Modeling and simulating the Zika Outbreak under Deep Uncertainty:
A Multi-Method Multi-Resolution Approach

by

Philipp Schwarz

in partial fulfillment of the requirements for the degree of

Master of Science
in Engineering and Policy Analysis

at the Delft University of Technology,
to be defended publicly on 29 August 2016.

Student Number:	4370058	
Chair of graduation committee:	Prof. dr. B.A. Van de Walle	TU Delft
First Supervisor	Dr. E. Pruyt	TU Delft
Second Supervisor	Dr. M. Oey	TU Delft

Acknowledgements

This thesis is the product of my graduation work for the program Master Engineering and Policy Analysis at Delft University of Technology. I aimed to conduct a thorough and innovative research not intended to provide a conclusive framework but to explore future modeling and simulation capabilities, building on current innovations in the modeling field.

I am very grateful for the support and encouragement of my first supervisor Erik Pruyt following this path. Erik motivated me with his amazing conviction and enthusiasm to start with this of research, believe in myself and keep going. He was and is a source of inspiration to think not only what can be done now, but also what may be possible.

Further, I would like to thank the rest of my committee for showing confidence in my abilities, taking time out of their busy schedule for each meeting and providing valuable feedback and guidance. I am grateful for the clear and critical comments from the chair of my committee Bartel van de Walle and similarly I deeply appreciate the feedback from Michel Oey, who helped me greatly to balance my targets and organize the research.

In addition, I enjoyed the stimulating environment in the faculty Technology, Policy and Management. Bouncing off ideas with my committee and other researchers was extremely valuable. Especially, I would like to thank Marc Jaxa-Rozen for sharing thoughts and code and Islam Tushith for taking the time to explain me how to make use of geodata.

For the research of this thesis, I heavily depended on data made available by other researchers worldwide. I would to thank all researchers which provide open-access to their work - it was absolutely indispensable. Specially, I thank Jane Messina and Moritz Kraemer from the Department of Zoology at the University of Oxford for emailing me the raster dataset on the Environmental Suitability for Zika they created. Their research has been a cornerstone and inspiration for my own work.

Finally, I would like to end these acknowledgements by thanking my dear friends, my soulmate Graciela and my family that have given me my confidence and so much understanding. Without any one of you, I would not have managed. Now, I am eager to see new changes in my life!

Philipp Schwarz
Delft

Executive Summary

Within infectious diseases, Zika virus infections (ZIKV) are a recently recognized threat for global health (Fauci & Morens, 2016). They are transmitted by infected mosquitoes from the *Aedes* species and are closely related to dengue, chikungunya, yellow fever, West Nile, and Japanese encephalitis. Due to the ongoing large outbreak in the Americas and its links to cases of microcephaly, Guillain–Barré syndrome and other severe neurologic conditions (Brasil et al., 2016; Calvet et al., 2016; Cao-Lormeau et al., 2016; Schuler-Faccini et al., 2016), there is a pressing need to design effective response strategies.

To address this need, the development and use of mathematical modeling-based approach was selected. While existing modeling methods, such as system dynamics, agent-based and geographic information systems (GIS) models can be used for this purpose, the Zika outbreak poses seven challenges that are hard to overcome using a single-method approach: (1) subnational heterogeneity, (2) seasonal dynamics, (3) complex feedback effects, (4) global geographic scope, (5) deep uncertainty, (6) human connectivity, and (7) recognized ignorance on the relation between environmental factors and vector.

Therefore, the objective of this study was to propose a new integrated design that is suitable to study the problem statement in a systemic manner and able to cope with the identified functional requirements. Accordingly, the research question was formulated as follows: *How can a design method suitable to identify response strategies to prevent the potential spread of Zika virus infections be integrated and applied, while only having sparse and uncertain information on the epidemiology and effectiveness of potential solutions?*

To answer the research question, this thesis proposes a novel integrated design method with three main elements: (1) multiple interacting region-specific models, (2) big (georeferenced) data and (3) advanced pre-processing and post-processing operations.

The challenge of subnational heterogeneity is addressed by dividing the world along administrative provincial divisions and joining regions with similar characteristics within nations. Subsequently, each regional model is parameterized by aggregating open source raster datasets and combining them with other data sets such as census-based data.

Human connectivity between regions is incorporated by making use of open-access model based predictions for global air-travel passenger flows. The multiple regional models are simulated synchronously and after each time step information updated.

Seasonal effects are assumed to be exogenous but are still incorporated by using monthly input data. Therefore, exogenous model parameters are dynamically updated during run time.

Complex feedbacks were identified in the model structure of each individual model and also in the interaction between models.

The conflict of target between relevant micro-behavior effects and global geographic scope were dissolved by starting the simulation with an agent-based model with a subset of the agents in the region and switching once a threshold condition is reached and the disease

dynamics have established to a differential equation model (System Dynamic). Therefore, for this thesis two largely equivalent simulation models were built, one the one hand a System Dynamic model and on the other hand an agent-based model. Both models describe disease transmission by the interaction between human and mosquito. The human population is divided into a SEIR structure and the mosquito population into a SEI structure. When a mosquito bites a human, infectious humans infect susceptible mosquitoes and vice versa. In addition, several self-regulating feedback mechanisms are present, which are produced by several social and psychological effects.

The presence of deep uncertainties give reason for adopting an explorative modeling and analysis approach. This approach allows informed decision-making, despite extremely scarce and uncertain information.

The relation between environmental conditions and vector abundance and competence is recognized but poorly understood. Simulation model-based approaches require minimal knowledge on the relation between variables to specify its equations and consequently are not a viable option. This thesis presents an alternative to ignoring the relevant issue or making arbitrary combinations of assumptions. Based on the single assumption that environmental suitability for Zika's vectors is equivalent to the ratio of vector to human. The projections of an existing species distribution model which maps global environmental suitability for Zika virus could be used that condense the complex multivariate relationships between environmental conditions and vector abundance in a single index value.

In short, this thesis provides a proof of concept showing how Zika can be modelled with the proposed integrated design method.

To gain policy insights to prevent further spread of Zika, four main policies were tested. First, vector control was found to be insufficient to control the Zika outbreak and should not be the only measure in place. Second, the size of the Zika outbreak was not found to be sensitive to efforts to promote and increase personal protection. While these measures should certainly remain in place, they by no means constitute a robust response strategy to prevent or control an outbreak. Third, efforts to upgrade a surveillance and reporting system for Zika pay off under many scenarios. These findings support the WHO's decision to release a public alert about Zika in the beginning of 2016. Fourth, the analysis revealed that travel advisories would have only little effect on the dissemination of Zika.

Further, consistently over the complete ensemble of models, the recovery period (time human is infectious and can infect mosquito when being bitten) is the most important single variable that determines the speed of transmission and the Zika cases in the first wave of the global outbreak. Because of the importance of this parameter for disease propagation, efforts should be made to determine it more accurately. In addition, the exploratory analysis reveals that the share of vertical transmission is also an epidemiological key parameter. However, this parameter is highly uncertain as very different values have been reported. Therefore, controlled field tests should be initiated to establish better estimates for this variable.

Overall, five lessons that may be applicable to these diseases were drawn from the study of the Zika outbreak. First, travel plays a crucial role in the diffusion of the disease and that early alert the measures that air travelers can conduct to combat exposure. Second, it was found that the recovery period of a population (time that a human remains infectious) has a large influence on the size of the outbreak. Third, there is a complex dynamic component in the occurrence of an outbreak. On the one hand, they are re-emergent. On the other hand, it is difficult to predict its time, location and speed of dissemination. Fourth, reporting and communicating incidences of disease cases is crucial as it triggers self-regulating balancing feedback loops in the population. Fifth, model structure and parameters are deeply uncertain and there is a wide range of plausible futures for the outbreak that cannot be ranked probabilistically. Decision makers ought to design policies that are robust to all of these plausible outcomes.

At the moment, public media attention is focused on Zika outbreaks in South America and particularly in Brazil. Additionally, some subject matter experts stress the potential threat of large outbreaks that could also occur in southern parts of the United States. The simulation study supports this concerns as legitimate; under many simulated assumptions more than 1 million cases occur in North America. Besides, the analysis reveals that western and central Africa could be a center of especially big outbreaks.

The proposed integrated design method has proven to be useful to study Zika and could be applied with relatively little effort also to other vector-borne diseases such as malaria and dengue, which have similar characteristics as Zika. In addition, parts of the design method indicate potential to be applied in other areas. Many complex problems in humanitarian and social science display similar characteristics.

Regarding the application of the design method to Zika, additional efforts are needed to refine the proposed concept and reach acceptance for the outcomes among decision-makers such as health authorities and parts of the research community. First, targets for further research should be to revisit the assumptions made for generating the input data and constructing the model and check them for consistency with reality. For this, an expert validation process should be carried out. In line with this, the model outcomes should be calibrated to real-world data. In addition, attempts should be initiated to incorporate further important effects for the disease transmission, in particular, seasonality of climate conditions.

Table of Contents

List of Figures

List of Tables

Part I

Defining purpose and direction of the study

1 Introduction

Outbreaks of infectious diseases are an important global policy issue. Infectious diseases such as smallpox, polio, cholera, HIV/AIDS, influenza, malaria, dengue are the second leading cause of death worldwide, especially in low- and middle-income countries (Fauci, Touchette, and Folkers 2005). Because of the adaptability of infectious pathogens to the environment, new or reemerging infectious diseases are a perceptual challenge for humanity (Fauci and Morens 2012). In addition, environmental changes and increasing international interconnectedness have intensified the health threat of explosive global epidemics (McMichael, 2013).

Within infectious diseases, Zika virus infections (ZIKV) are a recently recognized threat for global health (Fauci & Morens, 2016) (see Infobox 1 for additional information). They are transmitted by infected mosquitoes from the *Aedes* species and are closely related to dengue, chikungunya, yellow fever, West Nile, and Japanese encephalitis. While ZIKV have been detected for several decades in parts of Africa, Southeast Asia, and the Pacific Islands, until recent times they had been viewed as relatively harmless and unworthy of much attention (Al-Qahtani et al. 2016). However, the ongoing large outbreak in the Americas (Infobox 1) and its links to cases of microcephaly, Guillain–Barré syndrome and other severe neurologic conditions (Brasil et al., 2016; Calvet et al., 2016; Cao-Lormeau et al., 2016; Schuler-Faccini et al., 2016), has led to the reassessment of ZIKV's potential dissemination and health risks.

In particular, the primary population risk group are childbearing women. When infected with the Zika virus during pregnancy, the infant is at increased risk at being born with microcephaly, a condition that causes unusual small head size, due to irreversible brain under development (Schuler-Faccini et al. 2016). The burden of microcephaly on families is usually devastating, constant care and support is required and full treatment unavailable (Louis 2016). Although, other complications such as Guillain-Barré syndrome can hit anyone infected with ZIKV, these cases are rare.

Due to the risks that Zika poses to human health, there is a pressing need to prevent its further spread. However, the design of effective global and regional measures to combat this disease is complicated by two major knowledge gaps. First, there is a lack of insight regarding the size of a potential epidemic. Second, it is unknown whether the disease would disseminate at different rates and scales between world regions. Shedding light on these knowledge gaps and providing insights for policy design are the primary motivations of this research.

Infobox 1: Zika outbreak in the Americas

The Zika outbreak in the Americas

In May 2015, Brazil reported the first ever cases of infections in the Americas and in February 2016, the World Health Organization declared the cluster of cases of severe neurologic conditions associated with the Zika outbreak in the Americas a public health emergency (PAHO/WHO 2016). Currently, as depicted in Figure 1, outbreaks are occurring in at least 23 south and middle American countries and are also being reported in some parts of Southeast Asia, and the Pacific Islands.

The Zika outbreak in Brazil is currently in the global spotlight due to the Olympic Games that will be hosted by Rio de Janeiro this summer (5-21 August 2016). Some health officials have voiced concerns over potential risks that international athletes and tourists become infected and spread the virus to other parts of the world after leaving the country (ECDC 2016; Romero and Ruiz 2016). Nevertheless, the upcoming Olympic Games are not the only factor that could trigger the spread of Zika: international air passenger traffic originating from regions with endemic ZIKV could also spread the disease to other world regions (Fauci and Morens 2016; Lucey and Gostin 2016).

Figure 1: Reported active ZIKV transmission as of May 26, 2016 (CDC 2016b)

1.1 Problem statement

To make the problem of the Zika outbreak explicit and define the direction of this research, the identified lack of insight was summarized in the following problem statement:

> *There is a pressing need to design robust response strategies to prevent the potential spread of Zika virus infections to regions worldwide suitable for Zika, while only having sparse and uncertain information on the epidemiology and effectiveness of potential solutions.*

1.2 Research approach

To address the problem statement, a research perspective known as systems thinking was adopted. This perspective describes the real-world phenomena in terms of interrelations and feedback effects between their components (Richmond 1997). It aims at overcoming the limitations of reasoning in terms of dependent and independent variables. In addition, it focuses attention on causality instead of on correlation. Therefore, this approach is suitable for the design of policy interventions that aim at tackling problems that are persistent over time, complex, and nonlinear and that occur under uncertain conditions.

Following this perspective, potential Zika outbreaks are not studied in isolation with a mechanistic model of the epidemiological process. Instead, the dissemination of this infectious disease is viewed as a part in a larger system with elements and effects such as environmental conditions for its transmission, human actors, international traveling, self-regulating psychological effects and interventions, urbanization and social inequality.

Based on this research perspective, the development and use of mathematical modeling-based approach was selected to provide policy advice. From this approach, models are widely used to assess the dissemination of infectious diseases by quantitatively describing the epidemiological mechanisms that generate the dynamics of disease outbreaks. Then, these models are used to understand plausible futures (Grassly and Fraser 2008; Siettos and Russo 2013) and to identify and explore effective prevention and control measures without requiring intervention in reality (Grassly and Fraser 2008).

1.3 Need for a new integrated design method

In this study, the overarching methodology to develop and use a mathematical model followed the general procedure set out by van Dam (2013). First, the problem was made explicit by means of a problem statement, presented in the previous section. Then, the structure of the system and its boundaries were specified, based on a literature review. Further, functional requirements for a mathematical model were defined. Afterwards, the mathematical model was specified and data from diverse sources collected and integrated. After implementation in a programming environment, the model was verified and validated. Finally, experimental scenarios were determined and simulated, results interpreted and their policy implications made explicit.

While studying the structure of the system and defining the functional requirements for a mathematical model, seven aspects that make the Zika outbreak a particularly complex modeling case were identified: Complex feedback effects, Subnational heterogeneity, Global geographic scope, Human connectivity, Seasonal effects, Deep uncertainty and Recognized ignorance on the relation between environment and vector. Each of these aspects was translated into a functional requirement for a model that appropriately represents the problem and allows the exploration of policies. These requirements are detailed in section 2.2.

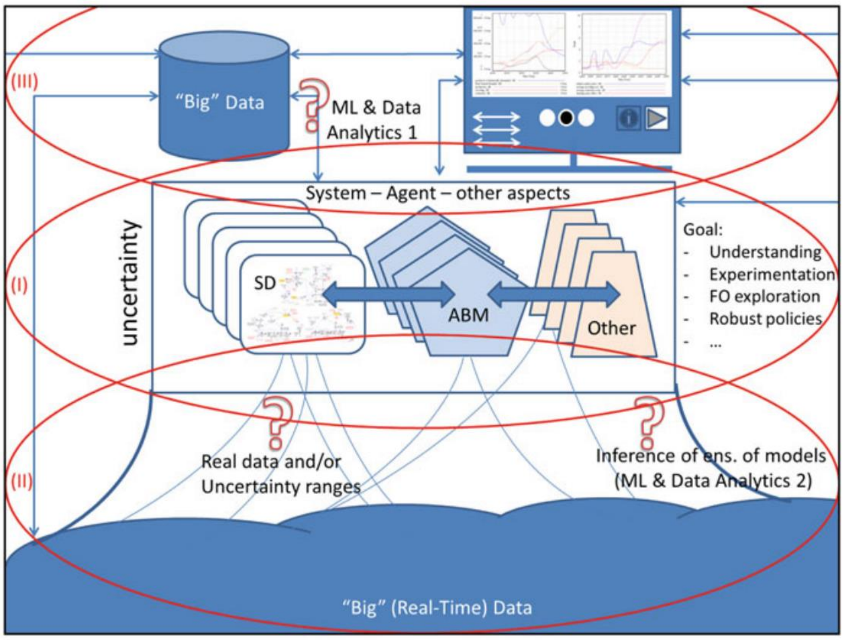
Modeling methods presently available are able to address these aspects individually; however, they fail to account for all relevant issues at the same time and are therefore insufficient to address the problem statement. Accordingly, the focus of this study was not only to design robust response strategies to prevent the potential spread of Zika, but to reach beyond the boundaries of traditional modeling and propose a new integrated design method. In the future, the products of this research can also be applied to study other vector-borne diseases, such as malaria[1] and dengue, and beyond.

The need for research on multi-method and hybrid modeling and simulation approaches to expand the modeling and simulation capabilities of computational models and enhance their usefulness as decision-making tools has been previously recognized (Pruyt 2015). However, existing approaches are still now limited when dealing with aspects such as disaggregated system characteristics, spatial and network aspects, deep uncertainty, georeferenced information, and Big Data, (Pruyt 2015). To tackle these challenges, modelers are beginning to adopt and integrate modeling methods and techniques from related disciplines such as computer science and geomatics. Although their development is yet in its infancy, multi-method simulation approaches promise to deliver significant breakthroughs in model based policy analysis and may provide the urgently needed toolsets to improve advice on the grand challenges of our time (Pruyt 2015).

---

[1] According to the World Health Organization, in 2015, malaria caused 450 million cases and dengue is known to be the world's fastest growing vector-borne disease with already 2.35 million cases annually (WHO 2016b; WHO 2016c).

Figure 2: Future state of the art of modeling and simulation (Pruyt 2015)



1.4 Research objective and questions

The objective of this study is to propose a new integrated design that is suitable to study the problem statement in a systemic manner and able to cope with Zika's 7 modeling challenges. This objective was translated into the following research question:

> *How can a design method[2] suitable to identify response strategies to prevent the potential spread of Zika virus infections be integrated and applied, while only having sparse and uncertain information on the epidemiology and effectiveness of solutions?*

To organize the research, this question was decomposed into seven sub-questions (Figure 3). Each of these sub-questions is answered in one of the remaining chapters of this document.

---

[2] The term design method is used to describe an exploratory research approach to generate new solutions to problems by going beyond the available standard set of frameworks and techniques.

Figure 3: Research Sub-Question

Sub-question	Aspect addressed	Chapter
From a systems thinking perspective, what are relevant aspects of the Zika case that a simulation model should integrate in order to appropriately represent the problem and enable policy analysis?	• *System boundaries* • *Criteria* • *Internal & external factors* • *Policy options*	*2*
How can three single modeling and simulation methods – system dynamics (SD), agent-based (AB) and geographic information systems (GIS) - represent the previously identified aspects and what are the methods' limitations?	• *Modeling methods for infectious diseases:* • *System dynamics* • *Agent-based* • *GIS*	*3*
How can system dynamics, agent-based and GIS be integrated into a new design method to address Zika's 7 functional requirements?	• *General concepts* • *Method integration and architecture*	*4*
What are the specific data requirements of the proposed design method and how can they be met?	• *Data collection* • *Data pre-processing* • *Exogenous model input* • *Handling missing data*	*5*
How can simple simulation models be formulated to represent the key dynamics of the problem?	• *Region-specific models:* • *Conceptualization* • *Specification*	*6*
What insights on the Zika case can generated by applying the new integrated design method?	• *Experimental set-up* • *Results and analysis* • *Conclusions for Zika*	*7*
How can the new integrated design method be used and improved in future work?	• *Recommendations for future use* • *Critical reflection* • *Future improvement*	*8*

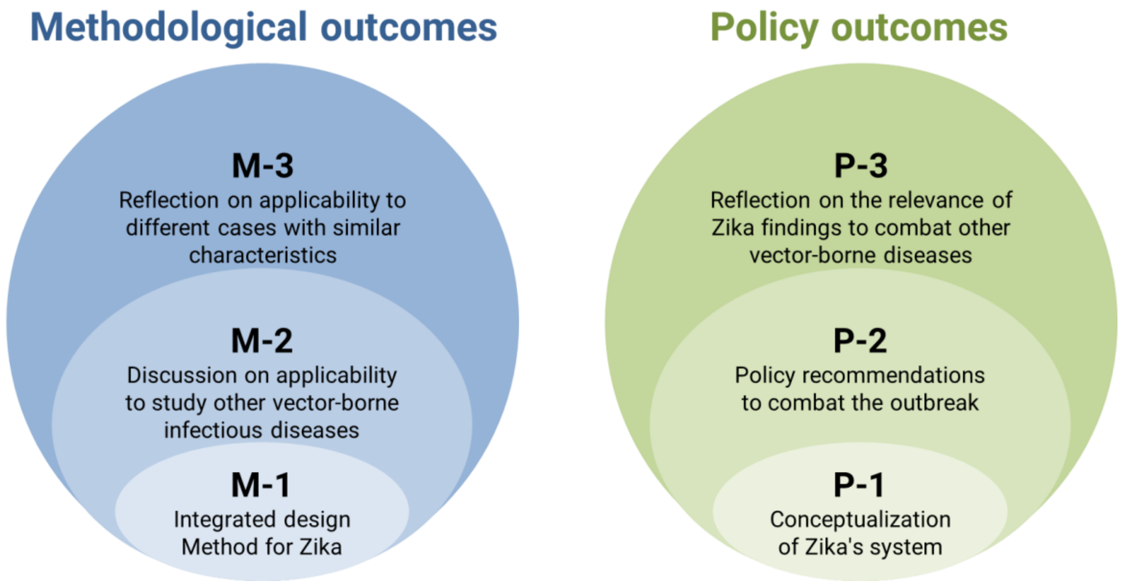
1.5 Expected outcomes

The six outcomes expected from this study are grouped in two categories: methodological and policy.

On the one side, there are three methodological outcomes (M): (M-1) the conceptualization and implementation of a new integrated design method to address the problem statement, along with a discussion on its capabilities and limitations and recommendations for future improvement, (M-2) a discussion on its applicability to similar vector-borne infectious diseases, and (M-3) a reflection on how a similar design method can be used to study different cases with similar characteristics.

On the other side, three policy outcomes (P) are expected: (P-1) a conceptual model that provides a systemic understanding of the dynamics of the ZIKV outbreak, (P-2) policy recommendations to combat this outbreak, and (P-3) a reflection on the relevance of these findings to combat other vector-borne diseases with similar characteristics.

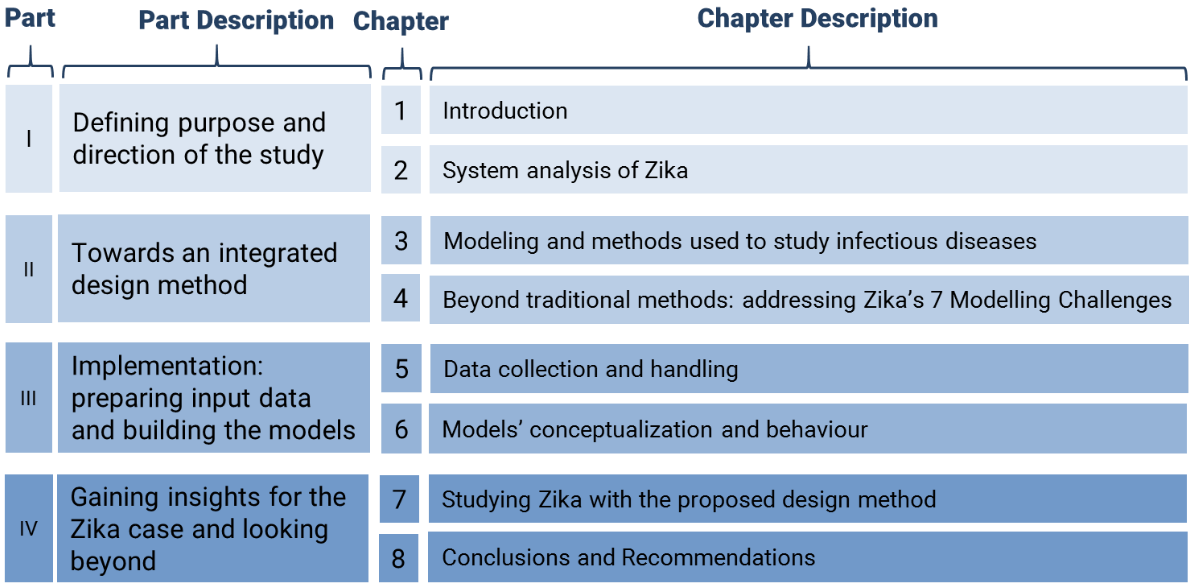
Figure 4 Methodological and policy outcomes from this research



1.6 Reading ahead: Organization of Thesis

The remaining parts of this thesis are structured in four parts and eight chapters (Figure 5).

Figure 5 Structure of this document



Part	Part Description	Chapter	Chapter Description
I	Defining purpose and direction of the study	1	Introduction
		2	System analysis of Zika
II	Towards an integrated design method	3	Modeling and methods used to study infectious diseases
		4	Beyond traditional methods: addressing Zika's 7 Modelling Challenges
III	Implementation: preparing input data and building the models	5	Data collection and handling
		6	Models' conceptualization and behaviour
IV	Gaining insights for the Zika case and looking beyond	7	Studying Zika with the proposed design method
		8	Conclusions and Recommendations

In Part I, purpose and direction of this study are defined. In the present chapter, the problem is made explicit and the motivation for an integrated design method is introduced. Then, a system analysis of the problem is presented in Chapter 2, and seven functional requirements for the modeling and simulation study are made explicit.

In Part II, the need for an integrated design method is presented, and its conceptualization is described. In Chapter 3, the capabilities and limitations of three methods commonly used for modeling vector-borne infectious diseases are discussed, and further motivation is provided for their integration. Then, in Chapter 4, the integrated design method is formally proposed. Further, its suitability to address each of Zika's seven functional requirements is detailed.

In Part III, the implementation of the design method is described. First, collection and handling of Big Data and georeferenced data are discussed in Chapter 5. Second, the formulation of the system dynamics and the agent-based model are presented in Chapter 6.

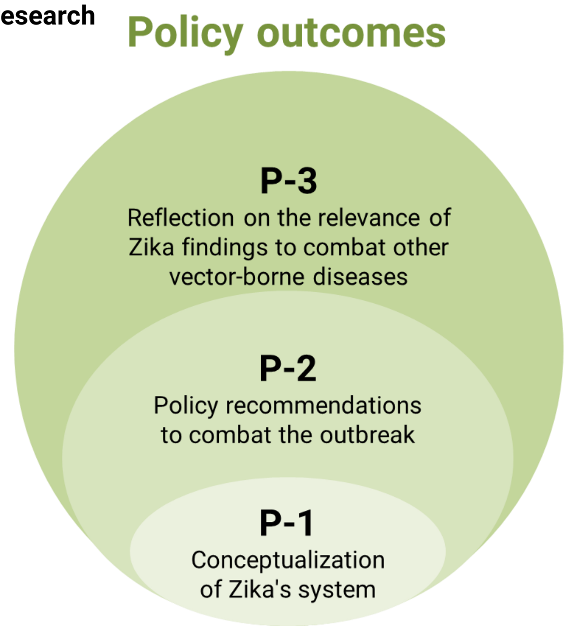
In Part IV, insights, conclusions and recommendations are drawn. In Chapter 7, further study of the problem statement using the design method is reported and results presented. Then, in Chapter 8, conclusions and recommendations are drawn from both, policy and methodological points of view.

2 System analysis of Zika

This chapter is divided in two sections. First, in section 1, further detail is presented on the Zika outbreak, and this problem is analyzed from a systems thinking perspective. This understanding of the problem consists the first policy outcome of this thesis, as depicted in Figure 6.

Then, in section 2, the implications of this problem analysis for the design of a model are discussed by making the features of Zika that make the modeling exercise a challenge explicit.

Figure 6: Policy outcomes expected from this research



2.1 System Demarcation

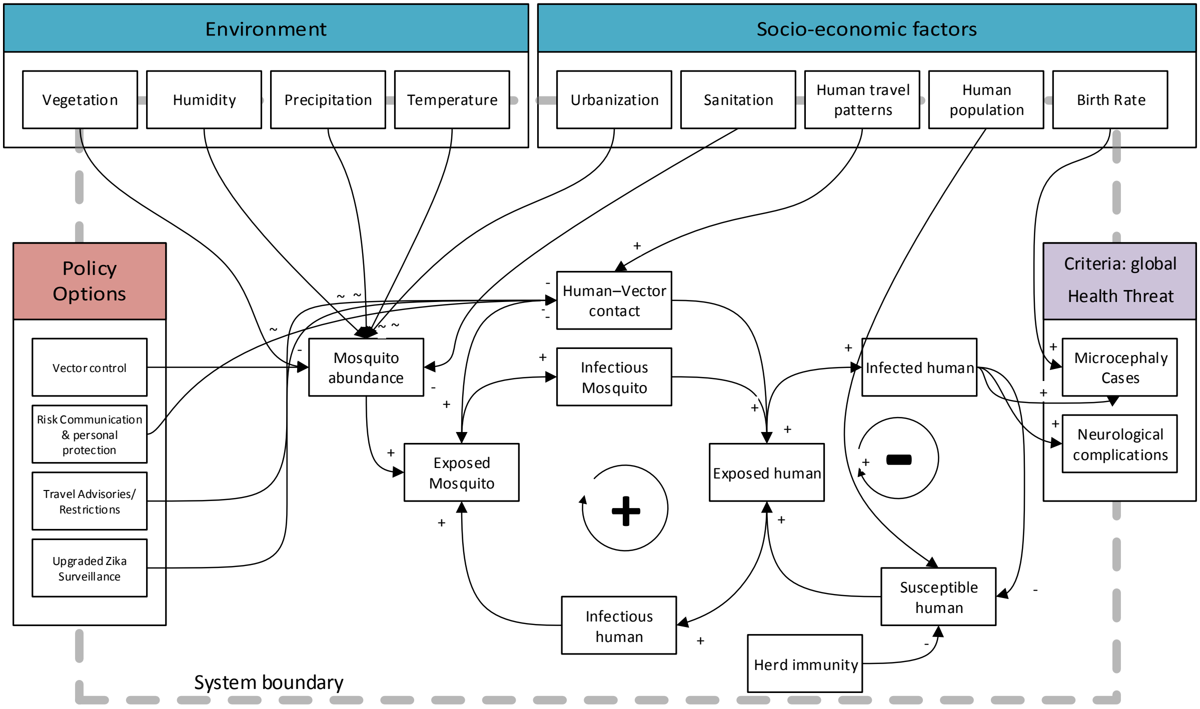
The Zika outbreak was studied from a systems thinking perspective in order to develop an adequate understanding of its driving mechanism and to define suitable boundaries for its study. To achieve this, a systems diagram of the problem was built.

