

INSTITUTO TECNOLÓGICO Y DE ESTUDIOS SUPERIORES DE  
MONTERREY  
CAMPUS ESTADO DE MÉXICO



***SoNA3BS: An Agent-Based Model to Study High-Resolution  
Spatial Heterogeneity in *Aedes aegypti*-Borne Diseases  
Through the Use of Network Theory***

*A thesis submitted in fulfilment of the requirements  
for the degree of Doctor of Philosophy in Computer Science*

Author: Héctor Manuel SÁNCHEZ  
CASTELLANOS

Advisors: Dr. Edgar Emmanuel VALLEJO CLEMENTE  
Dr. John M. MARSHALL

Thesis Committee: Dr. Nicholas C. Manoukis  
Dr. Jorge Adolfo Ramírez Uresti  
Dr. María de los Ángeles Junco Rey

Atizapán de Zaragoza, Estado de México  
3<sup>rd</sup> May, 2017

---

## Abstract

### ***SoNA3BS: An Agent-Based Model to Study High-Resolution Spatial Heterogeneity in *Aedes aegypti*-Borne Diseases Through the Use of Network Theory***

by Héctor Manuel SÁNCHEZ CASTELLANOS

Mosquitoes are responsible for transmitting pathogens that cause around 725,000 human deaths per year. Unfortunately, understanding the factors that facilitate the spread of these diseases is complicated; as studying human-mosquito interactions in the field is a difficult task to undertake. Both mosquitoes and humans have their own set of biological requirements and behaviours, which quickly scales up the number of variables involved in these epidemiological processes.

In this work, we present *SoNA3BS* (Social Networks *Aedes aegypti* Agent-Based Simulation); a computational model created to study the spatial relations of human-mosquito interactions in high-spatial resolution scenarios. This simulation is meant to provide framework to study how variables such as human movement, temperature and mosquito-control interventions affect the potential spread of diseases transmitted by the *Aedes aegypti* mosquito (such as Dengue, Zika and Chikungunya). Specifically, it allows us to study how humans connect to each other epidemiologically through mosquito bites with the use of tools borrowed from network theory analysis.

Throughout this thesis document we will demonstrate the usefulness of this approach by making use of the simulation to analyse a realistic scenario mapped after a location near Catemaco, Veracruz (México); in which we analyse the effect of spatial distribution in the way humans connect to each other epidemiologically through mosquito bites both in presence and absence of mosquito-control interventions. We also provide a thorough description of the way biological processes were mapped into the model and a validation set of experiments in which we compare our system to an independent differential equations model; in which we show that the expected accuracy of our results is within reasonable limits of error.

With these descriptions and experiments we wish to show the way epidemiological analysis and control could benefit from tools more traditionally associated with computer sciences than health disciplines; so that, in the future, we can use them to further our understanding of the way mosquito-borne diseases spread in human populations and how we can make the best use of our available resources in our efforts to bring epidemics to a halt.

*“We wish to pursue the truth no matter where it leads. But to find the truth, we need imagination and skepticism both. We will not be afraid to speculate, but we will be careful to distinguish speculation from fact. The cosmos is full beyond measure of elegant truths; of exquisite interrelationships; of the awesome machinery of nature.”*

Carl Sagan

*Para mi familia...*

*A mi padre, por enseñarme a siempre tener curiosidad sobre como funciona el mundo. Gracias por acompañarme a ver documentales y aprender sobre ciencia e ingeniería.*

*A mi madre, por enseñarme el valor de las cosas y a trabajar duro por lo que quiero. Gracias por alentarme y no dejar que me rindiera aún cuando las cosas fueran difíciles.*

*A mi hermana, por cuidar de mí y enseñarme a ser feliz. Gracias por ayudarme a entender a los demás; y quererme a pesar de mis desplantes y locuras.*

*...gracias por tenerme paciencia y darme todas las herramientas para llegar lejos.*



## *Acknowledgements*

This work would have been impossible without the help and support of many people that I have had the fortune to meet and work with throughout these five years of studies.

First and foremost, I would like to thank my thesis advisor Edgar E. Vallejo for guiding me throughout this difficult academic process. Without his insight and support I wouldn't have been able to achieve my goals. Thanks for making me part of so many research opportunities and for allowing me to develop in the research environment (even when it sometimes meant making mistakes on my own and learning from them).

Along these lines, I would also like to thank the Computer Science's faculty at Tecnológico de Monterrey for sharing their knowledge with us, the graduate students. Special thanks go to both program directors who led the program throughout these five years: Miguel Gonzalez and Raúl Monroy. Your patience and guidance is much appreciated by all of us.

I would also like to thank the people who have supported and hosted me throughout several research visits to US universities:

- Charles E. Taylor: for trusting me and giving me the opportunity to be part of a project that taught me how to work in an international environment.
- John M. Marshall: for the beginning of a collaboration which has sparked many other projects that have opened a whole world of opportunities for me.
- Allison Tatarsky: for taking a chance and supporting me, while leading us into doing awesome malaria research and showing us how to be passionate about what we do.
- David L. Smith: for believing in me and taking an interest in my development as a researcher; for which I will always grateful.

Special thanks go to all my academy friends: Vic, Benji, Christian (fanboy), Julio, Efrén, Acevedo, Ricardo, Sean, Michelle, Aram, Leonel, Joanna, Ari, Pablo, César, Pedro and Juanito. Thanks for putting up with all my rants, and for all the patience in answering my many academic and existential questions. You definitely kept me sane throughout these years.

Also thanks to some of my closest my friends, who have helped and supported me through some difficult times: Erik, Crimi, Charlie, Mojo, Adri, Montis, Checo, Chuy, Iván, Danton, Youss, Loren, David, Uriel and Esther (amongst many others). Thanks for standing up with me both in the best and the worst of times. Without all of you I would be pretty much lost.

And finally, thanks to my ex-students from which I learnt a lot; and with which I had more fun than they would know (even if they misbehaved from time to time... and probably still do).

**Financial Aid** This work was possible because of the financial aid provided by CONACyT (both through the PhD stipend and their “beca-mixta” program) and the scholarship provided Tecnológico de Monterrey. Another thanks goes to the University of California which funded our collaboration with UC-Berkeley, through the UC-Mexus initiative. Finally, thanks to Google for their support through the “Google Research Awards for Latin America”; which provided us with financial support throughout the final year of this project.

# Declaration of Authorship

I, Héctor Manuel SÁNCHEZ CASTELLANOS, declare that this thesis titled, “*SoNA3BS: An Agent-Based Model to Study High-Resolution Spatial Heterogeneity in *Aedes aegypti*-Borne Diseases Through the Use of Network Theory*” and the work presented in it are my own. I confirm that:

- This work was done wholly or mainly while in candidature for a research degree at this University.
- Where any part of this thesis has previously been submitted for a degree or any other qualification at this University or any other institution, this has been clearly stated.
- Where I have consulted the published work of others, this is always clearly attributed.
- Where I have quoted from the work of others, the source is always given. With the exception of such quotations, this thesis is entirely my own work.
- I have acknowledged all main sources of help.
- Where the thesis is based on work done by myself jointly with others, I have made clear exactly what was done by others and what I have contributed myself.

Signed:

---

Date:

---

# Contents

<b>Abstract</b>	i
<b>Acknowledgements</b>	iv
<b>Declaration of Authorship</b>	vi
<b>List of Figures</b>	xii
<b>Abbreviations</b>	xiv
<b>1 Introduction</b>	1
1.1 Background . . . . .	2
1.1.1 Epidemiological . . . . .	2
1.1.2 Modelling . . . . .	3
1.2 Hypotheses . . . . .	5
1.3 Scope . . . . .	6
1.3.1 Human Population Size . . . . .	6
1.3.2 Contacts Networks . . . . .	7
1.3.3 Mosquito Biology . . . . .	8
1.3.4 Pathogens . . . . .	8
1.3.5 <i>Aedes aegypti</i> Control Interventions . . . . .	8
1.4 Thesis Contribution . . . . .	8
1.5 Structure . . . . .	9
<b>2 Biological Background</b>	11
2.1 Mosquito Biology . . . . .	12
2.1.1 Aquatic Stages . . . . .	12
2.1.1.1 Egg . . . . .	13
2.1.1.2 Larva . . . . .	13
2.1.1.3 Pupa . . . . .	13
2.1.2 Adult Stage . . . . .	14
2.1.2.1 Male Adults . . . . .	14
2.1.2.2 Female Adults . . . . .	14
2.1.3 Metabolic Rate Temperature Dependence . . . . .	15

2.2	Mosquito-Borne Disease Transmission . . . . .	15
2.3	Mosquito Population Control and Transmission Reduction Interventions . . . . .	17
2.3.1	Traditional . . . . .	17
2.3.1.1	Insecticide Fogging . . . . .	17
2.3.1.2	Ovitraps . . . . .	18
2.3.1.3	Sugar Baits . . . . .	18
2.3.2	Novel . . . . .	18
2.3.2.1	Sterile Insect Technique (SIT) . . . . .	18
2.3.2.2	Female-Specific Release of Insects with Dominant Lethal Gene (fs RIDL) . . . . .	19
2.3.2.3	Release of <i>Wolbachia</i> -infected mosquitos . . . . .	20
2.3.2.4	Clustered Regularly Interspaced short Palindromic Repeats (CRISPR-CAS9) . . . . .	20
2.3.3	Integrated Vector Management . . . . .	21
2.4	Conclusions . . . . .	21
<b>3</b>	<b>Literature Review</b>	<b>23</b>
3.1	General Overview . . . . .	24
3.2	Mathematical Models . . . . .	25
3.2.1	Reviews . . . . .	25
3.2.2	Otero . . . . .	26
3.2.3	VCOM . . . . .	26
3.2.4	Differential Equations Remarks . . . . .	27
3.3	ABMs . . . . .	27
3.3.1	AGiLESim . . . . .	28
3.3.2	LAIS . . . . .	28
3.3.3	Jacintho . . . . .	29
3.3.4	Deng . . . . .	29
3.3.5	CIMSIM/DENSM & Skeeterbuster . . . . .	30
3.3.6	MASH . . . . .	30
3.3.7	ABM Remarks . . . . .	30
3.4	Network Theory . . . . .	31
3.4.1	Epidemics Dynamics in Networks Topologies . . . . .	32
3.4.2	Epidemic Prediction Through Social Networks Information . . . . .	32
3.4.3	Network Reconstruction . . . . .	34
3.4.4	Human Movement . . . . .	34
3.4.5	Network Theory Remarks . . . . .	35
3.5	Conclusions . . . . .	35
<b>4</b>	<b>Model Description</b>	<b>37</b>
4.1	Programming Languages Used . . . . .	40
4.2	Environment . . . . .	41
4.2.1	Boundaries . . . . .	41
4.2.2	Temperature . . . . .	42
4.2.3	Size . . . . .	42
4.2.4	Observers . . . . .	43
4.2.4.1	Real-Time . . . . .	43

4.2.4.2	Offline . . . . .	43
4.2.5	Simulation Cycle Description . . . . .	43
4.3	Mosquito Agents Description . . . . .	44
4.3.1	Aquatic . . . . .	44
4.3.1.1	Egg . . . . .	44
4.3.1.2	Larva . . . . .	46
4.3.1.3	Pupa . . . . .	48
Failed Emergence Death Probability	48	
4.3.2	Adults . . . . .	48
4.3.2.1	Male and Female Life Cycle . . . . .	49
Males	50	
Females	50	
4.3.3	Schoolfield Metabolic Development . . . . .	51
4.4	Other Agents and Objects . . . . .	52
4.4.1	Humans . . . . .	52
4.4.2	Landscape Objects . . . . .	53
4.4.2.1	Houses . . . . .	54
4.4.2.2	Work Zones . . . . .	54
4.4.2.3	Sugar Sources . . . . .	54
4.4.2.4	Egg Laying Site . . . . .	54
4.4.2.5	Mating Site . . . . .	55
4.5	Interventions . . . . .	55
4.5.1	Fogging . . . . .	56
4.5.2	Sugar Baits . . . . .	56
4.5.3	Ovitraps . . . . .	56
4.5.4	Sterile Males . . . . .	56
4.5.5	<i>Wolbachia</i> . . . . .	56
4.5.6	RIDL . . . . .	57
4.6	Computational Time . . . . .	57
4.7	Conclusions . . . . .	58
<b>5</b>	<b>Model Validation</b> . . . . .	<b>59</b>
5.1	Methods . . . . .	60
5.1.1	Stochastic Model Description . . . . .	60
5.1.2	Methodology . . . . .	62
5.1.2.1	Factorial Experiment Setup . . . . .	62
5.1.2.2	Seasonality Experiment . . . . .	63
5.2	Results . . . . .	64
5.2.1	Factorial Experiment . . . . .	65
5.2.2	Seasonality Experiment . . . . .	65
5.3	Discussion . . . . .	67
5.3.1	Density-Dependent Modelling . . . . .	67
5.3.2	Gonotrophic Cycles . . . . .	69
5.3.3	Remarks and Limitations . . . . .	70
5.4	Conclusions . . . . .	72
<b>6</b>	<b>Spatial Heterogeneity</b> . . . . .	<b>73</b>

6.1	Methods . . . . .	74
6.1.1	Simulated Scenarios . . . . .	74
6.1.1.1	Landscape and Humans . . . . .	74
6.1.1.2	Time and Temperature . . . . .	76
6.1.1.3	Mosquitoes . . . . .	76
6.1.2	Simulated Mosquito-Controlled Interventions . . . . .	77
6.1.3	Comparison Metrics . . . . .	78
6.1.3.1	Population Dynamics . . . . .	78
6.1.3.2	Networks . . . . .	78
6.2	Results . . . . .	80
6.2.1	Population Dynamics . . . . .	81
6.2.2	Vectorial-Contact Networks . . . . .	82
6.2.3	Interventions Effects . . . . .	86
6.2.3.1	Population Dynamics . . . . .	86
6.2.3.2	Vectorial-Contact Networks . . . . .	87
6.3	Discussion . . . . .	88
6.4	Conclusions . . . . .	90
<b>7</b>	<b>Conclusions and Future Work</b>	<b>92</b>
7.1	Future Work . . . . .	94
<b>A</b>	<b>Developed Software</b>	<b>98</b>
A.1	SoNA3BS . . . . .	98
A.2	EpiSoNA . . . . .	99
A.3	BehaviourSpaceInterface: . . . . .	99
<b>B</b>	<b>PajaroLoco</b>	<b>101</b>
B.1	Project's Description . . . . .	101
B.2	Publications . . . . .	102
<b>C</b>	<b>ZikViSoN: Zika Virus through Social Networks</b>	<b>103</b>
C.1	Goal . . . . .	103
<b>D</b>	<b>MASH: Modular Analysis and Simulation for human Health</b>	<b>105</b>
D.1	Project's Description . . . . .	105
D.1.1	Event-Driven ABM . . . . .	106
D.1.2	Movement Kernels . . . . .	107
D.1.3	Scope . . . . .	107
<b>E</b>	<b>VCOM: Vector Control Optimization Model</b>	<b>108</b>
E.1	Project's Description . . . . .	108
E.2	Publications . . . . .	110
<b>F</b>	<b>CRISPR-Cas 9</b>	<b>111</b>
F.1	Project's Description . . . . .	111
F.2	Publications . . . . .	112

<b>Bibliography</b>	<b>113</b>
---------------------	------------

# List of Figures

1.1	<i>SoNA3BS</i> Graphical User Interface screenshot . . . . .	5
1.2	Vectorial-contact networks example. . . . .	7
2.1	Photographs of <i>Aedes aegypti</i> mosquitos on their different life stages . . . . .	12
2.2	Mosquito-borne pathogen disease transmission cycle. . . . .	16
2.3	Female-specific RIDL Mendelian inheritance table. . . . .	19
2.4	Cytoplasmic incompatibility for <i>Wolbachia</i> pathogen transmission in mosquito populations. . . . .	20
3.1	Example of an analysis of the impact of vector-control intentions performed in VCOM. . . . .	27
3.2	Example of a response surface analysis of the efficacy of a combination of insecticide-treated nets and ovitraps as performed on MASH. . . . .	31
3.3	ZikViSoN Epidemics on different network topologies. . . . .	33
4.1	Activity diagram of a reactive agent's architecture. . . . .	39
4.2	Density-dependent death probabilities of aquatic mosquitoes . . . . .	46
4.3	Adult mosquito states diagram . . . . .	50
4.4	<i>Aedes aegypti</i> metabolic development parameters for Schoolfield's model. . . . .	51
4.5	Mosquito metabolic development dependence on temperature . . . . .	52
4.6	Agents icons used in <i>SoNA3BS</i> . . . . .	52
4.7	<i>SoNA3BS</i> environment example. . . . .	55
4.8	RIDL inheritance probabilities as modelled in <i>SoNA3BS</i> . . . . .	57
5.1	Weather model fitted to Catemaco's data (2006-2011). . . . .	64
5.2	Differences between ODE and ABM in factorial experiment . . . . .	66
5.3	Adult mosquitoes as a function of time and temperature in both ODE and ABM . . . . .	67
5.4	Modifications performed to the ODE to highlight differences with ABM. . . . .	69
5.5	Collapsing the ODE system into the ABM one by making equivalent modelling assumptions. . . . .	70
6.1	Original houses distribution of the studied population near Catemaco, Veracruz. . . . .	75
6.2	Two simulated spatially heterogeneous scenarios . . . . .	75
6.3	Adult mosquito population dynamics obtained in both levels of spatial heterogeneity . . . . .	81
6.4	Adult mosquitoes mean population sizes in spatial scenarios . . . . .	82
6.5	Vectorial contact transition matrices between humans . . . . .	83
6.6	Heterogeneous scenario transition networks . . . . .	84
6.7	Degree probability distributions of spatial scenarios . . . . .	84
6.8	Networks measures on both scenarios undergoing control interventions . . . . .	85

6.9	Network communities on spatially heterogeneous scenario . . . . .	85
7.1	Example of a preliminary analysis on human movement. . . . .	96
A.1	Screenshots of EpiSoNA's graphical user interface . . . . .	100
B.1	Screenshot of PajaroLoco's graphical user interface . . . . .	102
C.1	Example of an SIR epidemic on a scale-free network as performed in <i>ZikViSoN</i> . .	104
D.1	MASH's example of mosquito movement probabilities . . . . .	106
E.1	VCOM's web graphical user interface . . . . .	109
F.1	Dependence of population elimination probability on homing rate and resistant allele generation rate . . . . .	112

# Abbreviations

<b>ABM</b>	Agent-Based Model
<b>ABS</b>	Agent-Based Simulation
<b>ALife</b>	Artificial Life
<b>BS</b>	Breeding Sites
<b>CRISPR</b>	Clustered Regularly Interspaced Short Palindromic Repeats
<b>fsRIDL</b>	Female-Specific Release of Insects carrying Dominant Lethal Gene
<b>IBM</b>	Individuals-Based Model
<b>IBS</b>	Individuals-Based Simulation
<b>ODE</b>	Ordinary Differential Equations
<b>SDE</b>	Stochastic Differential Equations
<b>SIT</b>	Sterile Insects Technique

# Chapter 1

## Introduction

*“The best scientist is open to experience and begins with romance - the idea that anything is possible.”*

*Ray Bradbury*

Modelling is an inherent activity for humans. From the moment we are born we are continuously gathering information to try to understand our surroundings so that we can respond to them appropriately. Every time we catch a ball, jump over a rope or fall off a bike, we are updating our knowledge on how the world around us behaves and we use our experiences to create internal representation that allow us to perform better in the same, or similar conditions. In a sense, every time we make an attempt to understand or predict the outcome of some phenomenon we are using some model to extrapolate our past experiences into the new events (even if it is just a mental, unwritten one).

Throughout history, we have refined our models and extended their capabilities according to our needs and available resources. We moved from internal mental abstractions, to oral sharing of experiences, and to more sophisticated mathematical written representations of reality in a never-ending quest to understand our environment in a greater detail. As our abstractions became more complex, we faced challenges both in the way we represented and calculated them. We created and refined geometry, algebra and calculus to aid our efforts; but as we developed more precise models, the need for tools that can represent events in more accurate detail has pushed us towards approaches that combine traditional mathematical methods with new computational tools. This is specially true in biological systems where behaviours and

interactions between entities are complex, and in which emergent behaviours in large populations can be the consequence of simple interactions between individuals.

Creating models that take into account populations as composed by single entities provides us with a new level of detail that can help us explore how the interactions between individuals can affect the dynamics of the processes in large populations. These so called, “agent-based models” (ABMs), provide us with a new approach that can be used to simulate and understand phenomena with an unprecedented level of detail. This thesis project follows this idea and applies it to the particular context of diseases transmitted by the *Aedes aegypti* mosquito. In this document, we will describe an agent-based computational simulation in which we look at human-mosquito interactions so that we can generate a framework that allows us to look at how variables such as movement, distribution of breeding sites and the application of mosquito-control interventions could affect the transmission of *Aedes*-borne diseases such as Dengue, Zika, Chikungunya and Yellow Fever.

## 1.1 Background

### 1.1.1 Epidemiological

*Aedes aegypti* mosquitoes are responsible for the transmission of some of the most epidemiologically important vector-borne diseases in recent years: Dengue, Chikungunya and Zika. In the past 50 years, the incidence of Dengue has increased drastically [1, 2]. Recent estimates suggest around 390 million annual cases globally and rising; with nearly 96 million of these exhibiting clinical symptoms [3]. Similarly, Chikungunya and Zika are emerging diseases that are rapidly disseminating in Latin American countries, raising health concerns and placing a heavy burden upon their national health institutes [4–6].

These diseases are transmitted as a side-effect of *Aedes*’ need to obtain blood to complete its life cycle (gravid female mosquitoes require blood from a host for their eggs to be viable). Arboviruses, in turn, use these mosquitoes as vehicles to develop and travel between human hosts. Unfortunately, to date, there exists no effective vaccine to block transmission of any of these pathogens (Dengvaxia, the only available vaccine against dengue has been the subject of recent health concerns [7, 8]); so the control effort has largely focused on the disruption of the mosquito life cycle. This has proven to be a difficult endeavour.

Several mosquito-control interventions exist to date and more are being developed every year. Amongst the traditional *Aedes*-control interventions, spatial insecticide spraying (also known as fogging) is one of the oldest along with sugar baits and ovipositor traps (also known as ovitraps). Sadly, in spite of its long history and widespread use, the efficacy of fogging campaigns has been generally been limited [2, 9]. This, coupled with other traditional approaches' limitations, has created a pressing need for novel approaches to contain the pathogens' spread. In recent years, two of the most promising novel interventions have been: release of *Wolbachia*-infected mosquitoes and release of insects carrying a dominant lethal gene (RIDL). *Wolbachia*-based strategies work by infecting mosquitoes with a bacteria which has been shown to limit the potential of some arboviruses (such as Dengue and Chikungunya) to develop and subsequently be transmitted [10]. Female-Specific RIDL techniques, on the other hand, work by genetically modifying mosquitoes so that females carrying the dominant gene do not develop viable wings upon progression to adult stages [11].

Despite the fact that these novel interventions have shown promising results in field trials [12, 13]; evaluating their cost and effectiveness in a wide variety of different settings is both crucial and difficult. Evaluating interventions in the field is expensive both in time and economic resources. This, paired with the fact that mosquito-borne diseases usually affect low-income countries, makes it paramount to predict their impact before actually using them in practice.

### 1.1.2 Modelling

Historically speaking, most of our epidemiological modelling efforts for mosquito-borne diseases have been devoted towards the understanding of dynamics of diseases at a human population level [14]. This top-down approach (represented primarily by ordinary differential equations, or ODEs) has proven to be extremely useful and is considered to be the standard against which every new idea is to be tested. ODEs, however, have their limitations, as stated by Ruas et al.: “*Traditional analytical methods might not be suitable to deal with complex phenomena that are simply too complicated to be analytically tractable, especially when involving non-linear relationships*” [15]. They usually consider individuals in constant contact with each other, that they are immunologically similar and static in space. These approximations are useful for large-scale predictions but limit our understanding of the fine details of epidemic processes (such as the persistence of diseases in low prevalence settings or the existence of “super-spreader” individuals that transmit pathogens to large amounts of the population [16]). This is particularly important

in diseases that are transmitted by mosquitoes; in which the movement of both, humans and mosquitoes can affect the patterns of infection [17], the impact of disease-control interventions and the spread of their pathogens within human populations [18]. It is here where some of the methodological approaches more often used in computer science than epidemiology can come into play.

Agent-based models (ABMs) have been formulated and discussed for some decades now, but it has only been in recent years that advances in computational power have allowed us to start modelling epidemiological processes. This kind of artificial life system is different from traditional differential mathematical ones in the fact that they are built from the level of individuals to conform and simulate populations in computer systems (also known as *in silico* modelling). Their “bottom-up” approach makes them ideal for the study of individual level dynamics such as spatial distribution of individuals (also called spatial heterogeneity), movement, etcetera. In the particular case of vector-borne diseases, these spatial relations are known to be fundamental in the transmission of diseases [17, 18] and we believe that ABMs in combination with the use of network theory can provide profound insight into this kind of analysis.

***SoNA3BS*** As it was previously mentioned, throughout this work we will present *SoNA3BS* (Social Networks *Aedes aegypti* Agent-Based Simulation), an artificial life modelling tool designed to answer questions regarding spatial heterogeneity in the transmission of *Aedes*-borne diseases such as: Dengue, Zika and Chikungunya (a screenshot of its graphical-user interface is shown in figure 1.1). This model is being designed and coded with the purpose of providing a platform in which methodologies that are generally used in computer sciences can be applied in epidemiological contexts. Mainly, the use of agent-based modelling and network theory to provide insight on how movement, the presence of mosquito control interventions, and other related spatial heterogeneities (such as houses locations and mosquitoes breeding sites locations) can affect epidemic spread of mosquito-borne pathogens. This software, and its source code are publicly available on our github repository (<https://github.com/Chipdelmal/SoNA3BS>).

As a note, we should clarify that we will use the terms: agent-based model (ABM) and individual-based model (IBM), indistinctly. This is because, in the literature, the terms usually refer to the same type of simulation tools but in different fields of application (computational and ecological, respectively).

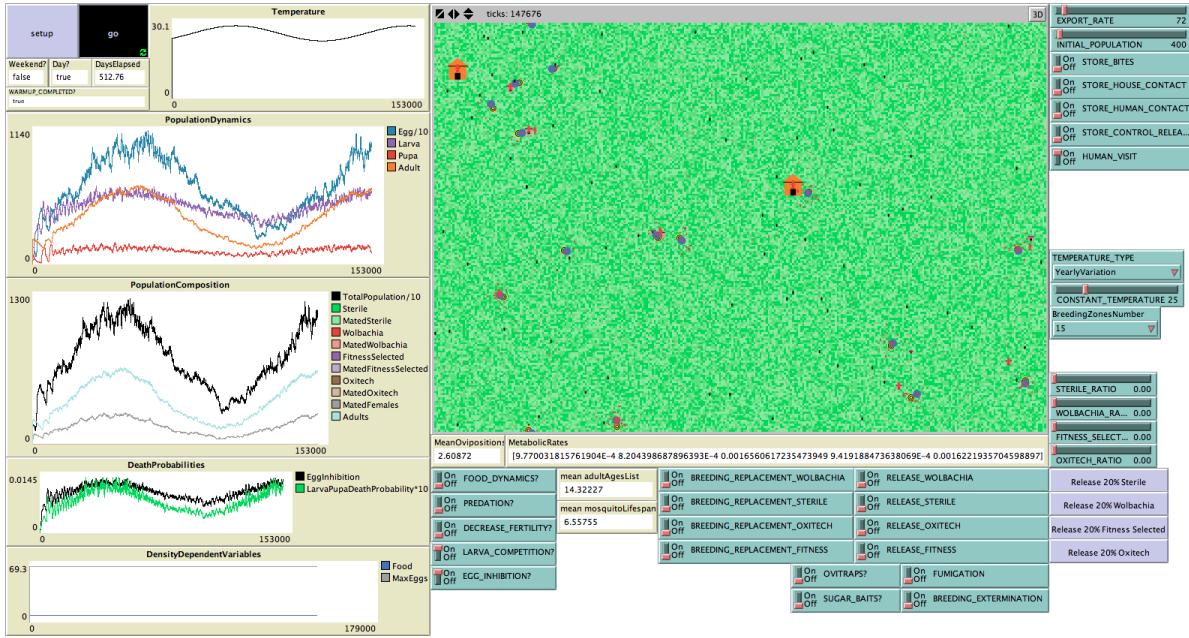


FIGURE 1.1: *SoNA3BS* Graphical User Interface (GUI) Screenshot. This capture shows an example of use of our ABM. The most often used parameters and behaviours can be modified directly on the GUI.

## 1.2 Hypotheses

In this work we will develop a simulation framework that should be capable of capturing the most important details of *Aedes aegypti*-human interactions in terms of spatial heterogeneity, and allow us to perform network analysis on the simulated scenarios to provide a test-bed upon which we can recreate realistic epidemiological scenarios. Summarising, the main hypothesis for this work is:

**Main Hypothesis:** The creation and use of an agent-based framework in conjunction with network theory analytical tools can provide epidemiological insight on how the spatial relations of human-mosquito interactions affect the potential spread of *Aedes aegypti*-transmitted diseases.

To narrow this idea into more specific objectives, we further divided this main hypothesis into three experimental sub-hypotheses that will be addressed throughout the different chapters of this document:

**Sub-hypothesis 1:** A computational agent-based model that adheres to existing knowledge of *Aedes aegypti* biology can be created and run in reasonable computational times on general purpose computers (on the order of days).

For this to be demonstrated we must provide extensive description of the biological processes and how they were mapped into a computational space; along with its rationale, limitations and sources (chapters 2 and 4). Additionally, we should provide a description of the computational times it took to run the experiments described in this document (chapters 5 and 6).

**Sub-hypothesis 2:** The model's adults mosquito population dynamics matches an independently-created model's number of adult females within the range of 15°C to 32.5°C temperatures and up to 20 breeding sites within an error rate of 1.5.

To test this hypothesis, the difference between the two models should be obtained in a factorial experimental framework to test the statistical differences between them in this operational range. Explanations of the discrepancies, if any, should be provided along with their implications (chapter 5).

**Sub-hypothesis 3:** The model can reconstruct the emergent vectorial-contact networks generated by mosquito bites upon humans, and these networks provide useful insight into the differences in epidemiological processes on different spatial layouts.

For this to be true, a comparison between the analysis of population dynamics and contact networks should be performed in two different spatial scenarios, along with a description of the epidemiological implications of the results (chapter 6).

## 1.3 Scope

To reduce our project's scope to a precise, useful size; we have defined a set of goals to achieve while creating the simulation.

### 1.3.1 Human Population Size

Our study will focus on high-resolution (spatial and temporal) interactions between humans and so, we will simulate small human population sizes (up to 20 houses, where each house harbours

one breeding site). This is because we want to understand the details of human behaviour that make differences in the spread of mosquito-borne pathogens, and as such we want to create highly detailed models of the individuals; which are computationally expensive.

### 1.3.2 Contacts Networks

The second goal of our modelling endeavour is to be able to reconstruct the contact networks that result from consecutive bites from mosquitoes on humans (where each node is a human and each edge a transitional bite). The rationale being that mosquito-borne pathogens spread as a consequence of mosquito bites between humans, so studying these bites in a network-theory framework would allow us to make comparisons not only in terms of population sizes or transmission quantities (such as prevalence) but to actually observe the connections between the persons as consequences of their behaviour and the changes in their environment. An example of this is shown in figure 1.2.

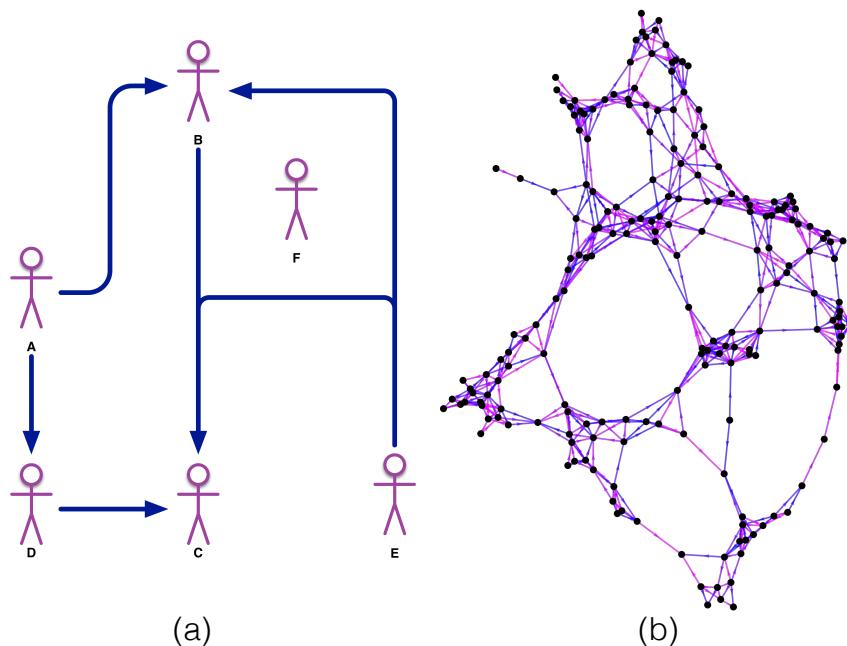


FIGURE 1.2: Vectorial-contact networks example. In figure a an index individual A is bitten by a mosquito in an initial egg-laying cycle. The mosquito then lays eggs and in a subsequent feeding cycle bites person D, which creates a link or “edge” between them. This process is repeated by several mosquitoes across time, creating all the links between persons. Figure b shows this idea in a random spatial network with 200 individuals (the colour of the links represents the frequency of the transition).

### 1.3.3 Mosquito Biology

Mosquito biology should be modelled as accurately as possible at a macroscopic level. In this work we will not go into molecular biology details of mosquitoes or pathogens, so we will mainly focus on characteristics that have a large impact in population dynamics. These characteristics and details will be described throughout chapters 2 and 4.

### 1.3.4 Pathogens

The pathogen transmission network is necessarily a sub-network of the one that emerges from bites between humans (which we will call “vectorial-contacts network” from this point on). For this reason, in the current stage of our research, our model should be able to obtain these vectorial-contact networks as an initial approach towards understanding the transmission patterns of pathogens in human populations. The next step in this line of work would be to simulate disease spread over these graphs. This idea is part of another project our research group is working on and that is described on appendix C.

### 1.3.5 *Aedes aegypti* Control Interventions

As it was mentioned before, the main *Aedes* population control interventions should be included in our simulation. Several different kinds of interventions exist and it is beyond the scope of this work to model them all. As such we will focus on: sterile insect technique, release of insects with dominant lethal gene, fogging, release of *Wolbachia*-infected mosquitoes, sugar baits and ovitraps; which are widely known to be the most relevant ones at the moment of publication of this document.

## 1.4 Thesis Contribution

The main contribution of this thesis project, is the creation (and validation) of an artificial life model that can work as an *in silico* testbed for the analysis of spatial relations in human-mosquito interactions through the study of human’s contact networks that emerge as a consequence of mosquito bites amongst them (in the presence and absence of mosquito control interventions). Computationally speaking, this is translated as the creation an agent-based

model that allows the study of a biological phenomenon in a computer-based framework, so that the interactions between its agents can be mapped and studied in the network theory domain of application.

