

Understanding mosquito vectors and methods for their control

Ben Lambert

Supervisors: Professor Sir Charles Godfray & Dr Ace North

A thesis presented for the degree of
Doctor of Philosophy



Somerville College
University of Oxford
UK
Trinity 2017

Acknowledgements

This work has been a labour of love for me although, for others, this is not an accurate description. This section is a meagre attempt to due justice to the patience and efforts of these people, in getting me from point A (a confused, slightly too old graduate student) to point B (a slightly less confused and, unfortunately, older graduate student).

First and foremost, I would like to thank my supervisors for the past three years, Ace North and Charles Godfray. Ace, I just checked my email history, and found 416 sent messages from me to you; to say you are a patient individual clearly does not suffice. Apart from answering my various queries through the years, I have learned much from you. I hope that my ineptitude has not put you off future supervisions, and that we can continue to work together in the future. Charles, thank you for our meetings. My scientific rigour and, I hope, the quality of my prose, have improved immeasurably since we met. Any good that comes from my work reflects on both of your efforts; any shortcomings reflects on my inability to listen to you.

There are various other people who have bettered my academic health throughout the years, including Armand Leroi, David Gavaghan, Astrid Iversen, everyone at Target Malaria, and Helen Byrne, and again, please accept my somewhat prosaic gratitude. I would like to especially thank Tom Churcher for including me in his work on near-infrared spectroscopy (on which Chapter 4 is based) after a chance meeting in Paris. Also, I would like to thank Tom for his excellent guidance, and for his contribution to this piece of work.

There are also those who have helped me, academically and sometimes less academically, throughout the years. Whilst this list is necessarily curtailed, I would like to thank Jean-michel Johnston, Hazel Tubman, Fergus Cooper, Joanna Raisbeck, Roman Walczak, and the various members of the Spartan team.

There are two others who have done more than their fair share of listening to me over

these DPhilled years. Charlotte King, without you, I would not have got to Oxford in the first place. Claire Wheeler, without you, my life in Oxford would have not been as fun-filled, and truly great, as it was. Please accept my humble thanks, apologies, and future drinks, as an attempt to bring our trade imbalance further away from the red.

My parents, David Robinson and Vivien Lambert, and the rest of my family, I owe most to you. Throughout the years you have helped me considerably and, one day, I hope to be able to repay you. For now, you'll have to accept these words as thanks.

Tatjana Detinova's mass mosquito dissection experiments, consisting of tens of thousands of mosquitoes in Soviet Russia, Derek Charlwood's seemingly global efforts, and the work of my experimental collaborators, Floyd Dowell and Maggy Sikulu-Lord, amongst other examples, make it clear that this work sits atop the broad shoulders of field entomologists past and present. I would like to dedicate this thesis to the arduous work of these insect people that do real science; I hope it does not disappoint.

The work of others

This section is an attempt to credit those people who directly contributed to the material within this thesis.

Ace North and Charles Godfray read through my introduction and provided useful comments throughout this process.

In Chapters 2 and 3, it should be recognised that Ace North and Charles Godfray met with me throughout the process of analysis, improving the content considerably. In Chapter 2, the original database of mark-release-recapture experiments was compiled by Guerra et al., 2014, and I thank them, and the work of field entomologists past. In Chapter 3, I compiled the database, but again owe much to the work of field entomologists in the past.

Tom Churcher was gracious enough to include me on his work on near-infrared spectroscopy, on which Chapter 4 is based. Many of the figures came about through our discussion and joint work on the manuscript, and I thank Tom for his efforts here. The near-infrared spectroscopy (NIRS) data was collected by my experimental collaborators, Floyd Dowell and Maggy Sikulu-Lord, and I thank them for their considerable work here.

Chapter 5, perhaps more than others, came from the joint work of myself and Ace North (Charles Godfray and Austin Burt also advised). My main contribution was the statistical model of rainfall, which we used to explore the effects of a homing endonuclease gene (HEG) release. Ace did much of the work of integrating my model into his spatial model of mosquito ecology (described in North, Burt, and Godfray, 2013). We both contributed equally to the running of simulations and figure generation.

A series of quotes illustrating malaria through the ages

...at the onset of his disease his temples bring fever to him, and afterwards they take the fever at sweat away...his are 'blown', his entrails...vertigo...his knees hurt him...[whatever] he eats and drinks, does not please him: that man is ill of the di'u illness.