This method consists in representing causal relations between key elements qualitative, following Enserink et al. (2010), and classifying these elements into four categories: (1) criteria (used to measure the problem and whether it has been solved solved); (2) external factors[3] (that cannot be influenced by policy makers but affect the outcome of the system); (3) means or policy options (available actions that could be performed and whose relative performance is assessed); and (4) internal factors (comprise of all other key factors that affect the criteria). Together, criteria, external factors and policy options define the boundaries of the system under study. The systems diagram for Zika, depicted in Figure 7, provides an overview of the most important factors and causal relations that need to be considered in the process of designing a computational model.

---

[3] In computational models, these factors are treated as exogenous variables: their values are independent from the states of other variables in the system

Figure 7: Systems Diagram[4]



.

2.1.1 Internal factors

Key internal factors and their interrelation are depicted in the center of the systems diagram. Vector-borne diseases are spread through mosquito blood feeding, which is carried out only by female mosquito who needs the blood to produce eggs. Typically, mosquitoes become infected while feeding on blood from an infected person. However, in rarer cases they may be infected through vertical transmission: mother-to-child. To a large extend, initial herd immunity -acquired immunity in human population through prior infection or vaccine- determines the potential spread of the infection in a population. Additionally, transmission dynamics are determined by a number of other parameters which were for simplicity not included in the systems diagram (e.g. extrinsic and intrinsic incubation period and vector-to-human and human-to-vector transmission rate).

Strongly simplified, the dynamics of infectious diseases can be classified as the *limits to growth* archetype, in which the exponential growth (of infected) is followed by eventual collapse (Senge 1990). One the one hand, in a positive feedback loop, as more humans become infectious more mosquitoes also become infected, which in turn infect more humans. On the other hand, with each newly infected human the susceptible population declines, which balances the increase of infected individuals.

2.1.2 External-factors

External factors, pictured at the top of the systems diagram and marked in cyan, influence the transmission of ZIKV but are not impacted by the system itself. They are classified into two

---

[4] Author’s illustration

categories, environmental and demographic factors, and it is assumed that no interplay occurs between them.

Environment

Variation in temperature, rainfall and other climate variables constraint the *Aedes* mosquitoes' persistence and competence for virus transmission, or ability to spread the disease (Brady et al. 2013; Hales et al. 2002; Rueda et al. 1990; Yang et al. 2009). The mechanism that temperature plays in imposing limits for ZIKV transmission is the most clearly understood one: it determines the development and survival of mosquitoes and their extrinsic incubation period (Brady et al. 2013). Further, evidence has been presented that higher precipitation and humidity promotes vector borne disease transmission (Hales et al. 2002; Rowley and Graham 1968), and that an association exists between local vegetation and mosquito reproduction (Linthicum et al. 1999). Nevertheless, since little is known about the internal physiological processes in *Aedes* mosquitoes, the mechanisms between environment conditions and mosquito abundance are difficult to specify quantitatively.

Socio-economic factors

Likewise, socio-economic factors affect the potential for ZIKV as a global health threat. One the one hand, the size of the susceptible population largely determines the number of possible neurological complication and microcephaly cases, associated with the ZIKV. In addition, socio-economic dynamics impact mosquito abundance: a link between urbanization and vector presence has been established for a long time. *Aedes* mosquitoes' habitats are found in close proximity to human dwellings, because breeding places (stagnant water bodies) and host species (humans) are available in abundance (Brown et al. 2011; Knudsen and Slooff 1992). In particular, slum housing are favorable habitats for mosquitoes, since the combination of poor sanitation (surface-water drainage and inadequate garbage collection) and dense population are perfect breeding places (Eisenstein 2016; Knudsen and Slooff 1992). Therefore, the outbreak also has a social dimension as the poorest demographics are more exposed to potential contagion than others, while at the same time they often lack the financial means for self-protection, such as insect repellent, mosquito net or air-conditioning units) (Eisenstein 2016).

On the other hand, the primary mode for large scale virus dissemination is human mobility rather than mosquito flights. In their whole lifetime, mosquitoes stay in or around the same neighborhood and rarely fly long distances above four hundred meters (Harrington et al. 2005; Trpis and Hausermann 1986). Individual human mobility is therefore directly responsible for the inter- and intra-continental spread of ZIKV. In particular, global air transport connecting different regions plays an important role in the dissemination of infectious diseases and their assessment as global health threat. In recent years, the increasingly extensive and intensive global interconnectedness has raised concerns regarding a rapid spread of deathly infectious diseases globally. Annually, global passenger travel alone is estimated to transport 0.7 - 3.3 billion people (Guimerà et al. 2005; IATA 2014).

2.1.3 Available policies

Policy options, marked in red in the systems diagram, represent actions that can indirectly influence the criteria. The available policies can be classified into four groups: (1) Vector Control, (2) Risk Communication & personal protection, (3) Upgraded Zika Surveillance and (4) Travel Advisories/Restrictions.

(1) **Vector control:** measures to reduce the mosquito population, it is arguable the most effective measure to control mosquito-borne epidemics (PAHO/WHO 2015). In practice, different strategies are possible.

First, potential mosquito breeding sites (stagnant water pools) should be identified and completely eliminated (Marcondes and Ximenes 2015). Second, in spite of, adverse health effects, some experts suggest to use insecticide to spray typical mosquito habitats (Lucey and Gostin 2016). Third, more recently, it was proposed to release genetically modified sterile male mosquitoes (Lucey and Gostin 2016; Thomé, Yang, and Esteva 2010). Although early successes in limited geographical regions have been reported, doubts on cost effectiveness and associated risks are being raised (Lucey and Gostin 2016; Snow et al. 2005; Subbaraman 2011). Fourth, another novel approach under development is to infect *Aedes* mosquitoes deliberately with *Wolbachia* - a bacterial parasite (Moreira et al. 2009; Aliota et al. 2016). The desired mode is that the *Wolbachia*-infected mosquitoes will live long enough to reproduce and this way keep spreading *Wolbachia* through the population but not long enough to transmit ZIKV to humans.

(2) **Measures for the personal protection:** standard measures to prevent vector transmitted diseases. Insect repellents and the use physical barriers (mosquito nets, screens, closed windows, long-sleeved shirts and long pants) are effective to prevent being bitten (Brazilian Ministry of Health 2016). Especially during the hours of highest mosquito activity (mid-morning and late afternoon to twilight), people should stay indoors and in air-conditioned rooms (ECDC 2016; Lucey and Gostin 2016). Furthermore, any possible mosquito-breeding site in and around the house should be removed. Nevertheless, a problem remains for those living under impoverished conditions because precautions to avoid mosquito bites may not be affordable (Fauci and Morens 2016).

(3) **Upgraded Zika Surveillance:** relevant measures to understand the magnitude of the epidemic and allocate resources. To ensure appropriate and targeted interventions, any mitigation action demand effective mosquito surveillance (Lucey and Gostin 2016). Public health authorities have to establish larger capacity to detect and confirm suspected ZIKV cases (ECDC 2016). In addition, health workers worldwide have to be trained to recognize and report Zika related diseases for an optimal response.

(4) **Travel Advisories and restriction**: travellers to countries where ZIKV is circulating are equally at risk of contracting the disease as their residents and should be informed about this risk (ECDC 2016).

Although it is not officially a policy, health experts in Brazil and Colombia advise women in high-risk areas to delay their pregnancies (Douglas 2015; The Guardian 2016).

2.1.4 Objectives and criteria

The objective of infection control is relative straightforward and two criteria for the assessment of policies were selected. First, the percentage of cumulative Zika cases in a population and second, the Zika virus infections in childbearing women, due to the link of ZIKV to microcephaly. Additionally, the dissemination of ZIKV is assessed, particularly, the number of countries with local transmission and affected continents.

2.2 Identification of 7 fundamental functional requirements for the model

Seven functional requirements were distilled from the system analysis. Each of them represents a key aspect of the problem that needs to be accounted for in the design of the modeling approach:

(1) **Complex feedback effects:** in addition to the core epidemiological mechanism depicted in the systems diagram, an appropriate model has to represent other important feedback effects. Parker and Epstein (2011) argue that the dynamic response of humans to the state of the outbreak play a key role in disease propagation. This dynamic response is driven by social and psychological effects.

(2) **Subnational heterogeneity:** local environmental conditions (e.g. temperature and precipitation) are linked to vector density and competence[5] and thus determine local capacity for transmission of vector-borne diseases (Johansson, Dominici, and Glass 2009). Further, socio-economic factors such as (susceptible) population are location-specific and likewise subnational heterogeneous. Therefore, modeling ZIKV calls for the representation of location-specific heterogeneity at subnational level (Wilson 2002).

(3) **Global geographic scope**: at a local level, the transmission of ZIKV depends on whether environmental conditions are suitable for *Aedes* mosquitoes. Messina et al. 2016 found that suitable areas are located worldwide, throughout tropical and subtropical regions. Although at the time of writing ZIKV outbreaks were mostly limited to the Americas, they

---

[5] Vector competence: Mosquitoes capability to acquire, maintain, carry and transmit Zika virus (Hardy et al. 1983)

are potentially threatening for large parts of the world. Therefore, a global modeling scope can shed light on plausible global ZIKV outbreaks.

(4) **Human connectivity:** Epstein et al (2007) reason that human travel is a central driver for infectious disease dissemination. They argue that in particular, global air-travel may cause a rapid spread of infectious diseases. If humans who become infected with Zika in epidemic regions travel to Zika-free regions with suitable conditions for the vector, outbreaks could occur also in their destination.

(5) **Seasonal effects**: Johansson et al. (2009) found that monthly changes in temperature and precipitation are correlated with local transmission. Though this study is focused on dengue, similar links can be hypothesized for Zika. More specifically, there are temperature and precipitation ranges that are optimal for mosquito reproduction, outside these ranges the egg mortality is greater and vector reproduction declines. These effects are non-linear and dependent on the duration of the less-than optimal period, e.g. long high temperature periods have widely deathly effects. In addition, Mao et al. (2015) argues that many diseases have seasonal dissemination patterns driven by seasonal effects in air-travel. In summary, annually aggregated data fails to appropriately represent key disease dissemination dynamics.

(6) **Uncertainty**: Villa (2016) argues that health authorities face the challenge of communicating uncertainty associated with Zika. For example, at the beginning of the outbreak, the link between Zika and microcephaly was not yet scientifically confirmed. Furthermore, as the recognition of this disease as a global health threat is recent, epidemiological information tends to be sparse and uncertain. Formulating and parameterizing a model in the presence of these enormous uncertainties calls for the adoption of methods for decision-making under deep uncertainty.

(7) **Recognized ignorance on the relation between environment and vector** – Hales et al. (2002) and Johansson et al. (2009) describe a link between local environmental conditions and vector abundance and competence. However, these complex multivariate relationships remain poorly understood. They based on evidence based but their statistical significance has been determined, in spite of decades of research efforts (Johansson, Dominici, and Glass 2009; Hales et al. 2002).

Part II

Towards an integrated design method

3 Modeling & simulation methods used to study infectious diseases

Over the years, multiple modeling methods have been proposed to study the spread of epidemics, three of the most prominent ones being system dynamics (SD), agent-based (AB) and spatial GIS models. While all three methods have been applied to various infectious diseases including Zika, each method focuses on different system features. Therefore, their inherent characteristics define their applicability to study certain aspects of this particular disease.

In this chapter, each method is shortly introduced and their respective applications to infectious diseases reviewed. Then, the methods' capabilities are contrasted. Finally, the need for integration of these methods to study Zika, based on the previously identified functional requirements is justified.

3.1 System Dynamics (SD)

SD models describe systems as a set of differential equations and are therefore commonly referred to as differential-equation models. The structure of these models is characterized by stock, flows and auxiliary variables (Sterman 2000). In this method, the spread of infectious diseases is modelled by dividing the population in compartments according to their health condition. Within each compartment the population is assumed to be homogenous and perfectly mixed (Rahmandad and Sterman 2008). Despite these apparently unrealistic assumptions, compartmental models in epidemiology are often able to capture the essence of dynamics in a simple and tractable manner (Siettos and Russo 2013)

SD is commonly used for rapid assessment of emerging situations, such as HIV/AIDS (Pedamallu et al. 2010; Roberts and Dangerfield 1990), Ebola (Pruyt, Auping, and Kwakkel 2015), and dengue (Kucharski et al. 2016; Ritchie-Dunham and Méndez Galván 1999). Many of these studies take a holistic view on the problem and account for a number of feedback effects that other methods typically neglect (e.g. psychological effects and the availability of health care infrastructure). Nonetheless, what these studies have in common is that their geographic scope is limited: they deal with only a specific city, community, region or island populations.

3.2 Agent-based (AB) models

AB models describe complex systems as networks of autonomous agents that interact with the environment, including other agents (Bonabeau, 2002; Epstein & Axtell, 1996; Phelan, 1999). Based on decision rules, agents reason upon their (limited) observed knowledge and take actions that affect their environment and their own internal state (Axtell et al. 1996; Jennings, Sycara, and Wooldridge 1998). Agents are heterogeneous and can differ in various ways (e.g. income, nationality, and behavioral rules) (Epstein 2006). Therefore, the propagation of a disease across a population is an emergent property, resulting from the interaction of many individuals (Bonabeau, 2002).

The strength of modeling infectious diseases with an AB perspective is that this method can be used to identify spatial patterns in disease transmission and differentiate risk groups. Some AB studies focus attention on micro behaviour patterns of disease spread, i.e. on the complex interaction between vector and human simulating vector-borne diseases (Jacintho et al. 2010; Manore et al. 2015; Medeiros et al. 2011). Moreover, AB models are applied to infectious diseases at global scale where transmission occurs directly from one person to another (Parker and Epstein 2011; Riley 2007). Due to the key role of human travel for disease dissemination, AB models have been linked to global air-travel data. For example, Epstein et al. (2007) explored scenarios on a possible influenza pandemic and the effects of travel restrictions adopting an AB perspective.

3.3 GIS models

Geographic information systems (GIS) are data models able to represent, structure, manipulate and analyse spatial data. GIS is a broad concept which may refer to many methods and processes. Here, the term GIS models focuses on multivariate spatial analyses. In this sense, GIS models are used for the analysis of spatial causal relations between variables and for integrating data from different sources including satellite images, censuses, and household surveys to produce high-spatial resolution maps.

An important analytical tool for the study of vector-borne diseases are species distribution models. They are used to produce predictive maps that depict where vectors are likely to occur, based on multiple layers of geo-referenced information. Species distribution models establish an empirical relations between environmental and socio-economic covariates hypothesized to affect the mosquito prevalence and observed mosquito occurrence (Bhatt et al. 2013; Messina et al. 2016). The outcome of these models are gridded maps. Each cell represents a geographic location and contains a value that indicates the suitability of the location for disease transmission. If data is available, these models can readily be applied to a global scale (Hay et al. 2006).

This technique has been applied to various vector-borne diseases including dengue, malaria and Zika (Bhatt et al. 2013; Semenza et al. 2013; Messina et al. 2016). Based on this results, GIS models can be used to provide spatially explicit projections of population at risk. Existing applications include a study of the ZIKV epidemics in the Americas, which concluded that 1.0 – 1.9 million pregnant women may be infected with Zika, and thousands of microcephaly cases can be expected (Perkins et al. 2016).

3.4 Contrasting the advantages and limitations of SD, AB and GIS models

SD, AB and GIS models have all been applied to study infectious diseases. However, each of them has been used to shed light on different aspects of the problem. In the following paragraphs, a reflection on these aspects is presented, and the methods' advantages and limitations are contrasted.

Recently, researchers have made multiple attempts to contrast the outcomes of AB and SD models (Parunak, Savit, & Riolo, 1998; Rahmandad & Sterman, 2008; Schryver, Nutaro, & Shankar, 2015), resulting in the identification of four key differences between AB and SD.

(1) They differ in the level of aggregation. In compartmental models the system state is aggregated to average values (Rahmandad 2004). In contrast, modeling at the level of individual agents allows the look into micro behavior and at any summary statistics describing the aggregate agents' states.

(2) Individual-based modeling inherently accounts for the heterogeneity of agents, but compartmental models require for each different property of individuals to increase the size of the model by adding more compartments (Milling and Schieritz 2003).

(3) Other than few exceptions, in SD the equation describing the dynamics from one compartment to another are formulated deterministic, in contrast in AB normally, decision rules, and initial properties are drawn from probability distributions (Bonabeau, 2002 Macal, 2010).

(4) Inherently AB models can represent spatial relations while its integration SD is very tedious and limited.

(5) AB models are computationally intensive compared to the other modeling methods. Large-scale models thousand and more agents require the availability of expensive parallel high-performance computing resources (Cockrell et al. 2015).

Both, ABM and SD are simulation methods that describe complex dynamics systems. In contrast, the GIS models discussed in this study are static. These models implicitly assume that the conditions that determine the transmission of the disease are homogenous in time. Therefore, these models fall short to account for dynamic feedback effects, i.e. changing behavior patterns in response to different states of the outbreak.

In Figure 8, the comparison of the three modeling methods is summarized.

Figure 8: Comparison of modeling methods of infectious diseases

Method	Main application	Key capabilities	Key Limitation
SD	• Capture essence of dynamics in a simple and tractable manner	• Low computational requirements • Dynamic simulation over time • Holistic system perspective	• Limited geographic scope • High aggregation
AB	• Study behaviour patterns of the system emerging from individual action • Large-Scale-Spatial-Transmission Models	• Agent heterogeneity • Spatially explicit • Dynamic simulation over time	• High computational requirements
GIS	• Species distribution modeling • Projection of population at risk	• High spatial resolution • Georeferenced multivariate relationships between observations and covariates	• Static • Feedback effects are difficult to represent

3.5 Capabilities and limitations of single-method approaches to address Zika's 7 functional requirements

In section 2.2, seven challenges associated to modeling the potentially worldwide Zika virus (ZIKV) outbreak have been presented. In the following paragraphs, an explanation of how each of these challenges can be addressed by SD, AB or GIS is presented. Furthermore, advantages and pitfalls of using these methods are set out.

(1) **Complex feedback effects** - SD can be used to readily encompass a large number of feedback effects. In principle, AB can account also for feedback effect, however, since it adopts a bottom-up perspective (local behaviour rules drives emergent behaviour), the causal relations are less tractable. On the other hand, species-distribution models (a type of GIS models) are in essence spatial regression models with dependent and independent variables but have no feedback effects.

(2) **Subnational heterogeneity** - SD has limitations dealing with geospatial aspects, in contrast AB and GIS models excel on dealing with these aspects. Both methods can account for regional differences and provide spatial visualization. GIS models can integrate multiple layers of geo-spatially specific data.

(3) **Global geographic scope** – SD can easily represent very large populations but aggregates them to few homogenous compartments. Assuming data availability, GIS models can readily encompass global scope. Detailed large-scale AB models are being developed but demand immense computational requirements.

(4) **Human connectivity** – As described above, global air-travel flows have been integrated into large-scale AB models. Although no studies where applying this method were found, in principle SD could be used to study the effect of travel on disease dissemination. However, by nature, static GIS models do not allow to incorporate interconnection between regions.

(5) **Seasonal effects –** While seasonal effects can be incorporated in ABM and SD models, it remains the modeller's choice whether these effects are considered exogenous or endogenous - meaning that the dynamics can be generated by the use of temporally resolved data or by internal feedback structures of the model. Producing dynamic intra-annual GIS models that account for seasonal effects is challenging. Nevertheless, research efforts in this direction have been made Brady et al. (2014) present recently gridded monthly maps of global temperature constraints on *Aedes* mosquito persistence and competence for dengue virus transmission based on spatial and monthly resolved data.

(6) **Uncertainty** – Sensitivity analysis is part of all three modeling methods, however, currently, only AB and SD support the study of robust policies under conditions of deep uncertainty using the Exploratory Modeling and Analysis methodology (EMA) (Kwakkel and Pruyt 2013).

(7) **Recognized ignorance on the relation between environment and vector** – Formulating the equations of a simulation model requires at least minimal knowledge on the mathematical relationship between relevant factors. Nonetheless, information on the complex multivariate relationship between environment conditions and vector abundance and competence is scarce. This situation of recognized ignorance[6] calls for a data-poor modeling approach instead of using traditional simulation models. Species-distribution modeling can be used for this purposes because it allows to link empirical spatially referenced data to generate predictive models of species distribution. Consequently, the relation of multiple variables can be condensed into a single index value (Messina et al. 2016).

In summary, while the three modeling methods can address different aspects of the seven key modeling challenges, tackling these problems at the same time calls for the integration of SD, AB and GIS into a new design method. This integration is the subject of the next chapter.

---

[6] Recognized ignorance is a term coined by Kwakkel (2010), which describes a situation where we are even unable to enumerate alternatives.

4 Beyond traditional methods: addressing the 7 modeling challenges

Combining modeling methods has received increasing attention during the last years because it promises to exploit the capabilities of single methods without balancing out their trade-offs. Motivated by the objective to generate insights into complex interdependent processes having very different natures, numerous attempts have been made to mix and align SD and ABM (Borshchev and Filippov 2004; Osgood 2007), and a number of different architectures coupling and integrating the methods have been proposed (Swinerd and McNaught 2012; Vincenot et al. 2011). However, a challenge remains on how to align and integrate these methods. In contrast, integrated ABM and GIS models are more established. Current commercial-off-the-shelf ABM software like Netlogo and Repast support at least a limited set of GIS processes and visualization capabilities (Wilensky and Rand 2015; North et al. 2013). Likewise, SD and GIS models have been coupled (Ahmad and Simonovic 2004; Neuwirth and Peck 2013), however, these integrated methods have not yet found wider adoption.

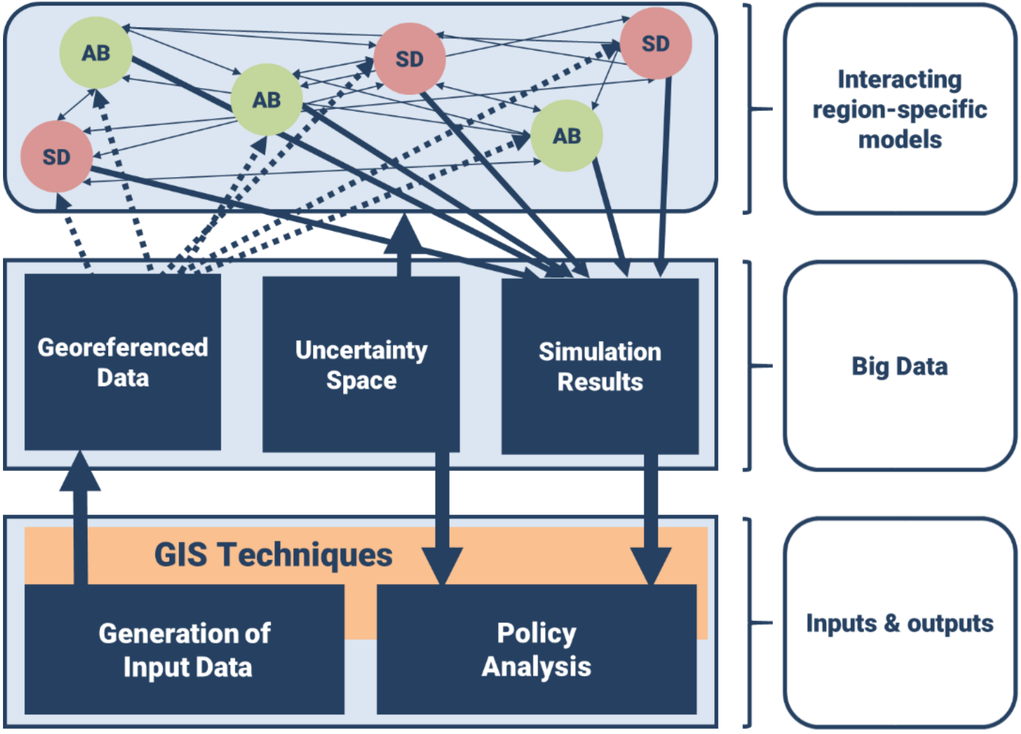
In this chapter, a new integrated design method to address Zika's 7 modeling challenges is presented. It consists of different modeling and simulation methods and adopts concepts from other discipline, namely data mining and GIS. In the first section of this chapter, the conceptualization of the design method is presented, and its components and interactions are introduced. Then, in the second section, the capabilities of the adopted methods to cope with each of the 7 modeling challenges are discussed.

4.1 A new integrated design method

The general concept of the integrated design method for modeling Zika draws from Pruyt's (2015) definition of the future state of the art of modeling and simulation (Figure 2). By incorporating four of its key elements -uncertainty, Big Data, data analytics and hybrid modeling methods- it is able to address Zika's 7 functional requirements.

The new integrated design method has three major components, depicted in Figure 9: interacting region-specific models, Big Data, and inputs and outputs. In the following paragraphs, each of these components is detailed.

Figure 9: General conceptualization of the integrated design method



4.1.1 Interacting region-specific models

As illustrated in the upper shaded box, multiple region-specific models are built which interact with one another. This is achieved by simulating the models synchronously and exchanging information after each time step. The models are based on the same core model structure but parameterized differently according to regional conditions. Another feature is that two simulation methods (SD and AB) are used to model the disease dynamics within a region. Each region begins in the AB model but switches during the simulation to the SD model, once the number of active agents has reached a threshold condition. It is worth noting that an SD model can interact with an AB model from a different region, and vice versa, in spite of their different system representations.

4.1.2 Big Data

Handling big data is one of the capabilities of the design method. Region specific input data for each model's parameterization is stored in large georeferenced databases. Based on these datasets, the model's parameters are set at the beginning of the simulation and can be even set dynamically - meaning that an exogenous model parameter is updated over the simulation period. Apart from this, certain parameters as well as model structure are the same in all models. However, in our highly unsteady and complex world, model parameters and an appropriate system structure are deeply uncertain. Thus, different sets of assumptions need to be considered and multiple plausible futures explored. Therefore, the new design method was developed in a way that allows to explore the uncertainty space and analysis thousands or more model ensembles, instead of simulating only one or few scenarios.

4.1.3 Inputs and outputs

Finally, operations performed before and after the simulation are major components of the design method. Generating inputs as well as analysing the outputs is time consuming and requires to utilize advanced data analytics to process very large multidimensional datasets. Further, for managing and evaluating the spatial data, several GIS techniques are used.

One the input side, the needed data for the regional model is commonly not in the required format, hence before data can be used for parameterization of the models the collected raw data has to be transformed in useful form.

On the output side, policy analysis requires to interpret the spatially and temporally resolved model ensembles consisting of thousands of individual runs. This demands to adopt and develop techniques to visualize, classify and aggregate the simulation results.

4.1.4 Implementation tools

Nearly the complete realization, from data processing, setting up the simulation and analyzing the results, was performed in the high-level programming language Python. Python was selected as implementation environment for several reasons: First, it achieves great balance between flexibility and ease of use (Pérez, Granger, and Hunter 2011). Second, Python allows mixing different programming paradigms (e.g. object orientation, procedural). Third, across many scientific disciplines Python is popular for scientific computing and quantitative data analysis. Further, Python programming was performed in Jupyter Notebooks, a web application where live code, equations, visualizations and text can be interactively embedded (Pérez, Granger, and Hunter 2011).

Nevertheless, the single most important reason is that using Python as open source reporting tool fosters transparency and reproducibility of research. These two objectives in mind, deliberately, for this project only open source or free available software was used. The SD model was developed in the software Vensim PLE and handling and processing the georeferenced data sets was done in Quantum GIS.

4.2 Coping with the 7 modeling challenges

The integrated design method addresses Zika’s 7 functional requirements: (1) multi-regional disease models at subnational level are used to cope with subnational heterogeneity; (2) dynamics implemented at multiple levels enable to account for human connectivity; (3) by feeding monthly data into the model it is possible to obtain more realistic representations of seasonal dynamics; (4) applying simulation methods designed to cope with complex systems allows to manage inherent complexities; (5) a hybrid threshold approach from AB to SD allows to expand the spatial scope studied to the whole world; (6) under the conditions of deep uncertainty an explorative modeling and analysis methodology is adopted; (7) the projections made by a spatial correlation model are used to account for the recognized ignorance on the relation between environmental factors and vector.

The following sections elaborate further on how each requirement was addressed.

4.2.1 Subnational heterogeneity: Multi-regional disease models at subnational level

Location specific factors largely determine the dynamics of Zika. However, there are trade-offs between high spatial resolution and computational requirements. A full AB model could embrace high-resolution spatial differences, however it requires immense computational capacity and data availability at desired resolution. On the other hand, modeling at the level of countries or even higher aggregation level would entirely fail to account for spatial differences and the diffusion process of the disease.

Therefore, a middle course was pursued and a global-multi-regional model built. A core model structure describing the epidemiological process within each region was formulated and by adopting an object-oriented approach, this model structure reused in each region. The boundaries of the subnational regions were defined by dividing the world along administrative provincial divisions and joining similar regions within nations based on environmental characteristics for Zika transmission. The data needed to parameterize each regional model was derived from aggregating high-resolution gridded dataset to the described level of spatial differences.