Additionally, we will show the utility of the system through the study of a realistic example of use. This scenario will simulate a human population near the Mexican town of Catemaco, Veracruz; in which we will make use of the full-fledged ABM to analyse the way the spatial distributions of humans and mosquitoes affect the way they connect epidemiologically, and the implications this could have in mosquito-borne pathogens' spread.

## 1.5 Structure

Up to this point we have described our background, general objectives, scope and hypotheses. We will now present an outline of the structure of this work to better guide the reader on its contents.

- Chapter 1 presented the introduction, in which we state the importance and reach of this research endeavour and present the biological phenomenon we are trying to model. We also stated our objectives, scope and hypothesis to answer.
- Chapter 2 will provide a general description of the biological processes involved in this modelling project. This chapter starts by describing mosquito biology, followed by pathogen transmission processes and finishes with the description of mosquito control interventions.
- Chapter 3 relates us to the state of the art on differential equations models, agent-based simulations and network epidemiology in *Aedes*-borne diseases.
- Chapter 4 describes how *SoNA3BS* incorporates the biological aspects described in chapter 2.1 were mapped into a computational framework.
- Chapter 5 will describe the model's validation experiment performed as a comparison with an independent ODE model at a population dynamics level.
- Chapter 6 presents the use of the model to analyse emergent properties of spatial phenomena in vector-borne diseases with the use of network theory. It also describes the epidemiological implications of the obtained results.
- Chapter 7 presents the conclusions and the future work to be done with the provided simulation.

In this document, we will also present a series of appendices with the description of the software created as part of this and other related collaboration projects in which the authors were involved with during the span of this PhD program.

- Appendix A describes related software created as part of *SoNA3BS* modelling project (ITESM).
- Appendix B provides a description of PajaroLoco (ITESM, UCLA); an animal vocalisations analysis package which is used extensively throughout the project and that was developed by our research group in collaboration with UCLA.
- Appendix C describes ZikViSoN (ITESM, UCB); a project being developed by the Bio-Informatics group at Tecnológico de Monterrey, supported by the Google Awards for Latin-America for Zika prediction through the use of social networks data.
- Appendix D describes MASH (UW, UCB, UCSF and ITESM; amongst others); an agent-based simulation created to study large-scale heterogeneity questions in mosquito-borne pathogens transmission.
- Appendix E presents VCOM (IFH, UCB, UCSF, UW and ITESM); an ordinary-differential equations model for the analysis and optimisation of vector control interventions for malaria control.
- Appendix F shows a model used for evaluation of CRISPR techniques for mosquito population crash (UCD, UCB and ITESM).

# Chapter 2

## Biological Background

*“For me, I am driven by two main philosophies: know more today about the world than I knew yesterday and lessen the suffering of others. You’d be surprised how far that gets you.”*

*Neil deGrasse Tyson*

This thesis represents a modelling work in computer science but it would be meaningless to start describing our computational system without giving a proper description of our object of study. In this spirit, this will provide the reader with a brief explanation of the *Aedes aegypti* life cycle, the factors that affect their development and the behaviours that affect epidemic processes. This, however, is not intended to be a thorough entomological description of mosquito biology, but merely a definition of the behaviours and characteristics that are generally regarded as the most important ones for vector-borne diseases transmission.

One thing to take into account while reading this section, is that the *Aedes aegypti* mosquito species is widespread across tropical regions of the world and can adapt quickly to changes in the environment, so the behaviours described here represent the ones that are most widely accepted by the scientific community. We based most of the main characteristics on the works by Christophers [19] and Becker et al. [20]; complementing and updating with current scientific literature where needed.

We will start by describing the details of mosquito biology that are needed for their individuals to be simulated (the way these individuals were mapped to a computational space will be described in chapter 4). Then, we will describe the general process in which mosquito-borne diseases are transmitted; and from that point we will describe the current *Aedes*-control interventions.

## 2.1 Mosquito Biology

We will begin this section by describing the different life-stages that mosquitos go through (divided in their aquatic and aerial life-phases, which are shown in figure 2.1).

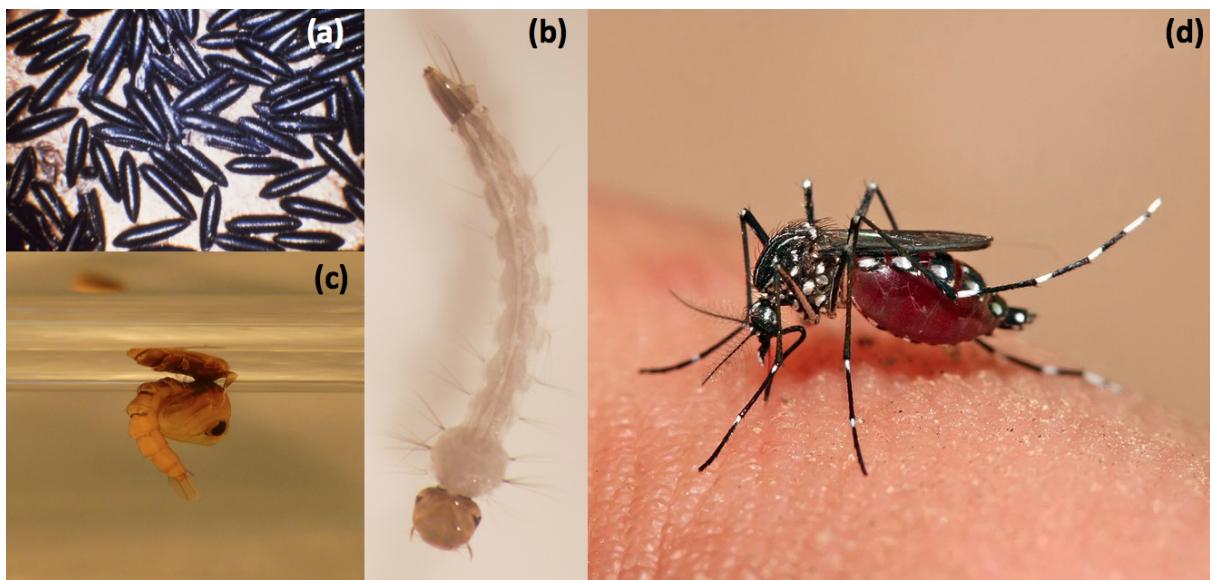


FIGURE 2.1: Photographs of *Aedes aegypti* mosquitos in their different life stages: egg (a), larva (b), pupa (c) and adult (d). Images licensed under creative commons.

### 2.1.1 Aquatic Stages

Aquatic stages comprise the mosquito life cycle from the moment they are laid as eggs in the water to the moment they emerge almost ready to fly. Throughout this period mosquitos transform from eggs, into larvae and then to pupae before becoming adults; and their survival depends greatly on the availability of food and water.

Aquatic phases are not epidemiologically relevant in the sense that individuals in these stages are unable to bite and transmit diseases. Despite of this, they are biologically important because *Aedes aegypti*'s environment's carrying capacity (the maximum amount of individuals that can be sustained at a given environment) is usually dictated by the number of breeding sites where females can lay eggs and their nutrients. Each female lays around 60 eggs in a single egg-laying bout (the exact number of eggs relates roughly to the female's body weight), so these water bodies are easily saturated to the point where the fight for resources starts to limit the population that the environment can handle. Nutrients scarcity, inter and even intra-species predation are part of what is known with the name of "density-dependent limiting factors" which

are behaviours that halt the population growth depending on the number of active individuals of the same species living at any given time point.

#### **2.1.1.1 Egg**

Mosquitos spend around four days in egg stage (figure 2.1a) at an environmental temperature of  $25^{\circ}\text{C}$ . Throughout these period the eggs float in the surface of the water while they develop, where they are vulnerable to desiccation and predators (although they have been found to resist desiccation for periods of up to a year).

Female *Aedes aegypti* mosquitoes are particularly selective about the water bodies in which they lay their eggs. They fly looking for stagnant but clean water containers and adult females usually engage in a behaviour called “skip oviposition” in which they do not lay all their eggs in a single water container but distribute them amongst several ones [21, 22]. This is generally believed to be an evolutionary trait to reduce extinction event probabilities.

#### **2.1.1.2 Larva**

During this phase the mosquitoes spend their time floating and feeding (figure 2.1b). One interesting thing about larvae of this species (as well as other mosquito species) is that, given nutrients scarcity, they may display cannibalistic behaviours to survive both by eating eggs or other larvae [23]. Mosquitoes spend around three days in this state at  $25^{\circ}\text{C}$  and they can die either by lack of nutrients, desiccation, low temperatures or predation.

#### **2.1.1.3 Pupa**

Mosquitoes in pupa phase (figure 2.1c) have limited mobility, although they can purposely sink in response to changes in the water’s surface or light (in contrast with pupae from other insects that are unable to move) [19]. Mosquitoes spend around six days in this state at  $25^{\circ}\text{C}$  and are extremely vulnerable to predators and environmental changes. Pupae have a certain probability to fail their emergence into adult stage in which case they die.

### 2.1.2 Adult Stage

The mosquito adult stage (figure 2.1d) is the most epidemiologically important as it is in this period in which females go into blood-feeding bouts to fulfil their egg-development requirements (and in doing so, they might transmit diseases to humans). This species is most active during the day and its individuals spend most of the night resting (in contrast with other species such as the *Anophelines* that transmit malaria and are mostly nocturnal).

As female and male individuals differentiate their behaviours from this stage on, we will now describe their life cycles separately.

#### 2.1.2.1 Male Adults

After emergence from pupa stage, adult males spend around one day resting, waiting for sexual maturity. After doing so they spend their lives alternating between mating and resting bouts, only interrupted by sugar feeding when needed.

For mating purposes, males fly in swarms to attract females. In the case of *Aedes aegypti* mosquitos, these swarms are usually located very close to their human hosts [24] (in contrast to Anopheles, which determine their swarming spots through contrast markings in the ground). After copulation males can go back to the swarm and mate as many times as needed or possible (only interrupting their cycles by resting and sugar-feeding bouts).

It is important to note that males do not bite humans so they are unable to transmit mosquito-borne diseases. This activity, as it will be described in the following section, is reserved to females.

#### 2.1.2.2 Female Adults

Female *Aedes* emerge ready to mate, in contrast to their male counterpart (although there is evidence suggesting that recently-emerged females might be refractory to insemination [24]). General consensus dictates that females only need to mate once throughout their lifespans (despite there being some debate about this in the literature [25]). Once they have obtained genetic material from males they spend their lives alternating blood-feeding, resting and egg-laying bouts; only interrupted by sugar feeding.

Females, as was described earlier, approach mating swarms and are almost immediately seized by males for mating purposes. Once females have mated, they start their search for blood, which they need for their eggs to develop (this egg-development period is called the female “gonotrophic cycle”). *Ae. aegypti* are highly anthropophilic which means that they prefer biting humans than other kinds of possible hosts (such as dogs or cats). Females can take several attempts at a host until they fulfil their blood need, and how much they are attracted to certain humans varies on carbon dioxide emissions, lactic acid production, pregnancy, etcetera. This makes some mosquitos more attracted to certain persons more than others (a behaviour known as “preferential biting”).

After obtaining blood for egg development, females rest for a determined period before egg-laying (around five days at a temperature of  $25^{\circ}\text{C}$ ). This period varies with temperature and is slightly longer the first time they lay eggs from the subsequent ones.

### 2.1.3 Metabolic Rate Temperature Dependence

One final remark to make is that both aquatic and aerial stages’ rate of the metabolic development depends on the temperature of the environment [26]. This means that the time an individual spends in each life phase depends on the temperature of its surroundings. In general terms, the warmer the weather, the faster do mosquitoes develop. This, however has its limits. Although literature varies wildly in the exact limits of individuals survivability it is fairly consistent that *Ae. aegypti* mosquitoes have trouble developing and surviving outside the range of  $15^{\circ}\text{C}$  to  $35^{\circ}\text{C}$  as individuals (although, mosquito populations as a whole, can be sustained as far north and south as the  $10^{\circ}\text{C}$  isotherms).

## 2.2 Mosquito-Borne Disease Transmission

As it was previously mentioned, female mosquitos are the ones responsible for vector-borne disease transmission. This infection cycle starts when a female mosquito bites and draws blood from an infected human. Once this happens, if the pathogen was ingested by the mosquito, its extrinsic incubation period begins. This period encompasses the time from ingestion to the moment the parasites are developed enough to be infectious to humans again. Once this extrinsic incubation period has passed, if a mosquito bites a non-infected human the disease can

be transmitted from the mosquito to the person. If this occurs, the pathogen starts its intrinsic incubation period until it can be transmitted to mosquitos again.

This process is exemplified in figure 2.2 where a mosquito bites an infected human, becomes infected, then transmits the pathogen to a naive human which, in turn infects another susceptible mosquito, starting the transmission cycle again.

All four *Aedes*-borne diseases: Yellow Fever, Dengue, Chikungunya and Zika; are transmitted this way (although Zika can also be transmitted through sexual contact between humans) and disease-elimination efforts focus on disrupting any part of this transmission process. Killing mosquitoes, preventing them from biting humans, vaccinating humans or even “vaccinating” mosquitoes with other pathogens that block transmissions; are all strategies that are being assessed and compared.

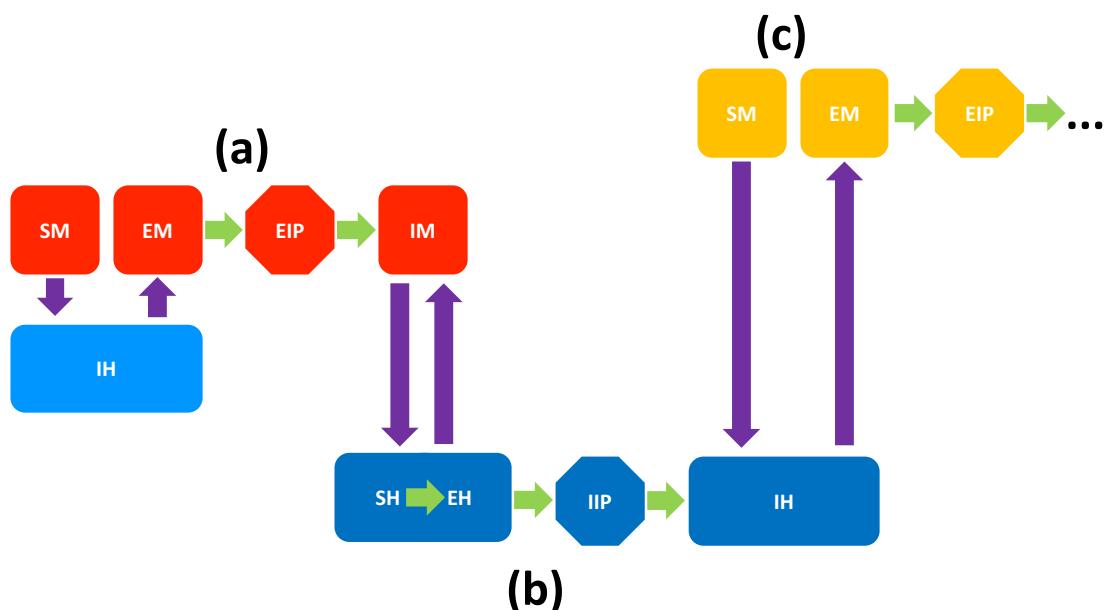


FIGURE 2.2: Example of a mosquito-borne disease transmission cycle. (a) A susceptible mosquito (SM) bites an infected human (IH) and absorbs the pathogen. The mosquito is now exposed (EM). After the extrinsic incubation period (EIP) has passed, if the mosquito bites a susceptible human (SH) this human becomes exposed to the pathogen. (b) The exposed human goes through the intrinsic incubation period (IIP) and after it has passed this human becomes infected. (c) Another susceptible mosquito now bites this infected human and the cycle repeats. In this cycle, humans follow an SIR (susceptible, infected, recovered) behaviour in which they get infected and either recover or die from the disease. Mosquitos, on the other hand, follow an SI (susceptible, infected) one; in which, once infected, they remain that way for the rest of their lives.

## 2.3 Mosquito Population Control and Transmission Reduction Interventions

As it was mentioned before, there are several different types of mosquito-control interventions currently in use or under development. In this section, we will describe briefly the mechanism of work of some of the most relevant interventions at the time of publication of this document.

We will divide this section into traditional and novel control measures, to differentiate between the ones that have been used for a long time and the ones that are currently under development and trial. We will also give a brief description of the idea behind integrated vector control, which involves the study of the synergies in the simultaneous application of several control interventions.

### 2.3.1 Traditional

Mosquitos have been a problem for a large part of human history. As a consequence, several mosquito control interventions have been devised over the years with varying degrees of success. In this section we will describe some of the control measures that have had widespread use in the past (and present) to reduce mosquito population sizes.

#### 2.3.1.1 Insecticide Fogging

Fogging campaigns are arguably the first organised effort to eliminate mosquitos in the modern age. In its most basic form, this technique consists on spraying insecticide on sites where mosquitos have been detected or where they usually gather so that they can be killed in large quantities.

Despite their widespread use, spatial insecticides have been largely unable to eliminate mosquito populations on their own, although they tend to temporarily work in reducing the amount of mosquito adults in the environment [27].

### 2.3.1.2 Ovitraps

Ovitraps attack female mosquitoes that are trying to lay eggs in aquatic habitats. There are several variations of these artefacts ranging from simple soda bottles to more elaborate recipients with long-lasting adhesive surfaces to trap both the adult female and the eggs being laid.

The effectiveness of this intervention is generally limited to the number of alternative oviposition sites so it is largely dependant on the environment in which it is used [28].

### 2.3.1.3 Sugar Baits

Sugar baits focus on the period in which mosquitos need to fulfil their energetic needs and work by laying poisoned sugar sources in the environment so that mosquitos feed on them and die [29, 30]. Again, this kind of trap ranges from extremely cheap and primitive versions to more sophisticated ones but their working principle remains the same.

These traps are not commonly used for *Aedes aegypti* mosquitos control due to the availability of many different sugar-sources in the environment they develop (both natural and artificial). Their use is more widespread in malaria-transmitting mosquitos [29], although they do get used from time to time for *Aedes* control.

## 2.3.2 Novel

Despite their relative inexpensiveness and ease of use, traditional control measures have not been able to halt the spread of *Aedes*-borne diseases (which is reflected in the recurrent epidemics in tropical areas of the world). To address this, several new interventions have been developed in latest years. Here we will describe three of the most promising and used ones: sterile insect technique, release of insects carrying dominant lethal gene, and release of *Wolbachia*-infected mosquitoes.

### 2.3.2.1 Sterile Insect Technique (SIT)

This intervention takes advantage on the fact that female *Aedes* usually mate only once throughout their lifespans. With this in mind, male mosquitos are captured and irradiated to sterilise

them. Once they are sterile they are released again so that females mate with them but have no genetic material to produce offspring.

In general terms, SIT has faced several drawbacks that have limited its success. One of the most important ones is that irradiated mosquitos have been repeatedly found to be less biologically fit with respect to their wild counterparts [31], which impacts their mating behaviours and, ultimately, the success of these campaigns.

### 2.3.2.2 Female-Specific Release of Insects with Dominant Lethal Gene (fsRIDL)

RIDL is a genetic variation transmitted by Mendelian inheritance (shown in figure 2.3). It works by transmitting genes that prevent female mosquitoes to develop proper wings after emerging from their pupa stage [13]. This allows males to further propagate their alleles while females perish upon adulthood. This has the advantage that not only do female mosquitos die before being able to bite humans; but they survive throughout their aquatic stages; which helps to maintain the environment saturated near its carrying capacity and to reduce the adult mosquito populations as a whole. This intervention has been tested in several countries although it's effectiveness and reach is still being studied [32].

		Female		
		RR	Rr	rr
Male	RR	RR   RR	RR   RR	Rr   Rr
	Rr	RR   Rr	RR   Rr	Rr   Rr
	rr	Rr   Rr	Rr   rr	rr   rr

FIGURE 2.3: Female-specific RIDL Mendelian inheritance table. This figure shows the possible genetic combinations of the offspring whenever a male and female mosquito mate ( $r$  being a wild gene and  $R$  the dominant lethal gene). Red states are not possible as females can not reach an adult stage if they present the dominant RIDL gene ( $R$ ). Each cell shows the probabilities of the offspring to inherit each genetic combination.

### 2.3.2.3 Release of *Wolbachia*-infected mosquitos

The *Wolbachia* endosymbiont is a bacteria that has been recently transmitted to *Aedes aegypti* mosquitoes. One of the interesting characteristics of this pathogen is that it has been found to block the transmission of the Dengue, Chikungunya and Zika viruses; while reducing the lifespan of adult females [12, 33] (which reduces the probability of pathogen transmission even further). One of the advantages of this method is that it is transmitted by cytoplasmic incompatibility [9], which allows it to fixate quickly in wild populations.

Female	Male	Offspring
Wild	Wild	Wild
<i>Wolbachia</i>	Wild	<i>Wolbachia</i>
Wild	<i>Wolbachia</i>	No Offspring
<i>Wolbachia</i>	<i>Wolbachia</i>	<i>Wolbachia</i>

FIGURE 2.4: Cytoplasmic incompatibility for *Wolbachia* pathogen transmission in mosquito populations. This table shows the inheritance patterns between wild (uninfected) and *Wolbachia*-infected mosquitos.

The efficacy of releasing mosquitos with the *Wolbachia* pathogen is also being currently studied and some isolated trials have taken place, which show promising results [12]. One of the drawbacks of this intervention, however, is that due to the nature of the transmission process, the pathogen could spread into non-targetted wild populations easily, so caution has to be exerted to avoid releasing a pathogen that could be harmful to other animal species.

### 2.3.2.4 Clustered Regularly Interspaced short Palindromic Repeats (CRISPR-CAS9)

CRISPR-CAS9 technique is mainly used for malaria-transmitting mosquitoes (*Anophelines*) at the time of publication of this work; but it is strongly suggested that could work for *Aedes* mosquitos. This technique will be defined briefly in this section, and a more in-depth description will be provided in appendix F; where we will talk about work performed as part of a collaboration with the University of California, Riverside and California, Berkeley .

There are several variations of this idea and one of the most promising ones is CRISPR-Cas9-based homing system for mosquito population suppression. In this genetic modification technique, mosquitos are altered to make them less fertile but maintaining high inheritance rates;

which, ultimately, might lead to population reduction and crashes [34]. This would allow the reduction and, possibly, complete disruption of transmission chains in human populations due to the lack of vectors that transmit them.

### 2.3.3 Integrated Vector Management

Throughout the years it has become clearer that if we are to stop the spread of diseases transmitted by mosquitos we have to take an integrated approach towards their control [35, 36]. This means that it is unlikely that we will be able to halt current epidemics by just applying one or two types of interventions in any given place, so we need to study the synergies between various control measures to optimise our use of the limited resources devoted to these tasks.

Integrated vector management, however, poses great challenges as the number of variables involved in each specific scenario is extremely large. Weather, coverage, efficacy, human behaviour, costs, etcetera; all affect the expected outcomes of each control measure and combining several at the same time causes the search space to grow exponentially. This is another reason why we need models such as the one presented in this work. With them, we can simulate and optimise the use of resources to control epidemics in an efficient way so that we can attack the spread of these diseases and revert their current trends.

## 2.4 Conclusions

In this chapter we have related the reader with a general description of the mosquito biology as well as some of the terminology that will be used throughout our project. As we have seen, *Aedes aegypti* mosquito life cycle is complex, partly because it involves several stages which have their own associated behaviours and characteristics. This makes it difficult to study the way changes in this cycle impact the grand scheme of vector-borne diseases transmission and how we can more efficiently stop these processes. Moreover, several control interventions are being developed and tested all throughout the world to try to lower the number of cases of the diseases these mosquitos transmit; which creates a pressing need to develop models that help us understand how to bring epidemics to a halt.

In the following chapter we will describe the state of the art of the mosquito-borne diseases modelling to place our model in scientific context and to describe the academic gaps we will attempt to address with it.

# Chapter 3

## Literature Review

*“Somewhere, something incredible is waiting to be known.”*

*Carl Sagan*

Due to the importance of its nature, the field of epidemiological modelling is extremely dynamic. Both the study of mosquito biology and the epidemiology of the diseases they transmit, evolve every day. As such, many types of models are created to alleviate the needs of specific areas of opportunity. This makes it impossible to cover all of the models that are available for the study of mosquito borne-diseases; and this is why, our state of the art we will focus on the most relevant and current developments of the field.

In this chapter we will relate the reader the main work that has been done in recent years in the modelling of mosquito-borne diseases. For clarity, we will divide this chapter into four sections: a general overview of the state of the field, a brief overview of the state of differential equation models, a more thorough description of current mosquito-borne ABM simulations; and, finally, the way network epidemiological models are being incorporated into the field (along with human movement and spatial heterogeneity studies). Throughout these descriptions we will try to focus as much as possible on *Aedes*-borne diseases but, as a large amount of the modelling effort has been devoted to malaria, it will ultimately be impossible to avoid making some comparisons and describing some of these *Anopheline*-oriented models. Similarly, the epidemiological networks models have focused mainly on direct-contact diseases (such as influenza) so we will also describe some of these studies and how they could be translated into the vector-borne scene.

This will serve as a reference point for further chapters to describe our model in the current context of the scientific understanding of the diseases we intend to eradicate.

### 3.1 General Overview

Many distinct models have been created to address specific epidemiological questions regarding mosquito-borne diseases at a large-scale population level. This has allowed the scientific community to understand the general dynamics of vector-borne diseases; but as our knowledge on these subjects has increased, we have come to realise that the interactions between individuals have an important impact on the transmission of these diseases in human populations [37–41]. This has caused that the modelling efforts to slowly shift towards the creation of frameworks that allow the analysis of different levels of interaction in vector-borne diseases.

The shift has been partly motivated by the realisation of the impact that human movement has on the spread of mosquito-borne diseases [38, 39], the interest in understanding the impact of behaviours such as preferential biting [42, 43] and how, in low-prevalence environments, pockets of infections might survive even when eradication is predicted by fully-mixed models.

Generally speaking, traditional differential equations models have been slowly adapted to incorporate heterogeneous behaviours in mosquito-human interactions. ABM's, on the other hand, include these phenomena intrinsically, but have trouble maintaining tractability and scalability. In a sense, each modelling branch has tried to address the inherent limitations of its specific approach, looking for niches of opportunity to cover where their strengths outmatch their weaknesses.

In addition to these two modelling currents, networks structures have slowly started to be integrated in both type of approaches [44–48]. This has been done through individuals contacts analysis (in ABMs) and meta-populations models (in ODEs). This integration, however, has focused mostly on direct-contact diseases and not so much on vector-borne ones, so there is still work to be done to make it useful in the mosquito-borne scene.

## 3.2 Mathematical Models

Differential equation models have the advantage of considerable formality and background over other approaches, as they have been used for many years as the standard for scientific prediction. Their tractability and clarity makes them ideal to provide tools for thorough exploration of epidemiological scenarios; and this is why they remain as the mainstream way to model complex phenomena such as diseases transmission.

In this section we will briefly describe the general state of the mosquito modelling field regarding ODE models through two recent surveys [14, 49] that state the main differences and areas in which these efforts have focused on, as well as the things that are still needed for a more realistic representation of these epidemic processes. We will also describe two systems that are particularly relevant for this research project.

### 3.2.1 Reviews

Reiner et al. provide an extensive review of mosquito-borne mathematical models [14] (including Malaria, Dengue, West-Nile virus, Filariasis, etcetera). In this review the authors identify the aspects in which each system differs from the original Ross-Macdonald representations (established by the year 1970), which are widely regarded as the beginning of epidemiological modelling theory (Smith, et al. provide a description of the historical of the development of these approaches [50]). According to the review, “Ross-Macdonald-like” models assume: homogeneous biting, homogeneous contacts, no spatial distribution, no waning immunity, no mosquito demography, no explicit aquatic population stages, no host population dynamics; amongst many others. In this review, the authors define an “RM” index with which they measure from 0 to 15 how much current models differ from the core Ross-Macdonald assumptions (with 15 being the most changes). A surprising 76% obtained a score of 2 or less, meaning that most of the systems hardly differ from the original assumptions. In particular, the authors identify three main aspects that should be addressed by new studies (as the current ones hardly cover them): heterogeneous biting, poorly mixed mosquito-host encounters, and spatial heterogeneity with seasonality variations. Another phenomenon that most models neglect is the inclusion of aquatic stages and multiple vector-control interventions, which are paramount for elimination efforts to be successful.

Andraud et al. provide a further discrimination and description of the different models that have been proposed specifically for dengue [49]. In this review, most of the systems, again, deal with the case of homogeneous mixing of populations and neglect vector-host dynamics along with stages of development previous to mature reproductive ones (assuming a constant or semi-constant rate of influx of female mosquitoes to the population), which are also identified as niches of opportunity for new models to cover.

Both these reviews highlight and re-iterate the necessity for new models to overcome the inherent limitations of differential equations systems. Namely, the need for the incorporation of heterogeneity analysis and the aquatic phases of development; as well of comprehensive frameworks that allow the simultaneous study of different aspects of the epidemiological processes of mosquito-borne diseases.

### 3.2.2 Otero

This is a stochastic Markov chain simulation was developed to investigate seasonality effects on mosquito population sizes in Argentina [51]. We highlight this model amongst the other ones, as we will use it in chapter 5 to compare our simulation for validation purposes. Because of this, a more thorough description will be provided in the validation section of this work (chapter 5).

This model simulates mosquitoes at the populations level but it includes both: aquatic and adult stages; and metabolic rate dependence on temperature so it is particularly suited for the study of temperature effects and extinction thresholds. The downside of this system is that it does not incorporate any spatial nor behavioural heterogeneity at any level so it is able to make analysis of the impact of these variables upon disease spread. The authors themselves state that: *“The incorporation of the spatial extension of the model seems to be the most urgent need. Dispersal strategies of mosquitoes might be a determining factor in their survival in temperate climates as well as in environments with a low density of breeding sites”* [51].

### 3.2.3 VCOM

Another relevant work is the one developed as a collaboration between researchers at the University of California San Francisco, University of California Berkeley, the Ifakara Health Institute and Tecnológico de Monterrey (amongst others) [52]. This model is being created to test current mosquito-control interventions in a coherent framework that allows the optimisation of their

coverage to minimise costs (more details are provided in appendix E). The project provides a web-based user interface for interventions analysis (figure 3.1) and is being designed for analysis of mosquito-control interventions effects in large homogeneous scenarios.

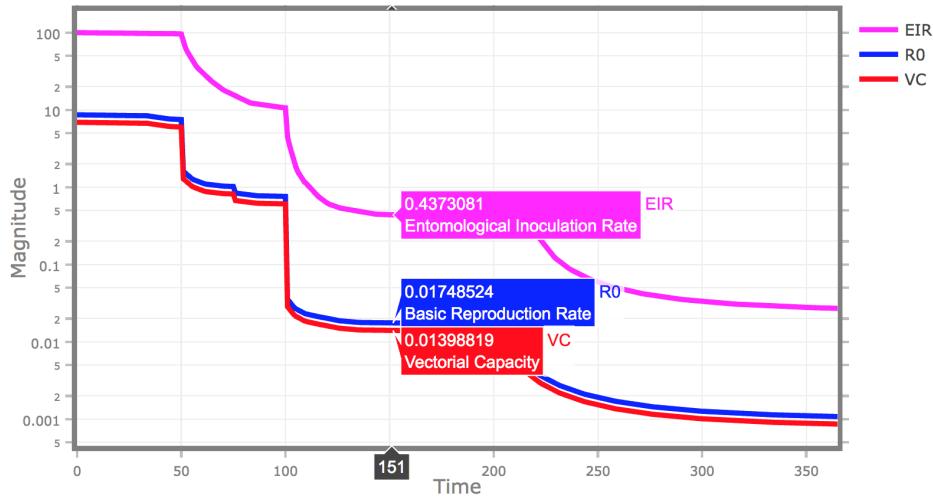


FIGURE 3.1: Example of an analysis of the impact of vector-control intentions performed in VCOM. In this analysis various mosquito-control interventions are applied at different times and coverage levels to observe the impact they have on epidemiological inoculation rate (*EIR*), basic reproductive number ( $R_0$ ) and vectorial capacity (*VC*).

### 3.2.4 Differential Equations Remarks

Despite this not being a thorough description of all the work done in mosquito-borne differential equations modelling field, it can be observed through the reviews that most of the work performed up to recent years has been lacking in incorporating important aspects such as human movement, spatial distribution, heterogeneous biting and aquatic life stages. This is something ABM systems can incorporate naturally and so, in the next section of this state of the art review, we will describe some of the most relevant work done in artificial-life modelling of these epidemic processes.

## 3.3 ABMs

Computational agent-based modelling is becoming more widespread in biological sciences and the epidemiological world is no exception to this tendency. Many frameworks with different objectives and characteristics have been proposed for the study of mosquito-borne diseases, and we will briefly describe some of the most relevant ones for our research project.

### 3.3.1 AGiLESim

One of the attempts to create a coherent framework for mosquito modelling (specifically, *Culex* and *Anopheles*) comes from Gentile et al. [53–55]. This framework was implemented in both *Java* and *C++* (UML classes diagrams are provided); and is aimed at providing a platform for mosquitoes simulations that allow parallelisation of processes. This is an extremely detailed simulation that includes aquatic stages, temperature effects, blood-meal seeking, blood-meal digesting and some vector-control interventions. In their publications, the authors provide an analysis of the mosquito population dynamics as effects of temperature and initial population sizes; in which they show the population stability of their model and the increases in population sizes as an effect of rises in temperature.

Despite the fact that this model is thorough on its description on mosquito biology, the main drawback of this work (in terms of the proposed contributions of this thesis) is the fact that humans are not considered as explicit agents but only through biting events. This makes it difficult to understand the impact of human and mosquito movement in the transmission of vector-borne diseases. Another important drawback, for scientific replication purposes, is that, to the best of our knowledge, the simulation can not be freely downloaded from a public repository. The final downside is that no comparison is provided with real-world data or other model so the conclusions that can be obtained from this simulation might be limited.

### 3.3.2 LAIS

Isidoro et al. [56, 57] created the multithreaded *LAIS* model to create small scale scenarios in which *Aedes* mosquitoes can be simulated. One of the interesting things in this framework is that it includes a version of the diffusion and degradation processes of pheromones, body heat and humidity. The main objective of this model is to provide a tool to analyse the release of RIDL mosquitoes of which they provide examples in their publications.

While it is not an absolute requirement, a graphical user interface is somewhat desired in this kind of simulation and, unfortunately, to the best of our knowledge, *LAIS* does not have one (which makes it somewhat difficult to approach for people in health sciences). Another downside is it that it develops in a discrete and relatively coarse cell environment (100 by 100) in which agents interact. Furthermore, it seems that the code is not published or documented in a public repository, which makes it difficult to replicate the results obtained in it.

### 3.3.3 Jacintho

Jacintho et al. created a simulation in the Swarm modelling framework that is aimed at integrating part of human dynamics within the environment of mosquitoes populations [58]. This simulation integrates basic behaviours of mosquitoes biology along with human and human exterminator agents. It integrates seasonality and compares two scenarios (winter and summer) within the same simulation. It also models all life stages of the mosquitoes along with males and females which is very important to be able to incorporate various vector-control interventions which deal with mating behaviours. The authors provide validation of the model at an epidemiological level as their simulation includes pathogens transmission.

