones. In fact, the formula presented above combines the situation when  $0 \leq p < 1$  and when  $p = 1$ . From the first equation of (7.23) we can see that when  $p = 1$ ,  $\frac{dS}{dt}$  is independent of  $A(t)$ . Thus we have the same behaviour as with an *SEIR* model. We present formulae for both cases:  $p = 1$  and  $0 \leq p < 1$ .

- when  $p(t)=1$ , we have

$$P\beta' - L\beta^2 - Q\beta = 0$$

This is a Bernoulli equation. By letting  $y(t) = \frac{1}{\beta(t)}$ , the Bernoulli equation can be rewritten as a first order linear differential equation

$P y'(t) + Q y(t) + L = 0$  with

$$\begin{aligned} P &= -Hf, \\ Q &= -(H'f - Hf' + gHf), \\ L &= -Hf^2. \end{aligned}$$

- when  $0 \leq p(t) < 1$ , we have  $M\beta''\beta^2 + N(\beta')^2\beta + P\beta'\beta^2 - L\beta^4 - Q\beta^3 = 0$  with

$$\begin{aligned} M &= -Hf^3, \\ N &= 2Hf^3, \\ P &= 2Hf'f^2 - 2H'f^3 - \left(\frac{p'}{1-p} + 2g + \delta\right)Hf^3, \\ Q &= -(H''f^3 - Hf''f^2 - 2H'f'f^2 + 2H(f')^2f + \left(\frac{p'}{1-p} + 2g + \delta\right)H'f^3 \\ &\quad - \left(\frac{p'}{1-p} + 2g + \delta\right)Hf'f^2 + g\left(\frac{p'}{1-p} + g + \delta\right)Hf^3), \\ L &= -(H'f^4 + \left(\frac{p'}{1-p} + g + \delta\right)Hf^4 - ag\delta(1-p)f^4). \end{aligned}$$

We omit the proof of this theorem since it is similar to that of Theorem 3. Same as before, the lack of post-vaccination prevalence data prevented us from testing the algorithm with real data. But as can be seen from the experiment with fake data demonstrated before, the prevalence algorithm works well.

## 7.11 Extracting the time dependent transmission rate $\beta(t)$ from incidence post vaccination data

Here, we derive the formula for  $\beta(t)$  from the SEIRA model with vaccination based on incidence data that can be used to construct an algorithm for extracting transmission rate from post-vaccination incidence data as with  $\beta(t)$  that we constructed for pre-vaccination data. As above, we will omit the algorithm for extracting  $\beta(t)$  from post-vaccination incidence data as the steps of the algorithm are the same as when extracting  $\beta(t)$  from incidence pre-vaccination data.

**Theorem 14.** *For the vaccinated SEIRA model with time-dependent vaccinated fraction  $p$ , the time-dependent transmission rate is  $\beta(t) = \frac{\omega(t)}{S(t)I(t)}$  where  $S(t)$  and  $I(t)$  are*

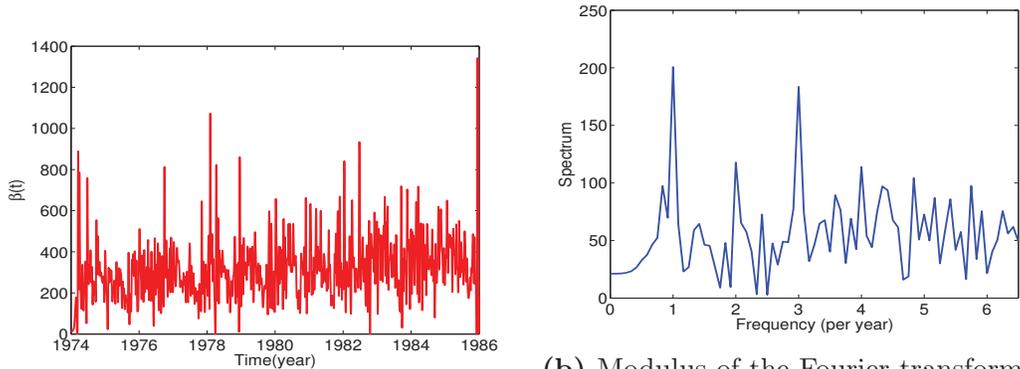
$$S(t) = S(0)e^{-gt} + \int_0^t (\delta(1-p(s))(A(0)e^{-(g+\delta)s} + \int_0^s g e^{(g+\delta)(\sigma-s)} d\sigma) - \omega(s)) e^{g(s-t)} ds'$$

$$I(t) = I(0)e^{-(\nu+g)t} + \int_0^t a(E(0)e^{-(a+g)s} + \int_0^s \omega(\sigma) e^{(a+g)(\sigma-s)} d\sigma) e^{(\nu+g)(s-t)} ds.$$

The only difference between this Theorem and Theorem 4, is the present of  $1-p(t)$  in the equation for  $I(t)$  above. We omit the proof of this theorem, as it is almost identical to the proof of Theorem 4.

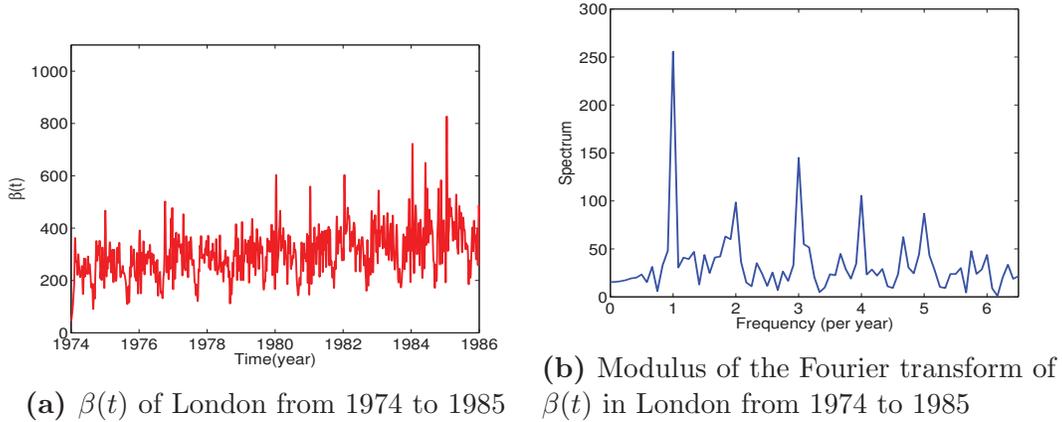
Since we have post-vaccination measles weekly notification data of Liverpool and London from 1974 to 1986 (see Figure 7.3), we use it to illustrate the efficiency of this algorithm.

Figure 7.6a presents  $\beta(t)$  extracted from post-vaccination measles weekly notification data of Liverpool from year 1974 to 1986. Figure 7.6b plots the modulus of Fourier transform of  $\beta(t)$  in Liverpool. Same as with pre-vaccination incidence algorithm, we observe dominant peaks of 1/year and 3/year periods. Figure 7.7a illustrates  $\beta(t)$  extracted from post-vaccination measles weekly notification data of London from year 1974 to 1986 by the incidence algorithm. Figure 7.7b shows the modulus of Fourier transform of  $\beta(t)$  in London. The modulus of the Fourier transform of  $\beta(t)$  in this city has the same number of dominant spectral peaks all having the same periods like that in the city of Liverpool before and after vaccination and in it prior to vaccination.



