Distribution Modelling of Gene Drive-Modified Mosquitoes and Their Effects on Wild Populations

Bartlomiej Piotr Kolodziejczyk

DISTRIBUTION MODELLING OF GENE DRIVE-MODIFIED MOSQUITOES
AND THEIR EFFECTS ON WILD POPULATIONS

by

BARTLOMIEJ PIOTR KOLODZIEJCZYK

A THESIS

Submitted in partial fulfillment of the requirements for the degree

MASTER OF SCIENCE

IN

GEOGRAPHICAL INFORMATION SCIENCE

DEPARTMENT OF PHYSICAL GEOGRAPHY AND ECOSYSTEM SCIENCE
LUND UNIVERSITY

2020

Supervisors:
Valentijn Venus and David Tenenbaum

Abstract

Emerging technologies have the potential to bring numerous new opportunities and solutions to the existing challenges, including addressing sustainable development goals (SDGs). Particular hopes are given to areas of the life that require our urgent action, these include but are not limited to medicine and food security. Scientists investigate how these technological advances can be applied for the benefit of humanity by making more robust crops, eliminating diseases, or trying to extend our longevity. One technology that attracted and kept on attracting attention is the so-called "*gene drives.*" Genes in sexually reproducing organisms have, on average, a 50% chance of being inherited by the offspring. There are, however, genes that have a higher chance of being inherited. In a long-term, such dominant genes can affect the entire population by adding, replacing, suppressing, or editing genetic traits. Being able to eradicate invasive local species, alter mosquito genomes to eliminate Zika, dengue fever, and malaria or produce more environmentally robust to plant species is something to aim for.

The study uses computational modelling techniques in which gene drive inheritance model are combined with distribution models of mosquito species to develop unified modelling approach to evaluate the factors related to gene drive altered species and their capability to eradicate population of wild species. The study shows how gene drive altered mosquitoes can influence wild mosquito populations to prevent them from vectoring malaria and other vector-borne diseases.

The study focuses on malaria spread that is associated with one specific species - *Anopheles* mosquitoes. The study area is Kenya due to a number of reported cases of malaria. The proliferation of malaria mosquitoes was selected due to a number of spatial distribution models that have been developed over the years, as well as the availability of existing remote sensing data.

Keywords: Geography, Geographical Information Systems, GIS, Physical Geography, Mosquito Distribution, Malaria Modelling, Gene Drives

Acknowledgments

This journey would not have been possible without the support of my family, professors, and friends. To my family, thank you for your support and encouragement in all my pursuits to follow my dreams. I am especially grateful to my wife, Ranthini. Thank you for your patience, understanding, and ongoing support. Thank you to my mother for guiding me as a person, for believing in me, and always pushing me to go a step further and achieve more.

Thank you to my professors, as well as teaching and research staff from both Lund University and the University of Twente. Your knowledge, expertise, and experience were essential in my learning process, while your patience, guidance, and advice were invaluable in completing courses. The thesis research process enabled me to further learn geographic information tools and natural resources management and enhanced my problem-solving abilities.

I want to give special thanks to my thesis advisor, Mr. Valentijn Venus, from the University of Twente, for guiding me through this master thesis research. I owe a debt of gratitude to Ms. Louise van Leeuwen - de Leeuw, also from the University of Twente for coordinating the master program and for being there to address every question I had throughout my journey. Louise, thank you for your enormous patience. I would also like to thank Prof. David Tenenbaum for coordinating the program from Lund University.

To my friends, thank you for listening, offering me advice, and supporting me through the entire process. Thank you to all the students that I have met and had the of privilege working with during this journey. Thank you for sharing your life stories with me and for your friendship.

'The extraordinary is in what we do, not who we are.'
— Lara Croft, Tomb Rider

Melbourne, 12 June 2020

Table of Contents

List of Figures

List of Tables

List of Abbreviations

SDG(s)	Sustainable Development Goal(s)
GIS	Geographic Information System
DNA	Deoxyribonucleic Acid
P element	
CRISPR	Clustered Regularly Interspaced Short Palindromic Repeats
Cas9	CRISPR associated protein 9
DIYBio	Do-It-Yourself Biology
GMO	Genetically Modified Organism
iGEON	Joint master program in geographic information systems between Lund University and the University of Twente
NLDA	Non-Linear Discriminant Analysis
R	Programming language and statistical computing software environment
MGDrivE	Mosquito Gene Drive Explorer, a simulation framework designed to investigate the population dynamics of various gene drive interventions for mosquito-borne diseases control
US	United States of America
°C	Degrees Celsius

1. Introduction

Gene drives are newly derived genetic engineering tool that can force a genetic trait throughout a population by using a variety of chemical and biological pathways, defying the usual rules of inheritance. Traditionally, genetic traits have a 50-50 chance of being passed along to the next generation.[1] A gene drive can push the gene inheritance rate to nearly 100 percent, making a gene dominant and continuing this trait in all future generations.[2,3] drives use is explicitly limited to sexually reproducing organisms. Hence, gene drives cannot be applied to re-engineer populations of bacteria or viruses.[4,5] Gene While being limited to sexually reproducing organisms, gene drives hold enormous potential to revolutionize disease control, species conservation, agriculture, and other fields.[6,7] Scientists expect to use gene drives to eradicate invasive species,[8-10] re-engineer mosquitoes to remove their ability in spreading Zika, malaria or dengue fever.[11-13] Some studies indicate that gene drives can be used as evolution-warping technology applied to invasive species to control their spread and to conserve biodiversity.[14,15]

Gene drives bring new opportunities to the old issues. However, technology brings also numerous doubts, concerns, and challenges. Gene drives raise severe ethical and practical concerns. Critics and environmentalists are skeptical of this technology, claiming that it may cause long-lasting or even irreversible harm to the environment, biodiversity, and human health. The views of critics and environmental groups are often shared by the scientists and developers of gene drive technology.[16-19] On the other hand, the supporters of the gene drive technology are very optimistic about the opportunities that this technology brings. Among some potential applications, they mention eradicating vector-borne malaria by making mosquitoes incapable of transmitting malaria virus, making coral reefs more robust to environmental conditions, including ocean acidification, bleaching, and rising water temperatures or eradicating invasive species and, hence preserving native biodiversity. Further, the supporters of the technology claim that human mistakes in gene drives have to undergo the strict evolutionary revision process and that evolution and time will eventually fix all wrong human decisions, especially if the gene drive is harmful to the organism, the organism will try to survive by breaking the parasitic gene drive.[4,20] A study by Messer et al. suggests that the gene drive resistance will evolve almost inevitably in standard gene drive systems.[21] While resistance can be foreseen as an important natural safety switch. This may, however, mean that the introduction of the gene drive was pointless in the first place.

Altering entire populations and being able to eradicate species, such as pests, raises ethical and regulatory issues that governmental, scientific, and environmental organizations are currently beginning to investigate.[4,5] Ecosystems are complex and highly unpredictable formations with multiple variables to consider.[4,7,11] Changing or removing one of these variables may bring lethal results.[6,8] For example, removing one species from the ecosystem may disrupt the entire food chain, leading to ecosystem collapse. Some gene drive applications explore the suppression or removal of certain species from their respective ecosystems.[14]

A number of studies have shown that the altered organisms tend to develop evolutionary resistance that can effectively shut down the unwanted gene drive.[20,21] These studies have been performed on smaller populations restrained to the boundary conditions of the lab environment.[4] It remains unknown whether similar behavior would be seen in large wild and uncontrolled groups, where numerous environmental factors must be taken into consideration.[20,21] Because of this gene drives

may not be working in the long-term as planned.[4,14,16] However, the development of the evolutionary resistance will most likely take generations allowing enough time to have desired effects.[15,17,18] Most of the studies showing the development of the evolutionary residence have been performed on organisms that are altered with genes that pose harm to the host.[21] There is no reason why evolutionary resistance would act in cases in which the introduced alterations are of benefit to the host. Many of the technological factors are yet to be understood.

  Due to the immaturity of the technology, gene drives are currently tested on a laboratory scale with small controlled populations.[4,5] None of the gene drive altered organisms have been released into the wild as of yet. Testing gene drives in large native and mobile populations is both unethical and nearly impossible to reverse.[1,2,14] However, scientists are already developing ways to address ethical concerns and overcome some of these reversibility limitations by incorporating certain "safety switches" to make the gene drive organisms more controllable.[15,22] Several scientists have suggested that there is a possibility of introducing a second gene drive to shut down the previous one.[22,23] The strategy is, however, strictly hypothetical.

  Observing the development of genetic traits over generations is a challenging task, even in organisms with a very short lifespan.[11,12,14] This forced researchers to find quicker tools to test their hypotheses. Many complex computer models have been built to approximate how wild populations behave with the introduction of gene drives.[11,23,24] These computer-based simulations allow evaluating population development over many generations in a matter of minutes. However, they have been developed using limited evidence collected from controlled laboratory populations, and as such, results of such simulations may not be reliable when extrapolated to larger wild populations.[1,3,11,23,24] The reliability of those extrapolations is yet to be determined.[11,24] Nevertheless, due to limited evidence and scientific data currently, these simulations are the only feasible evaluation tool. Validation of those models on wild populations would be a breach of ethical codes and many international protocols, as such, no one dares to perform such studies.[4,5]

  The origin of the gene drive idea can be traced back to 2003.[6] Professor Austin Burt of Imperial College London was the first to propose harnessing some of these dominant entities for a range of applications. Burt has described inserting a selfish element into the specific gene to influence the feature inheritance that would be passed onto the offspring. This selfish element would drive itself throughout generations altering the entire population. The study showed that in the 1950s, such a selfish element, called P-element, of unknown origin, altered the DNA of *Drosophila melanogaster* fruit flies in the 1950s, and less than five decades later the P-element managed to spread itself worldwide without human intervention obeying natural borders such as oceans.[25] Today, the P-element can be found in *Drosophila melanogaster* fruit flies throughout the globe. Proving the theory was challenging as genetic engineering tools back then were not as advanced as they are today. It took years, but finally, in 2011, Burt et al. announced in Nature[26] that they had created a homing endonuclease that could find and cut a gene in mosquitoes.

  Only the development of the so-called CRISPR/Cas9 allowed to harness the real power of the technology. CRISPR stands for "*Clustered Regularly Interspaced Short Palindromic Repeats*" and refers to sequences of viral DNA that bacteria have incorporated into their genomes. CRISPR, together with the enzyme referred to as Cas9, helped bacteria to develop defense approaches against viruses.[1,5,21] In 2012, a group of researchers announced that they have managed to modify the bacterial CRISPR/Cas9 defense system mechanism and use the mechanism of bacterial-origin as a gene-editing

technology. CRISPR/Cas9 brings a lot of flexibility, allowing scientists to literally cut and paste nearly any genetic information into any organism.[5,27] CRISPR gene editing has been widely used since. The ease of CRISPR application enables it to be used not only by professional researchers but also by students and citizen scientists, which resulted in further ethical and regulatory issues.

Together with growing interest and intensified research in CRISPR and gene drives, the very same researchers that developed the technology, are urging the governments to regulate this emerging field. In 2014, a group of scientists published a scientific paper titled "*Regulating gene drives*"[17], and a year later, even larger groups of the foremost pioneers in gene drives joined together to publish another scientific paper on "Safeguarding gene drive experiments in the laboratory."[4] Researchers are trying to warn that the accidental release of gene drive altered organisms may be deadly for the wild populations. Such accidental breaches could also be harmful to potential plans of using gene drives for malaria eradication and pest control.[28-30] The proposed guidelines aim to help researchers safeguard their experiments and apply precautionary principles in all their scientific endeavors. Many are worried that the lack of laboratory experience and bioethical knowledge of such individuals may accidentally or intentionally breach and gene drive release to the ecosystems.

The study aimed to apply the existing evidence, data collected from limited and controlled laboratory populations of gene drive altered species, as well as developed gene drive inheritance models together with a similar data collected over the years for vector-borne malaria spread and use remote sensing and spatial modelling to expand on existing knowledge and to combine the two models: mosquito distribution model and gene drive inheritance to establish how usable gene drives can be used for controlling or eradicating mosquito-borne malaria in Kenya. The study also examined how the distribution of gene drive altered mosquitoes, which are made incapable of vectoring malaria, will spread geographically within the selected study area and how this geospatial distribution of gene drive altered versus wild mosquito population will shape over a number of generations. These methodologies and techniques learned during the iGEON courses will be combined with other available techniques to make a complete model of gene drive mosquitoes distribution with the aim of eradicating malaria.

1.1 Research Problem Statement

Gene drives are solutions that are effective only when released among wild populations. However, current genetic engineering regulations are designed to keep genetically engineered organisms away from wild populations. It is unclear whether gene drives will ever gain enough approval from the governments, environmentalists, and the community to leave the lab. The government support for gene drives will not be easy unless current regulations will change. Since gene drive-modified organisms, just like any type of fauna and flora, do not respect international boundaries and treaties, a country that released gene drive organisms could be accused of violating the United Nations Biological Weapons Convention if gene drive-altered organisms affect native species in the neighbouring country.

Moreover, in such cases, another international treaty would be violated. The Cartagena Protocol on Biosafety governs the cross-border movement of genetically modified organisms. The release of gene drives would most likely violate the treaty causing at least intergovernmental conflict, if not an environmental catastrophe.

The study applies computational models, including geographic information systems, remote sensing, and distribution modelling, to evaluate the spread of gene-

drive modified mosquitoes and their effects on the wild population to eradicate malaria in Kenya. The models rely on limited information collected in the laboratory setting, in which gene drive altered organisms have been exposed to the wild population, and their effects and spread have been evaluated over several generations. These models have then been extrapolated onto larger, wild populations. The results of experiments performed on limited and controlled laboratory populations, may not be directly transferable to larger, wild populations. However, these results give a good indication and are currently the most reliable source of information. Releasing gene drives into the wild, would breach a number of protocols and could potentially pose harm to the environment, biodiversity, and human health, as such further understanding of the gene drive technology have to be established based on controlled laboratory experiments before gene drives altered organisms are released into the wild populations. The lack of validation data for models developed in the study has been identified as the main limiting factor.

Spatial distribution models for mosquito-borne malaria virus spread and gene drive inheritance showing how gene drives can spread within the wild populations have been separately researched and developed over the last decade. However, there is an obvious need for detailed models where the spatial distribution of mosquito-borne malaria virus spread and gene drive inheritance are simulated collectively to illustrate and visualize how gene drive altered mosquitoes that are incapable of vectoring malaria can spread over the study area and help to limit, control, or completely eradicate malaria. While these two models have been independently developed and applied, combining them closes a significant knowledge gap and allows to more precisely evaluate the applicability of gene drive technology for eradicating malaria. Previous gene drive models were limited and restricted. The main limitation of those models was the lack of geographic dimension. Gene drive models developed to date assumed that all organisms were limited to one spot, as such geographic location or species distribution was never considered before in those studies, making the models incomplete and inaccurate.

The current research adds a geographic dimension to the previous studies and allows the evaluation of geographic location to see how it affects and influences the spread of gene drives. As such, these two models combined allow for a more realistic scenario in which the geographic location of the species is one of the variables. Previous gene drive models allowed for the establishment of parameters such as a chance of inheriting gene drive, number of gene drive organisms introduced into the wild population, or number of generations influence the spread of gene drives. The current study expands on previous capabilities and, in addition to the above, adds geographic dimension, which allows establishing how the geographic location of existing wild species and geographic location of where gene drive altered species are introduced influences the spread of gene drives. This will allow establishing future most effective strategies for gene drive introduction and will allow determining whether it is best to introduce all initial, drive mosquitoes in one geographic location, or it is best to introduce it in different locations, as well as find the best practices for selecting the location.

The new model can have significant scientific and societal implications, as it will allow to simulate complex scenarios and find the best potential practices to fight one of the world's grand challenges, which is malaria disease. Similar models can be developed for other vector-borne diseases such as Chikungunya, dengue fever, lymphatic filariasis, Rift Valley fever, yellow fever, Zika, West Nile fever, Lyme

disease, and many more. Vector-borne diseases account for more than 17% of all infectious diseases, causing more than 700,000 deaths annually.

Furthermore, the study results are presented in a few different forms and are consulted with the public to evaluate socio-journalistic factors to communicate the findings as a widely accessible and understandable policy brief.

1.2 Research Questions and Objectives

The study aims to use the existing gene drive framework to model the distribution for mosquito-spread and apply the developed model to the selected study area to evaluate the spread of gene drive mosquitoes and the replacement rate of wild mosquito populations gene drive altered organisms.

The overall objective of this study is to establish how gene drive altered mosquitoes will distribute over study area within a given time limit, counted in days, and further whether this distribution over given generation will be suitable to eradicate malaria within the selected study area. This will allow testing how gene drives behave based on initial gene drive inheritance frequency, and a number of gene drive organisms as a percentage of the wild population at the release location (Malindi, Kenya). These simulations will close a significant knowledge gap as these kinds of experiments has never been performed in a real environment or even using computational simulations.

For this main study question has been developed: *Considering the outputs of developed distribution model for gene drive mosquitoes, can gene drives eradicate malaria?*

The hypothesis for the main study question was that based on initial parameters such as the location of release, number of released gene drive mosquitoes, landscape, gene drive frequency, etc. gene drives can be either effective or less effective.

A series of four supporting objectives have been derived to facilitate the overall objective, and corresponding four support questions, these are:

- To establish how an initial number of gene drive altered mosquitoes affects the rate of distribution within the wild population,
  Based on different initial numbers of gene drive mosquitoes introduced to wild population, how many days will be required to eradicate malaria?
  The hypothesis for the supporting question is that the larger number of gene drive mosquitoes relative to the number of mosquitoes in the wild population, the faster the gene drive organisms can replace wild mosquitoes. Hypothesis for this supporting question can be written as:
  H_0: *There is a no relationship between a number of gene drive mosquitoes to a number of mosquitoes in the wild population and the number of days to replace the entire wild population.*
  H_A: *There is a relationship between a number of gene drive mosquitoes to a number of mosquitoes in the wild population and the number of days to replace the entire wild population.*

- To evaluate the role of gene drive inheritance frequency in the distribution of gene drive-modified mosquitoes within the wild population,
  How different inheritance frequencies for gene drive mosquitoes introduced to wild population affect the number of populations required to eradicate malaria?

The hypothesis for the supporting question is that higher inheritance frequencies for gene drives will lead to a reduction in the number of days required to replace wild mosquitoes and vice versa. Lower inheritance frequencies will increase the number of days required for gene drive mosquitoes to replace wild mosquitoes. The hypothesis can be expressed as:

H_0: *There is a no relationship between inheritance frequencies and a number of days to replace the entire wild population.*

H_A: *There is a relationship between inheritance frequencies and a number of days to replace the entire wild population.*

- To compare how different initial factors influence the distribution of the gene drive-modified mosquitoes,
  What is the coverage of the gene drive mosquitoes, and do they spread within the entire study area?

The hypothesis for the third supporting question is that with the current semi-spatial model which does not include landscape, weather, or other limitations, mosquitoes should spread within the entire study area given sufficient time. Sufficient time, in this case, is considered to be five years. The hypothesis can be written as:

H_0: *Gene drive mosquitoes require five years or less to spread and be present throughout the entire study area.*

H_A: *Gene drive mosquitoes require over five years to spread and be present throughout the entire study area.*

- To understand whether gene drive technology will be an effective solution for eradicating malaria in Kenya within the next decade, based on the life cycle of mosquitoes and different initial factors.
  Are gene drives effective tool to completely eradicate mosquito-borne malaria in Kenya the next 10 years?

The hypothesis for the final supporting question is gene drive altered mosquitoes can replace the entire wild mosquito population within only 10 years, and by doing so eradicate malaria, given no resistance to gene drives is developed. Considering purely replacement of wild population within a decade, the hypothesis can be written as:

H_0: *Genetically altered organisms can replace the entire wild population within a decade.*

H_A: *Genetically altered organisms cannot replace the entire wild population within a decade.*

Previous attempts to address the above or similar research problem focused mostly on genetic aspects of the problem. While the inheritance ratios are well understood in terms of genetics and based on that predicting how many generations (iterations) will be needed to replace the wild population is relatively straightforward. However, a current scenario which aims to emulate real-world problem adds a significant degree of complexity to it. The current simulation does not consider the mosquito population as a whole but separates the population into smaller sub-populations with various size representing cities in Kenya. Further, the geographical factor between each of the nodes/cities influences the rate of inheritance.

A philosophical approach to the problem and research methods designed to provide answers all research questions are described in the following chapter.

2. Background Information

There are numerous factors to consider which influence malaria spread. Most importantly, the presence of female *Anopheles* mosquitoes is required for malaria transmission, as only female *Anopheles* mosquitoes are capable of vectoring malaria.[31] A person bitten by female *Anopheles* mosquito infected with malaria virus has a chance of being infected by the virus. The parasite passes from the mosquito into a new host's bloodstream, infecting the bitten host with malaria. Female mosquitoes can also contract the parasite from an infected human host and then infect a new host.[32] However, there are also other factors, including physical and climatic, as well as human behaviour and genetic factors.[33] This chapter describes factors that influence malaria spread, as well as factors that influence gene drive inheritance.

2.1 Physical and Climatic Factors Influencing Malaria Spread

Factors that influence the spread of the malaria virus can be divided into three categories: physical factors, human factors, and genetic factors. Physical factors that have been identified as ideal and supporting of the pattern and level of malaria transmission include temperature generally ranging from 15 to 40 °C.[34,35] Although, the minimum temperature for mosquito development is between 8 to 10 °C.[36,37] The optimum temperature has been identified to range from 25 to about 27 °C,[38,39] whereas the maximum temperature is around 40 °C.[35] The ranges of minimum and maximum temperature greatly affect the development of the mosquito, which determines malaria transmission. Studies showed that higher temperatures allow for quicker mosquito larva development.[36,39] Higher temperatures also have been linked to the increase in feeding habits and the number of eggs laid by mosquitoes.[40,41] Altogether effectively increasing the number of mosquitoes in a given area.

Different studies also link the temperature with the life cycle of the malaria parasite.[35,39,42,43] The average time required for the malaria parasite to complete its development in the gut of female mosquito is about ten days.[42] Depending on the temperature, the development time can be shorter or longer.[43] The number of days necessary to complete the development for a given *Plasmodium* species increases with the temperature decrease.[43] Malaria parasites *Plasmodium vivax* and *Plasmodium falciparum* has the shortest development cycles, making them more common than *Plasmodium ovale*, or *Plasmodium malariae*.[44-46] The time required for the malaria parasite to complete its development in the mosquito, decreases to less than 10 days for temperatures ranging from 21 to 27 °C, whereas 27 °C is considered to be the optimal development temperature.[42,47] Similarly to the mosquito host, the maximum temperature for the malaria parasite development is 40 °C.[38] The minimum temperatures for the parasite development are ranging from 14 to 19 °C.[42,47] However, in temperatures below 18 °C, the development of *Plasmodium falciparum* is limited. *Plasmodium vivax* species are more robust and can survive lower temperatures than *P. falciparum*.[39] Studies have shown cases of malaria transmission in areas colder than 18 °C. However, these cases include transmission by the *Anopheles* mosquitoes inside of households or other buildings, where the temperature tends to be warmer than the outside temperature.[48,49]

In certain countries where elevation changes significantly, the altitude has been indirectly linked to malaria transmission.[50,51] The altitude influences the distribution and transmission of malaria through its effect on temperature patterns.[50] With an increase in altitude, the temperature decreases, making lowlands warmer and highlands colder. For instance, in Kenya, the altitude varies from in Indian Ocean coast at the sea

level to over 5,000 meters above the sea level. As such, the malaria distribution and transmission patterns vary depending on the altitude and further depending on the seasonal temperature patterns. As such, Kenyan highlands, with altitudes between 1,800 to 2,400 meters, are exposed to malaria transmission only for a short period of time during the summer when the temperature is high.[52,53] The reason for this is that it is normally too cold for mosquitoes to develop in large numbers, or for the malaria parasite to develop inside the mosquito vector.[39,47] Similar pattern can be likened to geographical latitude. Areas closer to the equator are associated with higher and more uniform temperature patterns throughout the year.[54,55] Temperatures tend to decrease and be more prone to annual variations as individuals move away, both north and south, from the equator. As such, the warm and relatively constant temperature near the equator is favourable for both mosquito and malaria parasite development.[38,43]

Anopheles mosquitoes breed in the water environment, as such high and often rainfalls have been identified as a direct factor supporting mosquito growth, making the tropics ideal environment for malaria.[56,57] The right amount of water is essential for mosquitoes to breed. In some equatorial countries, water bodies to support *Anopheles* mosquito breeding appear mainly after the rain, and therefore the highest number of malaria cases occur during the rainy season.[56] Continuous as well as on-and-off rain patterns tend to flush away mosquito breeding habitats.[57,58] The flushing has more impactful in the highlands and hilly areas, making again highland landscape disadvantageous for mosquito development and malaria parasite transmission.[59,60] Not all types of water are suitable for mosquito development. Importantly, *Anopheles* mosquitoes that vector malaria parasites prefer to breed in freshwater or muddy water, but do not breed in polluted or foul-smelling water.[57]

Malaria vectoring mosquitoes breed mainly in stagnant water bodies, cases in which mosquitoes breed in slowly moving water has also been seen, but rapidly flowing water is a preventive factor for mosquito development.[56] Areas with low rainfall and drought but covered with vegetation throughout the year and flowing water bodies, such as rivers or streams, can favour mosquito development and malaria transmission.[61,62] This is because the interrupted flow of the water in the stream by delayed rainfall or drought, supports the formation of the local pools along the steam. These temporary pools enable favourable conditions for mosquito breeding.[62]

The relative humidity level is defined as the amount of moisture in the air and expressed as a percentage. High relative humidity levels, ideally of 60% and above, together with stagnant water, i.e. lake, pond, etc., provides ideal conditions for larval *Anopheles* mosquito development.[38] Relative humidity is directly linked to the activity and survival rates of mosquitoes, and mosquitoes need to live at least 8 to 10 days (depending on physical conditions), to be able to vector malaria.[38] Certain studies have shown that relative humidity below 60% is shortening mosquito lifespan, effectively preventing it from being able to become malaria vector.[36] In drier areas, rainfall can indirectly affect malaria vectoring through its effect on relative humidity.[36,63] Vegetation cover has been shown to increase the relative humidity of the surrounding area, and rainfall effectively allows for vegetation to grow.