Despite the thoroughness of this simulation, its source code is not readily available, so extending the model to accommodate more realistic human behaviour and mosquito bites tracking is, to the best of our knowledge, not possible. On top of this, the Swarm framework is not regularly updated or documented which makes it difficult to learn and use.

### 3.3.4 Deng

Deng et al. present a model that uses a more complex combination of natural phenomena simulations to represent the behaviour of mosquitoes [59]. In this model, a fitness function was created, which controls the behaviour of mosquito agents. This function takes into account factors such as wind, availability of oviposition sites, human attraction and roughness of the terrain to allow the mosquitoes to assess which part of the environment is more convenient for them to move to. This is a good approximation of how living beings tend to make decisions although unfortunately it requires a large amount of computational power and resorts to the use of super-agents (or cohorts) to group individuals born on the same day.

Even if the level of resolution of this simulation allows the exploration of how factors such as wind might affect the transmission of diseases, this extreme detail has a high computational cost that makes it unable to simulate detailed spatial resolutions (it uses a 30 by 30 spatial grid) and uses a cohort model agent-based model (not an individuals agent based model).

### 3.3.5 CIMSIM/DENSIM & Skeeterbuster

While not strictly an ABM, the most prominent Dengue model is the dynamic tables simulation CIMSIM (coupled with DENSIM) [60]; along with its improved version Skeeterbuster [61]. This simulation is aimed at providing a very biologically accurate simulation of the *Ae. aegypti* population dynamics. In Skeeterbuster the research group even performed a parametrisation with data from Iquitos, Peru; and included options to simulate genetic parameters to calculate dynamics produced by different control measures including a grid in which “mosquitoes” can move and reproduce. This simulation is available online and provides an accessible interface.

As good as this simulation is it not a full-fledged individual-based model. It follows mosquitoes through life stages in cohorts which limits the amount of detail that can be obtained from its use. Furthermore, humans are not modelled explicitly so studying heterogeneous biting and human movement patterns is difficult using this framework.

### 3.3.6 MASH

This model’s development is led by David L. Smith (University of Washington) as a collaboration between researchers at the Institute for Health Metrics and Evaluation, University of California Berkeley, University of California San Francisco and Tecnológico de Monterrey (amongst many others). In it, mosquitoes develop over their full life-cycle in a biological adaptation of the Gillespie algorithm [62]. This allows the simulation in semi-continuous time while still keeping computational costs to a minimum. Sites are considered as discrete “bins” distributed along space and mosquitoes move from site to site through movement probability kernels. This simulation is extremely detailed and provides a framework to test a large number of heterogeneity variables and control interventions (an example is shown in figure 3.2) but is still under heavy development and is not yet available. As this is an ongoing project, it will be described in more detail in appendix D.

### 3.3.7 ABM Remarks

Amongst these models, there are some common issues that can be identified and addressed to make a contribution in the field. Probably the most recurrent one was the lack of availability of the source code (along with documentation) to extend these models so that they can accommodate scenarios beyond their original scope. The second area of opportunity is the inclusion of

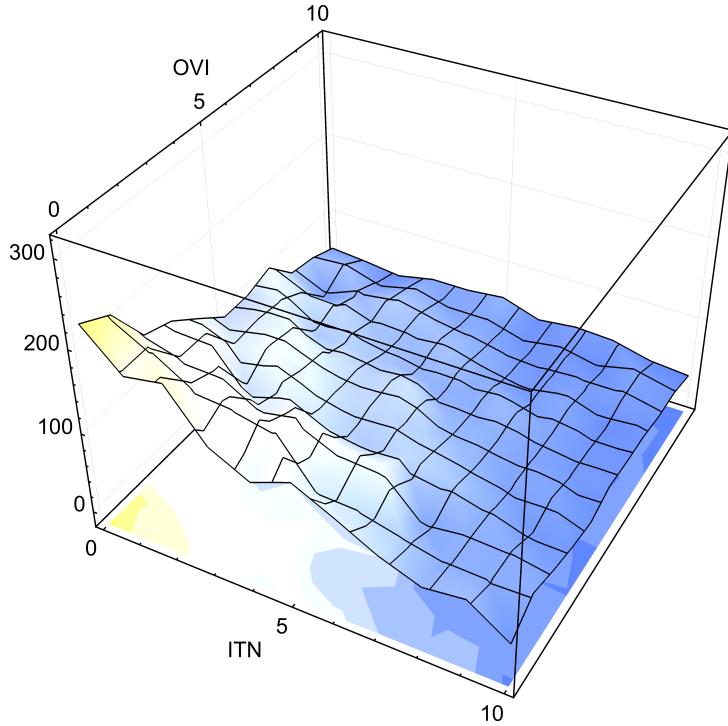


FIGURE 3.2: Example of a response surface analysis of the efficacy of a combination of insecticide-treated nets (ITN) and ovitraps (OVI) as performed on MASH. The  $x$  and  $y$  axis of each frame represent the level of coverage in a simulated environment (from 0 to 10 where 10 represents full coverage) while the  $z$  axis shows the number of mosquito adults present in the environment. This type of analysis would provide insight into integrated-vector management efficacy (as mentioned in chapter 2.3.3).

humans as part of the model, which few models do. Another aspect that is somewhat recurrent in these studies is that modelling agents with extreme detail might limit the usability of the simulations due to the high computational cost that running them entails so it is important to keep computational costs to a minimum while designing new agent-based models.

### 3.4 Network Theory

In this work, we will develop a model that allows us to analyse how human movement can affect vector-borne disease transmission dynamics through the use of network theory. This is because previous work has found that the way the contact networks emerge affects the reach of epidemic processes. An example of this shown in figure 3.3; in which, for demonstration purposes, we performed a simulation of the same pathogen spread upon three different network structures. We can see that the way the humans (nodes) contact each other (edges) affects the transmission dynamics of the epidemic. By understanding the way these relations emerge we can attempt to stop them more effectively.

To relate the reader to the work that has been done in this regard, in the following section we will describe some of the most relevant works in this field, along with some of the work that has been performed that could greatly benefit from the use of graph theory.

### **3.4.1 Epidemics Dynamics in Networks Topologies**

The use of network theory has been explored before to observe the different properties that epidemics would have in different structures for direct-contact diseases [44–48] such as SARS [63] and influenza [64].

Throughout these works much emphasis is placed on network topologies and how variables such as mean path length between individuals, clustering, small-worldness [65], density, etcetera; affect epidemiological parameters such as the basic reproductive number:  $R_0$  (defined as the number of infections an individual causes when introduced in a totally susceptible population) and prevalence.

Of particular interest are the so called “small-world” [65] and scale-free [66] networks topologies. The reason for this is that they scale logarithmically in path length when adding individuals to the network (which would keep humans epidemiologically close to one another) and that, in the case of scale-free networks, the  $R_0$  diverges [67]; which theoretically implies that elimination through vaccination is virtually impossible for these network structures (at least in the limit of infinite population sizes [46]).

Despite of these thorough explorations for direct-contact diseases, the study of the structure of human connections through mosquito bites is difficult to perform in the field and this kind of analysis has been largely neglected for mosquito-borne pathogens.

### **3.4.2 Epidemic Prediction Through Social Networks Information**

Our research group at Tecnológico de Monterrey recently obtained one of the Google Awards for Latin-America 2016 for the development of a platform capable of simulating and predicting Zika epidemics through the use of social networks information (such as Twitter). This work will make thorough use of network theory to simulate epidemic processes from internet contacts between people (to test if some correlation to real-world contacts and virtual contacts exists).

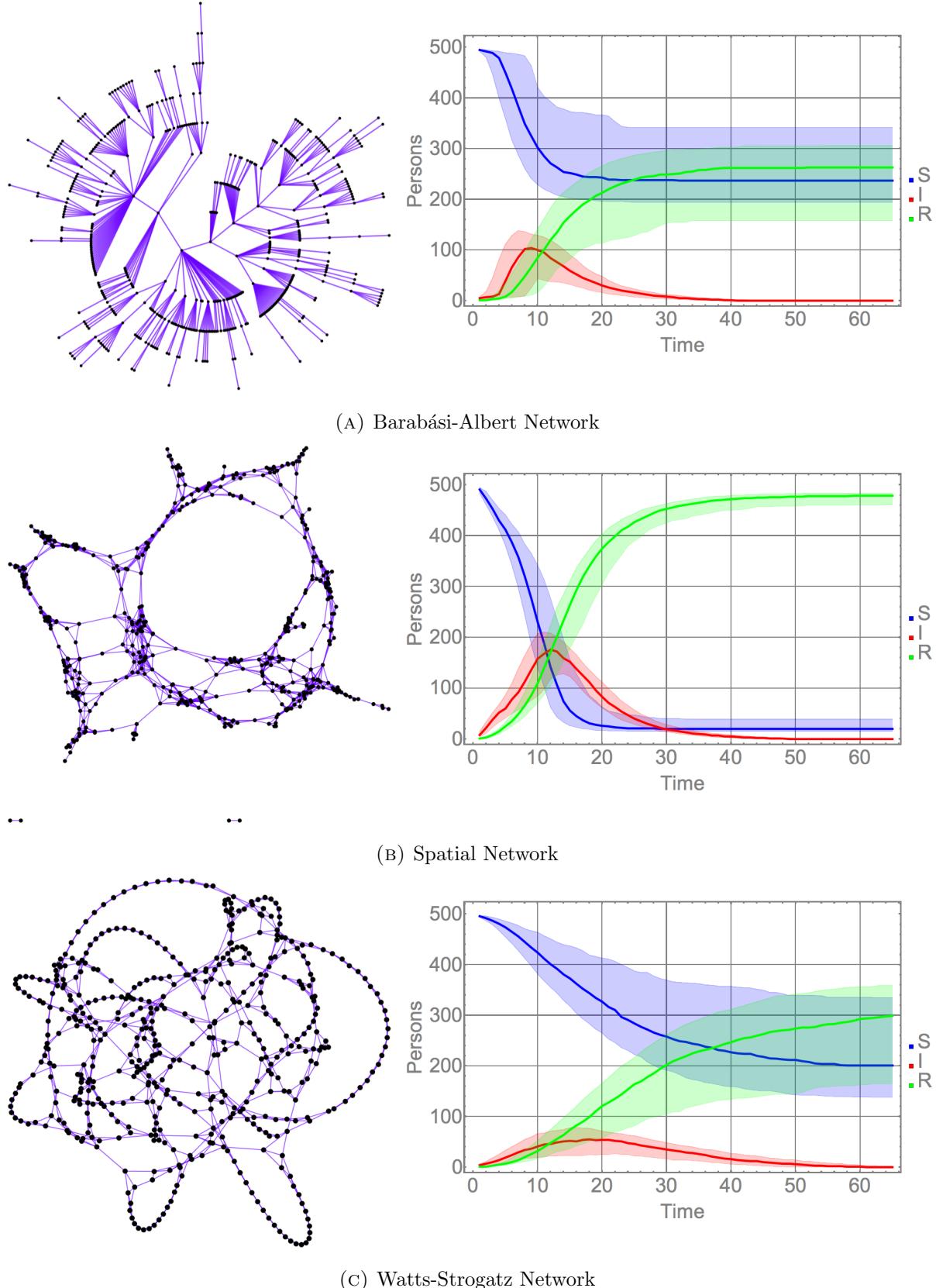


FIGURE 3.3: ZikViSoN Epidemics on different network topologies. In these examples we can see that changing the structure of the networks affects the dynamics of how diseases spread upon them. Figures on the left represent the networks structures in which each node is a person and an edge is a contact between them. Figures on the right show the way a simulated epidemic spread in the population through the count of: susceptible (blue), infected (red) and recovered (green), individuals. The shaded areas show the quartiles of the 500 repetitions of the stochastic simulation in which all the parameters were kept constant besides from the network topology.

This is still an ongoing project in its early stages that will benefit and build upon the findings of this thesis. A more detailed description of it can be found on appendix C.

### **3.4.3 Network Reconstruction**

Bryan Greenhouse's lab at the University of California San Francisco, is currently working on reconstructing malaria contagion networks from pathogen's genetic data obtained in Africa [68]. This poses significant challenges since mutations, missing links and noise in the samples can cause the results to vary wildly; but obtaining these networks would give the scientific community to understand an unprecedented level of detail on how these pathogens disperse in human settlements and how to make efficient use of resources to stop them.

This work is currently in progress and it is important for our project as it highlights the interest in this kind of analysis, but this kind of study could provide feedback to works like the one that will be described throughout this thesis and vice-versa.

### **3.4.4 Human Movement**

Human movement has been acknowledged as a crucial part of vector-borne diseases dynamics. This is partly because mosquitoes have a finite flight range and; even the species that have more exploratory tendencies, long-range transmission is more probably due to human travelling patterns than mosquito migratory behaviours.

To address these phenomena, several studies have been performed both in the field and through modelling. We will briefly describe some of their most relevant findings in the context of this thesis project.

- Marshal et al. [37] identified women travelling with kids to be an important factor in long-range malaria transmission in sub Saharan Africa, as children travelling with them usually have high parasite prevalence. They also highlight the movement of youth workers coinciding with seasonal rain as having a high probability of spreading this disease.
- Adams et al. [69] use a patches model to show how infection incidence might be independent of the duration of contact and how increased human movement might increase the “rescue effect” with which pathogens can persist in populations even when strong mosquito-control campaigns are underway. It also highlights the importance of movement

hubs and reservoirs can help spread the pathogens (in the case of malaria, transmission hubs have been detected in isolated work zones which allow the efficient dispersal of the pathogens in populations due to workers movement).

- Reiner et al. [41] perform a study analysing the importance of social structure and overlap of mosquito movement. One of their conclusions highlights the importance of the formation of communities of humans; suggesting that this social information could be used for optimisation of indoor-residual spraying techniques.
- Stoddard [38] creates a model and uses it to find that the impact of human movement in diseases spread depends greatly on its vector's behaviour. One thing to highlight is that the authors explicitly suggest networks analysis for the study of heterogeneity could enhance the results of these kind of study; and they propose a model to analyse the differences in the basic reproductive number ( $R_0$ ) in the presence of human movement and behavioural patterns.

It is examples like the previous ones, which motivate the study of the impact of these human movement either through field data or models and all of these analysis can be mapped to graphs so that more analysis can be performed on them using networks algorithms and measures such as mean path length (related to epidemiological distance between individuals), connectivity (used to measure epidemiological closeness) and clustering (that helps detect population pockets where infections could spread easily); amongst many others.

### 3.4.5 Network Theory Remarks

The epidemiology community has a general interest on understanding how heterogeneous behaviours affect disease transmission. In this regard, network theory is on the forefront of this kind of analysis; and, as this is still a relatively new and hot topic in epidemiology, new ideas and approaches are under development. This gives us an opportunity to make contributions in the field both through the use of field-data analysis and, as it will be shown throughout this work, through the use of artificial-life models.

## 3.5 Conclusions

In this chapter we have described the current state of the modelling of the transmission of *Aedes*-borne diseases. In it, we have seen that despite the extensive work done so far by the community,

there are still things to be done to further our understanding of the dynamics of vector-borne diseases. In particular, we have observed that even if ABMs have been created to alleviate some of the ODE models inherent limitations, most of them have not had widespread impact in their use (which can be partly attributed to their high computational costs, lack of documentation and availability). Finally, we have described the general interest of the epidemiology community into using tools borrowed from network theory to analyse in further detail the intricacies of heterogeneous contacts between humans and the impact of human movement in the spread of mosquito-borne diseases.

Throughout this thesis project, we will aim towards creating a well documented, open-source agent-based framework in which spatial heterogeneity questions can be answered. Towards this end, we will begin the description of our model in the following chapter by stating how the biological behaviours described in chapter 2.1 were translated into a computational framework.

## Chapter 4

# Model Description

*“I like the scientific spirit — the holding off, the being sure but not too sure, the willingness to surrender ideas when the evidence is against them: this is ultimately fine — it always keeps the way beyond open — always gives life, thought, affection, the whole man, a chance to try over again after a wrong guess.”*

*Walt Whitman*

All models are, in essence, limited versions of reality. This is why modelling is a difficult endeavour. We have to decide which factors are the most relevant for the desired tasks at hand; while trying to represent a phenomenon we make compromises on which variables and behaviours must be included for the study of a particular area of interest. For this reason, if we want a model to be useful we have to provide not only its validation and results; but a thorough description of how it works and the questions it is meant to answer.

With this in mind, we will describe in detail all of *SoNA3BS*’s components (environment, agents and mosquito-control interventions). To do this, we will adhere as much as reasonably possible to the “Overview, Design Concepts and Details” protocol suggested by Grimm et al. [70] as a template to describe ABMs in extreme detail. We will, however, allow ourselves some flexibility, as the protocol requires the thorough description of each variable of the agents, which would greatly increase the length of this document (although this information is available both in the code itself and on *SoNA3BS*’s website: <http://chipdelmal.github.io/SoNA3BS/index.html>).

We will begin this chapter with the definition of what will be considered an agent in the context of the project. We will then state the languages in which the simulation and analysis routines were programmed. Afterwards, we will describe the general environment in which the agents interact (along with its topology, weather and size). After doing so, we describe the mosquito agents in their own section (as they involve several sub-models); and then we will move on to the remaining agents and objects: humans, houses, work zones, sugar sources, egg laying and mating sites. Then we will define the way the *Aedes aegypti* control interventions were implemented in the simulation and finally we will describe the computational speed of the model.

## Agents' Definition

The literature of artificial life is full of definitions of what a computational agent is, and the characteristics that are expected from it. These definitions can usually be clustered in two main categories: the ones that focus on artificial intelligence (and problem solving), and the ones that are created to simulate tangible, real-world phenomena. This stems from the fact that, in general, these are usually regarded as the main areas of study of these kinds of models; as summarised by Taylor: “*In Artificial Life there is a major intellectual divide, similar to the one in the field of Artificial Intelligence, between “engineered” systems designed to accomplish some complex task by any means the designer can devise, even if only distantly related to the way natural systems accomplish it, and systems meant to accurately model biological systems and intended for testing biological hypotheses*” [71]. Our model is clearly placed in this second area of interest, as our main goal is to reproduce a biological phenomenon in a computational environment so that we can make conclusions about the real-world version of the system. This is why, for the purposes of this project, we will use Shiflet’s definition of an agent-based simulation: “(*in an individuals-based system*)... each entity is modelled as an autonomous, decision-making agent that has a state, which is represented by a set of state variables, or attribute values, and behaviours, which control its actions. A method or procedure, which is associated with a class, or breed or group, of agents, is a function that captures some or all of an agent’s behaviour. A simulation frequently includes several global simulation variables, which all agents can access. Agents often operate in an environment that arranges cells in a rectangular grid” [72].

In addition to this definition, there are some characteristics that are often referred to in the computational agent-based modelling community when defining the expected level of detail and complexity of the simulated entities (some of them being: architecture, autonomy, proactivity,

adaptiveness and social interactions [73]). Our agents were coded with a reactive architecture in mind (summarised in figure 4.1). According to this scheme, an agent perceives its environment, it then verifies its internal states and modifies them (if required) so that it can select and perform an action that has some repercussions in its surroundings. In this regard, we can conclude that our agents' level of autonomy is limited (as its actions greatly depend on the environment) and that it does not have a rational purpose, but merely performs actions as responses to needs and changes in its environment. Cooperation is limited to emergent behaviours and agents have no sense of memory or forecast. The reasons behind these decisions will be described when dealing with the descriptions of each of the actors of our model.

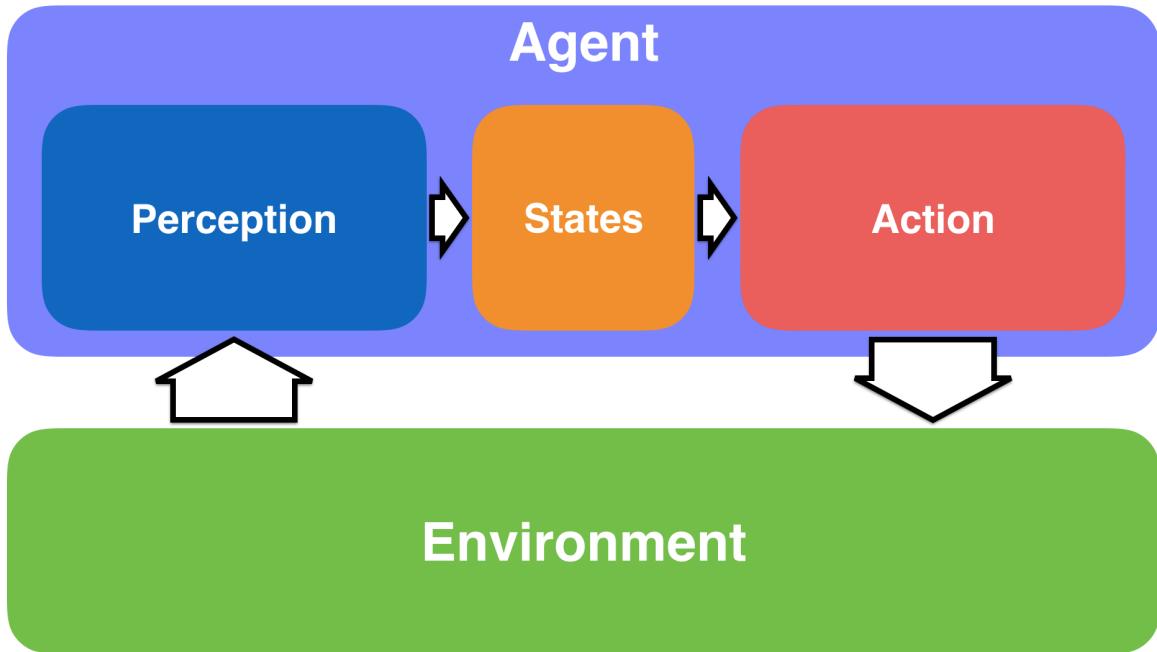


FIGURE 4.1: Activity diagram of a reactive agent's architecture. In this cycle we can observe the agent perceives its environment, modifies and reacts according to its internal states and then performs an action that might influence its surroundings.

As a final note, we should clarify that we will make the distinction between agents and objects by defining an agent as having some level of autonomy and decision-making capabilities (limited as they might be). Objects, on the other hand do not present these characteristics and, therefore, are only considered to be passive elements that exist for agents to perform their assigned tasks and behaviours.

## 4.1 Programming Languages Used

Several hundreds of programming languages exist, and arguments can be made for each one to be used in any number of applications, including agent-based modelling and data analysis. In the case of this thesis project we took into account that, despite it being developed by computer scientists, the model is of interest for people in life sciences fields. With this in mind, before selecting a development platform, we briefly reviewed several agent-based modelling frameworks such as: *Swarm*, *MASON*, *NetLogo* and *Jade*; along with scientific analysis programming languages and software: *Python*, *R*, *Mathematica*, *Stata* and *Matlab*.

After a lot of consideration and tests we decided on using *NetLogo* as the platform upon which we would develop our ABM and *Mathematica* as the language used for parameter setup and data analysis. The reasons behind these selections will now be described.

**NetLogo** One of the most often cited problems of ABMs is that they lack clarity and portability. It is easy to be tempted to create an agent-based framework from scratch in a mainstream language such as *C* or *C++*. This, however, might not be the best approach in all instances as suggested by Grimm and Railsback [74, 75]. While creating a framework from scratch provides more flexibility; it, more often than not, makes it more difficult for new users to use and extend the current capabilities of the simulation. With this in mind, using an existing framework for agent-based modelling is a better option to provide a common language on top of which other research groups can build upon. Railsback et al. [76] provide an analysis of several ABM frameworks in which *NetLogo* is highlighted because it gives us a good balance between documentation, portability and usability. This is important for our work because we ultimately want it to be available and used by the scientific community; including people who do not have an extensive computer programming background.

One final advantage of using *NetLogo* is that it provides the means to run parallel iterations experimental scenarios “out of the box”. This means that even though one single iteration of the simulation can’t be parallelised, several iterations can be run on different, independent threads without having to modify the simulation’s code.

**Mathematica** For the parametrisation of the ABM and analysis of results we used *Mathematica*. This decision was partially based on previous work we had done on networks analysis

on this platform (as part of an animal language project developed in partnership with UCLA which is described on appendix B). This, however, was not the only reason behind our selection.

*Mathematica* is extensively documented with many examples of use. It runs on different operating systems and has several symbolic calculus capabilities (of which we make use throughout the parametrisation and analysis routines). Despite this, parametrisation and analysis can be performed in other software if needed (such as *R* with *iGraph* for network analysis). The parameters needed for the simulation to run are ultimately exported to plain text files so they can be generated on any other computer programming language; and the results of the simulations are exported in *XML* files so they can also be analysed in a variety of different languages and software platforms.

## 4.2 Environment

Before describing our agents, we will describe the “arena” in which all the behavioural rules are laid out and interactions take place (an example shown in figure 4.7). This environment involves not only the spatial setting of the simulation but also the weather, scheduling, measuring devices, etcetera; all of which will be defined in the following section.

### 4.2.1 Boundaries

The “world’s” boundary was defined as non-toroidal and non-absorbent by default. This is important for our model because we are ultimately interested in understanding how contact-communities of humans emerge through mosquito bites. If we allowed agents to move through the environment’s limits in a toroidal way we would create false spatial relations between humans near the edges of the landscape; while allowing mosquitoes to be “absorbed” by the landscape’s border would require the consideration of migration patterns (which is something we are excluding from our model for the time being). In our current implementation, whenever a mosquito attempts to move outside of the world’s boundaries another direction is randomly sampled and the mosquito moves there if it is a valid one within its movement range.

In the case of the experiments purely concerning population dynamics, however, the boundary can be safely set to a toroidal configuration. This is because, if we were not really interested in human connections, the fact that mosquitoes move from one end to the other does not affect

noticeably the dynamics of the population as a whole (in chapter 6 we will show that in small-scale scenarios, such as the ones we will be testing throughout this work, spatial relations do modify the population dynamics in a significant manner).

### 4.2.2 Temperature

Environment's temperature influences greatly the developmental rates of mosquitoes (as it was described in chapter 2.1). In our simulation, we include the mosquitoes' metabolic response to temperature (the way mosquito agents incorporate this response will be defined in chapter 4.3) and to take advantage of this, we included three basic weather temperature models that can be modified and fitted to real-world data. These three models are: constant temperature throughout the year (equation 4.1a), sinusoidal yearly variation of temperature (equation 4.1b) and sinusoidal daily and yearly variation of temperature (equation 4.1c).

$$(a) \text{temperature}_{\text{constant}} = A$$

$$(b) \text{temperature}_{\text{yearly}}(\text{day}) = A + B \cos \left( C + \frac{2\pi * \text{day}}{365.25} \right)$$

$$(c) \text{temperature}_{\text{yearly+daily}}(\text{day}) = A + B \cos \left( C + \frac{2\pi * \text{day}}{365.25} \right) + D * \cos(E + F * \text{day})$$

EQUATION 4.1: Prototype equations to fit the yearly temperature variation in a simplified year way and a daily variation of temperature. Each parameter in the equations is a constant that can be adjusted to data.

In the current version of our model we do not include rainfall dynamics nor mosquito deaths as an effect of sun exposure.

### 4.2.3 Size

In the experiments presented in this thesis, the simulated world encompasses a total area of 1000 by 1500 meters, in which each discrete block (or "patch") is equivalent to ten meters. Despite patches being discrete, the agents can move in fractional quantities so distances between individuals and movements are calculated not in discrete but in continuous amounts.

The simulation allows the modification of these sizes parameters and provides with routines that make it possible to automatically rescale both the world's size and the agent's variables that are related to spatial phenomena.

#### 4.2.4 Observers

One advantage of using an ABM over other kinds of models is that we can trace the history and behaviour of individuals can be followed in real time and exported for further analysis if needed. With this in mind we included routines to saved data to XML files. This process occurs in two phases: a real-time one and an offline one.

##### 4.2.4.1 Real-Time

The real-time exporting routine is designed to record variables such as population sizes changes throughout time, average lifespans, number of ovipositions, density-dependent death probabilities, etcetera. The variables to be saved can be changed easily by adding and removing lines of code in the corresponding file. This allows flexibility to the kind of analysis that can be performed with the model and these variables are easy to analyse in languages such as *R*, *Mathematica* or *Python*.

##### 4.2.4.2 Offline

The offline exporting routine deals with information such as the biting transitions lists, interventions application; and contacts between humans and sites. This information is exported at the end of each run in the “tail” of the XML files.

#### 4.2.5 Simulation Cycle Description

Each “tick” of the simulation comprises a number of actions:

1. Agents are randomly assigned a priority to act
2. Each agent performs the actions that correspond to its defined behaviour
3. If the environment needs to be modified it is updated
4. The agent is “popped” from the queue and another agent performs its actions

This process ensures that no agent gets precedence of action over the other ones which would cause biased results. It is handled automatically by *NetLogo* and requires no modification by the programmer nor the user.

In our simulation we used a tick scale of  $5\text{min}/\text{tick}$ , although we allow a variable tick-size resolution and parametrisation routines adjust all the required parameters so that the simulation behaves correctly even whilst changing the value of this parameter.

## 4.3 Mosquito Agents Description

Mosquito agents are, by far, the most complicated individuals in the simulation. This is because, given that at this stage of our research we are interested on how different variables affect their population sizes and potential to transmit diseases, we wanted to add as much realism as computationally possible. As such, we devoted much of our modelling efforts to getting an accurate description of their behaviours and characteristics so that we could make inferences on how these individuals would respond to different stimuli.

In this section of our work, we will describe thoroughly the way we mapped the behaviours of these individuals into a computational space. We will begin by defining how we modelled their aquatic stages (egg, larva and pupa), then move to define their behaviours in adult phases, and finally describe the dependence of their metabolic developments with temperature.

### 4.3.1 Aquatic

The first three life stages of the mosquito life cycle are spent in aquatic habitats (as described in chapter 2.1). The way each of these phases were implemented in our system will be described in the following section of this document.

#### 4.3.1.1 Egg

Eggs are passive agents that spend their time developing into pupae. In this stage, mosquitoes undergo two probabilistic death processes: a constant mortality rate and a density-dependent hatching inhibition.

**Constant Egg Mortality Rate** In our model, this mortality rate is modelled by a binomial event that occurs at every tick for every egg with a probability of:  $e_{kdP} = 0.01(1/\text{day})$  [77]. With this parameter we encompass a wide range of phenomena that can destroy eggs such as: desiccation and failure to develop into viable larvae.

**Density-Dependent Hatching Inhibition** The first of the density-dependent processes that take place on the aquatic stages of the mosquito life cycle is the egg hatching inhibition. If the aquatic habitat is too densely populated by larvae; eggs have a lower probability to hatch and might enter a diapause stage [51, 78] (or they might even be eaten by hungry larvae that need nourishment [23]).

Adding a diapause effect on eggs in our ABM was contemplated but discarded because of two reasons. On one hand it has been previously found by other modelling work that it does not affect greatly the population dynamics as a whole in temperate climate situations [51]; and, on the other hand, it added a very high computational overhead to the model (as the number of eggs laid by each female is extremely high). Instead of getting eggs into diapause stage we decided to add a linear mortality effect as an approximation (its validity and working range will be discussed further in chapter 5.2). To create this function we defined the linear prototype shown in the equation 4.2.

$$e_{DdP'}(l) = EK * l$$

EQUATION 4.2: Egg hatching inhibition for one larval site ( $e_{DdP'}$ ) as a function of the number of larvae ( $l$ ) and a scaling factor ( $EK$ ) that controls the number of eggs that can be sustained in the environment.

Ideally speaking, each breeding site would have an independent carrying capacity corresponding to its width, height and sun exposure. As a first approximation, however, we defined a standardised container shape to account for all these variables and pooled them all together to maintain low computational costs (this is a fairly common assumption that is used in other ABMs [23]). This behaviour is modelled by the equation 4.3, which was created to maintain a low number of eggs as the environment scaled up.

$$EK(bs) = \frac{1}{2.5*bs+2}$$

EQUATION 4.3: Scaling behaviour of the  $EK$  value as a function of the number of breeding sites ( $bs$ ). The constants of the equation and its shape were defined to allow at least 30 eggs to survive each day (simulating a small water container [51]).

Combining these equations results in equation 4.4 and the response shown in figure 4.2(A); where the death probability of eggs is represented as a function to the number of breeding sites and available larvae in the environment.

It should be noted that the total number of eggs in the environment is not being modelled precisely to match the real biological phenomena, as scaling this behaviour with explicit egg

$$e_{DdP}(l, bs) = \frac{l}{2.5*bs+2}$$

EQUATION 4.4: Egg death probability ( $e_{DdP}$ ) as a function of the larvae quantity ( $l$ ) and the number of breeding sites ( $bs$ ). This equation was derived from the combination of equations 4.2 and 4.3.

agents is computationally unfeasible (at least in most “general-purpose” computer systems). This, however, should not pose a problem in the analysis of adult populations dynamics. The reason behind this is that a second density-dependent process takes place at the larval stage which, in turn, presents the most drastic reduction in the survival of mosquitoes and acts as a population “funnel” that limits the number of mosquitoes supported by the environment (this will be further described and discussed on chapter 5.3; where we will present the results from our validation experiments).

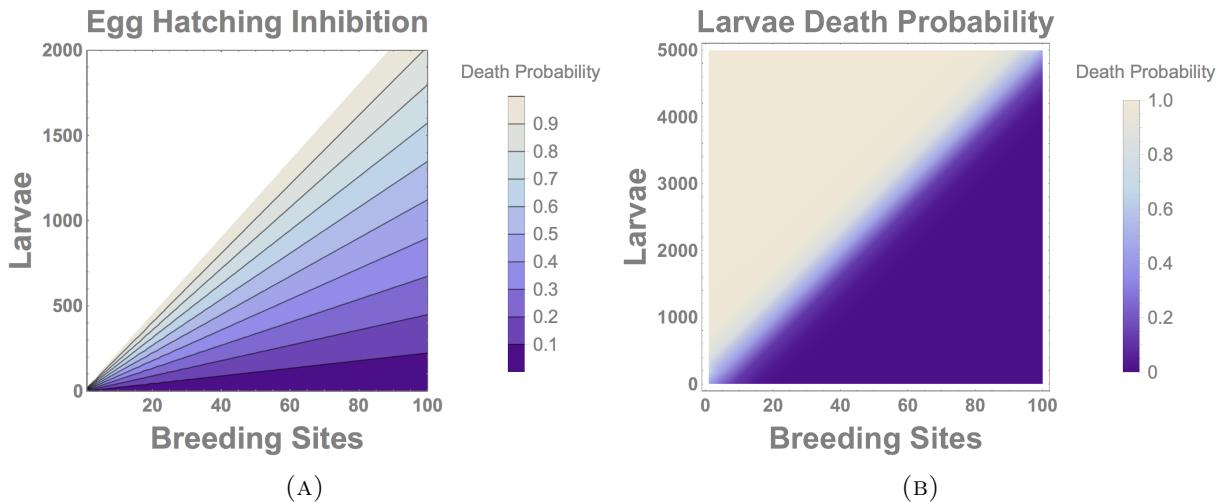


FIGURE 4.2: Density-dependent death probabilities of aquatic mosquitoes. (A) Scaling behaviour of egg hatching inhibition as a function of the breeding sites and larvae number. (B) Scaling behaviour of larvae death probability as a function of the breeding sites and larvae number.

#### 4.3.1.2 Larva

Biologically speaking, larvae individuals are more active than eggs as they search for nutrients, move to escape from predators and to avoid direct sunlight. This, however, is a level of detail not relevant for the scales treated in our ABM; and can be simplified by making larvae spatially static while using equations that simulate their approximated mortality rates. To this end, we included two mortality processes to simulate what happens in real life: a temperature dependent mortality probability and a density-dependent death rate.