– Medical description of a febrile illness called *di'u*, *Babylonian-Assyrian text, circa 1900-600 BC*

As one who has the shivering of the quartan so near,/ that he has his nails already pale/ and trembles all, still keeping the shade,/ such I became when those words were uttered.

– Discussion of quartan fevers (those appearing every four days) in *The Inferno*, *Dante Alighieri, 1471*

You are so very choleric of complexion./ Beware the mounting sun and all dejection,/ Nor get yourself with sudden humours hot;/ For if you do, I dare well lay a groat/ That you shall have the tertian fevers pain,/ Or some ague that may well be your bane.

– Discussion of tertian fevers (those appearing every three days) in *The Nun's Priest's Tale*, *Geoffrey Chaucer, c.1390s*

. . . (he) hath got, as I take it, an ague . . . hes in his fit now and does not talk after the wisest. He shall taste of my bottle: if he have never drunk wine afore it will go near to remove his fit . . . Open your mouth: this will shake your shaking . . . if all the wine in my bottle will recover him, I will help his ague.

– Discussion of ague (a term for an illness thought to be malaria in common use in England at the time) in *The Tempest*, *William Shakespeare, c.1611*

Beware beware, the Bight of Benin. One comes out where fifty went in.

– English jingle associated with slavery in West Africa, c.1800s

*The poor white man, faint and weary
Came to sit under our tree.
He has no mother to bring him milk;
No wife to grind his corn.
Let us pity the white man.*

– An African song, c.1800s

Malaria is the great enemy of the explorer, the missionary, the planter, the merchant, the farmer, the soldier, the administrator, the villager and the poor...

– From *The Prevention of Malaria*, Ronald Ross, 1913

No one who has studied the epidemiology of malaria can fail to be impressed by the extreme diversity of the data recorded regarding parasite prevalence, period of transmission, degree of endemicity, epidemic potential and amenability to control measures in different regions of the world and even in different parts of the same country.

– From *The Epidemiology and Control of Malaria*, George Macdonald, 1957

Considering their impact, you might expect mosquitoes to get more attention than they do. Sharks kill fewer than a dozen people every year and in the U.S. they get a week dedicated to them on TV every year. Mosquitoes kill 50,000 times as many people, but if there's a TV channel that features Mosquito Week, I haven't heard about it.

– From a web article titled *The Deadliest Animal in the World*, Bill Gates, 2014

We are planning to have humans settling on Mars by 2023, which is 225 million miles away from earth. But malaria is just 0km from earth. If we can solve problems at 225 million miles away from us, then there is no reason that we can't defeat malaria here, in our living room.

– From the *Target Malaria* presentation at the *Grand Challenges* meeting, Abdoulaye Diabate, 2016

Contents

1	Introduction	15
1.1	Abstract	15
1.2	Introduction	17
1.3	The largely forgotten impact of mosquito borne disease in the developed world throughout history	26
1.4	The impact of mosquito borne disease in the developing world today	37
1.5	Mosquitoes as vectors	49
1.6	Disease control since Ronald Ross	52
1.7	Why we need to understand mosquito vectors	64
1.8	The importance of mosquito senescence	72
1.9	Future vector control by genetic drive	75
1.10	Motivating this thesis	81
2	A meta-analysis of mark-release-recapture experiments of mosquitoes	83
2.1	Abstract	83
2.2	Introduction	84
2.3	Materials and methods	87
2.3.1	Data	87

2.3.2	Statistical model	87
2.3.3	Individual time-series estimates	91
2.3.4	Estimating lifespan at the species, genus and overall groupings	92
2.3.5	Testing for age-dependent mortality in wild mosquitoes	95
2.3.6	Model estimation by MCMC	95
2.3.7	K-Fold cross validation	98
2.4	Results	102
2.4.1	Mosquito lifespan varies across species although variation is less marked at the genus level	102
2.4.2	A small proportion of mosquitoes are likely responsible for the bulk burden of disease for malaria and Zika	104
2.4.3	The longevity of mosquitoes does not vary systematically over the range of temperatures of MRR study locations	106
2.4.4	Male mosquitoes have shorter lives than females	106
2.4.5	Blood- and sugar-fed mosquitoes do not live longer than those that were not fed before release	107
2.4.6	Wild mosquitoes do not appear to experience strong age-dependent mortality	107
2.5	Discussion	119
2.5.1	The statistical power of MRR studies	122
2.5.2	MRR design	124
2.6	Conclusion	127
3	A meta-analysis of mosquito dissection experiments	129
3.1	Abstract	129
3.2	Introduction	130