2.2 Climate Change and Land Use

As explained above, physical factors - predominantly temperature and rainfall - have a direct effect on the malaria vector and malaria parasite. A number of recent studies have shown a correlation of climate change, and increasing temperature tends to rise of malaria in certain parts of the world.[47,63] A research published in The Quarterly

Review of Biology reports that increasing temperature may not be the sole contributor to the increase in malaria cases, but other factors such as climate change-forced migration and land-use change are likely contributors too.[64] The research aim of this published study was to answer conflicting theories that try to explain why malaria has been spreading into high altitude areas of Indonesia, Afghanistan, East Africa, and elsewhere. As described before, high altitude due to different physical conditions was a preventing factor for malaria spread.[51]

Researchers considered many climate change-related factors that could enable malaria to spread. As shown in the paper,[65] climate change-forced human migration from lowlands into highland areas might be introducing the malaria parasite. In which case, humans are malaria vectors. Climate-lead changes in land use and farming practices also are enabling factors for mosquito development. Previously, mainly wild highland areas are effectively being converted into farming intensive areas where irrigation systems are being developed, allowing the mosquitoes to breed. Another study has correlated the increase in maize farming in Ethiopia to increase in local malaria cases. The maize pollen diet was an enabling factor for the growth of immature and aquatic stages of mosquito larva.[66]

2.3 Human, and Non-Climatic Factors Influencing Malaria Spread

Human factors often called behavioural factors, which include factors related to human activity and human behavioural habits that enable the spread of malaria. These factors are often dictated by socio-economic settings and include all sorts of dwelling developments that create additional shade for mosquitoes.[67-69] Poor sanitation in underdeveloped areas, i.e., shanty towns and slums, as well as progressing urbanisation are also contributing factors.[70,71] Urbanisation results in the development of large and congested settlement areas and as a result, an easy target for bloodthirsty mosquitoes.[70] Some other studies have concluded that rapid unplanned urban development can create new breeding habitats for malaria vectors.[72,73] However, certain studies are contrary to these claims, finding that the incidence of malaria is generally lower in urban areas compared to rural areas, and there are a number of factors for this. For example, mosquito breeding sites are limited in urban areas due to the space being covered by houses.[67,74] Further, certain malaria vectors prefer clean water, whereas most of the water collections in urban areas are polluted and unfavourable for mosquito breeding habitats.[62] Urban areas have more and better access to healthcare and malaria prevention strategies compared to rural areas.[75,76]

Human population movements are a significant contributing factor to malaria transmission.[67,77] Within the African continent, the majority of movements involve people relocating from less developed highlands to malaria-endemic lowlands. These relocations include both permanent and temporary movements, i.e., seasonal labourers. These labourers are often working in the agricultural industry, for example, harvesting or planting crops during mosquito peak season.[78,79] Poor living conditions and sanitation, often in temporary housing, and limited healthcare only worsen malaria problems.[70]

Migrants from malaria-free highlands, whether permanent or temporary, lack immunity against the malaria parasite and often do not have appropriate knowledge about malaria transmission process or steps can be taken to reduce the risk.[79] Further explanation about genetic immunity to malaria is described later in this chapter under genetic factors influencing malaria. Temporary migration from, often densely populated, highlands to fertile lowlands introduces malaria problem in the highland

areas from which migrants came and to which they return at the end of the season. Temporary migrant labourers often bring back malaria parasites to malaria-free highlands,[80,81] and although climatic conditions in the highlands make it more difficult for both malaria parasite and mosquitoes to develop, it can still be possible,[36] potentially resulting in sporadic and local epidemics that can affect a large percentage of the non-immune highland population.[60]

Migration of large population displacements caused by war, civil unrest, or natural causes such as flooding, drought, famine, etc. from areas with malaria can introduce or reintroduce malaria disease into malaria-free areas, and vice versa, population displacements from malaria-free into malaria-endemic areas.[82] Malnutrition, which is often a common issue, during such displacements only worsens malaria problem.[83,84] Global air travel contributes to the spread of malaria through so-called "*airport malaria*," a term coined by researchers to explain the more recent spread of malaria to areas such as the US and Europe.[85,86]

Human-made areas of stagnant water, i.e., irrigation ditches and channels, dams, ponds or burrow pits, provide additional breeding grounds for mosquitoes increasing the incidence of malaria in villages or urban areas that are located near water collections.[87] Previous studies resulted in contrary outcomes. Raising domestic animals in close approximation to households can provide an alternative source of blood for *Anopheles* mosquitoes, and as such decrease, the human risk.[88,89] On the contrary, this alternate source of blood can also enable increased breeding of *Anopheles* mosquitoes and thus increase the human exposure.[90] Agricultural work, especially in irrigated fields or agricultural areas with high water concentration, creating additional breeding sites for malaria mosquitoes and forcing increased exposure that may lead to increased malaria transmission.[91] For example, the use of irrigation and flooding of agricultural land for rice cultivation has long been linked to an increase in the vector number corresponding to increase malaria cases.[92]

Lack of appropriate protection, i.e., bed nets, mosquito replant, being a result of economic or knowledge reasons, is often listed as a direct cause of malaria infection.[93,94] Whereas the lack of appropriate knowledge to recognize and treat malaria promptly and appropriately often results in further spread.[95] Cultural beliefs and deeply rooted traditional knowledge about malaria prevention methods, as well as the use of natural medicine, often result in ineffective treatment of malaria.[96]

Human behaviour factors in malaria-endemic countries also determine the success rate of malaria prevention. Whereas some richer countries located in the area with preferential physical conditions for malaria vectoring have been able to decrease or eradicate malaria and other vector-borne diseases through various strategies, including water and sanitation conditions, as well as developing effective communication and education programs among others. The poorer malaria-endemic countries often lack appropriate financial resources.[97,98] The lack of financial resources often results in health workers being overworked and underpaid. The lack of drugs, appropriate equipment and training, and supervision is also a limiting factor.[98,99] Moreover, developing countries suffer from regulatory shortfall and corruption, both being enablers of creation of black markets for fake and ineffective malaria medicines.[100-102]

2.4 Mosquito and Parasite Factors Influencing Malaria

Other factors that influence malaria transmission, but which are not related to the climate, are often called non-climatic factors and can include a number of

parameters such as the type of vector, the type of parasite, drug resistance in parasite or insecticide resistance in mosquitoes, all influence the incidence and severity of malaria.[103] As a matter of fact, the environmental development and urbanisation, as well as population movement and migration, described above, also fall within non-climatic factors. However, since they are closely related to human behaviour and habits, they have been described as a separate group. The level of immunity to malaria in the human hosts is also considered a non-climatic factor. However, for the purpose of the study malaria resistance was categorized as a genetic factor and is described in the next sub-section. This sub-section focuses on mosquito and parasite factors that influence malaria spread.

As it was already briefly explained in the introduction chapter, not all mosquitoes are capable of vectoring malaria. Only females of *Anopheles* mosquitoes can carry the malaria parasite.[103]

Various species of *Anopheles* mosquitoes differ in their ability to vectoring malaria. This ability of a mosquito to be able to vector malaria parasite depends on many factors, including the biology and behavioural patterns of the mosquitoes.[103,104] Mosquitoes in the *Anopheles gambiae* group are known to be the most efficient malaria vectors in the world, and the reason for this is that it feeds predominantly on humans. This group of mosquitoes is found only in Africa, and there is a much higher incidence of malaria in Africa compared to the rest of the world and is associated to the efficiency of this mosquito group in vectoring malaria parasites.[38,103]

There are about 3,500 species of mosquitoes in the world, classification of major mosquito groups is shown in Figure 1. However, there are only about 460 *Anopheles* mosquitoes, of which only about 100 can transmit human malaria, only 30 to 40 commonly transmit parasites of the genus *Plasmodium*, which causes malaria in human endemic areas.[105-107] Other species of mosquitoes are responsible for vectoring other diseases, viruses, and parasites. For example, yellow fever, dengue fever, and Chikungunya are vectored mostly by *Aedes aegypti* mosquitoes.[105,108,109]
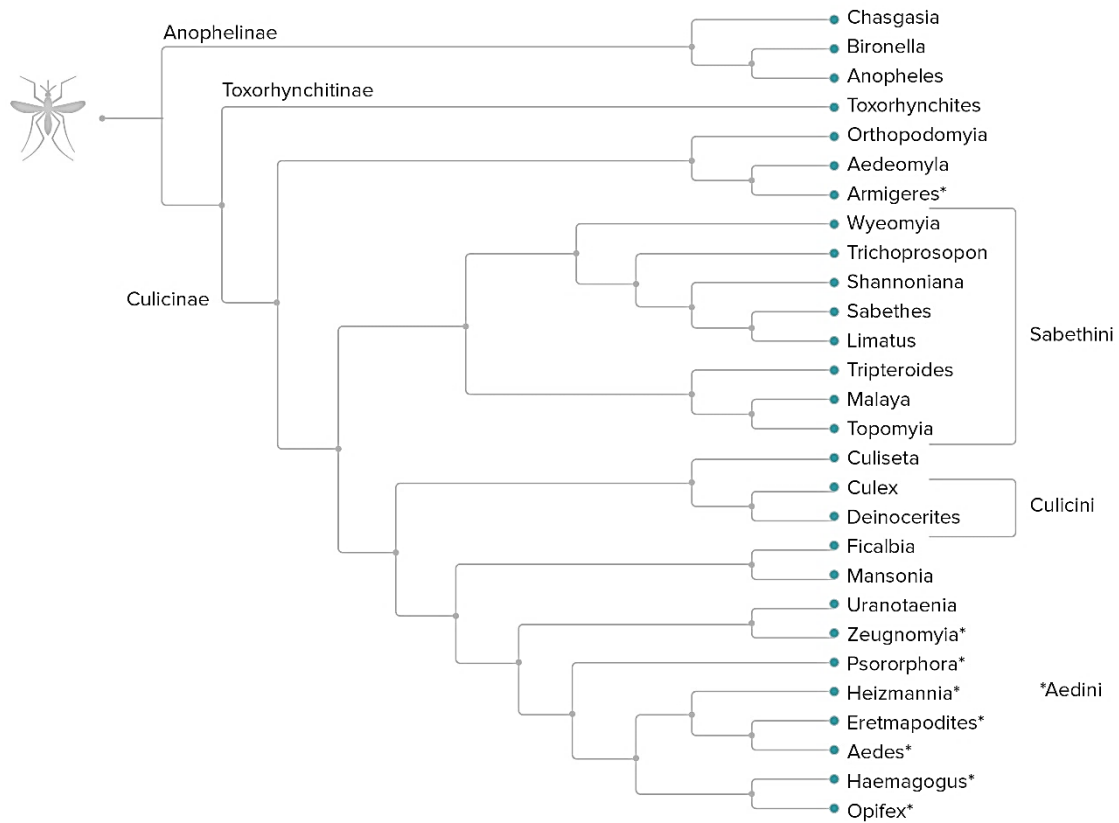
Figure 1 *Mosquito classification. Reproduced from Harbach.*[110]

Female mosquitoes need either human or animal blood to reproduce and develop their eggs.[103,111] Some mosquitoes prefer human blood over animal blood, and these, for obvious reasons, are more efficient malaria vectors.[111] Mosquito groups that feed exclusively on animals are not malaria vectors, whereas those that feed equality on humans and animals are considered weaker malaria vectors.[112] As such, mosquito type and their feeding behaviour is one of the main factors for malaria vectoring.

Mosquito breeding habits are also important factors. Some mosquitoes have adapted to breed in a wide range of water environments, including small sun-exposed pools, whereas others prefer larger stagnant, and shaded water bodies.[113] The mosquitoes that are able to breed in various environments are better malaria vectors than the mosquitoes that are more selective in their breeding habitats.[114,115] *A. gambiae* group is well adapted to breed in multiple settings and as such becomes even better malaria vectors, making *A. gambiae* mosquitoes responsible for much of malaria vectoring in African continent. *A. arabiensis* group, which is the primary malaria vector in Ethiopia, can breed in a variety of water environments and prefers breeding habitats closer to human settlements.[56,57] However, *A. arabiensis'* primary preferences are water collections created immediately after the rain.[116]

The type of malaria parasite (Figure 2) is another critical factor influencing the spread and its efficiency. There are four types of malaria parasite that can infect people. These parasites are single-cell protozoa that are invisible for a bare eye and can be seen only under a microscope.[117,118] These four different *Plasmodium* species have already been introduced before in relation to their temperature habits. However, *Plasmodium* species also differ in their efficiency.[119]

	P. falciparum	P. vivax	P. malariae	P. ovale
RINGS				
SCHIZONTS				
GAMETOCYTES				
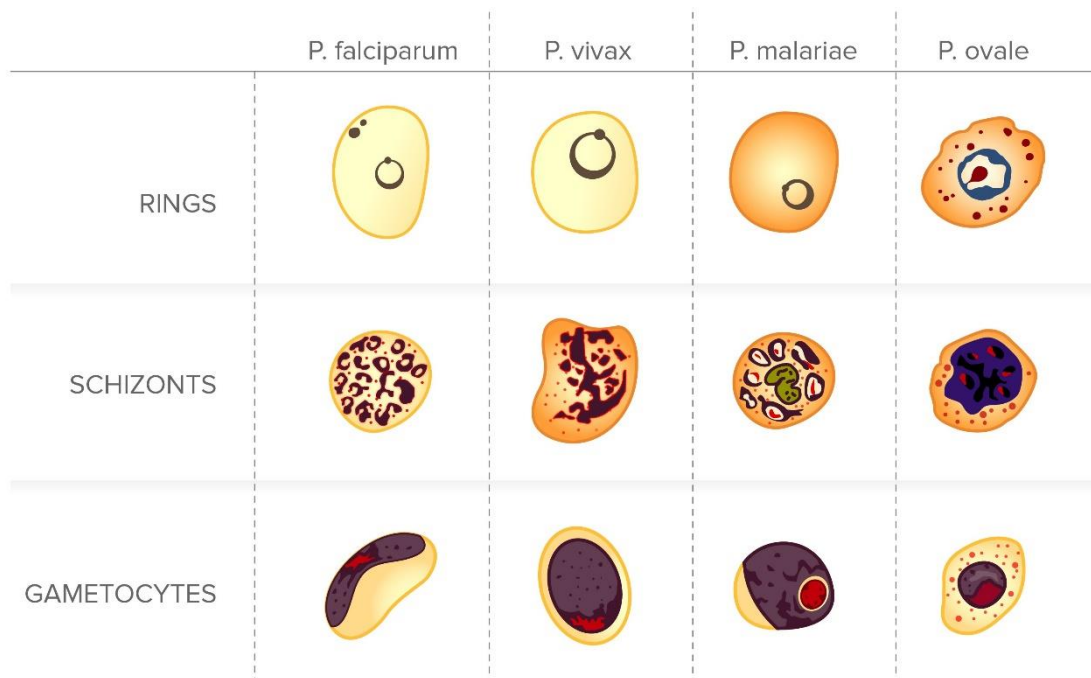
Figure 2 *Image of the malaria parasite drawn based on microscopic images. Reproduced from Nanoti et al.[120]*

Plasmodium falciparum malaria parasite is considered to be the most invasive and efficient type among the four parasites that can infect people.[121,122] *Plasmodium falciparum* has the shortest development cycles, making them more common than other *Plasmodium* types.[122,123] *Plasmodium falciparum* is more common in Africa compared to other parts of the world and is one of the reasons for more malaria-related deaths in Africa than elsewhere.[124,125]

2.4.1   Insecticide resistance in vectors

Some chemicals are used to kill mosquitoes and other insects to prevent from vectoring diseases, including malaria.[126] However, repeated application of insecticides can lead to the development of insecticide resistance in mosquitoes, meaning that these mosquitoes have complete or partial immunity to these insecticides.[127,128] The developed resistance will allow a large number of mosquitoes to survive, at the same time increasing the number of malaria cases.

2.4.2   Drug resistance in malaria parasites

Similarly, to insecticide resistance in the aforementioned vectors, malaria parasites can develop drug resistance.[129] The anti-malaria drugs are designed to kill the malaria parasite inside the human host body. However, after repeated use of anti-malaria medicine, malaria parasite can develop complete or partial resistance to that particular drug or similar medicines.[130] The developed drug resistance, makes anti-malaria treatment ineffective, preventing patients from being treated unless new drugs are developed.[130,131] This poses a more significant threat since the procedure is ineffective the malaria parasite can be contracted by mosquitoes and transmitted into a new host who then develops drug-resistant malaria. This can result in a spread of drug-

resistant malaria, resulting in more host deaths due to lack of effective treatment.[132,133,134]

2.5 Genetic Factors Influencing Malaria Spread

Genetic factors can also influence malaria spread.[135] These factors are present from birth and often inherited from one of the parents, effectively protecting against certain malaria types.[136,137] It is commonly agreed that two genetic traits associated with human red blood cells have specific epidemiologic importance against malaria.[138]

Potential human hosts who have a sickle cell disease, which is a group of disorders that affects hemoglobin, and are relatively protected against *Plasmodium falciparum* malaria, making their disorder a biological advantage.[138] People with sickle cell disease have atypical hemoglobin molecules called hemoglobin S, which can distort red blood cells into a sickle, or crescent shape. Sickle cell disease is more frequent in Africa and in-person of African ancestry compared to other ethnic groups, which over centuries lead to certain protection against *Plasmodium falciparum* malaria in the African continent.[139] Generally, different blood cell and hemoglobin-related disorders such as G6PD deficiency, hemoglobin C, and thalassemia, are more common in malaria-endemic areas and are believed to provide certain protection or immunity to malaria.[140,141]

Persons who lack antigens of the Duffy glycoprotein, a receptor for chemicals that are secreted by blood cells during inflammation, which also happens to be a receptor for *Plasmodium vivax*, are relatively resistant to invasion by *P. vivax*.[142,143] The majority of Africans have developed evolutionary lack of Duffy antigens, making *P. vivax* rare in Africa, specifically West Africa and south of the Sahara Desert.[144] The role of *P. vivax* in those areas was taken over by *P. ovale* which is capable of infecting people lacking Duffy antigens.[145]

To a lesser extent, some other genetic factors associated with red blood cells can influence partial malaria immunity. Various genetic factors, including the human leukocyte antigen complex, which is responsible for the regulation of the immune system in humans, can influence the risk of developing severe malaria.[146]

Interestingly, individuals can develop partial malaria immunity. So-called acquired immunity is commonly known and broadly accepted theory, in which a human target of repeated malaria attacks develops a partial protective immunity.[139] A person who develops partial immunity often can be infected by the malaria parasite but may not develop severe disease and often lacks any malaria symptoms.[139] In regions where *P. falciparum* infections are high and common, newborn babies are protected for the first few months of their life most likely due to antibodies transferred from the mother. However, the number of antibodies decreases with time, making these children vulnerable to malaria. If the children survive multiple infections, they will develop acquired malaria immunity.[147,148] In high malaria transmission areas, such as parts of Africa, newborns and young children are major risk group. In regions with lower malaria transmission rates, including Latin America and the Caribbean as well as Asia, malaria infections happens less frequently, and as such, a smaller number of the population have acquired malaria immunity.[107,149] Making these areas are more prone to epidemics.

2.5.1 Interruption of control and prevention measures

Malaria is both preventable and curable disease.[148] However, combating malaria requires the long-term and sustained implementation of prevention and control

measures to reduce or eliminate the problem from a given area.[104,122] Studies have shown that a result of successful long-term measures to reduce or eliminate malaria, a local population can slowly lose their immunity to malaria parasite in a given area where the measures have been successfully applied.[147] Repeated malaria infections are necessary to develop partial or complete immunity to malaria.[139] The immunity fades away or can be lost entirely, where a person is not exposed to malaria anymore, for example, is continuously protected from malaria infection for several years or moves out of a malaria-endemic area.[150] As such, continuous malaria prevention and control measures must be in place until malaria is eliminated.[95] Otherwise, the disease can surge back, affecting more severely and affecting larger population groups, due to lost malaria immunity.

2.6 Factors and Limitations Influencing Gene Drives

2.6.1 Limitations of gene drive technology

The normal inheritance and gene drive inheritance mechanisms are shown in Figure 3. In normal inheritance mechanisms, genes have a 50-50 percent chance (or close to that) of being inherited, resulting in genetic diversity. In gene drive inheritance, specific genes have a higher chance (often close to 100 percent) of being inherited. Hence, after several generations, gene drives are present in the majority of the population, given that gene drive resistance does not occur in the process.
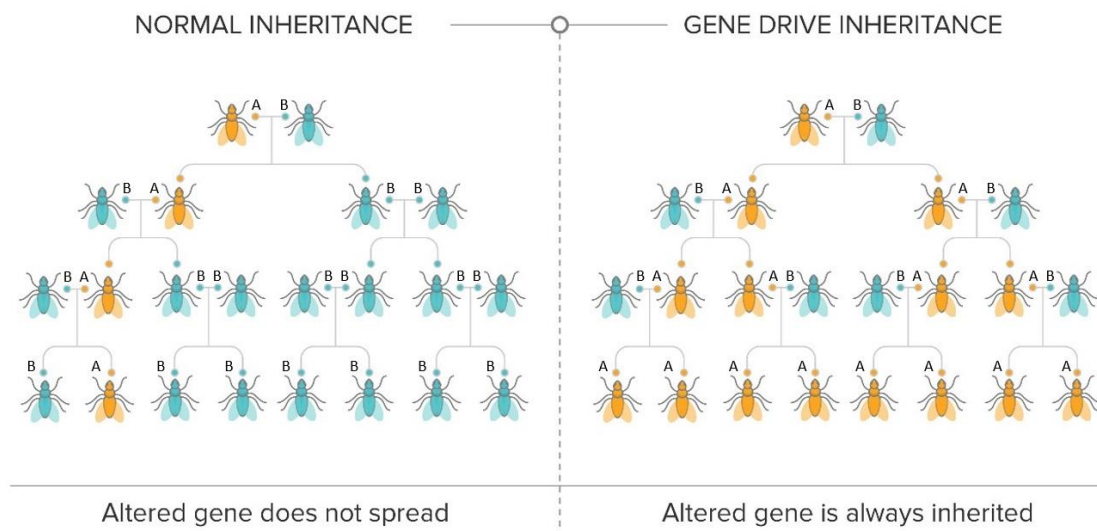
Figure 3 Comparison of ordinary inheritance and gene drive inheritance mechanisms. Letters A and B correspond to different types of inherited genes.

The main flaw in gene drive applicability is the need for altered species to reproduce sexually. This is because gene drives propagate by replacing other alleles that contain a cutting site and the corresponding homologies.[1,18] Asexually reproducing populations will not be affected by gene drives. The time required to spread the gene drive-altered trait to most individuals in particular population depends on the lifespan of the species, mating frequency, number of drive-carrying individuals that are released to the wild population, size of the native population, the efficiency of the drive, gene flow dynamics and more.[4,24,28] Generally, gene drives require many generations to spread through populations.[24] As such, gene drives will work faster in fast-reproducing species and will even faster if these altered species are released in large numbers.[1]

Previous computational studies have shown that introducing only a small percentage of gene drive organisms into a population of wild organisms would take several generations for gene drives to replace most of the wild population.[23,151,152]

Some types of genetic alterations will most likely need to be continuously reintroduced, as it is believed that traits that are harmful to the host will be effectively eliminated or broken by genetic mechanisms of the host.[19] As such, newly introduced gene drives would have to override the eliminated or broken version with new functional copies. Periodic introduction of new gene drives would support the efficiency of the entire process leading to desired results.[153]

Certain issues related to the application of gene drives have also been identified by experts. For example, mutations could occur mid-drive processes, which could potentially allow unwanted traits to be inherited by the offspring, effectively spreading into the entire population or its part.[151] Cross-breeding or gene flow can potentially allow a gene drive to move beyond its target population.[153,154] Whereas, inbreeding within the population could be a way to avoid the spread of gene drives.[152] Potential ecological impacts of gene drives are difficult to identify. Even when the direct impact of traits on a target species is well understood, the gene drives can have side effects on the ecosystem including other species.[155]

There are several factors that can affect the efficacy of CRISPR-based gene drives. Gene drive that remains stable over time will spread to more individuals, where gene drive that accumulates mutations can evolve resistance in the host to the drive mechanism. The fitness cost of a gene drive is also one of the factors that influences drive's spread within the population.[151]

2.6.2 Gene drive resistance

Since gene drives can never more than double in frequency with each generation, a gene drive introduced in a single individual typically requires dozens of generations to affect a substantial fraction of a population. A recent study by de Jong (2017) has shown that gene drives do not always increase in frequency.[3] When gene drives have a significant adverse effect on the fitness of the population, the spread of that gene drive depends on a threshold, or the gene drive may not spread at all. The spread of harmful gene drives can further be limited by developed resistance to the particular gene drive.[3]

While many things related to gene drives are still poorly understood, and performing a risk assessment can pose a risk itself, researchers are certain that wild populations will develop resistance to the modifications.[1,21] Multiple efforts have been initiated to identify and understand the mechanisms of the resistance and its development so that the problem can be addressed. In late 2015, researchers reported a CRISPR gene drive that caused an infertility mutation in female mosquitoes to be passed on to all their offspring.[27] Further laboratory experiments showed that the infertility mutation increased in frequency over several generations, but resistance to the gene drive was developed simultaneously, effectively preventing some mosquitoes from inheriting the modified gene.

Two other studies described in *Nature* have identified the CRISPR system being a source of resistance itself.[156] CRISPR mechanism uses an enzyme to cut a specific DNA sequence and insert a new genetic code in its place. However, rarely, cells sew the incision back together after adding or deleting random DNA letters. This rare occurrence can result in a situation in which the sequence is no longer recognized by the CRISPR gene-drive system, preventing the spread of the modified code. Natural

genetic variation can be another route to gene drive resistance. CRISPR-based gene drives mechanism works by recognizing short genetic sequences. However, variations or differences at these sites would effectively make the host immune to gene drive alteration. Research performed on 765 wild *Anopheles* mosquitoes across Africa analysed their genomes, finding extreme genetic diversity within the population. This would limit, or to some extent, prevent successful application of gene drives.[157]

In another study,[21] researchers from the University of Kansas and Cornell University have used computational models of gene drive inheritance and incorporated potential resistance mechanisms as well as random genetic drift in the model. The comprehensive framework for CRISPR-based gene drive modelling of population genetic showed that resistance to standard CRISPR-based gene drives should evolve almost inevitably in the majority of natural populations. The key factor influencing the probability of resistance evolution was found to be the overall rate at which resistance alleles arise at the population level by mutation. The study also concluded that factors such as the conversion efficiency of the gene drive, gene drive fitness cost, and introduction frequency of the gene drive have a minor impact on the gene drive resistance development. Figure A2 in Appendix A shows the conceptual diagram with gene drives introduced to the system.

2.7 Ontologies

The current study relies on the interplay of three distinctly different fields, namely spatial modelling or geographic information systems, synthetic biology and gene drives, and malaria study. All three of these fields have very little in common, and only in the last few years, researchers have been trying to employ spatial modelling to answer synthetic biology-related questions. As such, all three of the fields have very different origins and draw from very different ontologies. Spatial modelling originated from a border field of geography, whereas synthetic biology and gene drives are loosely related to genetics and biology fields. Because of the distinct differences, but also the maturity of both ontologies independently, developing a new ontology that could capture and categorize concepts and the properties of the new interdisciplinary field of spatial modelling of genetic behaviour would most likely not contribute much. Instead of developing new concepts and ontologies, the current work attempts to use existing and well-established concepts and nomenclature used by both respective fields.

In terms of spatial modelling, existing ontologies are well-established and regulated, although the ontologies and vocabulary may differ between different standardization attempts. The most recognized and broadly used ontology and geospatial vocabulary is the one done by the International Organization for Standardization (ISO). Published in 2003,[158] the ISO 19107:2003 standard titled '*Geographic information - Spatial schema*' is the most comprehensive document which describes geographic information and geospatial ontologies. The document specifies concepts for describing the spatial properties and characteristics of geographic features. It defines a set of spatial operations of geographic information for spatial objects and established over all ontologies in the field.

Another broadly recognized attempt to develop vocabulary and ontology for geospatial concepts, especially for web and internet use, was an attempt by the World Wide Web Consortium (W3C). W3C is an international community of member organizations, full-time staff to develop web-based standards. W3C Geospatial Ontologies published a report[159] developed by the W3C Geospatial Incubator Group (GeoXG) in which the group provided an overview and description of geospatial

ontologies to describe geospatial concepts and properties for use on the Worldwide Web. The report also emphasizes the for extending existing geospatial vocabulary, explains the specific geospatial ontologies and the need for their extensions, and identifies existing work from which definitions can be used to develop geospatial ontologies. It is also mentioned in the report that the W3C work might disagree with the previous work done by ISO.

There might be other ontologies and vocabularies related to geospatial modelling and analysis.[160] However, the work aims to apply the broadly accepted ontologies. For more information, please refer to the above-mentioned ISO document.

Ontologies and vocabularies in the fields of biology and genetics are also well-established and very consistent. While certain concepts and definitions from those border ontologies have been transferred and adapted by the synthetic biology and gene drive community, many new concepts, definitions, and vocabularies had to be developed to accommodate specific differences and extensions that these new and immature fields bring. As such, ontologies in synthetic biology and gene drives tend to be inconsistent and often somehow fuzzy. The literature review done for the current study has revealed numerous inconsistencies but also concept and nomenclature evolutions over the years. For example, what is known today as *gene drives* a decade or so ago was broadly referred to as '*selfish genes*,'[6] and '*genetic safety switches*' are replaceable with '*daisy gene drives*.'[23,154] A myriad of other ontological inconsistencies exists in this emerging field.

Because of the relative immaturity of the field, ontologies are either currently developed or initially development has been done, but further work is required to extend and unify these ontologies. One work that deserves a special mention because of its pioneering nature, but also extend to which it covers synthetic biology concepts is so-called Systems Biology Ontology (SBO). SBO is a set of controlled, relational vocabularies of definitions and terms commonly used in the field of systems biology, and in particular, in computational modelling of synthetic biology system. SBO ontology is curated by the European Molecular Biology Laboratory-European Bioinformatics Institute (EMBL-EBI) in the UK as part of the BioModels.net effort.[161]

Other more basic attempts to develop synthetic biology ontologies exist[162] (i.e., Synthetic biology ontology by the team at Imperial College London or the Synthetic Biology Open Language), ISO is yet to develop its standardized document. The vocabulary and concepts related to synthetic biology and gene drives used in the study have followed SBO ontology and other vocabularies generally accepted in the field and broadly used in the scientific literature.