**Temperature-Dependent Larva Mortality Rate** The larva population undergoes some variation in their death probabilities as the environment's temperature drops. This death probability is represented by the equation 4.5 [79–81].

$$l_{TdP}(T) = 0.9725e^{-0.369891(T-278)}$$

EQUATION 4.5: Larval death probability ( $l_{TdP}$ ) dependent of the environment's temperature ( $T$ ).

**Density-Dependent Death Rates** In a process similar to the one that occurs to eggs, as the number of larvae increases the fight for resources shifts the survival probabilities down [82]. We modelled this process with a sigmoidal function as an initial, empirical approximation to represent the real biological phenomena (as these curves can be flexible enough to accommodate hard, step-wise thresholds; to leaner, almost linear, responses). This function is shown in equation 4.6.

$$l_{DdP'}(l) = \frac{1}{e^{a-LK*l}+1}$$

EQUATION 4.6: Larval density death probability in one breeding site ( $l_{DdP'}$ ). The parameter  $LK$  was defined to scale linearly with the number of breeding sites in the environment as shown in equation 4.7. For one larval site the inflection point was set to 50 larvae ( $l$ ), which corresponds to a value of  $a = 5$

Once again, we wanted our model to be able to scale dynamically according to the number of larvae and breeding sites. Towards this end, we defined the larval death probability ( $LK$ ) value to scale linearly with the number of breeding sites as shown in equation 4.7.

$$LK(bs) = 0.25 * bs - 0.035$$

EQUATION 4.7: Value of the larval death probability ( $LK$ ) from the equation 4.6 as scaled with the number of breeding sites ( $bs$ ).

Combining equations 4.6 and 4.7 resulted in equation 4.8 and the response shown in figure 4.2(B). As it was the case with the density dependence equation of the egg stage, the validity and scope of this equation will also be discussed in chapter 5.3.

$$l_{DdP}(l, bs) = \frac{1}{e^{0.25*bs-0.005*l-0.035}+1}$$

EQUATION 4.8: Larval density death probability final function ( $l_{DdP}$ ). This equation is obtained by combining equations 4.6 and 4.7.

### 4.3.1.3 Pupa

Pupae individuals are, once again spatially static, passive agents. These individuals undergo two death processes: a temperature-dependent mortality probability and a failed emergence death probability.

**Temperature-Dependent Pupa Mortality Rate** As with the larva stage pupa individuals vary their survival probabilities according to the environment temperature. This probability is shown in equation 4.9 [79–81].

$$p_{Tdp}(T) = 0.9725e^{-0.369891(T-278)}$$

EQUATION 4.9: Pupae death probability ( $p_{Tdp}$ ) dependent of the environment's temperature ( $T$ ).

**Failed Emergence Death Probability** Pupae have a certain probability to die while in the process of emergence to adulthood from pupa stage. This is modelled as a binomial event with a survival probability of  $ef = 0.83$  [83] which is evaluated just once upon emergence attempt.

### 4.3.2 Adults

Once mosquitoes emerge as adults, they are able to fly and interact with their environment. This is where most of the advantages of using an ABM come into play, as we can now incorporate movement along with the interactions between humans and mosquitoes.

Throughout this final life period, mosquitoes were programmed to fly at a speed of  $0.5(m/s)$  [19] and a simulated action range of  $1/5(m/min)$ . In this stage, they face only one constant death rate (although they might also die due to mosquito population control interventions, starvation and interactions with humans).

**Constant Adult Mortality Rate** The regular adult mortality rate is a constant parameter in our simulation  $a_{Kdp} = 0.09(1/day)$  [51, 80] that encompasses predation, wing tattering and senescence.

#### 4.3.2.1 Male and Female Life Cycle

As it was described in chapter 2.1, mosquitoes differentiate their adult life cycles depending on whether they are males or females. To describe these behaviours, we will first define all the actions that can be performed by the agents and then describe the differences between the behaviours each gender can engage on.

**Bouts Description** Each adult mosquito can be in one of several life states with behaviours attached to them. These states and their corresponding actions (“bouts”) were programmed in our simulation as follows:

- Mating (M): Both males and females that enter this stage move towards a mating spot or “landmark” to form mating swarms and find a viable individual to copulate with. If females are close enough to a male they trigger a binomial event that might result in mating attempt success or failure.
- Sugar-Feeding (S): Individuals can enter this cycle spontaneously whenever their energy level drops below a given threshold (defined as 50% of their total feeding capacity for initial tests). Once energy level drops below that threshold a mosquito searches for the closest sugar source available and attempts to feed on it. If the mosquito is unable to find sugar and its energy level drops to zero the mosquito dies.
- Blood-Feeding (F): Females in this bout look for human blood so that they can develop eggs. These individuals “bite” persons based on proximity and can be killed probabilistically by their hosts (to simulate human reactions to these pests).
- Resting (R): These individuals stay fixed in place or move randomly in small flights. They might be killed by a person if they are close enough for them to act (through a binomial event).
- Egg-Laying (L): Females enter this stage after blood feeding and resting. In this first iteration of our model, females lay all their eggs in the same aquatic habitat (we are not modelling skip-oviposition for the time being; which is a fair approximation in conditions where high detail in extinction thresholds in heterogeneous egg-laying sites is not of interest).
- Dead (D): These mosquitoes are no longer alive and as such are removed from the simulation in the next time “tick” after their death.

*Aedes aegypti* are mainly diurnal so these bouts are mostly performed during daytime. Throughout the night both males and females rest and can probabilistically engage in some other behaviour but for the purposes of this work this probability is set to a level close to zero to adhere to mainstream scientific knowledge on *Aedes* behaviour [19].

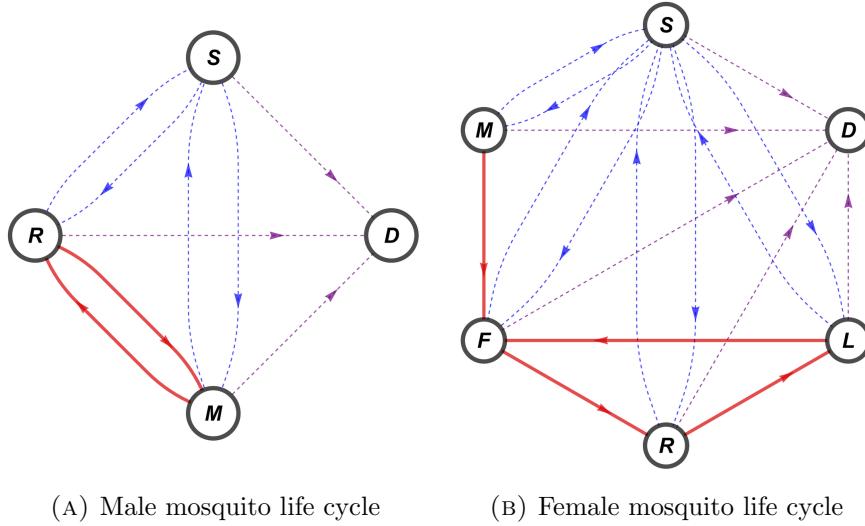


FIGURE 4.3: Male cycle (a) is comprised of repeated mating (M) and sugar feeding (S) cycles after their initial rest bout (R). Female cycle (b) involves iterations of blood-feeding (F), resting and egg laying (L) actions after they have mated. Both males and females can go into a sugar feeding bout if needed or die (D) due to biological or control driven phenomena.

We will now move on to describing in more detail the differences and specifics of male and female adults behaviours (which are summarised in figure 4.3).

**Males** Male behaviour involves three actions and a death stage (as shown in 4.3A). Upon emergence from pupa stage, these agents wait for a day and then they engage in mating routines throughout the day only interrupted to rest and sugar-feed sporadically.

Once male mosquitoes die (either because of their natural death probability or by an artificial process) they are removed from the simulation and males do not store any additional information into XML files.

**Females** Female life-cycle involves more stages and variables than the male one. When they emerge from pupa stage, these agents mate almost immediately. In doing so, they store the genetic kind of the mosquito they mated with. This information is saved in flags that determine the individual's type: RIDL, *Wolbachia*, SIT, regular; so that we can simulate the inheritance rules for these interventions (described in chapter 4.5).

Once they have mated, females look for a human to blood-feed upon; and when they are close enough, they attempt to bite. If they succeed, they store the ID of the human in their bites lists (so that the bites networks can be reconstructed after the simulation has ended) and they begin a resting bout waiting for egg maturation (called “gonotrophic cycle”). After this period has elapsed, they search for a viable breeding site, lay all their eggs and repeat the cycle starting from the human blood-feed action (highlighted in red in figure 4.3B).

Throughout these cycles, female mosquitoes sugar feed when needed and can die probabilistically either by natural processes or by human action (blue and purple dashed lines respectively on figure 4.3B). Upon death they add their bites lists information into the global bites lists of the simulation so that the human bite-contacts networks can be recreated in posterior analysis.

### 4.3.3 Schoolfield Metabolic Development

As it was mentioned earlier, the time a mosquito spends in each of its life stages is controlled by its metabolic development rates; which are correlated with the temperature of the environment the mosquito lives in. These development rates are modelled by equation 4.10 [84] and apply to each of the mosquitoes life stages, the only difference being the specific thermodynamic parameters used for the calculation of each stage’s rates (shown in table 4.4 [81]).

$$dR(T) = \frac{\frac{Ha}{R} \left( \frac{1}{298} - \frac{1}{T} \right)}{298 \left( e^{\frac{Hh}{R} \left( \frac{1}{T_{12}} - \frac{1}{T} \right)} + 1 \right)}$$

EQUATION 4.10: Development ratio ( $dR$ ) as a function of temperature ( $T$ ) given a set of thermodynamic parameters:  $RdK$  (development rate at  $25^\circ C$  assuming no enzyme inactivation),  $Ha$  (enthalpy of activation of the reaction catalysed by enzymes),  $Hh$  (change in enthalpy associated with enzymes),  $T_{12}$  (temperature at which the enzyme is half active and half low temperature inactive),  $R$  (universal gas constant).

	R	RdK	Ha	Hh	T12
Egg	1.98588	0.24	10798	100000	14184
Larva	1.98588	0.2088	26018	55990	304.6
Pupa	1.98588	0.384	14931	-472379	148
Adult G1	1.98588	0.216	15725	1756481	447.2
Adult G2	1.98588	0.372	15725	1756481	447.2

FIGURE 4.4: *Aedes aegypti* metabolic development parameters for Schoolfield’s model.

For the use these development rates we are assuming that the temperature is homogeneous over the simulated environment (which is a fair approximation for small-scale scenarios). This allows

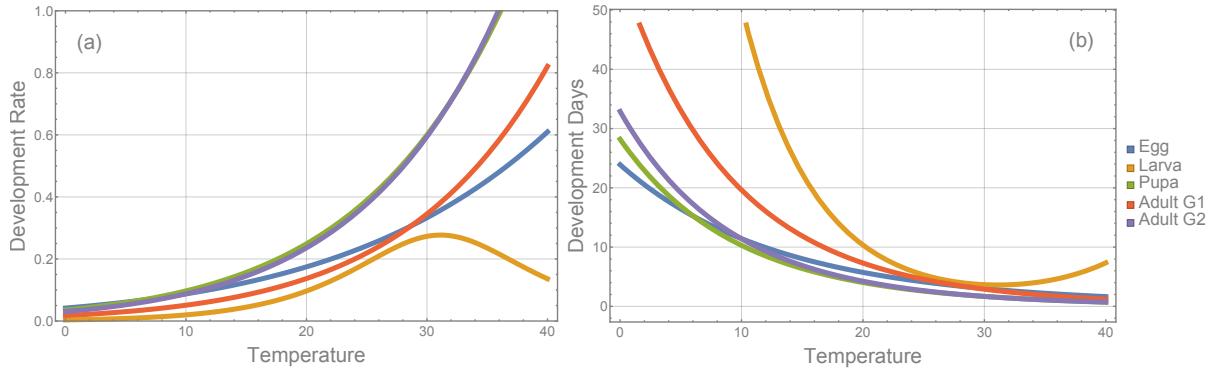


FIGURE 4.5: Mosquito metabolic development dependence on the temperature of the environment. Figure (a) shows the development rate of the mosquitoes at any given temperature while figure (b) shows the expected number of days a mosquito spends in each life stage at constant temperatures.

us to calculate a global development rate in each tick and add it to the mosquito agents to save computational power.

## 4.4 Other Agents and Objects

Despite mosquito agents being the main kind of individuals our simulation includes various other types of actors (shown in figure 4.6). In the following section we will describe: humans, houses, work zones, sugar sources, ovitraps, egg laying and mating sites.

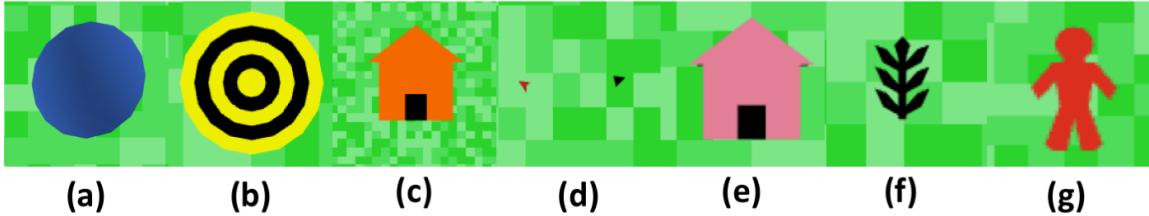


FIGURE 4.6: Agents and objects modelled in *SoNA3BS*. Egg-laying site/ovitrap(a), mating landmark (b), household (c), mosquito (d), work zone (e), sugar source (f) and human (g).

### 4.4.1 Humans

The second most important type of agent included in our simulation is the human class. Our humans were programmed with day and night activities to simulate realistic daily cycles. These agents move at a speed of  $2.5(km/h)$  in straight paths and they have three different kinds of behaviours to accommodate basic types of human behaviours:

- Homestead: This type of person stays at home throughout the day only moving in its vicinity.
- Working: This type of human goes to work by day to a designated zone in the landscape and goes back home at night.
- Visiting: These individuals make brief visits to other people's houses. The duration of these visits can be freely modified.

Two different algorithms were implemented for visiting humans to choose the house they would go to:

- Random Selection: The target house for the visit is selected in a random way through a uniform probability distribution.
- Gravity-Based: The visited house is selected by its size and proximity according to equation 4.11 [85]. The value of the gravity of each household is then normalised to add up to one so that their probabilities to be visited are scaled up with their sizes and relative distances.

$$\text{gravity}_n(\text{mass}, \text{distance}) = \text{mass}/\text{distance}$$

EQUATION 4.11: The probability of a given site “ $n$ ” to be selected from the pool of available sites is defined by its gravity, which is a function of its “ $\text{mass}$ ” (or size) and the  $\text{distance}$  to the current location of the agent.

These two visiting patterns encompass basic human behaviours in small geographical scale settings, but the simulation can be extended to accommodate other models such as radiation models [86] if necessary. In addition to this, humans can be set to react to large concentrations of mosquitoes by destroying the nearest aquatic habitat or spraying insecticide, although these reactive behaviours will not be explored in this present work.

As a final remark, direct human to human contact and human to building can also be tracked for further analysis and comparison to vectorial-contact networks.

#### 4.4.2 Landscape Objects

The landscape set comprises the environment's elements that are needed either for mosquitoes or humans to develop properly and perform their programmed actions. In this section we will

describe how we integrated them into the simulation along with the way they interact with the other individuals of the system.

It is worth noting that we do not consider these computational structures agents (in a strict sense) because they do not possess any decision-making capabilities, so we will refer to them simply as objects to highlight this distinction.

#### **4.4.2.1 Houses**

Houses are passive structures designed to accommodate persons. Each house has an identifier number to determine which humans belong to a given household, and an area that is usually used to determine the quantity of humans that live there (as well as the “attractiveness” or “mass” of the house in the case of gravity models).

#### **4.4.2.2 Work Zones**

Work Zones are similar to houses in the sense that they possess an identification number for humans to “know” to which work zone they have to attend day after day. They do not modify nor interfere with any variables of other agents.

#### **4.4.2.3 Sugar Sources**

Sugar sources exist for mosquitoes to fulfil their energetic needs. These elements simulate fruits and other artificial sugar sources that can usually be found where *Aedes aegypti* thrive. They possess an allotted amount of sugar that can be decreased each time a mosquito feeds on it and that replenishes over time (to simulate naturally occurring sugar sources). These objects can be created and destroyed dynamically; although, for the most part of this work these sugar sources were treated as static and constant in their sugar supply.

#### **4.4.2.4 Egg Laying Site**

Egg laying sites are required for female mosquitoes to oviposit. As it was discussed before, these sites were designed to have uniform carrying capacity; which is translated to each oviposition spot having an equivalent volume and level of nutrients. These sites can also be created and

destroyed dynamically while the simulation runs (this is meant to accommodate future models such as source management to reduce aquatic habitats, and increases in oviposition spots due to rainfall).

#### 4.4.2.5 Mating Site

These sites are landmarks on the environment upon which mosquitoes swarm for mating purposes. In the case of *Aedes aegypti* species, these sites are usually located in close proximity to their breeding locations so that is the way their location is selected in the setup routines.

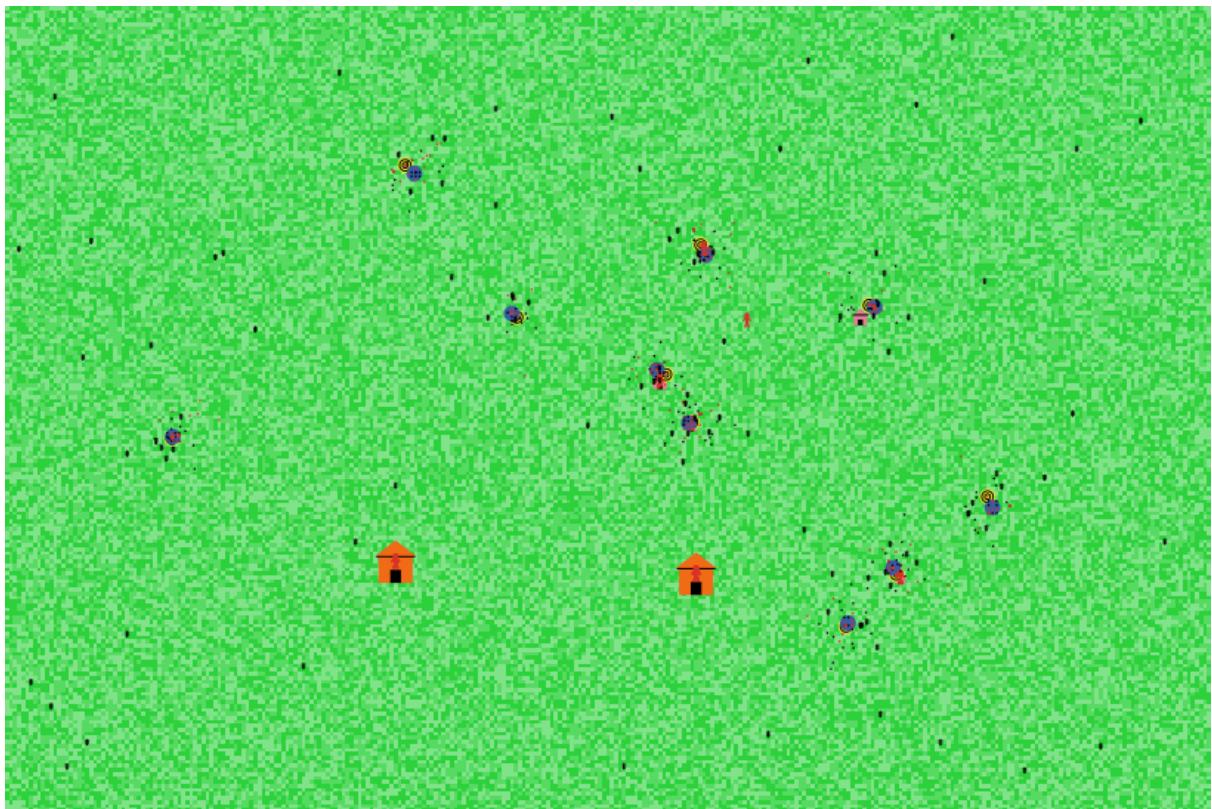


FIGURE 4.7: SoNA3BS environment example. This image shows one of the setups in which validation experiments took place (described in the chapter 6) and shows all the agents described in the current section of our document.

## 4.5 Interventions

The general way these interventions work in real-life was described in chapter 2.3. In the following section we will describe how they were mapped into our simulation.

#### 4.5.1 Fogging

This intervention is modelled as an instant-action layer that works uniformly over the environment. This insecticide has a fast exponential decay (starting with a death probability of  $6.25\%(1/tick)$  immediately after application and with a half-life of  $120min$ ) and is assumed to affect only adult mosquitoes. These parameters were defined empirically but could be fitted to any specific type of insecticide if data became available.

#### 4.5.2 Sugar Baits

Sugar baits are sugar sources that have been poisoned so that adult mosquitoes who feed upon them are killed. In our simulation this behaviour is mapped through a binomial event with a static death probability for each mosquito that feeds on a poisoned sugar source (which initially is set to 90%).

#### 4.5.3 Ovitraps

There are many types of ovitraps available; ranging from the ones designed to just kill the laid eggs to others that kill both the eggs and the female that lays them. In our simulation we modelled the latter. Females that lay eggs have some binomial probability of escaping (initially is set to 10%) but the eggs are always killed.

#### 4.5.4 Sterile Males

Sterile males releases work by saturating the environment with males that can mate but which's genetic material is unviable. In *SoNA3BS*, we include this intervention by keeping track of the kind of male with which a female mates. If a female mates with a male marked as sterile she goes through all her regular oviposition cycles but lays no eggs while attempting to do so.

#### 4.5.5 *Wolbachia*

This intervention involves mating and the transmission of a specific pathogen. For this to be modelled properly female mosquitoes include variables that store if they mated with a wild or

*Wolbachia*-infected mate. With this information the offspring's infection status follows the rules stated in table 2.4.

#### 4.5.6 RIDL

As it was the case with *Wolbachia*, this intervention requires that each female stores the type of male they mated with. With this information, we can simulate if their offspring is RIDL-infected or not. The Mendelian inheritance pattern is described in table 2.3, which is translated in the simulation into the probabilities of inheritance shown in table 4.8.

Male	Female	RR	Rr	rr
RR	RR	1	0	0
RR	Rr	0.5	0.5	0
RR	rr	0	1	0
Rr	Rr	0.25	0.5	0.25
Rr	rr	0	0.5	0.5
rr	rr	0	0	1

FIGURE 4.8: RIDL inheritance probabilities as modelled in SoNA3BS.  $R$  represents the presence of the RIDL allele while  $r$  represents absence.

## 4.6 Computational Time

As it was mentioned in our first sub-hypothesis (chapter 1.2), we want this simulation to be useful to the scientific community and, for this to be true, we considered it important that it was able to run in general-use computers in reasonable amounts of time.

We tested our model in a personal computer in a four-core  $3.5GHz$ ,  $i7$  processor ( $64bit$ ). This system has  $16Gb$  of RAM and a solid-state drive was used for additional virtual memory. In our tests, small mosquito population sized scenarios (five to ten breeding sites) ran in the order of two hours of computer time (for a one year simulated period). As we scaled up the number of mosquitoes, however, this time increased considerably. The largest scenario we tested consisted of 35 breeding sites and a temperature of  $35^{\circ}C$  (which increases the population size) in which each iteration took close to two days to finish. This is outside the operational range we have set for our model (chapter 1.3) but still well within the specification set by sub-hypothesis 1; which

states that the model should run in the order of days in a general-purpose computer system. It should be noted, though, that we can run several iterations of the system in parallel (4 iterations in the case of our four-core computer); which is relevant because due to the stochastic nature of the model, we need to repeat scenarios several times for statistical significance.

Overall, using computers with more processing cores and RAM would make the model useful in larger scenarios without having to make any modifications to its code. The main computational limitation faced for the experiments performed throughout this thesis was the number of nuclei available to run separate iterations of the experimental setups. Given more computational resources scenarios of about 50 breeding sites would be feasible in reasonable times (namely, about 16 physical computer cores).

As for scalability of the ABM, there are several points where optimisation can take place so that the simulation can be run with larger mosquito populations. Particularly, the egg stage poses a high computational cost due to the large number of individuals that live in this stage at any particular point in time. To alleviate this we could take two approaches. One would be to substitute egg survival binomial events to exponentially distributed survival probabilities (this optimisation could even be performed on adult stages). This would lower the computational cost of egg stages and allow larger populations to be run in reasonable times. The second approach would be to add a constant rate of larva emergence, which would skip the egg stages hoarding of memory altogether (albeit losing some biological precision in the process). Other improvements would be to make the data export routines and biting lists storage more efficient, which we plan to do in future work.

## 4.7 Conclusions

In this chapter we have described thoroughly the models and sub-models used in our simulation. By stating the reasons behind the modelled behaviours, their sources and the way they were represented in our simulation, we have addressed part of the first two sub-hypothesis of our thesis work (which states that our model should adhere to current mosquito biological knowledge, provide with a description of sub-models used along with their sources and be able to model current mosquito control interventions). The second part of this first sub-hypothesis, which stated that the limitations of the model should be discussed and stated clearly, will be completed in the next chapter of this document.

# Chapter 5

## Model Validation

*“The first principle is that you must not fool yourself and you are the easiest person to fool.”*

*Richard Feynman*

A model is just as good as the predictions it makes, and the insight it provides. This is why we need a way to assess how well our ABM represents reality. Although it would have been ideal to have field data to validate our framework; these data is, unfortunately, difficult to obtain. This is because it is expensive (both in time and economic resources) to collect information on mosquito counts at different life stages, while also keeping track of weather factors and breeding sites sizes. It would have been complicated to obtain these data and curate it to a level where we could use it to parametrise and validate the behaviour of our ABM in the allotted time for the project (as we would have needed really detailed descriptions not only of the data but of the environmental circumstances upon which these data was collected).

This, however, is no excuse to neglect validating our ABM’s behaviour. We need means to verify that our system is consistent with current scientific knowledge of the biological phenomena we are trying to model. For this reason, we tested our ABM against the differential equation approximation of a stochastic mathematical model that has gone through peer-review process (created by Otero et al. [51]). The reason behind the selection of this model in particular was twofold. On the one hand this model uses mosquito-life parameters that can be easily translated for their use in our simulation as well as similar environmental assumptions (death probabilities, metabolic development rates, etcetera); which ensures that we are using comparable settings in both systems. On the other hand, it is not exactly the same type of model so we are

not comparing it to a system created to answer the same type of questions; which it gives us a certain level of robustness in our assumptions and validity of the predictions. As a final remark we should mention that the mathematical model we used to compare our simulation was developed by a research group with which we have no collaboration so both models were created independently. This helps us avoid biases in the results and ensures that the models are comparable even with their inherent structural differences.

In this chapter, we will present two experiments in which we will compare the predictions of our ABM and the independent ODE: a factorial sweep of breeding sites numbers and environment temperatures; and a year-long simulation with realistic temperature pattern. With these simulations we will show that our model follows the behaviour of the mathematical model within a reasonable level of error within the range 15 to 32.5 Celsius degrees and up to 20 breeding sites (which falls in line with the scope set for our simulation in chapter 1.3).

## 5.1 Methods

To begin the description of our validation experiments we will define the methodologies to follow. This section will be divided in two parts. First we will briefly describe the model proposed by Otero et al. (as it is defined in depth in its publication [51]). After doing so, we will describe the two experiments we performed to compare the behaviours of the models.

### 5.1.1 Stochastic Model Description

This model was created for large-scale population dynamics analysis of mosquitoes living in temperate climate environments. This is reflected by the fact that it is used to analyse population sizes responses to weather in Buenos Aires, Argentina; along with the population extinction thresholds due to seasonal temperature changes in the same city. In this work, we worked with the ordinary differential equations approximation of the Markov model (both defined by the authors on their publication).

Otero's ODE approximation considers five mosquito life stages: egg ( $e$ ), larva ( $l$ ), pupa ( $p$ ) and females in first and second gonotrophic cycles ( $a_1$  and  $a_2$ , respectively). This system is defined by the set shown in equations 5.1 and 5.2.

- 
- (a)  $\frac{\partial e(t)}{\partial t} = egn * (ovr1 * a1(t) + ovr2 * a2(t)) - me * e(t) - elr * (1 - \gamma(l(t))) * e(t)$
- (b)  $\frac{\partial l(t)}{\partial t} = elr * e(t) * (1 - \gamma(l(t))) - ml * l(t) - \alpha * l(t)^2 - lpr * l(t)$
- (c)  $\frac{\partial p(t)}{\partial t} = lpr * l(t) - mp * p(t) - par * p(t)$
- (d)  $\frac{\partial a1(t)}{\partial t} = par * ef * 1/2 * p(t) - ma * a1(t) - ovr1 * a1(t)$
- (e)  $\frac{\partial a2(t)}{\partial t} = ovr1 * a2(t) - ma * a2(t)$

EQUATION 5.1: *Aedes aegypti* population model proposed by Otero.  $e(t)$  represents the number of eggs,  $l(t)$  the number of larva,  $p(t)$  the number of pupa,  $a1(t)$  the number of female adults that have not laid eggs and  $a2(t)$  the number of females that have already laid eggs. The parameters of the model are: eggs mortality ( $me$ ), larva mortality ( $ml$ ), pupa mortality ( $mp$ ), adult mortality ( $ma$ ), egg-hatching rate ( $elr$ ), larva pupation rate ( $lpr$ ), adult emergence ( $par$ ),  $a1$  oviposition rate ( $ovr1$ ) and  $a2$  oviposition rate ( $a2$ ).

This set of equations can be solved analytically for their deterministic behaviour, which will be used for the rest of this chapter (this solution was found using *Mathematica V11* and is available on the project's repository <https://github.com/Chipdelmal/SoNA3BS>).

- (a)  $\alpha = \alpha_0/BS$
- (b)  $ml(T) = 0.9725e^{-0.369891(T-278)}$
- (c)  $mp(T) = 0.9725e^{-0.369891(T-278)}$
- (d)  $\gamma(L) = \begin{cases} 0 & \text{if } L/BS < a_0 \\ 0.63 & \text{if } L/BS \geq a_0 \end{cases}$

EQUATION 5.2: Auxiliary functions for the mathematical model.  $ml(T)$  represents the larva mortality rate,  $mp(T)$  the pupa mortality rate,  $\alpha$  the carrying capacity factor and  $\gamma(L)$  the breeding sites threshold for larva inhibition.

Otero's model makes some similar assumptions to the ones we make in our ABM. The main ones being:

- Equal temperature in water and in the air.
- No rainfall dynamics.
- No death due to sun exposure.
- Metabolic rates controlled by temperature and modelled by Schoolfield's equations [84].

Despite of this, there are some differences between the models that should be pointed out:

- Predation processes are included in larval stages of mosquito populations in the ODE.

- The models for density-dependent processes are different.
- Our ABM currently incorporates only one static gonotrophic period which contrasts the two gonotrophic periods available in the ODE. The first gonotrophic period ( $a_1$ ) encompasses the time it takes for a female to develop eggs on the first oviposition bout and the second gonotrophic period ( $a_2$ ) all the subsequent oviposition bouts.
- The ABM does include day and night cycles, whilst the ODE does not.
- The ODE model does not include male individuals in the model.

It is worth noting that some of these differences will become relevant when we make the comparison of the results obtained from our experiments (particularly, the differences in modelling of density-dependent processes and gonotrophic cycles).

### 5.1.2 Methodology

To verify the response of our model in comparison to the ODE one, we performed two experiments: a factorial sweep of temperature values and breeding sites number (BS) in which we compared the female adults population sizes; and a seasonality experiment in which we compared the same output but with a realistic yearly temperature variation pattern. We will now describe in detail how these experiments were defined and performed.

#### 5.1.2.1 Factorial Experiment Setup

To compare the stable-state responses of the models we performed a five level factorial experiment in which we obtained the total population of female adults through a range of temperatures and number of breeding sites. This is important so that we can verify that our model performs correctly in stable environments, and that its responses are sensible regarding the predicted mosquito quantities.

**Temperatures** The populations were calculated in both models for temperatures between  $15^{\circ}C$  and  $35^{\circ}C$  (in increments of  $5^{\circ}C$ ). This range of temperatures was defined because it is generally regarded as the *Ae. aegypti* suitable temperature range for survival. Extinction thresholds are usually expected around constant temperatures of  $15^{\circ}C$  [51, 87] and survivability is difficult for this species above  $32.5^{\circ}C$  on the field [87] (directly due to heat and indirectly due to decreases in humidity and oviposition sites).

**Breeding Sites** To compare the response to population sizes changes due to increases in viable breeding sites we tested the models in a range between 5 and 25 oviposition spots. This range was selected because, despite the scope of our model being defined as capable of simulating up to 20 oviposition spots (chapter 1.3); we wanted to know more about the response of the system when this variable scaled up so we included the simulation of a slightly larger number of aquatic habitats in the experiment.

The location of the breeding sites was kept static throughout the simulations to reduce variations in the repetitions. This, however does not pose a problem in the generalisation of results as, for reasons that were described in chapter 6, population sizes were found to be resilient to spatial distribution modifications in small-scale scenarios.

Each of the 25 experimental scenarios was repeated 20 times for a total of 500 simulations (we stopped at this number of repetitions due to the low variance in the population sizes between iterations, which will be discussed when we present the results of our experiments); and the experiments were run for 200 days to allow complete stabilisation of the initial population dynamics (“burnin” period). This period was defined empirically after testing different settings of the simulation (the longest settling time was of about 150 days, so we allowed a 50 more days as an additional error buffer). From each run we extracted the last female adult population value on both the stochastic model and our ABM. With these values we calculated both the raw difference in population sizes and the error between them (equations set 5.3 shows the way these differences were calculated).

$$(a) \text{difference}(BS, T) = ABM_{ssv} - ODE_{ssv}$$

$$(b) \text{absDifference}(BS, T) = |ABM_{ssv} - ODE_{ssv}|$$

$$(c) \text{error}(BS, T) = \frac{ABM_{ssv} - ODE_{ssv}}{ODE_{ssv}}$$

$$(d) \text{absError}(BS, T) = \left| \frac{ABM_{ssv} - ODE_{ssv}}{ODE_{ssv}} \right|$$

EQUATION 5.3: Equations used to compare the ODE model with our ABM in steady state (*ssv*) through changes in breeding sites number (*BS*) and temperatures (*T*).

### 5.1.2.2 Seasonality Experiment

To test the models’ responses to temperature in a realistic weather scenario; we compared their female adult populations predictions using a simulated version of the climate in Catemaco,

Veracruz (Mexico). For this weather approximation we used real-world climatological data obtained from: <http://clicom-mex.cicese.mx/>; and we fitted it to the prototype function shown in equation 5.4 [88] (the resulting fitted model can be observed in figure 5.1).

$$\text{temperature}(day) = A + B \cos\left(C + \frac{2\pi \cdot day}{365.25}\right)$$

EQUATION 5.4: Prototype equation to fit the yearly variation of the temperatures with respect to the day number (*day*).  $A = 24.5$ ,  $B = -2.87$  and  $C = 0.0172$  are constants fitted to the experimental data.