3.3	Method	138
3.3.1	Collection of dissection data	138
3.3.2	Statistical analysis of dissection data	141
3.3.3	Data collection of gonotrophic cycle duration	149
3.3.4	Statistical analysis of gonotrophic cycle data	150
3.3.5	Conversion of lifespan from physiological to calendar age	152
3.4	Results	158
3.4.1	Wild mosquito lifespan in gonotrophic cycles	158
3.4.2	Gonotrophic cycle duration	159
3.4.3	Wild mosquito lifespan in calendar age	160
3.4.4	The evidence for age-dependent mortality	160
3.5	Discussion	167
3.5.1	Comparison with MRR estimates	167
3.5.2	Comparison with laboratory-based methods	172
3.5.3	Does age-dependence matter?	174
3.6	Conclusion	175
4	Using near-infrared spectroscopy to estimate mosquito population mean lifespan	179
4.1	Abstract	179
4.2	Introduction	180
4.3	Method	185
4.3.1	Data	185
4.3.2	Machine learning method: existent work	186
4.3.3	Machine learning method: our approach	189
4.3.4	Cross validation and bias correction	190

4.3.5	Surrogate Bayesian model	193
4.3.6	Population lifespan inference	194
4.4	Results	196
4.4.1	Accurate prediction of individual mosquito age within a given experiment	196
4.4.2	Individual prediction errors decrease with sample size	198
4.4.3	Inferring population mean lifespan is possible with modest sample sizes	199
4.4.4	The bulk of uncertainty in population lifespan estimates is due to sampling variability	201
4.4.5	Between-study prediction is poor	203
4.5	Discussion	207
4.6	Conclusion	210
5	Using a homing endonuclease to control vector populations in a landscape with temporal heterogeneity	213
5.1	Abstract	213
5.2	Introduction	214
5.3	Method	216
5.3.1	Mosquito population model	216
5.3.2	Linking rainfall with aquatic habitat density	220
5.3.3	A general model for rainfall	221
5.3.4	HEG releases	222
5.4	Results	223
5.4.1	Constant environment	223
5.4.2	HEG dynamics are more variable in a seasonal environment	224
5.4.3	The start of the wet season is the optimal time for release	225

5.4.4	In seasonal environments eradication is easier to achieve if the average aquatic habitat density is low or high rather than intermediate	225
5.4.5	Random variation in rainfall generally increases the probability that a HEG will induce population extinction	226
5.4.6	Releases during the first half of the wet season result in the highest chance of population extinction, although in some cases the time taken for extinction to occur may be lower by releasing HEGs during the dry season	227
5.5	Discussion	237
5.6	Conclusion	242
6	Conclusion	243
7	Appendix	257
7.1	A mathematical model of senescence	257
7.2	Studies included in the MRR meta-analysis	271
7.3	Studies included in the dissection study meta-analysis	272
7.4	Studies included in the gonotrophic cycle duration meta-analysis	273

Chapter 1

Introduction

1.1 Abstract

Mosquitoes spread diseases that shorten and worsen the lives of many people, chiefly children in poor countries, around the world. Since Ronald Ross' discovery at the end of the nineteenth century that mosquitoes transmit malaria, field entomologists have collected a great deal of information about mosquito ecology. Despite this tremendous effort, there still remain significant gaps in our knowledge of mosquito ecology, in part, reflecting the significant variation in mosquito ecology across species and geographies. The main aim of this thesis is an attempt to synthesise the substantial information that field entomologists have collected on mosquito lifespan. In Chapters 2 and 3, I conduct meta-analyses of the two predominant approaches used to estimate mosquito lifespan: mark-release-recapture experiments and female mosquito dissection-based studies, respectively. These analyses produce estimates of mosquito lifespan by species and genus, and more broadly, allow for an appraisal of these two experimental approaches. In Chapter 4, I describe a recently developed approach, known as *near-infrared spectroscopy*, which enables estimation of the

15 age of individual mosquitoes, and then perform an *in silico* analysis to explore the use of
16 this technology for estimating the average lifespan of wild populations of mosquitoes.

17 The emergence of mosquito resistance to the main insecticides used in vector control,
18 along with the concerning recent discovery that the malaria parasites in Asia are becoming
19 resistant to artemisinin - an important drug used to treat malaria - highlight the need
20 for novel approaches to control disease transmission. Some recently-proposed approaches
21 involve genetic modification of the mosquito vectors, for example, to render them incapable
22 of acting as hosts for disease or to reduce their fecundity. In Chapter 5, I model the impact
23 of a release of mosquitoes carrying a genetic construct known as a *homing endonuclease*,
24 which has been constructed to bias the sex of mosquito offspring towards males, in compu-
25 tational environments that capture some facets of the real life landscapes where mosquito
26 borne disease is rife.