(a)  $\beta(t)$  of Liverpool from 1974 to 1986 (b) Modulus of the Fourier transform of  $\beta(t)$  in Liverpool from 1974 to 1986

**Figure 7.6:** Time-dependent transmission rate  $\beta(t)$  of Liverpool from year 1974 to 1986 and the modulus of its Fourier transform. The parameters are  $\delta = 1/64/52/week$ ,  $a = 52/52/week$ ,  $\nu = 52/52/week$ ,  $g = 1/16/52/week$  and initial values are  $S(0) = 0.25$ ,  $E(0) = 0.001$ ,  $I(0) = 0.001$ ,  $A(0) = 0.7$ .



**Figure 7.7:** Time-dependent transmission rate  $\beta(t)$  of London from year 1974 to 1985 and the modulus of its Fourier transform. The parameters are  $\delta = 1/64/52/week$ ,  $a = 52/52/week$ ,  $\nu = 52/52/week$ ,  $g = 1/16/52/week$  and initial values are  $S(0)=0.15$ ,  $E(0)=1 \text{ e-}05$ ,  $I(0)=1 \text{ e-}04$ ,  $A(0)=0.7$ .

Although the dominant frequencies for pre-vaccination and post vaccination data from both cities are the same, the peak values are greater in the pre-vaccination case.

For the same reason as with the pre-vaccination data, the dominant frequencies in London have less noise than those in Liverpool. Also, as with pre-vaccination data, there is no significant difference in the values of the dominant frequencies in Liverpool as opposed to London where there is a noticeable difference in the values. The fact that vaccination does not change the dominant frequencies in both cities strengthens our belief that transmission of measles is driven by school seasons as well as other seasonal factors. We find the transmission cycles to be synchronized in different cities.

## 7.12 Discussion

We present an efficient model with and without vaccination for childhood infectious diseases. Our mathematical and numerical investigations have revealed a number of biologically and mathematically significant results that provides theoretical framework for public health interventions.

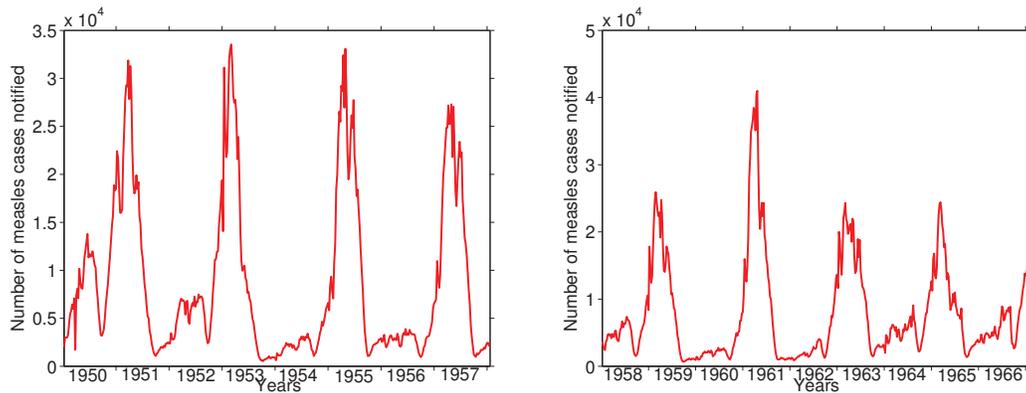
We conjecture that when  $R_0 \geq 1$ , the solutions of the *SEIRA* system goes to the endemic equilibrium and when  $R_0 < 1$ , they go to the disease-free equilibrium.  $R_0 = \frac{a\beta\delta(1-p)}{(a+g)(g+\delta)(\nu+g)}$  is the basic reproduction number.

Sensitivity analysis proves the importance of quarantining patients to prevent an epidemic outbreak. It reveals that birth and removal rate are the most important factors in controlling the endemic level of patients which in-dicts the importance of medication treatment. We equally find out from from our sensitivity analysis that the transmission rate is one of the most important parameter in controlling the endemic level of infectives, When an outbreak occurs and the number of people that are infected in an outbreak.

We equally present algorithms to compute the time-dependent transmission rate from pre and post vaccination prevalence and incidence data. We illustrate the efficiency of these algorithms using London and Liverpool measles data. The extracted transmission rate functions have two dominant spectral peaks with frequencies 1/year and 3/year. These dominant frequencies are neither affected by vaccination nor the city in question. The 1/year dominant frequency is consistent with common belief that measles is driven by seasonal factors such as environmental changes and immune system changes and the 3/year frequency indicates the superiority of school seasons in driving

measles transmission over other seasonal factors. The 1 per year and the 3 per year peaks are comparable for both pre and post-vaccination data from Liverpool, whereas the 1 per year peak is larger than the 3 per year peak for both pre and post-vaccination data from London. This is because London is a landlocked city and thus has large temperature variations which strongly affect Paramyxovirus (measles virus) seasonality with subsequent influence on its transmission. Weather variation thus is as important as school seasons in modulating the transmission of Paramyxovirus when it comes to landlocked cities. Liverpool on the other hand is a coastal city with stable temperature and thus the main modulator of the transmission of Paramyxovirus for this city was school seasons. The dominant frequencies in London have less noise than those in Liverpool because the population of London is greater than that of Liverpool and it is less sensitive to unexpected factors. The Matlab in built function `fft` was used to compute the Fast Fourier Transforms (FFT) of the data. The frequencies resolve by `fft` are  $\frac{1}{T}, \frac{2}{T}, \frac{3}{T}, \dots, \frac{N}{2T}$ ,  $T = N\Delta t$  where  $N =$  Sampled points,  $\Delta t$  sampling time interval. The first frequency is called the dominant frequency and the last one is the Nyquist critical frequency. By taking the absolute value of the FFT we obtained the amplitude spectrum shown in Figures (7.4b), (7.5b), (7.6b), and (7.7b).

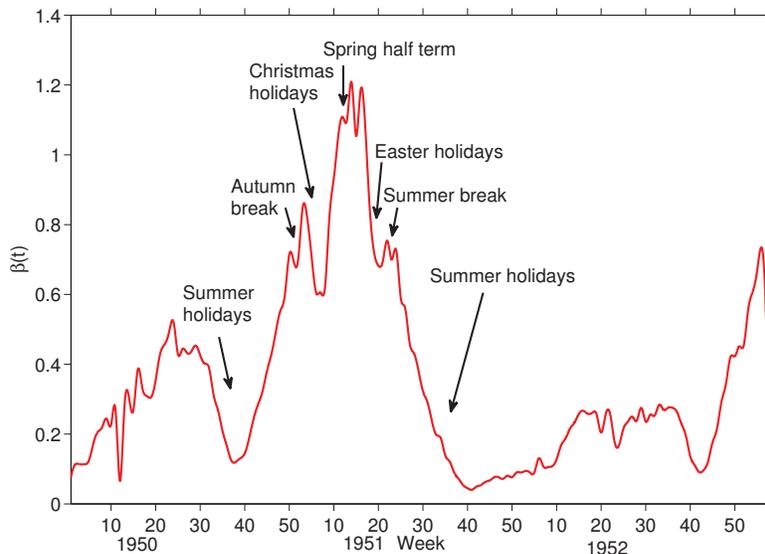
Fine P.E. and Clarkson J.A. [145], estimated the transmission parameter using the notification data from England and Wales from 1950-1979. To compare our results to those of these authors, we equally apply our algorithm to these data shown in Figure (7.8) The figure shows that the outbreaks within this period were biennial in nature, with the magnitude of those in even numbered years higher than those in the odd numbered years.