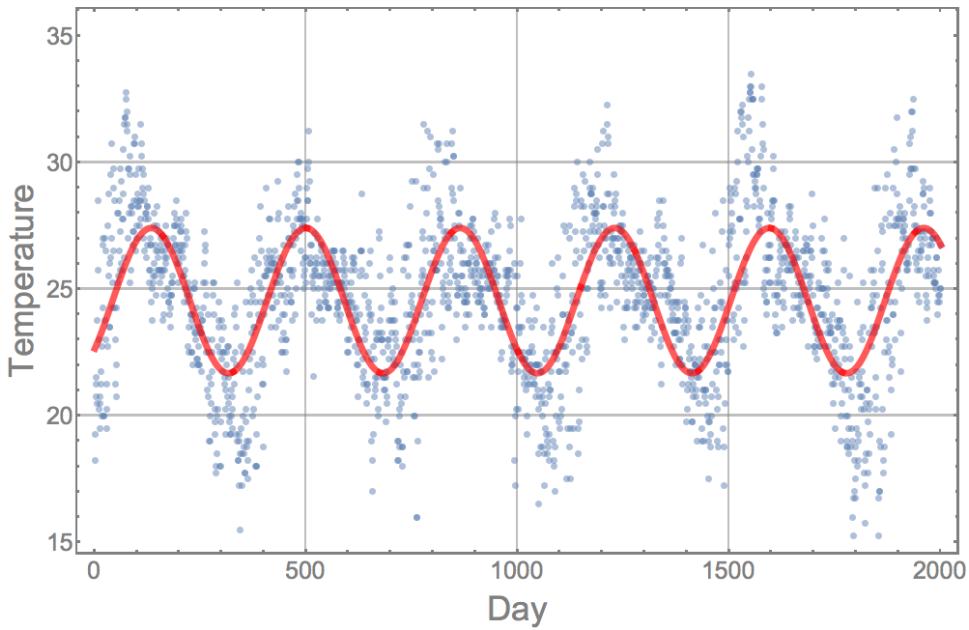


FIGURE 5.1: Weather model fitted to Catemaco's data from the year 2006 to 2011 and available on clicom.

Similar to the previous experiment, we defined three levels of population sizes by the number of available breeding sites. These levels were set to 5, 10, 15 and 20 oviposition spots; which were, again, selected because of the scope defined for our model (chapter 1.3). Each iteration was run for a total of 500 days and the adult mosquito population count was sampled 4 times every day. Each scenario was repeated 50 times for statistical significance.

## 5.2 Results

We will now present the results of both of the experiments along with the comparisons between the models.

These experiments took nearly 12 weeks to finish, which gave us an average of little over 7 iterations per day (with a total of 600 runs); complying with our design requirements laid down in chapter 1.3.

### 5.2.1 Factorial Experiment

The result of the factorial response comparison is presented in figure 5.2. In these plots we show the difference (5.2a), the absolute difference (5.2b), error (5.2c) and absolute error (5.2d) between the total adult population sizes in the models. The operational range of the ABM (as defined in the scope of this work) is contained within the purple rectangle of the contour plots. We can see that the discrepancy between these models becomes larger when we move from a lower to a higher number of breeding sites (particularly outside the 5 to 20 breeding sites range). The difference between the models is also higher whenever we simulate outside the range from  $15^{\circ}C$  to  $30^{\circ}C$  and it has a peak around the  $25^{\circ}C$  when we increase the breeding sites number. The ABM under-predicts the population size at a temperature of  $35^{\circ}C$ , although it should be noted that the  $35^{\circ}C$  is outside the population survivability temperature threshold which is closer to  $32.5^{\circ}C$ .

Regarding the variance in the repetitions of each data point, the difference between simulations was fairly stable. The maximum standard deviation across all scenarios was 19.92 mosquitoes with a mean of 3.72 and a minimum of 0.0045. Additionally to this, the average of the difference between the quartiles on the repetitions was 10.12 individuals while the average error difference was 0.1. These statistical measures gave us confidence in the number of repetitions performed to obtain the presented response surfaces and their validity.

### 5.2.2 Seasonality Experiment

Moving on to the results of experiment dealing with the realistic weather pattern; figure 5.3 shows the population size response from running a year on both models with the seasonality pattern described in equation 5.4. In particular, figure 5.3A shows the female adult population sizes responses as a function of temperature change. In this plot it can be seen that temperature change intrinsically includes the direction of the change in the weather as both models present hysteretic behaviour; which indicates that a change in population size from lower to higher temperature is different from the one obtained from the change in a high temperature to a

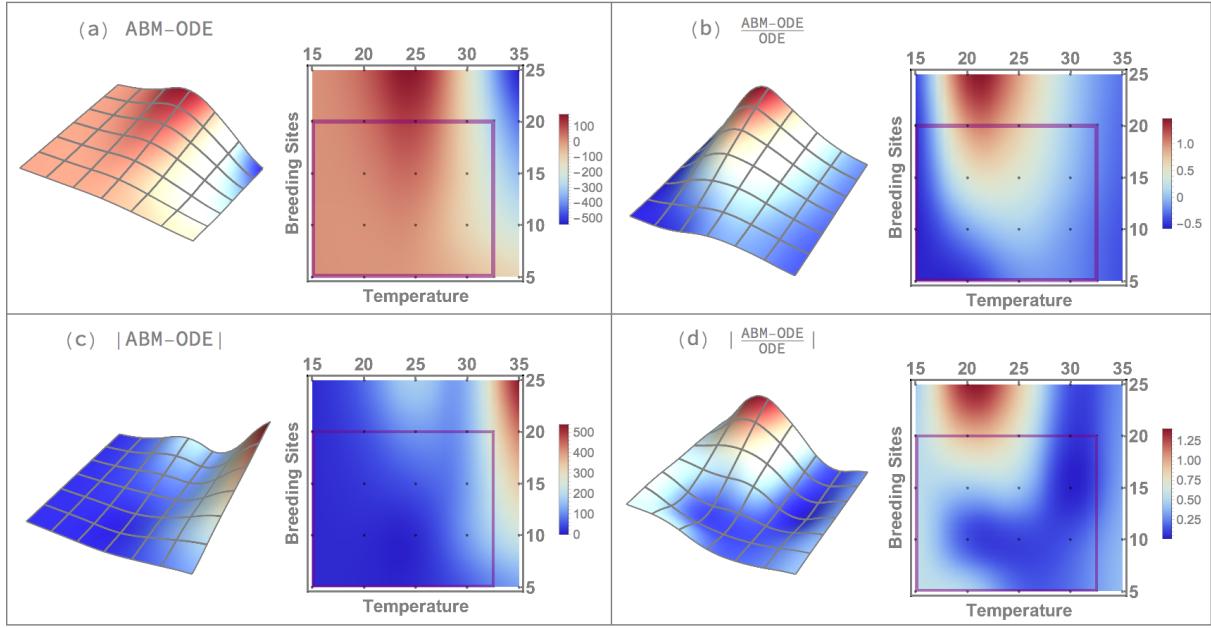


FIGURE 5.2: Calculated differences between the ABM and the stochastic model in terms of female adults in the population. Figure (a) shows the difference between models, figure (b) the absolute difference and (c) the error between them and (d) the absolute value of the error. The surfaces and contour plots represent the same data and scale while the surfaces are shown for visual reference. The purple rectangle shows the operative area for which the ABM was designed.

lower one (easily observed with the departure from a straight line to an elliptical shape in the temperature versus population size curve). This is an expected behaviour as population sizes do not respond immediately to temperature changes so the system presents a certain inertial resistance to change.

Throughout the tested period, our ABM predicted a larger amount of adult mosquitoes than the ODE in scenarios larger than the ones defined by 10 breeding sites (observed in figure 5.3B). This effect was magnified when we increased the number of oviposition spots, which confirms what was observed in the factorial responses to this variable (figure 5.2). The ODE model also had a less-linear response to the changes in temperature. This can be seen in figure 5.3A; where this model showed a less elliptical shape in the population sizes responses to temperature, while the ABM deviates less from this geometric pattern. In figure 5.3B we can also observe the difference in the response to the temperature dynamics as both models seem to be slightly out of phase when temperature rises (which corresponds to the lower part of the ellipse in figure 5.3B).

One thing worth mentioning, as a final remark, is that neither of the models presented “yearly memory”. This means that populations sizes cycle throughout the years in a stable way (without

increases or decreases in population sizes as years pass; which would be reflected in spiral shapes in population sizes of figure 5.3A).

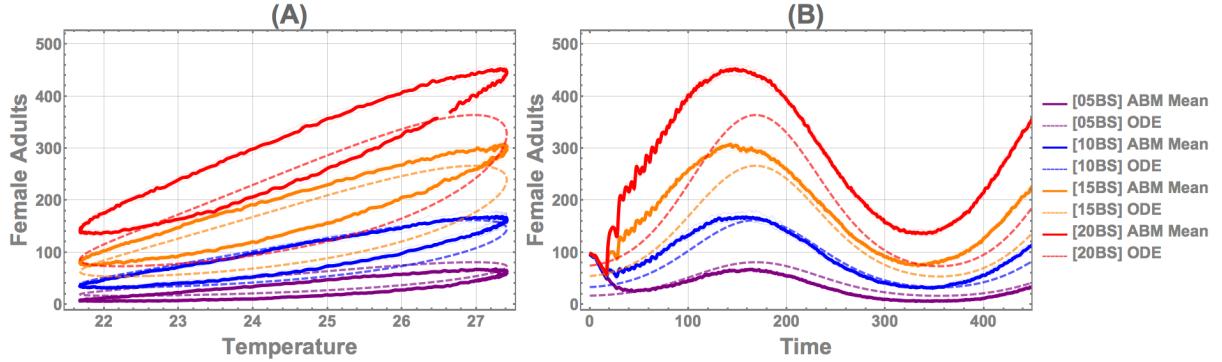


FIGURE 5.3: Change in adult mosquitoes number as a consequence of temperature. Figure (A) shows the population size changing as an effect of temperature changes (B) shows the same scenario but with population size change through time. In the first figure we can see the hysteretic behaviour of the population dynamics as an effect of seasonality's differential temperature change while the second one informs us of the amount of mosquitoes found at any point through the year.

## 5.3 Discussion

From the described results we can see that, although the models predict comparable results given some constraints and that the general shape of their dynamics is similar; they diverge when we increase the number of available oviposition sites. In this section we will discuss the reasons behind these differences and their implications.

### 5.3.1 Density-Dependent Modelling

The first difference between the models is the representation of density-dependent factors in aquatic habitats. The ODE model includes egg hatching inhibition and larva density death probability as a step and quadratic function (respectively). In contrast, our ABM models the egg hatching inhibition as a linear function and the larvae death probability as a sigmoid one.

The linear response to egg hatching inhibition is the one that presents the starker difference between the models. In the case of our ABM, the quantity of eggs in the environment varies continuously until the point where the death probabilities become too high for more eggs to be maintained in the aquatic habitats. In contrast, the ODE presents a step-function diapause where the eggs are prevented to hatch until enough environmental “room” is provided. This

modelling discrepancy, however, should not impact greatly the difference between the number of adults because the biological “bottleneck” is defined by the number of larvae that can be carried in the environment. As long as we have enough eggs in the habitats to saturate the aquatic population at the larval stage, sufficient individuals will compete for resources to cap the environment at this point. To verify this idea, we removed the diapause term from equation 5.1a; which resulted in no visible change in the tested scenarios. This difference in modelling of egg mortality/diapause, however, would become relevant if we come close to extinction thresholds as the deviations between the models could become meaningful.

In terms of larvae density dependence, both models use different functions to model the death probabilities of mosquitoes in this stage. Our ABM uses a sigmoid function for larvae death probabilities that scales with the number of breeding sites in the environment (described in chapter 4.3); whilst the ODE uses a quadratic density-dependent equation (equation 5.5). This difference can account for part of the divergence in the prediction of the adult mosquito quantities between the models. To show the effect of this difference, we perturbed the value of the  $\alpha$  parameter of the ODE (equation 5.1b) and observed the response in the dynamics curves (this parameter is related to the carrying capacity of the environment and is defined by the authors as an empirical constant  $\alpha_0 = 0.75$  over the number breeding sites in the environment). The result of this change is shown in figure 5.4A in which we can observe that reducing this parameter shifts the population size upward and changes the shape of the curve to a less-linear form. Furthermore, figures 5.4 A and C; show how, by modifying this parameter, we can obtain similar population baselines to the ones predicted by the ABM in the seasonality experiment. However, as can be seen in figure 5.4C, this is not enough to account for the discrepancies alone, which diverge in shape as we scale up the number of breeding sites. For this, we will address the second main difference between our models: the inclusion of two gonotrophic cycles in the ODE system but not in the ABM. Before doing so, however, it is important noting that the  $\alpha_0$  variable is, as described by the authors: “*the only adjustable parameter on the model*” [51]. By changing its value, we are not modifying an arbitrarily selected constant, but changing the amount of larvae that can be sustained in each aquatic habitat; which would be equivalent to having smaller or larger containers in the tested scenarios.

$$\omega'_5 = -\alpha * l(t)^2$$

EQUATION 5.5: Larval predation mortality factor ( $\omega'_5$ ) as a function of larvae quantity ( $l$ ), and scaling factor involving the standardised container carrying capacity to breeding sites fraction ( $\alpha$ ).

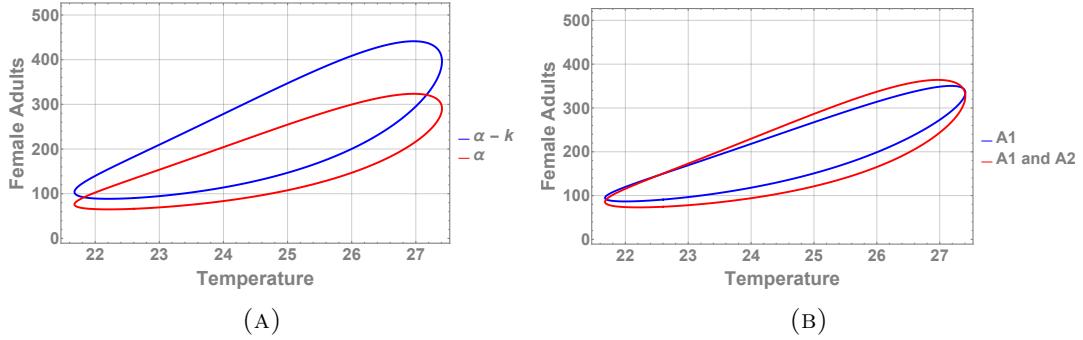


FIGURE 5.4: Modifications performed to the ODE to highlight differences with ABM. Figure A shows the response to slight changes in the  $\alpha$  parameter of the ODE (associated to the carrying capacity of the aquatic sites). It can be seen that this parameter shifts the curve upwards and changes the shape slightly (which mimics what happens with the output of our ABM that predicts more mosquitoes than the ODE). Figure B shows the results of the ODE model when taking into account one and two gonotrophic cycles ( $A_1$  and  $A_2$ ). In it we can see that the removal of the second gonotrophic cycle causes the system to have a more linear response (similar to the one obtained in our ABM).

### 5.3.2 Gonotrophic Cycles

The second big difference between the models is that we currently incorporate one gonotrophic cycle length, whilst the ODE includes two. This, as we will see in the following discussion, affects the responses of each of the models to the temperature changes in the environment.

In general, the ABM model predicted an elliptical pattern of responses to seasonality changes. In contrast, the ODE presented a less linear response to them. While other factors might certainly be influencing this result, the discrepancy can largely be explained by the implementation of two adult stages in the ODE. In figure 5.4B, we present this model's response when we include one and two gonotrophic cycles in its implementation. In this figure, the curve that corresponds to the use of two different stages presents a shape with a more pronounced hysteretic behaviour; while using only one gonotrophic period mimics more closely the behaviours obtained in the ABM model (figure 5.3). This is a result we were not expecting but that highlights the importance of including the models of these two adult stages to obtain accurate population dynamics in changing weather experiments.

Finally, we can see in figures 5.4B and 5.4D; that by removing the second gonotrophic cycle and perturbing the  $\alpha$  parameter in the ODE system, we can account for most of the differences in the responses of the models. However, there are still discrepancies in the predictions of the systems in larger population sizes (which can be seen in figure 5.4D) that can't be accounted for by just perturbing the  $\alpha$  parameter but that would require changes in the shape of the density-dependent function. For the scope of this work, however, this exploration is enough to

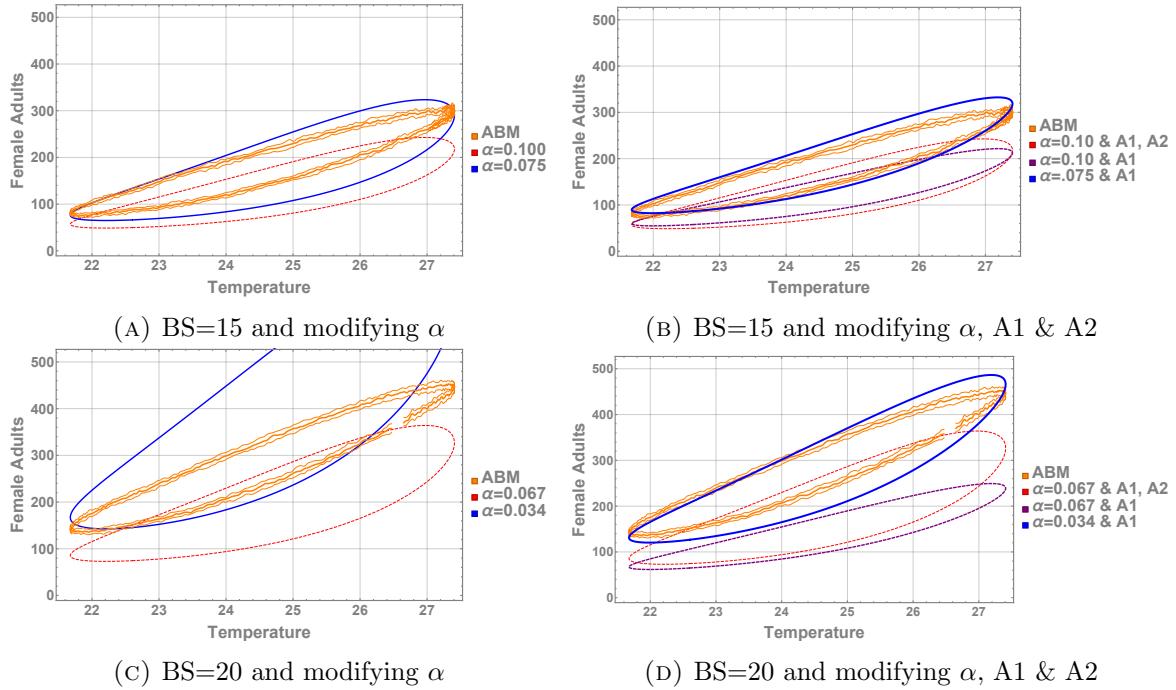


FIGURE 5.5: Collapsing the ODE system into the ABM one by making equivalent modelling assumptions. In all figures, orange curves show the population size response of the ABM and red curves show the ODE's response with its original  $\alpha$  parameter and both gonotrophic cycles. In figures A and B the blue curves show the result of modifying the  $\alpha$  parameter. In figures B and D purple curves use the original  $\alpha$  value with only one gonotrophic cycle; and blue lines show the result of using one gonotrophic cycle and a modified  $\alpha$ .

show that the systems are comparable and that by making modifications to the ABM we would be able to simulate larger populations with a lower error level if we needed to do so.

### 5.3.3 Remarks and Limitations

As we have seen in this validation experiment, the population size's response of our ABM can be thought of as a particular case of the published ODE system. For the experiments presented in this thesis project, the error obtained is considered acceptable between the ranges of temperatures and breeding sites defined in our scope (chapter 1.3). This is mainly because, at this stage of our research, we are interested in comparisons between scenarios with the same amount of aquatic habitats and stable temperature values. In these cases the system is internally consistent, meaning that even if we have a slight error in the prediction of individuals numbers, this level of error should fairly uniform across the compared scenarios. For example, in the case of the spatial heterogeneity experiment presented in chapter 6, we tested a scenario with 12 breeding sites at  $25^{\circ}\text{C}$ . By checking the level of error in the prediction of female population sizes (figure 5.2) we can see that its value is close to zero. Even if it was the case that we were

simulating a larger number of breeding sites (which would increase the error level) we could expect the error in the predicted population size to be uniform across the tested interventions.

Another point to consider is that we are specifically interested in the way human movement changes the way pathogens spread in their populations; which is unlikely to be affected in a significant way by the slightly increased number of mosquitoes predicted by our ABM. This would ultimately translate into uniformly denser networks all across the tested scenarios so the structures of the networks are likely to be resilient to small deviations in predicted population sizes.

There are situations, however, in which we do have to take into account differences in predicted population sizes. For example, where we compared scenarios with different numbers of breeding sites and precise values of extinction thresholds. This is because in both these situations the levels of error are different across the settings so we would be biasing results if we do not account for this divergence between predictions. Moving forward and thinking into increasing the operational range of our model and expanding the range of applications for which it would be useful we will have to make some changes in our ABM:

- The first of these changes would be to verify and modify the density dependent functions. Data of larval stages counts are easier to come by than adult stages so it should be able to model and parametrise these functions with real-world information. This would ensure that there is a biological foundation to these models and that they are accurate.
- The second modification would be to include the modelling of two separate gonotrophic stages to obtain dynamics that more closely resemble the real-life behaviour of the species.
- The final modification would be to handle the carrying capacity of each container separately. This would ensure that there are no errors being introduced as a function of carrying capacity scaling in the environment. An extension to this would be to add varying levels of carrying capacities according to the size of the water containers, which would add flexibility to the ABM.

Finally, we would like to acknowledge the limitations of comparing our model to another model instead of field data. By taking this approach we have made the implicit assumption that the ODE represents reality; and, while this system is compared to the patterns of population behaviours in Argentina, further analysis is required for this to be proven conclusively. We performed these validation experiments as an initial, wide-scope, verification that allowed us to

verify the general behaviour of the model within its application scope; as well as to identify the areas in which we need to improve our simulation to make it more robust. In the future, however, we would definitely aim towards parametrising and verifying predictions from our model with field data (more on this subject will be mentioned in chapter 7.1).

## 5.4 Conclusions

In these experiments we have compared an ODE model designed for the analysis of the dynamics of large mosquito populations with our ABM designed for the analysis of small-scale scenarios. By comparing both their responses in a factorial setting scaling the temperature and breeding sites number (correlated to the carrying capacity of the environment) we were able to quantify the difference in the response of a computationally and environmentally viable range of settings.

With these comparisons we can conclude that within the range of population sizes and temperature ranges it was designed for, our ABM predicts a reasonable amount of mosquito adults as compared to the independently published ODE model presented in this chapter. In the future, however, we would like to make changes to our model so that it can be used in a wider array of situations of epidemiological interest.

# Chapter 6

## Spatial Heterogeneity

*“We are just an advanced breed of monkeys on a minor planet of a very average star. But we can understand the Universe. That makes us something very special.”*

*Stephen Hawking*

As it has been stated throughout this work, the driving idea behind the creation of this simulation has been to develop a model that allows us to study how spatial variables in human-mosquito interactions can affect the epidemiological relations between humans in mosquito-borne diseases. We have proposed the use of network theory (also known as graph theory) as means to quantify and analyse these relations and their epidemiological importance. In this chapter we will describe the experiments we performed to show that these ideas are viable and useful.

The importance of spatial heterogeneities on pathogen transmission [14, 42, 89], has motivated the use of *SoNA3BS* to explore the effect of spatial arrangement on vector-borne disease transmission. Specifically, we used the ABM to test the hypothesis that spatial distribution of human houses and mosquito breeding sites has a significant effect on reshaping the way human contacts occurred both, in absence and in presence of three mosquito-control measures: spatial fogging, *Wolbachia*-releases and RIDL. To test these effects, we first simulated two different spatial scenarios: a homogeneous one (in which every household and human is placed in the same place) and a heterogeneous one (with a more natural spatial distribution obtained from a real human settlement). We then obtained both the population dynamics and the vectorial-contact networks that result from applying each of the interventions in the environment. This information allowed us to show that despite the fact that the mosquito population dynamics remain almost

identical in both situations on all cases, the networks that arise from them have different structural properties. This, in turn, shows that mosquito biting heterogeneities can arise solely from spatial distribution and highlights the importance of taking into account spatial information in the planning of mosquito-control interventions' deployment on the field.

## 6.1 Methods

To investigate how spatial distribution of hosts and egg-laying sites affects the ability of vector control interventions to disrupt pathogen transmission, we simulated two different scenarios under the presence of *Aedes* control campaigns. After doing so we analysed and compared both the population dynamics and vectorial-contact network structures.

### 6.1.1 Simulated Scenarios

In the following section we will discuss how these spatial settings were selected and implemented on the simulation; along with the measurements that were performed as part of the analysis.

#### 6.1.1.1 Landscape and Humans

We generated a simulated version of a location near the Mexican town of Catemaco, Veracruz (shown in figure 6.1) by obtaining approximate positions of houses' locations using *Google Maps* (geographical coordinates: 185°52.0'N 955°25.1'W). We chose this area because it is a small-sized region in Mexico in which the *Ae. aegypti* presence is widespread [90].

A population of 30 humans was distributed amongst 12 houses proportionally to the household area that was detected (viewed from the satellite image). This gives us an average value of 2.5 persons per house. For each household we assumed one viable egg-laying site in its vicinity, as most of the populations with no piped water supply use containers to provide for their needs [91]. Under these conditions, we simulated two different scenarios:

- Homogeneous (HOM): Every house and person was placed at the center of the environment as depicted in figure 6.2a. Humans remained static while mosquitoes were allowed to move to fulfil their biting, sugar feeding and eggs-laying needs.



FIGURE 6.1: Original houses distribution of the studied population near Catemaco, Veracruz. This population was encoded into our simulation through the use of computer vision algorithms.

- Heterogeneous (HET): Houses in the environment were placed as shown in figure 6.2b. A maximum of two humans per household were allowed to visit other households (with a probability of 10% per day), while the remaining humans stay at home. This house was chosen randomly from the pool of houses in the simulation each time a visiting event was triggered.

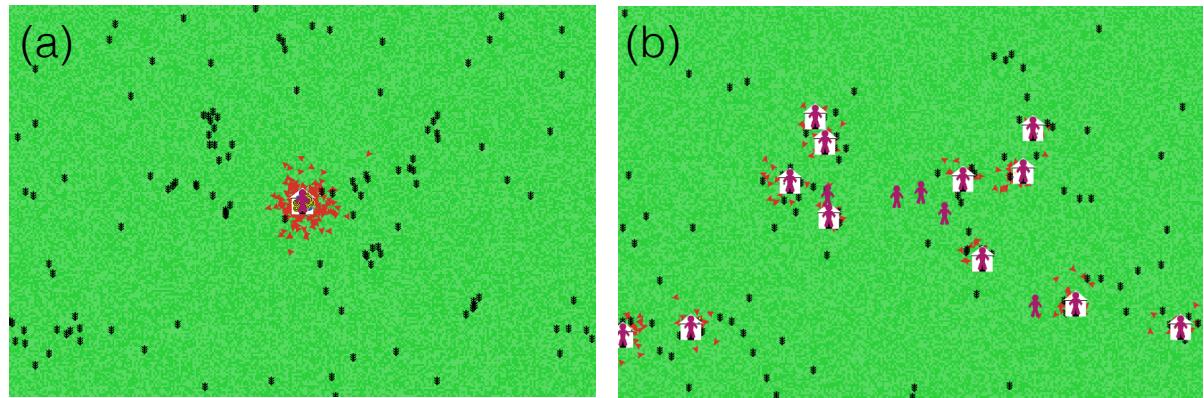


FIGURE 6.2: Simulated spatial scenarios. SoNA3BS screenshots representing the two modelled layouts: homogeneous (a) and heterogeneous (b). Humans are shown in purple, houses in white, sugar sources in black and mosquitoes in red (triangles). Breeding and mating sites were scaled down but are always near the location of the houses according to the methodology described while the other agents and landscape elements were scaled up for readability purposes.

### 6.1.1.2 Time and Temperature

In these scenarios we, once again, gravitated towards the high time resolution value of 5(*min/tick*) (although other similar ABMs generally step sizes of hours or days [59, 90, 92]). This is important for our analysis because we are interested in the analysis of fine-detail interactions between individuals so we need to be able to simulate them as precisely as computationally possible.

The simulated timespan is also important in this experiments as it can affect the analysis of the dynamics of the system (enough time must be given for transitory dynamics to settle). In our study we needed variations in the populations' sizes and, as a consequence, mosquito bites to achieve stability while still maintaining a low computational cost. To achieve this, each run comprised a period of 360 days with a burn-in period of 100 days. Both of these periods were defined empirically after running extensive preliminary tests on the simulation.

With the timing information of our experiments in place we moved on to defining its weather characteristics. For the purposes of the experiments presented in this work, we fixed the temperature's value to 25°C (close to the average temperature of the region which was calculated with data obtained from: <http://clicom-mex.cicese.mx>). We made this decision because we wanted to limit the effect of variables other than spatial distribution on the results (having a realistic seasonal pattern would have not only affected the mosquito population dynamics but it would have also added interactions with the timing of mosquito-control interventions [12, 13, 27]).

### 6.1.1.3 Mosquitoes

Simulating a realistic population of mosquitoes was a crucial part of our experiments and as such, special care had to be taken defining their quantities and their behaviours.

Under the described conditions, our ABM produced a baseline population at equilibrium of 30 adult mosquitoes per aquatic habitat. This number is close to densities observed in Cayman Islands (where RIDL field tests took place and with similar weather patterns to the ones found on Catemaco [13]). It is important to note, though, that this carrying capacity value is not a hard threshold in our ABM, but an emergent property that arises from interactions between mosquitoes and the environment as a whole.

In terms of mosquito behaviour, our simulation treats mosquitoes going through their life stages as finite state machines as described in chapter 4.3.

### 6.1.2 Simulated Mosquito-Controlled Interventions

Once the landscape and behaviour of the agents was defined we focused on the way the vector-control interventions affected mosquito individuals. Three different mosquito interventions were simulated along with a baseline scenario. Their effects on mosquitoes were defined as follows (which mimic the definitions described in chapters 2.3 and 4.5):

- Baseline (Base): Mosquito dynamics and death probabilities remained unchanged.
- *Wolbachia* (Wolb): The bacteria was inherited between generations according to cytoplasmic incompatibility's rules of inheritance [9]. Adult lifespan was halved on average and the ability of *Wolbachia* to block the disease's transmission was assumed to be 90% (although studies suggest that this percentage could be higher depending on the strain used [12, 33]).
- fsRIDL (RIDL): Mosquitoes carrying a lethal gene allele were allowed to mate and reproduce but their female offspring did not develop viable wings [13]. Only males could mate further and propagate their genes. These RIDL genes were transmitted to the offspring in accordance to Mendelian inheritance laws.
- Spatial Spraying (Fog): An instant-action layer of insecticide is applied to the environment. This insecticide was parametrised in the way described in chapter 4.5.

Control measures were applied uniformly in the environment. *Wolbachia* and RIDL had a fixed number of mosquitoes released uniformly over the landscape; while fogging was assumed to work with equal efficacy across the whole landscape. This decision was taken to focus on how the spatial distribution of individuals affects the effectiveness of the control campaigns, not the specific way in which interventions are applied in the landscape (which is known to be important [93, 94]).

In terms of time and density of the releases, intervention events took place on a weekly basis to match field campaigns; and the number of RIDL and *Wolbachia*-infected released individuals were also scaled to match field tests (225 individuals per release [12, 13]). The main difference between the way the interventions campaigns were simulated was that *Wolbachia*-infected

mosquitoes were released for five weeks (in this period the pathogen achieved fixation in all situations), while RIDL and spatial spraying were applied during the rest of the simulated time to make a fair comparison between the interventions (*Wolbachia* gets fixated in the population and continues to be propagated while RIDL is self-regulating, and fogging stops working almost as soon as campaigns finish).

Each combination of spatial scenario and control measure was repeated 30 times to reduce the variance on the analysis.

### 6.1.3 Comparison Metrics

With the simulation's settings defined we now describe the analysis we performed on the obtained data. As discussed earlier, we performed a contrast analysis to compare population dynamics and vectorial contact networks across different scenarios. We think that making these comparisons is a meaningful way to separate the effects of spatial location from simple population counts. This is because in a scenario in which the spatial distribution of individuals had little to no effect, we would expect the quantity of mosquitoes and the biting networks structures to change proportionally to each other. However, if the spatial effects are meaningful, some independence in the way the metrics behave is expected.

#### 6.1.3.1 Population Dynamics

To analyse the impact of the control measures on mosquito population sizes, we stored their demographics twice a day. We focused on the analysis of adult mosquitoes which were broken down into the following categories: total adults, adult females, adult males, RIDL-infected and *Wolbachia*-infected. It should be noted, though, that we did store information on other life stages (eggs, larvae and pupae) in case further analysis is deemed useful.

#### 6.1.3.2 Networks

The vectorial-contact networks are, in our opinion, the most novel part of our analysis. Despite the fact that network epidemiology has become more widespread in direct-contact diseases [16, 95, 96]; in vector-borne scenarios performing this kind of transmission analysis is difficult. The use of an ABM allows us to track the biting history of each mosquito, so we can recreate

not only the epidemiological transmission network but also the network that arises purely from mosquito bites (of which the epidemiological one is a sub-network). Networks were obtained according to the following procedure:

1. Mosquito bites were recorded along with the time in which they occurred.
2. If a person was bitten after another person was also bitten by the same mosquito (in a previous gonotrophic cycle) a vectorial transition was created between them (in the case of *Wolbachia*-infected mosquitoes we are assuming pathogens transmission reduction, so 90% of the bites from them were discarded [12, 33], as we are interested in the potentially-infective ones). If the same person gets bitten in subsequent gonotrophic cycles these links are discarded as these bites are not epidemiologically relevant (self-loops would not disperse the disease in the population).
3. The resulting network is created by translating vectorial transitions into directed edges between persons (which are the vertices).

Once these graphs were generated for each scenario we calculated a collection of network measures that are related to disease transmission within a population [46, 97–99]. We did this for both the weighted networks (with the weight being the number of transitional bites between two given individuals) and the binary networks (one or more transitional bites are treated simply as one transition).

The first metric we evaluated was the in-degree distribution of the persons (number of incoming bites after the first mosquito's gonotrophic cycle). This is related to the general risk of a person to contract a vector-borne disease and is of upmost importance to understand diseases transmission. We defined the in-degree as the total number of incoming bites for each person as we are working with weighted networks (also known as multigraphs).

In addition, we also analysed the following network measures:

- Graph Density: Represents the ratio between the number of edges of our graph and the number of edges of a fully connected graph with the same number of vertices. This measure indicates how our network compares to the absolute worst case scenario where the possibility for transmission between every human in the population exists.

- Mean Path Length: Average length of the shortest paths in the network. It is related to the speed and reach a disease could have in the population because it represents to how many “jumps” a pathogen would have to make to cross from one human host to another.
- Small World Coefficient (SW): Measures to what extent nodes are neighbours of one another in relation to how long are their paths to other nodes in the network. When this quantity is high, networks have the characteristic of having nodes that are separated by a low number of steps but that are not necessarily neighbours with each other. The small-world effect is of epidemiological concern because it allows a pathogen to spread on a network even when connections between individuals are sparse (the mean path length grows logarithmically with the number of nodes) [97].