27 About a century ago, the famous Italian Malariologist Giovanni Grassi declared that
28 malaria was a “giant with clay feet”; reflecting the optimism, in some academic circles at
29 the time, that eradication of this terrible disease would soon occur. Unfortunately, a cen-
30 tury of often unsuccessful attempts to control and eradicate malaria, and other mosquito
31 borne diseases, would follow Grassi’s statement, meaning that this fight is likely to con-
32 tinue throughout the twenty-first century. We now know much more about mosquitoes and
33 mosquito borne disease than we did a generation ago, but there is still crucial information
34 that we do not. In this thesis, I argue that in order to make significant inroads to disease
35 eradication, further research on mosquito ecology is crucial. Only when we better under-
36 stand our unwitting mosquito foe, can we design and implement effective disease control
37 measures that are so desperately needed in those most desperate parts of the world.

1.2 Introduction

In terms of the number of people they sicken and kill, mosquitoes are more dangerous than any other organism on earth, including humans (Gates, 2014). At the peak of their breeding season, mosquitoes outnumber every other animal on the planet, except ants and termites (Gates, 2014). Malaria is the most deadly of the mosquito borne diseases, which is vectored by anopheline mosquitoes and caused by micro organisms of the genus *Plasmodium*. Malaria has plagued humankind throughout our history: the 4700 year old Chinese tome ‘The Nei Ching’, probably the world’s oldest medical textbook, refers to conditions characterised by recurrent fevers and spleen enlargement, typical of malaria (Shampo and Kyle, 1989); it has been argued that Alexander the Great died of *P. falciparum* malaria on route to India (Carter and Mendis, 2002); Shakespeare mentions ague (the English word for the disease in common usage until the 19th Century) in eight of his plays (Reiter, 2000). Whilst likely exaggerated, it has been suggested that malaria accounts for half of all human deaths through history (Whitfield, 2002). In the twentieth century alone, it has been estimated that malaria killed 300 million people (Carter and Mendis, 2002), as much as five times the combined death toll from both world wars (Britannica, 2017a; Britannica, 2017b). Regardless of its historical importance, malaria today sickens around 300 million and likely kills between 0.5 and 1.2 million people each year (World Health Organisation, 2015; Murray et al., 2012), most of whom are children born into poverty in the tropics. Mosquitoes also vector other diseases that continue to inflict devastating human suffering worldwide. These include filariasis - a disease caused by macroscopic worms that often causes severe disfigurement in those it affects - which has an annual incidence of around 40 million cases worldwide (World Health Organisation, 2017a). *Aedes* mosquitoes vector a febrile disease called chikungunya whose global spread is widening (Charrel, 2007), having recently caused outbreaks in Italy (Rezza et al., 2007). Chikungunya, like many other

63 mosquito borne diseases, has no cure. The ecological disruption caused by World War
64 II expanded the geographic distribution of *Ae. aegypti* (Gubler, 2002), which are major
65 vectors of chikungunya, as well as the deadly dengue virus. Troop movements during this
66 period likely accelerated the spread of dengue virus throughout Asia and, today, it has
67 been estimated that there are 390 million dengue infections each year (Bhatt et al., 2013)
68 and that 3.9 billion people - over half the world's population - are at risk of contracting this
69 disease (Brady et al., 2012). Whilst dengue infection is often asymptomatic, it can result
70 in influenza-like symptoms known as *dengue fever* and also a deadly condition referred to
71 as *dengue hemorrhagic fever* (Guha-Sapir and Schimmer, 2005), where affected patients
72 bleed from their nose and mouth before falling into a state of shock from which they may
73 never recover. In February 2016, the World Health Organisation declared a 'Public Health
74 Emergency of International Concern', following the rapid spread of the Zika virus (mainly
75 vectored by *Ae. aegypti* and *Ae. albopictus* mosquitoes) through Latin America (World
76 Health Organisation, 2016b). The prevailing opinion is that Zika infection is behind the
77 tragic rise in cases of microcephaly in newborn children in Brazil observed recently (Mlakar
78 et al., 2016; Oliveira Melo et al., 2016).