**Figure 7.8:** 1950-1966 weekly notification data for England and Wales.

Figure (7.9) shows the results obtained by applying our algorithm to these data. As as in [145], the form of the transmission rate for odd and even numbered years' outbreaks differ only in magnitude. This difference in magnitude can be credited to the difference in the number of susceptibles available. Since the trend is similar, to analyze our extracted transmission parameter, we consider just the data for few outbreaks. The figure shows that the trend of the transmission rate is similar to that of the notified measles cases in Figure (7.8). English schools have three terms; Autumn, Spring and Summer. These terms respectively runs from September to mid December (followed by two weeks Christmas holidays), January to late March (followed by two weeks Easter holidays) and March to mid July (followed by six weeks of Summer holidays). Each term is divided into half by a half term break. The transmission parameter drops down whenever students were either on major holidays or midterm breaks and begin rising immediately as soon as school resumes, reaching a peak value sometime within the term. There are three main lengthy decline and steep rise. The lengthy decline could be attributed to major school hol-

idays and the steep rise to school openings after major school holidays. The lowest incidence were reported during period when students were on summer holidays. These support the assertion in [146, 149] that measles transmission is mostly driven by school contacts.



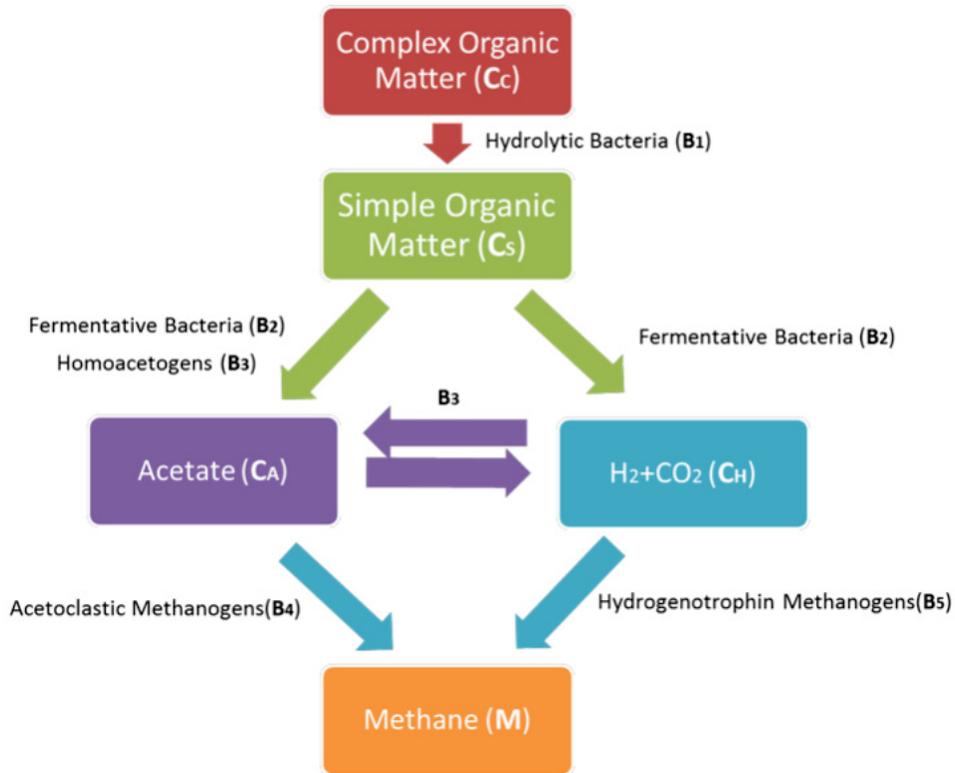
**Figure 7.9:** England and Wales time-dependent transmission rate  $\beta(t)$  from 1950 to 1952. The parameters are  $\delta = 1/64/52/\text{week}$ ,  $a = 52/52/\text{week}$ ,  $\nu = 52/52/\text{week}$ ,  $g = 1/16/52/\text{week}$  and initial values are  $S(0) = 0.2$ ,  $E(0) = 0.003$ ,  $I(0) = 0.003$ ,  $A(0) = 0.79$ .

We believe that our algorithms could be used to estimate the transmission rate of another infectious childhood disease. The choice of the algorithm to use depends on the available data and on whether or not vaccination has been applied to a percentage of children in the region. For almost any type of infectious diseases, the derivations of our prevalence and incidence formulas can be applied with necessary modifications in disease transmission models.

# Chapter 8

## Future directions

In designing our predictive model for methane emissions from oil sands tailings ponds (Chapter 4) we assumed that tailings' hydrocarbons are immediately converted to greenhouse gasses by indigenous microbes after a lag period. The population of the microbes was assumed to be homogenous. The lag period for each organic compound in Chapter 4 accounts for the time it takes to complete all the intermediate processes leading to methane generation. Biodegradation in that sense is only the process that immediately leads to the generation of methane. In general, biodegradation occurs through three main stages (see Figure (8.1)). At the first stage, complex organic matter ( $C_C$ ) are hydrolyzed by extracellular enzymes released from hydrolytic bacteria ( $B_1$ ) to relatively simple, soluble organic matter ( $C_S$ ) that can be transported through the cell membrane of a bacterium. Secondly, simple organic compounds are then fermented or oxidized to acetate  $C_A$  with carbon dioxide and hydrogen ( $C_H$ ) as by-products by either fermentative bacteria ( $B_2$ ) or homoacetogens ( $B_3$ ). At this point,  $H_2$  and acetate can be utilized and/or produced by several homoace-



**Figure 8.1:** General schematic biodegradation pathway of organic matter

togens. Finally for a complete biodegradation process, acetates are consumed by acetoclastic methanogens  $B_4$  and  $H_2$  by hydrogenotrophic methanogens producing methane and carbon dioxide as by-products. Thus with other assumptions that lead to the predictive model unchanged, the general model for the methane biogenesis from a single tailing hydrocarbon is given by the

system of equations below:

$$\begin{aligned}
\frac{dB_1}{dt} &= \mu_1 B_1 \min\{f_1(T_N), g_1(C_C)\} - d_1 B_1 \\
\frac{dC_C}{dt} &= \frac{-1}{r_1} \mu_1 B_1 \min\{f_1(T_N), g_1(C_C)\} + \beta_1 d_1 B_1 + \beta_2 d_2 B_2 + \beta_3 d_3 B_3 \\
&\quad + \beta_4 d_4 B_4 + \beta_5 d_5 B_5 + C_C^{\text{in}} \\
\frac{dC_S}{dt} &= \frac{1}{r_1} \mu_1 B_1 \min\{f_1(T_N), g_1(C_C)\} - \frac{1}{r_2} \mu_2 B_2 \min\{f_2(T_N), g_2(C_S)\} \\
&\quad - \frac{1}{r_3} \mu_3 B_3 \min\{f_3(T_N), g_3(C_S)\} \\
\frac{dB_2}{dt} &= \mu_2 B_2 \min\{f_2(T_N), g_2(C_S)\} - d_2 B_2 \\
\frac{dB_3}{dt} &= \mu_3 B_3 \min\{f_3(T_N), g_3(C_S)\} - d_3 B_3 \\
\frac{dC_A}{dt} &= \frac{\lambda_1}{r_2} \mu_2 B_2 \min\{f_2(T_N), g_2(C_S)\} + \frac{1}{r_3} \mu_3 B_3 \min\{f_3(T_N), g_3(C_S)\} \\
&\quad - \frac{1}{r_4} \mu_4 B_4 \min\{f_4(T_N), g_4(C_A)\} + h_1(C_H, B_3) - h_2(C_A, B_3) \\
\frac{dB_4}{dt} &= \mu_4 B_4 \min\{f_4(T_N), g_4(C_A)\} - d_4 B_4 \\
\frac{dC_H}{dt} &= \frac{\lambda_2}{r_2} \mu_2 B_2 \min\{f_2(T_N), g_2(C_S)\} - \frac{1}{r_5} \mu_5 B_5 \min\{f_5(T_N), g_5(C_H)\} \\
&\quad - h_1(C_H, B_3) + h_2(C_A, B_3) \\
\frac{dB_5}{dt} &= \mu_5 B_5 \min\{f_5(T_N), g_5(C_H)\} - d_5 B_5 \\
\frac{dG}{dt} &= \frac{1}{r_4} \mu_4 B_4 \min\{f_4(T_N), g_4(C_A)\} + \frac{1}{r_5} \mu_5 B_5 \min\{f_5(T_N), g_5(C_H)\} \\
CH_4 &= \eta \Gamma G \\
T_N &= T - \theta_1 B_1 - \theta_2 B_2 - \theta_3 B_3 - \theta_4 B_4 - \theta_5 B_5,
\end{aligned} \tag{8.1}$$

where the parameters and functions are defined in Table 8.1 and

Parameter/function	Definition
$C_C^{\text{in}}$ Constant inflow of $C_C$ $\mu_x$	Maximum growth rate of $B_x$
$d_x$	Per capita death rate of $B_x$
$\beta_x$	Fraction of $C_x$ recycled from dead $B_x$
$\theta_x$	Nitrogen to carbon ratio of $B_x$
$r_x$	Scaling constant for $B_x$
$T_N$	total nitrogen available for bacteria
$f_x(T_N)$	Nitrogen dependent $B_x$ per capita growth rate
$g_x(C_y)$	$C_y$ dependent $B_x$ per capita growth rate function
$h_1(C_H, B_3)$	Model for the conversion of $C_H$ to $C_A$
$h_2(C_A, B_3)$	Model for the conversion of $C_A$ to $C_H$
$\lambda_x$	Interspecific interference coefficient for $B_x$
$\Gamma$	expected $\text{CH}_4$ yield of $C_C$
$\eta$	maximum theoretical yield of $\text{CH}_4$

**Table 8.1:** Definition of the parameters of System (8.1)

$0 < \lambda_i < 1$ ,  $i = 1, 2$ ,  $x = 1, \dots, 5$ ,  $y = 1, \dots, 5$ , This can be generalized to any number of biodegradable tailing pond's hydrocarbons as in Chapter 4. We will analyze this model in the future and when data that quantify  $B_x$ , become available, we will fit it to the model and compare the results to those obtained in this thesis. Also, biodegradation is influenced by many factors such as the temperature, pH and moisture content of the microbes. In Chapter 4, assumptions were made to accommodate these factors. In the future, we will study the sensitivity of the emitted greenhouse gasses to all these factors.

The interaction between bacteria and algae in pelagic ecosystems is complex [12]. Bacteria are nutrient-rich organisms whose growth can be easily limited by nutrient supply and organic carbon produced by plants and algae, which have highly flexible stoichiometry [25]. Suspended algae, also called phytoplankton, grow in open water by taking up nutrients such as phosphorus and nitrogen from the water and capturing energy from light. While algae are an important source of dissolved organic carbon to bacteria, algae and bacteria compete with each other for nutrient elements [21]. These complicated interactions can be effectively modeled and explored by some simple yet powerful stoichiometric constraints.

We are in the process of formulating a model that incorporates nutrient cycling, energy flow, and the growth dependence on light and cellular nutrient contents. The main objectives here are twofold: Firstly, to examine how the key features of a stratified lake, such as the depths of epilimnion and hypolimnion, nutrient and light availability, affect the persistence, extinction and biodiversity of algae and bacteria in a stratified lake; Secondly, to discuss how human activities, such as nearby agriculture or industrial pollution, cause harmful algal blooms (HAB) in a stratified lake.

In a temperate deep lake, the water column is seasonally stratified by a thermocline into two zones, epilimnion and hypolimnion. The epilimnion is the upper zone which is warm with sufficient light and well mixed overnight. The hypolimnion is the bottom colder zone, which is relatively dark, quiescent and normally not mixed. As a consequence, the algal and bacterial densities are homogeneous along the water column in the epilimnion but nonhomogeneous in the hypolimnion. Species movements via turbulent mixing and gravitational

sinking will be explicitly modeled with stoichiometric constraints. This novel model is the hybrid of highly interconnected nonlinear partial and ordinary differential equations.

# Bibliography

- [1] Nidal Abu Laban, Anh Dao, Kathleen Semple, and Julia Foght. Biodegradation of c7 and c8 iso-alkanes under methanogenic conditions. *Environmental microbiology*, 17(12):4898–4915, 2015.
- [2] Martin Alexander. *Biodegradation and bioremediation*. Gulf Professional Publishing, 1999.
- [3] Teri C Balser and Mary K Firestone. Linking microbial community composition and soil processes in a california annual grassland and mixed-conifer forest. *Biogeochemistry*, 73(2):395–415, 2005.
- [4] Frederick Frost Blackman. Optima and limiting factors. *Annals of Botany*, 19(74):281–295, 1905.
- [5] Z Burkus, J Wheler, and S Pletcher. Ghg emissions from oil sands tailings ponds: Overview and modelling based on fermentable substrates. 2014.
- [6] Douglas S Clark and Harvey W Blanch. *Biochemical engineering*. CRC Press, 1997.
- [7] CT Codeco and JP Grover. Competition along a spatial gradient of resource supply: a microbial experimental model. *The American Naturalist*, 157(3):300–315, 2001.
- [8] David C Coleman, CV Cole, HW Hunt, and Donald A Klein. Trophic interactions in soils as they affect energy and nutrient dynamics. i. introduction. *Microbial Ecology*, 4(4):345–349, 1977.
- [9] CE Victoria Collins, Julia M Foght, and Tariq Siddique. Co-occurrence of methanogenesis and n<sub>2</sub> fixation in oil sands tailings. *Science of The Total Environment*, 565:306–312, 2016.
- [10] John P Connolly. Modeling carbon utilization by bacteria in natural water systems. *Modeling the Metabolic and Physiologic Activities of Microorganisms*, 1992.