Figure 10: Brazil's environmental suitability for Zika based on a species distribution model (Messina et al. 2016).
The high resolution raster dataset (left) could be used as input data for a Bayesian spatial model for disease mapping as well as a large-scale integrated ABM-GIS model with millions of agents. The data aggregated to national level (middle) would be useful for a single aggregated model in SD. For this research the data was condensed to similar regions within nations (right) and multiple regional model were simulated simultaneously.

Recent examples of similar approaches include Auping et al. (2016) who used a multi-regional model to study the impact of climate mitigation policies on social unrest. The key difference to our concept is that their model divides the world along national borders and regions are assumed to be independent with no interaction.

4.2.2 Human connectivity: Dynamics at multiple levels

The multi-regional models describe transmission dynamics within regions. However, at a higher-level the disease also propagates from one region to another. As explained, human travel is the primary driver for disease dissemination. Due to the importance of air-travel for potential rapid global spread, global air-travel was differentiated from passenger transport flows by other transport modes to adjacency regions.

Acquiring data about global air travel passenger flows is expensive and associated with restrictions on publications (Mao et al. 2015), therefore, open-access model based predictions for global air-travel passenger flows were used. At, at each time step, based on these passenger flows, the number of infectious individuals traveling to the destination is computed.

Data on other transport modes and passenger flows between neighbouring regions does not exists on global scale. If data becomes available, the same approach, used for dealing with air-travel, could be adopted. However, in the current revision of the model, adjacency regions are simply assumed to be equally connected and transmission occurs probabilistically, instead of being based on travel flows. At each time step, with some probability, single individuals are being infected in neighboring regions, to regions where ZIKVs are endemic.

4.2.3 Seasonal effects: Monthly data

Modeling the effect of seasonal climatic conditions on vector abundance and competence and as a result ZIKV propagation is challenging. In the current version of the model annual averaged data was used to consider the spatial differences in environmental suitability for Zika, but seasonal effects are currently not accounted for. In future work, monthly data may be used. Recently, Brady et al. (2014) presented gridded monthly maps of global temperature constraints on *Aedes* mosquito persistence and competence for dengue virus transmission. However, in contrast to the dataset used, their dataset does not incorporate other environmental variables linked to the mosquito distribution. Alternatively, seasonal climatic variations could be made endogenous to the model. However, this option was eventually rejected, since information on the effects are largely unknown.

Besides the effect of climatic conditions on vector abundance and competence, the seasonality of infectious disease outbreaks is also caused by variations in travel. Therefore, I used monthly air travel data, instead of annually aggregated data.

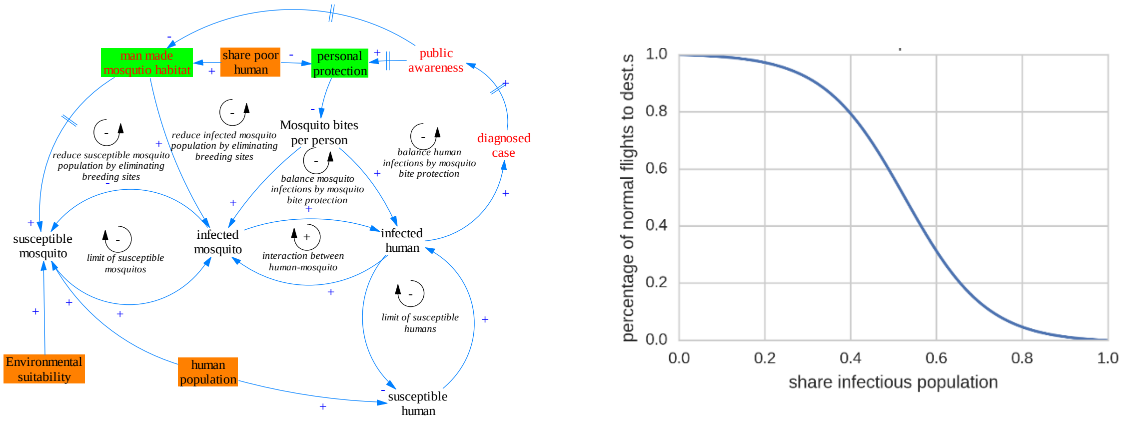
The impact of seasonality on infectious diseases is even more relevant, when studying reemerging infectious diseases. However, in this research, the scope was limited to the first wave of the outbreak. Studying a longer period would require to account for the effect that mosquito enter into diapause – a period of reduced development during the winter months and the complex process of hibernation. On the other hand, seasonal effects are significantly diminished, by the fact that the main vector *Aedes aegypti* breed indoors and indoor habitats are less susceptible to climatic variations (WHO 2016a).

4.2.4 Complex Feedback effects: Simulation models for complex systems

In the formulation of the core-model for each region, multiple feedback effects have been incorporated, besides the S(E)IR structure of epidemiology. Figure 11, left, indicates the multiple feedback loops in the system. For instance, as response to the perceived magnitude of the disease outbreak, it is assumed that the awareness for Zika as health threat raises and people increasingly protect themselves - apply insect repellents, and use physical barriers such as mosquito nets and screens. Furthermore, efforts to eliminate potential breeding sites are likewise driven by the perceived level of threat. The intensity of this effect depends on the policy actions performed.

In addition, also at higher level of interaction between regions through air-travel, relevant feedback loops were identified. It is to be expected that the number of incoming flights to a destination is impacted by the number of infected person in the destination. Therefore, air-travel passenger flows were weighted by the number of infected persons at the destination, following a (sigmoid) function normalized to a range from 0-1 (Figure 11, right).

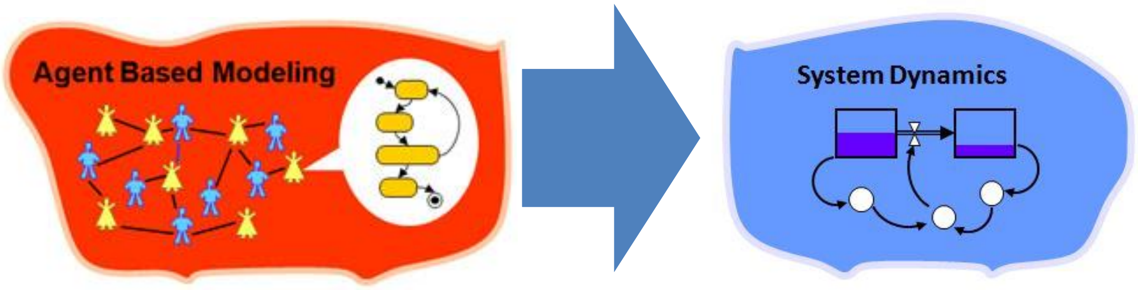
Figure 11: Aggregated causal loop diagram (left) which describes the transmission of the vector-borne disease as non-linear complex systems with multiple feedback loops – the diagram is presented in higher resolution in section 6.1. (right) Plausible relationship between share of infectious population at destination and the percentage of flights to destination the depicted function is controlling air travel to destination where disease outbreak is ongoing.



4.2.5 Global geographic scope: Hybrid-threshold model from ABM to SD

The dynamic interaction between human and vector is central for describing the problem. In addition, this process is the cause for the difficulty to combat disease spread. But detailed individual-based models come at high computational costs. A full agent based model at the level of individuals and worldwide scope would require to simulate several billions distinct agents and hence is not feasible without high-performance computing resources. As a result, in order to deal with the computational requirements, it is valuable to explore alternative simulation approaches such as switching from individual-based modeling (AB) to aggregated system descriptions (SD). The concept for switching during runtime between modeling methods originates from the observation that in the beginning of local transmission, the network topology between agents and few stochastic events determine whether this leads to sustained transmission. However, once the number of infected individuals is large, despite heterogeneity, the infection remains endemic until a large share of the population has been infected or other dynamic effects activate. Operational details and limitations of this pursued approach are presented in chapter 6.

Figure 12: Switch from ABM to SD ("Multimethod Modeling — AnyLogic Simulation Software" 2016)



4.2.6 Uncertainty: Exploratory Modeling and Analysis (EMA)

Only recently, the focus has turned to Zika as major health threat, consequently the virus has been only minimally studied. Thus, the course of the Zika outbreak and the effectiveness of suggested policies is deeply uncertain. The expression deep uncertainty was coined by Lempert (2003), referring to a situation in which the actors involved can neither agree on a particular model structure that describes the relationship between model parameters nor on the parameterization of model variables. Furthermore, beyond these rather conspicuous uncertainties, the method selection poses inherent uncertainty as well (Pruyt 2015). More specifically, the validity of the assumptions made to construct the epidemic model which incorporates spatial heterogeneity between regions and long-distance interaction is open to question. Therefore, wherever possible these uncertainties where likewise included in the analysis.

Designing robust policies – meaning policies which perform well over a broad range of possible futures- in a situation where so little information is available, requires an approach which explicitly accounts for the present uncertainty (Lempert et al., 2006). Therefore, instead of using the mathematical models for accurate prediction, following Bankes (1993), the model is used as explorative analysis tool to test hypothesis regarding how the world works under specific sets of assumptions. To identify robust policies, available options are tested over a broad set of scenarios and hypotheses.

To evaluate alternatives across a wide variety of plausible future many thousand simulation runs have to performed and hence models with a relative short runtime are required (Walker, Haasnoot, and Kwakkel 2013). Therefore, Walker et al. (2013) suggest the use of meta-models that mimic the behavior of more complex models. The presented multi-method approaches that swaps depending on a condition from AB to SD can be considered as such and is hence particular valuable in the field of exploratory modeling.

The Exploratory Modeling and Analysis (EMA) Workbench[7] provides a rich toolbox of methods and algorithms (e.g. LHS sampling, Random Forests and PRIM) to efficiently generate computational experiments and analyze complex and uncertain systems (Kwakkel and Pruyt 2013). Fully embracing this modeling perspective, the developed modeling framework has been integrated into the EMA Workbench from the start.

4.2.7 Recognized ignorance on the relation between environmental factors and vector: Spatial correlation model

Links between local environmental conditions and vector abundance and competence are assumed but remain little understood. Nevertheless, the factors are crucial for determining local transmission dynamics. Therefore, the outcomes of a species distribution models were utilized. Messina et al. (2016) generated a high-resolution gridded map, establishing empirical relationships between ZIKV vector occurrence and local environmental conditions. The dataset quantitatively describes as index value (from 0 to 1) six environmental conditions supporting the presence of *Aedes* mosquitoes and hence local transmission. The environmental factors considered are: (1) urbanization, (2) vegetation, (3) minimum relative humidity, (4) cumulative annual precipitation, and (5) temperature constraints for *Ae. Aegypti* and (6) *Ae. Albopictus,* respectively. For formulating of the model it is assumed that the index value of environmental conditions suitable for Zika virus corresponds to the mosquito-to-human ratio. In future work, the model could also incorporate the effects of environmental conditions on vector competence in addition to vector abundance.

---

[7] The Exploratory Modeling Workbench is available via https://github.com/quaquel/EMAworkbench

Part III

Implementation: preparing input data and building the models

5 Data collection and handling

The data used for the simulation model's input parameters was collected from blended publicly available sources. Geospatial data, spreadsheet databases, research articles and open-access modeling outcomes of reported research were integrated. The task of combining data from multiple sources required to perform transformation and aggregation steps to organize it into a manageable form. Furthermore, before usage, the datasets were checked for plausibility and consistency through several verification processes.

5.1 Data collection

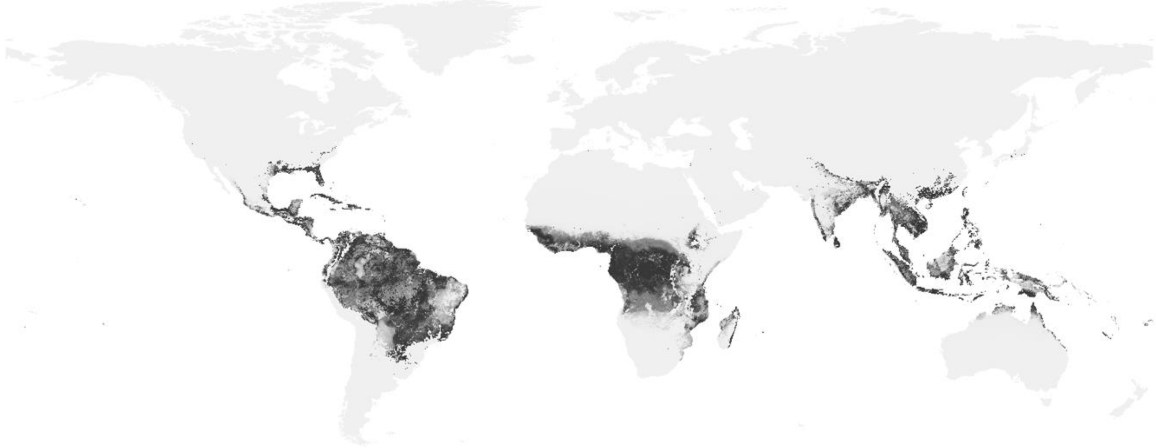
The data collection process can roughly be distinguished in two steps. First, a systematic literature review was conducted to obtain epidemiological information on ZIKV. Second, georeferenced data to account for the diversity of epidemic conditions across regions was compiled. The paragraphs below specify each of the GIS data sources of this study. Because epidemiological information about ZIKV transmission is scarce and the transmission cycle of ZIKV is widely unknown, estimates in the literature vary strongly. Therefore, some estimates used in the model were derived from epidemiological studies related to arboviruses dengue and chikungunya. Nonetheless, its resemblance with ZIKV enables the use of this data.

5.1.1 Environmental suitability for Zika virus

A high spatial resolution (5km x 5km) raster dataset mapping global environmental suitability for ZIKV transmissions was obtained from Messina (2016). The dataset describes quantitatively a combination of environmental conditions supporting the presence of the main ZIKV vectors *Ae. Aegypti* and *Ae. Albopictus*. The suitability is defined as standardized index value on a continuous scale from 0 to 1. The geospatial dataset was produced by compiling known records for ZIKV occurrence and a comprehensive set of high-resolution environmental covariates that presumptively impact the global suitability for ZIKV transmissions.

Six environmental factors were considered: (1) urbanization, (2) vegetation, (3) minimum relative humidity, (4) cumulative annual precipitation and (5) temperature suitability for *Ae. Aegypti* and (6) *Ae. Albopictus* respectively. Based on location, empirical relationships between the probability of ZIKV occurrence and environmental conditions were established. The authors reported that the cumulative precipitation (65%) and the temperature suitability for *Ae. Albopictus* (14.6%) are the main factors contributing to the variation in the vector occurrence. With an ensemble boosted regression trees approach, the suitability for Zika in areas where it has not yet been reported was inferred (Messina et al. 2016). The approach predicts that worldwide large parts of tropical and sub-tropical regions have suitable conditions for Zika (Figure 13).

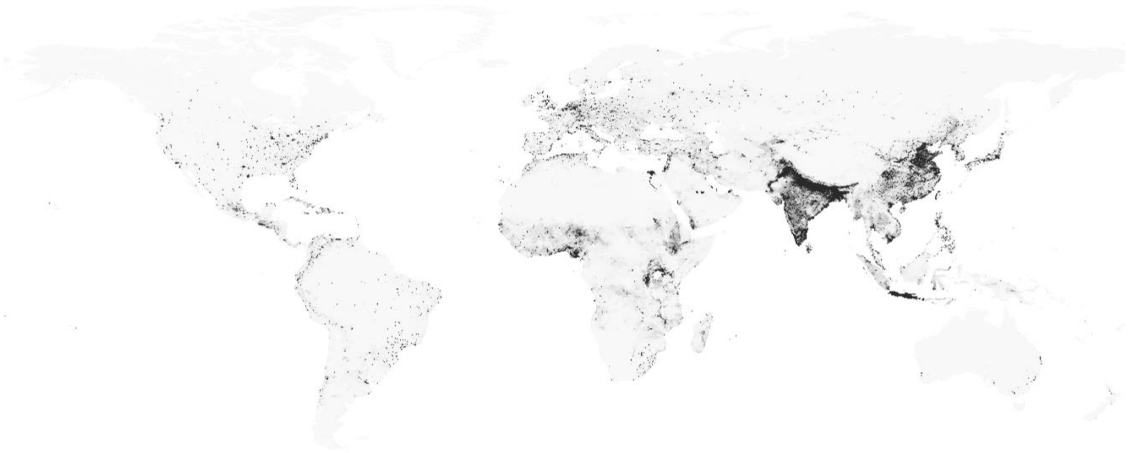
Figure 13: Raster dataset Environmental suitability for Zika virus (Messina et al. 2016)



5.1.2 Human Demography - Population and births

Population estimates by region were obtained from the most recent high-resolution gridded population dataset of the world, version 4 (GPWv4), released by the Center for International Earth Science Information Network (CIESIN 2015). Figure 14 visualizes the population counts for the world and shows how big China's and India's population is to the rest of the world. The number of potential Microcephaly cases have been estimated based on the crude birth rate (per 1000 people) annually published in the World Bank 's World Development Report (The World Bank 2016). Estimates were last updated for the year 2014 and are provided aggregated to country level.

Figure 14: Raster dataset population counts (CIESIN 2015)

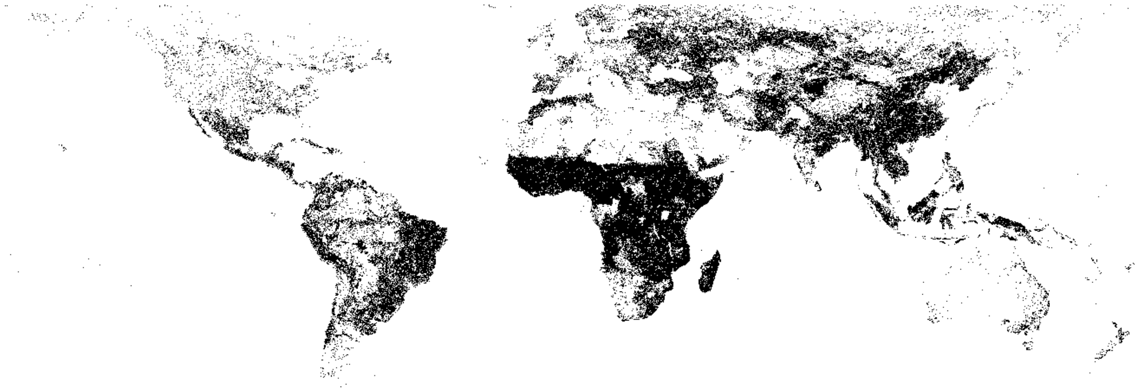


5.1.3 Poverty

The World Bank demarcates poverty as living on less than \$1.90 per day and \$2.50 a day, adjusted for purchasing power parity for global comparability (World Bank. 2015). However, for the purpose of this study using the World Bank's dataset on poverty is problematic, because the World Bank conducts its surveys only in the poorest countries of the world and does not report subnational heterogeneity. Hotez (2008) argues that also in economically developed countries (such as the US), there are a not negligible shares of impoverished people, which are unable to take appropriate measures for the protection from vector-borne diseases. Furthermore, these people live often in neighbourhoods that mosquitoes prefer as breeding habitat as a consequence of reduced municipal services (poor garbage collection) and low-quality housing (Hotez 2008; Phillips 2008). Therefore, high-resolution rasters on poverty levels were obtained and utilized (NOAA 2009). The map has been produced by combining satellite data on population and night-time lights, where the share of light per population is being used as proxy for wealth (Elvidge et al. 2009). Where data was available, the dataset was calibrated on the World Bank poverty line of people living from less than \$2.50 a day.

Figure 15 illustrates that the highest poverty levels are found in Africa and low levels of poverty are located primarily in Western Europe, USA, Japan and Australia.

Figure 15: High-resolution raster dataset on share of people living in poverty (NOAA 2009)



5.1.4 Air-travel passenger flows

Data on global air-travel passenger flows was obtained from open-access predictions of monthly flows of global air-travel passengers (Mao et al. 2015). The authors used a gravity model to estimate the air passenger travel between origin-destination pairs that are connected by nonstop passenger flights. The model relies on the well-established assumption that air passenger volume is proportional to the population size of the origin and destination region, and inversely proportional to the geographic distance between origin and destination (Gardner et al. 2012; Grosche, Rothlauf, and Heinzl 2007). Utilizing a set of Poisson regression models, the quantitative model was fitted to real travel data from 2010. Additional explanatory variables consist of flight routes, population counts, economic index, and several climate characteristics. The compiled dataset contains monthly passenger predictions between 3416 airports across the world and the airport's geographic coordinates.

5.2 Data pre-processing

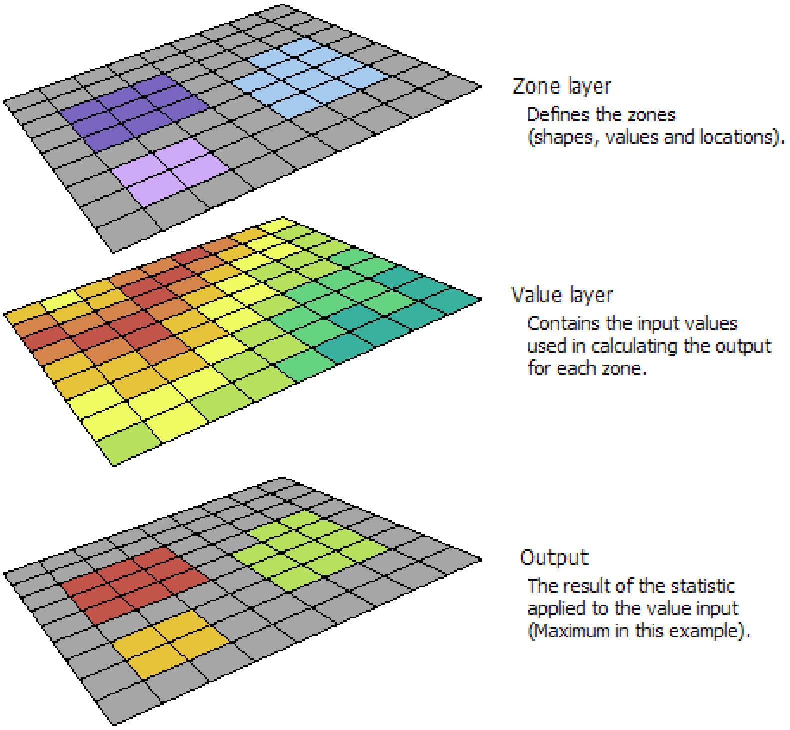
The described geospatial raster datasets were transformed into the common reference coordinate system WGS84 (EPSG 4326) for data alignment with the vector data of global administrative divisions on provincial and national level, which was obtained from (Natural Earth Data 2016; Natural Earth Data 2016)[8].

The raster dataset of global environmental suitability for ZIKV transmissions could not successfully be aligned in Python, QGIS or ArcGIS. As a workaround, the georeferencing tool of QGIS was utilized to (nearly perfectly) align the geo data. Derived from the vector dataset, thirty landmarks across the world, such as characteristic edges of islands, were used as reference points and their coordinates assigned to the pixel of the raster dataset.

5.2.1 Aggregation of raster data to appropriate resolution

Next, an appropriate spatial resolution for modeling was identified. From the disaggregated raster datasets, both national and sub-national estimates of poverty levels, population counts and environmental suitability for ZIKV were produced. Figure 16 demonstrated the basic concept of geo-processing the so called zonal statistics.

Figure 16: Illustration of computing zonal statistics (ArcGIS 2016)



---

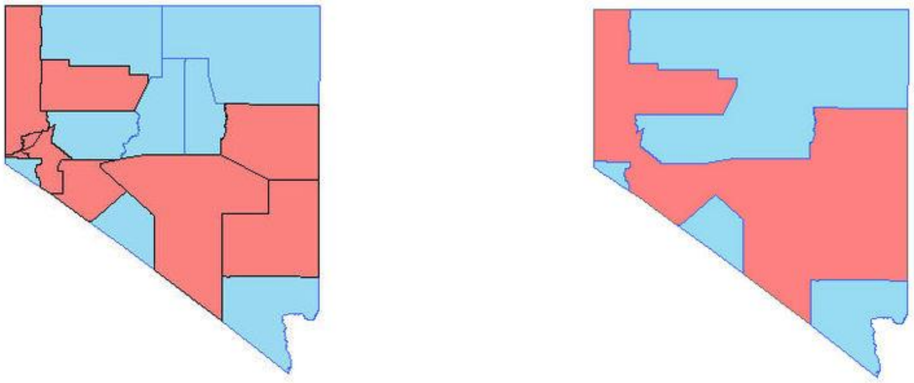
[8] The geospatial vector data representing the shape of the regions are stored in shapefile format, which is the de facto standard for this data type.

One the one hand, aggregating the raster data to national level is problematic, since some countries such as Brazil and the United States of America are very large. Consequently, within nations there are very different conditions for the propagation of the epidemic ZIKV. While Florida, Texas and California are suitable for ZIKV vectors, large parts of the US are not expected to be affected by the outbreak. A map aggregated to national level will predict that Wyoming and Florida have the same low suitability for Zika while in fact for Wyoming it is practically zero and in Florida it is high.

One the other hand aggregating the raster data to subnational level results in a very large number of elements (4647) and with each additional model element the computational requirements raise. Therefore, the idea evolved to group the lowest administrative provinces by shared environmental characteristics reflected in the index for Zika suitability but keep countries together for analysis purposes. This is achieved in two steps, first by discretizing the environment suitability into five buckets (one for all zero values and the rest of the data into quartiles). Second by grouping the data by (i) country and (ii) the new created bins for the environmental suitability.

Grouping geospatial data implies merging the geometric areas and computing summary statistics of the associated properties. In the present case, the sum of the population and for poverty and suitability the arithmetic mean. This aggregation process merging adjacent regions that have a common value for a specified attribute is referred to in GIS vocabulary as dissolve and illustrated in Figure 17.

Figure 17: Dissolve (http://wiki.gis.com/wiki/index.php/Dissolve)



Polygons before the dissolve operation.

Dissolving on the attribute (differently colored) results in fewer polygons with common values for all attributes.

After performing the described aggregation process, the produced map contains of 478 vector elements. Its result is depicted in Figure 18, Figure 19 and Figure 20:

Figure 18:Vector Environmental Suitability (index value)

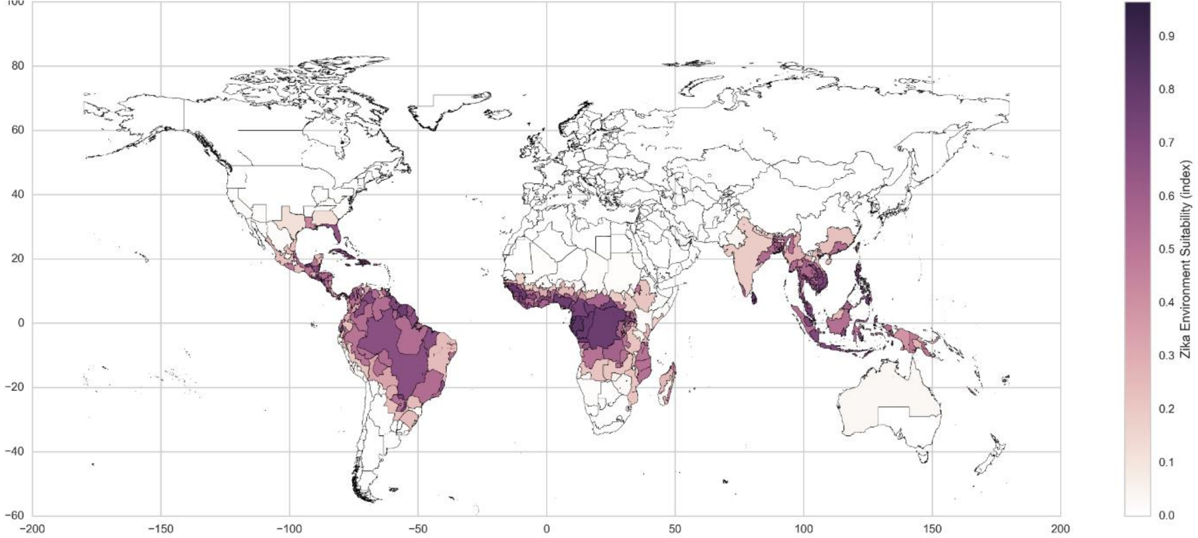


Figure 19: Population count in million

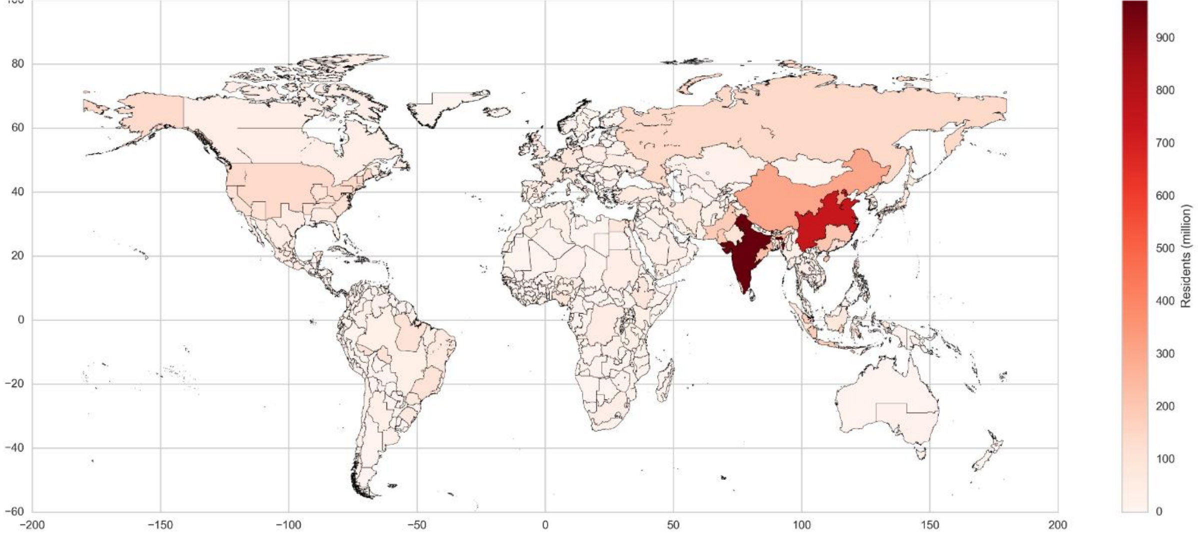
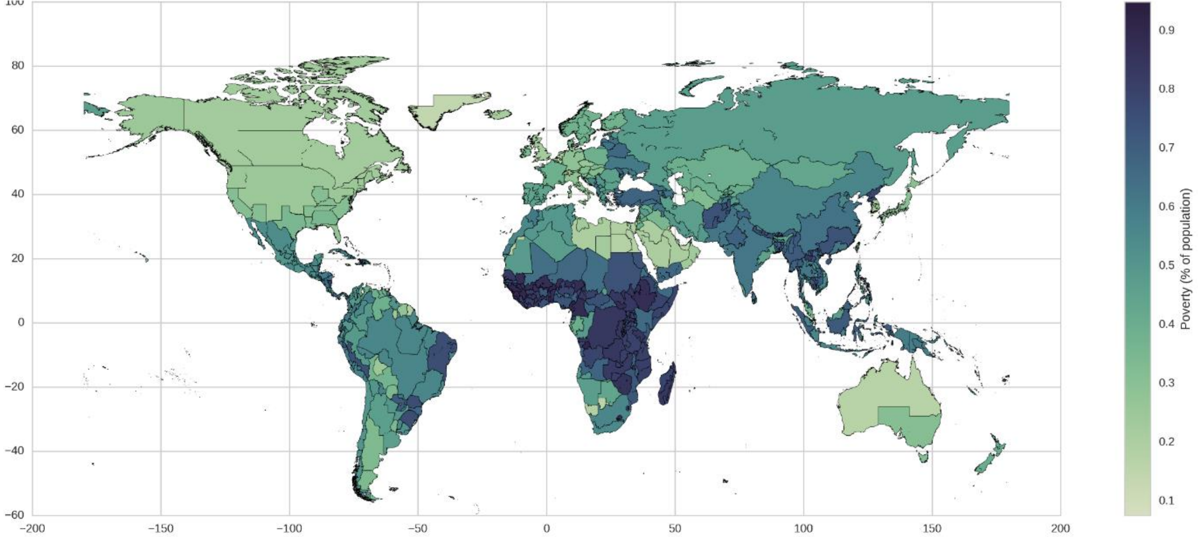


Figure 20: Poverty % of population



5.2.2 Integrate data on birth rate

Subsequently, the dataset on crude birth rate was merged to the prepared geo data. Missing values where filled with the average of the regions continent. The expected pregnancies per year in each region were estimated applying the formula:

$$\frac{1}{1000} * \frac{9}{12} * population * birth\ rate = \text{pregnancies per year}$$

The first element of the equation is needed to convert the crude birth rate (per 1000 people) to a per capita figure and the second element is used to account for the nine-month period of pregnancy.