79 Set against this backdrop of illness and suffering - both historical and present - it is
80 surprising that mosquitoes do not get more media attention than they do. Indeed, as
81 a researcher in the field, it is not uncommon to encounter views from members of the
82 public in the developed world, to the tune of 'I thought we had solved malaria, so why
83 are you studying mosquitoes?'. Academically, there has been over one hundred years of
84 research into the organism, initially catalysed by Ronald Ross' discovery in 1897 that
85 anopheline mosquitoes could transmit the malarial parasite, and implicated mosquitoes
86 in the transmission of human malaria (Cox, 2010). Throughout the twentieth century,
87 the pendulum has swung frequently between the two predominant approaches used to

control malaria (paralleled in other mosquito borne diseases): targeting the human stages of the disease using drugs (for malaria, quinine then chloroquine then artemisinin) and vector control methods (Packard, 2007). In the 1980s, advances in our understanding of immunology and vaccination technology lead to a shift towards interventions that targeted the human stages of vector borne disease (Godfray, 2013). Indeed, in their landmark text on the mathematical epidemiology of infectious diseases Anderson and May, 1992 declared, ‘Today, with the intense interest in the development of vaccines and the application of new molecular and biochemical techniques to the study of acquired immunity in humans, the research emphasis has changed from entomology to infections in humans.’ More recently, the popularity of vector control methods has been resurgent (and largely responsible for the recent gains against malaria in Sub-Saharan Africa), partly due to the difficulty in developing vaccines and drugs that can effectively target pathogenesis across the diverse life stages of malarial and filarial parasites in humans (Greenwood et al., 2008). In Cambodia in 2009, forms of the malarial parasite resistant to artemisinin (the main chemical compound from which drugs used to treat human malaria are derived) were reported (Dondorp et al., 2009). These resistant forms of the parasite have subsequently spread throughout Asia (Ashley et al., 2014) and now its potential to invade Africa threatens to inflict as much damage to human life as resistance to chloroquine caused 50 years ago (White, 2010). These recently emergent parasitological hurdles underscore the continuing importance of vector control in the twenty-first century and have, in part, lead to calls for a generation of novel approaches to vector control.

To design the most effective vector control interventions, it is essential to have a detailed knowledge of mosquito ecology. The transmission of vector borne disease is particularly sensitive to the amount of time that adult mosquitoes live in the wild (Ross, 1910; Macdonald, 1957; Smith, Dushoff, and McKenzie, 2004). Mosquitoes that live longer bite more humans

and, hence, have a greater chance to become infected with the aetiological agent, survive its extrinsic incubation period, and transmit the disease. Whilst it is fairly straightforward to estimate the mean lifespan of mosquitoes laboratory colonies, this task is nontrivial for wild mosquitoes. Historically, two approaches have been used to estimate the lifespan of wild mosquitoes. The first is mark-release-recapture experiments (see, for example, Silver, 2007) where a number of lab-reared or wild-caught mosquitoes are marked, typically with a fluorescent dye, then released into the wild. By monitoring the number of marked mosquitoes that are recaptured over time, it is possible to estimate the combined rate of mosquito mortality and dispersal (if the number recaptured rapidly decays, mosquitoes die and fly away from the study area at a correspondingly fast rate). The second approach is based on the dissection of female mosquitoes. Female mosquitoes undergo changes in their reproductive anatomy that occur at roughly uniform intervals throughout their lifetime, providing a biological ‘clock’ which can be used to assess the physiological age of individual specimens by dissection (Polovodova, 1949; Detinova, 1962). By determining the age of a large number of specimens this allows estimation of mosquito lifespan. Whilst both of these approaches provide valuable (and complimentary) information about mosquito ecology, the individual experiments are costly and require large numbers of mosquitoes to produce accurate estimates of mosquito lifespan. Further, there is considerable variation in the estimates of mosquito lifespan produced across studies, presumably due to differences in experimental protocol, the particular mosquito species studied, and the climate and geography of the study region (Guerra et al., 2014). In Chapters 2 and 3, I present the results of analyses of data from a large number of mark-release-recapture (MRR) experiments and dissection-based studies, respectively. These meta-analyses synthesise the results of the arduous work done by field entomologists and provide robust estimates of wild mosquito lifespan. A benefit of analysing the results of many studies is that this allows us (my col-

laborators and I) to produce estimates of mosquito lifespan at the genus and species levels. To our knowledge, this is the first attempt to do this, and we hope this will prove useful for researchers and public health officials who study the epidemiology of various mosquito borne diseases. The simultaneous analysis of data from the two types of experiments also allows us to appraise both approaches in a way that was previously not possible.