- [11] DE Contois. Kinetics of bacterial growth: relationship between population density and specific growth rate of continuous cultures. *Microbiology*, 21(1):40–50, 1959.
- [12] James B Cotner and Bopaiah A Biddanda. Small players, large role: microbial influence on biogeochemical processes in pelagic aquatic ecosystems. *Ecosystems*, 5(2):105–121, 2002.
- [13] MR Droop. Vitamin b 12 and marine ecology. iv. the kinetics of uptake, growth and inhibition in *monochrysis lutheri*. *Journal of the Marine Biological Association of the United Kingdom*, 48(03):689–733, 1968.
- [14] Paul Edmonds. *Microbiology: an environmental perspective*. Macmillan, 1978.
- [15] Mahsa Ghovvati, Gholam Khayati, Hossein Attar, and Ali Vaziri. Comparison across growth kinetic models of alkaline protease production in batch and fed-batch fermentation using hybrid genetic algorithm and particle swarm optimization. *Biotechnology & Biotechnological Equipment*, 29(6):1216–1225, 2015.
- [16] RF Grant, M Nyborg, and JW Laidlaw. Evolution of nitrous oxide from soil: I. model development. *Soil Science*, 156(4):259–265, 1993.
- [17] Sam Griffiths-Jones, Harpreet Kaur Saini, Stijn van Dongen, and Anton J Enright. mirbase: tools for microrna genomics. *Nucleic acids research*, 36(suppl 1):D154–D158, 2008.
- [18] JB Haldane. s.(1930) enzymes. *London, Longmann and Green*, 1930.
- [19] Keehyun Han and Octave Levenspiel. Extended monod kinetics for substrate, product, and cell inhibition. *Biotechnology and Bioengineering*, 32(4):430–447, 1988.
- [20] Dag O Hessen, James J Elser, Robert W Sterner, and Jotaro Urabe. Ecological stoichiometry: An elementary approach using basic principles. *Limnol. Oceanogr*, 58(6):2219–2236, 2013.
- [21] Dag O Hessen, Kari Nygaard, Kalevi Salonen, and Ansii Vähätalo. The effect of substrate stoichiometry on microbial activity and carbon degradation in humic lakes. *Environment international*, 20(1):67–76, 1994.
- [22] J Hrynyshyn. End pit lakes guidance document 2012: Cumulative environmental management association, fort mcmurray, ab, cema contract no. 2010-0016 rwg. 434 p, 2012.

- [23] David Kirchman. *Processes in microbial ecology*. Oxford University Press, 2012.
- [24] Joseph G Leahy and Rita R Colwell. Microbial degradation of hydrocarbons in the environment. *Microbiological reviews*, 54(3):305–315, 1990.
- [25] W Makino, JB Cotner, RW Sterner, and JJ Elser. Are bacteria more like plants or animals? growth rate and resource dependence of bacterial c: N: P stoichiometry. *Functional Ecology*, 17(1):121–130, 2003.
- [26] Eric Masse, Hubert Salvail, Guillaume Desnoyers, and Melina Arguin. Small rnas controlling iron metabolism. *Current opinion in microbiology*, 10(2):140–145, 2007.
- [27] Daryl L Moorhead and Robert L Sinsabaugh. A theoretical model of litter decay and microbial interaction. *Ecological Monographs*, 76(2):151–174, 2006.
- [28] Jorge J Moré. The levenberg-marquardt algorithm: implementation and theory. In *Numerical analysis*, pages 105–116. Springer, 1978.
- [29] Hermann Moser et al. The dynamics of bacterial populations maintained in the chemostat. *The dynamics of bacterial populations maintained in the chemostat.*, 1958.
- [30] Andrew H Paterson, John E Bowers, Remy Bruggmann, Inna Dubchak, Jane Grimwood, Heidrun Gundlach, Georg Haberer, Uffe Hellsten, Therese Mitros, Alexander Poliakov, et al. The sorghum bicolor genome and the diversification of grasses. *Nature*, 457(7229):551–556, 2009.
- [31] Viktor Popov and Henry Power. *Landfill emission of gases into the atmosphere: boundary element analysis*. Computational Mechanics/WIT Press, 1999.
- [32] Joshua P Schimel and JAY Gullede. Microbial community structure and global trace gases. *Global change biology*, 4(7):745–758, 1998.
- [33] Kathleen Semple and Julia Foght. Unpublished data.
- [34] Mohd Faidz Mohamad Shahimin, Julia M Foght, and Tariq Siddique. Preferential methanogenic biodegradation of short-chain n-alkanes by microbial communities from two different oil sands tailings ponds. *Science of the Total Environment*, 553:250–257, 2016.

- [35] Mohd Faidz Mohamad Shahimin and Tariq Siddique. Methanogenic biodegradation of paraffinic solvent hydrocarbons in two different oil sands tailings. *Science of The Total Environment*, 583:115–122, 2017.
- [36] Mohd Faidz Mohamad Shahimin and Tariq Siddique. Sequential biodegradation of complex naphtha hydrocarbons under methanogenic conditions in two different oil sands tailings. *Environmental Pollution*, 221:398–406, 2017.
- [37] Tariq Siddique, Phillip M Fedorak, and Julia M Foght. Biodegradation of short-chain n-alkanes in oil sands tailings under methanogenic conditions. *Environmental science & technology*, 40(17):5459–5464, 2006.
- [38] Tariq Siddique, Phillip M Fedorak, Michael D MacKinnon, and Julia M Foght. Metabolism of btex and naphtha compounds to methane in oil sands tailings. *Environmental science & technology*, 41(7):2350–2356, 2007.
- [39] Tariq Siddique, Rajender Gupta, Phillip M Fedorak, Michael D MacKinnon, and Julia M Foght. A first approximation kinetic model to predict methane generation from an oil sands tailings settling basin. *Chemosphere*, 72(10):1573–1580, 2008.
- [40] Tariq Siddique, Mohd Faidz Mohamad Shahimin, Saima Zamir, Kathleen Semple, Carmen Li, and Julia M Foght. Long-term incubation reveals methanogenic biodegradation of c5 and c6 iso-alkanes in oil sands tailings. *Environmental science & technology*, 49(24):14732–14739, 2015.
- [41] Tariq Siddique, Tara Penner, Kathleen Semple, and Julia M Foght. Anaerobic biodegradation of longer-chain n-alkanes coupled to methane production in oil sands tailings. *Environmental Science & Technology*, 45(13):5892–5899, 2011.
- [42] Tariq Siddique, Kathleen Semple, Carmen Li, and Julia Foght. Biodegradation of iso- and cyclo-alkanes in oil sands tailings under methanogenic conditions. *Preprint*.
- [43] J Six, RT Conant, E A. Paul, and K Paustian. Stabilization mechanisms of soil organic matter: implications for c-saturation of soils. *Plant and soil*, 241(2):155–176, 2002.
- [44] Michael K Stenstrom and Stephen S Song. Effects of oxygen transport limitation on nitrification in the activated sludge process. *Research Journal of the Water Pollution Control Federation*, pages 208–219, 1991.