5.2.3 Identify model elements

While usual human travelling patterns will bring the infection to all regions in the world, due to the absence of the vector in some regions, certain locations will remain in any scenario ZIKV outbreak free, despite imported infections and limited sexual transmissions. Thus, regions where the suitability for ZIKV transmission is vanishingly small were not included in the model. To reflect for potential error, a threshold cut-off of 1 % was chosen. Only regions with at least a suitability over this threshold were considered. Furthermore, justified by the global scope of the study, minor islands with less than 1000 inhabitants were also excluded.

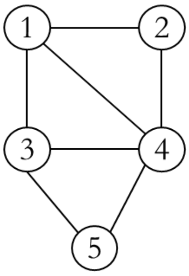
5.2.4 Compute adjacency regions

The network of adjacency regions (with a minimum suitability for Zika) was determined based on the regions' geospatial location. Adjacent regions where initially computed by touching regions' polygons. Using this approach alone however produced problems, since polygons included only land borders. Consequently, despite close proximity, neighbouring regions separated by lakes or seaside were not featured as adjacent, although naturally intense non-air travel is expected between those. For example, France and Britain are strongly connected by tunnels, bridges and ferry traffic, although these nations do not have shared land borders. To account for this issue, a Delaunay triangulation (DT)[9] was computed from the centroids of the regions. Combining the two approaches yielded satisfactory results. Alternatively, and avoiding to compute the DT of regions' centroids, shape files which include marine boundaries could be obtained and used to compute adjacency regions based on touching polygons alone. For processing and representation of the adjacency regions the data was saved into a symmetric adjacency 0-1-matrix which is schematically depicted in Figure 21.

---

[9] In computational geometry, a Delaunay triangulation is defined as set of points in a plane being connected by edges such that no point is inside the circumcircle of any triangle (Lee and Schachter 1980)

Figure 21: Schematic illustration of 0-1-adjacency matrix as undirected graph (left) and as symmetric matrix (right)

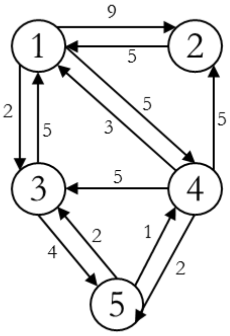


$$\begin{array}{c|ccccc|} & 1 & 2 & 3 & 4 & 5 \\ \hline 1 & 0 & 1 & 1 & 1 & 0 \\ 2 & 1 & 0 & 0 & 1 & 0 \\ 3 & 1 & 0 & 0 & 1 & 0 \\ 4 & 1 & 1 & 1 & 0 & 1 \\ 5 & 0 & 0 & 0 & 1 & 0 \\ \hline \end{array}$$

5.2.5 Air-travel passenger flows

The dataset on monthly passenger air-travel between airports (Mao et al., 2015) contained several duplicates, which were removed. Further, the raw dataset provided passenger flows between airports, but not the required travel flows between regions. Therefore, a spatial join[10] between the coordinates of the airports with the polygons of the regions was performed. In other words, the airports were mapped to the specified regions. Since the spatial join's execution time increases exponentially with the number of spatial objects, the computation time of this operation was more than four hours. R-trees[11] were applied for more efficient execution of the operation and shortening the processing time. Finally, the computed air-travel flows between regions' origin-destination pairs were saved as a bidirectional graph for further use. Figure 22 provides a schematic illustration of this data representation.

Figure 22: Schematic illustration of air travel flows as bidirectional graph (left) and as asymmetric matrix (right)



$$\begin{array}{cc|ccccc|} & & & & \text{To} & & \\ & & 1 & 2 & 3 & 4 & 5 \\ \hline & 1 & 0 & 9 & 2 & 5 & 0 \\ & 2 & 5 & 0 & 0 & 0 & 0 \\ \text{From} & 3 & 5 & 0 & 0 & 0 & 4 \\ & 4 & 3 & 5 & 5 & 0 & 2 \\ & 5 & 0 & 0 & 2 & 1 & 0 \\ \hline \end{array}$$

---

[10] Spatial join is a GIS operation that affixes an attribute from one feature layer to another based on the spatial relationship.

[11] R-trees are smart data structures that organize geospatial data such that indexing of data is considerably accelerated (Brinkhoff, Kriegel, and Seeger 1993).

5.3 Data verification approach

After conduction the pre-processing steps, a quick-and-dirty verification of the input data was performed. The process heavily relied on graphical representations that facilitates checking a large number of data points for their plausibility. Figure 23 shows the power of this approach visualizing the air-travel passenger flows across the world of all 3416 airports in the dataset.

Figure 23: Global air-travel passengers for June

Furthermore, data verification facilitates to identify errors in the data pre-processing and helps to correct them. For example, initially, the Delaunay triangulation did not span around the globe because the spherical shape of the earth had not been considered. As a direct result of the verification process, the adjacency graph was corrected by adding also the edges computed for the furthest-site of the Delaunay triangulation. With respect to the worldwide air-travel passenger network, the generated network was found consistent to the structure of a scale-free small-world network reported by (Guimerà et al. 2005). As measure for centrality, in addition, the eigenvectors of the network matrix were computed. This way another programming error was identified. After correction, now as anticipated, the United States is the most central region in the air-travel network of passenger flows. Graphs and detailed steps of the verification are presented in the Appendix.

6 Models’ conceptualization and behaviour

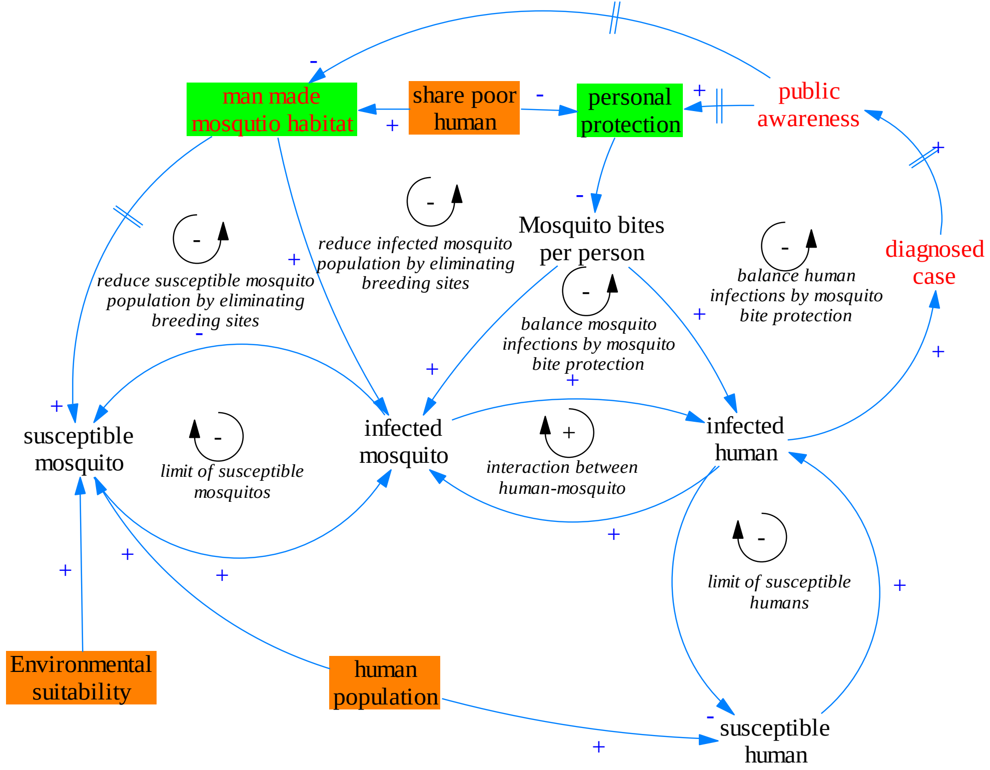
This chapter presents the core simulation model that was used for each region. First, the most important feedback effects are conceptualized in a causal loop diagram. Second, the qualitative model is transformed into quantitative ABM and SD models. Because the system description of the SD model is more tractable it is presented before the AB model. Third, the main behavior modes of the two models are presented and verification and validation steps are described.

6.1 Conceptualization of feedback structure

Model conceptualization aims at determining model boundaries, important variables, and feedback effects. Figure 24 shows an aggregated causal loop diagram that captures the main causal relations. In the diagram, causal links are indicated by arrows from one variable to another, and labeled as positive or negative. System structure that causes output from one parameter to eventually influence input of the same are in the following referred to as feedback loops. The loops were labeled and categorized into reinforcing positive loops (even number of negative causal links) and negative balancing loops (uneven number of negative links). Delays between causal relations are denoted by drawing two short lines across the link.

Furthermore, the system parameters were organized into exogenous local conditions (orange highlighted), policy-induced parameters (green highlighted) and parameters which are only present in the SD model formulation (red font). Further exogenous parameters that are globally equal, such as the recovery period, are not included in the diagram. The generated qualitative model consists of seven key feedback effects. On the one hand, the interaction between susceptible and infected mosquitoes and human drive the positive feedback loop of disease transmission from human-to-vector-to-human. On the other hand, there are multiple negative feedback loops balancing the trend of exponential growing infected individuals. The infectious disease dies eventually out, because the number of susceptible mosquito and the number of susceptible human diminishes when more become infected and acquire immunity. In addition, other psychological balancing feedback effects were identified. With the diagnosis of increasingly more Zika cases the public awareness for Zika raises and as a result four more negative feedback loops become active. This effects can be subdivided into the elimination of man-made mosquito habitats and personal protection. The former reduces over time the mosquito population while the later cuts down the mosquito bites per person.

Figure 24: Aggregated causal loop diagram for single region

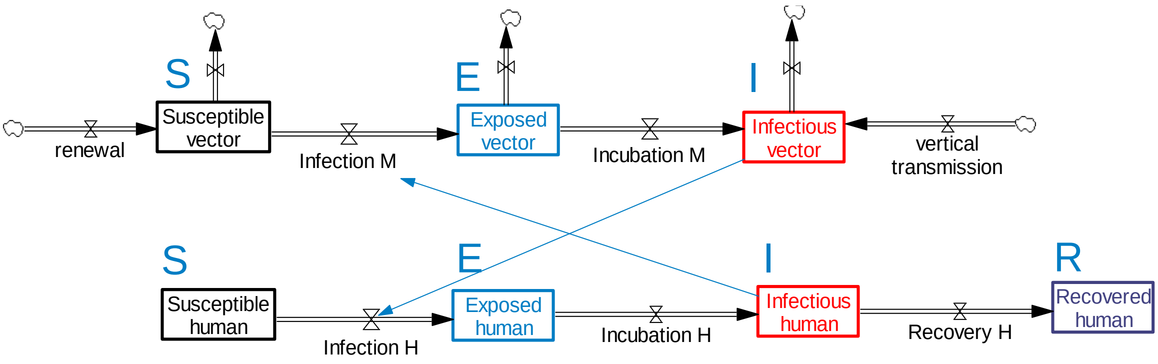


With respect to the number of persons infected, the expected behaviour of this model is exponential growth followed by eventual collapse, and corresponding to this a S-shaped growth of the cumulative cases of infected persons. Due to the cycle of birth and death of mosquitoes and only partial vertical transmission (mother-to-child transmission), new susceptible mosquitoes become continuously available. Therefore, the number of susceptible mosquitoes will first fall but then overtime recover to the level at the start. The pace of recovery dependents upon the speed of transmission relative to the replacement of mosquitoes (birth and death cycle). The dynamics of infectious and exposed mosquitoes follow generally the disease dynamics in the human population.

6.2 System Dynamic model

To define the quantitative SD model, the depicted causal loop diagram was transformed into a stock-flow structure. The spread of Zika is modelled by dividing the population in compartments according to the population's health condition. To the basic SIR compartmental model in epidemiology (Grassly and Fraser 2008), the exposed class was added to account for the incubation period in which individuals are infected but not yet infectious. Once infectious, mosquitoes remain infectious over their complete lifetime, therefore vectors do not have the class recovered (R). Figure 25 depicts the compartments of humans and vector and the interaction of susceptible and infectious in a simple stock flow diagram.

Figure 25: Human-vector transmission as schematic compartmental model



Because of the relative short simulation time of few months, for this model the human population size is assumed to be stable. In contrast, the lifespan of mosquitoes is only few days, therefore the model accounts explicitly for mosquitoes' birth and death. When no vector control measures are applied, it is assumed that the total vector population remains constant overtime. The flow of vertical transmissions relative to the renewal flow of susceptible vectors is determined by the share of mother-to-child transmissions.

All mosquitoes can bite all humans, but only the interaction between infectious and susceptible is relevant for disease transmission dynamic. When infectious vectors bite susceptible humans, the later become with a certain likelihood infected with Zika. As a result, those humans flow from the susceptible class to the exposed. After the expiration of the intrinsic incubation period, the exposed humans flow to the infectious class. Finally, the recovery period determines how long infectious humans remain infectious before they flow to the recovered class. Since it is assumed that once infected humans acquire immunity to Zika for life, humans, that have reached the recovered state dwell for the rest of the simulation time in this state. On the side of the vector, the dynamics are similar, susceptible vectors that bite infectious humans become infected with a specified probability and as a result, flow to the exposed class. After the expiration of the extrinsic incubation period this vectors flow to the infectious class and remain there until their death.

Figure 24 illustrates the three local conditions that determine disease potential in specific regions. The spatially explicit parameters are (i) environmental suitability, (ii) size of human population and (iii) share of people living in poverty. For this model the fundamental hypothesis is made that the vector-to-human ratio is equivalent to the normalized environmental suitability for Zika reported by Messina et al. (2016) as normalized index value between 0 and 1.

The regional parameters determine fundamentally the region's disease dynamics. The local human population sets the number of susceptible individuals and hence the maximally infected individuals in the region. The product of human population and the environmental suitability yields the number of susceptible mosquitoes. This parameter drives the speed of disease propagation and when low provokes that the disease dies out early, though a large number of susceptible individuals are available. The share of people living in poverty is relevant to account for the different vector exposure of population groups and the ability to take actions for personal protection and vector control. Thus, the epidemiological compartments have been further divided into poor and rich persons in the different health conditions.

In the current revision of the model it is assumed that both population groups interact with a single mosquito population. However, in future versions of the model different model structures or perspective of the world could be formulated and structural uncertainty explicitly addressed similar to recent research on Ebola outbreak and the future of copper. In this work, multiple SD models were used to account for deep uncertainty surrounding not only parameters but also fundamental feedback effects (Pruyt, Auping, and Kwakkel 2015; Auping, Pruyt, and Kwakkel 2014). For instance, it could be speculated that each population group has a mosquito population associated to it that only little interact with the mosquito population of the different income group. Fewer interaction between the poor and the rich may have significant impact on expected disease dissemination and potentially also on derived policy advise.

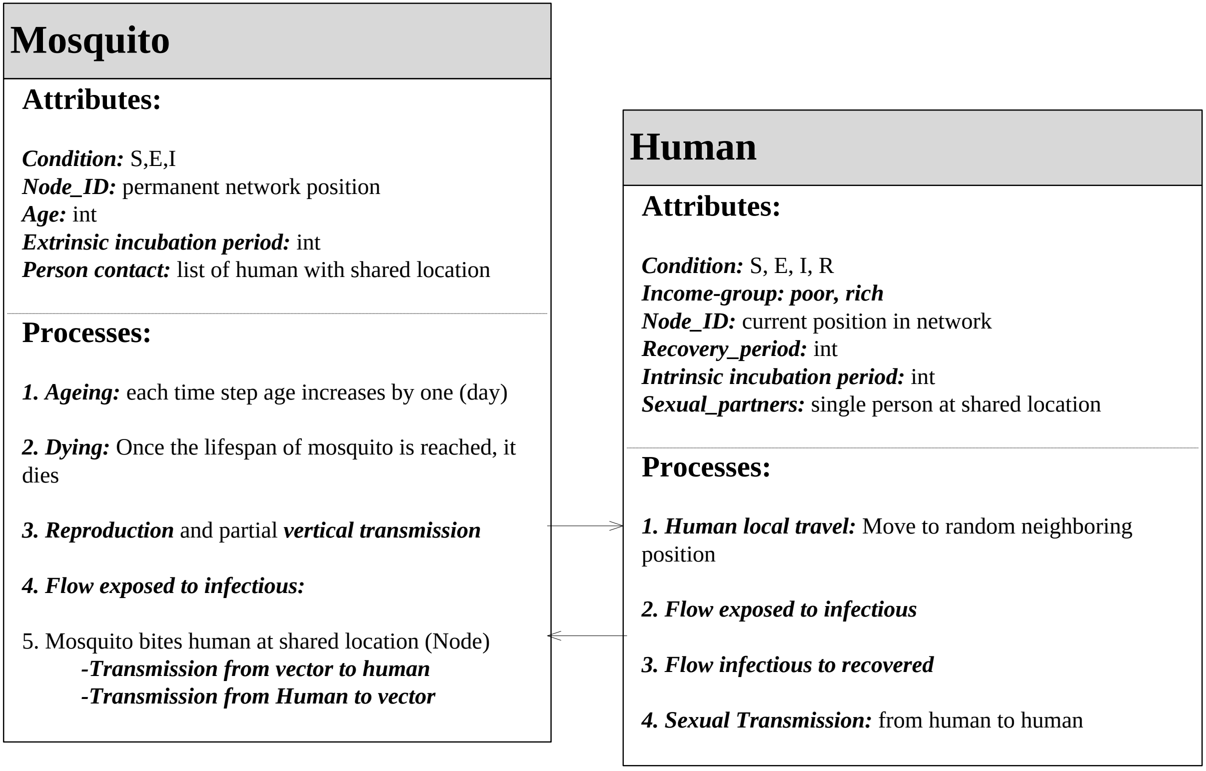
For simplicity, several causalities in the formulated model have been neither included in Figure 24 nor in Figure 25. Besides, vector-to-human disease transmission, direct human-to-human transmission is possible through blood transfusion and sexual contact (Musso et al. 2015; Patiño-Barbosa et al. 2015). Thus, a second flow from susceptible to exposed human is present, which is unaffected by the dynamics in the mosquito population. Furthermore, Fauci and Morens (2016) note that 4 in 5 with Zika infected individuals are asymptomatic and that it is likely that most of the asymptomatic cases will not be diagnosed. In the model, therefore, it is differentiated between asymptomatic and symptomatic cases, since many of the balancing feedback effects dependent on the number of detected cases. Furthermore, it can be presumed that persons without symptoms will not change their behavior pattern despite being likewise infectious.

While formulating, the SD model, several limitations were forced by the choice of using PySD to run the model. At the time writing, PySD did not allow to incorporate higher order delay functions and other complex functions. Hence, the model had to be simplified so that only first order delay functions are used. Nevertheless, it was achieved that the generated behaviour was correspondent to the behaviour of the more complex model that was formulated first.

6.3 Agent-based model

The AB model was constructed largely based on the same assumptions as the differential equation model. The AB model consists of two heterogeneous agent classes: human and mosquito. But instead of different compartments, health condition and income group are associated to individuals by agent parameters. Furthermore, in contrast to the deterministic averaged values SD model, in the AB model agent processes are stochastically and individual agent attributes are drawn based on distribution functions. Besides conceptual differences, the main distinction of the two models is that the balancing feedback effects driven by the awareness of Zika as global health threat are not included in the AB model. The reasoning behind this is that these effects only activates when a large share of the population has become infected and for this study the ABM model was only used to simulate the beginning of the outbreak. An overview of the agent attributes and processes is presented in Figure 26, further details including used distributions for parameterization and rules are specified in the Appendix.

Figure 26: Agent attributes and processes



Mosquito

Attributes:

Condition: S,E,I
Node_ID: permanent network position
Age: int
Extrinsic incubation period: int
Person contact: list of human with shared location

Processes:

1. Ageing: each time step age increases by one (day)

2. Dying: Once the lifespan of mosquito is reached, it dies

3. Reproduction and partial ***vertical transmission***

4. Flow exposed to infectious:

5. Mosquito bites human at shared location (Node)
-Transmission from vector to human
-Transmission from Human to vector

Human

Attributes:

Condition: S, E, I, R
Income-group: poor, rich
Node_ID: current position in network
Recovery_period: int
Intrinsic incubation period: int
Sexual_partners: single person at shared location

Processes:

1. Human local travel: Move to random neighboring position

2. Flow exposed to infectious

3. Flow infectious to recovered

4. Sexual Transmission: from human to human

Critical assumptions in AB model different than in the SD model:

The chance of human-mosquito interaction by mosquito bites are determined by the agent's location in the network. It is assumed that at every time step, mosquitoes bite humans which are at the same location. Therefore, disease dissemination is largely defined by the underlying network structure. Clustered and heterogeneous network topologies can create richer behavior patterns than homogenous networks and perfect mixing (Grassly and Fraser 2008). On the one hand, clustered and heterogeneous networks may cause very rapid spread of infectious diseases, on the other hand, clustered networks can lead to local saturation and dying out of infectious diseases, though a large number of remaining susceptible individuals. Nevertheless, a realistic network in which human and mosquito interact can only be approximated. Thus, to account for the deep uncertainty surrounding the network topography, we used four classes of network often used for the modeling infectious diseases: scale-free, random, 2D grid and small-world. The network characteristics are contrasted and illustrated in the appendix.

Inspired by Medeiros (2011), who built a cellular automata model to simulate the spread of dengue fever, it is assumed that humans move to neighbouring nodes in the network but mosquitoes remain at a fixed position. This modeling choice is motivated by flight range studies of *Aedes* mosquitoes which suggest that most mosquitoes spend their complete lifetime around a relative small area and that human travel primarily causes the wide spatial dissemination of diseases (WHO, 2016a).

Though evidence on the importance of the agent activation scheme has been reported (Caron-Lormier et al. 2008), often in AB simulation studies the selection of the agent activation scheme is not made explicit and impacts of choosing different activation schemes not explored. Since highly uncertain but with great impact on simulation results, for this study the two most popular activation schemes: asynchronous and synchronous agent updating in random order have been addressed in the uncertainty analysis. In the appendix the two agent activation mechanisms are shortly explained and the differences in generated behavior illustrated.

6.4 Behaviour modes and model testing

The two models –AB and SD- were tested to identify the different modes of behaviour that they were able to produce. For illustrative purposes, the separation of the population into a group of rich and a group of poor was overlooked. Instead, one single figure was tracked for the health condition of the human population.

In Figure 27 to the left, the behaviour of the human population according to their health condition is displayed. It can be observed that the expected behaviour is matched: infectious and exposed compartments follow exponential growth and subsequent collapse. On the other hand, the recovered immune population shows an S-shape growth.

In Figure 27 to the right, the reference behaviour of the mosquito population is depicted. Due to the cycle of birth and death and only partial vertical transmission, new susceptible mosquitoes become continuously available and the disease dynamics becomes more complex. Along with the propagation of the disease amongst the human and mosquito population, the susceptible mosquitoes decline. However, once the disease dies out due to the lack of susceptible humans, the susceptible mosquito population recovers to its initial level.

Figure 27: Disease dynamics in human and mosquito populations – SD model

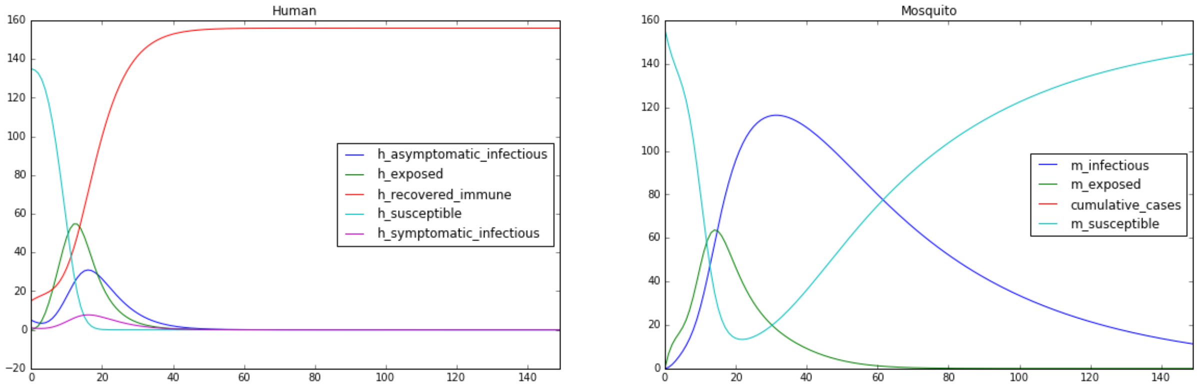
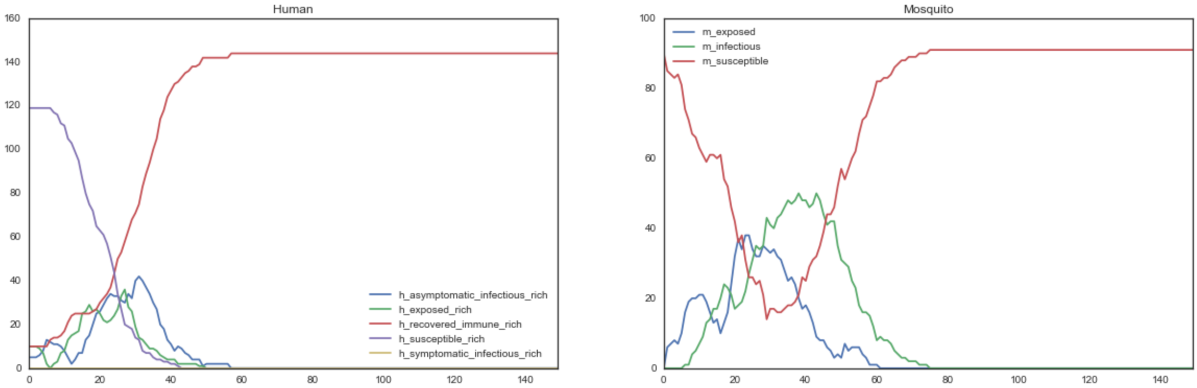


Figure 28 shows that the SD and AB models are equivalent and exhibit the same reference behaviour mode. The difference between these models lies in the stochasticity of the latter.

Figure 28: Disease dynamics in human and mosquito populations – AB model



Iteratively, the models were verified and validated by undertaking a number of different tests. These processes are reported in Appendix C.

Part IV

Gaining insights for the Zika case and looking beyond

7 Studying Zika with the proposed design method

This chapter presents the analysis of simulation experiments that were conducted. To allow reproduction of research results, first, a detailed description of the experimental setup is provided. Secondly, the obtained results are analyzed on multiple levels of aggregation: regional, national, continental and finally global level. The focus of the analysis is on gaining insights on how the design method is able to generate insights with respect to plausible dissemination of Zika and policies that can prevent larger epidemics. Because of deep uncertainty present a large uncertainty space is explored and for this purpose the experiments defined.