2.7.1 Justification of the work on ontology

Combining several different scientific disciplines comes with discrepancies and differences in their respective ontologies. For example, MGDrivE package uses the word '*landscape*' to describe the study area. The same artefact in spatial modelling would likely be given the name of '*environment*' or '*geospatial environment.*' MGDrivE package was developed by biologists and programmers familiar with gene drive technology. This is why vocabulary related to biology and gene drives is generally correct in the MGDrivE package but not necessarily correct or aligned with ontologies broadly accepted in spatial modelling. This example makes it clear that researchers performing multidisciplinary research are better off borrowing ontologies, vocabularies, and definitions from existing fields, rather than creating their own.

The problem of ontologies is not purely a multi- or interdisciplinary one. Ontologies are organic and evolving datasets where definitions, nomenclature, and vocabularies can evolve, generally as the field matures. This phenomenon is often seen in new fields, where vocabularies and ontologies of immature fields evolve to describe better specific elements. One example of such evolution was clearly shown in the thesis, where back in 2003 when gene drives were discovered or more broadly studied and were initially called '*selfish genes,*' but as the field matured the nomenclature had been adapted to what individuals know today as '*gene drives*' or '*gene drive technology.*' The reason for this ontological shift is, most likely, a trivial one, while the term 'selfish genes' is very descriptive, genes lack personality, and as such, they cannot be selfish. Gene drives without any doubt sound more scientific and professional. Another example from the thesis is '*gene drive safety switch*' and '*daisy gene drives*' which are used interchangeably.

The ontological aim of the thesis was to stick to existing ontologies and borrow from them to develop this multidisciplinary research. Because this thesis relies on the MGDrivE package '*landscape*' nomenclature was copied from the original research, however, the author acknowledges that this nomenclature is wrong. Other than using '*landscape*' same as '*geospatial environment*' the author did not develop or borrowed other ontologies that are considered to be outside of generally accepted standards.

Building clear ontologies is of importance as it helps experts in multidisciplinary research teams to better communicate with each other. In this current research having clear ontologies allows to interface software developers (i.e., MGDriveE) with domain experts.

2.7.2 Attempt to quantify the extent of ontologies used in this study

To compare ontologies, vocabularies, and their similarities, three groups of scientific papers on synthetic biology, spatial modelling, and malaria spread. The online word cloud generator (www.wordclouds.com) was used to import thematic papers and count repeating words. The non-specialist words have then been removed, and ontologies of these three scientific fields compared. Once words and their frequency have been derived from the papers, the words with frequencies lower than three have been removed. Other non-specialist vocabulary, numbers, and surnames have also been removed from the vocabularies. Because of space limitation, Table 1, Table 2, and Table 3 display only the most frequent words from derived vocabulary.

Table 1 *Vocabulary list and word-frequency derived from synthetic biology and gene drive papers.*

Genetic engineering to eradicate invasive mice on islands: modelling the efficiency and ecological impacts (Gregory Backus and Kevin Gross, 2016)[14]		Gene drives thwarted by the emergence of resistant organisms (Ewen Callaway,2017)[156]		Gene drives do not always increase in frequency: from genetic models to risk assessment (Tom J. de Jong, 2017)[3]		Safeguarding gene drive experiments in the laboratory (Omar S. Akbari et al., 2015)[4]		Safeguarding CRISPR-Cas9 gene drives in yeast (James E DiCarlo, 2015)[5]		Site-specific selfish genes as tools for the control and genetic engineering of natural populations (Austin Burt, 2003)[6]		Synthetic biology: discovering new worlds and new words. (Víctor de Lorenzo and Antoine Danchin, 2008)[162]	
155	mice	21	Gene	79	model	33	gene	150	gene	103	gene	51	biology
112	population	17	Drive	51	allele	32	drive	144	drive	66	population	45	biological
87	Rate	9	Genetic	50	fitness	10	Cas9	62	wild-type	37	target	43	synthetic
81	release	9	Mosquito	41	population	10	genetics	34	Cas9	27	species	31	engineering
78	eradication	8	Resistance	38	gene	10	laboratory	33	haploid	26	fitness	21	cell
50	construct	5	Wild	25	drive	10	medicine	26	element	26	sequence	19	system
31	gene	4	Modified	25	homing	10	populations	25	confinement	22	control	17	life
28	Density	4	Mutation	21	frequency	9	confinement	23	plasmid	22	resistant	16	gene
28	model	4	Population	21	model	8	biology	22	diploid	20	engineered	13	circuit
27	ecological	3	Code	20	embryo	7	risk	22	experiment	19	frequency	11	science
25	Species	3	CRISPR	19	genetic	7	sgRNA	21	wild	16	genetic	10	bacteria
24	Genetic	3	gene-drive	16	fixation	7	spread	20	cell	16	selfish	10	component
23	engineered	3	Inheriting	16	heterozygote	6	molecular	20	target	15	generations	10	DNA
23	Invasion	3	Offspring	15	spread	6	science	18	population	14	endonuclease	10	machine
23	wild-type	3	Organism	14	rate	6	strategy	17	genome	14	spread	10	protein
21	Eradicate	3	Spread	13	gametes	6	stringent	17	molecular	13	construct	9	genetic
20	impact			12	CRISPR-Cas9	5	drosophila	17	synthetic	13	DNA	9	regulatory
19	rate			11	life	5	experiment	16	CRISPR-Cas9	13	drive	8	molecular
18	Dynamics			10	species	5	release	16	organism	12	allele	7	artificial
14	control			10	threshold	5	wild	15	colony	11	host	7	molecular
14	Musculus			9	offspring	3	biosafety	15	homing	10	chromosome	7	ribosome
12	spread			9	wild-type	3	cell	15	locus	10	element	7	RNA
11	capacity			8	cell	3	DNA	14	mating	10	homing	6	construct
10	drive			8	homozygous	3	encoding	14	strain	10	mutation	6	evolution
10	fitness			8	meiosis	3	genome	11	DNA	10	recessive	6	metabolic
10	offspring			8	organism	3	health	11	gRNA	9	natural	5	mRNA
10	population			8	recessive	3	organism	11	inheritance	9	release	5	network
9	ecosystem			8	target	3	probability	11	RNA	8	promoter	5	paleome
9	non-target			7	dominance	3	RNA-guided	10	sequence	8	protein	5	polymerase

22

Table 2 *Vocabulary list and phrase-frequency derived from malaria spread studies.*

Acquired immunity to Malaria (Denise L. Doolan et al., 2009)[139]		Antimalarial drug resistance (Nicholas J. White, 2004)[129]		Assessing the risk factors associated with Malaria in the highlands of Ethiopia: What do we need to know? (Élodie Anne Vajda and Cameron Ewart Webb, 2017)[81]		The demographics of human and malaria movement and migration patterns in East Africa (Deepa K Pindolia et al., 2013)[79]		Effect of climate on malarial vector distribution in Monsoon Asia (Shunji Ohta and Takumi Kaga, 2012)[35]		Genetic factors in malaria (L. Luzzatto, 1974)[135]	
308	malaria	152	resistance	229	malaria	135	malaria	73	malaria	40	malaria
227	immunity	94	malaria	89	transmission	75	movement	40	model	29	parasite
185	falciparum	79	parasite	29	population	47	travel	34	distribution	23	genetic
178	plasmodium	77	drug	29	malaria	41	migrant	34	climate	23	cell
90	exposure	73	antimalarial	28	climate	35	migration	33	anopheles	21	gene
85	infection	46	falciparum	26	vector	34	transmission	30	data	19	resistance
80	protection	46	parasite	22	health	33	demographic	24	vector	18	plasmodium
78	disease	44	infection	21	vector	30	network	22	mosquito	18	host
70	parasite	37	treatment	19	resistance	28	movement	20	health	17	gene
69	clinical	28	infection	19	outbreak	28	pattern	20	water	14	infection
68	sporozoite	24	resistant	19	mosquito	26	flow	19	species	14	blood
61	parasite	24	spread	18	falciparum	25	connectivity	17	generation	13	falciparum
60	parasitemia	23	probability	17	malar	23	stratified	13	population	11	susceptibility
55	immune	20	combination	17	vivax	23	malaria	13	condition	11	erythrocyte
52	transmission	20	novo	16	anopheles	23	net	12	incidence	9	mechanism
50	strain	19	transmission	16	infect	22	gender	11	observe	8	parasitaemia
49	infection	19	mefloquine	14	plasmodium	20	source	11	factor	8	parasitology
45	infect	19	plasmodium	13	distribution	20	group	10	temperature	8	genotype
44	heterologous	18	mutation	11	arabiensis	17	endemicity	10	vector	8	vitro
43	antigenic	17	chloroquine	11	dynamics	17	control	9	temporal	7	tropical
41	protective	16	elimination	11	species	16	population	9	spatial	7	medicine
39	vivax	16	gene	10	intervention	15	infection	9	period	7	factor
36	homologous	15	concentration	10	incidence	15	sink	8	variation	7	cycle
36	parasitol	14	antimalarial	9	chloroquine	13	elimination	8	habitat	7	data
36	knowlesi	13	multiplication	9	elimination	13	falciparum	8	risk	7	DNA
34	strain	12	artemisinin	9	epidemic	13	map	7	parasite	6	G6PD-deficient
33	vaccine	12	symptomatic	9	disease	11	plasmodium	6	variation	6	population
32	immunization	11	clinical	8	entomological	11	dataset	6	simulate	6	hypothesis
32	host	11	patient	7	surveillance	10	parasite	6	emergence	6	frequency

Table 3 *Vocabulary list and word-frequency derived from spatial modelling studies.*

Deforestation and malaria on the Amazon frontier (Fábio S. M. Barros and Nildimar A. Honório, 2015)[61]		Gene drive through a landscape (Andrea Beaghton et al., 2016)[24]		Impact of climate change on global malaria distribution (Cyril Caminade et al., 2014)[66]		The links between agriculture, Anopheles mosquitoes, and malaria risk in children younger than 5 years in the Democratic … (Mark M Janko et al., 2018)[91]		Diversity in breeding sites and distribution of Anopheles mosquitoes in selected urban areas of southern Ghana (Precious A. Dzorgbe Mattah et al., 2017)[57]		Spatial change in the risks of Plasmodium vivax and Plasmodium falciparum malaria in China, 2005-2014 (Samuel Hundessa, 2016)[55]		Effects of habitat fragmentation on birds and mammals in landscapes with different proportions of suitable habitat (Henrik Andrén, 1994)[163]	
183	forest	140	population	88	climate	42	agriculture	130	habitat	33	change	213	habitat
178	malaria	134	wave	65	model	36	probability	70	site	28	latitude	133	landscape
92	water	99	model	55	transmission	25	Data	54	urban	28	country	89	patch
85	deforestation	94	speed	40	change	21	Mosquito	39	permanent	27	longitude	86	size
69	area	64	state	38	population	21	Land	33	sample	23	study	83	species
60	collection	63	equilibrium	28	distribution	20	Study	33	habitat	22	period	60	fragmentation
52	transitional	55	density	27	global	19	population	32	city	18	area	54	forest
47	transmission	40	rate	20	impact	19	Site	32	area	13	incidence	53	isolation
46	distance	38	asymptotic	19	scenario	19	Cover	27	study	13	parasitol	45	random
46	stream	38	spread	19	region	17	relationship	25	metropolitan	12	climate	43	population
42	fringe	37	release	19	change	17	Estimate	25	farm	12	control	38	proportion
40	hotspot	32	dispersion	17	data	16	transmission	23	temporary	11	transmission	35	suitable
37	site	28	spatial	16	endemic	15	temperature	21	density	11	temperature	35	area
35	area	28	system	15	observe	15	increased	20	mean	11	model	34	original
33	pond	26	density-dependent	15	northern	14	Indoor	19	distribution	11	rate	34	density
30	cluster	26	travel	13	estimate	14	Effect	19	point	10	province	34	effect
29	study	26	case	13	epidemic	13	surveillance	18	species	10	analyse	32	hypothesis
27	location	24	landscape	13	factor	13	associated	16	observe	9	case	30	sample
26	sample	24	stable	12	endemicity	13	Model	13	construction	8	affected	23	relationship
25	proximity	23	relation	12	control	11	Survey	13	place	8	country	20	diversity
24	risk	23	solution	12	health	10	Rural	12	variation	8	trend	19	community
23	habitat	23	immobile	10	temperature	9	abundance	12	natural	8	data	17	dynamics
23	settler	23	species	10	simulate	9	expansion	12	count	7	situation	17	pattern
23	cluster	21	parameter	10	suitable	9	Variable	12	data	7	expansion	16	scale
22	frontier	21	life	10	impact	8	entomological	11	transmission	7	spatial	14	range
21	model	19	diffusivity	9	socioeconomic	8	environmental	11	proportion	6	regression	13	metapopulation
21	data	19	unstable	9	uncertainty	8	Analyse	10	population	6	indicate	13	biogeography
19	land	18	extinction	9	condition	8	Factor	8	combination	6	mainland	13	individuals
18	analyse	17	factor	8	projection	8	Spatial	7	environmental	6	system	13	population

3. Prior State-of-the-Art

Computational modelling and simulation of both spatial mosquitos spread to evaluate malaria risk and gene drive modelling to estimate the population replacement or suppression have become fast, cost-effective, and often reliable methodology to determine how different system parameters affect the outcomes. Performing similar field studies would in most cases be expensive, very time consuming, and often ethically wrong. It could also pose severe threats to biodiversity and human health. As such, multiple models have been developed to evaluate these systems.

3.1 Spatial Modelling of Malaria Spread

Mathematical models of mosquito-borne pathogen transmission originated in the early twentieth century.[164] The primary role of these models was to establish the most effective ways to combat or mitigate malaria. Since 1970, growing demand in reducing vector-borne disease has put a lot of emphasis on mathematical modelling and simulation of different case scenarios to guide the control measure and often policy developments. In a study published in 2013 by Reiner et al., the authors assess how theory and mathematical tools have changed to address evolving public health challenges related to malaria.[164] The authors compiled a list of 325 publications from 1970 through 2010 including mathematical models of mosquito-borne pathogen spread and subsequently developed a 79 item-long questionnaire to classify each of 388 identified mathematical models in accordance with their purpose. The study concluded that although geographical, epidemiological, and ecological scope, extent, and complexity have changed significantly over the past 40 years, the majority of the models resemble the original, over a century old, Ross-Macdonald approach.[165,166] Ronald Ross and George Macdonald are credited with developing a first mathematical model of mosquito-borne pathogen transmission. Moreover, authors of the study argue that the modern theory would benefit from an expansion of the concepts such as variation of individual host attributes and their consequences for heterogeneous mosquito biting, spatial heterogeneity and temporal variation in the transmission process, as well as poorly mixed mosquito-host encounters. Reiner et al., acknowledge that these concepts have been at times successfully addressed, they are not widely used or appreciated. Reiner and co-authors believe that emphasizing more on these efforts would lead to more accurate, robust, and useful tools for addressing global health challenges posed by vector-borne diseases. The comprehensive study provides a complete list and further arguments for or against identified modelling frameworks.

3.2 Modelling of Gene Drive Inheritance

Gene drive models focus predominantly on three major scenarios: replacement of the wild population with gene drive-engineered organisms, suppression of the wild population by introduction of gene drive organisms or reversibility of gene drive effects to the wild population.[151,154,167,168]

MGDrivE framework is described in more detail in later chapters. Some other common approaches to modelling gene drives include a well-mixed finite population model, where gene drives are modelled in finite populations, often by introducing a Moran-type model with sexual reproduction. The finite population of diploid individuals is created and simulated, generally focusing on three allelic classes wild (W), gene drive element (H) and drive-resistant (R), and possible genotypes of WW, WD, WR, DD, DR, and RR. Each genotype has assigned a reproductive rate, and the

simulation process occurs in discrete time-steps.[151] This modelling framework is commonly applied thanks to its simplicity and computational flexibility.

Another often applied modelling scheme for gene drive systems is the so-called finite population model with population structure, which allows examining the effects of population structure on drive containment. The framework is an extension of the well-mixed, which additionally considers the well-mixed subpopulations of individuals. Just like the well-mixed finite population model, finite population modelling is performed in discrete time steps. In each iteration, individuals are either migrated between subpopulations or subpopulation is chosen to proceed through mating and replacement procedure.[151,167]

Among multiple other modelling frameworks, one that deserves mention is Medusa.[169] Medusa model relies on a stochastic discrete-time framework, which, due to its complexity, allows simulating complex scenarios while maintaining a high degree of flexibility of the framework. Medusa was developed by the Marshall team at the University of California, Berkeley, the same team that developed MGDrivE.

Further comparison of existing models for the spatial distribution of gene drive organisms is presented in Table 4.

Table 4 *Comparison of spatially-explicit gene drive models. Reproduced from Reproduced from bioRxiv.*[168]

	Inheritance Patterns	Life History Ecology	Spatial and landscape details	Software
MGDrivE	Very flexible, can be user-specified	Egg-larva-pupa-adult, density-dependence at larval stage, not responsive to environmental variables at present	Metapopulations distributed in space, connected by migration	R package, open source
EMOD	Homing-based gene drive could be extended with effort	Egg-larva-pupa-adult, density-dependence at larval stage, responsive to environmental variables	Populations arranged on a grid, each representing 1 km^2, connected by migration	Java Script Open Notation (JSON) feeds into executable file, open source
Skeeter Buster	Homing-based gene drive, release of insects carrying a conditional lethal, etc., cannot be user-specified	Egg-larva-pupa-adult, density-dependence at larval stage, responsive to environmental variables	Households and containers modelled explicitly, connected by migration	Executable file, not open source
SLiM	Very flexible, can be user-specified	Discrete generations, no life history at present	Can model either connected metapopulations or cells on a grid	Scripting environment with graphical user interface, open source

4. Research Methods

Since the current research relies strictly on existing model adaptation and simulation, four-step system development has been applied. The four steps include system specification methods, system design methods, system implementation methods, and system evaluation methods. A simplified study flowchart is shown in Figure 4. Map of the study area is presented in Figure 5.

STEP 1

Selecting input variables and data (topographic images)

- Collection of viable intial parameter data based on existing literature

- Selecting appropriate study map from online resources

STEP 2

Model development and optimization

- Model developement and adaptation
- Model optimization and troouble shooting

STEP 3

Selecting scenarios for simulation and output data evaluation

- Scenario simulations and data collection

- Evaluation of the simulation outcomes

STEP 4

Data visulaization and presentation, discussion and research conclusions

- Data visualization

- Discussion, conclussions and future work

Figure 4 *Simplified study development flowchart.*

Figure 5 *Map of the study area – Kenya.*

The thesis and corresponding simulations rely on the assumption that the gene drive-modified mosquitos will have no significant competitive advantage or disadvantage over their malaria-carrying kin. This assumption has been made for two reasons, first at the current stage, gene drive studies have shown no significant competitive advantage or disadvantage, and second, the assumption is necessary to simplify the analysis.[20,22,33] The assumption is not going to be testable with data until some of these gene drive-modified organisms live and die in the environment. The researcher cannot know, a priori, whether the modification will produce a more or less hardy mosquito in the Kenyan milieu, even if the intent is to create one where the modification has no significant effect in this regard. Likewise, the researcher cannot

know whether or not deleting the symbiotic relationship (between the parasite and host) is going to be somehow subtly deleterious to the host in that environment. Presumably, any small effect of the sort, for either reason or in either direction, could have a significant impact on the results. Attempt to model such an advantage or disadvantage could be the focus of a separate study.

4.1 MGDrivE Model Description

The MGDrivE (Mosquito Gene Drive Explorer) gene drive inheritance simulation framework was used to simulate genetic inheritance within the mosquito population. The framework was developed by John Marshall's laboratory at the University of California, Berkeley.[168] The MGDrivE framework is different from other gene drive modelling frameworks because of its ability to simulate a wide array of user-specified inheritance-modifying systems at the population level within a single, computationally efficient framework that also incorporates mosquito life history and landscape ecology. MGDrivE framework was used to build a simulation model as described in the original publication.[168]

4.2 Working Principle of the Model

The overarching working principle of the MGDrivE model is to consider the inheritance matrix of genotypes as a three-dimensional structure, in which each single intersection point in the matrix determines the probability of an offspring genotype (z-axis) provided a certain combination of male-to-female genotypes (x-axis and y-axis respectively). The arrangement enables using tensors for the calculations, at the same time offering a number of advantages, including model scalability, excellent computation speed, and model transparency.

Second, the further novelty of the framework relies on consideration of the spatial layout as a network of interconnecting breeding habitats, where each habitat is considered as a network node. By doing so, the landscapes and topographic maps are transformed into distance matrices and subsequently transformed into transition probability matrices. Consequently, the novel framework enables modelling of arbitrary topologies, where mosquito populations, their mating, and migration can be simulated in realistic geographical settings. The entire model will rely on R working environment and associated toolboxes.

The MGDrivE framework consists of three modules, namely inheritance module, life history module, and landscape module. The genetic inheritance module is a fundamental module of the MGDrivE framework. In MGDrivE, genetic inheritance is expressed by a three-dimensional tensor called "inheritance cube." In the inheritance cube, the first and second dimensions refer to the maternal and paternal genotypes, respectively, whereas the third dimension corresponds to the offspring genotype. The inheritance cube variables correspond to the proportion of offspring with the combination of all possible parental genotypes. This combination matrix excludes offspring fitness and viability, and all the values in the matrix should sum up to one.

The mosquito life history module is based on previous studies by Hancock and Godfray adapted by Deredec et al.[170,171] In this model, the mosquito life cycle is divided into four distinct stages: egg, larva, pupa, and adult. Besides, adult mosquitoes can be either male or female: the life history module is presented in more detail in Figure 6. Further information about the module can be found in the original study.[168]
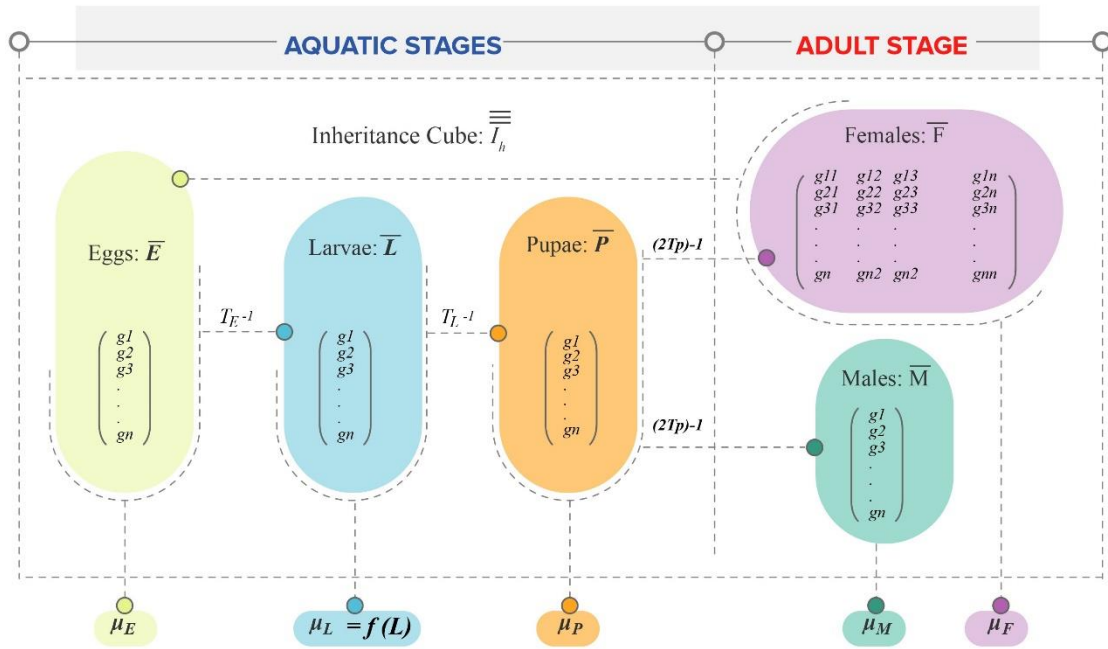
Figure 6 *Schematics of mosquito life history module from MGDrivE framework showing aquatic states, including eggs, larvae, and pupae as well as adult mosquito stage for males and females. Parameters associated with each of the model adaptation stage are shown in the figure. Figure was reproduced from bioRxiv.[168] For more information about the module refer to the original study.[168]*

Finally, the landscape module is responsible for the distribution of mosquito metapopulations in selected space, and for calculating the movement through the resulting network determined by calculating the mosquito dispersal kernels. The framework assumes that the metapopulations of mosquitoes are mixing randomly, based on the lumped age-class model equations. The MGDrivE simulation workflow is shown below.
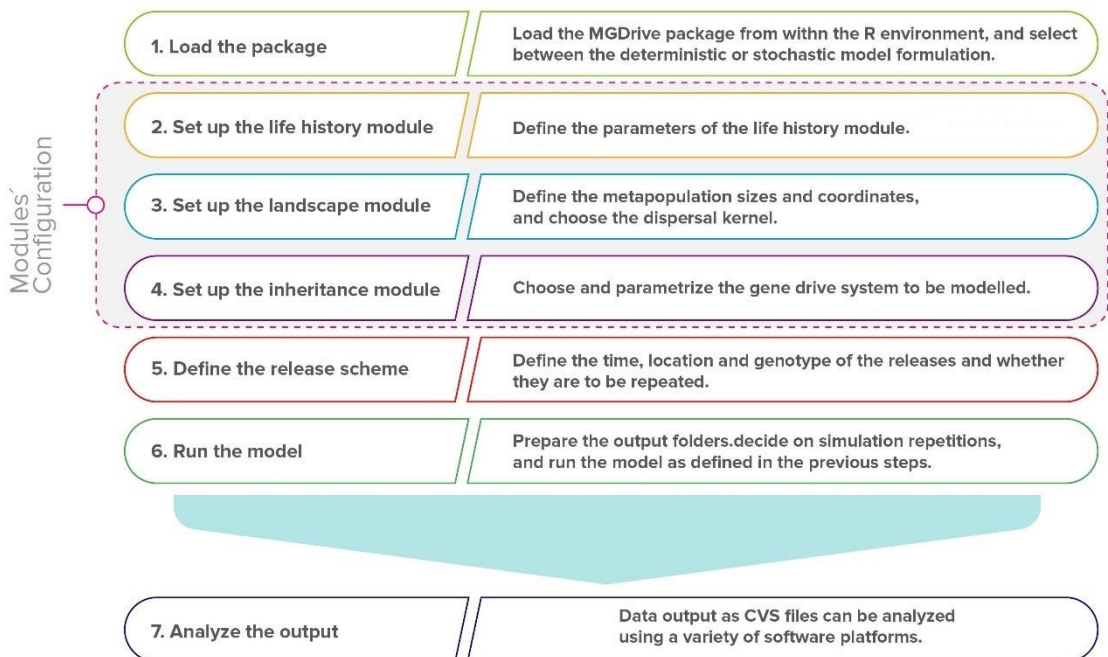
Figure 7 *The MGDrivE simulation workflow. Reproduced from Sánchez et al.[168]*

4.3 Semi-spatial Nature of the MGDrivE Environment

While the study often refers to the distribution as a spacial distribution, the MGDrivE package is not entirely spatial. Because of computational power limitations and time required to compute large populations of mosquitoes over extensive time intervals, often counted in years, the authors of the package have introduced certain assumptions and limitations, making MGDrivE semi-spatial environment. While this is not desired, modelling and simulation of large populations of mosquitoes would otherwise be impossible.

Instead of modelling each mosquito as an independent node with its geographic coordinates, MGDrivE models larger groups of mosquitoes and treats these groups as separate nodes with the assigned number of mosquitoes in it. Using the metapopulation concept where the entire population is divided into smaller sets or subpopulations, each consisting of a certain number of mosquitoes is a significant simplification, which allows performing simulations involving a large number of mosquitoes. Each node in MGDrivE interacts with other nodes, creating an interconnected network. Creating the same network having millions of nodes would otherwise be impossible or at least very difficult.

As such, the nature of the MGDrivE package is not truly spatial, but because it contains significant aspects of a spatial system with certain limitations, MGDrivE can be considered a semi-spatial environment. With the increase in computing capabilities, the MGDrivE package can easily be extended to a fully spatial environment.

4.4 Methodology and Steps to Answer Research Questions

The MGDrivE is a model designed to be a reliable testbed for various gene drive scenarios for mosquito-borne disease control. The model has been developed to accommodate the use of different mosquito-specific gene drive systems within a population dynamics framework to allow simulation migration of individuals between nodes in the selected study landscape.