- [45] Robert Warner Sterner and James J Elser. *Ecological stoichiometry: the biology of elements from molecules to the biosphere*. Princeton University Press, 2002.
- [46] GE Symons and AM Buswell. The methane fermentation of carbohydrates1, 2. *Journal of the American Chemical Society*, 55(5):2028–2036, 1933.
- [47] Zeli Tan and Qianlai Zhuang. Arctic lakes are continuous methane sources to the atmosphere under warming conditions. *Environmental Research Letters*, 10(5):054016, 2015.
- [48] N Uemura, J Takahashi, and K Ueda. Fundamental studies on cultivation of hydrocarbon-utilizing microorganisms. 2. determination of transfer rate of n-pentane into culture medium. *JOURNAL OF FERMENTATION TECHNOLOGY*, 47(3):220–+, 1969.
- [49] Diana H Wall and John C Moore. Interactions underground: soil biodiversity, mutualism, and ecosystem processes. *BioScience*, 49(2):109–117, 1999.
- [50] Hans V Westerhoff, Juke S Lolkema, Roel Otto, and Klaas J Hellingwerf. Thermodynamics of growth non-equilibrium thermodynamics of bacterial growth the phenomenological and the mosaic approach. *Biochimica et Biophysica Acta (BBA)-Reviews on Bioenergetics*, 683(3-4):181–220, 1982.
- [51] TC Whitmore. A review of some aspects of tropical rain forest seedling ecology with suggestions for further enquiry. *Man and the Biosphere Series*, 17:3–40, 1996.
- [52] Gregory P Zogg, Donald R Zak, David B Ringelberg, David C White, Neil W MacDonald, and Kurt S Pregitzer. Compositional and functional shifts in microbial communities due to soil warming. *Soil Science Society of America Journal*, 61(2):475–481, 1997.
- [53] Andersen OK, Goldman JC, Caron DA, Dennett MR(1986). *nutrient cycling in a microflagellate food chain: III. Phosphorus dynamics*. Mar Ecol-Prog Ser 31: 47-55.
- [54] Barsdate RJ, Prentki RT, Fenchel T(1974). *Phosphorus cycle of model ecosystems: Significance for decomposer food chains and effect of bacterial grazers*. Oikos 25: 239-251.

- [55] Bratbak G (1985). *Bacterial Biovolume and Biomass Estimations*. Appl Environ Microbiol 49(6):1488-1493.
- [56] Caron DA, Goldman JC, Anderson OK, Dennett MR (1985). *Nutrient cycling in a macroflagellate food chain: II. Population dynamics and carbon cycling*. Mar Ecol-Prog Ser 24: 243-254.
- [57] Cross WF, Benstead JP, Frost PC, Thomas SA (2005). *Ecological stoichiometry in freshwater benthic systems: recent progress and perspectives*. Freshwater Biol 50: 1895-1912.
- [58] Chitnis N, Hyman JM, Cushing JM (2008). *Determining Important Parameters in the Spread of Malaria Through the Sensitivity Analysis of a Mathematical Model*. Bull. Math. Biol. (2008) 70:1272-1296.
- [59] Dhooge A, Govaerts W, Kuznetsov, YA, Mestrom W, Riet AM, and Sautois B(2006). *MATCONT and CL-MATCONT: Continuation tool boxes in MATLAB*. <http://www.matcont.ugent.be/manual.pdf>
- [60] Ekelund F, Rønn R(1994). *Notes on protozoa in agricultural soil, with emphasis on heterotrophic flagellates and naked amoebae and their ecology*. FEMs Microbiol. Rev. 15:321-363.
- [61] Fenchel T(1975). *The qualitative importance of the benthic microfauna of an arctic tundra pond*. Hydrobiologia 46: 445-464.
- [62] Fenchel T, Harrison P (1976). *The significance of bacterial grazing and mineral cycling for the decomposition of particulate detritus. The Role of Terrestrial and Aquatic Organisms in Decomposition Processes*. (Anderson JM & Macfadyen A, eds), pp. 285-299. Blackwell Scientific Publications, Oxford.
- [63] del Giorgio PA, Cole JJ(1998). *Bacterial growth efficiency in natural aquatic systems*. Annu Rev Ecol Syst 29: 503-541. doi: 10.1146/annurev.ecolsys.29.1.503.
- [64] Goldman JC, Caron DA, Anderson OK, Dennett MR (1985). *Nutrient cycling in a macroflagellate food chain: I. Nitrogen dynamics*. Mar Ecol-Prog Ser 24: 231-242.
- [65] Habte M, Alexander M (1975). *Protozoa as agents responsible for the decline of Xanthomonas campestris in soil*. Appl. Microbiol. 29: 159-164.

- [66] Herbert D, Elsworth R, Telling RC(1956). *The continuous culture of bacteria: a theoretical and experimental study*. J. Gen Microbiology 4: 601:622.
- [67] Hessen DO, Elser JJ, Sterner RW, Urabe J (2013). *Ecological stoichiometry: An elementary approach using basic principles*. Limnol. Oceanogr 58, 6, 2219-2236.
- [68] Hunt HW, Cole CV, Klein DA, and Coleman DC(1977). *A simulation Model for the Effect of Predation On Bacteria in Continuous Culture*. Microbial Ecology 3:259-278.
- [69] Jiang L (2007). *Negative selection effects suppress relationships between bacterial diversity and ecosystem functioning*. Ecology 88: 1075-1085.
- [70] Kooi BW, Boer MP, Kooijman SALM (1997). *Complex dynamic behaviour of autonomous microbial food chains*. J. Math. Biol. 36: 24-40.
- [71] Kong DJ, Foght J, Siddique T, Lewis M, Wang H. *A methane generation model for oil sands basins and end pit lakes*. Manuscript in preparation.
- [72] Marinescu M, Dumitru M, Lacatusu A (2009). *Biodegradation of Petroleum Hydrocarbons in an Artificial Polluted Soil*. Research Journal of Agricultural Science 41(2).
- [73] Monod J (1949). *The growth of bacterial cultures*. Ann. Rev. Microbiol. 3: 371-394.
- [74] Nisbet B (1984). *Nutrition and Feeding Strategies in Protozoa*. Croom Helm, London.
- [75] Novick A, Szilard L(1950). *Description of the chemostat*. Science 112: 715-716
- [76] Rønn R, McCaig AE, Griffiths BS, Prosser JI (2002). *Impact of Protozoan Grazing on Bacterial Community Structure in Soil Microcosms*. AEM,68:12.
- [77] Salceanu PL (2011). *Robust uniform persistence in discrete and continuous dynamical systems using Lyapunov Exponents*. Math. Biosci. Eng. 8(3):807-825.
- [78] Sherr BF, Sherr EB and Hopkinson CS (1988). *Trophic interactions within pelagic microbial communities: indications of feedback regulation of carbon flow*. Hydrobiologia 159: 19-26.

- [79] Singleton P, Sainsbury D (1981). *Introduction to Bacteria for students in the biological sciences*. Ch. 1, John Wiley & Sons Ltd. [editors], Vail-Ballou press, Inc., Binghamton, N.Y.
- [80] H. L. Smith and H. Thieme, “Dynamical Systems and Population Persistence,” Graduate Studies in Mathematics, **118**, Amer. Math. Soc., Providence, RI, 2011.
- [81] Smith HL, Waltman P(1995). *The Theory of the Chemostat: Dynamics of Microbial Competition*. Cambridge Studies in Mathematical Biology.
- [82] Sterner RW, Elser JJ (2002). *Ecological stoichiometry: the biology of elements from molecules to the biosphere*. Princeton University Press.
- [83] Stout JD (1974). *Protozoa*. In Dickinson D.H., and Pugh G.J.F., (eds.), *Biology of Plant Litter Decomposition*, Academic Press, New York, 2: 385-420.
- [84] Wang H, Jiang L, and Weitz JS(2009). *Bacterivorous grazers facilitate organic matter decomposition: a stoichiometric modeling approach*. FEMS Microbiol Ecol, 69:170-179.
- [85] Wang H, Smith HL, Kuang Y, Elser JJ (2007). *Dynamics of stoichiometric bacteria-algae interactions in the epilimnion*. SIAM J Appl Math 68:503-522.
- [86] Yang RD, Humphrey AE(1975). *Dynamics and steady state studies of phenol biodegradation in pure and mixed cultures*. Biotechnol. Bioeng. 17:1211-1235.
- [87] B.S. Fields, R.F. Benson, R.E. Besser, *Legionella and legionnaires’ disease: 25 years of investigation*, Clin. Microbiol. Rev., 15(3) (2002), pp. 506-526.
- [88] C.T. Codeço, *Endemic and epidemic dynamics of cholera: The role of the aquatic reservoir*, BMC Infect. Dis. 1 (2001), p. 1.
- [89] D.M. Hartley, J.G.Morris, D.L Jr, Smith, *Hyperinfectivity: A critical element in the ability of V. cholerae to cause epidemics?*, PLoS Med 3, 2006.
- [90] H.L. DuPoint, C.L. Chappell, C.R. Sterling, P.C. Okhuysen, J.B. Rose, W. Jakubowski, *The infectivity of cryptosporidium parvum in healthy volunteers*, N. Eng. J. Med., 332 (1995), pp. 855-859.