As a final step of analysis of the relation between spatial distribution and the frequency of transitional bites we used spectral clustering on the networks [100]. This allowed us to identify patterns that were arising on the biting behaviour of the mosquitoes and it is important because, if no clustering patterns could be found on the heterogeneous layouts, then we could conclude that there was no direct relation between person’s location and the transitions amongst them.

All of these networks analyses were performed on *Mathematica 11* using the functions provided with the software and extending its capabilities by using *PajaroLoco* (Héctor M. Sánchez C. [101]), a package developed for networks’ structural analyses.

## 6.2 Results

In the following section we will show the results obtained from the interactions of mosquito and human agents in our simulation. First we will show how population dynamics behaved in the presence and absence of spatial heterogeneity. Then we will make the same comparison on the vectorial contact networks. After making these contrasts we will describe briefly the differential effect of each intervention in terms of population sizes and efficacy on the disruption of vectorial contact networks.

As a note, we will mention that the whole simulations set (240 simulations) took around four weeks of computer time to finish. This gives us an average of 8 iterations per day which falls within the scope set for this project (chapter 1.3).

### 6.2.1 Population Dynamics

Most of the interventions behaved similarly in terms of the way they affected mosquito population sizes across their homogeneous and heterogeneous cases, falling in line with our initial hypothesis that they should be similar given the conditions of the experiments. The one notable exception was RIDL. We can observe in figures 6.3a and 6.3c a slight difference on the long term effect of the releases. In the long run, the heterogeneous scenario suggests that RIDL alleles had a harder time getting transmitted in the mosquitoes population; but presenting firm conclusions on RIDL dynamics requires longer simulation times (an objective for future research). For the purposes of this particular experiment we can say that, at least in the simulated timespan, the population sizes in both RIDL settings were very much alike (something that will be discussed looking at the mean population sizes of the experiments).

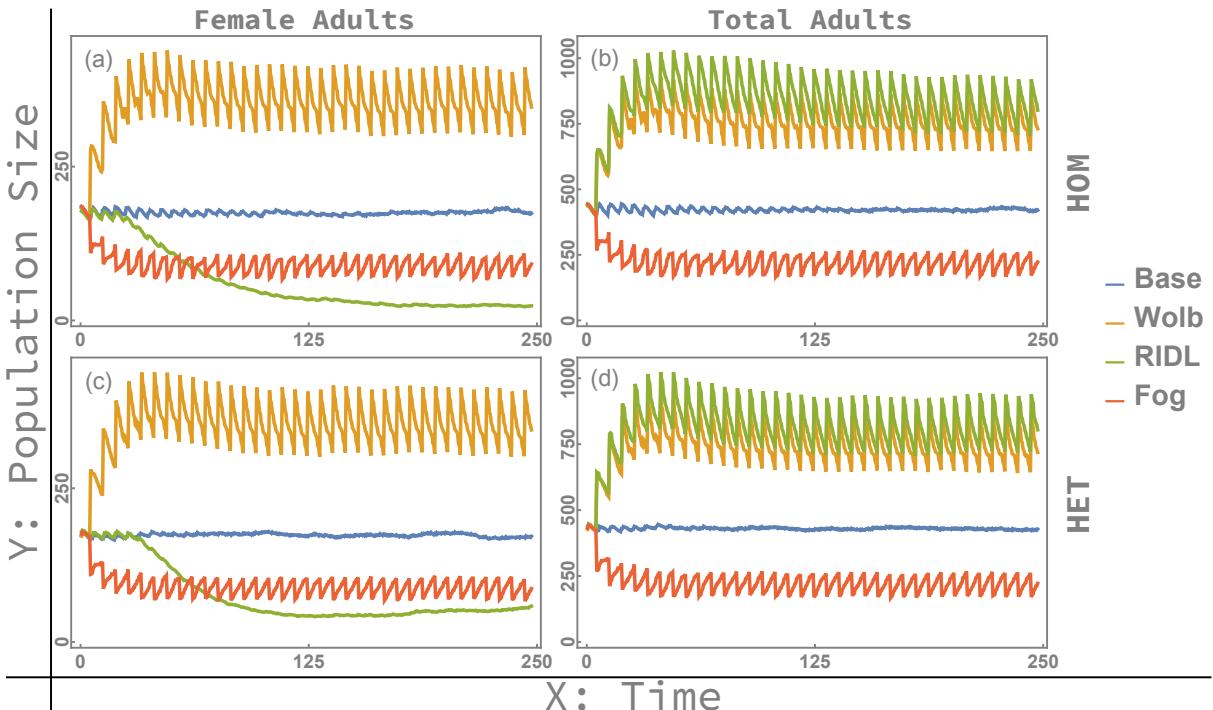


FIGURE 6.3: Population dynamics of adult mosquitoes. Figures (a) and (b) show the obtained populations in the homogeneous cases while figures (c) and (d) the heterogeneous ones. Given the experimental settings, the differences between the two spatial settings are barely noticeable through visual inspection.

Besides RIDL's specific case, visual inspection of the plotted data was not enough to establish any significant differences on population sizes between spatial layouts, so we calculated the average population size from each experiment (area under the curve divided by the number of time points). The results of these calculations are shown in figure 6.4, confirming that no meaningful deviations on population sizes were observed between spatial settings. This lack

of differences is most likely due to the fact that the tested environment was relatively small, interventions were applied uniformly on sites, and there was little environmental pressure on the mosquitoes (abundance of sugar food sources and human hosts). We would expect the dynamics to change if either environmental or behavioural variables become a significant external stressor on mosquito survivability.

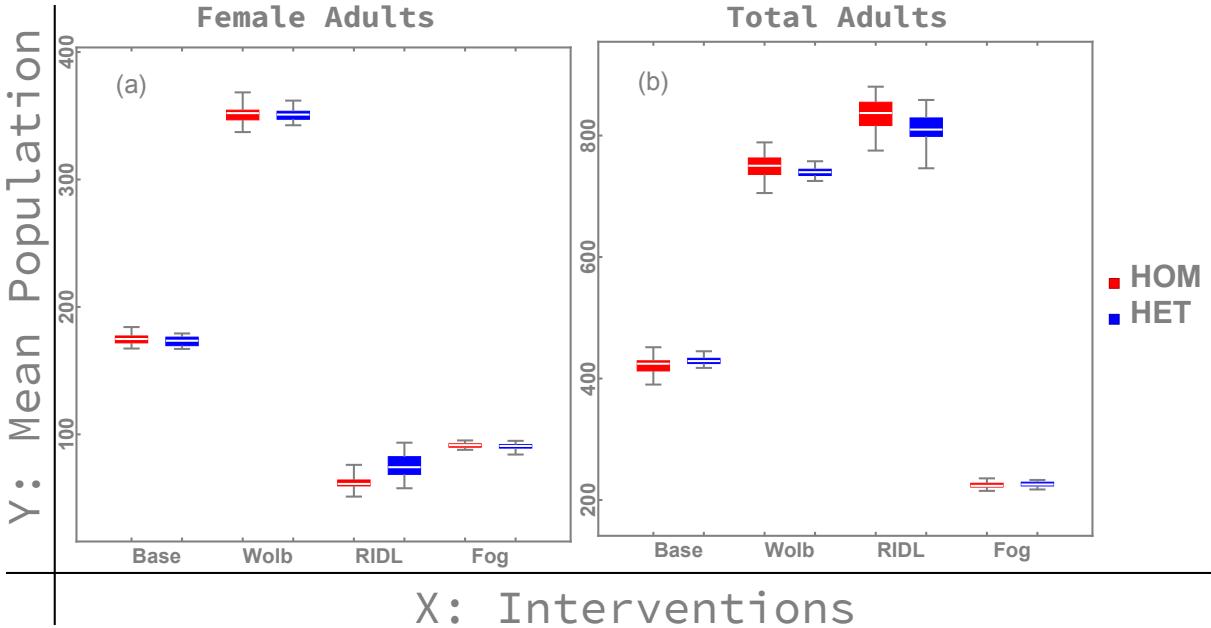


FIGURE 6.4: Adult mosquitoes mean population sizes. The area under the curve of the number of adult mosquitoes was calculated for each repetition of each scenario and was divided by the simulated time to obtain these quantities. Figure (a) Shows the number female adults while (b) shows the total adults count. Given this scenarios there was almost no change in the mosquito population sizes of the interventions in their homogeneous setting compared to the heterogeneous one.

Overall, the absence of significant differences in population dynamics due to spatial heterogeneity indicates that difference in mosquito-human vectorial contact networks can be attributed to changes to the spatial arrangement of the environment rather than change in population sizes alone.

### 6.2.2 Vectorial-Contact Networks

Given that we have established that there were no significant differences on mosquito population sizes due to spatial distribution changes, we move on to analyse the effect of the spatial layout upon the resulting contact networks. We did find significant differences in the networks between experimental scenarios, so we will divide this section into the analysis of the homogeneous and heterogeneous settings to better highlight the obtained behaviours.

## Homogeneous Layout

Under this spatial distribution we expected the contact networks to be uniformly distributed as a consequence of each human having the same probability to be bitten. To visually investigate this hypothesis we present the networks' transition matrices on figures 6.5a through 6.5d. To confirm the lack of distinct hosts communities we performed spectral clustering analysis on the networks. No structures on the connections between individuals were found, implying that no bias existed in the way mosquitoes selected their victims. The small-worldness values of these networks also confirm this result as they approach a value of 1 (figure 6.6c), corresponding to the case where mean path length is equal to the clustering coefficient (which is expected in a uniform network).

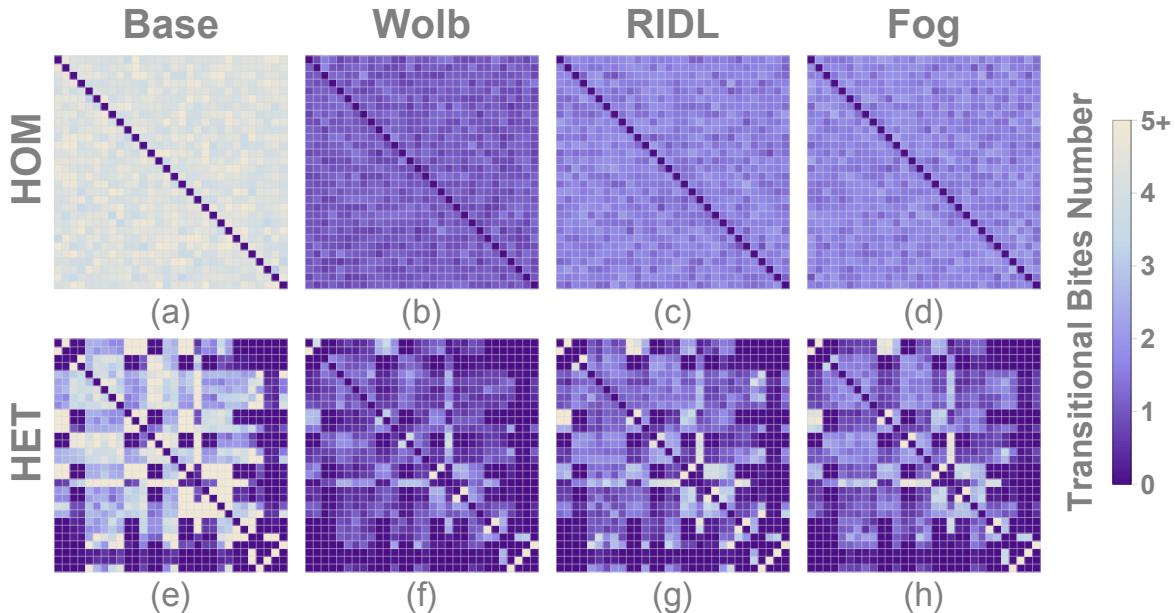


FIGURE 6.5: Vectorial contact transition matrices between individuals across different experimental settings. Each element is a count of total consecutive bites where the first bite is on the individual indexed by row and the second bite is on the individual indexed by column.

As a consequence of these results, the degree probability distributions showed a concentrated peak in their of the probability distribution frequencies (PDF) and a sharp transition in the cumulative distribution frequencies (CDF). These outcomes are represented by the solid lines in figure 6.7. Epidemiologically speaking, this is relevant because these peaks raise the herd immunity threshold to untenably high levels. Under the homogeneous setting of a fully connected network, quarantine ceases to be an effective method to halt transmission as all individuals are highly connected and a large number of edges would need to be removed to disconnect the network (figures 6.6a and 6.6b).

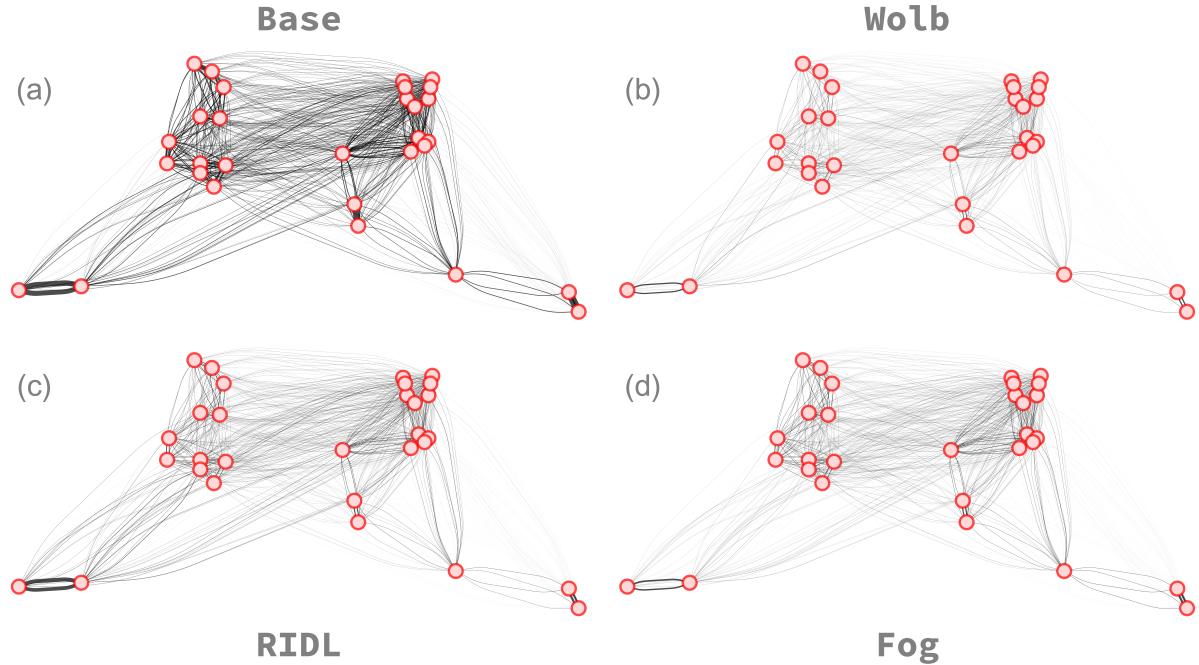


FIGURE 6.6: Heterogeneous scenario transition networks. Each node represents a human host and they are spatially distributed according to their location in the simulation. It can be observed that people who spend more time together tend to form stronger links between one another creating tighter clusters of people that live in proximity.

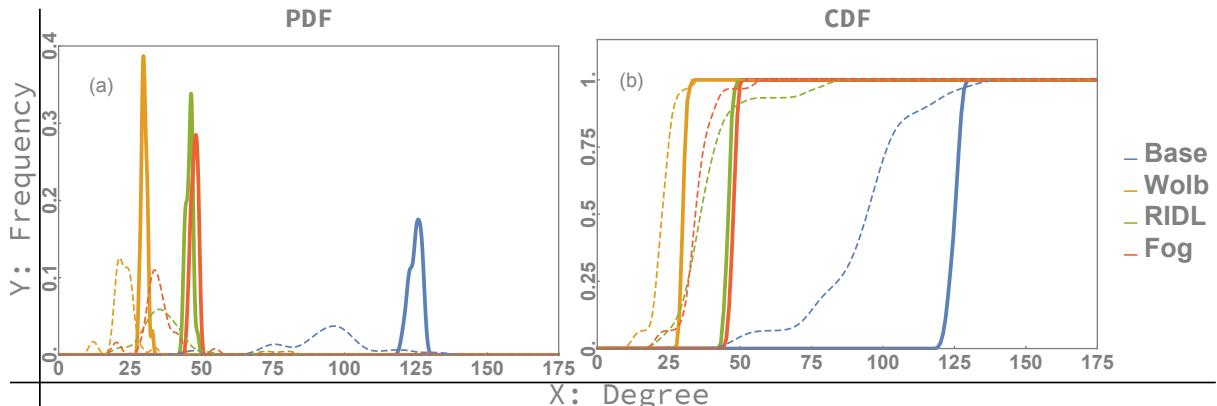


FIGURE 6.7: Degree probability distributions of spatial scenarios. Dashed lines represent the heterogeneous scenario while solid lines represent the homogeneous one.

### Heterogeneous Layout

In contrast with the homogeneous settings, these scenarios produced clear patterns in their transition frequencies matrices. The transition matrices shown on figures 6.5e through 6.5h clearly show the existence of clusters of individuals (this is a natural effect of *Aedes aegypti* mosquitoes being relatively weak flyers as compared to other mosquito species). We can also see this on figure 6.8, where some transitions connect individuals more frequently than others (represented by darker lines on the network visualisations). Performing spectral clustering

on these networks did find communities that correlate strongly with the spatial distribution of individuals (shown in figure 6.9). These results allow us to infer that a pathogen would be able to spread with relative ease within these communities, and that targeting the inter-community connections is a better approach to reducing transmission in a population (further demonstrating the idea that human movement plays a major role in dispersing *Aedes*-borne diseases [18, 102]). The close relationship between spatial arrangement of individuals on a landscape and communities embedded in the vectorial network structure hints at the small-world feature to the vectorial contact network. Investigating these network structures should be a priority in larger settings as they are highly relevant to infectious disease epidemiology [97].

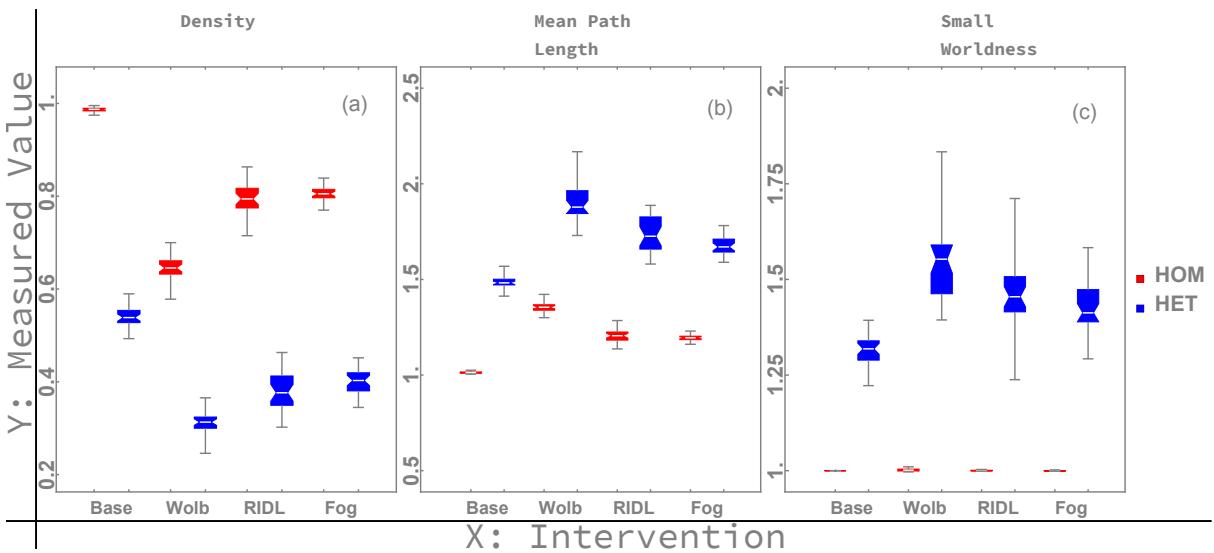


FIGURE 6.8: Networks measures on both scenarios undergoing control interventions.

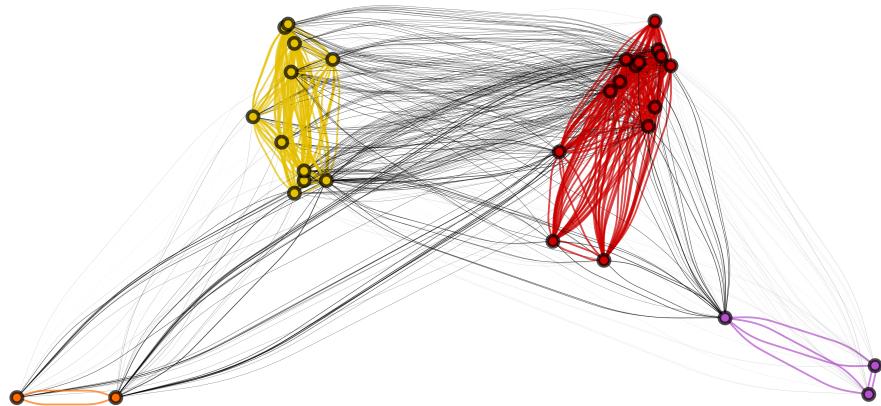


FIGURE 6.9: Network communities example. The colour of the nodes represents the community identity as identified by spectral clustering. The homogeneous case contains only one community (omitted) while the heterogeneous one shows that the spatial distribution affects the frequency of transitions and ultimately how the communities are structured. Applying spectral clustering returned the same result on all the interventions' networks.

Moving on to the degree distributions, the PDF showed a flatter, more platykurtic shape (and the CDF a leaner slope); in the heterogeneous settings than the homogeneous ones (dashed lines in figure 6.7). Lower network densities and lower connectivities are also related to this effect (figures 6.6a and 6.6b), which are relevant because the networks' capabilities to transmit diseases less robust. Furthermore, all of the interventions' distribution curves shifted towards the left mainly as a result of the general decrease of the number of transitional bites between individuals. This result is confirmed by the networks densities calculations shown in figure 6.6a in which in every setting has a lower number of transitions on the heterogeneous spatial setting (due to mosquitoes taking more time finding hosts and producing a higher number of self-loops which were discarded according to our proposed methodology).

### 6.2.3 Interventions Effects

As the last part of our results description, we will briefly describe the differences on the effects of the control interventions on both the population sizes and networks structures. It should be noted, though, that this is not intended to be a thorough description of the differences between the effects of mosquito control interventions. More variables would be needed to do an analysis of such nature (such as: release distributions, efficacy uncertainties, weather effects, etcetera), but we can make some general assertions of what to expect in a broad sense with the experiments we performed as part of this work. As in previous sections, we will first describe effects on population dynamics and then move on to the networks analyses.

#### 6.2.3.1 Population Dynamics

We can observe on figure 6.3, fogging rapidly decreased population sizes from the moment of first application, but that this decrease quickly stabilised to a new equilibrium point after a few treatment repetitions. RIDL releases, on the other hand, showed slower initial decrease of female population size but achieved near total population suppression. In the case of *Wolbachia*, both male and female populations grew, as mixed releases are required due to cytoplasmic incompatibility's transmission mechanism (it is important to note though, that with *Wolbachia* the goal is not so much to eliminate the mosquitoes population as it is to achieve fixation of the pathogen; so its effects are better described by examining its effect on the vectorial contact networks).

As a consequence of this analysis we can say that, in the face of sudden epidemic episodes, fogging might be a viable alternative towards quickly reducing the opportunities for the pathogen to spread. To combat endemic pathogen transmission, however, we would want to shift towards either the eradication of mosquitoes through RIDL or the fixation of *Wolbachia* to disrupt pathogen transmission. This can be done after reducing population sizes through more traditional approaches such as source reduction or fogging (which falls in line with how these two interventions are usually applied on the field or designed to work [12, 13]).

### 6.2.3.2 Vectorial-Contact Networks

Networks were sparser in all the heterogeneous scenarios with respect to the homogeneous ones, but *Wolbachia* produced the largest effect overall on lowering their densities and degrees (figures 6.6a and 6.7). Despite this, it is interesting to point out the behaviour of its small-worldness value. Although *Wolbachia* showed good qualities in disrupting the transmission network (along with higher mean path length values, as shown in figure 6.8b), it also showed the highest small-world coefficient; which would imply that the mean path length of the network would scale as the logarithm of its number of vertices, keeping the persons epidemiologically “close” to each other even while human population grew in number (given that they scale in similar spatial and behavioural patterns).

In terms of degree probability distributions, we can see the emergence of several interesting behaviours. The baseline scenario presented more heterogeneity in the transitional biting behaviour between landscapes (a flatter shape on the distribution on figure 6.7a). This is probably due to the fact that more mosquitoes were able to survive and create some sporadic long distance transitions between humans (effects which are dampened in the cases where the interventions are applied). RIDL managed to reduce its PDF peak to a lower value than fogging and *Wolbachia*; while the latter was the one with the largest change between spatial settings. In general terms, a more heterogeneous the number of bites would mean that the bites are concentrated amongst certain individuals in the network, individuals which could be targeted to reduce diseases’ spread (by using it’s centrality as a proxy measure of this “importance” in the epidemiological structure). Taking this into account RIDL could be the intervention with greater effect, although more repetitions would probably be required to make the distribution frequencies converge into more stable shapes for definite conclusions to be made.

### 6.3 Discussion

After comparing mosquito population dynamics and vectorial-contact networks it is evident that while overall population counts are useful to produce rough estimates of expected level of potential disease transmission they are less useful to examine how a pathogen can spread through a host population in a spatially heterogeneous scenario. This insufficiency of mean population counts to provide useful information in the face of spatial and other heterogeneities becomes even more evident when considering the effect of vector control interventions. Vectorial contact networks, on the other hand, are able to give a precise mathematical description of how an *Aedes*-borne pathogen might spread in a spatially distributed host population. However, calculation of these vectorial-contact networks in the field is operationally unfeasible, motivating our proposal to use detailed agent-based simulations to further our understanding of how epidemic processes may occur on real landscapes.

The vectorial-contact transition matrices derived from our simulations provide a precise mathematical description of how hosts are epidemiologically connected through vector contact. These matrices therefore give detailed individual level form of classic transmission metrics such as  $R_0$  and vectorial capacity [17]. While under certain limiting circumstances transmission dynamics could be well described by mean-field approximations based on systems of ordinary differential equations, finite population sizes, heterogeneous biting, and spatial aggregation patterns found in real transmission settings might invalidate these mean-field assumptions. Sophisticated mathematical techniques such as spatial moment-equations could be used to incorporate spatial effects into a deterministic model of transmission, assumptions must still be made in order to keep the models analytically tractable. Especially in settings characterised by heterogeneities of host behaviour and spatial distribution, as well as small population sizes, commonly encountered in residual transmission scenarios, it is paramount to capture emergent properties of the transmission dynamics which highly depend on these peculiarities of the setting. In these cases agent-based simulation provides an effective means by which transmission dynamics on real landscapes can be easily simulated and analysed.

Results from our spatially-explicit agent-based simulations strongly indicate that heterogeneous spatial distribution of hosts and mosquito breeding sites greatly impacts how a pathogen may invade a human population when mediated by *Aedes* mosquitoes. These differences in the epidemiological relations between individuals is clear from figures 6.5 and 6.7 where the inclusion of spatial heterogeneity produced drastically different epidemiological settings. While in all

cases the spatially homogeneous scenario produced the worst-case scenario across all interventions (figures 6.6 and 6.7) this simulated scenario is of marginal use in planning vector or host based interventions in the field. In many cases interventions aimed to mitigate the worst-case scenario may be much more costly and have much less impact per dollar spent than a targeted intervention informed through analysis of realistic spatially heterogeneous simulations. Furthermore, analyses of simulations under the assumption of spatial homogeneity lose relevance when considering scenarios of low-prevalence and residual transmission. The importance of spatial distribution to vectorial-contact may be observed in figure 6.9. The network structure can be observed to be characterised by several dense clusters of individuals that correlated strongly with the spatial distribution of hosts and breeding sites. This partitioning of the network into tightly connected clusters suggests that vector-borne pathogens can spread efficiently within clusters. These clusters may support residual pathogen transmission and provide a reservoir for disease even when other clusters or areas of the terrain are successfully targeted by transmission control campaigns. This is most relevant when considering elimination scenarios because these pockets of disease provide the pathogen a means of persistence and possible re-emergence even if inter-cluster transition probability is low (due to the small-world nature of the contact networks in the heterogeneous setting as shown in figure 6.6c).

With respect to comparison of different vector control interventions applied to the spatially heterogeneous scenario there was no evidence of substantial differences that were only attributable to spatial effects, although one notable difference was the small-worldness of the vectorial-contact networks, which seemed to present different behaviour on each intervention, shown in figure 6.6c. Most of the variance in calculated measures when compared between interventions can be solely attributed to reduction in population size (this can be seen in figure 6.6; measures closely follow the distributions in figure 6.7 which themselves strongly depend on overall vector population density). This is to be expected according to our experimental design where vector control interventions were applied uniformly to the simulated landscape. We simulated the somewhat unrealistic assumption of uniform application of interventions in order to compare their effects on population density and network measures without potential confounding from spatial distribution of the interventions themselves. In future research we plan to perform a more thorough analysis of each intervention including realistic spatial applications (targeting mosquito mating swarms, hotspot releases, rearing releases, etcetera). Analysing the interventions under realistic operational constraints should provide a better picture of how vector control interventions can

be targeted to take advantage of spatial heterogeneity in host distribution with respect to specific properties of each intervention to maximise their impact on fragmenting vectorial-contact networks.

We performed the aforementioned analysis to demonstrate the importance of acknowledging spatial distribution of hosts and breeding sites when planning vector-control interventions for *Aedes*-borne pathogens. We note however, that much work is still required to produce definite conclusions of how disease spread is affected by spatial heterogeneities. In particular, we plan on extending our model to accommodate larger human population sizes, more realistic mosquito-control releases, data-informed human movement and pathogen models; to be able to make location-specific analyses on how to control epidemic processes efficiently.

As a final remark, we should re-iterate that at this stage of our research we are interested in the networks that emerge from the biting patterns in human populations. If we want to take our analysis a step further, however, we would like to be able to perform epidemiological analysis of different spatial scenarios. To do it, we could take one of two approaches. The first one would be to include the pathogens within the ABM while keeping track of their spread amongst humans. The second one would be to simulate the pathogens' spread on the vectorial-contact networks that result from the ABM's simulations. This second approach is part of another project currently taking place in our research group (described in appendix C); and it has the benefit of being less computationally expensive in comparison to the first one. In this project, we would like to further our understanding of how different network structures affect the dynamics of epidemiological processes so by simulating epidemics on human network contacts we could calculate relevant measures such as infection frequencies, prevalence, incidence, entomological inoculation rate, etcetera. As it might be inferred by now, this project would benefit greatly from the conclusions drawn from our ABM so it might be thought of as a companion, but fully independent, way to interconnect mosquito bite dynamics to classic epidemiological models (susceptible-infected-recovered, for example) on vectorial-contact networks.

## 6.4 Conclusions

Understanding the effects of spatial heterogeneity in mosquito-borne diseases is a difficult task, but with the use of agent-based models and network theory we have shown that it has a significant effect on how humans connect to each other through *Aedes aegypti* mosquito bites both in

absence and in presence of three different mosquito-control interventions. This highlights not only the fact that spatial heterogeneity is an extremely important element of the transmission of mosquito-borne diseases, but also the need of new tools to further our understanding of the implications and effects it has on epidemic processes and vector-control interventions.

These initial conclusions are meant to serve as a guide for future research, as much work is still needed to get a bigger picture of how these heterogeneous contacts dynamics emerge from human-mosquito interactions; and how to take advantage of them to limit diseases spread. In particular, we want to simulate larger geographical regions with more realistic behaviours both in human behaviour and in weather patterns, to have a more robust model of the networks that result as a consequence of their interactions with mosquitoes. We are also planning on making a more thorough analysis of how spatial heterogeneity in the application/release of vector-control interventions affects the contact-networks. All of these analyses would help us move towards the efficient use of the limited resources dedicated to the eradication of often neglected tropical diseases transmitted by *Aedes aegypti* mosquitoes.

# Chapter 7

## Conclusions and Future Work

*“Fall in love with some activity, and do it! Nobody ever figures out what life is all about, and it doesn’t matter. Explore the world. Nearly everything is really interesting if you go into it deeply enough. Work as hard and as much as you want to on the things you like to do the best. Don’t think about what you want to be, but what you want to do. Keep up some kind of a minimum with other things so that society doesn’t stop you from doing anything at all.”*

*Richard Feynman*

As it was shown throughout this work, human-mosquito interactions are complex as they involve a large number of variables and non-linear relations. This is partly because mosquitoes’ life cycle involves several stages and partly because it depends on so many different phenomena and behaviours (spatial distribution, weather, movement, nutrients, etcetera). Throughout this research work we have presented and described *SoNA3BS*; a model created to serve as an *in silico* testbed for the use of computational methods applied to a biological and epidemiological context. In specific, we created an artificial life model capable of analysing the impact of human movement through the use of network theory in small-scale scenarios. This model was created with as much biological accuracy as possible to ensure that the adult population sizes and behaviours matched their mosquito real-life counterparts.

We started by describing thoroughly the way mosquitoes and humans were modelled along with the sources from which these behaviours were obtained and parametrised. Then, we provided a comparison with an independent, published model at the mosquito adults population dynamics levels. This provided us with the level of expected error in the amount of mosquitoes at different

temperatures and population sizes. And finally, we simulated two spatially explicit scenarios and performed population dynamics and network analysis on the results. In these analyses we highlighted the importance of taking into account the spatial aspect of transmission and not only the amount of mosquitoes and humans; and proposed the use of network theory to perform these kind of explorations. We also discussed the epidemiological implications of some of the calculated network measures and how they could be used to understand the transmission of *Aedes*-borne pathogens.

With these experiments we have addressed the hypotheses of our project and showed the value of both: using artificial life models and network theory; in the context of the analysis of spatial heterogeneity in *Aedes aegypti*-borne diseases. In this regard, we can frame the contributions of our project in the context of computer science, within two main stems: an engineering and a scientific application point of view.

**Engineering** From an engineering standpoint, it is easy to see this computational model's usefulness in the analysis of the impact of vector control interventions and spatial relations. With this framework we provide the tools so that people outside the modelling community can evaluate and compare the efficacy of mosquito-control measures; along with the way human and mosquito movement affect the epidemiological connections between individuals. The value of this kind of study should not be underestimated as much of the interest of the health and epidemiological communities is currently focused on studying these specific kinds of relations [52, 94, 103–105].

**Scientific** From a scientific point of view, this modelling work provides the tools to understand the intricacies of a complex biological phenomenon through the use computational methodologies. This is done by translating the results of simulated biological processes to an equivalent representation in the graph-theory domain. By mapping the problem of human-mosquito interactions into a network equivalent, we were able to take it to a higher level of abstraction in which we could study it with tools that have proven useful in other applications (such as transportation, electrical circuits design, maps colouring, etcetera). This is important not only for epidemiological purposes but also from a computational perspective, as the study of how a particular system can be transformed into an equivalent one, is a common and often desired approach [106, 107].