Some of the most effective methods of vector control, including insecticide treated nets and indoor residual spraying, aim to kill adult mosquitoes. At a population level, these approaches should lead to a reduction in the average lifespan of mosquitoes, resulting in lower disease transmission. In the past, the main way to assess vector control measures at an operational level was by surveying the percentage of the human population protected by interventions. However, resistance to pyrethroids - the main insecticide used in mosquito vector control and the only class used for treating bednets - is now widespread throughout Africa (Ranson et al., 2011; Basilua Kanza et al., 2012; Mulamba et al., 2014; Ranson and Lissenden, 2016), likely weakening the effect of traditional vector control methods. This resistance also means that intervention coverage is no longer a strong indicator of vector control. As I discussed above, MRR experiments and mosquito dissection studies can be used to estimate the lifespan of wild mosquito populations. Whilst these methods are currently the main approaches used to study many aspects of mosquito ecology, these experiments are time consuming, costly, and require specialist knowledge of mosquito biology to be effectively implemented. In Chapter 4, I discuss a novel approach to estimating mosquito lifespan using near-infrared spectroscopy (NIRS).

As a mosquito ages, it undergoes changes in its body chemistry. These changes can be detected by shining near-infrared light through the head and thorax of mosquito specimens and examining the resultant spectrum. Further, previous work using laboratory mosquitoes has demonstrated that the changes to the near-infrared spectrum throughout a mosquito's

lifetime are regular enough to be able to predict the age of individual specimens (Sikulu et al., 2010; Sikulu et al., 2011; Dowell, Noutcha, and Michel, 2011; Sikulu et al., 2014; Liebman et al., 2015). In Chapter 4, I present an analysis of a dataset which we compiled from a number of previously-published NIRS experiments where, for each sample, we have the near-infrared spectrum of a mosquito of known age. In the previously-published work, the statistical models used to predict the age of mosquitoes have been relatively basic. In this chapter, I apply more recently-developed machine learning models to the task of predicting the age of individual specimens in our database, resulting in a greater predictive accuracy than has previously been reported. Since our database consists of various previous studies, this also allows us, for the first time, to determine how heterogeneity in study protocol impacts the predictive accuracy of NIRS.

To gauge the effectiveness of vector control methods that aim to kill adult mosquitoes, we wish to know how the mean lifespan of the mosquito population is reduced as the interventions are implemented. This focus on the *population* level has been neglected by previous NIRS studies, which tend to focus on predicting the age of *individual* specimens. In Chapter 4, I use *in silico* populations of mosquitoes of a known age distribution, to demonstrate that NIRS can be used to infer changes in the population mean age across those age ranges likely to be encountered in applied vector control. Since the cost of NIRS is mostly upfront (the initial cost of the machine) with it being relatively easy to scan individual mosquito samples, I argue that this approach may provide a viable means of surveying mosquito vector populations in the future.

Since the turn of the twenty-first century, there has been unprecedented reduction in the human costs of malaria driven largely by increasing coverage of insecticide treated nets (Bhatt et al., 2015). It has been estimated that 663 million disease cases may have been averted due to the combination of insecticide treated nets, indoor residual spraying

and artemisinin-based combination therapy from 2000-2015 (Bhatt et al., 2015). The increasingly widespread resistance of mosquitoes to insecticides, climate change, along with the emergent threat of malaria parasite resistance, however, threatens to halt and, perhaps, even reverse these trends. Worryingly, in 2016, progress against malaria appears to have stalled, with an increase of around 5 million cases over 2015 estimated (World Health Organisation, 2017c). This hints that the future gains against malaria may be harder to attain and suggests that new strategies are required to continue to alleviate the disease burden in the future.

Endonucleases are genes that code for proteins that cleave DNA at specific recognition sites in the genome and occur in a variety of organisms. A class of these genes known as homing endonucleases (HEGs) are thought to be ‘selfish genes’, which use a trick to rapidly spread through populations. In heterozygotes for this gene (where one chromosome carries the gene but the other does not), a HEG protein cuts the opposite chromosome and the cell sets about repairing it. The cell’s repair mechanism, however, uses the uncut chromosome - the one carrying the HEG - as a blueprint for the repair operation. From the HEG’s perspective, this is fortunate because it is embedded in the blueprint, meaning it will be copied over to the other chromosome and, in doing so, convert the heterozygote into a homozygote. This huge selective advantage means that a newly-introduced HEG will spread from rare to fixation in a population (Goddard and Burt, 1999). Burt, 2003 first argued that HEGs could be used to control mosquito vector populations. Burt imagined that there were a number of ways that this could be done. For example, if the HEG recognised a gene essential for the pathogen but of low importance for the mosquito host, this could prevent the pathogen’s transmission, halting the spread of disease. Alternatively, if the HEG cut a site that was crucial for mosquito survival or reproduction, then the spread of this genetic element could knock down the population density of mosquitoes to