- [91] H. Smith, *Monotone Dynamical Systems: An Introduction to the Theory of Competitive and Cooperative Systems*, Math. Surveys Monogr. 41, AMS, Providence, RI, 1995.
- [92] J.B. Rose, *Environmental ecology of cryptosporidium and public health implications*, Annu. Rev. Public Health, 18 (1997), pp. 135-161.
- [93] J.B. Kaper, J.G. Morris Jr, M.M. Levine, *Cholera* Clin. Microbiol. Rev., 8 (1995), pp. 48-86.
- [94] J.H. Tien, D.J.D. Earn, *Multiple Transmission Pathways and Disease Dynamics in a Waterborne Pathogen Mode*, Bulletin of Mathematical Biology, 72 (2010), pp. 1506-1533.
- [95] J.P. Tian and J. Wang, *Global stability for cholera epidemic models*, Mathematical Biosciences, 232 (2011), pp. 31-41.
- [96] J. Snow, *On the Mode of Communication of Cholera* London, John Churchill, New Burlington Street, England, 1855.
- [97] K.M. Murphy, P. Travers, M. Walport, *Janeway's Immunobiology*, seventh ed., Garland Science, 2007.
- [98] L.A. Sanchez, *Dynamics of the modified Michaelis-Menten system*, J. Math. Anal. Appl., 317 (2006), pp. 71-79.
- [99] L. Chitsulo, D. Engles, A. Montresor, and L. Savioli, *The global status of schistosomiasis and its control*, Acta Trop., 77 (2000), pp. 41-51.
- [100] M.A. Jensen, S. M. Faruque, J. J. Mekalanos, B.R. Levin, *Modeling the role of bacteriophage in the control of cholera outbreaks*, PNAS, 103 (2006), pp. 4652-4657.
- [101] M.K. Estes, E.L. Palmer, and J.F. Obijeski, *Rotavirus: a review*, Curr. Top. Microbiol. Immunol., 105 (1983), pp. 123-184.
- [102] M. M. Levine, R.E. Black, M.L. Clements, D.R. Nalin, L. Cisneros, R.A. Finkelstein, *Volunteer studies in development of vaccines against cholera and enterotoxigenic escherichia coli: a review. In: Holme, T., Holmgren, J., Merson, M.H., Mollby, R. (Eds.), Acute Enteric Infections in Children. New Prospects for Treatment and Prevention*, Elsevier/North-Holland Biomedical Press, Amsterdam(1981), pp. 443-459.

- [103] M.S. Islam, B.S. Drasar, B. Sack, *Probable Role of Blue-green Algae in Maintaining Endemicity and Seasonality of Cholera in Bangladesh: a Hypothesis*, J Diarrhoeal Dis Res 1994 Dec, 12 (1994), pp. 245-256.
- [104] M.S. Wolfe, *Giardiasis*, Clin. Microbiol., 5(1) (1992), pp. 93-100.
- [105] M.W. LeChevallier, W.D. Norton, R.G. Lee, *A contribution to the mathematical theory of epidemics*, Proc. R. Soc. Lond. A, 115 (1995), pp. 48-86.
- [106] M.Y. Li, J.S. Muldowney *Global stability for the SEIR model in epidemiology*, Math. Biosci., 125 (1995), pp. 155-164.
- [107] M.Y. Li, L. Wang, *Global stability in some SEIR epidemic models*, IMA Math. Appl, 126 (2000), pp. 295-311.
- [108] P.O. White, F.J. Fenner, *Medical Virology*, 4th edn. 1994, Academic, San diego.
- [109] P.R. Brayton, M.L. Tamplin, A. Huq, R.R. Colwell, *Enumeration of vibrio cholerae O1 in Bangladesh waters by fluorescent-antibody direct viable count*, Appl. Environ. Microbiol., 53 (1987) pp. 2862-2865.
- [110] R.A. Cash, S.I. Music, J.P. Libonati, M.J. Snyder, R.P. Wenzel, R.B. Hornick, *Response of man to infection with Vibrio cholerae I. Clinical, serologic, and bacteriologic responses to a known inoculum*, J. Infect. Dis., 129 (1974), pp. 45-52.
- [111] R.B. Hornick, S.I. Music, R. Wenzel *The Broad Street pump revisited: response of volunteers to ingested cholera Vibrio*, Bull. N. Y. Acad. Med., 47 (1971), pp. 1181-1191.
- [112] R.C. Rendtorff, *The experimental transmission of human intestinal protozoan parasites: Giardia lamblia cyst given in capsules*, Am. J. Hyg., 59 (1954), pp. 209-220.
- [113] R.I. Joh , H. Wang, H. Weiss, J. S. Weitz, *Dynamics of Indirectly Transmitted Infectious Diseases with Immunological Threshold*, Bull Math Biol., 71 (2009), pp. 845-862.
- [114] R.L. Ward, D.L. Bernstein, E.C. Young, J.R. Sherwood, D.R. Knowlton, G.M. Shiff, *Human rotavirus studies in volunteers: determination of infectious dose and serological respons to infection*, Infectious Dis., 154 (1986), pp. 871-880.