7.1 Experimental Setup

Table 1 provides a comprehensive overview of the experimental setup used. The reporting of the simulation experiments is based on the requirements formulated by (Rahmandad and Sterman 2012).

Pre-processing, simulation and post-processing operations were performed in the high-level programming language Python. The steps executed to generate the exogenous model inputs are roughly set out in chapter 5 and further detailed in the appendix. The Python scripts used for simulating and transforming the simulation outcomes to reported results are likewise fully documented and provided for reference in the appendix. Though alternatives are conceptually possible, deliberately simulation results for each region were collected individually and aggregation steps only performed after completing the simulation. This approach enables full flexibility for post-processing in which the simulation outputs is transformed to reported results. For each region, five parameters were stored for each time step of the simulation: the number of infectious, exposed, susceptible, recovered individuals and the number of cumulative cases.

Running a single experiment for all 307 interacting regions takes on average around 2 hours. The reason for this is the large number of interacting agents and other internal processes including sampling and initializing the models. Each region begins to simulate the initial outbreak as stochastic AB model and switches once the disease dynamics have established (>30 infectious individuals) to a differential equation based representation. Every regional AB model simulates until the switch point the interaction between 250 persons and a corresponding mosquito population. This subset of agents is assumed to be representative for the complete population in each region at the beginning of the outbreak.

The large simulation time was reduced by utilizing parallel processing. In addition, for running the simulations, two different computers were used. The reported performance difference is primarily caused by the number of available cores.

To study the first attack of the emerging disease, a simulation period of 180 days was chosen, where one day corresponds to a single time step. Seasonal exogenous parameters such as the monthly flight data are dynamically feed into the simulation model. For this process it was assumed, that the simulation begins in June.

Deep uncertainties associated to model parameters are addressed by sampling with Latin hypercube sampling over the complete range of plausible parameters. In addition, structural uncertainty was addressed on four aspects: Agent activation scheme, Agent network topology, Intensity of travel to adjacency regions and On/Off of air travel. In the appendix, in table form a complete overview of the ranges for the deeply uncertain factors, their best estimates and source is stated.

Table 1: Overview of experimental set up

Software platform used for the simulation	Python 32bit Simulation of SD model: pysd 0.3.2 Simulation of AB model: mesa 0.7.5 (customized)[12]
Hardware platform used for the simulation	Hardware Platform I: Intel Core i7-4500 CPU 1.80GHz - 2.40GHz, 2 Core(s), 4 Logical Processor(s), RAM 8GB Hardware Platform II: Intel Xeon CPU W3690 3.47GHz, 6 Core(s), 12 Logical Processor(s), RAM 24 GB
Computational costs of Simulation[13]	Hardware Platform I: 108 Scenarios: 49h ~ 27min/scenario Hardware Platform II: 324 Scenarios: 53h ~ 10min/scenario
Integration method and time step (differential equation model)	Stiff and non-stiff solvers[14] Dynamically varying time step[15] Reported time step = 1
Agent activation scheme (agent-based model)	Random order and Synchronous or asynchronous agent activation
Sampling method	Latin hypercube sampling (Ford 1990)
Iterations	no iterations of scenarios
Number of experiments (Scenarios)	1114
Simulation period	180 days, Start June 2015
Number of simulated regions	307
Initial infected Region	Single region in north eastern Brazil around Recife
Sampled parameter distribution	Uniform distribution
Parameter ranges	reported in Appendix
Pre-processing to generate exogenous model inputs	reported in Appendix
Post-processing to transform simulation outputs to reported results	reported in Appendix

---

[12] The Python library mesa was updated to include different network topologies and distinguish the activation of agent types.

[13] Simulation time obtained using parallel computing.

[14] Dynamic switch of integration method Adams-BDF (Hindmarsh 1983).

[15] Algorithm selects time step according to the local accuracy requirement (Hindmarsh 1983).

Policies were designed with the objective of combating Zika and preventing the disease's further dissemination. Since the dynamics occur on multiple levels accordingly the fight against disease spread can set in at different levels. On the one hand, measures can be undertaken to slow down regional transmission between human and vector. On the other hand, policies can be implemented that aim to prevent trans-regional spread.

(1) **Vector control:** By eliminating man-made mosquito habitats and breeding sites overtime vector control diminishes mosquito reproduction and reduces as a result the mosquito population. Actions can target the speed at which water ponds are destructed or focus on breeding sites specifically in high-income neighborhoods or low-income neighborhoods. In experimental setup, therefore three continuous parameters were defined that represent this possibilities: pond destruction time, release of measures breeding site rich and breeding site poor

(2) **Personal protection:** Personal protection reduces the mosquito bites per person and consequently the number of interactions between human and vector. As a result, transmission speed slows done, as less transmission events in both directions occur, one the one hand from vector to human and other hand from human to vector. Policies on personal protection can likewise target either the poor or rich or both. In the experimental setup, hence two continuous parameters were specified that represent the intensity of personal protection measures respectively for the two income groups.

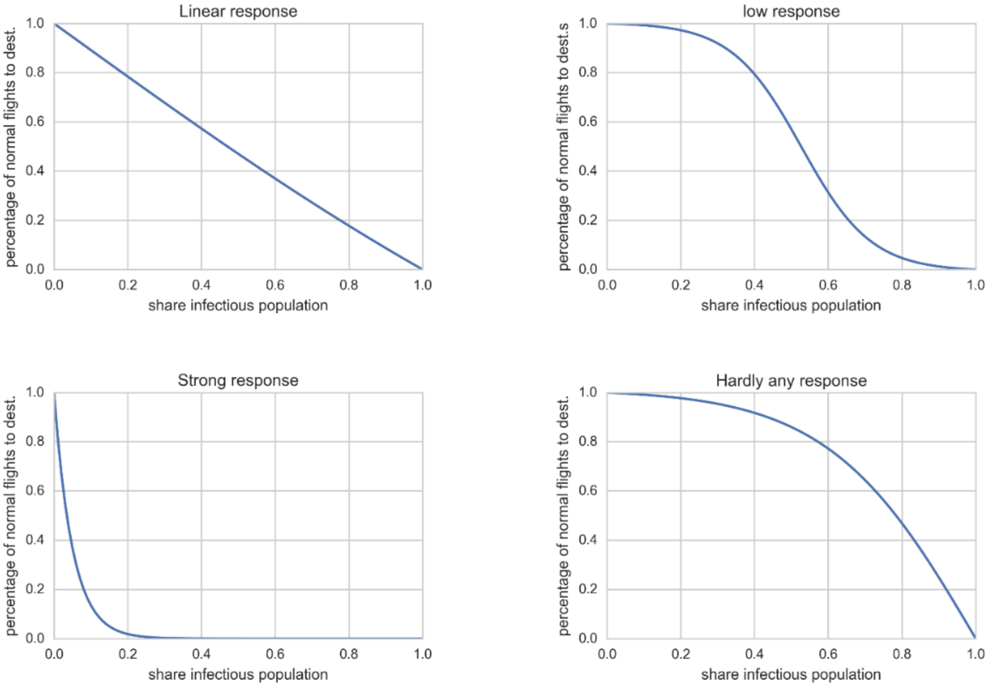
(3) Upgraded Zika Surveillance: The conceptual model for each regional model depicted as aggregated causal loop diagram in 6.1 demonstrates that the strength of the self-regulating balancing feedback effects depends on the awareness for the ongoing outbreak. Therefore, share of reported cases and the delay time of reporting are important policy parameters whose influence are explored with the selected experiments.

(4) Travel Advisories and restriction: Trans-regional dissemination potentially can be prevented by advising against travel to regions with ongoing Zika outbreak or more radically by declaring travel restrictions for this regions. Due to the differentiation made between air travel and other travel (to adjacent regions), travel advisories and restrictions are applied to the different travel modes.

First regarding air travel, possible advice by health authorities against travelling to regions with large Zika outbreaks are addressed in the experimental setup by sampling over different functions that represent the relation between the share of infectious in a region and the percentage of the normal air travel flows into this region. Implicitly it is assumed that travel advisories would augment the feedback effect that people tend to avoid travelling to regions with ongoing epidemics. Hence a sharply declining function applies. Figure 29 displays four basic function that may describe this relationship. When travel advisories prompt travelers to avoid traveling to regions with ongoing outbreaks, a strong response (lower-left) is expected.

Full air travel restrictions are modelled with a switch that turns off all air travel.

Figure 29: Illustrating the impact of air travel advisories on the travel behavior of individuals to regions with ongoing epidemic



Second, travel advisories may affect the travel between neighboring regions. Therefore, in the experimental setup the intensity of travel to adjacent region is variable between turning it off completely and three further categories from very intense to little.

Furthermore, simulation models may be classified by simulation period. On the one hand, rapid outbreaks of emerging infectious diseases that occur in less than one year. On the other hand, over a longer period of time the study of patterns of reemerging infectious diseases (Siettos and Russo 2013). Since, Zika is an emerging infectious disease and already the near future consequences are deeply uncertain, this thesis focuses on the (simpler) assessment of the population at risk in the first wave.

7.2 Multi-level Analysis of Results

After performing 1114 experiments the obtained simulation results were rigorously analysed with two purposes in mind. First to build confidence in the design method, and second to perform preliminary policy analysis. Therefore, outcomes were also checked for plausibility and consistency with the hypotheses of the conceptual model and examined for policy relevance.

The analysis was performed on different levels of aggregation. Starting from the raw disaggregated simulation results for each region. Subsequently, the data was aggregated to national-, continental- and global-level. This multi-level analysis approach was pursued to facilitate interpretation of simulation results and support the communication of research insights.

7.2.1 Crude results fully disaggregated (subnational level)

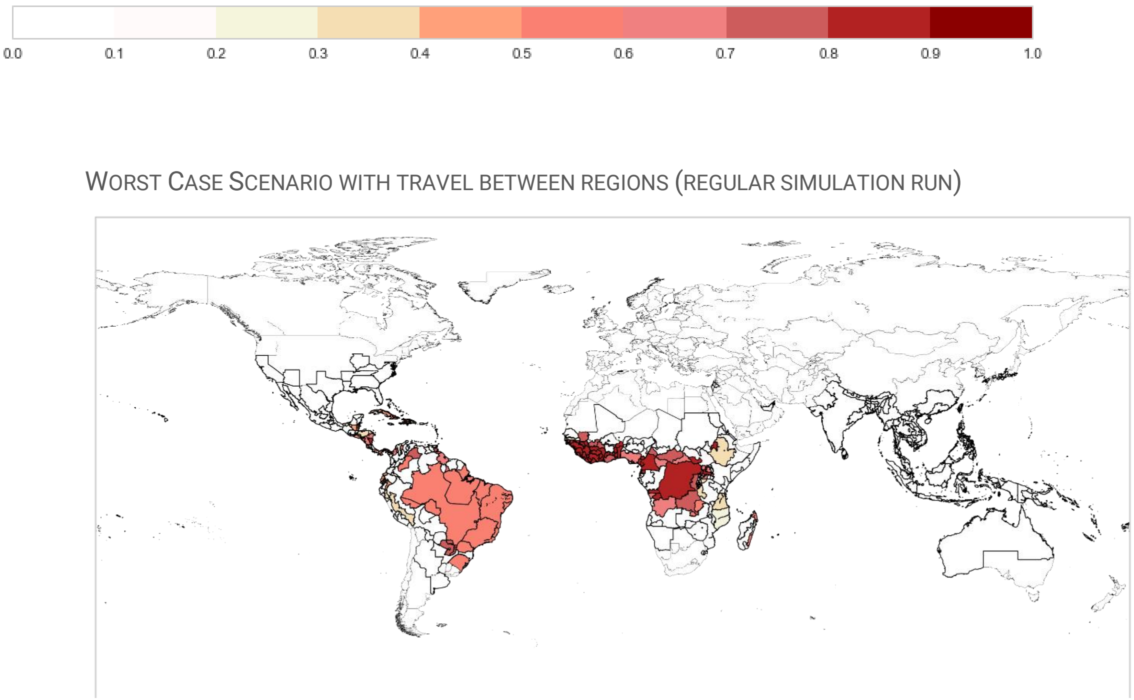
To analyse plausible dissemination patterns of Zika, the simulation results were projected on a world map. However, unlike, regular plotting of time series, were readily a large number of simulation runs can be displayed at the same time, for projecting scenarios on a map a choice has to be made which specific scenario is shown. For the purpose of validating the results, here scenarios were chosen, based on whether the switches for air travel and travel between adjacent regions were On or Off, entailing four different input combinations. The four map projections are illustrated in Figure 30. The colour coding represents the percentage of the population that has become infected until the end of the simulation time in the worst case scenario. Regions with minimal local transmission are highlighted by dark contour shape. A second issue encountered was that absolute values were incomparable across nations since their population sizes are so different. Hence, better indicators had to be developed. The cumulative cases as percentage of the population was found a useful parameter for visualization as well as interpreting the results.

It is apparent, that big Zika outbreaks affecting large shares of the population hit worldwide primarily equatorial regions particular western and central Africa and to a lesser extend south and Central America.

Furthermore, comparing the four maps across different experimental setups with respect to interregional travel demonstrates widely diverging outcomes. In the simulation run with no travel between regions (2), as expected, no transmission occurs across regions. Considering global air travel alone (4) lead to worldwide dissemination of the virus but transmission speed is slow and in none of the simulated scenarios large outbreaks occur. Without air travel but with travel to adjacent regions (3), Zika spreads to regions worldwide and the projection suggests that especially equatorial Africa is affected. In the scenario of air travel and other travel combined (1) the outbreaks are most common with larger outbreaks particular in Africa and in the Americas.

The spread of the disease from regions in South America to Africa and finally the rest of the world, in spite of the absence of air travel can be explained by the assumptions made generating the input data and specifically the adjacency matrix. Up to 3000km, regions are considered neighbours even across seas. From an ex-ante perspective this decision when generating the input data seems problematic as at the same time it is assumed that all adjacent regions are equally connected.

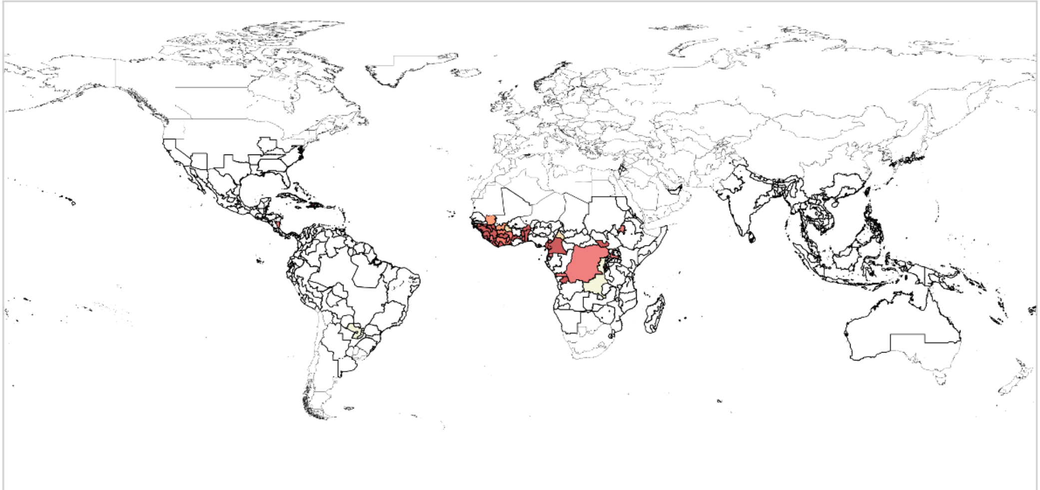
Figure 30: Cumulative cases as percentage of population classified by whether switch for air travel and switch for other travel (to adjacent regions) are ON and OFF.



WORST CASE SCENARIO WITH TRAVEL BETWEEN REGIONS (REGULAR SIMULATION RUN)

WORST CASE SCENARIO WITHOUT ANY TRAVEL BETWEEN REGIONS

WORST CASE SCENARIO WITHOUT AIR TRAVEL



WORST CASE SCENARIO WITHOUT TRAVEL TO ADJACENCY REGIONS

7.2.2 Results aggregated to national level

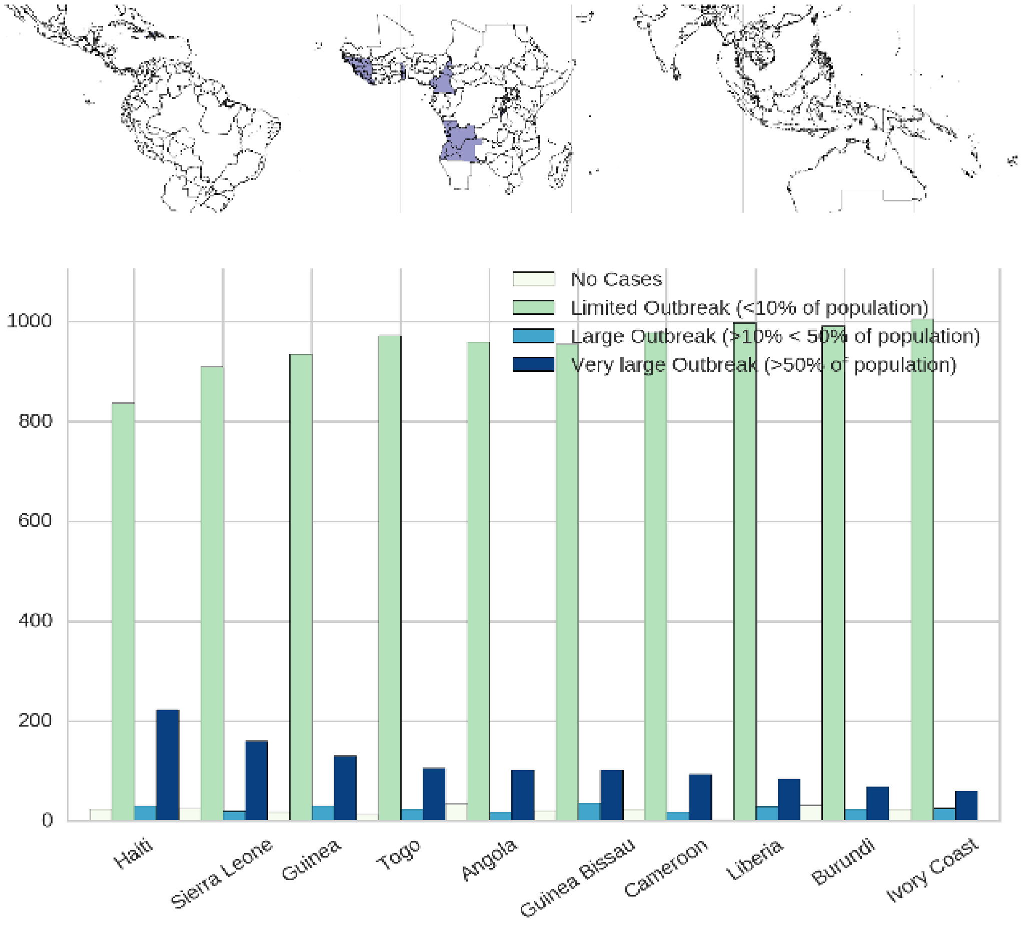
For further analysis, next, simulation outcomes were aggregated to national level. Analysing only one (random or extreme) scenario, as in the previous section, says little about the full ensemble of runs. Moreover, for exploration study it is likewise inappropriate to use summary statistics. Hence, to overcome this limitation cross tabulation was used.

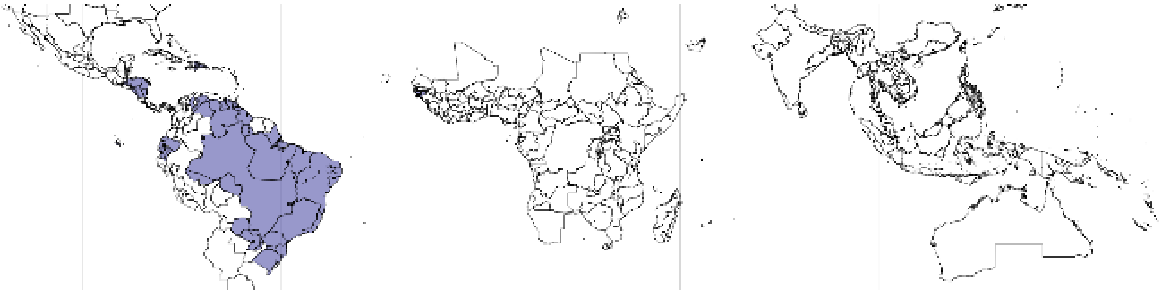
First, the severity of outbreak in each country was classified into two self-defined categories. On the one hand the severity of the outbreak was divided based on the total number of cases into: (i) few cases, (ii) limited outbreak with less than 1000 cases, (iii) large outbreak up to 1 million cases and (iv) very large outbreaks with more than 1 cases. On the other hand, the severity of the Zika outbreak was split by the share of the population that becomes infected over the simulation: (i) no cases, (ii) limited outbreak with less than 10%, (iii) large outbreak between 10% and 50% and (iv) very large outbreaks with more than 50% cases per population. Subsequently, cross tabulation allows to illustrate the relationship between nation’ outbreak category and run.

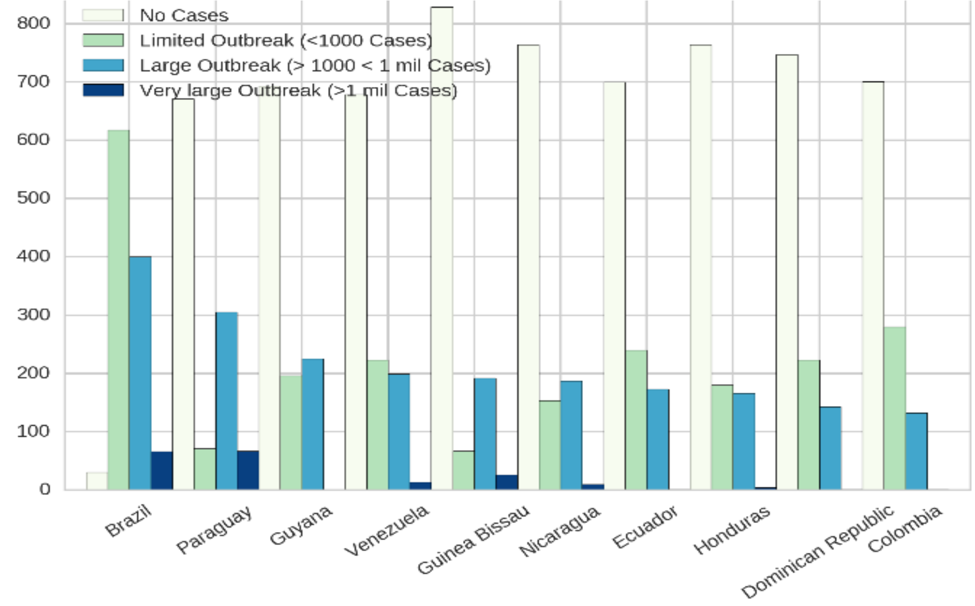
Figure 31 presents at the right side the top-10 nations that are most frequently hit by very large outbreaks with more than 1 million Zika cases. They are primarily located in south and middle America near the origin of the outbreak in a region of Brazil. Nevertheless, the figure at the at the bottom, demonstrates that measured by the population size other nations are even more impacted. Nations in equatorial Africa and Haiti are in many scenarios hit by outbreaks that infect or more than 50% of the population. Further analysis revealed that also measured by the number of cases these countries belong to the more threatened regions.

Besides, the graph suggests, that in most scenarios, only small outbreaks occur with few cases. This result is generated by the use of initially an AB model. Thus, in the majority of runs only the AB model is active. However, once the model switches to the SD model, in regions with highly suitable conditions, very large outbreaks are inevitable.

Figure 31: Top-10 threatened nations measured on cases per population (left) and total cases (right)



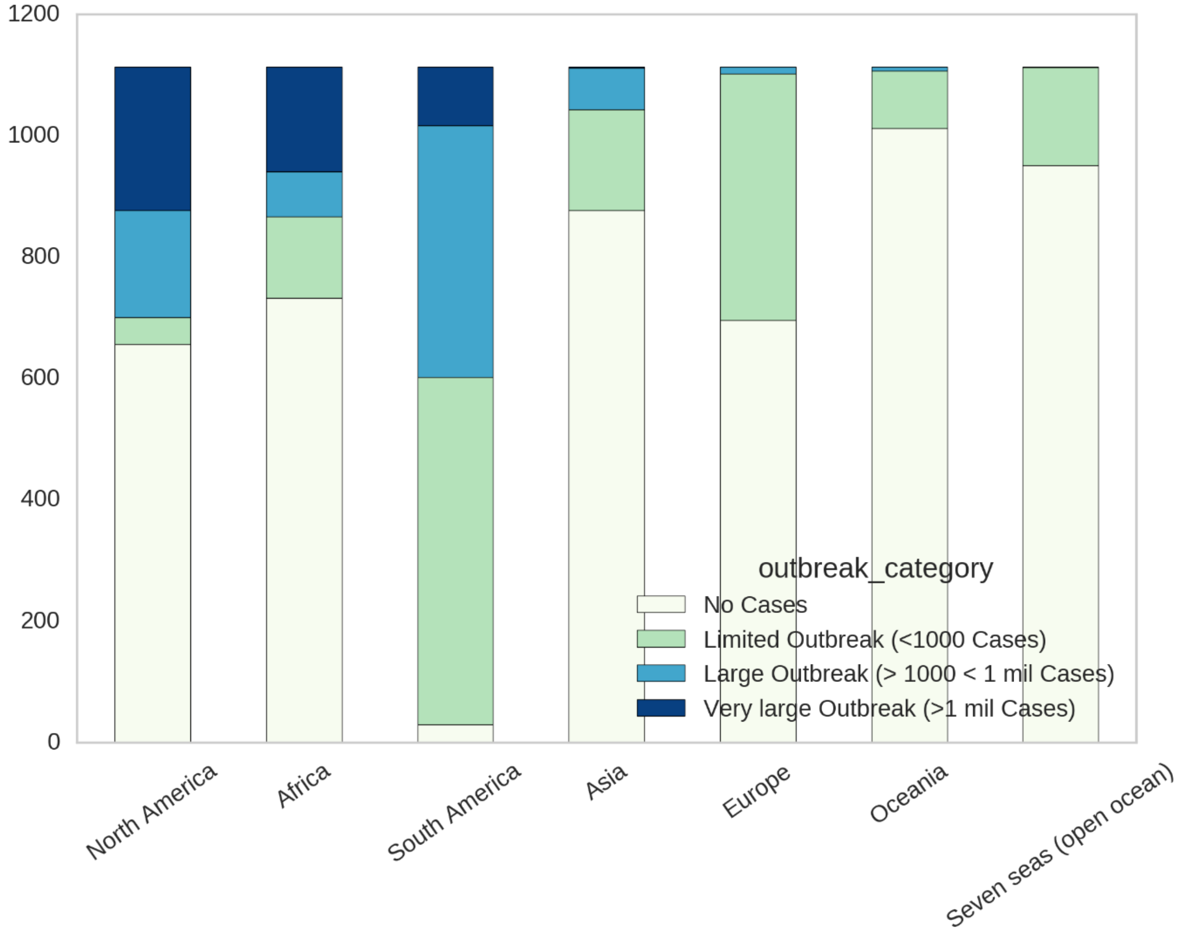




7.2.3 Results aggregated to macro geographical regions

Next, the results were aggregated to the level of continents and similar to the procedure described in the previous section, cross-tabulation was utilized to produce fruitful insights over all 1114 experiments. Figure 32 depicts that in nearly all runs in South America there is at least a limited outbreak of up to one thousand cases. In about half of the simulation results, the outbreak is larger and only in one-tenth of the runs more than 1 million persons become infected in South America. World regions that are on aggregated level more suitable for local Zika transmission are North America (including Mexico and other middle American countries) and Africa. If disease spread reaches this macro geographical regions large outbreak occurs and millions of people are infected.

Figure 32: Frequency of classified simulation results on continent level



7.2.4 Exploratory Analysis on global level

To further explore the parameter ranges responsible for specific outcome, eventually simulation results were summarized to global outcomes.

Figure 33 shows that global disease transmission with respect to the number of persons being infected occurs in large and small waves which superimpose. Regarding the infectious individuals, each region goes through a process of exponential growth followed by eventual collapse. When these dynamics are overlapped, complex nonlinear behaviour modes are produced. Further the figure suggests that at the time when the simulation stops, in some scenarios the disease dissemination is still ongoing and has not reached its peak. In addition, it is worth noting that the sharp bend visible in the right graph are caused by the switching mechanism from AB to SD.

Figure 33: Number of infectious individuals over the complete simulation period (left) and zoom into the first 30 time steps (right). Figures based on twenty-five representative simulation outcomes aggregated to global level.

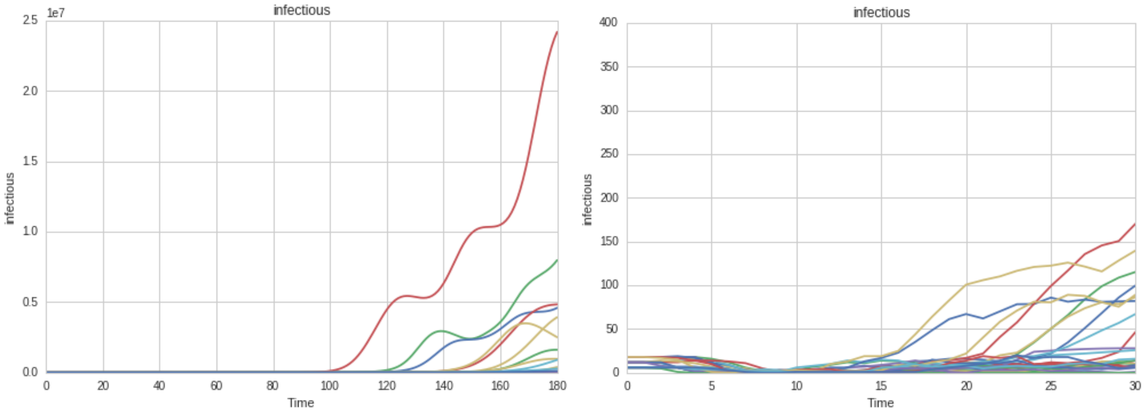
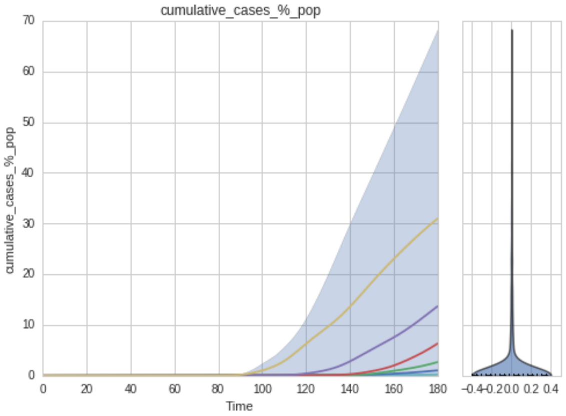


Figure 34 illustrates three major insights, first, the spread of the Zika is exponential, thus at the beginning, transmission speed is slow and it takes several month before larger outbreaks.