In the study, major cities in Kenya have been modelled as nodes with different population sizes on the nodes of the studied network (Figure 8). The geographic data, including latitude and longitude, as well as a population of 36 largest cities in Kenya, have been obtained from the World Cities Database, released for free under an MIT license. City data in this dataset comes from the National Geospatial-Intelligence Agency, whereas the population data comes from Natural Earth Data.
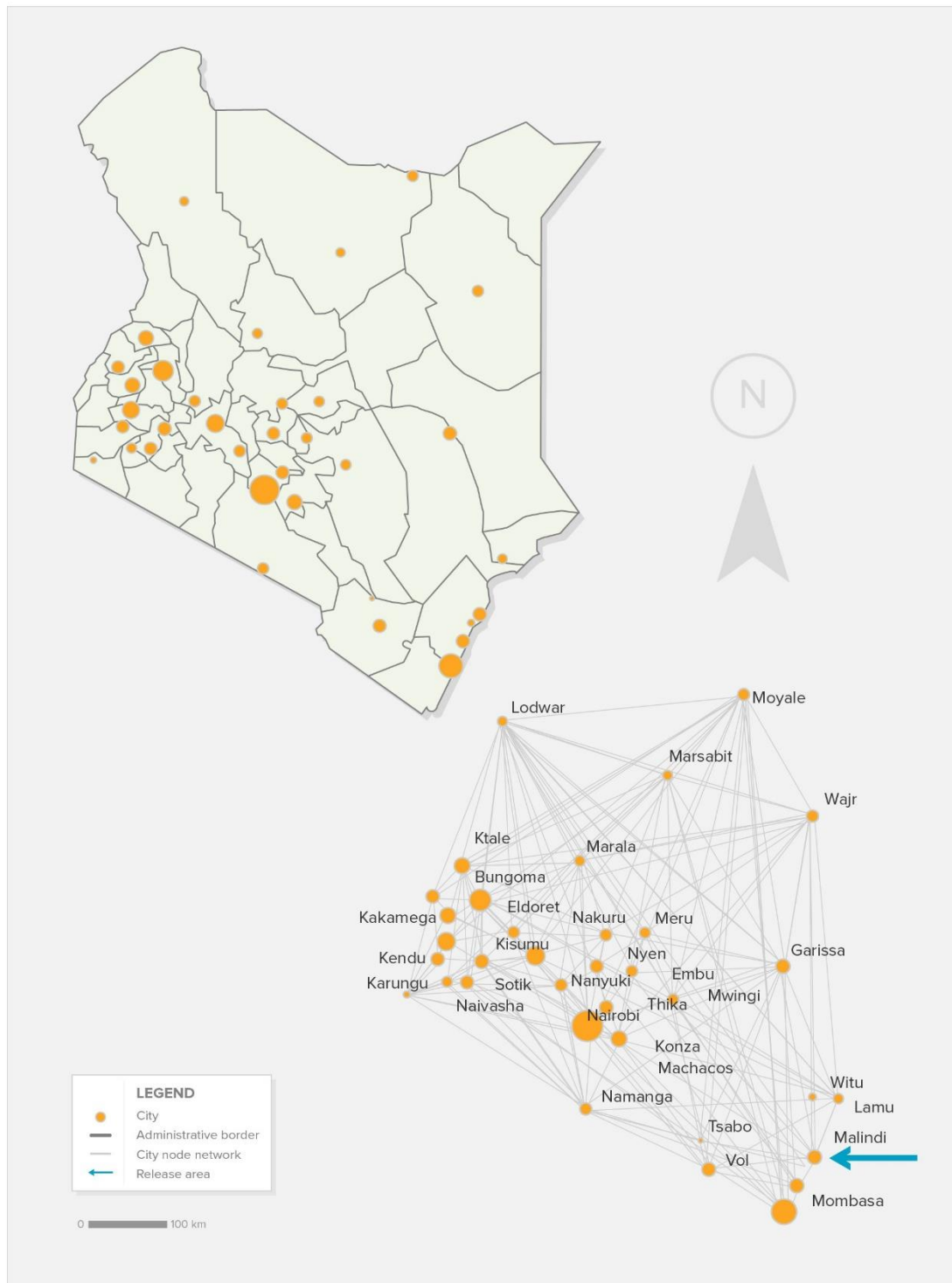
Figure 8 *Kenya map with administrative borders and 36 cities used in the study (top). Cities used in the study together with a network of connections between cities/nodes (bottom). The size of the node is scaled based on the existing human population, which is modelled one-to-one with initial mosquito population. The size of the node corresponds to the cubic root of the total human population. The release site (Malindi) is indicated by the blue arrow.*

The simulated scenarios assume that the study area is a closed system, meaning that no mosquitoes can cross the Kenya border. This assumption is far from real-life

scenarios where mosquitoes do not recognize national borders or international protocols and freely travel between the countries. However, simulating such a scenario is computationally not feasible because of the scale of such a study and the availability of computational power.

Latitude and longitude data have been used to create an Excel file with input patches, and to calculate the distance matrix in Excel as another input file. The population data were used to emulate the initial mosquito population at those nodes, and the mosquito population at each location was scaled proportionally to the number of humans present in this location.

Simulation time has been set to 3,650 days, which corresponds to 10 years. However, in some cases the simulation time was extended to a number of days necessary to replace the entire wild mosquito population. Each scenario was repeated only once because having 36 patches and 3,650 days already takes about one hour per scenario to compute and create output data. Mosquito lifespan probability parameter has been set up to 0.9 for most of the scenario, whereas other mosquito biology-related parameters, such as duration of egg stage (*tEgg*), duration of larval stage (*tLarva*), duration of pupa stage (*tPupa*), daily population growth rate (*popGrowth*) and daily mortality risk of adult stage (*muAd*) were 1 day, 13 days, 1 day, 1.096 days[-1] and 0.123 days[-1] respectively. Egg production per female (*betaK*) was 32 days[-1]. These parameters have been taken from Table 5, reproduced from the original MGDrivE publication.[168]

Table 5 *Life history module parameter values for three species of interest (at a temperature of 25 Celsius). Reproduced from Sánchez et al.[168]*

Parameter	Anopheles gambiae	Aedes aegypti	Ceratitis capitata
Egg production per female (day[-1])	32	20	20
Duration of egg stage (days)	1	5	2
Duration of larval stage (days)	13	6	6
Duration of pupa stage (days)	1	4	10
Daily population growth rate (day[-1])	1.096	1.175	1.031
Daily mortality risk of adult stage (day[-1])	0.123	0.090	0.100

Gene drive setup parameters of male homing rate (*eM*) and female homing rate (*eF*) were set up to 0.9 and 0.5 respectively unless otherwise specified. While parameters for *Anopheles gambiae* have been used in this current study, parameters for two other mosquitoes are shown for comparison only.

The gene drive-modified mosquito release was performed on day 100 from the beginning of the simulation and was performed only once during the simulation cycle. The release on day 100 from the beginning of the simulation is reasoned by the fact time is initially allowed for simulation to stabilise and also to be able to illustrate how release of the gene drive mosquitoes influences the shape of the graph. The release proportion of wild to modified mosquitoes was 1% unless otherwise specified. The scenarios assume that gene drive-modified mosquitoes have been released at one node only, which for the purpose of this exercise is a town of Malindi in south-eastern Kenya.

4.5 Data

Since most of the work relied on existing models from literature, which was developed based on data collected by other authors. The only data that was used in the initial input data was taken from literature. Certain model parameters can be selected

randomly and tested to evaluate a specific scenario. The data processing and plotting were performed in Microsoft Excel.

Other than the above data that were in the study area, which in this case was the map of Kenya, and the database of major cities with their coordinates and population in Kenya, the map was obtained from resources available on the Internet, whereas the coordinate and population database was obtained from the World Cities Database, released for free under an MIT license. The latitude, longitude, and population data used to create nodes is available in Table B1 in Appendix B.

5. Results

The series of output plots generated based on different initial parameters described in the methods section, allowed to derive answers to research questions. Different initial parameters were used to separately answer four supporting questions and based on that the answer to the main research question was derived.

To answer support question "*Based on different initial numbers of gene drive mosquitoes introduced to wild population how many days will be required to eradicate malaria?*", the simulation was rerun several times with the same initial parameters except for the number of gene drive mosquitoes released. This parameter was adjusted each time according to Table 6, and a number of days required to completely replace wild populations with gene drive altered population of mosquitoes incapable of vectoring malaria were recorded. Because of the extensive time needed to complete each simulation, one set of parameters was simulated only once. It could be of benefit to rerun these simulations more time times and calculate the average.

Table 6 *Values of the initial number of gene drive mosquitoes parameter to answer the first support question "Based on different initial numbers of gene drive mosquitoes introduced to wild population, how many days will be required to eradicate malaria?" The initial number of gene drive mosquitoes as a percentage of wild mosquito population in Malindi.*

Simulation number	The initial number of gene drive mosquitoes (as a percentage of wild population in Malindi)	Number of days required to replace wild mosquitoes with gene drive mosquitoes
1	1	4,009
2	2	3,972
3	3	3,813
4	5	3,638
5	7	3,561
6	10	3,479
7	12	3,431
8	15	3,389
9	20	3,326
10	25	3,271

Based on the results, it becomes clear that to achieve complete replacement of wild mosquito population, at least 5% of gene drive mosquitoes have to be released with their inheritance frequency (*eM*) of 90%. Any number higher than 5% of the gene drive-altered mosquitoes introduced to the wild population at the specific release point, in this case Malindi, will slightly reduce the time required for complete replacement. The above table is visualized below in the form of a graph.
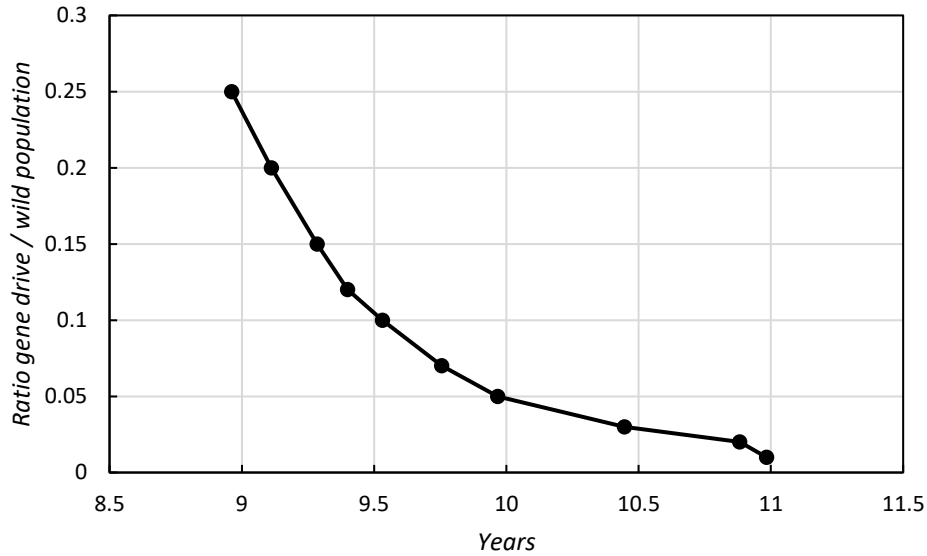
Figure 9 *Relationship between the ratio of introduced gene drive organisms and wild population to a number of years to entirely replace the wild population. The traits for each of the simulations that contributed to this figure are available in Appendix C.*

The graph above shows highly non-linear behaviour. It also becomes clear that the number of gene drive-altered mosquitoes to the size of the wild population is not critical. Changing the parameter from one up to 25% reduces the time required for complete replacement from nearly 11 years to only less than nine years. The number of days required to replace the wild population for the ratio of 1% behaves out of the pattern. The simulation for this point was rerun three times, giving similar results. It is unclear why this behaviour occurs.

The second supporting question ("*How different inheritance frequencies for gene drive mosquitoes introduced to wild population affect the number of days required to eradicate malaria?*" was answered in a similar way as the previous question. All parameters were fixed except for the male mosquito inheritance ratios, which was changed for each rerun. Generally, gene drives are designed to have 100% or close to 100% inheritance efficiency, as such, the numbers were selected to emulate this. It is very unlikely that the gene drive system will have a frequency of 100%. However, this value was also explored in the simulations. The male inheritance values used to answer the second supporting research question are shown in Table 13.

The effective numbers of inheritance frequency of gene drive altered mosquitoes does not exist. This is mainly because of the immaturity of the field but also due to the ethical and legal mechanisms that prevent the release of these organisms into the wild. All of the current knowledge of these systems comes from observations of limited and closed laboratory populations or computational models developed based on the outcomes of these limited experimental data. However, based on the general understanding of the genetics, the higher the inheritance frequency is, the faster the spatial replacement of the wild populations by gene drive altered ones will be.

36

Table 7 *Initial values of the initial inheritance ratio parameter to answer the second research support question "How different inheritance frequencies for gene drive mosquitoes introduced to wild population affect the number of days required to eradicate malaria?"*

Simulation number	Initial inheritance frequency / chance (percent) – male (*eM*)	Number of days required to replace wild mosquitoes with gene drive mosquitoes
1	100	2,297
2	99	2,376
3	97	2,552
4	95	2,758
5	93	3,004
6	91	3,303
7	90	3,479
8	89	3,664
9	87	4,152
10	85	4,811

Similarly to the above, Table 7 was converted into a graph to better visualize the results and to draw conclusions.
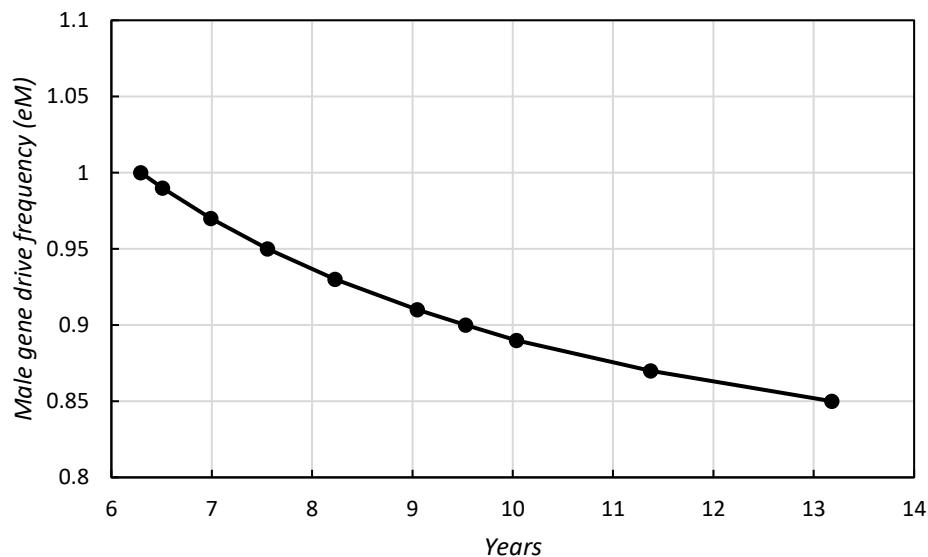
Figure 10 *Relationship between gene drive frequency in male mosquitoes to a number of years required to entirely replace the wild population. The traits for each of the simulations that contributed to this figure are available in Appendix D.*

Based on Table 7 and corresponding Figure 10, it becomes clear that the inheritance frequency factor has a significant influence on the system. A small adjustment in the male inheritance frequency results in a considerable time change required to replace the wild population. The inheritance frequency for male mosquitoes was changed from 85 to 100% resulting in time extension from slightly over six years to more than 13 years. The change of this parameter by 15% resulted in doubling the time. The relation is also non-linear.

The third support question ("*What is the coverage of the gene drive mosquitoes and do they spread within the entire study area?*") was be answered based on the output

maps from two previous exercises. Each node, which corresponds to a specific city in Kenya, generated its graphs. While these are not presented in the study, due to extensive space required to illustrate all 36 cities for each of the scenarios, the superposition of all these nodes resulted in general graphs for Kenya, which shows that spatial distribution throughout the study area after 10 years is uniform. These graphs are presented in Appendix C and Appendix D.

Finally, the fourth support question ("*Are gene drives effective tool to completely eradicate mosquito-borne malaria in Kenya the next 10 years?*") is a result of the previous exercises. In this case, all studied nodes/cities in Kenya were considered collectively. Depending on the value of the parameters replacing the wild mosquito population, it can be either doable or not. Generally, a high number of mosquitoes and high inheritance frequencies enable the reduction of the replacement time. However, the graphs also show the emergence of a significant number of genes drive resistant mosquitoes. These mosquitoes developed resistance to the gene drive system, and as such are not affected. However, their vectoring capabilities are still present. Based on this, complete eradication of malaria in Kenya within a decade using gene drive solution solely is not feasible. Gene drives could be used together with other, more traditional malaria prevention techniques.

Based on the conclusions from the above support questions, it becomes clear that gene drives are not capable of eradicating malaria, due to evolved gene drive resistance that occurs during the population replacement. These might be ways to reduce the effect of the gene drive resistance. This, however, would require further studies. As such, gene drive technology brings a grand promise for reducing the burden of malaria.

5.1 Fulfillment of Assumed Expectations

The last part of the discussion chapter reverts to the assumed expectations and checks whether they are fulfilled. Starting from the supporting question hypotheses, hypothesis for supporting question: *Based on different initial numbers of gene drive mosquitoes introduced to the wild population, how many days will be required to eradicate malaria?* stated that the larger the number of gene drive mosquitoes relative to number of mosquitoes in wild population, the faster gene drive organisms could replace wild mosquitoes. This relationship is already shown in Figure 9 and Table 6, but Spearman's rank coefficient test[172] will be used to further prove it. The expectation for the supporting question is that the larger number of gene drive mosquitoes relative to the number of mosquitoes in the wild population, the faster the gene drive organisms can replace wild mosquitoes. Hypothesis for this supporting question can be written as:

H_0: *There is a no relationship between a number of gene drive mosquitoes to a number of mosquitoes in the wild population and the number of days to replace the entire wild population.*

H_A: *There is a relationship between a number of gene drive mosquitoes to a number of mosquitoes in the wild population and the number of days to replace the entire wild population.*

Table 8 *Spearman's rank correlation test for the first supporting question.*

Simulation number	The initial number of gene drive mosquitoes (as a percentage of wild population)	Rank	Number of days required to replace wild mosquitoes with gene drive mosquitoes	Rank	Difference between ranks (*d*)
1	1	10	4,009	1	9
2	2	9	3,972	2	7
3	3	8	3,813	3	5
4	5	7	3,638	4	3
5	7	6	3,561	5	1
6	10	5	3,479	6	-1
7	12	4	3,431	7	-3
8	15	3	3,389	8	-5
9	20	2	3,326	9	-7
10	25	1	3,271	10	-9

The squared difference between ranks (d^2) equals 330 and Spearman's correlation coefficient (R_s) equals -1. This means that a perfect negative correlation occurs between two datasets. To test for a significant association and to reject H_0, $\sqrt{N-1}(R_s) > z_\alpha$. The z_α was taken from Ramsey[173] and equals 0.903 for quantiles of 0.9995. The test for a significant association has failed (-3 < 0.903), so we fail to reject H_0.

The expectation for the second supporting question was that the higher inheritance frequencies for gene drives would lead to a reduction in the number of days required to replace wild mosquitoes. The hypothesis can be expressed as:

H_0: There is a no relationship between inheritance frequencies and a number of days to replace the entire wild population.

H_A: There is a relationship between inheritance frequencies and a number of days to replace the entire wild population.

Spearman's rank correlation test will be used to accept or reject H_0.

Table 9 *Spearman's rank correlation test for the second supporting question.*

Simulation number	Initial inheritance frequency / chance (percent) – male (*eM*)	Rank	Number of days required to replace wild mosquitoes with gene drive mosquitoes	Rank	Difference between ranks (*d*)
1	1	1	2,297	10	-9
2	0.99	2	2,376	9	-7
3	0.97	3	2,552	8	-5
4	0.95	4	2,758	7	-3
5	0.93	5	3,004	6	-1
6	0.91	6	3,303	5	1
7	0.9	7	3,479	4	3
8	0.89	8	3,664	3	5
9	0.87	9	4,152	2	7
10	0.85	10	4,811	1	9

The squared difference between ranks (d^2) once again equals 330 and Spearman's correlation coefficient (R_s) equals -1. Meaning that a perfect negative correlation occurs between two datasets.

The third supporting expectation hypothesis stated that in the current semi-spatial model, where there is no landscape, weather, or other limitations, gene drive mosquitoes should be able to spread within the entire study within a time of fewer than five years. Answering the question requires looking at mosquito distribution at each of the cities in Kenya. The hypothesis can be written as:

H_0: *Gene drive mosquitoes require five years or less (≤ 5) to spread and be present throughout the entire study area.*

H_A: *Gene drive mosquitoes require more than five years (>5) to spread and be present throughout the entire study area.*

We fail to reject the null hypothesis and conclude that gene drive altered mosquitoes requires more than five years (>5) to be present within the entire study area consisting of 36 Kenyan cities considered in this study. However, the hypothesis is highly sensitive to initial conditions. For example, very low initial gene drive mosquito number introduced into wild population and very low gene drive inheritance frequency will result in the spread of gene drive mosquitoes within the entire study area, but this will take more than five years. For example, one gene drive mosquito with an inheritance frequency of 56% will result in gene drive mosquitoes in each of the cities but only after 3,729 days. To statistically check the assumed expectation, simulation results have been combined into one dataset (Table 10) and a number of years required for gene drive mosquitoes to be present in all Kenyan cities was derived.

Table 10 Simulation results showing the number of years required for gene drive mosquitoes to be present in all Kenyan cities. The initial number of gene drive mosquitoes refers to a percentage of gene drive mosquitoes to wild mosquitoes in Malindi. The initial number of gene drive mosquitoes has been altered from 1 to 25 percent while keeping the initial inheritance number at a constant rate of 0.9. Initial inheritance frequency was then altered from 0.85 to 1 while keeping the initial percentage of gene drive mosquitoes unchanged (1 percent). Both experiments have been merged to create the table below.

Simulation number	The initial number of gene drive mosquitoes (as a percentage of wild population in Malindi)	Initial inheritance frequency / chance (percent) – male (eM)	Number of years required for gene drive spread throughout the study area	Gene drive spread within up to five years? (TRUE/FALSE)
1	1	0.9	2.37	TRUE
2	1	1	1.36	TRUE
3	1	0.99	1.38	TRUE
4	1	0.97	1.48	TRUE
5	1	0.95	1.62	TRUE
6	1	0.93	1.74	TRUE
7	1	0.91	1.92	TRUE
8	1	0.9	2.02	TRUE
9	1	0.89	2.15	TRUE
10	1	0.87	2.42	TRUE
11	1	0.85	2.81	TRUE
12	2	0.9	2.31	TRUE
13	3	0.9	2.22	TRUE
14	5	0.9	2.12	TRUE
15	7	0.9	2.07	TRUE
16	10	0.9	2.04	TRUE
17	12	0.9	2.01	TRUE
18	15	0.9	1.97	TRUE
19	20	0.9	1.95	TRUE
20	25	0.9	1.91	TRUE

The sign test[174] performed on simulation results reveals the p-value of 1 based on the binomial distribution.[175] Because the p-value is larger than the alpha level of 0.05, we fail to reject null hypothesis and cannot accept the alternative hypothesis for research.

The final supporting hypothesis stated that gene drive altered mosquitoes can replace the entire wild mosquito population within only 10 years, and by doing so, eradicate malaria, given that no resistance to gene drives is developed. Unfortunately, in each of the simulations, resistance to gene drive was developed. Considering purely replacement of wild population within a decade, the hypothesis can be written as:

H_0: *Genetically altered organisms cannot replace the entire wild population within a decade.*

H_A: *Genetically altered organisms can replace the entire wild population within a decade.*

To statistically test this assumed expectation both datasets (Table 6 and Table 7) have been merged together (Table 11) and a sign test was performed to test the hypothesis. The test sign revealed the p-value of 0.979. As the p-value is larger than the alpha level of 0.05, we fail to reject the null hypothesis and conclude that genetically altered organisms cannot replace the entire wild population within a decade.

Table 11 *Simulation results showing the number of years that are required to replace wild mosquito population.*

Simulation number	The initial number of gene drive mosquitoes (as a percentage of wild population)	Initial inheritance frequency / chance (percent) – male (*eM*)	Number of years required to replace wild mosquitoes	Wild mosquito population replaced in a decade? (TRUE/FALSE)
1	1	0.9	10.98	FALSE
2	1	1	6.29	TRUE
3	1	0.99	6.51	TRUE
4	1	0.97	6.99	TRUE
5	1	0.95	7.56	TRUE
6	1	0.93	8.23	TRUE
7	1	0.91	9.05	TRUE
8	1	0.9	9.53	TRUE
9	1	0.89	10.04	FALSE
10	1	0.87	11.38	FALSE
11	1	0.85	13.18	FALSE
12	2	0.9	10.88	FALSE
13	3	0.9	10.45	FALSE
14	5	0.9	9.97	TRUE
15	7	0.9	9.76	TRUE
16	10	0.9	9.53	TRUE
17	12	0.9	9.40	TRUE
18	15	0.9	9.28	TRUE
19	20	0.9	9.11	TRUE
20	25	0.9	8.96	TRUE

Finally, the main hypothesis for the research stated that based on initial parameters such as the location of release, number of released gene drive mosquitoes, landscape, gene drive frequency, etc. gene drives can be either effective or less effective. Because majority of supporting null hypotheses have failed to be rejected, we

conclude that parameters tested in this thesis have had an effect on the effectiveness of gene drive technology. However, further simulations with different starting parameters and conditions may change this outcome. Due to time limitations, not all potential initial parameters, and their influences on gene drive mosquito-spread have been tested. For example, mosquitoes in all simulations have been released from the same city in Kenya, and no landscape consideration was given in the simulations. This hypothesis is fulfilled for number of released gene drive mosquitoes and gene drive inheritance frequency. The uncertainty analysis for the model has been tested using several different scenarios. This analysis is presented in Appendix E.

6. Discussion

Presented results align closely with previously published studies for gene drive altered mosquitoes.[2,3,7,11,22] It was expected that increase in the relative number of gene drive modified mosquitoes into the wild mosquito population. This relation has been previously shown in other studies, including laboratory-based tests.[2,3,24,26] The percentage ratio set to 1% of gene drive altered mosquitoes to the initial wild population released at Malindi in Kenya resulted in 4,009 days to replace the entire wild population of native mosquitoes with gene drive altered ones, including mosquitoes that developed gene drive resistance. However, this initial percentage ratio's increase to 25% decreased the time to replace the wild mosquito population to 3,271 days. This decrease is insignificant, given how much the percentage number increased. This is because considerable time is required for gene drive altered mosquitoes to travel from Malindi and spread throughout the study area. Another reason might be the inheritance ratio set up in those simulations to be 0.9, meaning that 0.1 of offspring still have a chance of not inheriting the gene drive from its parents. The time requirement to phase out wild mosquito population oscillating between 10 to 12 years is reasonable for short-lived organisms like Anopheles gambiae with an average life span of 15 days, which aligns with the previous studies.[14,167,168]

The inheritance ratio test outcomes have shown that the higher inheritance ratios, the shorter time required for wild mosquito replacement. This relationship is also well reasoned and supported by previous studies. The replacement of the wild mosquito population in Kenya took 2,297 days, assuming 100% inheritance frequency for gene drive and 10% ratio between released gene drive mosquitoes and the native population in Malindi. A decrease in inheritance frequency of gene drives increased the time required to replace wild mosquitoes in Kenya to 4,811 days. This has also been shown in previous studies. [2,3,7,22,167,168]

The above results clearly show that the impact of inheritance frequency on end results and time required to replace the wild mosquito population is significantly greater than the impact of the number of released gene drive mosquitoes. However, an important consideration, in this case, needs to be understood. The number of gene drive altered mosquitoes equal to 10% of the wild population in Malindi corresponds to 9,402 gene drive altered mosquitoes released in Kenya. This, however, corresponds to only 0.00144% of the entire Kenyan mosquito population of 6,530,363 mosquitoes.

Previous studies which explored the release of genetically modified organisms relied on a more decentralized approach,[6,8,9,11,14] where genetically modified organisms have been released from multiple nodes or locations throughout the study area. This approach has proven to be more effective in achieving more effective outcomes.[6,9,14] This was also confirmed via MGDrivE simulations (see Appendix E). Other potential considerations to make simulation results more aligned with real-world outcomes are described in chapter '*Future Work*'.

6.1 Extending the Further Capacity of the MGDrivE Package

While the MGDrivE package is very intuitive, robust and gives a user a significant degree of flexibility, currently the package is not fully spatial. The so-called landscape module is quite limited. Currently, a landscape module accommodates the distribution of insect metapopulations connected by migration within the study area and facilitates the movement through the resulting network determined by dispersal kernels. Whereas metapopulation is considered as a population of populations, or a group of populations, made up of the same species.