- [115] R.R Colwell, P. Brayton, D. Herrington, B. Tall, A. Huq, M.M Levine, *Viable but non-culturable Vibrio cholerae O1 revert to a cultivable state in the human intestine*, World Journal of Microbiology & Biotechnology 12 (1996), pp. 28-31.
- [116] S.M. Faruque, I. B. Naser, M.J Islam, A.S.G. Faruque, A.N. Ghosh, G.B. Nair, D.A. Sack, and J.J. Mekalanos, *Seasonal epidemics of cholera inversely correlate with the prevalence of environmental cholera phages*, PNAS, 102 (2005), pp.1702-1707.
- [117] S.R. Zaki, P.W. Greer, L.M. Coffield, C.S. Goldsmith, K.B. Nolte, K. Foucar, R.M. Feddersen, R.E. Zumwalt, G.L. Miller, A.S. Khan, P.E. Rollin, T.G. Ksiazek, S.T. Nichol, B.W.J. Mahy and CJ Peters, *Hantavirus pulmonary syndrome-pathogenesis of an emerging infectious disease*, Am. J. Pathol., 146 (1995), pp. 552-579.
- [118] V. Capasso, S.L. Paveri-Fontana, *A mathematical model for the 1973 cholera epidemic in the European Mediterranean region*, Rev. Epidemiol. Sante Publique., 27(2) (1979), pp. 121-32.
- [119] W.O. Kermack and A.G. McKendrick, *Contributions to Mathematical Theory of Epidemics-I*, Proceedings of the Royal Society, 115A, 1927.
- [120] Z. Mukandavire, S. Liao, J. Wang, H. Gaff, D.L. Smith, J.G. Morris Jr., *Estimating the reproductive numbers for the 2008-2009 cholera outbreaks in Zimbabwe*, PNAS, 108 (2011), pp. 8767-8772.
- [121] Asheshov, I & Lahiri M. N. (1931). *The treatment of cholera with bacteriophage*. Ind. Med. Gaz., 66, 179184.
- [122] Baker R. M., Singleton F. L. & Hood, M. A. (1983). *Effects of nutrient deprivation on Vibrio cholerae*. Appl. Environ. Microbiol., 46, 930-940.
- [123] Capasso V. & Paveri-Fontana S. L. (1979). *A mathematical model for the 1973 cholera epidemic in the European Mediterranean region*. Rev. Epidemiol. Sante Publique, 27, 121-132.
- [124] Cash, R.A., Music, S. I., Libonati, J. P., Snyder, M. J., Wenzel, R. P. & Hornick, R. B.(1974). *Response of man to infection with Vibrio cholerae. I. Clinical., serologic, and bacteriologic responses to a known inoculum*. J. Infect. Dis., 129, 45-52.
- [125] Codeço C. T. (2001). *Endemic and epidemic dynamics of cholera: The role of the aquatic reservoir*. BMC Infect. Dis., 1, 1.

- [126] Colwell, R. R., Brayton, P., Herrington, D., Tall, B., Huq, A., & Levine, M. M.(1996). *Viable but non-culturable Vibrio cholerae O1 revert to a cultivable state in the human intestine*. World Journal of Microbiology & Biotechnology, 12, 28-31.
- [127] Davis, W.(2012). *Unpublished Master's thesis*. University of Alberta. 49-95.
- [128] Emch, M., Feldacker, C., Islam, M. S., & Ali, M.(2008). *Seasonality of cholera from 1974 to 2005: a review of global patterns*. International Journal of Health Geographics, 7, 31.
- [129] Faruque, S. M., INaser, B., Islam, M. J , Faruque, A. S. G., Ghosh, A. N., Nair, G.B Sack, D. A. & Mekalanos, J. J. (2005). *Seasonal epidemics of cholera inversely correlate with the prevalence of environmental cholera phages*. PNAS, 102, 1702-1707.
- [130] Hartley D. M., Morris J. G. & Jr, Smith, D. L.(2006). *Hyperinfectivity: A critical element in the ability of V. cholerae to cause epidemics*. PLoS Med 3, e7.
- [131] Islam, M. S., Drasar, B. S. & Sack, B. (1994). *Probable Role of Blue-green Algae in Maintaining Endemicity and Seasonality of Cholera in Bangladesh: a Hypothesis*. J Diarrhoeal Dis Res, 12, 245-256.
- [132] Jensen, M. A., Faruque,S. M., Mekalanos, J. J., & Levin, B. R. (2006). *Modeling the role of bacteriophage in the control of cholera outbreaks*. PNAS 103, 4652-4657.
- [133] Joh, R. I., Wang, H., Weiss, H. & Weitz, J. S. (2009). *Dynamics of Indirectly Transmitted Infectious Diseases with Immunological Threshold*. Bull Math Bio., 71, 845-862.
- [134] McCluskey C. C. & van den Driessche P. (2004) *Global Analysis of Two Tuberculosis Models*. Journal of Dynamics and Differential Equations, 16, 139-166.
- [135] Mukandavire Z., Liao, S., Wang, J., Gaff, H., Smith, D. L. & Jr. Morris, J. G. (2011). *Estimating the reproductive numbers for the 2008-2009 cholera outbreaks in Zimbabwe*. PNAS, 108, 8767-8772.
- [136] Murphy, K. M., Travers, P. & Walport, M. (2007). *Janeways Immunobiology*. Seventh ed., Garland Science.

- [137] Pasricha C. L., de Monte, A. J. H, & OFlynn E. G. (1936). *Bacteriophage in the treatment of cholera*. Ind. Med. Gaz., 71, 6168.
- [138] Singleton, F. L, Atwell, R. W., Jangi, M. S. & Colwell, R. R. (1982). *Effects of temperature and salinity on Vibrio cholerae growth*. Appl. Environ. Microbiol., 44, 10471058.
- [139] Snow, J. (1985). *On the Mode of Communication of Cholera*. London, John Churchill, New Burlington Street, England.
- [140] Tian, J. P. & Wang , J. (2011). *Global stability for cholera epidemic models*. Mathematical Biosciences, 232, 3141.
- [141] Anderson, R.M., May, R.M.: *Infectious Diseases of Humans*. Oxford University Press, Oxford, (1991)
- [142] Bauch, C.T., Earn, D.J.D.: *Transients and Attractors in Epidemics*. Proceedings of the Royal Society of London B, **270** :1573 – 1578, (2003)
- [143] Becker, N.G.:*Analysis of Infectious Disease Data*. Chapman and Hall, New York, (1989).
- [144] Earn, D.J.D., Rohani, P., Bolker, B.M., Grenfell, B.T.: *A simple model for complex dynamical transitions in epidemics*. Science, **287** :667 – 670, (2000)
- [145] Fine, P.E., Clarkson, J.A., :*The efficiency of Measles and Pertussis Notification in England and Wales*. International Journal of epidemiology, **14(1)** :153 – 168, (1985)
- [146] Grassly, N.C., Fraser, C.:*Seasonal infectious disease epidemiology*. Proc R. Soc. B Biol. Sci.**273** :2541 – 2550, (2006).
- [147] Glendinning, P., Perry, L.P.: *Melnikov analysis of chaos in a simple epidemiological model*. Journal of Mathematical Biology, **35(3)** :359, (1997)
- [148] Haderl, K.P.: *Parameter identification in epidemic models*. Mathematical biosciences, **229(2)** :185 – 189, (2011)
- [149] Keeling, M.J., Rohani, P.: *Modeling infectious diseases in humans and animals*. Princeton University Press, (2008)
- [150] Keeling, M.J., Rohani, P., Grenfell, B.T.:*Seasonally forced disease dynamics explored as switching between attractors*. Phys. D., **148** :317 – 335 (2001)

- [151] Lombard, M., Pastoret, P.P., Moulin, A.M.: *A brief history of vaccines and vaccination*. Revue Scientifique et Technique-Office International des Epizooties, **26(1)** :29, (2007)
- [152] Pollicott, M., Wang, H., Weiss, H.: *Recovering the time-dependent transmission rate from infection data*. arXiv:0907.3529v3, (2010)
- [153] Pollicott, M., Wang, H., Weiss, H.: *Extracting the time-dependent transmission rate from infection data via solution of an inverse ODE problem*. Journal of biological dynamics, **6(2)** :509 – 523, (2012)
- [154] Rohani, P., Keeling, M.J., Grenfell, B.T.: *The Interplay Between Determinism and Stochasticity in Childhood Diseases*. Am. Nat., **159** :469–481, (2002)
- [155] The weekly OPCS (Office of Population Censuses and Surveys) reports, the Registrar General's Quarterly or Annual Reports & various English census reports 1948-1967