## Final Remarks

Tropical countries currently face a large amount of challenges while trying to eliminate mosquito-borne diseases. Throughout the latest decades it has become quite apparent that traditional vector-control campaigns and reactive control interventions alone are unlikely to stop recurring epidemics from happening [9]. It is also clear that we have to take into account both the mosquitoes and human behaviours [18, 38], along with other number of phenomena such as: pathogens mutation, insecticide resistance evolution and shifts in weather patterns; to be effective in the control of pathogens' spread. This forces us to acknowledge that it is unrealistic to think that any one field of study alone can address the questions we need to answer to fight the spread of these terrible diseases. We need new, integrated, approaches that take into account the full context of the interactions that promote the transmission of these pathogens [36]; and we need innovative, multidisciplinary initiatives that take advantage of the strengths of various fields of study. This is why, we believe it important to develop the kind of projects such as the one presented in this document. By combining knowledge from entomology, epidemiology, computer science and mathematics; we can perform complex analysis that helps us inform health institutes in their decision making, which is still a pressing need in the larger scheme of global health [108, 109].

## 7.1 Future Work

This thesis represents the work performed throughout five years of research. However, as it can be perceived through reading this document, it has focused largely on the creation, parametrisation and documentation of the model. Our long-term goal is for this simulation to be useful and to use more computational techniques that, we believe, would greatly benefit the epidemiology community. In this spirit, we are contemplating a number of future lines of research to take advantage of the system we have created.

**Refinement of Population Dynamics and Further Validation** As it was discussed in the validation experiments, we need to improve the accuracy of the population dynamics in our simulation for it to scale up more realistically. For this, we need to include better approximations to density-dependent behaviours, calibrate our aquatic stages models with field data, handle the

carrying capacity on a per-container basis, and include two adult gonotrophic cycles. This is a non-trivial task but one that must be done for our model to make more trustworthy predictions.

Another step we would like to take towards improving the reliability of our model would be to compare our predictions to field data that is available in the literature. One of these approaches could be to use epidemiological measures such as the entomological inoculation rate, vectorial capacity and prevalence (amongst others). This would involve making some enhancements on our system (such as implementing pathogens and optimising to make larger scale simulations viable) but would greatly improve the quality of our predictions and conclusions that could be drawn from the system.

**Optimisation** One of the ultimate goals of these kind of modelling projects is to reduce the transmission of diseases whilst making the most effective use of available resources. This is a situation in which we can definitely take advantage of computational optimisation techniques. With this into account, we would like to explore optimisation techniques applied to our ABM. We have performed some initial exploration of this by using *BehaviorSearch* (a platform developed in *Java* that connects to *NetLogo* and implements some common optimisation techniques such as simulated annealing and genetic algorithms) but these experiments have been merely to test the frameworks and general ideas. In the future we would like to pursue this endeavour further to examine combinations of vector-control interventions that could help reduce pathogens spread. This falls in line with current elimination efforts that focus on integrated vector management as the most promising approach to eliminate pathogens from human populations (as mentioned in chapter 2.3.3); and which oftentimes use modelling efforts to inform and guide the viability of different approaches [35, 94, 110].

**Human Movement** Another phenomenon we would like to explore is the effect of human movement on the structure of the vectorial-contact networks. In figure 7.1 we present a preliminary example in which we vary the human's visiting probability per day from zero to one and record the networks that result from this parameter sweep. Although this is just an exploratory experiment, we can observe that the networks properties change as a consequence of the change in human behaviour. In the future we would like to explore this idea further and perform proper experiments which allowed us to come to definite conclusions on these subject. Additionally, this is also one of the goals of a project that we are currently developing in our bioinformatics group, in which we analyse social networks interactions to obtain information on how humans

relate to one another in the real-world to gain insight on how epidemics could spread amongst them (this project will be described in more detail in appendix C).

## Human Movement Effects of Vectorial-Contact Networks

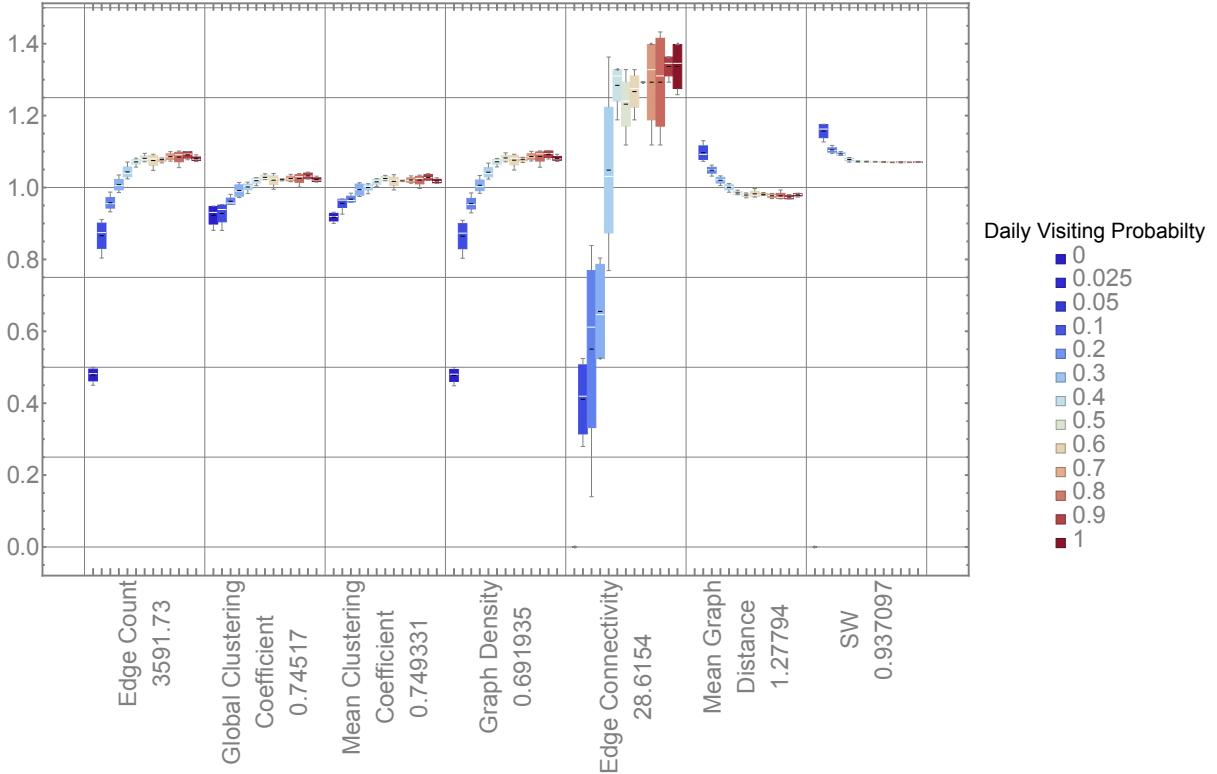


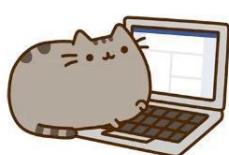
FIGURE 7.1: Example of a preliminary analysis on human movement. In this figure we observe the changes in the contact networks as we increase the humans probabilities of visiting other houses based on the heterogeneous spatial setting shown in chapter 6. We can observe the decrease in mean path length of the contacts between humans and the increase of the density of their networks, along with other graph measures that might be relevant in epidemiological analysis.

**Temporal Analysis** Another aspect we would want to explore in depth is the the way temporal aspects of seasonality and control interventions might be used in our favour to optimise the use of economic resources. In the future, we would like to test both seasonality and the temporal distance between the applications of interventions to perform computational optimisation techniques to reduce diseases spread as efficiently as possible.

**Pathogens** One of the extensions we have to do on our model is the inclusion of *Aedes aegypti*-transmitted pathogens; namely: Yellow Fever, Dengue, Zika and Chikungunya. This is important because, although the transmission networks are a subset of the vector-contacts

networks, we ultimately want to describe the epidemiological implications of spatial analysis and, as such, we would like to provide epidemiological measures such as the basic reproductive number ( $R_0$ ), prevalence, vectorial capacity, etcetera. As it was mentioned earlier, this also forms part of a project currently under develop by our research group which will benefit directly from the work performed so far in this thesis project.

**Preferential Biting** We also want to perform analysis of the effects of the preference of mosquitoes to bite certain humans and what effects this has on transmission. This could be done easily by adding an attractiveness parameter to the humans that made them more prone to be bitten by mosquitoes. Once that is done we can perform sensitivity analysis and response calculations of how this modifies the way the bites networks emerge from these scenarios.



# Appendix A

## Developed Software

As part of this research we developed several pieces of software that, despite being mostly independent and self-sustained, come together as a suite of tools for analyses such as the one proposed in our thesis objectives.

This section is meant to serve as a brief description for each of these software packages. Each of the following are freely available as we believe this to be important for scientific development.

### A.1 SoNA3BS

The main program that is provided as a result of this project is *SoNA3BS*. This computational model was developed under the *NetLogo* platform. This program was thoroughly described throughout the document and, as such, we will only provide the links to the download and project's website in this section.

We should note that we will provide the source code for our simulation and analysis routines; as well as our datasets on our github repository (<https://www.youtube.com/watch?v=DM9TF4IMjIo>) as soon as our first publication goes through peer-review process. This is to ensure replicability of the results presented in this document and for the simulation to be useful to the scientific community.

#### SoNA3BS Information:

- Research Groups: Bioinformatics @ Tecnológico de Monterrey + Modelling @ Berkeley

- Repository: <https://github.com/Chipdelmal/SoNA3BS>
- Licence: Creative Commons Attribution-NonCommercial 4.0 International License
- Project's Website: <http://chipdelmal.github.io/SoNA3BS/index.html>

## A.2 EpiSoNA

This project was created to simulate SIR, SEIR and SIS epidemics on contact networks. It is meant to be a companion package to *SoNA3BS* as it can be used to run epidemics on top of the networks that emerge from it. Despite of this it can be used to run epidemics on randomly generated networks that follow certain degree distributions or that come from experimental data.

### EpiSoNA Information:

- Research Group: Bioinformatics @ Tecnológico de Monterrey
- Repository: <https://bitbucket.org/chipdelmal/episona>
- Licence: Creative Commons Attribution-NonCommercial 4.0 International License
- Project's Website: NA

## A.3 BehaviourSpaceInterface:

*BehaviourSpaceInterface* is a companion package to our main simulation and, as its name suggests, it is meant to be an interface between *NetLogo*'s *Behaviour Space* features and the previously described analysis packages developed in *Mathematica*.

This package was programmed on *Wolfram's Mathematica v9* and its purpose is to translate the exported *XML* data from SoNA3BS's experiments into structures that can be used for mathematical and computational analysis.

### BehaviourSpaceInterface Information:

- Research Group: Bioinformatics @ Tecnológico de Monterrey
- Repository: <https://bitbucket.org/chipdelmal/behaviorspaceinterface>

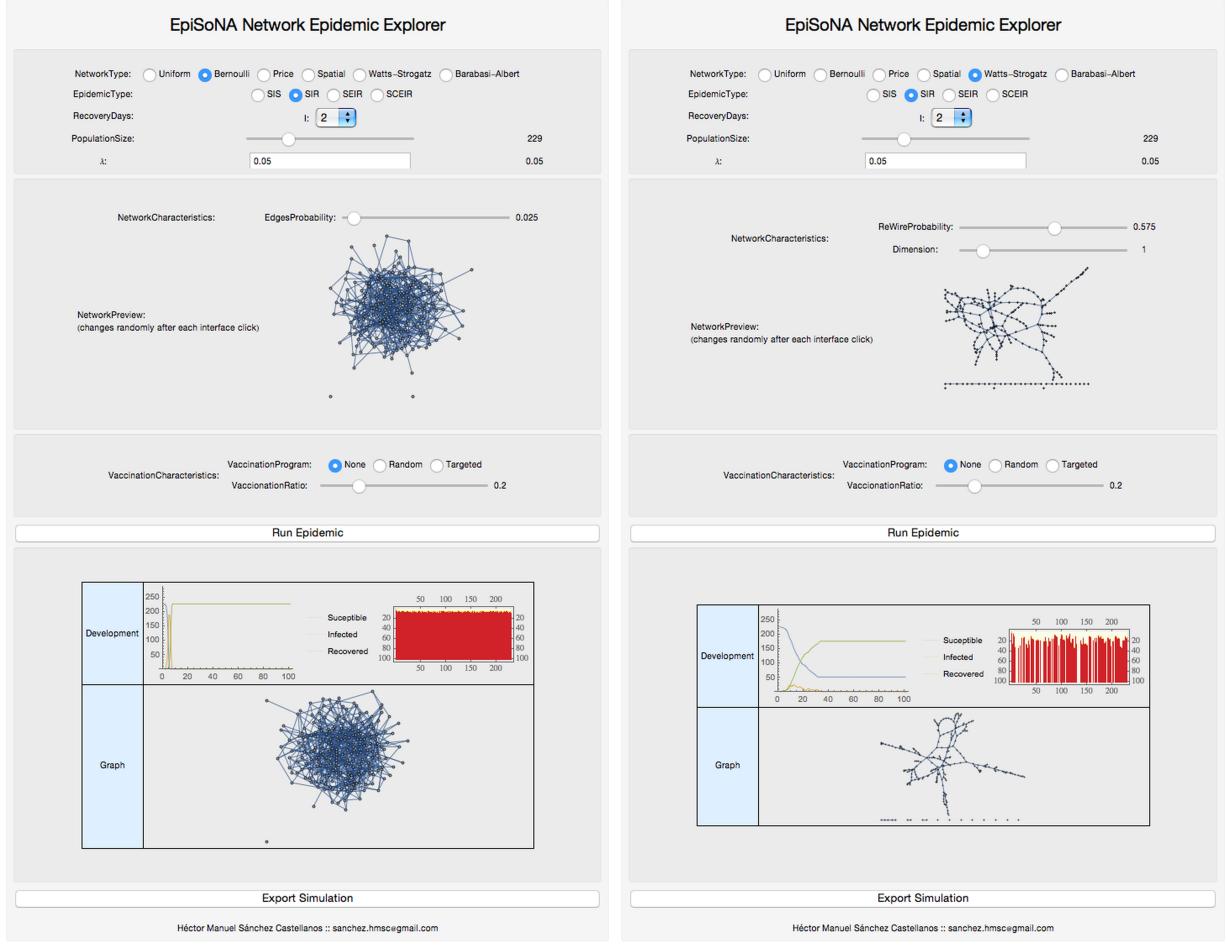


FIGURE A.1: Screenshots of EpiSoNA’s graphical user interface. Network topologies can be easily selected as well as the type of epidemic to be simulated while results can be exported in PNG format. In this examples we can see a Bernoulli topology (A) compared with a small-world one (B).

- Licence: Creative Commons Attribution-NonCommercial 4.0 International License
- Project’s Website: NA

## Appendix B

### PajaroLoco

The development of *PajaroLoco* started being totally independent from *SoNA3BS*. This package was developed by a collaboration between our research group at ITESM and the research group at UCLA to analyse animal vocalisations in the *Mathematica* platform. Shortly after the development started, though, we realised that it could be used for more general networks analyses and visualisations so we continued its development making it as general as possible so that it could be used in this thesis and elsewhere.

As it was thoroughly used in this thesis project we will briefly describe some of *PajaroLoco*'s capabilities in this appendix.

#### B.1 Project's Description

*PajaroLoco* can take lists of elements or transitions matrices and use them to reconstruct transition networks. From this point on network analysis can be performed with the functions provided by *Mathematica* and even with functions provided by the iGraph (by using a *Python* and/or *R* interface included with the package). A graphical user interface is provided along with extensive documentation for ease of use.

In the case of this thesis project, we used to display and analyse all the networks generated in the simulation. Every network and transition matrix in this document was created with the package or some slight extension of its capabilities and a revised version for more general applications (such as epidemiology) is being contemplated for the near future.

### PajaroLoco Information:

- Research Group: Evolutionary Ecology @ UCLA + Bioinformatics @ ITESM
- Repository: <https://bitbucket.org/chipdelmal/pajarolocopublic>
- Licence: GNU General Public License
- Project's Website: <https://sites.google.com/site/pajarolocopublic/>

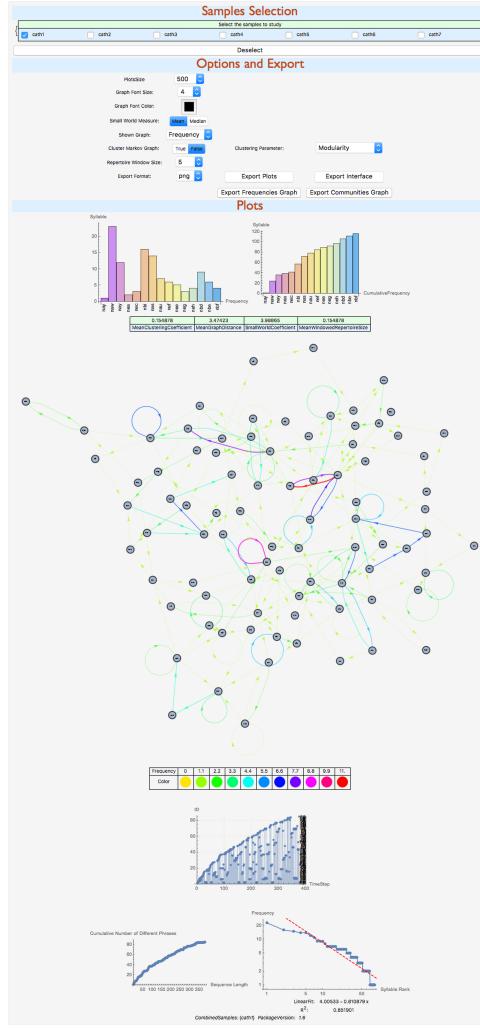


FIGURE B.1: Screenshot of PajaroLoco's graphical user interface.

## B.2 Publications

*PajaroLoco* has been presented in the *Twentieth International Symposium on Artificial Life and Robotics* [101] and *Artificial Life XV* [111] as a poster. Our package has also been used in animal identification publications [112] for syntactic network's features extraction.

## Appendix C

# ZikViSoN: Zika Virus through Social Networks

As it was mentioned throughout this work, it is of great interest to us to further the way we understand epidemic processes on heterogeneous-contacts situations in vector-borne diseases. To further our work in this area, we created *ZikViSoN*. A project oriented towards the study and prediction of Zika epidemics through the use of social media connections.

This endeavour is led by Edgar E. Vallejo and Héctor M. Sánchez C. (Tecnológico de Monterrey) and was awarded one of the Google Research Awards for Latin America in 2016. As the time of the writing of this document this project is being developed by the Bioinformatics group at Tecnológico de Monterrey with involvement of a student from UC-Berkeley (Sean L. Wu).

### C.1 Goal

The goal of this project is to develop a system capable to predict Zika epidemics using social networks data obtained from social networks such as Twitter. The idea behind this concept is that people who interact frequently in virtual spaces might relate to each other in real-life; which, in turn, would provide us with information on how epidemic might spread and be better controlled.

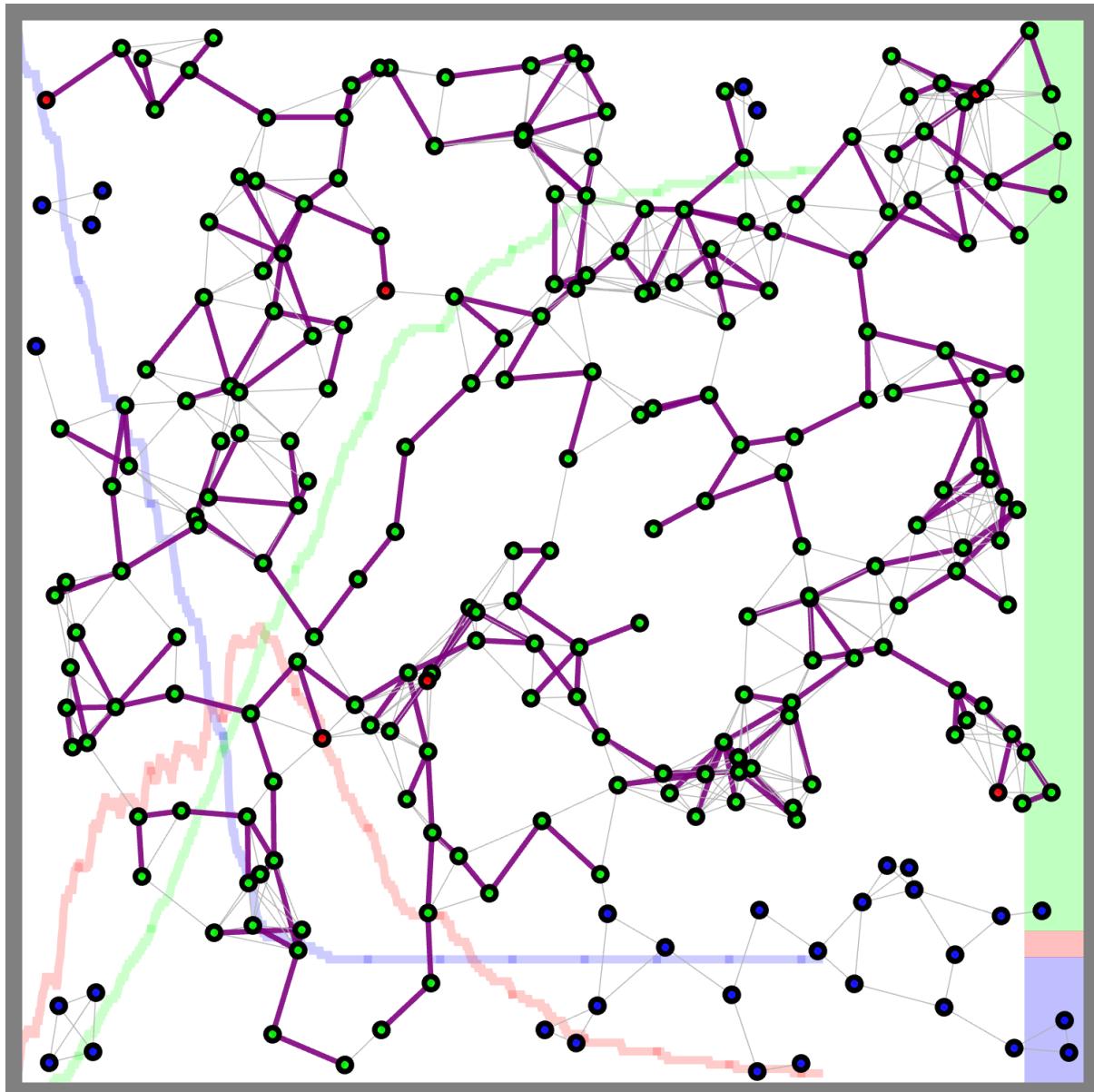


FIGURE C.1: Example of an SIR epidemic on a scale-free network as performed in *ZikViSoN*. Nodes represent persons and connections represent epidemiological interactions between them. Susceptible individuals are represented in blue, infected in red and recovered in green. The network is presented in foreground while the count of individuals is shown in the background.

## Appendix D

# MASH: Modular Analysis and Simulation for human Health

This project is led David L. Smith (University of Washington) as collaboration between researchers at various universities and health institutes. The idea behind the project was, initially, to develop a model that allowed the study and transmission of malaria and other vector-borne diseases in realistic scenarios with varying degrees of complexity. This scope, however, has been extended to provide a framework in which a wide array of diseases can be tested and simulated.

In this appendix, we will briefly describe *MASH* (as the author of this thesis has been thoroughly involved in its development). We will mention some of its capabilities and development principles; along with the similarities and differences with *SoNA3BS*.

### D.1 Project's Description

As it was mentioned, this project is a multidisciplinary endeavour in which a large amount of different agents and sub-models are involved. In this framework, humans, mosquitos, landscape, infectious diseases and interventions; are all being developed independently with modularity in mind so that various types of scenarios and analyses can be performed on the simulation.

With respect to vector-borne disease modelling, this work's is not unlike the one presented in this thesis project in the sense that it models explicit human-mosquito interactions to study

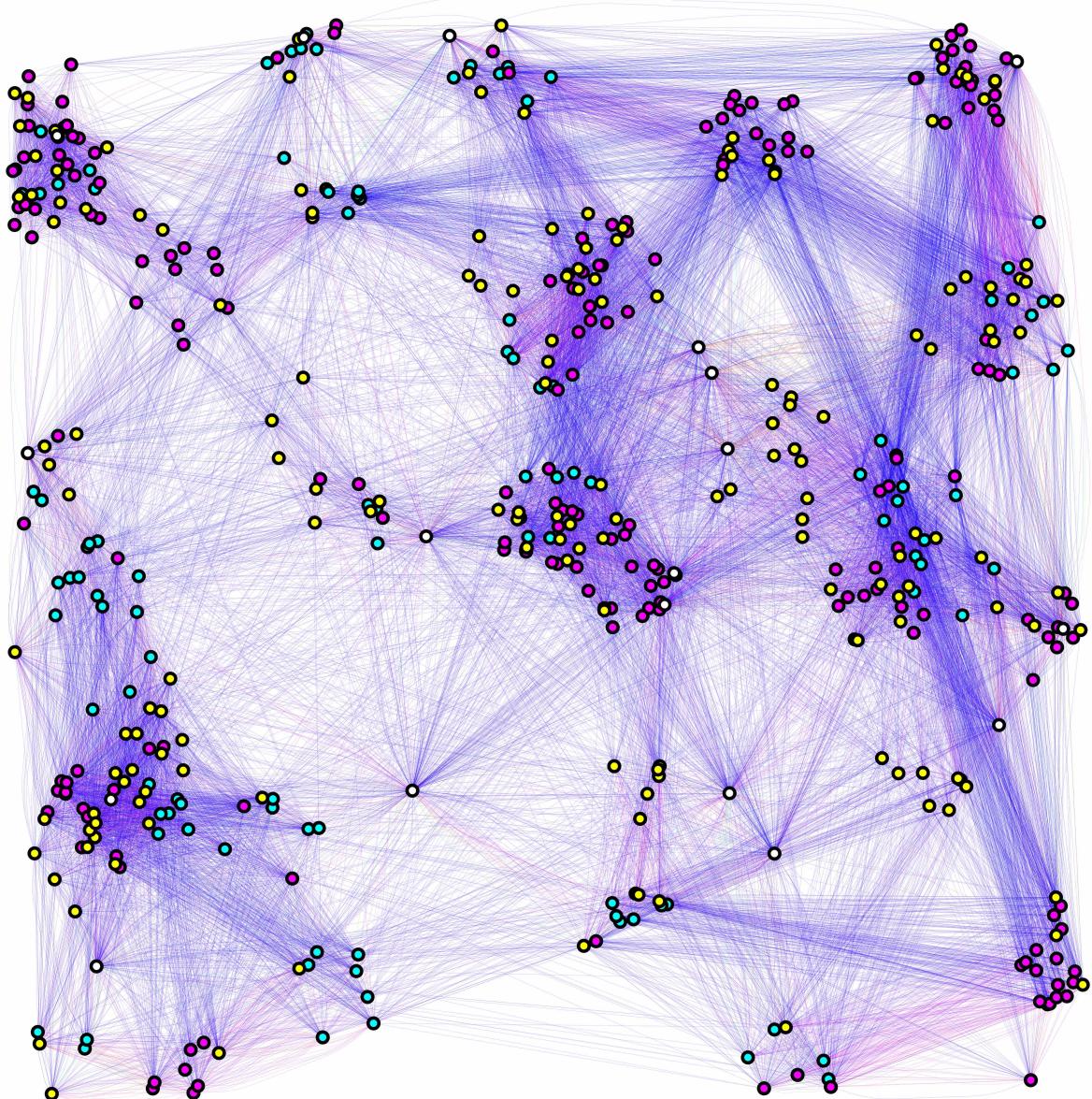


FIGURE D.1: Visual example of movement kernels used in MASH. Vertices represent different types of sites: laying (blue), sugar (yellow) and blood feeding (purple); and the connections the probabilities of moving towards another site (the colour represents the strength of the probability).

the way various heterogeneous behaviours affect pathogen's spread. However, this framework is different to *SoNA3BS* in modelling principles; some of which will be described in this section.

### D.1.1 Event-Driven ABM

*SoNA3BS* runs in discrete time with every agent interacting with each other in real-time. This means that in each “tick” of the simulation every agent performs an action. In *MASH*, in

contrast, agents' actions are triggered by probability functions with exponential shapes that determine when the next activity will take place (this idea is similar to an adaptation of the Gillespie algorithm for agent-based models [113]). This is computationally important because this modification using exponentially distributed event queues only takes one cycle to generate a random process time instead of evaluating a binomial event at each tick of the simulation, which would make it more scalable for the simulation of larger epidemiological settings.

### D.1.2 Movement Kernels

A second difference between the two models is that in *SoNA3BS*, agents are able to move in a continuous way between each site (both humans and mosquitoes). In contrast, *MASH* permits the movement of agents only between predefined sites. While this is a quite obvious simplification of the real movement of mosquitos, it is a useful one as it allows us to reduce the computing requirements substantially. This is because mosquitos move from site to site according to the probabilities dictated by pre-computed movement kernels. One example of this can be seen in figure D.1.

### D.1.3 Scope

Being developed by a larger group of experts, the reach and scope of this project are expected to be large. As it was mentioned before, this framework focuses largely on human individuals so that it is designed to answer a wide array of health-related questions (one of them being vector-borne diseases). In a sense, *MASH* is human-centric to accommodate behaviours such as reactions to diseases (such as resting or going to the doctor), travelling, diseases interactions, etcetera.

## Appendix E

# VCOM: Vector Control Optimization Model

*VCOM* is an ODE model designed for the quick and thorough evaluation of the effectiveness of mosquito-control interventions at a large scale. It is being developed as part of the Malaria Elimination Initiative project at University of California, San Francisco.

One of *VCOM*'s advantages over other available models is that it provides a platform in which most of the current *Anopheles*-control interventions efficacies can be tested. Modelled interventions are: insecticide treated nets, indoor residual spraying, ovitraps, larviciding, biological control, source reduction, house improvement, odor-baited traps, spatial repellants, personal protection and endocticide. The idea behind this is to take advantage of population control interventions that focus on different stages of *Anopheles* life cycle. In this way, elimination campaigns could make the most of probable synergistic interactions between them.

### E.1 Project's Description

*VCOM* is a differential equations model capable of calculating various relevant epidemiological measures such as: entomological inoculation rate (EIR), basic reproduction number ( $R_0$ ), vectorial capacity (VC). It can also output the susceptible, infected, exposed individuals (both human and mosquito) and total amount of mosquitos in each life-stage.

As it was mentioned earlier, one of *VCOM*'s objectives is to provide an accessible platform for people involved with global health to explore combinations of interventions which could decrease malaria levels in human populations. For this to become a reality, *VCOM* provides two web interfaces in which users with varying degrees of expertise can interact and explore different control measures and levels of coverage. A “basic” interface one provides the user with sliders and text boxes to control coverage levels and starting dates of the interventions. The “advanced” one can import XLS files for a more in-depth control of the intervention coverages, effects and mosquito biological parameters for the simulation of more specific scenarios. Both these GUIs are being programmed for end users with no programming experience in addition to the code which will be readily available for research groups interested in doing more in-depth analysis.

The lead researchers of this work are John M. Marshall (UC-Berkeley) and Samson Kiware (Ifakara Health Institute).

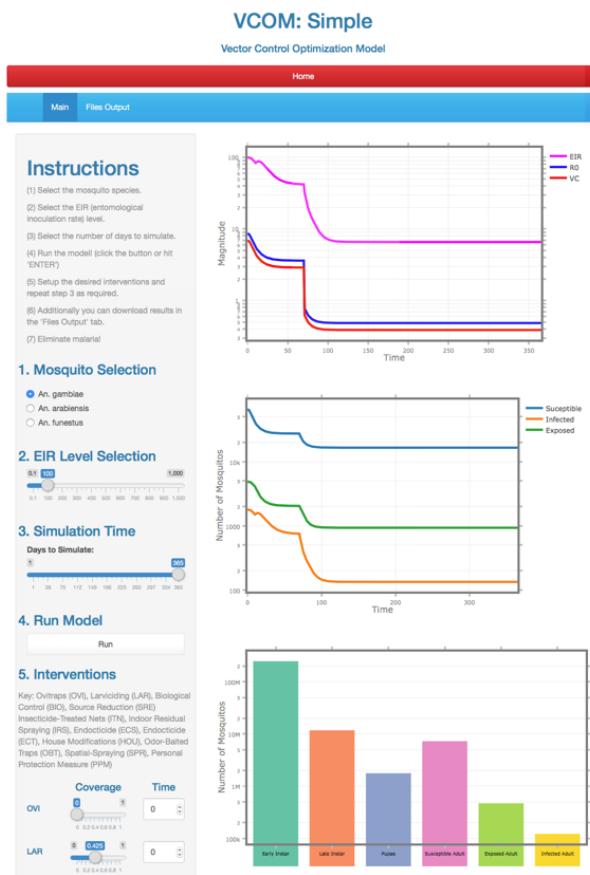


FIGURE E.1: VCOM's web graphical user interface.

## E.2 Publications

The publication of a paper on the journal *PLOS One* is planned in the first half of 2017 and preliminary work has been presented at the *ASTMH 2016* conference (as a poster) [114] and at the second *Annual Malaria Research Conference* [110] (as a talk).

## Appendix F

# CRISPR-Cas 9

Clustered regularly interspaced short palindromic repeats (CRISPR) is a promising genetic engineering technique being studied for population modification and suppression. This collaboration, led by Omar S Akbari (UC, Riverside), focused on CRISPR-Cas9-based homing system for *Anopheles gambiae* population suppression, by inducing fertility reductions into female mosquito populations while keeping inheritance rates high to compensate for the decrease in reproduction rate.

### F.1 Project's Description

One problem that arises from current CRISPR-Cas9 designs for population suppression is the emergence of homing-resistant alleles (this is, alleles that nullify the intervention's effect and that could spread rapidly throughout the populations due to their fitness advantage over population-suppressing homing alleles). To lower the probability of these events from happening, a homing system architecture in which guide RNAs are multiplexed, is proposed; which would result in a more robust system because of the redundancy that would be introduced to the system.

In this work we evaluated the effect of this intervention through the use of a discrete population, stochastic framework incorporating density-dependence at the larval stage. The rationale behind the use of a stochastic model is that we need to capture random events that could lead to population crashes near the crash threshold levels. The model incorporates the egg, larval, pupal and adult life stages with overlapping generations.