Moreover, each subpopulation, or sub-group, is separated from all other subpopulations. However, the movement of individuals between subpopulations is allowed and occurs regularly. In other words, MGDrivE is working with a limited number of larger subpopulations rather than individuals. This is mainly due to the computing power and associated computing limitations, considering each individual to be a separate node within the network, which is capable of moving around and interacting with other individuals, as well as of course being able to transmit malaria parasite would require significant computing capabilities, especially when mosquito populations can easily reach thousands of units in size.

A significant body of previous work has shown that populations are often vulnerable to fragmentation. When a larger uniform population becomes fragmented into smaller populations, small and isolated local populations are often too small to sustain. However, when these local populations remain connected with each other via the ecological network, such metapopulation can often maintain local populations.[163,176] Large local populations within the metapopulation have a low probability of extinction and become a driving force for metapopulation sustainability. These large populations are called '*key populations*' because of their persistent role.[177] Future work could focus on evaluating how environmental barriers prevent or support the formation of metapopulations. The simplest way to perform such simulation would be to manually break the network between one or more populations and evaluate the impact. The exercise would have on population survival. The previous study shows that for large vertebrates, the standard number of reproductive units for a key patch is 20.[177] The future study could establish similar numbers for different mosquito species. Independent studies have shown that metapopulation is sustainable if the chance of local extinction is less than 5% in 100 years.[177-179]

In addition, future work could evaluate the role of dispersal range on metapopulation sustainability. Previous studies have shown that less mobile species should form habitat patches more densely packed and situated closer together to form coherent ecological networks.[180] The question that remains open is sustainability versus extinction of malaria vectoring mosquitoes. If we let these mosquito populations go extinct, malaria disease will be gone too. Gene drive systems can be used to suppress mosquito populations. While current simulation was designed to model population replacement, several studies looked at suppressing malaria mosquito populations using gene drive systems. Future work could also evaluate such scenarios.[12,24,169]

To deal with the computing power issue, MGDrivE is designed to cluster mosquitoes into subpopulations forming a metapopulation. This limits the number of nodes, where nodes often correspond to cities, villages, or other human settlements where mosquito populations thrive. The solution, however, limits the capabilities of MGDrivE and makes the simulations as well as corresponding results less reliable. Another limitation of the solution is that mosquitoes within the metapopulation do not have their own geographic coordinates, as such all the mosquitoes in the specific metapopulation are in the exact same location, which corresponds to the geographic latitude and longitude of the simulated node.

In the MGDrivE, the metapopulations are populations that are mixing randomly. The equations of the lumped age-class model are applied to those populations. In terms of size, the resolution of the metapopulations is selected accordingly to the dispersal properties of the investigated insect species.

Addressing the issues described above and making MGDrivE a fully spatial package is of significant interest as it would improve the outcome quality. There are a

number of steps and modifications to consider. For example, it is believed that with the current code, the package could still maintain and support scenarios where limited, probably up to 100 individuals are modelled as single network nodes, each of which would have its geographic coordinates. In this case, the code modification would be minimal and would include only additional code lines to allow the nodes to be mobile within the given study area to simulate mosquito mobility. While a population of 100 or so individuals is not much, it could still provide valuable outcomes for very small study areas, such as a village.

Further modifications to expand MGDrivE capabilities to make the package more functional and to emulate more realistic environmental conditions could focus on forcing certain limitations to the current algorithm. The MGDrivE computes the distance between each of the nodes and performs a number of other calculations to compute different parameters for each pair of nodes. The procedure is very time-consuming from a computational point of view. To prevent this, a simple '*if*' statement could be put in place once the distance between the pair of mosquitoes/nodes is calculated and using logical expressions checks whether the distance is larger than the distance initially assumed by the user. If the distance is larger, nothing happens, and the algorithm moves to the next mosquito pair. However, if the calculated distance is equal or smaller to the distance initially agreed upon, the algorithm computes all these other parameters. In this case, the distance could be as little as 500 metres to one kilometre. It is unlikely that a pair of mosquitoes further away from each other will have any interaction (Figure 11).
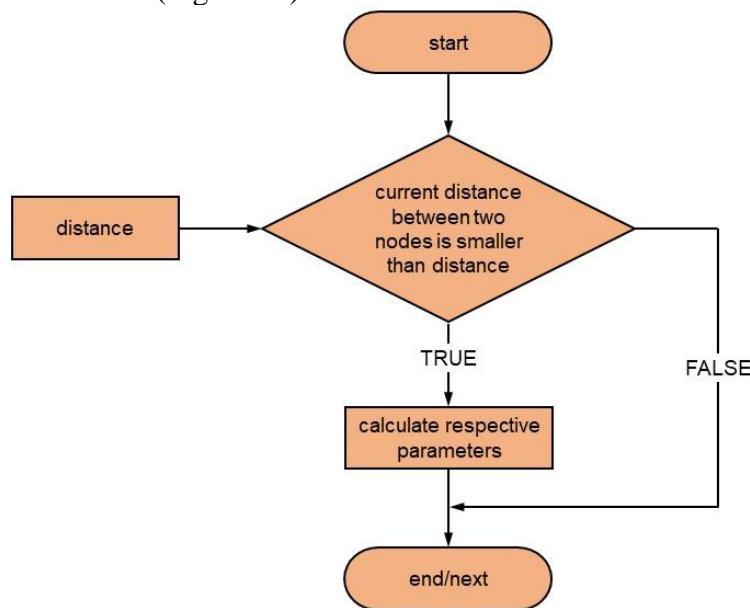

Figure 11 *Software system diagram illustrating potential distance restriction.*

Another modification of the existing code towards making it fully spatial would be to divide the entire mosquito population into smaller subpopulations based on geographic coordinates. Every single mosquito within the subpopulation would then be treated as a separate individual with its coordinates and capability to move around. Regular calculations, as per the current algorithm would be limited to calculate the distance only between pairs of mosquitoes within this subpopulation. This would create a separate network of nodes within the subpopulations. However, the subpopulations would by itself also be considered as a node of a higher instance and create its own

network, allowing for single mosquitoes to travel between. All of this is done to reduce the computational burden of the simulation. This case is shown in Figure 12.
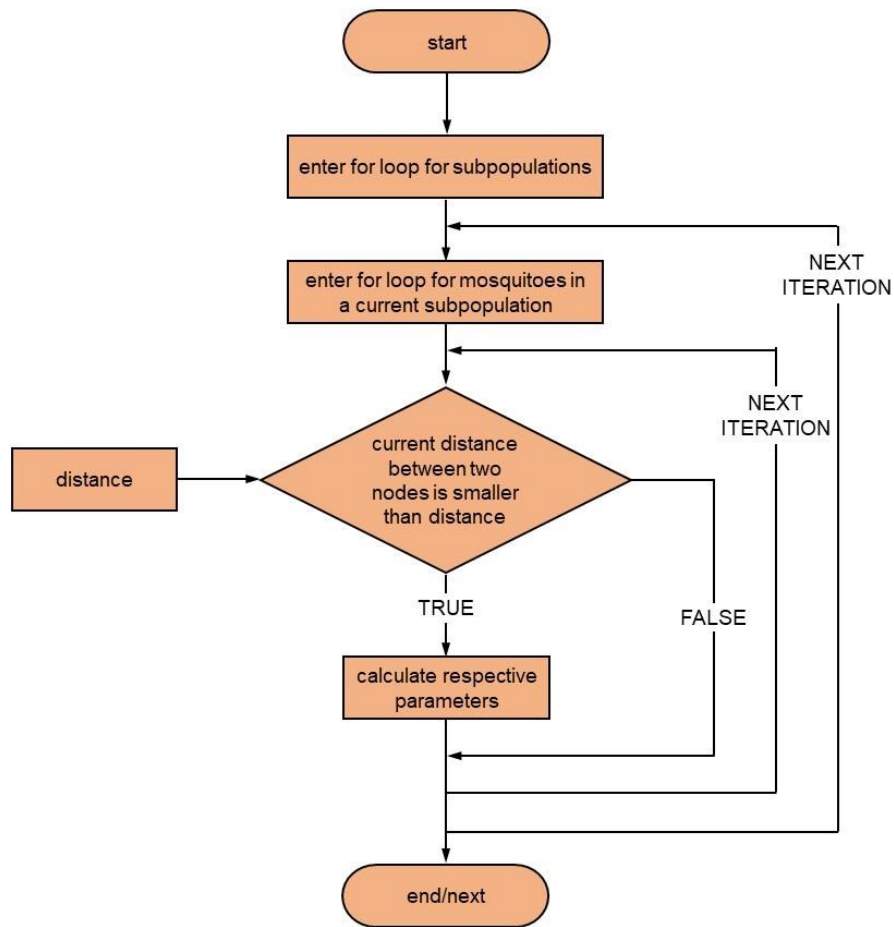

Figure 12 *Software system diagram illustrating potential distance restriction within separate subpopulations.*

Because interaction between those subpopulations is possible to reduce the computational burden further, once every five or so iteration, the distance between each one of the mosquitoes could be calculated and based on the result specific decision made. For example, it can be revealed that mosquitoes travelled so far within this given interval that now they should be considered as part of a different subpopulation, which is geographically closer to their original one. In such a case, sub-populational migration would occur. This case is presented in Figure 13.
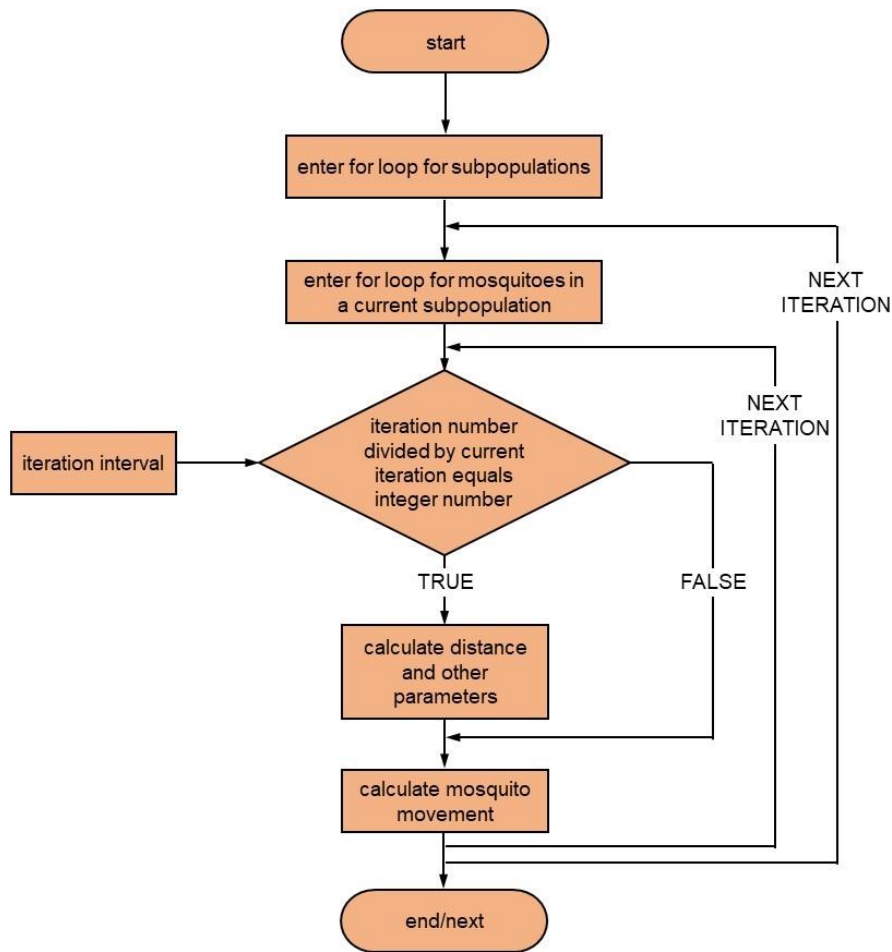
Figure 13 *Software system diagram illustrating distance calculation between mosquitoes every given simulation iteration number.*

Another way of reducing the computational burden required to model each of the mosquitoes independently and making the simulation spatial would be to limit the number of potential interactions of each mosquito to, for example, five. Meaning that every single mosquito within its own lifetime can interact with only a certain number of other mosquitoes, in this case, the number would be five mosquitoes. Considering a very short lifespan of mosquitoes, it is fair to assume that it will interact with only a limited number of other mosquitoes. These other mosquitoes would be selected based on the calculated distance. While this is a significant limitation of the study, it would allow extensively reducing the computing power (and time) required to perform such simulation and at the same time enabling the extension of the capabilities of the current environment from semi-spatial to spatial (Figure 14).
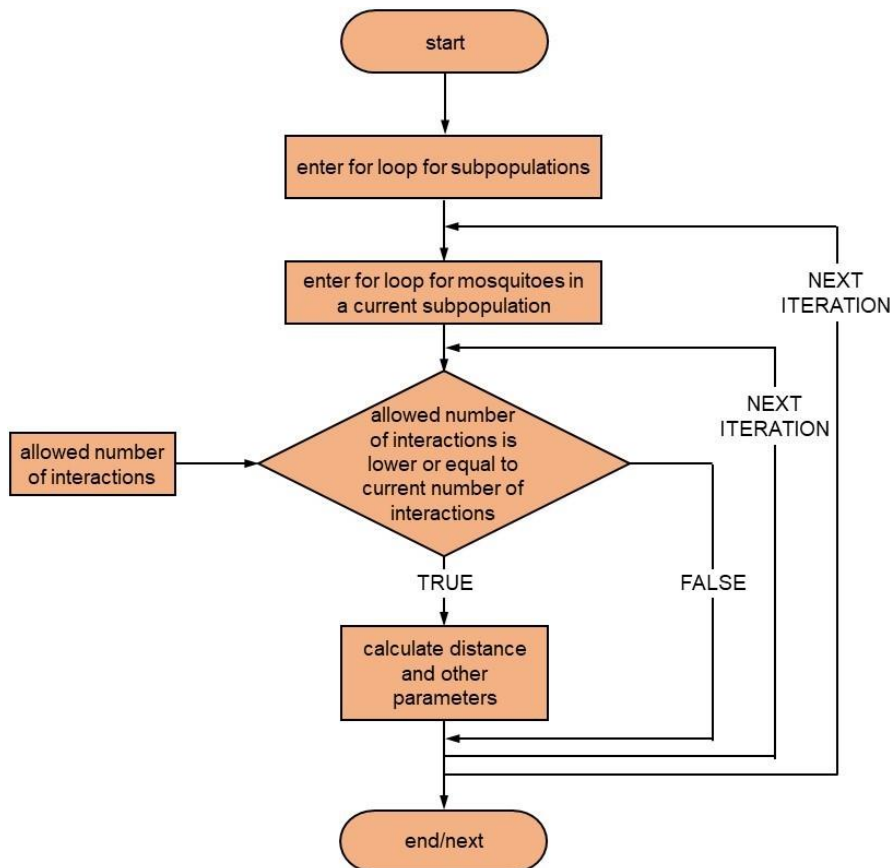
Figure 14 *Software system diagram illustrating potential mosquito limitation to interact with a given number of other mosquitoes.*

Introducing directed networks and clustering these networks with geographical and landscape constraints could be another step allowing for the reduction of the computational power required to calculate such simulations and contribute to making MGDrivE more spatial environment.

Introducing geographical and landscape constraints, as well as environmental factors such as temperature, humidity, rainfall, elevation, etc., would prevent some of the mosquitoes from interacting with each other and at the same time forming clustered populations which interact only with each other as such the larger network would be disrupted decreasing the number of potential interaction combinations between mosquitoes.

One other way of reducing the time required to perform such simulation and the same time to reduce computational power required would be to migrate the existing code to a low-level programming language. A low-level programming language is a programming language that provides little or no abstraction from computer instructions, commands or functions to processor instructions or machine code. Generally, this allows skipping a layer or number of layers that are required to translate the given commands to those understood by the processor, the same time significantly reducing the time necessary to perform such instructions.

There are also other ways of modelling large populations of mobile mosquitoes as independent nodes. However, this would require significant computing power. Supercomputers could quickly deal with large populations. Unfortunately, supercomputers are not easily accessible, and their computing time is either very

expensive or limited and often by application only. At the current stage modelling a network of 36 nodes with over 6.5 million mosquitoes over a period of ten years (3,650 iterations) took over 90 minutes using an average computer. Performing the same simulations using supercomputer would reduce this time to single minutes, maybe even seconds.

In addition, as the mosquito movement within the study area is modelled with complex migration networks, these networks can serve as a valuable source of further information regarding modelled systems. It would be of value to further extend the capabilities of the package and include a tool that will allow studying the movement of mosquitos and spatial gene-flows within the studied area. The new tool could also perform network analysis to identify any routines between mosquitoes but also between generations.

Being able to study these networks in more detail would reveal certain landscape-related factors influencing genes and diseases transmission. This would enable individuals to design more complex research questions. It could, for example, be possible to design a scenario with two separate populations and networks. One of which would correspond to mosquito populations, whereas the other one would simulate the human population. Having in-depth and more detailed information on the network and different routines within two populations would allow establishing how the movement of humans and mosquitoes influence malaria and gene-drive patterns.

6.2 Case Study Specific Aspects of Future Work

Other work could include multiple avenues and a number of smaller sub-studies. For example, it would be interesting to evaluate how quickly gene drive altered mosquitoes will replace existing, wild populations when introduced from multiple nodes/cities. The current arrangement which involves the introduction of engineered species at only one node, which is a coastal town in Kenya, is not preferential because cities or towns that are further away from the coast and will take longer for modified mosquitoes to migrate to the area. It would also be interesting to see how introduction into larger cities, such as Mombasa or Nairobi, influences the spatial distribution and the time required to replace the wild population. In the current scenario, mosquitoes are introduced in a coastal town with a small population, as such, from the beginning the spread is limited. Having a release node in a major city could give very different results.

Further, it would be of interest to introduce more nodes to the existing network of 36 cities. Additional nodes would potentially shorten the time required for mosquitoes to travel between the cities and as such, would reduce the time required to replace wild populations throughout all nodes in the network. However, the introduction of further nodes would have a negative effect on the computing time needed to perform the simulation. Some of those new nodes could be small rural areas and villages. The introduction of additional node parameter, which specifies whether the area is urban or rural and assigns specific probability parameter to it would also be an interesting experiment.

It would be interesting to introduce geographic and climatic parameters that influence the scenario. Temperature, latitude, altitude, rainfall, humidity, and a number of other parameters influence the mosquito and parasite life cycle. Including these parameters in the simulation would add a level of complexity to the model but would also make it more realistic and more precise. For example, it would allow linking the results to a specific time of the year. For example, the temperature in Kenya changes significantly throughout the year. Temperature is one of the major factors enabling mosquito and parasite development. Having altitude as a contributing factor to the

simulation would allow certain areas where climatic conditions are not favourable for mosquito development to be excluded from the studies.

Future environmental and climatic parameters have to be selected so that simulation can be performed without putting further burden on computing capabilities. Individual climatic and environmental layers might be correlated or compliment with each other. For example, altitude and temperature could be reduced to altitude layer only, since the temperature can be modelled as a function of latitude, longitude, and altitude. A simple equation that would translate geographic parameters to temperature could be used instead of a separate temperature layer.

Another effort to expand modelling capabilities could focus on the introduction of natural barriers. For example, it is known that mosquitoes do not thrive at higher altitudes due to significantly lower average temperatures at those elevated levels. This could create natural barriers where mosquitoes cannot live in or pass through. Other types of natural obstacles would be large water bodies and deserts, among others. This scenario is presented in form of a schematic in Figure 15.
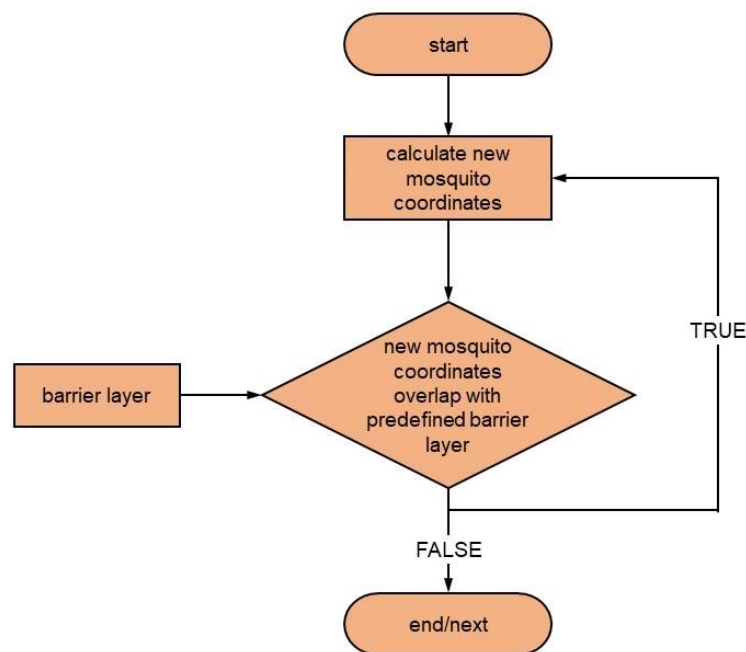

Figure 15 *Software system diagram illustrating natural barriers and limitations for mosquito travel.*

The environmental barriers not only have an immense effect on single metapopulations but also on the population as a whole. Factors such as species characteristics, their key lifecycle processes (i.e., larval or egg cycle), amount, shape as well as the area and spatial configuration of habitat patches (species connectivity to the landscape) all indicate whether an ecological network, like the network created between metapopulations in the current simulation, can sustain a persistent population. Modelling of environmental barriers could be a primary focus of future work as these barriers influence species change of extinction and survival.

All of the cases presented in the current study focus on population replacement. It would be interesting to run a series of similar scenarios where the main goal is population suppression to reduce the malaria effect and probability rather than complete replacement. It would be further interesting to compare how suppression compares to

50

replacement in terms of time, also because new entrants from natural populations may occur naturally at a later stage and then a suppression scenario is more cost-effective.

The current study assumes that gene drive modified mosquitoes are released only once during the entire simulation. It could be interesting to see how continuous, for example, once a year release of gene drive mosquitoes in smaller batches affects the spatial distribution and the time required for replacement of wild populations. It could be interesting to see the effect of different release intervals on the overall effectiveness of the strategy.

Finally, it would be interesting to continue the current two studies with different parameters to establish whether they always follow a certain trend. Extending the time of the studies would also be an interesting experiment. The data gathered thus far shows that once the wild population is entirely replaced by gene drive altered population, the number of gene drive mosquitoes slowly decreases while the mosquitoes with gene drive resistance lowly but effectively increase. As such, one would think that having long enough time the resistance could take over and make mosquitoes immune to gene drive, same time allowing further spread of malaria.

7. Conclusions

The undertaken computational study has further confirmed the complexity of the studied system, and the role of each of the studied system parameters on the overall outcome. The results have shown that it is feasible to alter the wild mosquito population using gene drives to introduce specific traits. However, the study has also demonstrated that complete eradication of malaria using gene drives is potentially not achievable because a number of individuals in the population would develop gene drive resistance, which would make the strategy ineffective for a fraction of the population while at the same time allowing them to further the vector malaria parasite.

The number of individuals in the population that would develop gene drive resistance is relatively small and is estimated to be 10% on average. Gene drive-altered mosquitoes would be an effective tool for reducing the burden of malaria. However, complete eradication is not possible based on the results of the computational study. Gene drive technology could potentially be combined with traditional preventive methods to further lower the probability of malaria contraction.

The study has also revealed that altering wild population is achievable within less than a decade depending on the gene drive system frequency and release strategy, namely the number of gene drive mosquitoes released into the wild population. Whereas the number of gene drive mosquitoes initially related to interacting with the wild population can be relatively small, the frequency of the gene drive system must be very high to make the replacement feasible within a decade. Generally, it seems that gene drive systems with inheritance frequencies of 90% or above would be sufficient enough to achieve the goal within the next ten years., given the resistance to gene drives is not developed. If gene drives resistance occurs, gene drives mosquitoes could also bring desired effects and reduce the malaria spread, but it wouldn't eliminate malaria.

Readers must bear in mind that the results presented here are based on a mathematical model, which aims to emulate the real case scenarios. However, a number of factors have not been included in the model as such the actual release of gene drive modified mosquitoes to the environment may show significantly different outcomes.

8. Research Ethics

The potential ethical issues associated with the study are non-existent. The study is based on previously developed models that have been adapted for the study. Neither mosquitoes or malaria parasites nor gene drives have been tested in the contained laboratory or broader environment as part of the study. The study was designed to avoid any potential issues arising from the use of living or living modified organisms as defined in the Cartagena Protocol on Biosafety to the Convention on Biological Diversity.

All results presented in the study are purely theoretical and are a result of mathematical models and computational simulations of the models in the study. Whereas, the models used in the study and developed in other studies are based on validated laboratory data for gene drives or real field data for mosquitoes, this researcher did not influence how the validation data was collected, and whether all biosafety and bioethics rules and regulations were in place. However, the researcher believes that the original data collected to develop the models met all bioethics and biosafety requirements.

Further, the data presented in the study and obtained in the simulation process may not represent real-life scenarios. However, the researcher put adequate effort to emulate, as closely as possible, the scenarios that can happen in the real world. In either case, the researcher does not encourage anyone to try to reproduce the presented simulations in the real-life scenario using gene drive-modified mosquitoes. Such action could result in breaching the Cartagena Protocol on Biosafety to the Convention on Biological Diversity as well as other country-specific or international protocols and laws.

The researcher behind this study would like to remind everyone that the Precautionary Principle, which states *"when an activity raises threats of harm to human health or environment, precautionary measures should be taken even if some cause and effect relationships are not fully established scientifically,"* should always be a priority.

While the introduction of gene drive altered mosquitoes shows a promise in decreasing malaria burden in Kenya, one needs to remember that the release of gene drive-altered organisms would be ethically wrong and could pose serious threats to Kenyan biodiversity and human health.

9. References

1   Noble, C., Olejarz, J., Esvelt, K. M., Church, G. M. & Nowak, M. A. Evolutionary dynamics of CRISPR gene drives. *Science Advances* **3**, e1601964, doi:10.1126/sciadv.1601964 (2017).

2   Hamilton, W. D. Extraordinary Sex Ratios. *Science* **156**, 477-488, doi:10.1126/science.156.3774.477 (1967).

3   de Jong, T. J. Gene drives do not always increase in frequency: from genetic models to risk assessment. *Journal of Consumer Protection and Food Safety* **12**, 299-307, doi:10.1007/s00003-017-1131-z (2017).

4   Akbari, O. S. *et al.* Safeguarding gene drive experiments in the laboratory. *Science* **349**, 927-929, doi:10.1126/science.aac7932 (2015).

5   DiCarlo, J. E., Chavez, A., Dietz, S. L., Esvelt, K. M. & Church, G. M. Safeguarding CRISPR-Cas9 gene drives in yeast. *Nature Biotechnology* **33**, 1250, doi:10.1038/nbt.3412 (2015).

6   Burt, A. Site-specific selfish genes as tools for the control and genetic engineering of natural populations. *Proceedings of the Royal Society of London. Series B: Biological Sciences* **270**, 921-928, doi:doi:10.1098/rspb.2002.2319 (2003).

7   Deredec, A., Burt, A. & Godfray, H. C. J. The Population Genetics of Using Homing Endonuclease Genes in Vector and Pest Management. *Genetics* **179**, 2013-2026, doi:10.1534/genetics.108.089037 (2008).

8   Alphey, L. *et al.* Malaria Control with Genetically Manipulated Insect Vectors. *Science* **298**, 119-121, doi:10.1126/science.1078278 (2002).

9   Curtis, C. F. Possible Use of Translocations to fix Desirable Genes in Insect Pest Populations. *Nature* **218**, 368-369, doi:10.1038/218368a0 (1968).

10   Lloyd, A. L. & May, R. M. How Viruses Spread Among Computers and People. *Science* **292**, 1316-1317, doi:10.1126/science.1061076 (2001).

11   Eckhoff, P. A., Wenger, E. A., Godfray, H. C. J. & Burt, A. Impact of mosquito gene drive on malaria elimination in a computational model with explicit spatial and temporal dynamics. *Proceedings of the National Academy of Sciences* **114**, E255-E264, doi:10.1073/pnas.1611064114 (2017).

12   Marshall, J. M., Pittman, G. W., Buchman, A. B. & Hay, B. A. Semele: A Killer-Male, Rescue-Female System for Suppression and Replacement of Insect Disease Vector Populations. *Genetics* **187**, 535-551, doi:10.1534/genetics.110.124479 (2011).