213 levels where the stochastic loss of the parasite is possible. In Chapter 5, I present the
 214 results from computer simulations of a different sort of HEG; again originally proposed by
 215 Burt, 2003: suppose a HEG was inserted onto the Y chromosome but was engineered to
 216 recognise a site on the X chromosome (and was expressed at meiosis), then only the Y
 217 gametes would be viable, resulting in entirely male offspring (mosquitoes have the same
 218 sex determination mechanism as humans). These so-called ‘Y drive’ genes would spread
 219 throughout the population and, in doing so, increasingly skew the population sex ratio
 220 towards males. Much like a party all-too-typical of human adolescence, eventually the
 221 few remaining females would be stochastically lost and the population would be locally
 222 eliminated.

223 In Chapter 5, I build on previous modelling work of Y-drive HEGs (Burt, 2003; Dere-
 224 dec, Burt, and Godfray, 2008; Deredec, Godfray, and Burt, 2011; North, Burt, and God-
 225 fray, 2013) to consider its spread in more realistic models of the underlying landscape.
 226 Specifically, I extend a spatial model introduced by North, Burt, and Godfray, 2013 that
 227 incorporates key aspects of mosquito ecology. North, Burt, and Godfray, 2013 considers
 228 how spatial characteristics of a landscape influence the impact of a Y-drive HEG on local
 229 mosquito populations. Here, I use computational simulation to investigate how temporal
 230 characteristics of a landscape - the availability of mosquito aquatic habitats, driven by
 231 short-term fluctuations in rainfall - affect the use of such a HEG to control vector pop-
 232 ulations. Encouragingly, the results that I present indicate that if a Y-drive HEG were
 233 deployed at the optimal time of year (also determined through simulation) then the re-
 234 sultant intervention may actually be bolstered by temporal heterogeneity. I conclude this
 235 chapter by using computational simulation to investigate the spread of a Y-drive HEG in
 236 two case study locations: one in Kenya, in tropical east Africa, the other in Mali, in the
 237 Sahel region in West Africa, and demonstrate that a HEG may an effective tool for vector

control in these areas.

In Chapter 1, I provide a historical narrative of the impact of vector borne disease on humans and the methods used to control it. This torrid history highlights the paramount importance of these diseases and conveys the practical difficulty of eradicating these scourges of humanity. I also believe that this historical perspective suggests that there is likely to be no silver bullet for controlling, and later eradicating, mosquito borne diseases. I do argue, however, that vector control methods are a crucial part of those strategies most likely to put an end to such manifold human suffering.

The role of the mosquito in human malaria transmission was first determined by Ronald Ross and other researchers at the turn of the twentieth century. In this chapter, I discuss the Ross-Macdonald mathematical model of malaria that underlies the main approaches to vector control and use it to highlight the need to better understand mosquitoes; specifically, the importance of knowing how long mosquitoes live. I also discuss how the nature of mosquito mortality matters. Historically, it has been assumed that mosquitoes, unlike humans, very rarely reach ages where they experience rapid physiological decline leading to death. If this is incorrect, mathematical models have hinted that we need to re-evaluate our current approaches to vector control. Whilst, in Chapters 3 and 4 I present evidence that mosquito senility is not a strong factor in wild mosquito mortality, in this introductory chapter I describe the previously-published attempts to look for senescence and the mathematical models that motivate this hunt.

The remaining part of this chapter is a short survey of gene-drive mechanisms that have been proposed to control mosquito populations. This area is expansive and I cannot do it full justice here. Instead, I focus on the mathematical models that have been used to investigate the performance of a few genetic constructs that are currently being considered for vector control.