Figure 34: Full ensemble of cumulative cases per population with outcome at the end of simulation period shown as a violin plot[1]



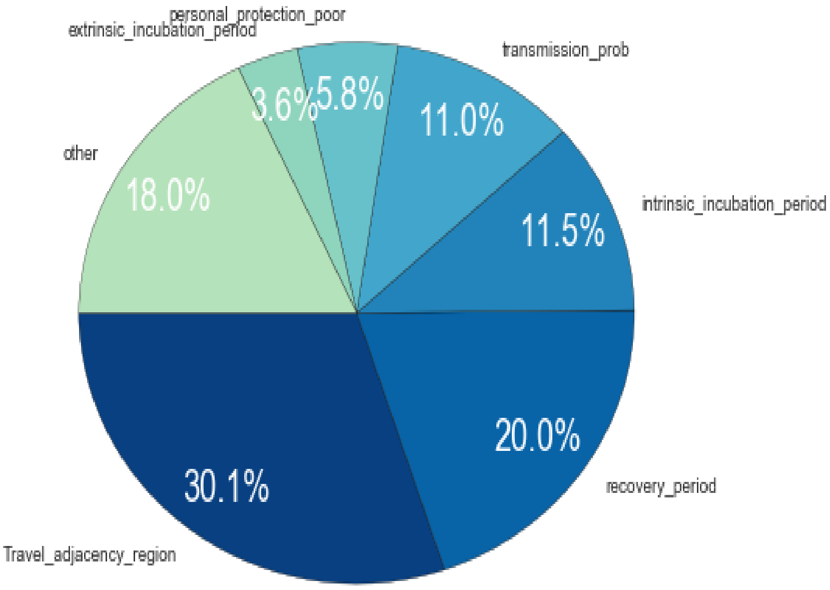
Second, the sparse and deeply uncertain information creates a very large input parameter space thus plausible scenarios are widely diverging, the blue area represents the complete ensemble of plausible outcomes.

Third, in most scenarios, the cumulative cases are low reaching not more than 5% of the population.

.

In a next step, the results were no longer analyzed alone but combined with the input parameters that produced this results. Consequently, what-if questions can be asked systematically and policy options examined based on the influence on the result.

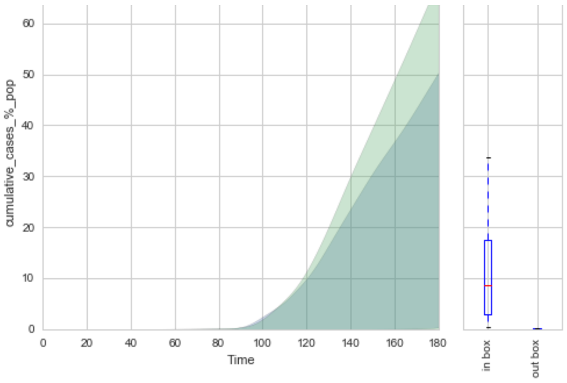
Figure 35: Relative variable's influence on ultimate cumulative cases



Feature selection algorithms allow to statistically determine which variable has the most influence on an outcome's parameter. Figure 35 depicts the relative importance of all input parameters for the variance in the cumulative cases at the end of the simulation period. The graph suggests that travel to adjacency region and the recovery period alone explain more than half of the variance observed in the cumulative cases per population. Variables such as agent activation, and agent network topology grouped all together into other play only a very small role.

The analysis can be extended by identifying the parameter combinations within the uncertainties which produce specific outcomes of interest. Therefore, the Patient Rule Induction Method (PRIM) was applied (Friedman and Fisher 1999; Lempert 2008). In a first step scenarios of interest were defined. For this research on Zika, two classifications were made. On the one hand, scenarios were classified as critical if more than 25% of the theoretically threatened world population becomes infected over the simulation period. On the other hand, experiments were defined as critical when more than 20 million total Zika cases have occurred at the end time of the simulation.

Figure 36: Envelopes of simulation results split by PRIM classification



The PRIM algorithm allows to identify then regions in the input space that produce a high density of scenarios of interest.

From the 1114 experiments, the algorithm found 264 case of interest were more than 25% of the population becomes infected. Figure 36 illustrates based on the shaded envelopes of both classes of runs the persisting great variety of simulation outcomes after applying the classification.

Figure 37 illustrates the trade-off to be made of selecting a particular box in the high-dimensional uncertainty space. In the upper left area, the boxes are very pure and the density of cases of interests up to 1. On the other hand, their coverage of the cases of interests is lower, meaning that there are critical cases which are outside the box and are produced by different parameter combinations.

Figure 37: Peeling trajectory I Trade-off between coverage and density

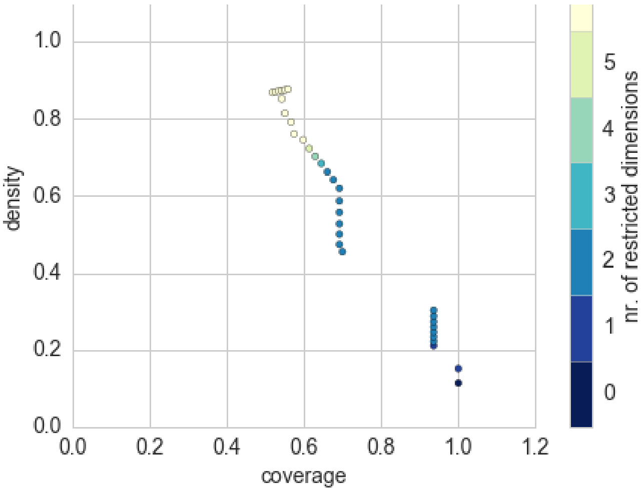
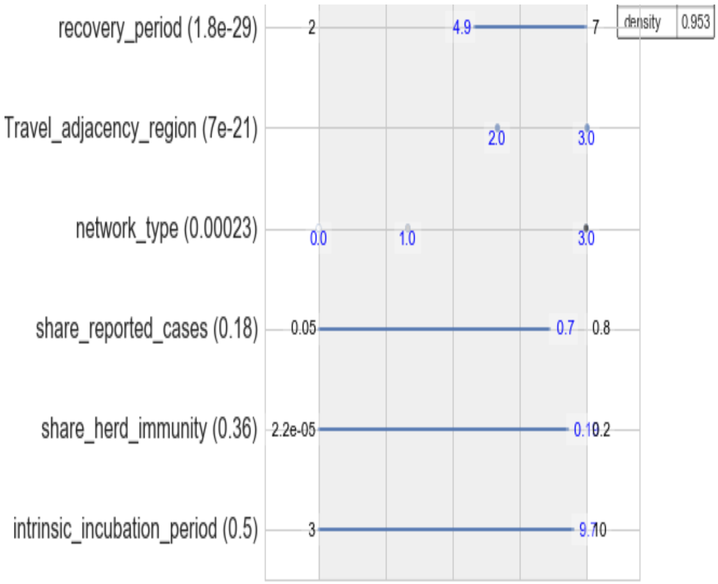


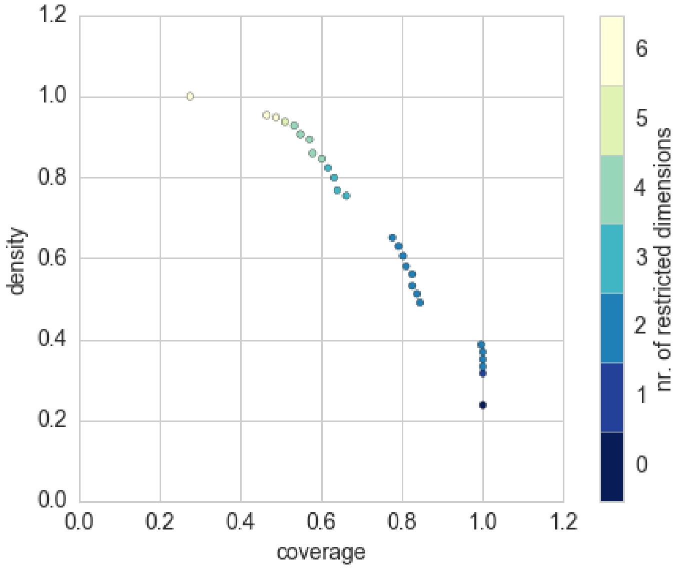
Figure 38: PRIM box for cumulative cases as % population



The PRIM box with a coverage of 47% and a density of 95% was selected. Figure 38 shows the combination of uncertainties for which the outbreaks are critical. Further assessment indicates that long recovery periods and intense travel is responsible for large worldwide epidemics. Surprisingly also the network type is identified to be critical, though feature selection did not found it as a large impact.

Small-world networks found the only network that does not generate larger outbreaks that could be explained by its property being highly clustered and therefore a high likelihood that the outbreaks stop in the beginning before many individuals have become infected.

Figure 39: Peeling trajectory II Trade-off between coverage and density

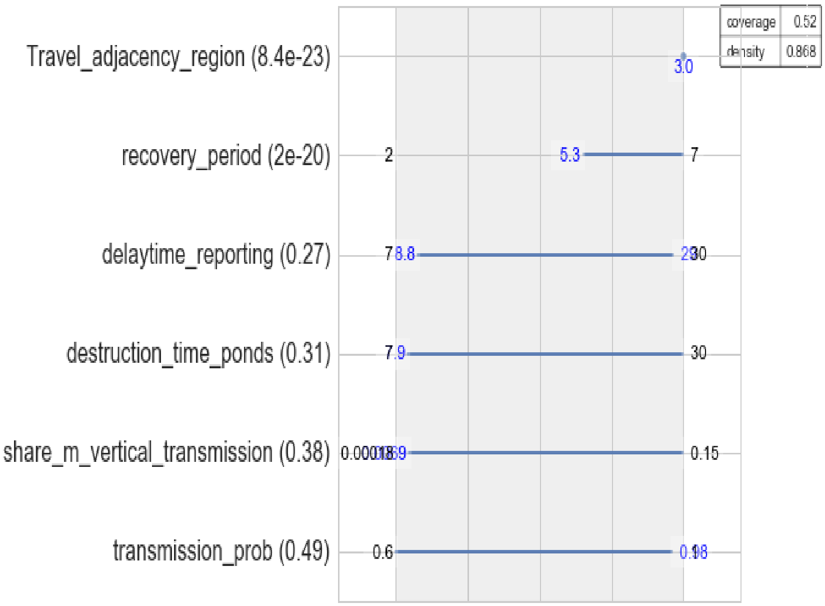


To check for consistency, also the total number of cases was explored following the same procedure. Based on the threshold criteria of more than 20 million cases 127 cases of interests were identified.

Just like before the trade-off between density and coverage has to be balanced. Several boxes were examined and eventually the PRIM box with a coverage of 52% and a density of 86% selected.

Figure 40 shows the uncertain parameter combinations which generates the critical cases. Again the intensity of travel to adjacent region and the recovery period are the two most relevant factors. In addition, and from a policy context interesting: when the delay time from occurrence of case until reporting is small between 7 to 9 days. In none of the scenarios a large outbreak with more than 20 million cases occurs. That is because under this conditions several self-regulating feedback mechanisms activate and the speed of transmission is significantly reduced. Further PRIM analysis suggests that a very fast destruction of the artificial water ponds would reduce mosquito population so that least this extremely critical scenarios are prevented.

Figure 40: PRIM box for cumulative cases

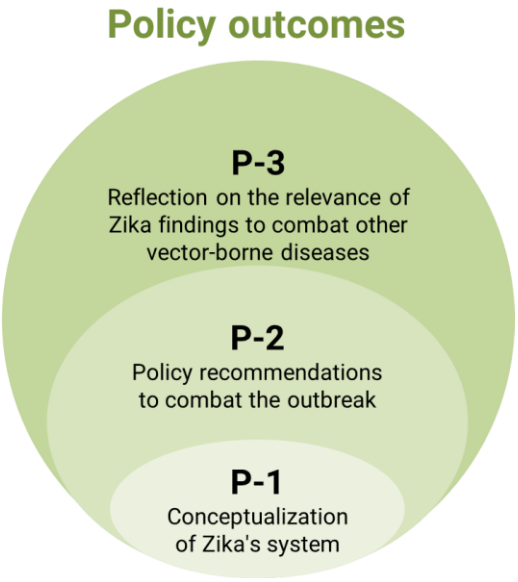


7.3 Policy recommendations

Due to the conditions of deep uncertainty that surround the epidemiology Zika outbreak, an exploratory modeling and analysis approach was applied to make a transparent identification of plausible futures under a certain set of assumptions. Though scenarios cannot be ranked by their likelihood, the aim for decision-makers is to be prepared for all or at least most critical futures.

Based on this explorative analysis, policy recommendations to combat the Zika outbreak and reflections on their relevance to combat other vector-borne diseases were drawn. They constitute of the second and third policy outcome expected from this research (Figure 41) and are presented in the following two sections, respectively.

Figure 41: Policy outcomes expected from this research



7.3.1 Findings to combat the Zika outbreak

Four main policies were tested to assess their effectiveness to prevent further spread of Zika. In this section, key findings that can be used to combat the Zika outbreak are summarized.

(1) **Vector control:** generally, vector control was found to be insufficient to control the Zika outbreak and should not be the only measure in place. Nevertheless, when undertaken, low-income neighborhoods should be the target areas, due to their vector-prone conditions. More specifically, these strategies should aim for short pond destruction times, as in a high density of scenarios, the absence of a critical outbreak came with short pond destruction times.

(2) **Personal protection:** the size of the Zika outbreak was not found to be sensitive to efforts to promote and increase personal protection. While these measures should certainly remain in place, they by no means constitute a robust response strategy to prevent or control an outbreak.

(3) **Upgraded Zika Surveillance:** policy analysis suggests that efforts to upgrade a surveillance and reporting system for Zika pay off under many scenarios. If the local reporting system is quick and Zika cases are correctly detected and reported, awareness about the threat raises and self-regulating feedback mechanisms activate. In affected regions, locals and travelers protect themselves better and make more

efforts to eliminate potential breeding habitats of mosquitoes. This implies that false positive (type I error) reporting is by far better than false negative (type II error) and generate positive side effects. These findings support the WHO's decision to release a public alert about Zika in the beginning of 2016.

(4) **Travel Advisories and restriction:** the analysis revealed that travel advisories would have only little effect on the dissemination of Zika. On the one hand, the negative effect on travel flows is limited. On the other hand, alternative dissemination modes exist which would remain unaffected. Furthermore, the economic and social costs would far outweigh the generated benefits.

Consistently over the complete ensemble of models, the recovery period (time human is infectious and can infect mosquito when being bitten) is the most important single variable that determines the speed of transmission and the Zika cases in the first wave of the global outbreak. This period has been reported being between 2 and 7 days for dengue. The PRIM analysis showed that when the recovery period is smaller than 5 days under any other made assumption no large outbreak with millions of cases occur. Because of the importance of this parameter for disease propagation, efforts should be made to determine it more accurately. In addition, PRIM analysis reveals that the share of vertical transmission is also an epidemiological key parameter. However, this parameter is highly uncertain as very different values have been reported. Therefore, controlled field tests should be initiated to establish better estimates for this variable.

It is worth noting that that some of the assumptions made in this study should be revisited in order to make decisions regarding robust response strategies. Further, it was not possible to quantify all relevant uncertainties. Nevertheless, the model produced plausible outcomes that can inform decisions in the combat against Zika.

7.3.2 Reflection on the relevance of Zika's findings to combat other vector-borne diseases

The spatial analysis disclosed that Africa not the Americas is the most threatened region based the share of the population that could become infected. More general, it could be studied whether this finding applies also to other vector-borne diseases. The simulation could support the argument that global attention is being gained when high-income nations are being affected by only small outbreaks. However, emerging and re-emerging vector-borne infectious diseases in other parts of the world reach only minimal attention if not extremely large and devastating.

One of the problems with the assessment of Zika is that it is unknown what share of people that have gained immunity in other parts of the world. One explanation that the link between Zika and microcephaly could be established is Brazil's relative good local health infrastructure and monitoring system.

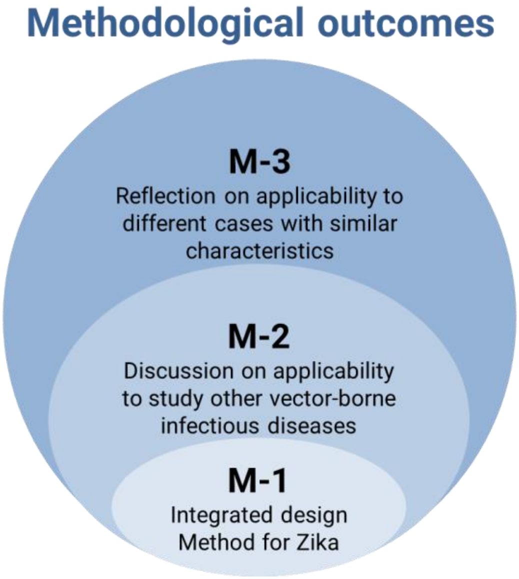
Overall, five lessons that may be applicable to these diseases were drawn from the study of the Zika outbreak:

> First, travel plays a crucial role in the diffusion of the disease and that early alert the measures that air travelers can conduct to combat exposure.
>
> Second, it was found that the recovery period of a population (time that a human remains infectious) has a large influence on the size of the outbreak. By empirically tracking this parameter and therefore reducing uncertainty, modelers can make better estimates of the future and design better policies to combat the disease.
>
> Third, there is a complex dynamic component in the occurrence of an outbreak. On the one hand, they are re-emergent. On the other hand, it is difficult to predict its time, location and speed of dissemination.
>
> Fourth, reporting and communicating incidences of disease cases is crucial as it triggers self-regulating balancing feedback loops in the population, i.e. people in the region and travelers make efforts to protect themselves better. Moreover, timely information enables the early detection of the outbreak, its study, and the design of effective strategies to prevent its further spread. Thus, producing early warnings may positively influence the control of the outbreak without other measures being taken.
>
> Fifth, model structure and parameters are deeply uncertain and there is a wide range of plausible futures for the outbreak that cannot be ranked probabilistically. Decision makers ought to design policies that are robust to all of these plausible outcomes.

8 Conclusions and Recommendations

This chapter is structured in two sections. In section 6.1, conclusions are provided by answering the research question and sub-questions. Then, in section 6.2, the three methodological outcomes of this study (Figure 42) are made explicit. First, critical reflections regarding the integrated design method are presented by discussing some of its fundamental choices and providing recommendations for its improvement. Second, its application to study other vector-borne diseases is discussed. Third, a reflection on its applicability to different cases with similar characteristics is presented.

Figure 42: Methodological outcomes from this research



8.1 Conclusions: answering the research sub-questions

The objective of this study was to answer the primary research question:

> *How can a design method suitable to identify response strategies to prevent the potential spread of Zika virus infections be integrated and applied, while only having sparse and uncertain information on the epidemiology and effectiveness of potential solutions?*

To answer this question, sub-research research questions were formulated and answered in each chapter. Thereby, qualitative and quantitative research methods were applied. In the following sections, the main research findings of each chapter are summarized.

Research Question 1 From a systems thinking perspective, what are relevant aspects of the Zika case that a simulation model should integrate in order to appropriately represent the problem and enable policy analysis?

This thesis presents a comprehensive system analysis of Zika by adopting a research perspective known as system thinking. The created qualitative model reveals that the disease dynamics of Zika are characterized mainly by the complex interaction between vector and human. However, the internal disease dynamics are additionally influenced by multiple spatially heterogeneous exogenous factors. Based on the system analysis for a simulation model with global scope that should appropriately represent the problem and enables policy analysis, seven fundamental functional requirements were identified: (1) subnational heterogeneity, (2) seasonal dynamics, (3) complex feedback effects, (4) global geographic scope, (5) deep uncertainty, (6) human connectivity, and (7) recognized ignorance on the relation between environmental factors and vector.

Research Question 2 How can the three single modeling and simulation methods (SD, AB, and GIS) represent the previously identified aspects and what are the methods' limitations?

This thesis argues that the three popular modeling methods agent-based modeling, System Dynamics and Geographic information systems are able to address each of the identified functional requirements individually but fail to account for all relevant issues at the same time. This finding motivates the integration of all three modeling methods.

Research Question 3 How can system dynamics, agent-based and GIS be integrated into a new design method to address Zika's 7 functional requirements?

This thesis proposes a novel integrated design method to study based on a mathematical model the potential for worldwide Zika outbreaks and use the simulation model to identify robust policies under deep uncertainty. The concept of the integrated design method is based on three elements: (1) Multiple interacting region-specific models, (2) big (georeferenced) data and (3) advanced pre-processing and post-processing operations.

Subnational heterogeneity is addressed by dividing the world along administrative provincial divisions and joining regions with similar characteristics within nations. Subsequently, parameterization of each regional model is performed based on aggregating raster datasets to this spatial unit.

Human connectivity between regions is incorporated by making use of open-access model based predictions for global air-travel passenger flows. The multiple regional models are simulated synchronously and after each time step information updated on how many contagious persons travel to a specific region.

Seasonal effects are assumed to be exogenous but are still incorporated by using monthly input data. Therefore, exogenous model parameters are dynamically updated during run time.

Complex feedbacks were identified in the model structure of each individual model and also in the interaction between models. For instance, it is hypothesized that travelers will avoid travelling to regions with ongoing large outbreaks. Besides, the used simulation methods System Dynamics and agent-based modeling were found suitable to represent the complex feedback effects of the system.

The conflict of target between relevant micro-behavior effects and global geographic scope were dissolved by starting the simulation with an agent-based model with a subset of the agents in the region and switching once a threshold condition is reached and the disease dynamics have established to a differential equation model (System Dynamic).

The presence of deep uncertainties give reason for adopting an explorative modeling and analysis approach. This approach allows informed decision-making, despite extremely scarce and uncertain information.

The relation between environmental conditions and vector abundance and competence is recognized but poorly understood. Simulation model-based approach require minimal knowledge on the relation between variables to specify its equations and consequently are not a viable option. This thesis presents an alternative to ignoring the relevant issue or making arbitrary combinations of assumptions. Based on the single assumption that environmental suitability for Zika's vectors is equivalent to the ratio of vector to human. The projections of an existing species distribution model which maps global environmental suitability for Zika virus could be used that condense the complex multivariate relationships between environmental conditions and vector abundance in a single index value.

Research Question 4 *What are the specific data requirements of the proposed design method and how can they be met?*

This thesis describes an in this form new approach to obtain georeferenced input data for multiple regional simulation models by transforming high-resolution raster datasets to vector data by applying several common GIS techniques. Further, this input data for the simulation model was combined by diverse other data sets, including open source maps on population estimates and census-based data.

Research Question 5 *How can simple simulation models be formulated to represent the key dynamics of the problem?*

For this thesis two largely equivalent simulation models were built, one the one hand a System Dynamic model and on the other hand an agent-based model. Both models describe disease transmission by the interaction between human and mosquito. The human population is divided into a SEIR structure and the mosquito population into a SEI structure. When a mosquito bites a human, infectious humans infect susceptible mosquitoes and vice versa infectious mosquitos infect susceptible humans. In addition, several self-regulating feedback mechanisms are present, which are produced by several social and psychological effects of people trying to protect themselves against the disease.

Research Question 6 *What insights on the Zika case can generated by applying the new integrated design method?*

This thesis demonstrates how Zika can modelled with the proposed integrated design method. Moreover, four policies were tested, gaining the following insights.

First, vector control was found to be insufficient to control the Zika outbreak and should not be the only measure in place. Second, the size of the Zika outbreak was not found to be sensitive to efforts to promote and increase personal protection. While these measures should certainly remain in place, they by no means constitute a robust response strategy to prevent or control an outbreak. Third, efforts to upgrade a surveillance and reporting system for Zika pay off under many scenarios. These findings support the WHO's decision to release a public alert about Zika in the beginning of 2016. Fourth, the analysis revealed that travel advisories would have only little effect on the dissemination of Zika.

Further, consistently over the complete ensemble of models, the recovery period (time human is infectious and can infect mosquito when being bitten) is the most important single variable that determines the speed of transmission and the Zika cases in the first wave of the global outbreak. Because of the importance of this parameter for disease propagation, efforts should be made to determine it more accurately. In addition, the exploratory analysis reveals that the share of vertical transmission is also an epidemiological key parameter. However, this parameter is highly uncertain as very different values have been reported. Therefore, controlled field tests should be initiated to establish better estimates for this variable.

At the moment, public media attention is focused on Zika outbreaks in South America and in particular on Brazil. Additionally, some subject matter experts stress the potential threat of large outbreaks that could also occur in southern parts of the United States. The simulation study supports this concerns as legitimate; under many simulated assumptions more than 1 million cases occur in North America. Besides, the analysis reveals that western and central Africa could be a center of especially big outbreaks, infecting very large shares of the population. Over all simulations, the top-10 most threatened nations, measured on cases per population, are except Haiti, all located in Africa. Only when making the assessment based on the total cases, nations in south and central America are more hit. Most frequently more than 1 million cases are projected for Brazil, Paraguay, Guyana, and Venezuela.

The used approach accounting specifically for subnational heterogeneity demonstrates that countries like Brazil, the United States as well as other are not equally affected throughout the country. Rather, transmission speed is differing and large outbreaks occur most often in regions with medium and high suitability for transmission. Nevertheless, under some set assumptions the differences between regions are smaller and there is little difference between regions.

Research Question 7 *How can the new integrated design method be used and improved in future work?*

The proposed integrated design method has proven to be useful to study Zika and could be applied with relatively little effort also to other vector-borne diseases such as malaria and dengue, that have similar characteristics as Zika. In addition, parts of the design method indicate potential to be also applied in other areas. Many complex problems in humanitarian and social science are characterized by deep uncertainty, subnational heterogeneity, complex feedback effects on multiple levels and other aspects which have been addressed in this research.

Regarding the application of the design method to design responses to Zika, many extra efforts are needed to refine the proposed concept and reach acceptance for the outcomes among decision-makers such as health authorities and parts of the research community. First targets for further research should be to revisit the assumptions made for generating the input data and constructing the model and check them for consistency with reality to build more confidence in the model. For this, an expert validation process should be carried out. In line with this the model outcomes should be calibrated to real-world data. In addition, attempts should be initiated to incorporate further important effects for the disease transmission, for example representing adequately seasonality of climate conditions including precipitation and temperature.

Additional recommendations to improve the design method are presented in the following section.

8.2 Discussion on methodological outcomes

The main contribution of this study is a new integrated design that is suitable to identify response strategies to prevent the potential spread of Zika virus infections while only having sparse and uncertain information on the epidemiology and effectiveness of potential solutions. The conceptualization and specification of this design method were presented, and it was applied to test multiple response strategies. In the following sub-sections, its use and improvement are discussed for the case of Zika, for other vector-borne diseases and for different cases with similar characteristics.

8.2.1 Using the design method to learn about Zika

Policy insights for the case of Zika were presented in Chapter 5 and summarized in the previous section. While the design method proved to be suitable to produce valuable insights, future work should focus on improving its shortcomings and seizing its capabilities.

On the one side, the assumptions made for generating the input data and constructing the model should be revisited. For this, an expert validation process should be carried out and the model outcomes should be calibrated to real-world data. In addition, attempts should be initiated to incorporate further important effects for the disease transmission, for example representing adequately seasonality of climate conditions including precipitation and temperature. Further, the hypothesis made that only air travel but not travel to neighbouring regions is affected by the perceived threat to travel to a county or region with ongoing outbreak should be revised.

By addressing these points of improvement, further versions of the design methods to study the Zika outbreak would become more accurate. Nevertheless, special attention should be paid to the collection of data, as this process can be challenging and time consuming and assumptions made in other studies to generate the data are transferred to the own model and may not be transparent. Moreover, an investigation on whether the values computed by the model correspond to known historical data should be conducted to build confidence in the model, keeping in mind that a good fit with data from the past does not guarantee a good fit in the future. Likewise, sufficient resources should be allocated to the verification of the design method as its implementation in a programming environment makes it error-prone.

An additional challenge is the computational requirements and need to handle extremely large data sets. In this study, the simulation results consist of nearly 250 million individual data points, representing 1114 simulation runs, 4 KPIs, 307 regions and 180 time steps. This pushes the available hardware (64bit Python, 8GB Memory) quickly to the absolute limits. As a result, post-processing operations like loading, saving, grouping and aggregating data takes minutes up to hours to execute. In addition, the design method has high computational requirements for its simulation, in spite of several implementation fixes to enhance its performance. Thus, future research should explore alternative ways to improve code efficiency and memory usage.

On the bright side, the choice of modeling local conditions, i.e. poverty and population, in a subnational disaggregated manner by adopting a region-specific multi-model architecture proved to be a viable and promising approach. Future improvement of the design method should pursue this direction.

Similarly, the visual components of the design method should be maintained and when possible, improved further. Projecting data on maps enhances the viewer's understanding. It allows to study the data that is initially fed to the model. Further, it makes possible to interpret simulation results in a more transparent manner by depicting time series.