13   Crain, P. R. *et al.* Wolbachia infections that reduce immature insect survival: Predicted impacts on population replacement. *BMC Evolutionary Biology* **11**, 290, doi:10.1186/1471-2148-11-290 (2011).

14   Backus, G. A. & Gross, K. Genetic engineering to eradicate invasive mice on islands: modeling the efficiency and ecological impacts. *Ecosphere* **7**, e01589, doi:doi:10.1002/ecs2.1589 (2016).

15   Min, J., Noble, C., Najjar, D. & Esvelt, K. M. Daisyfield gene drive systems harness repeated genomic elements as a generational clock to limit spread. *bioRxiv*, 104877, doi:10.1101/104877 (2017).

16   Arcieri, M. Spread and Potential Risks of Genetically Modified Organisms. *Agriculture and Agricultural Science Procedia* **8**, 552-559, doi:https://doi.org/10.1016/j.aaspro.2016.02.072 (2016).

17   Oye, K. A. *et al.* Regulating gene drives. *Science* **345**, 626-628, doi:10.1126/science.1254287 (2014).

18   National Academies of Sciences, E. & Medicine. *Gene Drives on the Horizon: Advancing Science, Navigating Uncertainty, and Aligning Research with Public Values.* (The National Academies Press, 2016).

19   Esvelt, K. M., Smidler, A. L., Catteruccia, F. & Church, G. M. Concerning RNA-guided gene drives for the alteration of wild populations. *eLife* **3**, e03401, doi:10.7554/eLife.03401 (2014).

20   Marshall, J. M., Buchman, A., Sánchez C, H. M. & Akbari, O. S. Overcoming evolved resistance to population-suppressing homing-based gene drives. *Scientific Reports* **7**, 3776, doi:10.1038/s41598-017-02744-7 (2017).

21   Unckless, R. L., Clark, A. G. & Messer, P. W. Evolution of Resistance Against CRISPR/Cas9 Gene Drive. *Genetics* **205**, 827-841, doi:10.1534/genetics.116.197285 (2017).

22   Dhole, S., Vella, M. R., Lloyd, A. L. & Gould, F. Invasion and migration of spatially self-limiting gene drives: A comparative analysis. *Evolutionary Applications* **11**, 794-808, doi:doi:10.1111/eva.12583 (2018).

23   Noble, C. *et al.* Daisy-chain gene drives for the alteration of local populations. *bioRxiv*, 057307, doi:10.1101/057307 (2016).

24   Beaghton, A., Beaghton, P. J. & Burt, A. Gene drive through a landscape: Reaction–diffusion models of population suppression and elimination by a sex ratio distorter. *Theoretical Population Biology* **108**, 51-69, doi:https://doi.org/10.1016/j.tpb.2015.11.005 (2016).

25   Brennecke, J. *et al.* An Epigenetic Role for Maternally Inherited piRNAs in Transposon Silencing. *Science* **322**, 1387-1392, doi:10.1126/science.1165171 (2008).

26   Windbichler, N. *et al.* A synthetic homing endonuclease-based gene drive system in the human malaria mosquito. *Nature* **473**, 212, doi:10.1038/nature09937 (2011).

27   Hammond, A. *et al.* A CRISPR-Cas9 gene drive system targeting female reproduction in the malaria mosquito vector Anopheles gambiae. *Nature Biotechnology* **34**, 78, doi:10.1038/nbt.3439 (2015).

28   Thresher, R. *et al.* Sex-ratio-biasing constructs for the control of invasive lower vertebrates. *Nature Biotechnology* **32**, 424, doi:10.1038/nbt.2903 (2014).

29   Bull, J. J. Lethal gene drive selects inbreeding. *Evolution, Medicine, and Public Health* **2017**, 1-16, doi:10.1093/emph/eow030 (2016).

30   Alphey, N. & Bonsall Michael, B. Interplay of population genetics and dynamics in the genetic control of mosquitoes. *Journal of The Royal Society Interface* **11**, 20131071, doi:10.1098/rsif.2013.1071 (2014).

31   Moller-Jacobs, L. L., Murdock, C. C. & Thomas, M. B. Capacity of mosquitoes to transmit malaria depends on larval environment. *Parasites & Vectors* **7**, 593, doi:10.1186/s13071-014-0593-4 (2014).

32   Churcher, T. S. *et al.* Probability of Transmission of Malaria from Mosquito to Human Is Regulated by Mosquito Parasite Density in Naïve and Vaccinated Hosts. *PLOS Pathogens* **13**, e1006108, doi:10.1371/journal.ppat.1006108 (2017).

33   Okell, L. C. *et al.* Factors determining the occurrence of submicroscopic malaria infections and their relevance for control. *Nature Communications* **3**, 1237, doi:10.1038/ncomms2241 (2012).

34   Haque, U. *et al.* The Role of Climate Variability in the Spread of Malaria in Bangladeshi Highlands. *PLOS ONE* **5**, e14341, doi:10.1371/journal.pone.0014341 (2010).

35   Ohta, S. & Kaga, T. Effect of climate on malarial vector distribution in Monsoon Asia coupled model for Ecophysiological and Climatological Distribution of mosquito generations (ECD-mg). *Climate Research* **53**, 77-88 (2012).

36   Marinho, R. A. *et al.* Effects of temperature on the life cycle, expansion, and dispersion of Aedes aegypti (Diptera: Culicidae) in three cities in Paraiba, Brazil. *Journal of Vector Ecology* **41**, 1-10, doi:doi:10.1111/jvec.12187 (2016).

37   Shapiro, L. L. M., Whitehead, S. A. & Thomas, M. B. Quantifying the effects of temperature on mosquito and parasite traits that determine the transmission potential of human malaria. *PLOS Biology* **15**, e2003489, doi:10.1371/journal.pbio.2003489 (2017).

38   Beck-Johnson, L. M. *et al.* The effect of temperature on Anopheles mosquito population dynamics and the potential for malaria transmission. *PloS one* **8**, e79276-e79276, doi:10.1371/journal.pone.0079276 (2013).

39 Mitzmain, M. B. The Malaria Parasite in the Mosquito: The Effects of Low Temperature and Other Factors on Its Development. *Public Health Reports (1896-1970)* **32**, 1400-1413, doi:10.2307/4574614 (1917).

40 Ciota, A. T., Matacchiero, A. C., Kilpatrick, A. M. & Kramer, L. D. The effect of temperature on life history traits of Culex mosquitoes. *Journal of medical entomology* **51**, 55-62 (2014).

41 Farjana, T. & Tuno, N. Multiple Blood Feeding and Host-Seeking Behavior in Aedes aegypti and Aedes albopictus (Diptera: Culicidae). *Journal of Medical Entomology* **50**, 838-846, doi:10.1603/me12146 (2013).

42 Aly, A. S. I., Vaughan, A. M. & Kappe, S. H. I. Malaria parasite development in the mosquito and infection of the mammalian host. *Annual review of microbiology* **63**, 195-221, doi:10.1146/annurev.micro.091208.073403 (2009).

43 Blanford, J. I. *et al.* Implications of temperature variation for malaria parasite development across Africa. *Scientific Reports* **3**, 1300, doi:10.1038/srep01300 (2013).

44 Antinori, S., Galimberti, L., Milazzo, L. & Corbellino, M. Biology of human malaria plasmodia including Plasmodium knowlesi. *Mediterranean journal of hematology and infectious diseases* **4**, e2012013-e2012013, doi:10.4084/MJHID.2012.013 (2012).

45 Singh, B. & Daneshvar, C. Human Infections and Detection of *Plasmodium knowlesi*. *Clinical Microbiology Reviews* **26**, 165-184, doi:10.1128/cmr.00079-12 (2013).

46 Kantele, A. & Jokiranta, T. S. Review of Cases With the Emerging Fifth Human Malaria Parasite, Plasmodium knowlesi. *Clinical Infectious Diseases* **52**, 1356-1362, doi:10.1093/cid/cir180 (2011).

47 Fialho, R. F. & Schall, J. J. Thermal Ecology of a Malarial Parasite and its Insect Vector: Consequences for the Parasite's Transmission Success. *Journal of Animal Ecology* **64**, 553-562, doi:10.2307/5799 (1995).

48 Mzilahowa, T. *et al.* Risk factors for Anopheles mosquitoes in rural and urban areas of Blantyre District, southern Malawi. *Malawi medical journal : the journal of Medical Association of Malawi* **28**, 154-158 (2016).

49 Jatta, E. *et al.* How house design affects malaria mosquito density, temperature, and relative humidity: an experimental study in rural Gambia. *The Lancet Planetary Health* **2**, e498-e508, doi:10.1016/S2542-5196(18)30234-1 (2018).

50 Shayo, D. *et al.* Relationship Between Altitude and Intensity of Malaria Transmission in the Usambara Mountains, Tanzania. *Journal of Medical Entomology* **40**, 706-717, doi:10.1603/0022-2585-40.5.706 (2003).

51 Attenborough, R. D., Burkot, T. R. & Gardner, D. S. Altitude and the risk of bites from mosquitoes infected with malaria and filariasis among the Mianmin people of Papua New Guinea. *Transactions of The Royal Society of Tropical Medicine and Hygiene* **91**, 8-10, doi:10.1016/s0035-9203(97)90373-4 (1997).

52 Alsop, Z. Malaria returns to Kenya's highlands as temperatures rise. *The Lancet* **370**, 925-926, doi:10.1016/S0140-6736(07)61428-7 (2007).

53 Shanks, G. D., Simon, I. H., Judy, A. O. & Robert, W. S. Malaria in Kenya's Western Highlands. *Emerging Infectious Disease journal* **11**, 1425, doi:10.3201/eid1109.041131 (2005).

54 Paaijmans, K. P., Read, A. F. & Thomas, M. B. Understanding the link between malaria risk and climate. *Proceedings of the National Academy of Sciences* **106**, 13844-13849, doi:10.1073/pnas.0903423106 (2009).

55 Hundessa, S. *et al.* Spatial change in the risks of Plasmodium vivax and Plasmodium falciparum malaria in China, 2005;2013;2014. *Infection, Disease & Health* **21**, 89-96, doi:10.1016/j.idh.2016.08.001 (2016).

56 Hamza, A. M. & Rayah, E. A. E. A Qualitative Evidence of the Breeding Sites of Anopheles arabiensis Patton (Diptera: Culicidae) in and Around Kassala Town, Eastern Sudan. *International Journal of Insect Science* **8**, IJIS.S40071, doi:10.4137/IJIS.S40071 (2016).

57   Mattah, P. A. D. *et al.* Diversity in breeding sites and distribution of Anopheles mosquitoes in selected urban areas of southern Ghana. *Parasites & Vectors* **10**, 25, doi:10.1186/s13071-016-1941-3 (2017).

58   Bashar, K. & Tuno, N. Seasonal abundance of Anopheles mosquitoes and their association with meteorological factors and malaria incidence in Bangladesh. *Parasites & Vectors* **7**, 442, doi:10.1186/1756-3305-7-442 (2014).

59   Nabi, S. & Qader, S. Is Global Warming likely to cause an increased incidence of Malaria? *The Libyan journal of medicine* **4**, 18-22, doi:10.4176/090105 (2009).

60   Wanjala, C. L. & Kweka, E. J. Impact of Highland Topography Changes on Exposure to Malaria Vectors and Immunity in Western Kenya. *Frontiers in Public Health* **4**, doi:10.3389/fpubh.2016.00227 (2016).

61   Barros, F. S. M. & Honório, N. A. Deforestation and Malaria on the Amazon Frontier: Larval Clustering of Anopheles darlingi (Diptera: Culicidae) Determines Focal Distribution of Malaria. *The American Journal of Tropical Medicine and Hygiene* **93**, 939-953, doi:https://doi.org/10.4269/ajtmh.15-0042 (2015).

62   Rufalco-Moutinho, P., Schweigmann, N., Bergamaschi, D. P. & Mureb Sallum, M. A. Larval habitats of Anopheles species in a rural settlement on the malaria frontier of southwest Amazon, Brazil. *Acta Tropica* **164**, 243-258, doi:https://doi.org/10.1016/j.actatropica.2016.08.032 (2016).

63   Lee, C.-Y. & How, Y.-F. Effects of Temperature and Humidity on the Survival and Water Loss of Cimex hemipterus (Hemiptera: Cimicidae). *Journal of Medical Entomology* **47**, 987-995, doi:10.1603/me10018 (2014).

64   Chaves, Luis F. & Koenraadt, Constantianus J. M. Climate Change and Highland Malaria: Fresh Air for a Hot Debate. *The Quarterly Review of Biology* **85**, 27-55, doi:10.1086/650284 (2010).

65   Ngarakana-Gwasira, E. T., Bhunu, C. P., Masocha, M. & Mashonjowa, E. Assessing the Role of Climate Change in Malaria Transmission in Africa. *Malaria Research and Treatment* **2016**, 7, doi:10.1155/2016/7104291 (2016).

66   Caminade, C. *et al.* Impact of climate change on global malaria distribution. *Proceedings of the National Academy of Sciences* **111**, 3286-3291, doi:10.1073/pnas.1302089111 (2014).

67   De Silva, P. M. & Marshall, J. M. Factors Contributing to Urban Malaria Transmission in Sub-Saharan Africa: A Systematic Review. *Journal of Tropical Medicine* **2012**, 10, doi:10.1155/2012/819563 (2012).

68   Prothero, R. M. Malaria in Latin America: Environmental and Human Factors. *Bulletin of Latin American Research* **14**, 357-365, doi:10.1111/j.1470-9856.1995.tb00015.x (1995).

69   Fortin, A., Stevenson, M. M. & Gros, P. Susceptibility to malaria as a complex trait: big pressure from a tiny creature. *Human Molecular Genetics* **11**, 2469-2478, doi:10.1093/hmg/11.20.2469 (2002).

70   Unger, A. & Riley, L. W. Slum Health: From Understanding to Action. *PLOS Medicine* **4**, e295, doi:10.1371/journal.pmed.0040295 (2007).

71   Reyes, R., Ahn, R., Thurber, K. & Burke, T. F. in *Challenges in Infectious Diseases* (ed I. W. Fong) 123-146 (Springer New York, 2013).

72   Samson, D. M. *et al.* New baseline environmental assessment of mosquito ecology in northern Haiti during increased urbanization. *Journal of Vector Ecology* **40**, 46-58, doi:doi:10.1111/jvec.12131 (2015).

73   Castro, M. C. *et al.* The Importance of Drains for the Larval Development of Lymphatic Filariasis and Malaria Vectors in Dar es Salaam, United Republic of Tanzania. *PLOS Neglected Tropical Diseases* **4**, e693, doi:10.1371/journal.pntd.0000693 (2010).

74   Schmidt, W.-P. *et al.* Population Density, Water Supply, and the Risk of Dengue Fever in Vietnam: Cohort Study and Spatial Analysis. *PLOS Medicine* **8**, e1001082, doi:10.1371/journal.pmed.1001082 (2011).

75   O'Meara, W. P. *et al.* The impact of primary health care on malaria morbidity – defining access by disease burden. *Tropical Medicine & International Health* **14**, 29-35, doi:doi:10.1111/j.1365-3156.2008.02194.x (2009).

76   Hay, S. I., Guerra, C. A., Tatem, A. J., Atkinson, P. M. & Snow, R. W. Urbanization, malaria transmission and disease burden in Africa. *Nature Reviews Microbiology* **3**, 81, doi:10.1038/nrmicro1069 (2005).

77   Padilla, J. C., Chaparro, P. E., Molina, K., Arevalo-Herrera, M. & Herrera, S. Is there malaria transmission in urban settings in Colombia? *Malaria Journal* **14**, 453, doi:10.1186/s12936-015-0956-0 (2015).

78   Schicker, R. S. *et al.* A Venue-Based Survey of Malaria, Anemia and Mobility Patterns among Migrant Farm Workers in Amhara Region, Ethiopia. *PLOS ONE* **10**, e0143829, doi:10.1371/journal.pone.0143829 (2015).

79   Pindolia, D. K. *et al.* The demographics of human and malaria movement and migration patterns in East Africa. *Malaria Journal* **12**, 397, doi:10.1186/1475-2875-12-397 (2013).

80   Aschale, Y., Mengist, A., Bitew, A., Kassie, B. & Talie, A. Prevalence of malaria and associated risk factors among asymptomatic migrant laborers in West Armachiho District, Northwest Ethiopia. *Research and reports in tropical medicine* **9**, 95-101, doi:10.2147/RRTM.S165260 (2018).

81   Vajda, É. A. & Webb, C. E. Assessing the Risk Factors Associated with Malaria in the Highlands of Ethiopia: What Do We Need to Know? *Tropical Medicine and Infectious Disease* **2**, 4 (2017).

82   Council, N. R. *Malaria Control During Mass Population Movements and Natural Disasters.* (The National Academies Press, 2002).

83   HUGHES, S. & KELLY, P. Interactions of malnutrition and immune impairment, with specific reference to immunity against parasites. *Parasite Immunology* **28**, 577-588, doi:doi:10.1111/j.1365-3024.2006.00897.x (2006).

84   Shikur, B., Deressa, W. & Lindtjørn, B. Association between malaria and malnutrition among children aged under-five years in Adami Tulu District, south-central Ethiopia: a case-control study. *BMC public health* **16**, 174-174, doi:10.1186/s12889-016-2838-y (2016).

85   Tatem, A. J., Rogers, D. J. & Hay, S. I. Global transport networks and infectious disease spread. *Advances in parasitology* **62**, 293-343, doi:10.1016/S0065-308X(05)62009-X (2006).

86   Tatem, A. J., Rogers, D. J. & Hay, S. I. Estimating the malaria risk of African mosquito movement by air travel. *Malaria journal* **5**, 57-57, doi:10.1186/1475-2875-5-57 (2006).

87   Kibret, S., Lautze, J., McCartney, M., Wilson, G. G. & Nhamo, L. Malaria impact of large dams in sub-Saharan Africa: maps, estimates and predictions. *Malaria Journal* **14**, 339, doi:10.1186/s12936-015-0873-2 (2015).

88   Franco, A. O., Gomes, M. G. M., Rowland, M., Coleman, P. G. & Davies, C. R. Controlling malaria using livestock-based interventions: a one health approach. *PloS one* **9**, e101699-e101699, doi:10.1371/journal.pone.0101699 (2014).

89   Hasyim, H. *et al.* Does livestock protect from malaria or facilitate malaria prevalence? A cross-sectional study in endemic rural areas of Indonesia. *Malaria journal* **17**, 302-302, doi:10.1186/s12936-018-2447-6 (2018).

90   SOTA, T. & MOGI, M. Effectiveness of zooprophylaxis in malaria control: a theoretical inquiry, with a model for mosquito populations with two bloodmeal hosts. *Medical and Veterinary Entomology* **3**, 337-345, doi:doi:10.1111/j.1365-2915.1989.tb00240.x (1989).

91   Janko, M. M. *et al.* The links between agriculture, Anopheles mosquitoes, and malaria risk in children younger than 5 years in the Democratic Republic of the Congo: a population-based, cross-sectional, spatial study. *The Lancet. Planetary health* **2**, e74-e82, doi:10.1016/S2542-5196(18)30009-3 (2018).

92   Jarju, L. B. S. *et al.* Agriculture and the promotion of insect pests: rice cultivation in river floodplains and malaria vectors in The Gambia. *Malaria journal* **8**, 170-170, doi:10.1186/1475-2875-8-170 (2009).

93   Pulford, J., Hetzel, M. W., Bryant, M., Siba, P. M. & Mueller, I. Reported reasons for not using a mosquito net when one is available: a review of the published literature. *Malaria journal* **10**, 83-83, doi:10.1186/1475-2875-10-83 (2011).

94   Taremwa, I. M. *et al.* Knowledge, attitude and behaviour towards the use of insecticide treated mosquito nets among pregnant women and children in rural Southwestern Uganda. *BMC public health* **17**, 794-794, doi:10.1186/s12889-017-4824-4 (2017).

95   Tanner, M. *et al.* Malaria eradication and elimination: views on how to translate a vision into reality. *BMC medicine* **13**, 167-167, doi:10.1186/s12916-015-0384-6 (2015).

96   Singh, R., Musa, J., Singh, S. & Ebere, U. V. Knowledge, attitude and practices on malaria among the rural communities in aliero, northern Nigeria. *Journal of family medicine and primary care* **3**, 39-44, doi:10.4103/2249-4863.130271 (2014).

97   Peters, D. H. *et al.* Poverty and Access to Health Care in Developing Countries. *Annals of the New York Academy of Sciences* **1136**, 161-171, doi:doi:10.1196/annals.1425.011 (2008).

98   Ricci, F. Social implications of malaria and their relationships with poverty. *Mediterranean journal of hematology and infectious diseases* **4**, e2012048-e2012048, doi:10.4084/MJHID.2012.048 (2012).

99   Maiteki-Sebuguzi, C. *et al.* Community case management of malaria: exploring support, capacity and motivation of community medicine distributors in Uganda. *Health Policy and Planning* **30**, 451-461, doi:10.1093/heapol/czu033 (2014).

100   Diap, G., Amuasi, J., Boakye, I., Sevcsik, A.-M. & Pecoul, B. Anti-malarial market and policy surveys in sub-Saharan Africa. *Malaria Journal* **9**, S1, doi:10.1186/1475-2875-9-S1-S1 (2010).

101   Group, A. C., Michael, D. & Mkunde, S. P. The malaria testing and treatment landscape in mainland Tanzania, 2016. *Malaria journal* **16**, 202-202, doi:10.1186/s12936-017-1819-7 (2017).

102   Tipke, M. *et al.* Substandard anti-malarial drugs in Burkina Faso. *Malaria Journal* **7**, 95, doi:10.1186/1475-2875-7-95 (2008).

103   Medicine, I. o. *Saving Lives, Buying Time: Economics of Malaria Drugs in an Age of Resistance*. (The National Academies Press, 2004).

104   Eckhoff, P. A. A malaria transmission-directed model of mosquito life cycle and ecology. *Malaria journal* **10**, 303-303, doi:10.1186/1475-2875-10-303 (2011).

105   Smith, R. C., Vega-Rodríguez, J. & Jacobs-Lorena, M. The Plasmodium bottleneck: malaria parasite losses in the mosquito vector. *Memorias do Instituto Oswaldo Cruz* **109**, 644-661, doi:10.1590/0074-0276130597 (2014).

106   Orfano, A. S. *et al.* Species-specific escape of Plasmodium sporozoites from oocysts of avian, rodent, and human malarial parasites. *Malaria Journal* **15**, 394, doi:10.1186/s12936-016-1451-y (2016).

107   Carter, R. & Mendis, K. N. Evolutionary and Historical Aspects of the Burden of Malaria. *Clinical Microbiology Reviews* **15**, 564-594, doi:10.1128/cmr.15.4.564-594.2002 (2002).

108   Lozano-Fuentes, S. *et al.* The dengue virus mosquito vector Aedes aegypti at high elevation in Mexico. *The American journal of tropical medicine and hygiene* **87**, 902-909, doi:10.4269/ajtmh.2012.12-0244 (2012).

109   Coffey, L. L., Failloux, A.-B. & Weaver, S. C. Chikungunya virus-vector interactions. *Viruses* **6**, 4628-4663, doi:10.3390/v6114628 (2014).

110   Harbach, R. *The Culicidae (Diptera): A Review Of Taxonomy, Classification And Phylogeny *. Vol. 1668 (2007).

111   Chaves, L. F., Harrington, L. C., Keogh, C. L., Nguyen, A. M. & Kitron, U. D. Blood feeding patterns of mosquitoes: random or structured? *Frontiers in zoology* **7**, 3-3, doi:10.1186/1742-9994-7-3 (2010).

112   Braack, L. *et al.* Biting behaviour of African malaria vectors: 1. where do the main vector species bite on the human body? *Parasites & vectors* **8**, 76-76, doi:10.1186/s13071-015-0677-9 (2015).

113   Dida, G. O. *et al.* Presence and distribution of mosquito larvae predators and factors influencing their abundance along the Mara River, Kenya and Tanzania. *SpringerPlus* **4**, 136-136, doi:10.1186/s40064-015-0905-y (2015).

114   Dejenie, T., Yohannes, M. & Assmelash, T. Characterization of mosquito breeding sites in and in the vicinity of tigray microdams. *Ethiopian journal of health sciences* **21**, 57-66 (2011).

115   Gunathilaka, N., Abeyewickreme, W., Hapugoda, M. & Wickremasinghe, R. Species Composition and Diversity of Malaria Vector Breeding Habitats in Trincomalee District of Sri Lanka. *BioMed research international* **2015**, 823810-823810, doi:10.1155/2015/823810 (2015).

116   Lehmann, T. *et al.* Aestivation of the African malaria mosquito, Anopheles gambiae in the Sahel. *The American journal of tropical medicine and hygiene* **83**, 601-606, doi:10.4269/ajtmh.2010.09-0779 (2010).

117   Ortiz-Ruiz, A. *et al.* Plasmodium species differentiation by non-expert on-line volunteers for remote malaria field diagnosis. *Malaria journal* **17**, 54-54, doi:10.1186/s12936-018-2194-8 (2018).

118   Zimmerman, P. A., Mehlotra, R. K., Kasehagen, L. J. & Kazura, J. W. Why do we need to know more about mixed Plasmodium species infections in humans? *Trends in parasitology* **20**, 440-447, doi:10.1016/j.pt.2004.07.004 (2004).

119   Müller, I. *et al.* Three different Plasmodium species show similar patterns of clinical tolerance of malaria infection. *Malaria journal* **8**, 158-158, doi:10.1186/1475-2875-8-158 (2009).

120   Nanoti, A., Jain, S., Gupta, C. & Vyas, G. in *2016 International Conference on Inventive Computation Technologies (ICICT).* 1-6.

121   Cowman, A. F. & Crabb, B. S. Invasion of Red Blood Cells by Malaria Parasites. *Cell* **124**, 755-766, doi:https://doi.org/10.1016/j.cell.2006.02.006 (2006).

122   Medicine, I. o. *Malaria: Obstacles and Opportunities*. (The National Academies Press, 1991).

123   Cowman, A. F., Berry, D. & Baum, J. The cellular and molecular basis for malaria parasite invasion of the human red blood cell. *The Journal of Cell Biology* **198**, 961-971, doi:10.1083/jcb.201206112 (2012).

124   Sherling, E. S. *et al.* The Plasmodium falciparum rhoptry protein RhopH3 plays essential roles in host cell invasion and nutrient uptake. *eLife* **6**, e23239, doi:10.7554/eLife.23239 (2017).

125   *Disease and Mortality in Sub-Saharan Africa. 2nd Edition*. (World Bank., 2006).

126   Bird, W. A. Net Loss? Agrochemicals and Insecticide Resistance in the Fight against Malaria. *Environmental health perspectives* **125**, A50-A57, doi:10.1289/ehp.125-A50 (2017).

127   Liu, N. Insecticide Resistance in Mosquitoes: Impact, Mechanisms, and Research Directions. *Annual Review of Entomology* **60**, 537-559, doi:10.1146/annurev-ento-010814-020828 (2015).

128   Fouet, C., Atkinson, P. & Kamdem, C. Human Interventions: Driving Forces of Mosquito Evolution. *Trends in Parasitology* **34**, 127-139, doi:10.1016/j.pt.2017.10.012 (2018).

129   White, N. J. Antimalarial drug resistance. *The Journal of clinical investigation* **113**, 1084-1092, doi:10.1172/JCI21682 (2004).

130   Legros, M. & Bonhoeffer, S. Evolution of drug resistance in malaria parasites. *Malaria Journal* **11**, P63, doi:10.1186/1475-2875-11-S1-P63 (2012).

131 Imwong, M., Hien, T. T., Thuy-Nhien, N. T., Dondorp, A. M. & White, N. J. Spread of a single multidrug resistant malaria parasite lineage (PfPailin) to Vietnam. *The Lancet Infectious Diseases* **17**, 1022-1023, doi:10.1016/S1473-3099(17)30524-8 (2017).

132 Pongtavornpinyo, W. *et al.* Spread of anti-malarial drug resistance: Mathematical model with implications for ACT drug policies. *Malaria Journal* **7**, 229, doi:10.1186/1475-2875-7-229 (2008).

133 Zuber, J. A. & Takala-Harrison, S. Multidrug-resistant malaria and the impact of mass drug administration. *Infection and drug resistance* **11**, 299-306, doi:10.2147/IDR.S123887 (2018).

134 WHO. Basic malaria microscopy – Part I: Learner's guide. Second edition. *World Health Orgaization* (2010).

135 Luzzatto, L. Genetic factors in malaria. *Bulletin of the World Health Organization* **50**, 195-202 (1974).

136 Driss, A. *et al.* Genetic polymorphisms linked to susceptibility to malaria. *Malaria journal* **10**, 271-271, doi:10.1186/1475-2875-10-271 (2011).