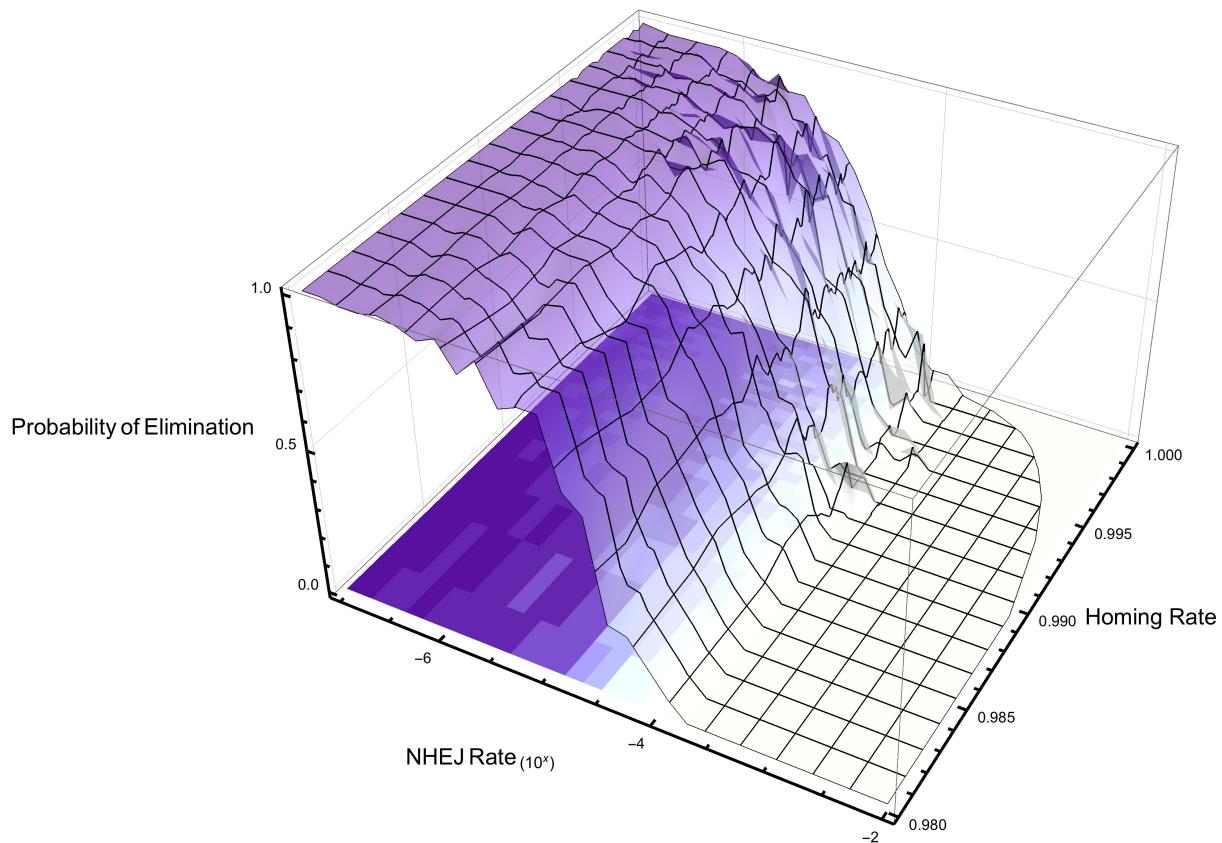


FIGURE F.1: Dependence of population elimination probability on homing rate and resistant allele generation rate. In this example we show the population elimination probability as a function of the homing rate and the resistant allele generation rate (NHEJ).

## F.2 Publications

This work with is currently under revisions for publication in *Nature Scientific Reports*. A preliminary version of a manuscript describing this model in detail can be found in: <http://www.biorxiv.org/content/early/2016/12/07/088427>.

# Bibliography

- [1] María G. Guzman and Gustavo Kouri. Dengue and dengue hemorrhagic fever in the Americas: lessons and challenges. *Journal of Clinical Virology*, 27:1–13, 2003.
- [2] Jennifer L. Kyle and Eva Harris. Global spread and persistence of dengue. *Annual Review of Microbiology*, 62:71–92, 2008.
- [3] Samir Bhatt, Peter W. Gething, Oliver J. Brady, Jane P. Messina, Andrew W. Farlow, Catherine L. Moyes, John M. Drake, John S. Brownstein, Anne G. Hoen, Osman Sankoh, Monica F. Myers, Dylan B. George, Thomas Jaenisch, William Wint, Cameron P. Simmons, Thomas W. Scott, Jeremy J. Farrar, and Simon I. Hay. The global distribution and burden of dengue. *Nature*, 496(7446):504–7, apr 2013.
- [4] Lin H. Chen and Davidson H. Hamer. Zika Virus: Rapid spread in the western hemisphere. *Annals of Internal Medicine*, 164(9):613–615, 2016.
- [5] Katherine B. Gibney, Marc Fischer, Harry E. Prince, Laura D. Kramer, Kirsten St George, Olga L. Kosoy, Janeen J. Laven, and J. Erin Staples. Chikungunya fever in the United States: A fifteen year review of cases. *Clinical Infectious Diseases*, 52:4–9, 2011.
- [6] Edward W. Campion, Scott C. Weaver, and Marc Lecuit. Chikungunya Virus and the Global Spread of a Mosquito-Borne Disease. *New England Journal of Medicine*, 372(13):1231–1239, 2015.
- [7] Tikki Pang. SAGE committee advice on dengue vaccine. *The Lancet Infectious Diseases*, 16(8):880–882, 2016.
- [8] Maíra Aguiar, Nico Stollenwerk, and Scott B. Halstead. The risks behind Dengvaxia recommendation. *The Lancet Infectious Diseases*, 16(8):882–883, 2016.
- [9] Elizabeth McGraw and Scott L. O'Neill. Beyond insecticides: new thinking on an ancient problem. *Nature Reviews. Microbiology*, 11(3):181–93, mar 2013.

- [10] Luciano A. Moreira, Iñaki Iturbe-Ormaetxe, Jason A. Jeffery, Guangjin Lu, Alyssa T. Pyke, Lauren M. Hedges, Bruno C. Rocha, Sonja Hall-Mendelin, Andrew Day, Markus Riegler, Leon E. Hugo, Karyn N. Johnson, Brian H. Kay, Elizabeth A. McGraw, Andrew F. van den Hurk, Peter A. Ryan, and Scott L. O'Neill. A Wolbachia symbiont in *Aedes aegypti* limits infection with dengue, Chikungunya, and Plasmodium. *Cell*, 139(7):1268–78, dec 2009.
- [11] Megan Wise de Valdez R., Derric Nimmo, John Betz, Hong-Fei Gong, Anthony A. James, Luke Alphey, and William C. Black. Genetic elimination of dengue vector mosquitoes. *Proceedings of the National Academy of Sciences of the United States of America*, 108(12):4772–5, mar 2011.
- [12] A. A. Hoffmann, B. L. Montgomery, J. Popovici, I. Iturbe-Ormaetxe, P. H. Johnson, F. Muzzi, M. Greenfield, M. Durkan, Y. S. Leong, Y. Dong, H. Cook, J. Axford, A. G. Callahan, N. Kenny, C. Omodei, E. A. McGraw, P. A. Ryan, S. A. Ritchie, M. Turelli, and S. L. O'Neill. Successful establishment of Wolbachia in *Aedes* populations to suppress dengue transmission. *Nature*, 476(7361):454–7, aug 2011.
- [13] Angela F. Harris, Derric Nimmo, Andrew R. McKemey, Nick Kelly, Sarah Scaife, Christl A. Donnelly, Camilla Beech, William D. Petrie, and Luke Alphey. Field performance of engineered male mosquitoes. *Nature Biotechnology*, 29(11):1034–1037, 2011.
- [14] Robert C. Reiner, T. Alex Perkins, Christopher M. Barker, Tianchan Niu, Luis Fernando Chaves, Alicia M. Ellis, Dylan B. George, Arnaud Le Menach, Juliet R. C. Pulliam, Donal Bisanzio, Caroline Buckee, Christinah Chiyaka, Derek A. T. Cummings, Andres J. Garcia, Michelle L. Gatton, Peter W. Gething, David M. Hartley, Geoffrey Johnston, Eili Y. Klein, Edwin Michael, Steven W. Lindsay, Alun L. Lloyd, David M. Pigott, William K. Reisen, Nick Ruktanonchai, Brajendra K. Singh, Andrew J. Tatem, Uriel Kitron, Simon I. Hay, Thomas W. Scott, and David L. Smith. A systematic review of mathematical models of mosquito-borne pathogen transmission: 1970–2010. *Journal of the Royal Society, Interface*, 10(81):20120921, apr 2013.
- [15] Terry Lima Ruas, Maria das Gracas, Bruno Marietto, Andre Filipe de Moraes Batista, Robson dos Santos Franca, Alexandre Heideker, Emerson Aguiar Noronha, and Fabio Ara-gao da Silva. Modeling artificial life through multi-agent based simulation. *Multi-Agent Systems*, 2011.

- [16] A.H. Dekker. Network Centrality and Super-Spreaders in Infectious Disease Epidemiology. In *20th International Congress on Modelling and Simulation*, number December, Adelaide, Australia, 2013.
- [17] Miguel A. Acevedo, Olivia Prosper, Kenneth Lopiano, Nick Ruktanonchai, T. Trevor Caughlin, Maia Martcheva, Craig W. Osenberg, and David L. Smith. Spatial heterogeneity, host movement and mosquito-borne disease transmission. *PloS One*, 10(6):e0127552, 2015.
- [18] Steven T. Stoddard, Brett M. Forshey, Amy C. Morrison, Valerie A. Paz-Soldan, Gonzalo M. Vazquez-Prokopec, Helvio Astete, Robert C. Reiner, Stalin Vilcarromero, John P. Elder, Eric S. Halsey, Tadeusz J. Kochel, Uriel Kitron, and Thomas W. Scott. House-to-house human movement drives dengue virus transmission. *Proceedings of the National Academy of Sciences of the United States of America*, 110(3):994–9, jan 2013.
- [19] S. Rickard Christophers. *Aedes Aegypti: The Yellow Fever Mosquito: Its Life History, Bionomics and Structure*. Cambridge University Press, London, 1st edition, 1960.
- [20] Norbert Becker, Dusan Petric, Marija Zgomba, Clive Boase, Minoo Madon, Christine Dahl, and Achim Kaiser. *Mosquitoes and Their Control*. Springer Berlin Heidelberg, Berlin, Heidelberg, 1st edition, 2010.
- [21] Y. M. Colton, D. D. Chadee, and David W. Severson. Natural skip oviposition of the mosquito Aedes aegypti indicated by codominant genetic markers. *Medical and Veterinary Entomology*, 17(2):195–204, 2003.
- [22] P. Reiter, M. A. Amador, R. A. Anderson, and G. G. Clark. Short report: dispersal of Aedes aegypti in an urban area after blood feeding as demonstrated by rubidium-marked eggs. *The American journal of tropical medicine and hygiene*, 52(2):177–9, feb 1995.
- [23] Philip A. Eckhoff, Edward A. Wenger, H. Charles J. Godfray, and Austin Burt. Impact of mosquito gene drive on malaria elimination in a computational model with explicit spatial and temporal dynamics. *Proceedings of the National Academy of Sciences*, page 201611064, 2016.
- [24] W. K. Hartberg. Observations on the mating behaviour of Aedes aegypti in nature. *Bulletin of the World Health Organization*, 45(6):847–850, 1971.

- [25] Sebastien Boyer, Celine Toty, Maxime Jacquet, Guy Lempérière, and Didier Fontenille. Evidence of multiple inseminations in the field in *Aedes albopictus*. *PLoS ONE*, 7(8):1–5, 2012.
- [26] L. M. Rueda, K. J. Patel, R. C. Axtell, and R. E. Stinner. Temperature-dependent development and survival rates of *Culex quinquefasciatus* and *Aedes aegypti* (Diptera: Culicidae). *Journal of medical entomology*, 10:892–898, 1990.
- [27] Mika Oki, Toshihiko Sunahara, Masahiro Hashizume, and Taro Yamamoto. Optimal timing of insecticide fogging to minimize dengue cases: Modeling dengue transmission among various Seasonalities and transmission intensities. *PLoS Neglected Tropical Diseases*, 5(10), 2011.
- [28] Ratana Sithiprasasna, Pradith Mahapibul, Chumnong Noigamol, Michael J. Perich, Brian C. Zeichner, Bob Burge, Sarah L. W. Norris, James W. Jones, Sonya S. Schleich, and Russell E. Coleman. Field evaluation of a lethal ovitrap for the control of *Aedes aegypti* (Diptera: Culicidae) in Thailand. *Journal of medical entomology*, 40:455–462, 2003.
- [29] John M. Marshall, Michael T. White, Azra C. Ghani, Yosef Schlein, Gunter C. Muller, and John C. Beier. Quantifying the mosquito’s sweet tooth: modelling the effectiveness of attractive toxic sugar baits (ATSB) for malaria vector control. *Malaria journal*, 12(1):291, jan 2013.
- [30] Lin Zhu, Whitney A. Qualls, John M. Marshall, Kris L. Arheart, Donald L. DeAngelis, John W. McManus, Sekou F. Traore, Seydou Doumbia, Yosef Schlein, Günter C. Müller, and John C. Beier. A spatial individual-based model predicting a great impact of copious sugar sources and resting sites on survival of *Anopheles gambiae* and malaria parasite transmission. *Malaria Journal*, 14:1–15, 2015.
- [31] Michelle E. H. Helinski, Andrew G. Parker, and Bart G. J. Knols. Radiation-induced sterility for pupal and adult stages of the malaria mosquito *Anopheles arabiensis*. *Malaria journal*, 5:41, 2006.
- [32] Helena R. C. Araújo, Danilo O. Carvalho, Rafaella S. Ioshino, André L. Costa-da Silva, and Margareth L. Capurro. *Aedes aegypti* Control Strategies in Brazil: Incorporation of New Technologies to Overcome the Persistence of Dengue Epidemics. *Insects*, 6(2):576–594, 2015.

- [33] Guowu Bian, Yao Xu, Peng Lu, Yan Xie, and Zhiyong Xi. The endosymbiotic bacterium Wolbachia induces resistance to dengue virus in *Aedes aegypti*. *PLoS Pathogens*, 6(4):e1000833, 2010.
- [34] Andrew Hammond, Roberto Galizi, Kyros Kyrou, Alekos Simoni, Carla Siniscalchi, Dimitris Katsanos, Matthew Gribble, Dean Baker, Eric Marois, Steven Russell, Austin Burt, Nikolai Windbichler, Andrea Crisanti, and Tony Nolan. A CRISPR-Cas9 gene drive system targeting female reproduction in the malaria mosquito vector *Anopheles gambiae*. *Nature Biotechnology*, 34(1):78–83, 2016.
- [35] M. Rafikov, E. Rafikova, and H. M. Yang. Optimization of the *aedes aegypti* control strategies for integrated vector management. *Journal of Applied Mathematics*, 2015, 2015.
- [36] DVI Dengue Vaccine Initiative. Report of the Americas Dengue Prevention Board Meeting. Technical report, México, 2014.
- [37] John M. Marshall, Mahamoudou Touré, André Lin Ouédraogo, Micky Ndhlovu, Samson S. Kiware, Ashley Rezai, Emmy Nkhama, Jamie T. Griffin, T. Deirdre Hollingsworth, Seydou Doumbia, Nicodem J. Govella, Neil M. Ferguson, and Azra C. Ghani. Key traveller groups of relevance to spatial malaria transmission : a survey of movement patterns in four sub Saharan African countries. *Malaria Journal*, pages 1–12, 2016.
- [38] Steven T. Stoddard, Amy C. Morrison, Gonzalo M. Vazquez-Prokopec, Valerie Paz Soldan, Tadeusz J. Kochel, Uriel Kitron, John P. Elder, and Thomas W. Scott. The role of human movement in the transmission of vector-borne pathogens. *PLoS neglected tropical diseases*, 3(7):e481, jan 2009.
- [39] Ben Adams and Durrell D. Kapan. Man bites mosquito: Understanding the contribution of human movement to vector-borne disease dynamics. *PLoS ONE*, 4(8), 2009.
- [40] Nick W. Ruktanonchai, Patrick DeLeenheer, Andrew J. Tatem, Victor A. Alegana, T. Trevor Caughlin, Elisabeth Zu Erbach-Schoenberg, Christopher Lourenço, Corrine W. Ruktanonchai, and David L. Smith. Identifying Malaria Transmission Foci for Elimination Using Human Mobility Data. *PLoS computational biology*, 12(4):e1004846, 2016.
- [41] Robert C. Reiner, Steven T. Stoddard, and Thomas W. Scott. Socially structured human movement shapes dengue transmission despite the diffusive effect of mosquito dispersal. *Epidemics*, 6:30–36, 2014.

- [42] T. Alex Perkins, Thomas W. Scott, Arnaud Le Menach, and David L. Smith. Heterogeneity, mixing, and the spatial scales of mosquito-borne pathogen transmission. *PLoS Computational Biology*, 9(12), 2013.
- [43] Jennifer E. Simpson, Paul J. Hurtado, Jan Medlock, Goudarz Molaei, Theodore G. Andreadis, Alison P. Galvani, and Maria A. Diuk-Wasser. Vector host-feeding preferences drive transmission of multi-host pathogens: West Nile virus as a model system. *Proceedings. Biological sciences / The Royal Society*, 279(1730):925–33, mar 2012.
- [44] Romualdo Pastor-Satorras and Alessandro Vespignani. Epidemic Spreading in Scale-Free Networks. *Physical Review Letters*, 86(14):3200–3203, apr 2001.
- [45] Romualdo Pastor-Satorras and Alessandro Vespignani. Epidemic dynamics and endemic states in complex networks. *Physical review. E, Statistical, nonlinear, and soft matter physics*, 63(6 Pt 2):066117, 2001.
- [46] Robert M. May. Network structure and the biology of populations. *TRENDS in Ecology and Evolution*, 21(7), 2006.
- [47] Alison L. Hill, David G. Rand, Martin A. Nowak, and Nicholas A. Christakis. Infectious disease modeling of social contagion in networks. *PLoS computational biology*, 6(11):e1000968, jan 2010.
- [48] Moez Draief. Epidemic processes on complex networks: The effect of topology on the spread of epidemics. *Physica A: Statistical Mechanics and its Applications*, 363(1):120–131, 2006.
- [49] Mathieu Andraud, Niel Hens, Christiaan Marais, and Philippe Beutels. Dynamic epidemiological models for dengue transmission: a systematic review of structural approaches. *PloS One*, 7(11):e49085, jan 2012.
- [50] David L. Smith, Katherine E. Battle, Simon I. Hay, Christopher M. Barker, Thomas W. Scott, and F. Ellis McKenzie. Ross, Macdonald, and a theory for the dynamics and control of mosquito-transmitted pathogens. *PLoS Pathogens*, 8(4), 2012.
- [51] Marcelo Otero, Hernán G Solari, and Nicolás Schweigmann. A stochastic population dynamics model for *Aedes aegypti*: formulation and application to a city with temperate climate. *Bulletin of mathematical biology*, 68(8):1945–74, nov 2006.

- [52] Samson Kiware, Kiware N. Chitnis, Allison Tatarsky, Sean Wu, Héctor M. Sánchez C., Roly Gosling, David L. Smith, and John M. Marshall. Attacking the mosquito on multiple fronts : Insights from the Vector Control Optimization Model (VCOM) for malaria elimination. *PLoS computational biology (In Review)*, 2017.
- [53] James E. Gentile, Gregory J. Davis, Brandy St. Laurent, and Steve Kurtz. A Framework for Modeling Mosquito Vectors. *Proceedings of the 2010 Summer Computer Simulation Conference*, pages 110–117.
- [54] Ying Zhou, S.M. Niaz Arifin, James Gentile, Steven J. Kurtz, Gregory J. Davis, and Barbara A. Wendelberger. An Agent-based Model of the *Anopheles gambiae* Mosquito Life Cycle.
- [55] S.M. Niaz Arifin, Ying Zhou, Gregory J. Davis, James E. Gentile, Gregory R. Madey, and Frank H. Collins. An agent-based model of the population dynamics of *Anopheles gambiae*. *Malaria journal*, 13(1):424, 2014.
- [56] Carlos Isidoro, Nuno Fachada, Fábio Barata, and Agostinho Rosa. Agent-based model of *aedes aegypti* population dynamics. *Progress in Artificial Intelligence*, pages 53–64, 2009.
- [57] Carlos Isidoro, Nuno Fachada, Fábio Barata, and Agostinho Rosa. Agent-based model of dengue disease transmission by *Aedes aegypti* populations. *Advances in Artificial Life*, 2011.
- [58] Luís F. O. Jacintho, André F. M. Batista, Terry L. Ruas, Maria G. B. Marietto, and Fábio A. Silva. An agent-based model for the spread of the Dengue fever: a swarm platform simulation approach. *Proceedings of the 2010 Spring Simulation Multiconference on - SpringSim '10*, pages 1–8, 2010.
- [59] Chengbin Deng, Haiyan Tao, and Zhiwei Ye. Agent-Based modelling to simulate dengue spread. In *Geoinformatics 2008 and Joint Conference on GIS and Built Environment: Geo-Simulation and Virtual GIS Environments*, volume 7143, pages 71431O–71431O–10, oct 2008.
- [60] Dana A. Focks, E. Daniels, D. G. Haile, and J. E. Keesling. A simulation model of the epidemiology of urban dengue fever: literature analysis, model development, preliminary validation, and samples of simulation results. *The American journal of tropical medicine and hygiene*, 53(5):489–506, nov 1995.

- [61] Krisztian Magori, Mathieu Legros, Molly E. Puente, Dana A. Focks, Thomas W. Scott, Alun L. Lloyd, and Fred Gould. Skeeter Buster: a stochastic, spatially explicit modeling tool for studying *Aedes aegypti* population replacement and population suppression strategies. *PLoS neglected tropical diseases*, 3(9):e508, jan 2009.
- [62] Daniel T. Gillespie. Exact Stochastic Simulation of couple chemical reactions. *The Journal of Physical Chemistry*, 81(25):2340–2361, 1977.
- [63] Gilberto González-Parra, Rafael-J. Villanueva, Javier Ruiz-Baragaño, and Jose-A. Moraño. Modelling influenza A(H1N1) 2009 epidemics using a random network in a distributed computing environment. *Acta tropica*, 143:29–35, 2015.
- [64] M. Small and C. K. Tse. Complex network models of disease propagation: modelling, predicting and assessing the transmission of SARS. *Hong Kong medical journal = Xianggang yi xue za zhi / Hong Kong Academy of Medicine*, 16(5 Suppl 4):43–4, jan 2010.
- [65] M. E. J. Newman. Models of the small world. *Journal of Statistical Physics*, 2000.
- [66] Albert-László Barabási and Réka Albert. Emergence of Scaling in Random Networks. *science*, 286(October):509–512, 1999.
- [67] Marian Boguna, Romualdo Pastor-Satorras, and Alessandro Vespignani. Absence of epidemic threshold in scale-free networks with connectivity correlations. pages 1–4, aug 2002.
- [68] John H. Huber, Geoffrey L. Johnston, Bryan Greenhouse, David L. Smith, and T. Alex Perkins. Quantitative, model-based estimates of variability in the generation and serial intervals of *Plasmodium falciparum* malaria. *Malaria Journal*, 15(1):490, 2016.
- [69] Ben Adams and Durrell D Kapan. Man bites mosquito: understanding the contribution of human movement to vector-borne disease dynamics. *PloS one*, 4(8):e6763, jan 2009.
- [70] Volker Grimm, Uta Berger, Finn Bastiansen, Sigrunn Eliassen, Vincent Ginot, Jarl Giske, John Goss-Custard, Tamara Grand, Simone K. Heinz, Geir Huse, Andreas Huth, Jane U. Jepsen, Christian Jørgensen, Wolf M. Mooij, Birgit Müller, Guy Pe'er, Cyril Piou, Steven F. Railsback, Andrew M. Robbins, Martha M. Robbins, Eva Rossmanith, Nadja Rüger, Espen Strand, Sami Souissi, Richard A. Stillman, Rune Vabø, Ute Visser, and Donald L. DeAngelis. A standard protocol for describing individual-based and agent-based models. *Ecological Modelling*, 198(1-2):115–126, sep 2006.
- [71] Christopher G. Langton. *Artificial Life: An Overview*. MIT Press, 1995.

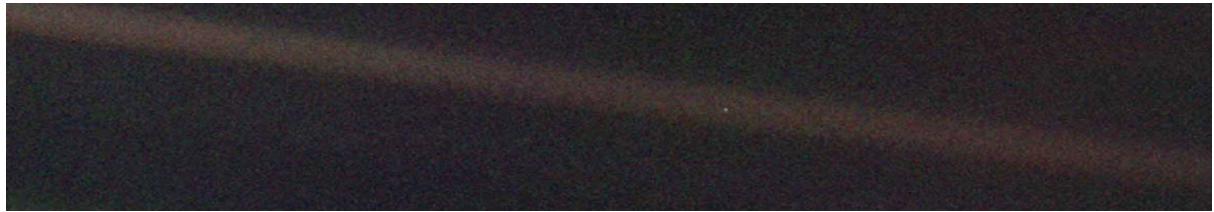
- [72] Angela B Shiflet and George W.. Shiflet. Introduction to Computational Science: Modeling and Simulation for the Sciences. *System*, pages 1–20, 2006.
- [73] Michael Wooldridge. *An Introduction to MultiAgent Systems*. John Wiley, 2009.
- [74] Steven F. Railsback and Volker Grimm. *Agent-Based and Individual-Based Modeling: A Practical Introduction*. Princeton University Press, 2011.
- [75] Volker Grimm and Steven F. Railsback. *Individual-based Modeling and Ecology: (Princeton Series in Theoretical and Computational Biology)*. Princeton University Press, 2005.
- [76] Steven F. Railsback, Steven L. Lytinen, and Stephen K. Jackson. Agent-based Simulation Platforms: Review and Development Recommendations. *Simulation*, 82(9):609–623, sep 2006.
- [77] M. Trpis. Dry season survival of Aedes aegypti eggs in various breeding sites in the Dar es Salaam area, Tanzania. *Bulletin of the World Health Organization*, 47(3):433–437, 1972.
- [78] T. Livdahl. The complex hatching response of Aedes eggs to larval density. *Ecological Entomology*, pages 437–442, 1984.
- [79] Micha Bar-Zeev. The Effect of Temperature on the Growth Rate and Survival of the Immature Stages of Aëdes aegypti (L.). *Bulletin of Entomological Research*, 49(01):157, 1958.
- [80] W. E. Horsfall. *Mosquitoes. Their Bionomics and Relation to Disease*. The Eonalid Press Co., 1955.
- [81] L. M. Rueda, K. J. Patel, R. C. Axtell, and R. E. Stinner. Temperature-dependent development and survival rates of Culex quinquefasciatus and Aedes aegypti (Diptera: Culicidae). *Journal of medical entomology*, 27(5):892–898, 1990.
- [82] Christopher Dye. Intraspecific competition amongst larval Aedes aegypti : food exploitation or chemical interference ? *Ecological Entomology*, 7(1 982):39–46, 1982.
- [83] T. R. Southwood, G. Murdie, M. Yasuno, R. J. Tonn, and P. M. Reader. Studies on the life budget of Aedes aegypti in Wat Samphaya, Bangkok, Thailand. *Bulletin of the World Health Organization*, 46(2):211–226, 1972.

- [84] R. M. Schoolfield, P. J. Sharpe, and C. E. Magnuson. Non-linear regression of biological temperature-dependent rate models based on absolute reaction-rate theory. *Journal of Theoretical Biology*, 88(4):719–31, feb 1981.
- [85] Kingsley E. Haynes and A. Stewart Fotheringham. Gravity and spatial interaction models. (November), 1984.
- [86] Chaogui Kang, Yu Liu, Diansheng Guo, and Kun Qin. A generalized radiation model for human mobility: Spatial scale, searching direction and trip constraint. *PLoS ONE*, 10(11):1–11, 2015.
- [87] S. Rickard Christophers. *Aedes aegypti the Yellow Fever Mosquito*. Cambridge University Press, London, 1st edition, 1960.
- [88] Andrea Király and Imre M. Jánosi. Stochastic modeling of daily temperature fluctuations. *Physical Review E - Statistical, Nonlinear, and Soft Matter Physics*, 65(5):1–6, 2002.
- [89] David L. Smith, T. Alex Perkins, Lucy S. Tusting, Thomas W. Scott, and Steven W. Lindsay. Mosquito population regulation and larval source management in heterogeneous environments. *PloS one*, 8(8):e71247, jan 2013.
- [90] Javier Escobar-Mesa and Héctor Gómez-Dantés. Determinantes de la transmisión de dengue en Veracruz: Un abordaje ecológico para su control. *Salud Publica de Mexico*, 45(1):43–53, 2003.
- [91] Cassie C. Jansen and Nigel W. Beebe. The dengue vector *Aedes aegypti*: what comes next. *Microbes and Infection*, 12(4):272–279, 2010.
- [92] Stephan Karl, Nilimesh Halder, Joel K. Kelso, Scott A. Ritchie, and George J. Milne. A spatial simulation model for dengue virus infection in urban areas. *BMC Infectious Diseases*, 14(1):447, jan 2014.
- [93] Laura C. Harrington, Thomas W. Scott, Kriangkrai Lerdthusnee, Russell C. Coleman, Adriana Costero, Gary G. Clark, James J. Jones, Sangvorn Kitthawee, Pattamaporn Kittayapong, Ratana Sithiprasasna, and John D. Edman. Dispersal of the dengue vector *Aedes aegypti* within and between rural communities. *The American Journal of Tropical Medicine and Hygiene*, 72(2):209–20, feb 2005.

- [94] Mathieu Legros, Chonggang Xu, Kenichi Okamoto, Thomas W. Scott, Amy C. Morrison, Alun L. Lloyd, and Fred Gould. Assessing the feasibility of controlling *Aedes aegypti* with transgenic methods: a model-based evaluation. *PLoS One*, 7(12):e52235, jan 2012.
- [95] Thomas Hladish, Eugene Melamud, Luis Alberto Barrera, Alison Galvani, and Lauren Ancel Meyers. EpiFire: An open source C++ library and application for contact network epidemiology. *BMC Bioinformatics*, 13:76, jan 2012.
- [96] Marc Barthélemy, Alain Barrat, Romualdo Pastor-Satorras, and Alessandro Vespignani. Dynamical patterns of epidemic outbreaks in complex heterogeneous networks. *Journal of Theoretical Biology*, 235(2):275–88, jul 2005.
- [97] Duncan J. Watts and Steven H. Strogatz. Collective dynamics of small-world networks. *Nature*, 393(6684):440–2, jun 1998.
- [98] M. Girvan and M. E. J. Newman. Community structure in social and biological networks. *Proceedings of the National Academy of Sciences of the United States of America*, 99(12):7821–6, jun 2002.
- [99] Y. Moreno, R. Pastor-Satorras, and A. Vespignani. Epidemic outbreaks in complex heterogeneous networks. *European Physical Journal B*, 26(4):521–529, 2002.
- [100] Scott Smyth and White Padhraic. A spectral clustering approach to finding communities in graphs. *Proceedings of the 5th SIAM International Conference on Data Mining*, pages 76–84, 2005.
- [101] Héctor M. Sánchez C., Edgar E. Vallejo, and Charles E. Taylor. PajaroLoco: A suite of programs to study the grammatical structure of bird songs. In *The Twentieth International Symposium on Artificial Life and Robotics*, volume 2015, pages 148–153. ISAROB, 2015.
- [102] Ruth Aralí Martinez-Vega, Rogelio Danis-Lozano, Fredi Alexander Diaz-Quijano, Jorge Velasco-Hernandez, Rene Santos-Luna, Susana Roman-Perez, Pablo Kuri-Morales, and Jose Ramos-Castaneda. Peridomestic infection as a determining factor of Dengue transmission. *PLoS Neglected Tropical Diseases*, pages 1–18, 2016.
- [103] Amy C. Morrison, Emily Zielinski-Gutierrez, Thomas W. Scott, and Ronald Rosenberg. Defining challenges and proposing solutions for control of the virus vector *Aedes aegypti*. *PLoS Medicine*, 5(3):0362–0366, 2008.

- [104] John M. Marshall, Anna Buchman, Héctor M. Sánchez C., and Omar Akbari. Overcoming Evolved Resistance to Population-Supressing Homing-Based Gene Drives. In *ASTMH*, 2016.
- [105] Harriet Hughes and N. F. Britton. Modelling the use of Wolbachia to control dengue fever transmission. *Bulletin of mathematical biology*, 75(5):796–818, may 2013.
- [106] Michael. Sipser. *Introduction to the theory of computation*. Cengage Learning, 1997.
- [107] Peter Bürgisser. *Completeness and Reduction in Algebraic Complexity Theory*, volume 7 of *Algorithms and Computation in Mathematics*. Springer Berlin Heidelberg, Berlin, Heidelberg, 2000.
- [108] Malaria Elimination Initiative. Using mathematical models to optimize vector control. Technical report, University of California, San Francisco, San Francisco, California, 2016.
- [109] Malaria Elimination Initiative. Identifying the drivers and distribution of residual transmission in Africa using geospatial modeling. Technical report, University of California, San Francisco, San Francisco, California, 2016.
- [110] Samson Kware, Allison Tatarsky, Sean Wu, Héctor M. Sánchez C., Nakul Chitnis, and John Marshall. Attacking the Mosquito on Multiple Fronts: Insights on Optimal Combinations of Vector Control Interventions for Malaria Elimination from a Mathematical Model. In *Annual Malaria Research Conference II*, 2016.
- [111] Héctor M. Sánchez C., Edgar E. Vallejo, and Charles E. Taylor. PajaroLoco: A suite of programs to study complex adaptive properties of animal language. An example of a Cassin’s vireo syntax network. In *Artificial Life XV*, pages 578–579, 2016.
- [112] Julio G. Arriaga, Hector M. Sanchez C., Edgar E. Vallejo, Richard Hedley, and Charles E. Taylor. Identification of Cassin’s Vireo (*Vireo cassini*) individuals from their acoustic sequences using an ensemble of learners. *Neurocomputing*, pages 1–14, 2015.
- [113] Christian L. Vestergaard and Mathieu Génois. Temporal Gillespie Algorithm: Fast Simulation of Contagion Processes on Time-Varying Networks. *PLoS Computational Biology*, 11(10):1–28, 2015.

- [114] Samson Kiware, Allison Tatarsky, Sean Wu, Héctor M. Sánchez C., Nakul Chitnis, and John M Marshall. Attacking the Mosquito on Multiple Fronts: Insights on Optimal Combinations of Vector Control Interventions for Malaria Elimination from a Mathematical Model. In *American Society of Tropical Medicine and Hygiene*, 2016.



*"Look again at that dot. That's here. That's home. That's us. On it everyone you love, everyone you know, everyone you ever heard of, every human being who ever was, lived out their lives. The aggregate of our joy and suffering, thousands of confident religions, ideologies, and economic doctrines, every hunter and forager, every hero and coward, every creator and destroyer of civilisation, every king and peasant, every young couple in love, every mother and father, hopeful child, inventor and explorer, every teacher of morals, every corrupt politician, every "superstar", every "supreme leader", every saint and sinner in the history of our species lived there-on a mote of dust suspended in a sunbeam. The Earth is a very small stage in a vast cosmic arena. Think of the endless cruelties visited by the inhabitants of one corner of this pixel on the scarcely distinguishable inhabitants of some other corner, how frequent their misunderstandings, how eager they are to kill one another, how fervent their hatreds. Think of the rivers of blood spilled by all those generals and emperors so that, in glory and triumph, they could become the momentary masters of a fraction of a dot. Our posturing, our imagined self-importance, the delusion that we have some privileged position in the Universe, are challenged by this point of pale light. Our planet is a lonely speck in the great enveloping cosmic dark. In our obscurity, in all this vastness, there is no hint that help will come from elsewhere to save us from ourselves. The Earth is the only world known so far to harbour life. There is nowhere else, at least in the near future, to which our species could migrate. Visit, yes. Settle, not yet. Like it or not, for the moment the Earth is where we make our stand. It has been said that astronomy is a humbling and character-building experience. There is perhaps no better demonstration of the folly of human conceits than this distant image of our tiny world. To me, it underscores our responsibility to deal more kindly with one another, and to preserve and cherish the pale blue dot, the only home we've ever known."*

Carl Sagan