Likewise, using an entity-based global-multi-regional model enables the computation of key performance indicators on different aggregation levels: provinces can be aggregated to countries and these in regions and eventually to global parameters. This capability increases the relevance of this design method for policy analysis and decision-making.

An additional line of research consists of possible extensions to the design method. For instance, this study did not consider seasonality as part of the model's environmental conditions. However, its patterns play an important role in the long-term persistence of the virus and on the size of an initial outbreak by dynamically limiting mosquito density. Other diseases, such as chikungunya and dengue, which may be co-circulating in a population were not considered. Moreover, demographic data involving details about households, location of schools and workplaces, and population mobility were not taken into account. Extensions to the design method could also include different data collection techniques, such as geo-data sets derived from digital satellite imagery, web scraping, and social media feed in real-time.

Finally, it should be noted that alternatives to the integrated design method exist to design response strategies to control vector-borne diseases. Massive agent-based models in distributed platforms are being developed for high-resolution agent-based modeling on global scale (Parker and Epstein 2011). Another example is Eubank et al. (2004) who addressed the use of EpiSims, a detailed agent-based simulator which incorporates data from population mobility and epidemic models.

8.2.2 Transferring the design method to study other problems

As previously explained, Zika exhibits characteristics similar to other vector-borne diseases, such as dengue and malaria. By revisiting assumptions and updating parameters to case-specific values, the design method can be applied to study other vector-borne diseases.

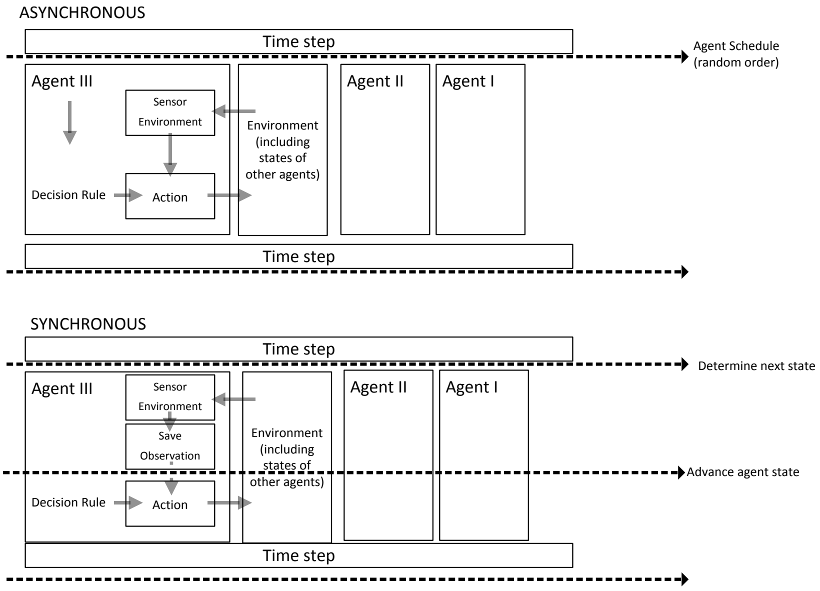
Similarly, the design method could be applied to cases beyond vector-borne and infectious diseases that share characteristics similar to Zika's 7 functional requirements: (1) subnational heterogeneity, (2) seasonal dynamics, (3) complex feedback effects, (4) global geographic scope, (5) deep uncertainty, (6) human connectivity, and (7) recognized ignorance on the relation between environmental factors and vector. More specifically, problems that can benefit from the use of case-specific multi-model architectures to represent heterogeneity and different levels of aggregation, Big Data and high-resolution raster data, GIS techniques to handle input and output data, and techniques to analyze deep uncertainty.

Appendix A: Details of Agent-based model

Contrasting synchronous and asynchronous agent scheduling

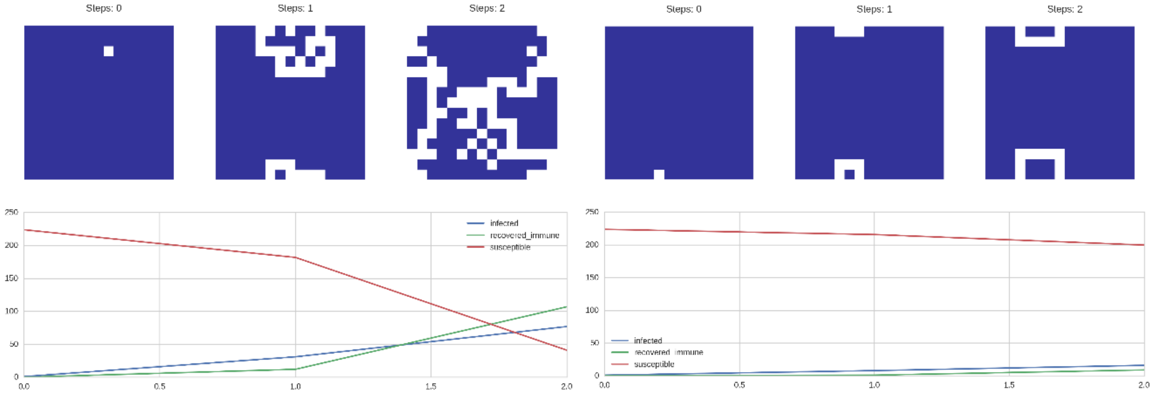
Both agent activation schemes emulate real-life processes which are naturally concurrent. Asynchronous agent activation is based on the hypothesis that random agent activation, where at each time step the agents are activated in a different order is sufficient to emulate reality. In contrast, agent updating, splits observation and action.

Synchronous agent scheduling scheme is implemented by recording each agent's decision one at a time, but not altering the state of the model. All agents are sequentially activated and assess their environment including the other agents' states. Upon this each individual decides their next activity. Thereafter, all agents are activated a second time to perform their prior decided action.

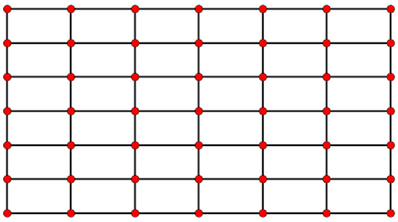
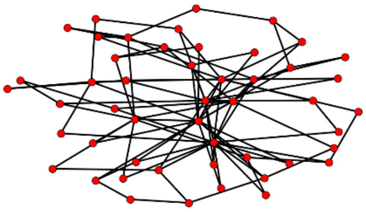
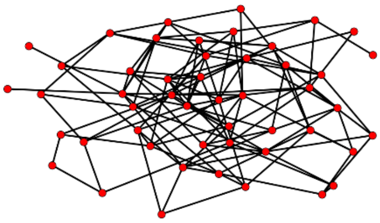
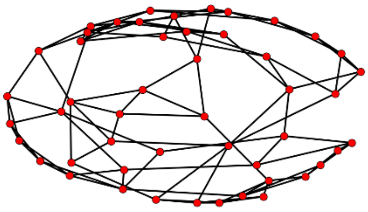


The figure below illustrates the different behavior modes produced when changing the agent activation scheme but keeping all other parameters the same. For simplicity, here, the underlying AB model is a simple SIR model of individuals which are located in a lattice network with non-empty cells. It is assumed that every susceptible neighbor of an infected agent becomes infected but recovers within one tick. The asynchronous agent activation displays a runaway disease spread, that is, because the neighbor becomes immediately infected and if not already being activated before within the current tick becomes activated and itself infects its neighbor cells.

Figure 43: Contrasting behavior modes and produced simulation results of asynchronous agent activation (left) and synchronous agent activation (right)



Contrasting Network Topologies for AB model

2D Grid of cells:	**Scale-free networks:**
Definition: Regular grid of cells is the most widely (and often only) used network in agent based modeling. This structure is also referred to as 2D cellular automaton and in mathematics classified as lattice. (Keeling and Eames 2005).	*Definition*: Graphs created by preferential attachment, where those nodes are more likely to be connected that have already more links (Barabasi 2002).
Characteristics: Highly clustered, Totally homogeneous, No variation in the number of links per node	*Characteristics*: Heterogeneous structure, Few nodes strongly linked, the others little
Algorithm:	*Algorithm:*
Standard grid structure that wraps around	Barabasi Albert algorithm
	
Random Networks:	**Small-world networks:**
Definition: Graphs in which connections are formed at random independently of the spatial proximity to other nodes or already existing edges (Keeling and Eames 2005).	*Definition:* Hybrid form of Graphs between the rigid arrangement of lattices and the unstructured connections of random networks , widely used to describe large social networks (Watts and Strogatz 1998).
Characteristics: Homogeneity, No apparent structure Little variation in the number of links per node	*Characteristics*: High clustering, Small average shortest path length Little variation in the number of links per node
Algorithm:	*Algorithm*:
erdos renyi algorithm	Watts strogatz algorithm
	

Appendix B: Switching from ABM to SD

Facilitated by the rapidly growing computational power in the last decades, individual based modeling has despite the computational burden become increasingly popular. Nevertheless, very large simulation models, with thousands, millions or even more individuals, require computing power that are and will remain extremely expensive in the short and medium term. Computational limitations are in particular relevant, due to the need for exhaustive sensitivity analysis and undergo rigorous validation processes. As a result, ten thousand and more simulation runs need to be conducted. Rahmandad & Sterman, 2008 conclude that a conflict in objectives arises between the richness of feedback effects captured endogenously, the scope of the simulation study (number of agents), and the elaborateness of performed sensitivity analysis.

Therefore, in this thesis alternative approaches were explored. Specifically, beginning with an AB model and switching to a compartmental model once the number of infectious individuals reaches a threshold.

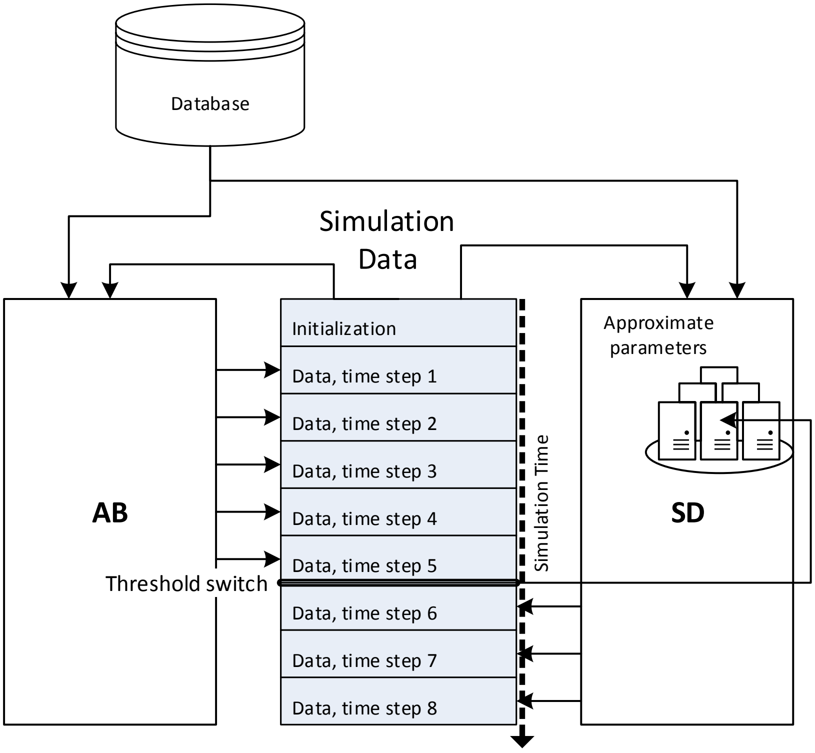
While working on this thesis, similar research by Bobashev et al. (2007) was discovered, which confirms the relevance of the approach. The threshold criteria for the model switch has to be chosen such that the patterns of disease dynamics have been established.

In the present study, simply, once, for the first time thirty persons in the population are infected the model switches to differential equation, however more rigorous exploration is needed to make decision rules for the selection of criteria or combination of criteria for the switch moment. By combining the advantages of the two methods the computational requirements can be dramatically reduced without conceding on detail, in the beginning of the outbreak, when it is of importance. Moreover, the different system representations allow to explicitly study different dynamics. The effects captured endogenously in the two models can different based upon, they are relevant at this stage of the outbreak. As a consequence, the impact of policies at the local and global level can be explored using one and the same model.

Strongly clustered networks lead to local saturation and dying out of infectious diseases. Arguably, however, in moderately clustered large populations the hypothesis of the mean field approximations and perfect mixing holds and thus despite heterogeneity at the individual level, an compartmental models produces very similar results as an individual based model (Rahmandad and Sterman 2008).

The concept for multi-method model is based on interfacing to a single data set with aggregated system values and initializing the SD based by prior AB simulation results. Figure 44 depicts the realization of the interface between AB and SD. It is based on four building blocks: (1) AB model, (2) an equivalent SD model, (3) simulation data (aggregated KPIs), (4) background database.

Figure 44: Conceptual model for threshold switch from ABM to SD[16]



1. Model is advanced only a single time step.
2. At the beginning of each time step the threshold condition is checked based on the last computed simulation values.
3. If few agents are infected, an AB model is initiated. If the threshold criterion is exceeded, then instead SD model is set up.
4. The models' external parameters are parameterized by the shared database
5. At the end of each step, (aggregated) system values are written to the shared data frame.

When the switch point is reached, the produced result of the individual based simulation model is used to parameterize the aggregated system representation. With recursive least squares parameter estimation the unknown variables of the compartmental model are approximated. In the course of the optimization, guesses of parameter combinations are simulated in the SD model and compared to the results of the individual based model until convergence. Two popular optimization algorithms levenberg-marquardt and limited-memory-BFGS (Byrd et al. 1995; Lourakis 2005) have been assessed for their suitability for this optimization task. Although no significant performance difference between the algorithm could be identified, the ability to handle bounds constraints and thus to exclude unrealistic values from the start make the limited-memory-BFGS generally advantageous.

---

[16] Source: author's illustration

In Table 2, the parameters used to verify the switching mechanism between both models are specified. Then, Figure 45 depicts the system dynamics model used and Figure 46 illustrates the fitting mechanism.

Table 2: Fitting AB to SD Optimization set up

Software used for the simulation	scipy.optimize (Byrd et al. 1995)
Payoff function	$f = \sum_{n=1}^{steps\_fitting} (\mathrm{Tseries\_SD_n} - \mathrm{Tseries\_AB_n})$
Parameter space	Contact infectivity
search algorithm	Limited-memory-BFGS Levenberg-marquardt
Average number of iterations needed for convergence	35

Figure 45: Simplified SIR model used to test the switching mechanism

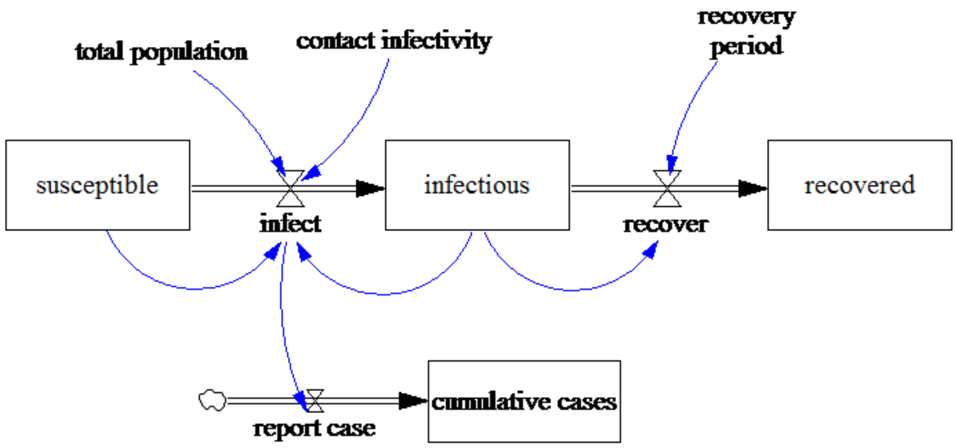
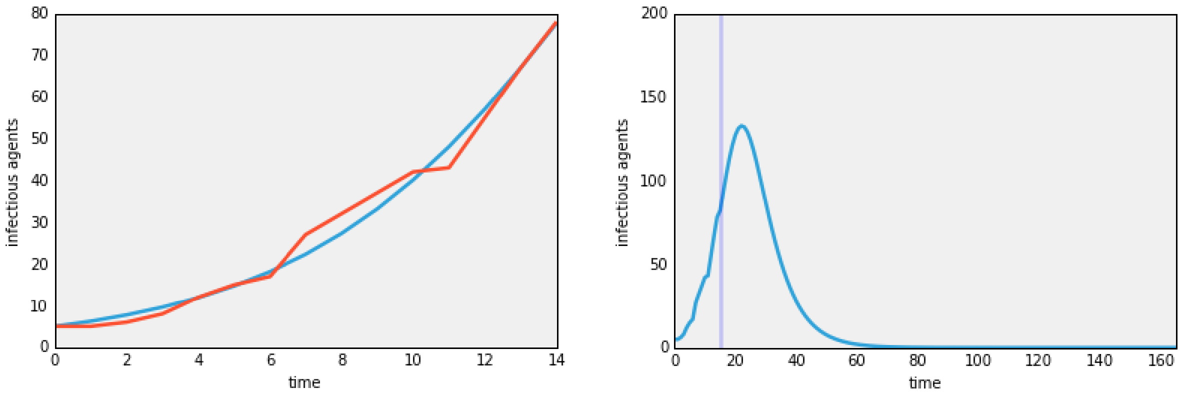


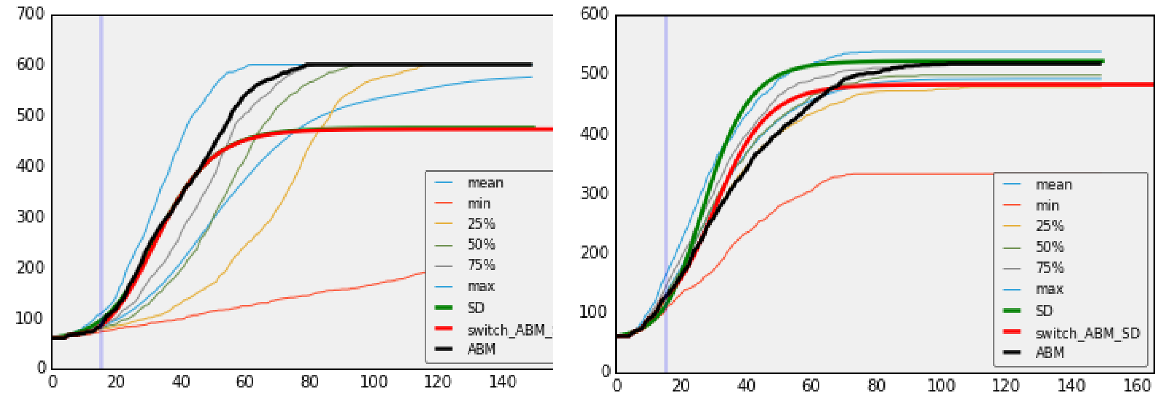
Figure 46: Illustration of fitting SD model to ABM



In the simulation period where the equation based model was fitted to the agent based model, very high R^2 coefficients (>0.95) were achieved which demonstrate an adequate goodness of fit. However, if comparing simulation results until the end of the simulation period, the alignment was poor. On average, the compartmental model infected a significantly larger share of the population than the agent base model.

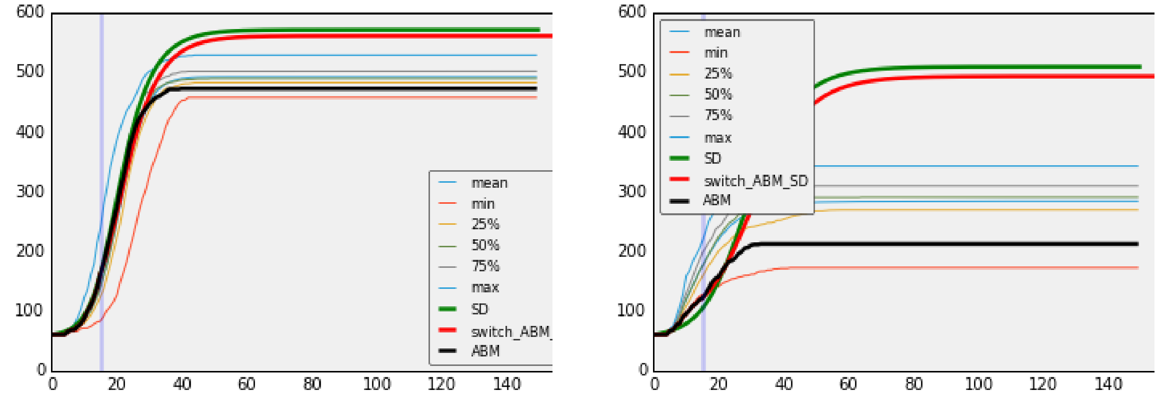
Figure 47 illustrates the outcome of the model when different network topologies are used. Homogeneous networks, random network and 2D Grid show relatively high model fit. In contrast, heterogeneous networks such as Scale-free and Small-world produce fundamentally different outcomes.

Figure 47: Comparison of model fit over different network topologies



Small-world network

2D Grid



Random network

Scale-free network

Appendix C: Verification and Validation of models

Verification

The presented behaviour modes generally correspond to the expected dynamics. In addition, the simulation model was exhaustively tested for consistency. In the model building process, complexity was incrementally added and iteratively verification performed. Model verification ensures that the model is coded correctly and the model represents the intended system. Since conceptually and operational different ABM and SD model were verified separately with adequate verification techniques.

Integration method (SD): Simulation results were found not sensitive to the use of different integration methods. In Vensim DSS ® the numerical methods Euler, RK2, RK4 and RK-Auto were tested. In addition, the results produced with Vensim DSS ® and PySD (Houghton, and Siegel 2015) were compared. Using PySD to simulate the SD model in Python predefined the choice of integration method. Underlying, PySD utilizes lsoda from the FORTRAN library odepack to solve systems of ordinary differential equations. This library uses a variable-step solver able to solve both stiff and non-stiff differential systems.

Time step (SD): To enable writing simulation values to a shared data frame, the simulation methods needs to have the same time step. Inherently, AB models have a fixed time step of 1, consequently the differential-equation model ought to report its results in the same interval Comparative analysis revealed no parameter sensitivity in the simulation outcomes when cutting the time step from 1 to 0.5, 0.125 and smaller.

Visualization of agent interaction (AB): A major strength of ABM is that agents' behaviour can be graphically displayed and doing so, single agent behaviour rigorously checked. The mesa library used to develop and simulate the AB model in Python supports visualizing basic agent behaviour. However, its capabilities to display agents is (yet) limited and therefore the final model with multiple agents, agent properties and agent actions could not be adequately visualized. Nevertheless, a visual verification of the AB was performed for a simplified earlier version of the model. In the appendix a notebook is presented which allows modellers and other users to visually explore the disease propagation overtime through the agent network by playing with interactive widgets.

Validation

Validation is conducted to check that the model is suitable to study a specific problem.

Extreme value test: To discover potential errors in the model structure extreme-conditions were tested. In addition, this validation step aimed at preparing the model for use in the exploratory modeling and analysis. Since broad uncertainty spaces are explored, it is a necessary precondition that even under extreme conditions no mathematical errors in equations occur.

Sensitivity Analysis: The objective of conducting the sensitivity analysis was to determine the parameters that have a major influence on the generated model behavior. A basic sensitivity analysis was applied to explore the robustness of the model to large uncertainty spaces.

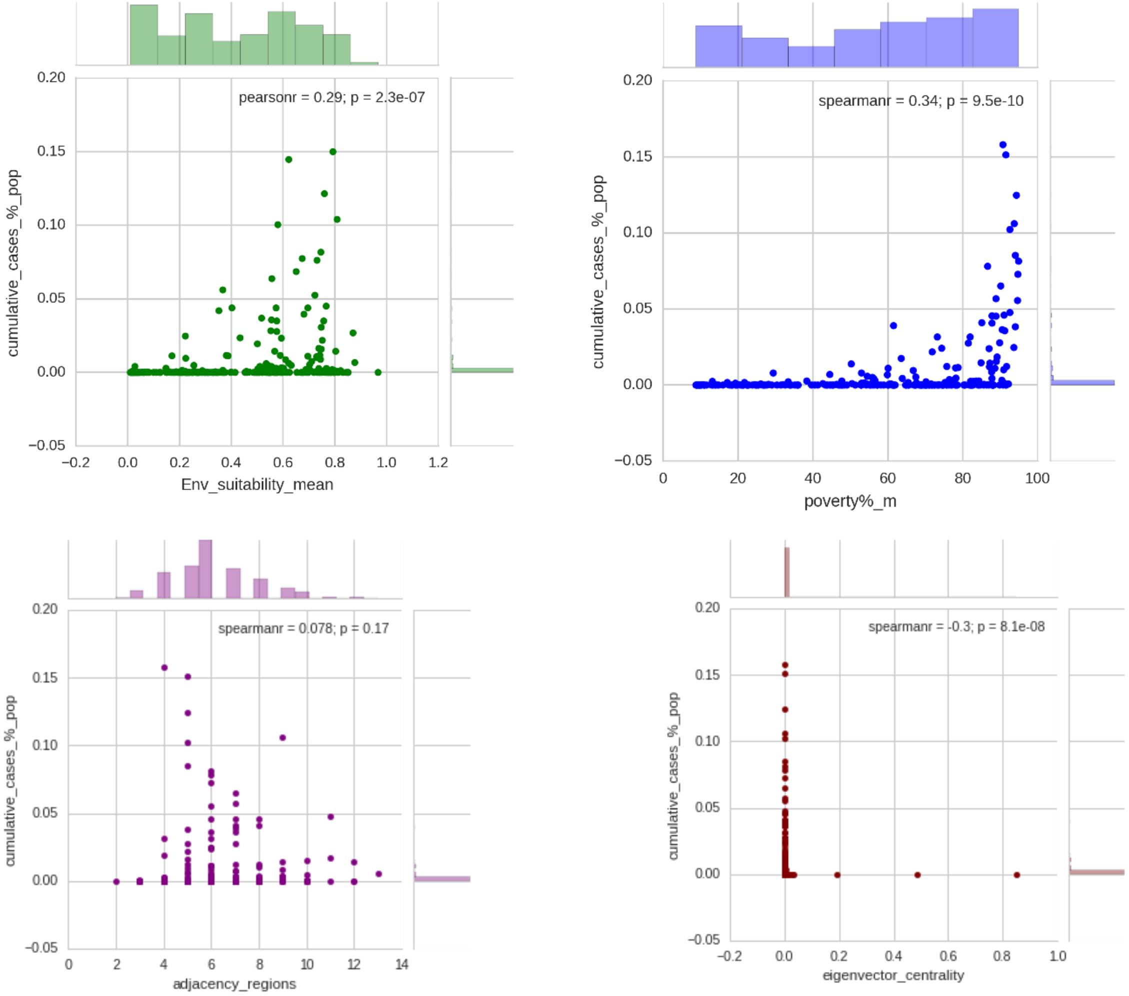
Model replication: In order to switch from AB to SD, two models developed in different modeling methods were constructed. Van Dam et al. (2013) describe model replication with different system representation as powerful tool for model validation. Nevertheless, in this study from the beginning the AB model was built with the aim to produce similar behaviour modes as the SD model and vice versa, which naturally introduces potential bias. Nevertheless, the two models are to an extent validated as they produce similar behaviour.

Appendix D: Further Analysis of Simulation results

One of the hypothesis considered was that the heterogeneous local conditions drive regional disease transmission and consequently largely determine whether and in which regions (large) outbreak happen. Therefore, this section elaborates on the relation between system performance and several local conditions.

A significant but weak linear correlation between environmental suitability and cases per population was found (Pearson-correlation = .29, p = 2.3e-07). Similarly, a weak exponentially growing correlation between the share of the poor people and the cases per population was identified (Spearman-correlation = .34, p =9.5e-10[17]). For the number of adjacent regions and the region's eigenvector centrality in flight graph no relation could be established with regard to the number of local Zika cases.

Figure 48: Visual representation (Scatterplot) of correlation between regional disaggregated average cumulative cases as % of population and local conditions: Environmental suitability (upper-left), Poverty as % of population (upper-right), Number of adjacency regions (lower-left), and eigenvector centrality in flight network (lower-right).



---

[17] In contrast to Pearson, Spearman can also be applied to non-linear monotonic relationships

The later results are unexpected as it was suspected that the virus would spread to regions that are more connected more easily and hence cause in this regions in many scenarios large outbreaks. One of the reason explaining this observation, is that regardless of how central a region’s network position is, actually only its position and links to regions with ongoing outbreaks are the relevant estimating the health threat for the population in the first wave of an emerging infectious disease.

Appendix E: Parameter intervals for Uncertainty Analysis

Table 3: Parameter intervals for Uncertainty Analysis

Parameter	Units	Reference	Interval	Sources/remarks
Intrinsic incubation period	[day]	5	(3,10)	(CDC 2016c)
Human recovery period	[day]	5	(2,7)	(Duffy et al. 2009)
Probability of infection	[Dmnl]	0.726	(.6, 1)	(Gilotra and Shah 1967), transmission of Chikungunya virus by mosquitoes
Share vertical transmission	[Dmnl]	0..1	(0, .15)	(Rajapakse, Rodrigo, and Rajapakse 2010)
Extrinsic incubation period	[day]	6	(5, 8)	(Hayes 2009)
Global share herd immunity	[Dmnl]	0.1	(0, 0.2)	Estimate
Switch_flights	[Dmnl]	1	[0, 1]	Validation
Switch agent activation	[Dmnl]	1	[0, 1]	Structural Uncertainty
Switch network type	[Dmnl]	1	[0, 1, 2, 3]	Structural Uncertainty
Intensity of travel	[Dmnl]	1	[0, 1, 2, 3]	Structural Uncertainty
Delaytime reporting	[day]	10	(7, 30)	Estimate
Share reported cases	[Dmnl]	0.2	(0.05, .8)	Estimate
Destruction time ponds	[day]		(7, 30)	Estimate
Network connectivity	[Links per node]	4	(2.5, 8)	Estimate
Share of asymptomatic	[Dmnl]	0.8	(0.7, 0.9)	(Duffy et al. 2009)
Maximal mosquito blood meals	[1/day]	4	(2, 5)	Estimate
Aedes aegypti survival rate to adult stage	[]	0.03	(0.03, 0.92)	(Rueda et al. 1990), dependent on humidity and temperature
Personal protection poor	[Dmnl]	0.2	(0.01, .5)	Estimate
Personal protection rich	[Dmnl]	0.2	(0.8, 1)	Estimate
Breeding site recreation poor	[Dmnl]	0.2	(0.01, .0.02)	Estimate
Breeding site recreation rich	[Dmnl]	0.2	(0.005 , 0.15)	Estimate

Bibliography

Ahmad, Sajjad, and Slobodan P. Simonovic. 2004. “Spatial System Dynamics: New Approach for Simulation of Water Resources Systems.” *Journal of Computing in Civil Engineering* 18 (4): 331–40. doi:10.1061/(ASCE)0887-3801(2004)18:4(331).