137 de Mendonca, V. R. R., Goncalves, M. S. & Barral-Netto, M. The Host Genetic Diversity in Malaria Infection. *Journal of Tropical Medicine* **2012**, 17, doi:10.1155/2012/940616 (2012).

138 Hedrick, P. W. Population genetics of malaria resistance in humans. *Heredity* **107**, 283-304, doi:10.1038/hdy.2011.16 (2011).

139 Doolan, D. L., Dobaño, C. & Baird, J. K. Acquired immunity to malaria. *Clinical microbiology reviews* **22**, 13-36, doi:10.1128/CMR.00025-08 (2009).

140 Eridani, S. Sickle cell protection from malaria. *Hematology reports* **3**, e24-e24, doi:10.4081/hr.2011.e24 (2011).

141 Mbanefo, E. C. *et al.* Association of glucose-6-phosphate dehydrogenase deficiency and malaria: a systematic review and meta-analysis. *Scientific reports* **7**, 45963-45963, doi:10.1038/srep45963 (2017).

142 Orlando Bordin, J. Duffy blood group and malaria AU - Langhi, Dante M. *Hematology* **11**, 389-398, doi:10.1080/10245330500469841 (2006).

143 Ntumngia, F. B., Thomson-Luque, R., Pires, C. V. & Adams, J. H. The role of the human Duffy antigen receptor for chemokines in malaria susceptibility: current opinions and future treatment prospects. *Journal of receptor, ligand and channel research* **9**, 1-11, doi:10.2147/JRLCR.S99725 (2016).

144 Zimmerman, P. A., Ferreira, M. U., Howes, R. E. & Mercereau-Puijalon, O. Red blood cell polymorphism and susceptibility to Plasmodium vivax. *Advances in parasitology* **81**, 27-76, doi:10.1016/B978-0-12-407826-0.00002-3 (2013).

145 Gunalan, K. *et al.* Role of Plasmodium vivax Duffy-binding protein 1 in invasion of Duffy-null Africans. *Proceedings of the National Academy of Sciences of the United States of America* **113**, 6271-6276, doi:10.1073/pnas.1606113113 (2016).

146 Crux, N. B. & Elahi, S. Human Leukocyte Antigen (HLA) and Immune Regulation: How Do Classical and Non-Classical HLA Alleles Modulate Immune Response to Human Immunodeficiency Virus and Hepatitis C Virus Infections? *Frontiers in immunology* **8**, 832-832, doi:10.3389/fimmu.2017.00832 (2017).

147 Dechavanne, C. *et al.* Acquisition of natural humoral immunity to P. falciparum in early life in Benin: impact of clinical, environmental and host factors. *Scientific reports* **6**, 33961-33961, doi:10.1038/srep33961 (2016).

148 Arbeitskreis Blut, U. B. B. K. Malaria. *Transfusion medicine and hemotherapy : offizielles Organ der Deutschen Gesellschaft fur Transfusionsmedizin und Immunhamatologie* **36**, 48-60, doi:10.1159/000197327 (2009).

149 Recht, J. *et al.* Malaria in Brazil, Colombia, Peru and Venezuela: current challenges in malaria control and elimination. *Malaria journal* **16**, 273-273, doi:10.1186/s12936-017-1925-6 (2017).

150   Gomes, P. S., Bhardwaj, J., Rivera-Correa, J., Freire-De-Lima, C. G. & Morrot, A. Immune Escape Strategies of Malaria Parasites. *Frontiers in microbiology* **7**, 1617-1617, doi:10.3389/fmicb.2016.01617 (2016).

151   Noble, C., Adlam, B., Church, G. M., Esvelt, K. M. & Nowak, M. A. Current CRISPR gene drive systems are likely to be highly invasive in wild populations. *eLife* **7**, e33423, doi:10.7554/eLife.33423 (2018).

152   Drury, D. W., Dapper, A. L., Siniard, D. J., Zentner, G. E. & Wade, M. J. CRISPR/Cas9 gene drives in genetically variable and nonrandomly mating wild populations. *Science Advances* **3**, e1601910, doi:10.1126/sciadv.1601910 (2017).

153   Smidler, A. L., Najjar, D. & Esvelt, K. M. Harnessing gene drive AU - Min, John. *Journal of Responsible Innovation* **5**, S40-S65, doi:10.1080/23299460.2017.1415586 (2018).

154   Min, J., Noble, C., Najjar, D. & Esvelt, K. Daisy quorum drives for the genetic restoration of wild populations. *bioRxiv*, 115618, doi:10.1101/115618 (2017).

155   Sarkar, S. Researchers Hit Roadblocks with Gene Drives. *BioScience* **68**, 474-480, doi:10.1093/biosci/biy060 (2018).

156   Callaway, E. *Gene drives thwarted by emergence of resistant organisms*. Vol. 542 (2017).

157   Miles, A. *et al.* Natural diversity of the malaria vector Anopheles gambiae. *bioRxiv*, 096289, doi:10.1101/096289 (2016).

158   ISO 19107:2003 Geographic information - Spatial schema. *International Organization for Standardization* (2003).

159   Barnaghi, J. T. L. v. d. B. P. Spatial Data on the Web Best Practices. *W3C Working Group Note* (2017).

160   Lazoglou, M. & Angelides, D. Development of an Ontology for Modeling Spatial Planning Systems. *Current Urban Studies* **04**, 241-255, doi:10.4236/cus.2016.42016 (2016).

161   EMBL. BioModels. *European Molecular Biology Laboratory* http://www.ebi.ac.uk/biomodels/ (2019).

162   de Lorenzo, V. & Danchin, A. Synthetic biology: discovering new worlds and new words. *EMBO reports* **9**, 822-827, doi:10.1038/embor.2008.159 (2008).

163   Andrén, H. Effects of Habitat Fragmentation on Birds and Mammals in Landscapes with Different Proportions of Suitable Habitat: A Review. *Oikos* **71**, 355-366, doi:10.2307/3545823 (1994).

164   Reiner Robert, C. *et al.* A systematic review of mathematical models of mosquito-borne pathogen transmission: 1970–2010. *Journal of The Royal Society Interface* **10**, 20120921, doi:10.1098/rsif.2012.0921 (2013).

165   Ross, R. S. *Report and the prevention of malaria in Mauritius / by Ronald Ross*. (Waterlow & Son, 1908).

166   Ross, R. S. *The Prevention of Malaria / by Ronald Ross*. (Murray, 1910).

167   Edgington, M. & Alphey, L. Modeling the mutation and reversal of engineered underdominance gene drives. *bioRxiv*, 257253, doi:10.1101/257253 (2018).

168   Sanchez C., H. M., Wu, S. L., Bennett, J. B. & Marshall, J. M. MGDrivE: A modular simulation framework for the spread of gene drives through spatially-explicit mosquito populations. *bioRxiv*, 350488, doi:10.1101/350488 (2018).

169   Marshall, J. M. & Hay, B. A. Medusa: A Novel Gene Drive System for Confined Suppression of Insect Populations. *PLOS ONE* **9**, e102694, doi:10.1371/journal.pone.0102694 (2014).

170   Deredec, A., Godfray, H. C. J. & Burt, A. Requirements for effective malaria control with homing endonuclease genes. *Proceedings of the National Academy of Sciences* **108**, E874-E880, doi:10.1073/pnas.1110717108 (2011).

171   Hancock, P. A. & Godfray, H. C. J. Application of the lumped age-class technique to studying the dynamics of malaria-mosquito-human interactions. *Malaria Journal* **6**, 98, doi:10.1186/1475-2875-6-98 (2007).

172   in *The Concise Encyclopedia of Statistics*   502-505 (Springer New York, 2008).

173   Ramsey, P. H. Critical Values for Spearman's Rank Order Correlation. *Journal of Educational Statistics* **14**, 245-253, doi:10.2307/1165017 (1989).

174   Dixon, W. J. & Mood, A. M. The Statistical Sign Test. *Journal of the American Statistical Association* **41**, 557-566, doi:10.2307/2280577 (1946).

175   Laudański, L. M. in *Between Certainty and Uncertainty: Statistics and Probability in Five Units with Notes on Historical Origins and Illustrative Numerical Examples* (ed Ludomir M. Laudański) 87-127 (Springer Berlin Heidelberg, 2013).

176   Molofsky, J. & Ferdy, J.-B. Extinction dynamics in experimental metapopulations. *Proceedings of the National Academy of Sciences of the United States of America* **102**, 3726-3731, doi:10.1073/pnas.0404576102 (2005).

177   Verboom, J., Foppen, R., Chardon, P., Opdam, P. & Luttikhuizen, P. Introducing the key patch approach for habitat networks with persistent populations: an example for marshland birds. *Biological Conservation* **100**, 89-101, doi:https://doi.org/10.1016/S0006-3207(00)00210-X (2001).

178   Opdam, P. & Steingröver, E. Designing Metropolitan Landscapes for Biodiversity: Deriving Guidelines from Metapopulation Ecology. *Landscape Journal* **27**, 69-80 (2008).

179   Zeigler, S. L., De Vleeschouwer, K. M. & Raboy, B. E. Assessing Extinction Risk in Small Metapopulations of Golden-headed Lion Tamarins (Leontopithecus chrysomelas) in Bahia State, Brazil. *Biotropica* **45**, 528-535, doi:10.1111/btp.12037 (2013).

180   van der Sluis, T., Pedroli, B. & Kuipers, H. *Corridors for LIFE; ecological network analysis Regione Emilia-Romagna - the plains of Provincia di Modena & Bologna.* (2001).
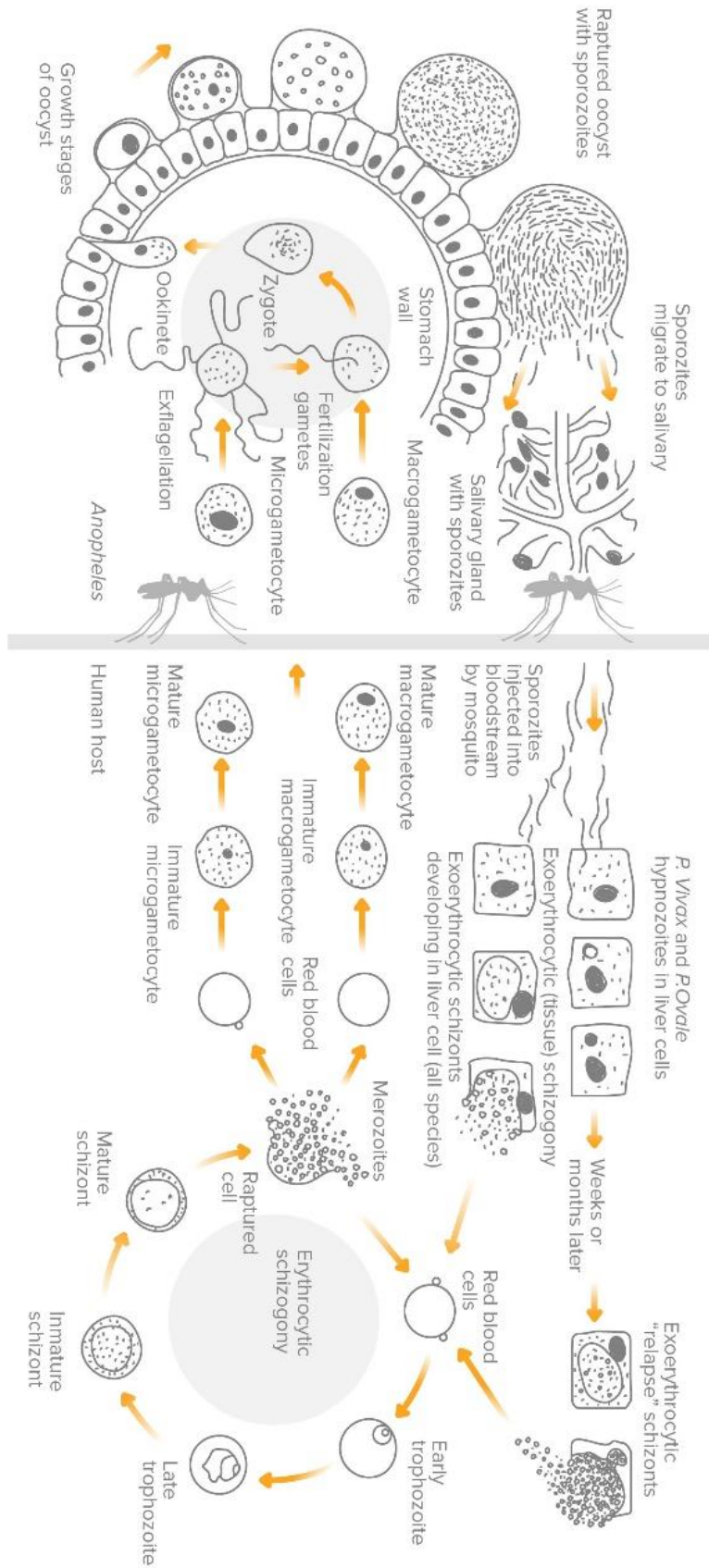
Appendix A

Figure A1 *Developmental stages of malaria parasites (Plasmodium species). Reproduced from World Health Organization.*[134]
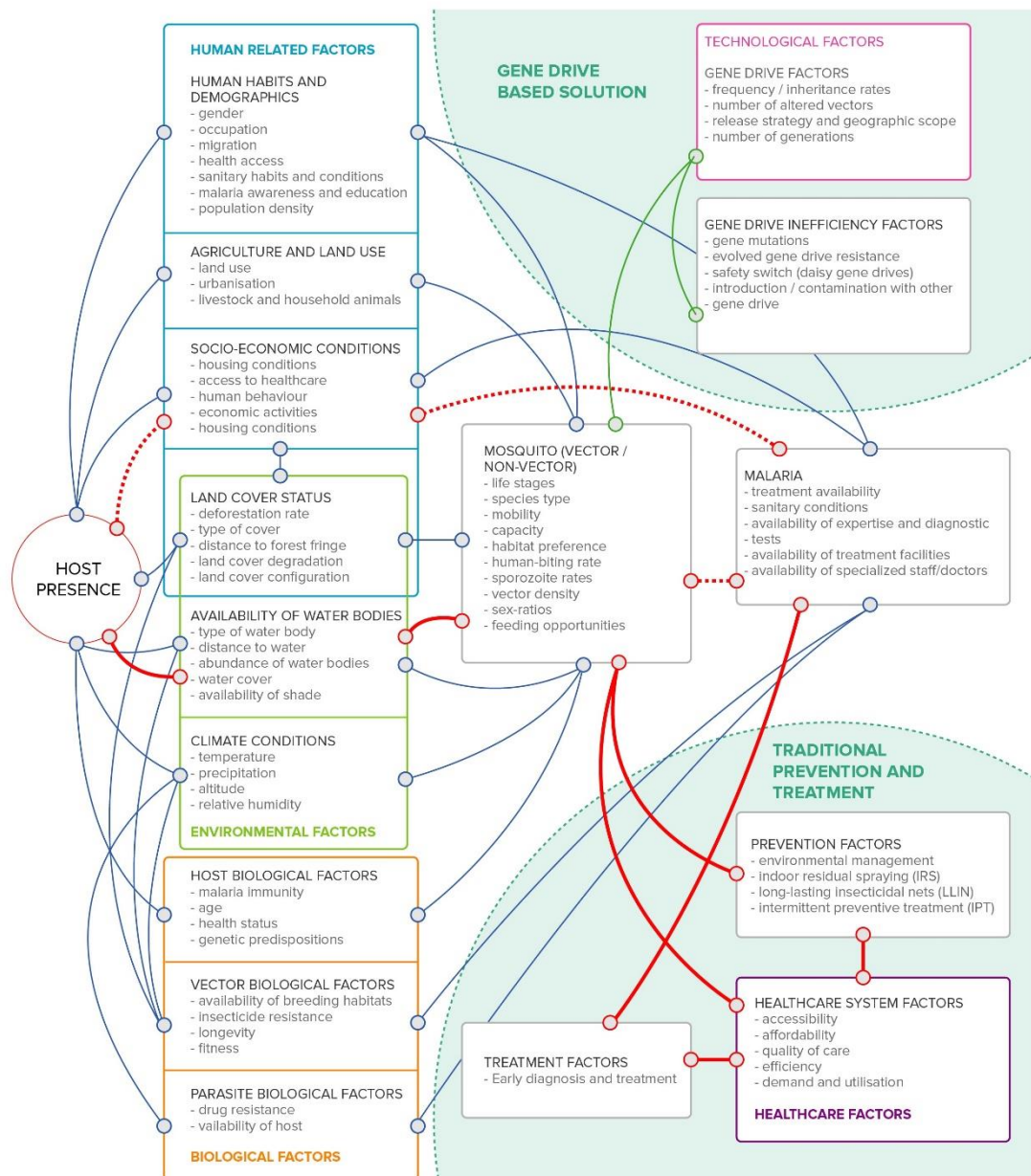
67

Figure A2 *Expanded conceptual diagram of factors related to malaria spread with gene drives introduced to the system. Boxes correspond to specific factors responsible for malaria spread. Circles group various factors responsible for the traditional prevention and treatment of malaria and factors for gene drive-based solution. Lines and dots represent relations between different factors. The author created the figure based on a literature review of various factors responsible for malaria spread. All presented factors have been described in the main text of this document.*

Appendix B

Table B1 *Database of 36 major cities in Kenya with their latitude, longitude, and population.*

| City	Latitude	Longitude	Country	Administrative division	Population
Nairobi	-1.28333	36.81667	Kenya	Nairobi City	3,010,000
Mombasa	-4.05466	39.66359	Kenya	Mombasa	882,000
Kisumu	-0.10221	34.76171	Kenya	Kisumu	395,615
Nakuru	-0.28333	36.06667	Kenya	Nakuru	364,727
Eldoret	0.520356	35.26993	Kenya	Uasin Gishu	353,381
Kitale	1.015725	35.00622	Kenya	Trans Nzoia	150,495
Machakos	-1.51667	37.26667	Kenya	Machakos	144,925
Thika	-1.05	37.08333	Kenya	Kiambu	99,322
Kericho	-0.36774	35.28314	Kenya	Kericho	98,852
Malindi	-3.21748	40.1191	Kenya	Kilifi	94,016
Kendu Bay	-0.35892	34.64924	Kenya	Homa Bay	91,248
Kilifi	-3.63045	39.84992	Kenya	Kilifi	80,339
Sotik	-0.68333	35.11871	Kenya	Kericho	71,285
Garissa	-0.45355	39.64011	Kenya	Garissa	67,861
Kakamega	0.284219	34.75229	Kenya	Kakamega	63,426
Embu	-0.53112	37.45061	Kenya	Embu	58,620
Bungoma	0.563504	34.56055	Kenya	Bungoma	55,962
Nyeri	-0.42013	36.94759	Kenya	Nyeri	51,084
Meru	0.046256	37.65587	Kenya	Meru	47,226
Wajir	1.747102	40.05732	Kenya	Wajir	45,771
Naivasha	-0.71667	36.43591	Kenya	Nakuru	43,983
Voi	-3.39452	38.56304	Kenya	Taita/Taveta	36,487
Nanyuki	0.016667	37.07283	Kenya	Laikipia	36,142
Kisii	-0.67394	34.77225	Kenya	Kisii	28,547
Moyale	3.5167	39.05842	Kenya	Marsabit	24,837
Lamu	-2.26925	40.89915	Kenya	Lamu	24,525
Maralal	1.09679	36.69799	Kenya	Samburu	20,841
Lodwar	3.119881	35.59642	Kenya	Turkana	20,219
Eldama Ravine	0.051578	35.73078	Kenya	Baringo	17,581
Marsabit	2.328394	37.98986	Kenya	Marsabit	16,460
Namanga	-2.54327	36.79053	Kenya	Kajiado	13,193
Mwingi	-0.93435	38.06005	Kenya	Kitui	11,219
Witu	-2.38886	40.43821	Kenya	Lamu	5,380
Karungu	-0.84641	34.15479	Kenya	Migori	2,376
Konza	-1.7426	37.12941	Kenya	Machakos	2,004
Tsavo	-2.99208	38.46188	Kenya	Taita/Taveta	414

Appendix C
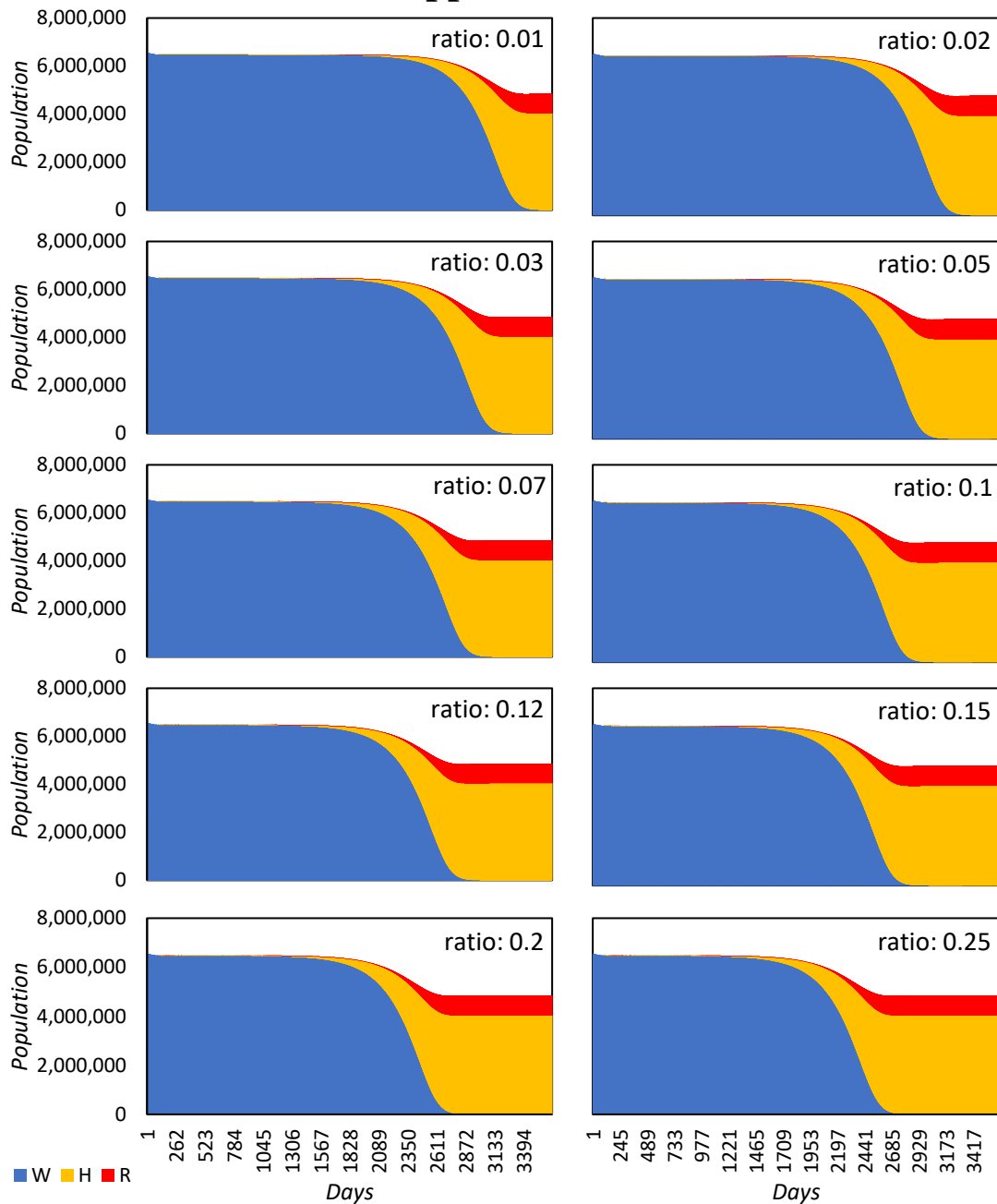
Figure C1 *Examples of MGDrivE simulations for CRISPR-based homing constructs. Simulation parameters were unchanged and are the same as described in the main body of the thesis, except for the parameter responsible for the ratio between a number of introduced gene drive organisms to the wild population. The ratio is listed on the plot for each of the simulations. Red color corresponds to organisms that developed the resistance to gene drive (R), orange is a gene drive (homing gene) population (H), and blue corresponds to wild population (W). The inheritance frequency (eM) was fixed at 0.9.*
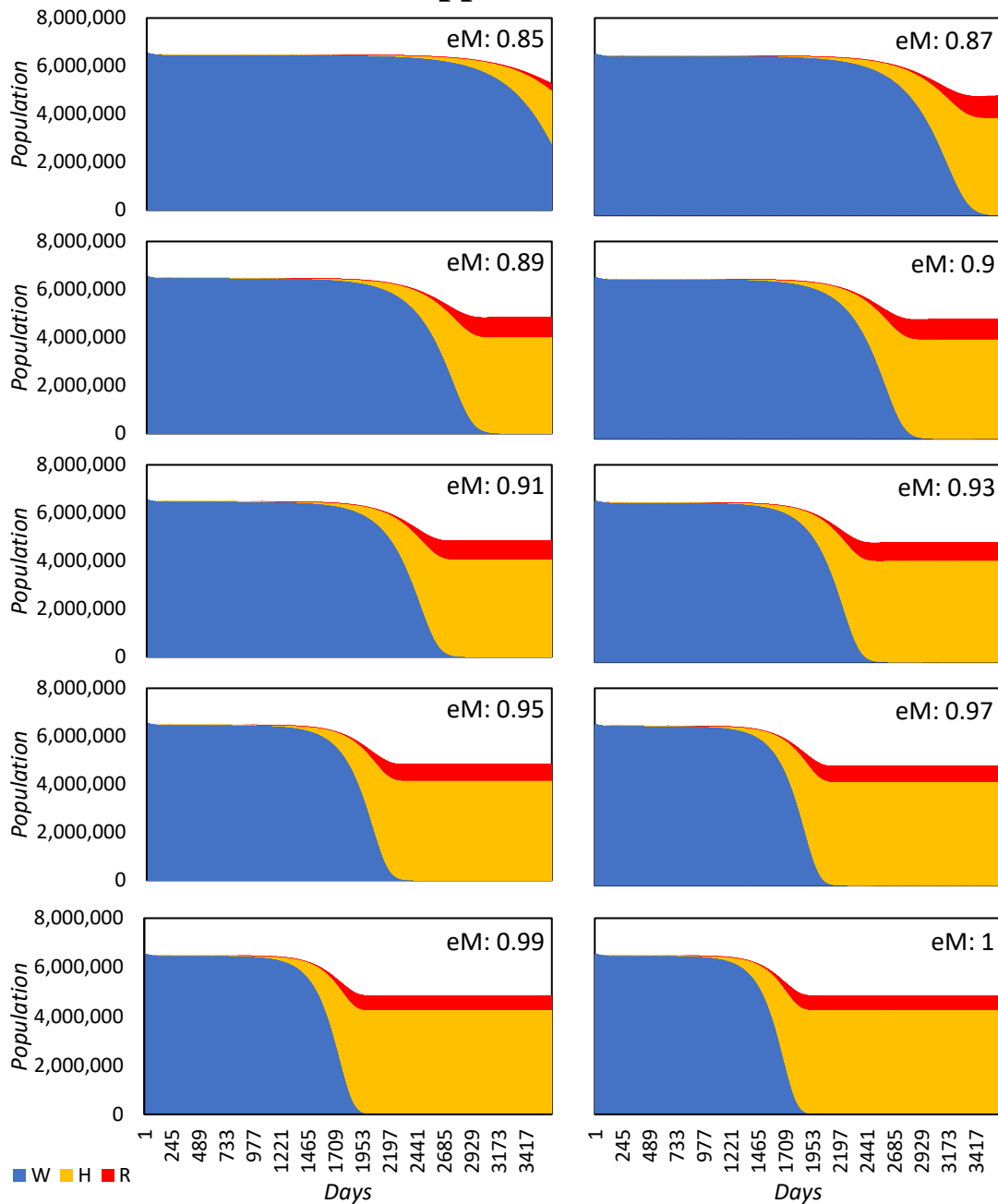
Appendix D

Figure D1 *Examples of MGDrivE simulations for CRISPR-based homing constructs. Simulation parameters were unchanged and are the same as described in the main body of the thesis, except for male gene drive frequency parameter (eM) responsible for the frequency of gene drive inheritance onto the offspring. The frequency is listed on the plot for each of the simulations. Red color corresponds to organisms that developed the resistance to gene drive (R), orange is a gene drive (homing gene) population (H), and blue corresponds to wild population (W). The initial number of gene drive mosquitoes as a percentage of wild population was fixed at 1%.*

Appendix E

To further evaluate how input parameters affect mosquitoes' distribution and the time required for gene drive mosquitoes to replace the wild population, additional modelling experiments have been developed.