Aliota, Matthew T., Stephen A. Peinado, Ivan Dario Velez, and Jorge E. Osorio. 2016. “The wMel Strain of Wolbachia Reduces Transmission of Zika Virus by Aedes Aegypti.” *Scientific Reports* 6 (July): 28792. doi:10.1038/srep28792.

ArcGIS. 2016. *Spatial-Analyst-Toolbox: How-Zonal-Statistics Work*. http://desktop.arcgis.com/en/arcmap/10.3/tools/spatial-analyst-toolbox/h-how-zonal-statistics-works.htm.

Auping, Willem L., Erik Pruyt, and Jan H. Kwakkel. 2014. “Dealing with Multiple Models in System Dynamics:: Perspectives on the Future of Copper.” *International Journal of System Dynamics Applications* 3 (4): 17–35. doi:10.4018/ijsda.2014100102.

Axtell, Robert, Robert Axelrod, Joshua M. Epstein, and Michael D. Cohen. 1996. “Aligning Simulation Models: A Case Study and Results.” *Computational & Mathematical Organization Theory* 1 (2): 123–41. doi:10.1007/BF01299065.

Barabasi, Albert-laszlo. 2002. *Linked: How Everything Is Connected to Everything Else and What It Means for Business, Science, and Everyday Life*. Cambridge, MA, USA: Perseus.

Bhatt, Samir, Peter W. Gething, Oliver J. Brady, Jane P. Messina, Andrew W. Farlow, Catherine L. Moyes, John M. Drake, et al. 2013. “The Global Distribution and Burden of Dengue.” *Nature* 496 (7446): 504–7. doi:10.1038/nature12060.

Borshchev, A, and A Filippov. 2004. “From System Dynamics and Discrete Event to Practical Agent Based Modeling: Reasons, Techniques, Tools. Presented at the The.” In *July 25 - 29, 2004*. Oxford, England.

Brady, Oliver J., Michael A. Johansson, Carlos A. Guerra, Samir Bhatt, Nick Golding, David M. Pigott, Hélène Delatte, et al. 2013. “Modelling Adult Aedes Aegypti and Aedes Albopictus Survival at Different Temperatures in Laboratory and Field Settings.” *Parasites & Vectors* 6 (1): 351. doi:10.1186/1756-3305-6-351.

Brazilian Ministry of Health. 2016. “Ministério da Saúde investiga 3.448 casos suspeitos de microcefalia.” *Portal da Saúde – Ministério da Saúde – www.saude.gov.br*. January 27. http://portalsaude.saude.gov.br/index.php/cidadao/principal/agencia-saude/21890-ministerio-da-saude-investiga-3-448-casos-suspeitos-de-microcefalia.

Brinkhoff, Thomas, Hans-Peter Kriegel, and Bernhard Seeger. 1993. “Efficient Processing of Spatial Joins Using R-Trees.” In *Proceedings of the 1993 ACM SIGMOD International Conference on Management of Data*, 237–246. SIGMOD ’93. New York, NY, USA: ACM. doi:10.1145/170035.170075.

Brown, Julia E., Carolyn S. McBride, Petrina Johnson, Scott Ritchie, Christophe Paupy, Hervé Bossin, Joel Lutomiah, et al. 2011. “Worldwide Patterns of Genetic Differentiation Imply Multiple ‘Domestications’ of Aedes Aegypti, a Major Vector of Human Diseases.” *Proceedings. Biological Sciences / The Royal Society* 278 (1717): 2446–54. doi:10.1098/rspb.2010.2469.

Byrd, R., P. Lu, J. Nocedal, and C. Zhu. 1995. “A Limited Memory Algorithm for Bound Constrained Optimization.” *SIAM Journal on Scientific Computing* 16 (5): 1190–1208. doi:10.1137/0916069.

Caron-Lormier, Geoffrey, Roger W. Humphry, David A. Bohan, Cathy Hawes, and Pernille Thorbek. 2008. “Asynchronous and Synchronous Updating in Individual-Based Models.” *Ecological Modelling* 212 (3–4): 522–27. doi:10.1016/j.ecolmodel.2007.10.049.

CDC. 2016. “West Nile Virus: Symptoms & Treatment.” *West Nile Virus*. January 27. http://www.cdc.gov/westnile/symptoms/.

CIESIN. 2015. “Gridded Population of the World, Version 4 (GPWv4): Population Count Adjusted to Match 2015 Revision of UN WPP Country Totals.” Center for International Earth Science Information Network - CIESIN - Columbia University. NASA Socioeconomic Data and Applications Center (SEDAC). http://dx.doi.org/10.7927/H4PR7SX1.

Cockrell, Robert Chase, Scott Christley, Eugene Chang, and Gary An. 2015. “Towards Anatomic Scale Agent-Based Modeling with a Massively Parallel Spatially Explicit General-Purpose Model of Enteric Tissue (SEGMEnT_HPC).” *PLoS ONE* 10 (3). doi:10.1371/journal.pone.0122192.

Douglas, Bruce. 2015. “Brazil Warns Women Not to Get Pregnant as Zika Virus Is Linked to Rare Birth Defect.” *The Guardian*, December 4, sec. Global development. http://www.theguardian.com/global-development/2015/dec/04/brazil-zika-virus-pregnancy-microcephaly-mosquito-rare-birth-defect.

Duffy, Mark R., Tai-Ho Chen, W. Thane Hancock, Ann M. Powers, Jacob L. Kool, Robert S. Lanciotti, Moses Pretrick, et al. 2009. “Zika Virus Outbreak on Yap Island, Federated States of Micronesia.” *New England Journal of Medicine* 360 (24): 2536–43. doi:10.1056/NEJMoa0805715.

ECDC. 2016. “Rapid Risk Assessment: Zika Virus Disease Epidemic: Potential Association with Microcephaly and Guillain-Barré Syndrome (First Update).” European Center for Disease Prevention and Control: ECDC. http://ecdc.europa.eu/en/publications/Publications/rapid-risk-assessment-zika-virus-first-update-jan-2016.pdf.

Eisenstein, Michael. 2016. “Disease: Poverty and Pathogens.” *Nature* 531 (7594): S61–63. doi:10.1038/531S61a.

Elvidge, Christopher D., Paul C. Sutton, Tilottama Ghosh, Benjamin T. Tuttle, Kimberly E. Baugh, Budhendra Bhaduri, and Edward Bright. 2009. “A Global Poverty Map Derived from Satellite Data.” *Computers & Geosciences* 35 (8): 1652–60. doi:10.1016/j.cageo.2009.01.009.

Epstein, Joshua M. 2006. *Generative Social Science: Studies in Agent-Based Computational Modeling*. STU-Student edition. Princeton University Press. http://www.jstor.org/stable/j.ctt7rxj1.

Fauci, Anthony S., and David M. Morens. 2016. “Zika Virus in the Americas — Yet Another Arbovirus Threat.” *New England Journal of Medicine* 0 (0): null. doi:10.1056/NEJMp1600297.

Fauci, Anthony S., Nancy A. Touchette, and Gregory K. Folkers. 2005. “Emerging Infectious Diseases: A 10-Year Perspective from the National Institute of Allergy and Infectious Diseases.” *Emerging Infectious Diseases* 11 (4): 519–25. doi:10.3201/eid1104.041167.

Friedman, Jerome H., and Nicholas I. Fisher. 1999. “Bump Hunting in High-Dimensional Data.” *Statistics and Computing* 9 (2): 123–43. doi:10.1023/A:1008894516817.

Gardner, Lauren M., David Fajardo, S. Travis Waller, Ophelia Wang, Sahotra Sarkar, Lauren M. Gardner, David Fajardo, S. Travis Waller, Ophelia Wang, and Sahotra Sarkar. 2012. “A Predictive Spatial Model to Quantify the Risk of Air-Travel-Associated Dengue Importation into the United States and Europe, A Predictive Spatial Model to Quantify the Risk of Air-Travel-Associated Dengue Importation into the United States and Europe.” *Journal of Tropical Medicine, Journal of Tropical Medicine* 2012, 2012 (March): e103679. doi:10.1155/2012/103679, 10.1155/2012/103679.

Gilotra, S. K., and K. V. Shah. 1967. “Laboratory Studies on Transmission of Chikungunya Virus by Mosquitoes.” *American Journal of Epidemiology* 86 (2): 379–85.

Grassly, Nicholas C., and Christophe Fraser. 2008. “Mathematical Models of Infectious Disease Transmission.” *Nature Reviews Microbiology* 6 (6): 477–87. doi:10.1038/nrmicro1845.

Grosche, Tobias, Franz Rothlauf, and Armin Heinzl. 2007. "Gravity Models for Airline Passenger Volume Estimation." *Journal of Air Transport Management* 13 (4): 175–83. doi:10.1016/j.jairtraman.2007.02.001.

Guimerà, R., S. Mossa, A. Turtschi, and L. a. N. Amaral. 2005. "The Worldwide Air Transportation Network: Anomalous Centrality, Community Structure, and Cities' Global Roles." *Proceedings of the National Academy of Sciences* 102 (22): 7794–99. doi:10.1073/pnas.0407994102.

Hales, Simon, Neil de Wet, John Maindonald, and Alistair Woodward. 2002. "Potential Effect of Population and Climate Changes on Global Distribution of Dengue Fever: An Empirical Model." *Lancet (London, England)* 360 (9336): 830–34. doi:10.1016/S0140-6736(02)09964-6.

Hardy, James L., Edward J. Houk, Laura D. Kramer, and William C. Reeves. 1983. "Intrinsic Factors Affecting Vector Competence of Mosquitoes for Arboviruses." *Annual Review of Entomology* 28 (1): 229–262.

Harrington, Laura C., Thomas W. Scott, Kriangkrai Lerdthusnee, Russell C. Coleman, Adriana Costero, Gary G. Clark, James J. Jones, et al. 2005. "Dispersal of the Dengue Vector Aedes Aegypti Within and Between Rural Communities." *The American Journal of Tropical Medicine and Hygiene* 72 (2): 209–20.

Hay, S. I., A. J. Tatem, A. J. Graham, S. J. Goetz, and D. J. Rogers. 2006. "Global Environmental Data for Mapping Infectious Disease Distribution." *Advances in Parasitology* 62: 37–77. doi:10.1016/S0065-308X(05)62002-7.

Hayes, Edward B. 2009. "Zika Virus Outside Africa." *Emerging Infectious Diseases* 15 (9): 1347–50. doi:10.3201/eid1509.090442.

Hotez, Peter J. 2008. "Neglected Infections of Poverty in the United States of America." *PLOS Negl Trop Dis* 2 (6): e256. doi:10.1371/journal.pntd.0000256.

IATA. 2014. "New IATA Passenger Forecast Reveals Fast-Growing Markets of the Future." October 16. http://www.iata.org/pressroom/pr/Pages/2014-10-16-01.aspx.

Jacintho, Luís F. O., André F. M. Batista, Terry L. Ruas, Maria G. B. Marietto, and Fábio A. Silva. 2010. "An Agent-Based Model for the Spread of the Dengue Fever: A Swarm Platform Simulation Approach." In *Proceedings of the 2010 Spring Simulation Multiconference*, 2:1–2:8. SpringSim '10. San Diego, CA, USA: Society for Computer Simulation International. doi:10.1145/1878537.1878540.

Jennings, Katia Sycara, and Michael Wooldridge. 1998. "A Roadmap of Agent Research and Development." *Autonomous Agents and Multi-Agent Systems* 1 (1): 7–38. doi:10.1023/A:1010090405266.

Johansson, Michael A., Francesca Dominici, and Gregory E. Glass. 2009. "Local and Global Effects of Climate on Dengue Transmission in Puerto Rico." *PLOS Negl Trop Dis* 3 (2): e382. doi:10.1371/journal.pntd.0000382.

Keeling, Matt J, and Ken T.D Eames. 2005. "Networks and Epidemic Models." *Journal of the Royal Society Interface* 2 (4): 295–307. doi:10.1098/rsif.2005.0051.

Knudsen, A. B., and R. Slooff. 1992. "Vector-Borne Disease Problems in Rapid Urbanization: New Approaches to Vector Control." *Bulletin of the World Health Organization* 70 (1): 1–6.

Kucharski, Adam J., Sebastian Funk, Rosalind M. Eggo, Henri-Pierre Mallet, W. John Edmunds, and Eric J. Nilles. 2016. "Transmission Dynamics of Zika Virus in Island Populations: A Modelling Analysis of the 2013–14 French Polynesia Outbreak." *PLOS Negl Trop Dis* 10 (5): e0004726. doi:10.1371/journal.pntd.0004726.

Kwakkel, Jan H., and Erik Pruyt. 2013. "Exploratory Modeling and Analysis, an Approach for Model-Based Foresight under Deep Uncertainty." *Technological Forecasting and Social Change* 80 (3): 419–31. doi:10.1016/j.techfore.2012.10.005.

Lee, D. T., and B. J. Schachter. 1980. “Two Algorithms for Constructing a Delaunay Triangulation.” *International Journal of Computer & Information Sciences* 9 (3): 219–42. doi:10.1007/BF00977785.

Lempert, Robert. 2008. “A Toolbox for Policy Planners.” *Nature* 455 (7210): 169–70. doi:10.1038/455169a.

Linthicum, K. J., A. Anyamba, C. J. Tucker, P. W. Kelley, M. F. Myers, and C. J. Peters. 1999. “Climate and Satellite Indicators to Forecast Rift Valley Fever Epidemics in Kenya.” *Science (New York, N.Y.)* 285 (5426): 397–400.

Louis, Catherine Saint. 2016. “Microcephaly, Spotlighted by Zika Virus, Has Long Afflicted and Mystified.” *The New York Times*, January 31. http://www.nytimes.com/2016/02/01/health/microcephaly-spotlighted-by-zika-virus-has-long-afflicted-and-mystified.html.

Lourakis, Manolis IA. 2005. “A Brief Description of the Levenberg-Marquardt Algorithm Implemented by Levmar.” *Foundation of Research and Technology* 4: 1–6.

Lucey, DR, and LO Gostin. 2016. “The Emerging Zika Pandemic: Enhancing Preparedness.” *JAMA*, January. doi:10.1001/jama.2016.0904.

Manore, Carrie A., Kyle S. Hickmann, James M. Hyman, Ivo M. Foppa, Justin K. Davis, Dawn M. Wesson, and Christopher N. Mores. 2015. “A Network-Patch Methodology for Adapting Agent-Based Models for Directly Transmitted Disease to Mosquito-Borne Disease.” *Journal of Biological Dynamics* 9 (1): 52–72. doi:10.1080/17513758.2015.1005698.

Mao, Liang, Xiao Wu, Zhuojie Huang, and Andrew J. Tatem. 2015. “Modeling Monthly Flows of Global Air Travel Passengers: An Open-Access Data Resource.” *Journal of Transport Geography* 48 (October): 52–60. doi:10.1016/j.jtrangeo.2015.08.017.

Marcondes, Carlos Brisola, and Maria de Fátima Freire de Melo Ximenes. 2015. “Zika Virus in Brazil and the Danger of Infestation by Aedes (Stegomyia) Mosquitoes.” *Revista Da Sociedade Brasileira De Medicina Tropical*, December. doi:10.1590/0037-8682-0220-2015.

Medeiros, Líliam César de Castro, César Augusto Rodrigues Castilho, Cynthia Braga, Wayner Vieira de Souza, Leda Regis, and Antonio Miguel Vieira Monteiro. 2011. “Modeling the Dynamic Transmission of Dengue Fever: Investigating Disease Persistence.” *PLOS Negl Trop Dis* 5 (1): e942. doi:10.1371/journal.pntd.0000942.

Messina, Jane P, Moritz UG Kraemer, Oliver J Brady, David M Pigott, Freya M Shearer, Daniel J Weiss, Nick Golding, et al. 2016. “Mapping Global Environmental Suitability for Zika Virus.” Edited by Mark Jit. *eLife* 5 (April): e15272. doi:10.7554/eLife.15272.

Milling, Peter, and Nadine Schieritz. 2003. “Modeling the Forest or Modeling the Trees - A Comparison of System Dynamics and Agent-Based Simulation.” In *Proceedings of the 21st International Conference of the System Dynamics Society : July 20 - 24, 2003, New York City, USA*, 1–15. Albany, NY: System Dynamics Soc. http://www.systemdynamics.org/conferences/2003/proceed/PAPERS/140.pdf.

Moreira, Luciano A., Iñaki Iturbe-Ormaetxe, Jason A. Jeffery, Guangjin Lu, Alyssa T. Pyke, Lauren M. Hedges, Bruno C. Rocha, et al. 2009. “A Wolbachia Symbiont in Aedes Aegypti Limits Infection with Dengue, Chikungunya, and Plasmodium.” *Cell* 139 (7): 1268–78. doi:10.1016/j.cell.2009.11.042.

“Multimethod Modeling — AnyLogic Simulation Software.” 2016. Accessed August 4. http://www.anylogic.com/multimethod-modeling.

Natural Earth Data. 2016. “Admin 0 - Countries, Vers. 2.0.0.” http://www.naturalearthdata.com/downloads/50m-cultural-vectors/50m-admin-0-countries-2/.

Neuwirth, Christian, and A Peck. 2013. “A Conceptual Spatial System Dynamics (SSD) Model for Structural Changes in Grassland Farming.” In . Adelaide, Australia.

NOAA. 2009. “A Global Poverty Map Derived From Satellite Data.” *NOAA, Earth Observation Group*. http://ngdc.noaa.gov/eog/dmsp/download_poverty.html.

North, Michael J., Nicholson T. Collier, Jonathan Ozik, Eric R. Tatara, Charles M. Macal, Mark Bragen, and Pam Sydelko. 2013. “Complex Adaptive Systems Modeling with Repast Simphony.” *Complex Adaptive Systems Modeling* 1 (1): 3. doi:10.1186/2194-3206-1-3.

Osgood, Nathaniel. 2007. “Using Traditional and Agent Based Toolsets for System Dynamics: Present Tradeoffs and Future Evolution.” In . http://www.systemdynamics.org/conferences/2007/proceed/papers/OSGOO488.pdf.

PAHO/WHO. 2015. “Epidemiological Alert: Neurological Syndrome, Congenital Anomalies, and Zika Virus Infection. Implications for Public Health in the Americas.” Pan American Health Organization / World Health Organization: PAHO/WHO. http://www.paho.org/hq/index.php?option=com_docman&task=doc_view&Itemid=270&gid=32405&lang=en.

———. 2016. “Epidemiological Update: Neurological Syndrome, Congenital Anomalies, and Zika Virus Infection.” Pan American Health Organization / World Health Organization: PAHO/WHO. http://www.paho.org/hq/index.php?option=com_docman&task=doc_download&Itemid=&gid=32879&lang= en.

Parker, Jon, and Joshua M. Epstein. 2011. “A Distributed Platform for Global-Scale Agent-Based Models of Disease Transmission.” *ACM Transactions on Modeling and Computer Simulation : A Publication of the Association for Computing Machinery* 22 (1): 2. doi:10.1145/2043635.2043637.

Pedamallu, Chandra Sekhar, Linet Ozdamar, Erik Kropat, and Gerhard-Wilhelm Weber. 2010. “A System Dynamics Model for Intentional Transmission of HIV/AIDS Using Cross Impact Analysis.” *Central European Journal of Operations Research* 20 (2): 319–36. doi:10.1007/s10100-010-0183-2.

Pérez, F., B.E. Granger, and J.D. Hunter. 2011. “Python: An Ecosystem for Scientific Computing.” *Computing in Science Engineering* 13 (2): 13–21. doi:10.1109/MCSE.2010.119.

Perkins, Alex, Amir Siraj, Corrine Warren Ruktanonchai, Moritz Kraemer, and Andrew Tatem. 2016. “Model-Based Projections of Zika Virus Infections in Childbearing Women in the Americas.” *bioRxiv*, February, 39610. doi:10.1101/039610.

Phillips, Melissa Lee. 2008. “Dengue Reborn: Widespread Resurgence of a Resilient Vector.” *Environmental Health Perspectives* 116 (9): A382–88.

Pruyt, Erik. 2015. “From Building a Model to Adaptive Robust Decision Making Using Systems Modeling.” In *Policy Practice and Digital Science*, edited by Marijn Janssen, Maria A. Wimmer, and Ameneh Deljoo, 75–93. Public Administration and Information Technology 10. Springer International Publishing. http://link.springer.com/chapter/10.1007/978-3-319-12784-2_5.

Pruyt, Erik, Willem L. Auping, and Jan H. Kwakkel. 2015. “Ebola in West Africa: Model-Based Exploration of Social Psychological Effects and Interventions.” *Systems Research and Behavioral Science* 32 (1): 2–14. doi:10.1002/sres.2329.

Rahmandad, Hazhir. 2004. “Heterogeneity and Network Structure in the Dynamics of Contagion: Comparing Agent-Based and Differential Equation Models.” In . Oxford, England: System Dynamics Society.

Rahmandad, Hazhir, and John Sterman. 2008. “Heterogeneity and Network Structure in the Dynamics of Diffusion: Comparing Agent-Based and Differential Equation Models.” *Management Science* 54 (5): 998–1014. doi:10.1287/mnsc.1070.0787.

Rahmandad, Hazhir, and John D. Sterman. 2012. “Reporting Guidelines for Simulation-Based Research in Social Sciences.” *System Dynamics Review* 28 (4): 396–411. doi:10.1002/sdr.1481.

Rajapakse, Senaka, Chathuraka Rodrigo, and Anoja Rajapakse. 2010. “Atypical Manifestations of Chikungunya Infection.” *Transactions of the Royal Society of Tropical Medicine and Hygiene* 104 (2): 89–96. doi:10.1016/j.trstmh.2009.07.031.

Richmond, Barry. 1997. “The ‘thinking’ in Systems Thinking: How Can We Make It Easier to Master.” *The Systems Thinker* 8 (2): 1–5.

Riley, Steven. 2007. “Large-Scale Spatial-Transmission Models of Infectious Disease.” *Science* 316 (5829): 1298–1301. doi:10.1126/science.1134695.

Ritchie-Dunham, James L., and Jorge F. Méndez Galván. 1999. “Evaluating Epidemic Intervention Policies with Systems Thinking: A Case Study of Dengue Fever in Mexico.” *System Dynamics Review* 15 (2): 119–38. doi:10.1002/(SICI)1099-1727(199922)15:2<119::AID-SDR163>3.0.CO;2-G.

Roberts, Carole, and Brian Dangerfield. 1990. “Modelling the Epidemiological Consequences of HIV Infection and AIDS: A Contribution from Operational Research.” *The Journal of the Operational Research Society* 41 (4): 273–89. doi:10.2307/2583798.

Romero, Simon, and Rebecca R. Ruiz. 2016. “Researchers Weigh Risks of Zika Spreading at Rio Olympics.” *The New York Times*, January 28. http://www.nytimes.com/2016/01/29/world/americas/brazil-zika-rio-olympics.html.

Rowley, Wayne A., and Charles L. Graham. 1968. “The Effect of Temperature and Relative Humidity on the Flight Performance of Female Aedes Aegypti.” *Journal of Insect Physiology* 14 (9): 1251–57. doi:10.1016/0022-1910(68)90018-8.

Rueda, L. M., K. J. Patel, R. C. Axtell, and R. E. Stinner. 1990. “Temperature-Dependent Development and Survival Rates of Culex Quinquefasciatus and Aedes Aegypti (Diptera: Culicidae).” *Journal of Medical Entomology* 27 (5): 892–98.

Schuler-Faccini, Lavinia, Erlane M. Ribeiro, Ian M.L. Feitosa, Dafne D.G. Horovitz, Denise P. Cavalcanti, André Pessoa, Maria Juliana R. Doriqui, et al. 2016. “Possible Association Between Zika Virus Infection and Microcephaly — Brazil, 2015.” *MMWR. Morbidity and Mortality Weekly Report* 65 (3): 1–4. doi:10.15585/mmwr.mm6503e2er.

Semenza, Jan C., Bertrand Sudre, Tolu Oni, Jonathan E. Suk, and Johan Giesecke. 2013. “Linking Environmental Drivers to Infectious Diseases: The European Environment and Epidemiology Network.” *PLOS Negl Trop Dis* 7 (7): e2323. doi:10.1371/journal.pntd.0002323.

Senge, Peter M. 1990. *The Fifth Discipline: The Art and Practice of the Learning Organization*. Doubleday/Currency.

Siettos, Constantinos I., and Lucia Russo. 2013. “Mathematical Modeling of Infectious Disease Dynamics.” *Virulence* 4 (4): 295–306. doi:10.4161/viru.24041.

Snow, A. A., D. A. Andow, P. Gepts, E. M. Hallerman, A. Power, J. M. Tiedje, and L. L. Wolfenbarger. 2005. “Genetically Engineered Organisms and the Environment: Current Status and Recommendations.” *Ecological Applications* 15 (2): 377–404. doi:10.1890/04-0539.

Sterman, John D. 2000. *Business Dynamics. Systems Thinking and Modeling for a Complex World*. Har/Cdr. Boston: Mcgraw-Hill Education Ltd.

Subbaraman, Nidhi. 2011. “Science Snipes at Oxitec Transgenic-Mosquito Trial.” *Nature Biotechnology* 29 (1): 9–11. doi:10.1038/nbt0111-9a.

Swinerd, Chris, and Ken R. McNaught. 2012. “Design Classes for Hybrid Simulations Involving Agent-Based and System Dynamics Models.” *Simulation Modelling Practice and Theory* 25 (June): 118–33. doi:10.1016/j.simpat.2011.09.002.

The Guardian. 2016. “Rare Brain Defect in Babies in Brazil Rises after Sudden Outbreak of Zika Virus.” *The Guardian*, January 20, sec. World news. http://www.theguardian.com/world/2016/jan/20/microcephaly-babies-brain-defect-brazil-zika-virus.

The World Bank. 2015. “World Development Indicators.” *Mortality Rate, Adult, Female, Male [Data File]*. http://data.worldbank.org/indicator/SP.DYN.AMRT.FE.

———. 2016. “Birth Rate, Crude (per 1,000 People).” The World Bank, Washington, D.C. http://data.worldbank.org/indicator/SP.DYN.CBRT.IN.

Thomé, Roberto C. A., Hyun Mo Yang, and Lourdes Esteva. 2010. “Optimal Control of Aedes Aegypti Mosquitoes by the Sterile Insect Technique and Insecticide.” *Mathematical Biosciences* 223 (1): 12–23. doi:10.1016/j.mbs.2009.08.009.

Trpis, M., and W. Hausermann. 1986. “Dispersal and Other Population Parameters of Aedes Aegypti in an African Village and Their Possible Significance in Epidemiology of Vector-Borne Diseases.” *The American Journal of Tropical Medicine and Hygiene* 35 (6): 1263–79.

Vincenot, Christian Ernest, Francesco Giannino, Max Rietkerk, Kazuyuki Moriya, and Stefano Mazzoleni. 2011. “Theoretical Considerations on the Combined Use of System Dynamics and Individual-Based Modeling in Ecology.” *Ecological Modelling* 222 (1): 210–18. doi:10.1016/j.ecolmodel.2010.09.029.

Walker, Warren E., Marjolijn Haasnoot, and Jan H. Kwakkel. 2013. “Adapt or Perish: A Review of Planning Approaches for Adaptation under Deep Uncertainty.” *Sustainability* 5 (3): 955–79. doi:10.3390/su5030955.

Watts, Duncan J., and Steven H. Strogatz. 1998. “Collective Dynamics of ‘small-World’ Networks.” *Nature* 393 (6684): 440–42. doi:10.1038/30918.

WHO. 2016a. “Dengue and Severe Dengue.” *Media Centre, Fact Sheet*. April. http://www.who.int/mediacentre/factsheets/fs117/en/.

———. 2016b. “Malaria.” *Media Centre, Fact Sheet*. April. http://www.who.int/mediacentre/factsheets/fs117/en/.

Wilensky, Uri, and William Rand. 2015. *An Introduction to Agent-Based Modeling: Modeling Natural, Social, and Engineered Complex Systems with NetLogo*. MIT Press.

Wilson, Mark L. 2002. “Emerging and Vector-Borne Diseases: Role of High Spatial Resolution and Hyperspectral Images in Analyses and Forecasts.” *Journal of Geographical Systems* 4 (1): 31–42. doi:10.1007/s101090100072.

World Bank. 2015. “The World Bank Annual Report 2015.” Washington, DC: © World Bank. License: CC BY 3.0 IGO. https://openknowledge.worldbank.org/handle/10986/22550.

Yang, H. M., M. L. G. Macoris, K. C. Galvani, M. T. M. Andrighetti, and D. M. V. Wanderley. 2009. “Assessing the Effects of Temperature on the Population of Aedes Aegypti, the Vector of Dengue.” *Epidemiology and Infection* 137 (8): 1188–1202. doi:10.1017/S0950268809002040.