Instead of releasing gene drive mosquitoes in one city only (Malindi), mosquitoes have been released in four different cities, including Malindi, Wajr, Lodwar, and Karungu. The results have been then compared to gene drive mosquitoes results from a single release point (Malindi). The percentage of gene drive altered mosquitoes released at Malindi to wold mosquito population was 3%. Because the assumption is made that the mosquito population in each of the cities equals the human population in those cities, the initial number of gene drive mosquitoes released at Malindi was 2,820. For the other case, where gene drive mosquitoes are released from four cities, the total number of gene drive mosquitoes released was kept the same. The number of gene drive mosquitoes released from those four nodes was 705, giving a total of 2,820. The results of this experiment are shown in Figure E1.
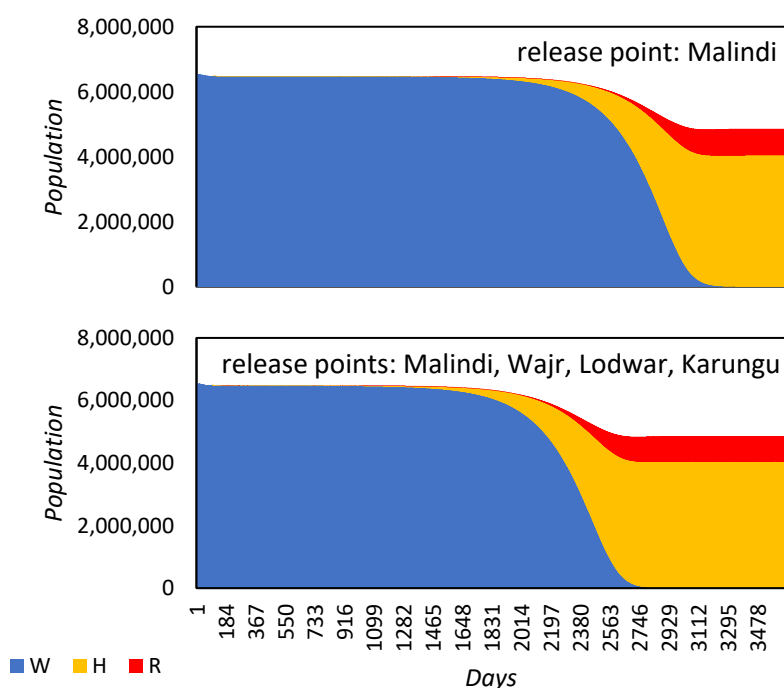

Figure E1 *Examples of MGDrivE simulations for CRISPR-based homing constructs. The comparison of gene drive release from one location versus gene drive mosquito release from four different locations. The total number of gene drive mosquitoes release in each of the cases was 2,820. In the scenario where gene drive mosquitoes are released from four different locations, the number of gene drive mosquitoes released at each location was equal (705 mosquitoes). Red color corresponds to organisms that developed the resistance to gene drive (R), orange is a gene drive (homing gene) population (H), and blue corresponds to wild population (W). The inheritance frequency (eM) was fixed at 0.9.*

The results shown in Figure E1 indicate that it is more efficient to release the same number of gene drive altered mosquitoes from multiple locations than from one single location. Release of gene drive mosquitoes from multiple locations reduces the time required for gene drive mosquitoes to replace the wild mosquito population.

Further uncertainty analysis looked at the influence of the release point on mosquito numbers. The results of this exercise are shown in Figure E2. In the first simulation, 2,820 mosquitoes (3% of Malindi population) have been released at Malindi. In two subsequent simulations, the same number of mosquitoes have been released in Nairobi and Witu. The reason behind selecting Nairobi is that it is the largest and centrally located city in Kenya. While Witu is significantly smaller yet centrally located. The initial release of 2,820 gene drive mosquitoes is significant compared to Witu's population of 5,380.
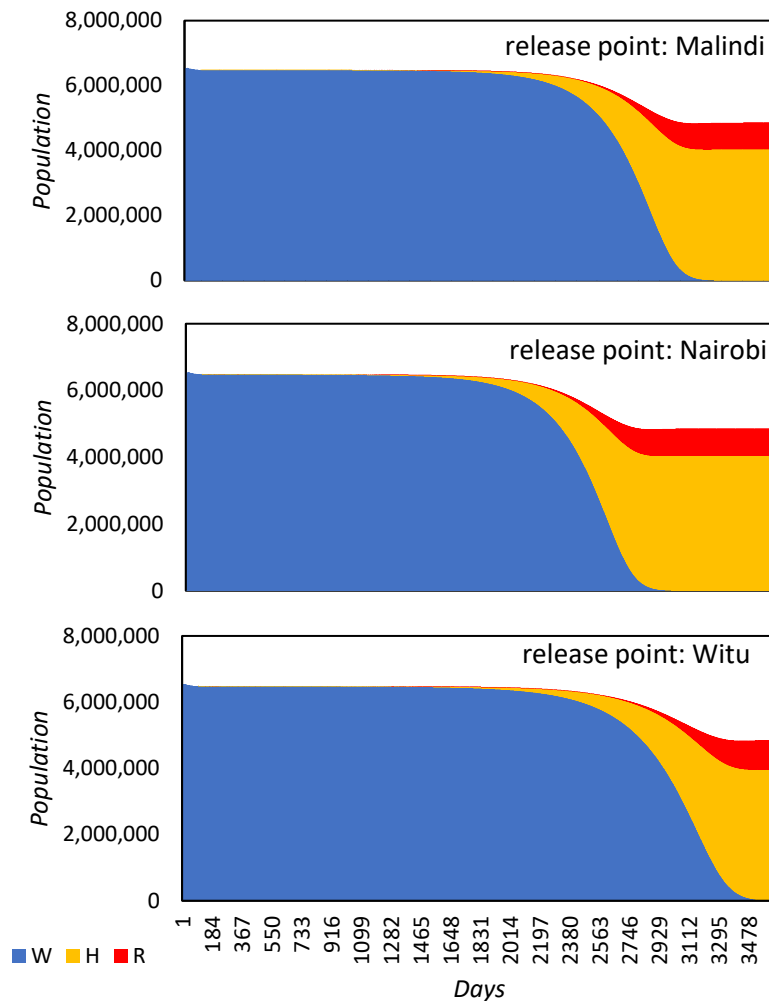

Figure E2 Uncertainty analysis of MGDrivE simulations for CRISPR-based homing constructs. The comparison of gene drive release from different locations and its influence on time required to replace wild mosquito populations. The total number of gene drive mosquitoes release in each of the simulations was 2,820. Red color corresponds to organisms that developed the resistance to gene drive (R), orange is a gene drive (homing gene) population (H), and blue corresponds to wild population (W). The inheritance frequency (eM) was fixed at 0.9.

Sensitivity analysis presented in Figure E2 indicates that release point significantly impacts the time required for gene drive mosquitoes to replace the wild mosquito population. Nairobi is large and centrally located. Malindi is of a smaller population and located on the coast, further away from other nodes. Witu is a small size town that is centrally located. This impact is, however, convoluted. The initial number of gene drive mosquitoes released in different cities was the same, while population-wise, those cities are of significantly different population size. Due to population size

and central location release of gene drive modified mosquitoes in Nairobi reduces the time required for gene drive mosquitoes to spread within the study area and replace the wild population.

The final uncertainty analysis relies on the release of gene drive mosquitoes in Malindi and a comparison of time required for gene drive mosquitoes to spread to cities that are further away, such as Nairobi and Lodwar. This case is presented in Figure E3. Please note that the population axis in each of the cities has a different scale.
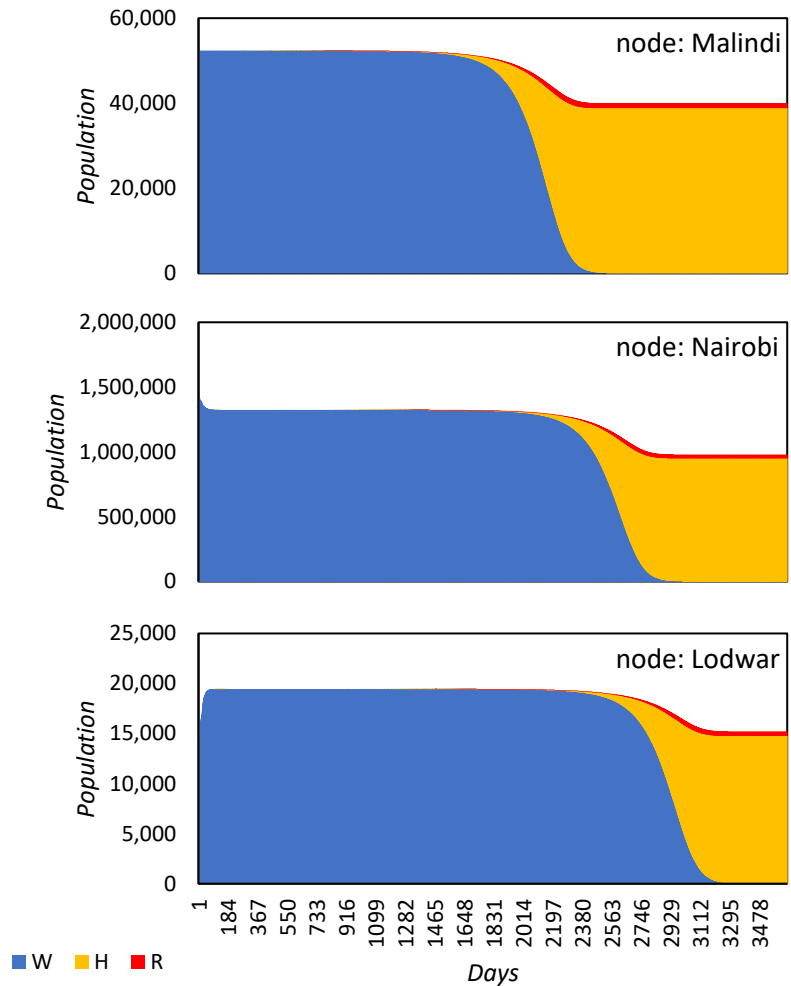

Figure E3 *Uncertainty analysis of MGDrivE simulations for CRISPR-based homing constructs. The comparison of gene drive release from Malindi to compare time required for gene drive mosquitoes to reach other cities (Nairobi, Lodwar). Red color corresponds to organisms that developed the resistance to gene drive (R), orange is a gene drive (homing gene) population (H), and blue corresponds to wild population (W). The inheritance frequency (eM) was fixed at 0.9. The percentage ratio of gene drive mosquitoes released at Malindi to wild mosquitoes at Malindi was fixed at 3%.*

The final uncertainty analysis clearly illustrates that based on the distance between simulation nodes (cities), the spread of gene drive mosquitoes is slower in the cities further away from the release location.

Series from Lund University
Department of Physical Geography and Ecosystem Science

Master Thesis in Geographical Information Science

1. *Anthony Lawther:* The application of GIS-based binary logistic regression for slope failure susceptibility mapping in the Western Grampian Mountains, Scotland (2008).
2. *Rickard Hansen:* Daily mobility in Grenoble Metropolitan Region, France. Applied GIS methods in time geographical research (2008).
3. *Emil Bayramov:* Environmental monitoring of bio-restoration activities using GIS and Remote Sensing (2009).
4. *Rafael Villarreal Pacheco:* Applications of Geographic Information Systems as an analytical and visualization tool for mass real estate valuation: a case study of Fontibon District, Bogota, Columbia (2009).
5. *Siri Oestreich Waage:* a case study of route solving for oversized transport: The use of GIS functionalities in transport of transformers, as part of maintaining a reliable power infrastructure (2010).
6. *Edgar Pimiento:* Shallow landslide susceptibility – Modelling and validation (2010).
7. *Martina Schäfer:* Near real-time mapping of floodwater mosquito breeding sites using aerial photographs (2010).
8. *August Pieter van Waarden-Nagel:* Land use evaluation to assess the outcome of the programme of rehabilitation measures for the river Rhine in the Netherlands (2010).
9. *Samira Muhammad:* Development and implementation of air quality data mart for Ontario, Canada: A case study of air quality in Ontario using OLAP tool. (2010).
10. *Fredros Oketch Okumu*: Using remotely sensed data to explore spatial and temporal relationships between photosynthetic productivity of vegetation and malaria transmission intensities in selected parts of Africa (2011).
11. *Svajunas Plunge:* Advanced decision support methods for solving diffuse water pollution problems (2011).
12. *Jonathan Higgins:* Monitoring urban growth in greater Lagos: A case study using GIS to monitor the urban growth of Lagos 1990 - 2008 and produce future growth prospects for the city (2011).
13. *Mårten Karlberg:* Mobile Map Client API: Design and Implementation for Android (2011).
14. *Jeanette McBride:* Mapping Chicago area urban tree canopy using color infrared imagery (2011).
15. *Andrew Farina:* Exploring the relationship between land surface temperature and vegetation abundance for urban heat island mitigation in Seville, Spain (2011).
16. *David Kanyari*: Nairobi City Journey Planner: An online and a Mobile Application (2011).
17. *Laura V. Drews:* Multi-criteria GIS analysis for siting of small wind power plants - A case study from Berlin (2012).
18. *Qaisar Nadeem:* Best living neighborhood in the city - A GIS based multi criteria evaluation of ArRiyadh City (2012).

19. *Ahmed Mohamed El Saeid Mustafa:* Development of a photo voltaic building rooftop integration analysis tool for GIS for Dokki District, Cairo, Egypt (2012).
20. *Daniel Patrick Taylor*: Eastern Oyster Aquaculture: Estuarine Remediation via Site Suitability and Spatially Explicit Carrying Capacity Modeling in Virginia's Chesapeake Bay (2013).
21. *Angeleta Oveta Wilson:* A Participatory GIS approach to *unearthing* Manchester's Cultural Heritage '*gold mine'* (2013).
22. *Ola Svensson:* Visibility and Tholos Tombs in the Messenian Landscape: A Comparative Case Study of the Pylian Hinterlands and the Soulima Valley (2013).
23. *Monika Ogden:* Land use impact on water quality in two river systems in South Africa (2013).
24. *Stefan Rova:* A GIS based approach assessing phosphorus load impact on Lake Flaten in Salem, Sweden (2013).
25. *Yann Buhot:* Analysis of the history of landscape changes over a period of 200 years. How can we predict past landscape pattern scenario and the impact on habitat diversity? (2013).
26. *Christina Fotiou:* Evaluating habitat suitability and spectral heterogeneity models to predict weed species presence (2014).
27. *Inese Linuza:* Accuracy Assessment in Glacier Change Analysis (2014).
28. *Agnieszka Griffin:* Domestic energy consumption and social living standards: a GIS analysis within the Greater London Authority area (2014).
29. *Brynja Guðmundsdóttir:* Detection of potential arable land with remote sensing and GIS - A Case Study for Kjósarhreppur (2014).
30. *Oleksandr Nekrasov:* Processing of MODIS Vegetation Indices for analysis of agricultural droughts in the southern Ukraine between the years 2000-2012 (2014).
31. *Sarah Tressel:* Recommendations for a polar Earth science portal in the context of Arctic Spatial Data Infrastructure (2014).
32. *Caroline Gevaert:* Combining Hyperspectral UAV and Multispectral Formosat-2 Imagery for Precision Agriculture Applications (2014).
33. *Salem Jamal-Uddeen:* Using GeoTools to implement the multi-criteria evaluation analysis - weighted linear combination model (2014).
34. *Samanah Seyedi-Shandiz:* Schematic representation of geographical railway network at the Swedish Transport Administration (2014).
35. *Kazi Masel Ullah:* Urban Land-use planning using Geographical Information System and analytical hierarchy process: case study Dhaka City (2014).
36. *Alexia Chang-Wailing Spitteler:* Development of a web application based on MCDA and GIS for the decision support of river and floodplain rehabilitation projects (2014).
37. *Alessandro De Martino:* Geographic accessibility analysis and evaluation of potential changes to the public transportation system in the City of Milan (2014).
38. *Alireza Mollasalehi:* GIS Based Modelling for Fuel Reduction Using Controlled Burn in Australia. Case Study: Logan City, QLD (2015).
39. *Negin A. Sanati:* Chronic Kidney Disease Mortality in Costa Rica; Geographical Distribution, Spatial Analysis and Non-traditional Risk Factors (2015).

40. *Karen McIntyre:* Benthic mapping of the Bluefields Bay fish sanctuary, Jamaica (2015).
41. *Kees van Duijvendijk:* Feasibility of a low-cost weather sensor network for agricultural purposes: A preliminary assessment (2015).
42. *Sebastian Andersson Hylander:* Evaluation of cultural ecosystem services using GIS (2015).
43. *Deborah Bowyer:* Measuring Urban Growth, Urban Form and Accessibility as Indicators of Urban Sprawl in Hamilton, New Zealand (2015).
44. *Stefan Arvidsson:* Relationship between tree species composition and phenology extracted from satellite data in Swedish forests (2015).
45. *Damián Giménez Cruz*: GIS-based optimal localisation of beekeeping in rural Kenya (2016).
46. *Alejandra Narváez Vallejo:* Can the introduction of the topographic indices in LPJ-GUESS improve the spatial representation of environmental variables? (2016).
47. *Anna Lundgren:* Development of a method for mapping the highest coastline in Sweden using breaklines extracted from high resolution digital elevation models (2016).
48. *Oluwatomi Esther Adejoro:* Does location also matter? A spatial analysis of social achievements of young South Australians (2016).
49. *Hristo Dobrev Tomov:* Automated temporal NDVI analysis over the Middle East for the period 1982 - 2010 (2016).
50. *Vincent Muller:* Impact of Security Context on Mobile Clinic Activities A GIS Multi Criteria Evaluation based on an MSF Humanitarian Mission in Cameroon (2016).
51. *Gezahagn Negash Seboka:* Spatial Assessment of NDVI as an Indicator of Desertification in Ethiopia using Remote Sensing and GIS (2016).
52. *Holly Buhler:* Evaluation of Interfacility Medical Transport Journey Times in Southeastern British Columbia. (2016).
53. *Lars Ole Grottenberg*: Assessing the ability to share spatial data between emergency management organisations in the High North (2016).
54. *Sean Grant:* The Right Tree in the Right Place: Using GIS to Maximize the Net Benefits from Urban Forests (2016).
55. *Irshad Jamal:* Multi-Criteria GIS Analysis for School Site Selection in Gorno-Badakhshan Autonomous Oblast, Tajikistan (2016).
56. *Fulgencio Sanmartín:* Wisdom-volkano: A novel tool based on open GIS and time-series visualization to analyse and share volcanic data (2016).
57. *Nezha Acil:* Remote sensing-based monitoring of snow cover dynamics and its influence on vegetation growth in the Middle Atlas Mountains (2016).
58. *Julia Hjalmarsson:* A Weighty Issue: Estimation of Fire Size with Geographically Weighted Logistic Regression (2016).
59. *Mathewos Tamiru Amato:* Using multi-criteria evaluation and GIS for chronic food and nutrition insecurity indicators analysis in Ethiopia (2016).
60. *Karim Alaa El Din Mohamed Soliman El Attar:* Bicycling Suitability in Downtown, Cairo, Egypt (2016).
61. *Gilbert Akol Echelai:* Asset Management: Integrating GIS as a Decision Support Tool in Meter Management in National Water and Sewerage Corporation (2016).
62. *Terje Slinning:* Analytic comparison of multibeam echo soundings (2016).

63. *Gréta Hlín Sveinsdóttir:* GIS-based MCDA for decision support: A framework for wind farm siting in Iceland (2017).
64. *Jonas Sjögren:* Consequences of a flood in Kristianstad, Sweden: A GIS-based analysis of impacts on important societal functions (2017).
65. *Nadine Raska:* 3D geologic subsurface modelling within the Mackenzie Plain, Northwest Territories, Canada (2017).
66. *Panagiotis Symeonidis*: Study of spatial and temporal variation of atmospheric optical parameters and their relation with PM 2.5 concentration over Europe using GIS technologies (2017).
67. *Michaela Bobeck:* A GIS-based Multi-Criteria Decision Analysis of Wind Farm Site Suitability in New South Wales, Australia, from a Sustainable Development Perspective (2017).
68. *Raghdaa Eissa*: Developing a GIS Model for the Assessment of Outdoor Recreational Facilities in New Cities Case Study: Tenth of Ramadan City, Egypt (2017).
69. *Zahra Khais Shahid*: Biofuel plantations and isoprene emissions in Svea and Götaland (2017).
70. *Mirza Amir Liaquat Baig*: Using geographical information systems in epidemiology: Mapping and analyzing occurrence of diarrhea in urban - residential area of Islamabad, Pakistan (2017).
71. *Joakim Jörwall*: Quantitative model of Present and Future well-being in the EU-28: A spatial Multi-Criteria Evaluation of socioeconomic and climatic comfort factors (2017).
72. *Elin Haettner*: Energy Poverty in the Dublin Region: Modelling Geographies of Risk (2017).
73. *Harry Eriksson*: Geochemistry of stream plants and its statistical relations to soil- and bedrock geology, slope directions and till geochemistry. A GIS-analysis of small catchments in northern Sweden (2017).
74. *Daniel Gardevärn:* PPGIS and Public meetings – An evaluation of public participation methods for urban planning (2017).
75. *Kim Friberg:* Sensitivity Analysis and Calibration of Multi Energy Balance Land Surface Model Parameters (2017).
76. *Viktor Svanerud:* Taking the bus to the park? A study of accessibility to green areas in Gothenburg through different modes of transport (2017).
77. *Lisa-Gaye Greene*: Deadly Designs: The Impact of Road Design on Road Crash Patterns along Jamaica's North Coast Highway (2017).
78. *Katarina Jemec Parker*: Spatial and temporal analysis of fecal indicator bacteria concentrations in beach water in San Diego, California (2017).
79. *Angela Kabiru*: An Exploratory Study of Middle Stone Age and Later Stone Age Site Locations in Kenya's Central Rift Valley Using Landscape Analysis: A GIS Approach (2017).
80. *Kristean Björkmann*: Subjective Well-Being and Environment: A GIS-Based Analysis (2018).
81. *Williams Erhunmonmen Ojo*: Measuring spatial accessibility to healthcare for people living with HIV-AIDS in southern Nigeria (2018).
82. *Daniel Assefa*: Developing Data Extraction and Dynamic Data Visualization (Styling) Modules for Web GIS Risk Assessment System (WGRAS). (2018).
83. *Adela Nistora*: Inundation scenarios in a changing climate: assessing potential impacts of sea-level rise on the coast of South-East England (2018).

84. *Marc Seliger*: Thirsty landscapes - Investigating growing irrigation water consumption and potential conservation measures within Utah's largest master-planned community: Daybreak (2018).

85. *Luka Jovičić*: Spatial Data Harmonisation in Regional Context in Accordance with INSPIRE Implementing Rules (2018).

86. *Christina Kourdounouli*: Analysis of Urban Ecosystem Condition Indicators for the Large Urban Zones and City Cores in EU (2018).

87. *Jeremy Azzopardi*: Effect of distance measures and feature representations on distance-based accessibility measures (2018).

88. *Patrick Kabatha*: An open source web GIS tool for analysis and visualization of elephant GPS telemetry data, alongside environmental and anthropogenic variables (2018).

89. *Richard Alphonce Giliba*: Effects of Climate Change on Potential Geographical Distribution of Prunus africana (African cherry) in the Eastern Arc Mountain Forests of Tanzania (2018).

90. *Eiður Kristinn Eiðsson*: Transformation and linking of authoritative multi-scale geodata for the Semantic Web: A case study of Swedish national building data sets (2018).

91. *Niamh Harty*: HOP!: a PGIS and citizen science approach to monitoring the condition of upland paths (2018).

92. *José Estuardo Jara Alvear*: Solar photovoltaic potential to complement hydropower in Ecuador: A GIS-based framework of analysis (2018).

93. *Brendan O'Neill*: Multicriteria Site Suitability for Algal Biofuel Production Facilities (2018).

94. *Roman Spataru*: Spatial-temporal GIS analysis in public health – a case study of polio disease (2018).

95. *Alicja Miodońska*: Assessing evolution of ice caps in Suðurland, Iceland, in years 1986 - 2014, using multispectral satellite imagery (2019).

96. *Dennis Lindell Schettini*: A Spatial Analysis of Homicide Crime's Distribution and Association with Deprivation in Stockholm Between 2010-2017 (2019).

97. *Damiano Vesentini*: The Po Delta Biosphere Reserve: Management challenges and priorities deriving from anthropogenic pressure and sea level rise (2019).

98. *Emilie Arnesten*: Impacts of future sea level rise and high water on roads, railways and environmental objects: a GIS analysis of the potential effects of increasing sea levels and highest projected high water in Scania, Sweden (2019).

99. *Syed Muhammad Amir Raza*: Comparison of geospatial support in RDF stores: Evaluation for ICOS Carbon Portal metadata (2019).

100. *Hemin Tofiq*: Investigating the accuracy of Digital Elevation Models from UAV images in areas with low contrast: A sandy beach as a case study (2019).

101. *Evangelos Vafeiadis*: Exploring the distribution of accessibility by public transport using spatial analysis. A case study for retail concentrations and public hospitals in Athens (2019).

102. *Milan Sekulic*: Multi-Criteria GIS modelling for optimal alignment of roadway by-passes in the Tlokweng Planning Area, Botswana (2019).

103. *Ingrid Piirisaar*: A multi-criteria GIS analysis for siting of utility-scale photovoltaic solar plants in county Kilkenny, Ireland (2019).

104. *Nigel Fox*: Plant phenology and climate change: possible effect on the onset of various wild plant species' first flowering day in the UK (2019).

105. *Gunnar Hesch*: Linking conflict events and cropland development in Afghanistan, 2001 to 2011, using MODIS land cover data and Uppsala Conflict Data Programme (2019).

106. *Elijah Njoku*: Analysis of spatial-temporal pattern of Land Surface Temperature (LST) due to NDVI and elevation in Ilorin, Nigeria (2019).

107. *Katalin Bunyevácz*: Development of a GIS methodology to evaluate informal urban green areas for inclusion in a community governance program (2019).

108. *Paul dos Santos*: Automating synthetic trip data generation for an agent-based simulation of urban mobility (2019).

109. *Robert O' Dwyer*: Land cover changes in Southern Sweden from the mid-Holocene to present day: Insights for ecosystem service assessments (2019).

110. *Daniel Klingmyr*: Global scale patterns and trends in tropospheric NO2 concentrations (2019).

111. *Marwa Farouk Elkabbany*: Sea Level Rise Vulnerability Assessment for Abu Dhabi, United Arab Emirates (2019).

112. *Jip Jan van Zoonen*: Aspects of Error Quantification and Evaluation in Digital Elevation Models for Glacier Surfaces (2020).

113. *Georgios Efthymiou*: The use of bicycles in a mid-sized city – benefits and obstacles identified using a questionnaire and GIS (2020).

114. *Haruna Olayiwola Jimoh*: Assessment of Urban Sprawl in MOWE/IBAFO Axis of Ogun State using GIS Capabilities (2020).

115. *Nikolaos Barmpas Zachariadis*: Development of an iOS, Augmented Reality for disaster management (2020).

116. *Ida Storm*: ICOS Atmospheric Stations: Spatial Characterization of CO2 Footprint Areas and Evaluating the Uncertainties of Modelled CO2 Concentrations (2020).

117. *Alon Zuta*: Evaluation of water stress mapping methods in vineyards using airborne thermal imaging (2020).

118. *Marcus Eriksson*: Evaluating structural landscape development in the municipality Upplands-Bro, using landscape metrics indices (2020).

119. *Ane Rahbek Vierø*: Connectivity for Cyclists? A Network Analysis of Copenhagen's Bike Lanes (2020).

120. *Cecilia Baggini*: Changes in habitat suitability for three declining Anatidae species in saltmarshes on the Mersey estuary, North-West England (2020).

121. *Bakrad Balabanian*: Transportation and Its Effect on Student Performance (2020).

122. *Ali Al Farid*: Knowledge and Data Driven Approaches for Hydrocarbon Microseepage Characterizations: An Application of Satellite Remote Sensing (2020).

123. *Bartlomiej Kolodziejczyk*: Distribution Modelling of Gene Drive-Modified Mosquitoes and Their Effects on Wild Populations (2020).