263 1.3 The largely forgotten impact of mosquito borne disease 264 in the developed world throughout history

265 *When unable to defend herself by the sword Rome could defend herself by means*
266 *of the fever.*

267 – The poet *Godfrey of Viterbo*, 1167

268 Ancient Rome was frequently paralysed by outbreaks of malaria (Celli, 1925). Being
269 millennia before Louis Pasteur and Robert Koch provided the final proof of the germ
270 theory of disease, the Romans turned to celestial origins. A temple to *Dea Febris* - the
271 goddess of fever (see Figure 1.1) - was constructed on the Palatine hill, possibly the site of
272 the original settlement of Rome by Romulus. According to Valerius Maximus (an author
273 of historical anecdotes living during the reign of Tiberius), the worship of the goddess of
274 fever was driven by fear, “When they worship other deities, they expect to receive a benefit.
275 However fever is worshipped, so that she will cause less harm...” (Maximus and Valerius,
276 1888). The terror of malaria was so widespread that there even existed a *febris* cult devoted
277 to the deity. Lest we imagine that belief in the divine providence of malaria only existed
278 during Classical antiquity, the febris cult was active in Italy well into the twentieth century
279 (Sallares, 2002).

280 The historical presence of malaria has also affected our language. Interestingly, the
281 word *abracadabra*, ubiquitous in children’s pantomimes, may have originated as a Roman
282 treatment for semitertian fever. Quintus Serenus Sammonicus, physician to the Roman
283 emperor Caracalla, dictated that patients suffering fever wear an amulet inscribed with an
284 incantation based on the word *abracadabra* for nine days (see Figure 1.2; Serenus and Pépin,
285 1950; Hempelmann and Krafts, 2013). Whilst some scholars believe that *abracadabra* has
286 no meaning, others believe it may have been derived from the Hebrew words *Abrai seda*



Figure 1.1: **A statue of Dea Febris, goddess of fever, by unknown artist.** Reproduced from Tanogabo, 2015.

brai meaning “Out, bad spirit, out” or, alternatively, from the Hebrew phrase *Ab, Ruach, Dabar* meaning “Father, Holy Ghost, Word” (Wootton, 1910).

Human efforts to transform the land for agriculture in the 1600s abetted the northwards expansion of malaria from southern Europe, where it was present from at least the 4th century BC (Packard, 2007). The seasonality and cooler temperatures likely thwarted *P. falciparum* from establishing, but *P. vivax* and *P. malariae* could and did spread into England and as far north as Russia (Packard, 2007). During seventeenth century England, the expanding population of well-off folk in London and other cities demanded more food. This rising demand for agricultural produce lead to efforts to drain the marshlands of southeast England and the Fenlands of Cambridgeshire and Lincolnshire to make way for farmland. These drainage schemes created the ideal habitats for anopheline mosquitoes. In attempts to hold back the sea, embankments were constructed that altered the flow of water over the land and lead to the development of brackish pools of water and swamps

Whilst this migration may have benefited the local economy, the large influx of individuals lacking any immunity to malaria could have worsened the marshland malarial situation (Packard, 2007). It is unfamiliar for us to associate what is today the prime commuter belt for London with malaria, but in the sixteenth till the late nineteenth century this was conceived wisdom,

In Kent, and all along on Essex side A troupe of cruel fevers did reside: Of severall Country agues lay an Hoast. And most of them, who had this place forsooke Were either slaine by them, or prisoners tooke.

– The poet *George Wither*, 1588-1667

One of the strongest pieces of evidence that ague (the common term for the condition causing intermittent fevers in England at the time) was indeed malaria, is the identity of its cure.

Peruvian Indians chewed the bark of the cinchona tree - possibly to avoid shivering in Spanish-owned mines because of quinine's muscle relaxant effects - but it is not believed that they did this to treat malaria (Meshnick and Dobson, 2001). The countess of Chinchon and her husband are falsely credited with bringing the bark back to Europe, and it is after her that Linneaus in 1742 named the tree, cinchona, although the more common name for the powder used to treat ague was Jesuits' powder (Jesuit priests were the first to bring cinchona bark from the New World) or Peruvian bark (Bruce-Chwatt, 1988).

The active ingredient of cinchona bark is quinine, possibly the drug responsible for alleviating more suffering in human history than any other single chemical. Quinine is still used to treat malaria today, but most people do not know that its clinical treatment was pioneered on patients suffering from ague in the salt marshes of Essex. Robert Talbor - a relatively uneducated man - developed an effective treatment regimen for ague that he published in 1672 (Reiter, 2000). His concoction, perhaps the first patent medicine,

338 was essentially a mix of white wine and cinchona powder. The success of his treatment,
339 however, was enough to propel him to fame and fortune, and he was knighted in 1678 after
340 he cured the King Charles II of ague.

341 The increasing cattle population between 1840 and 1910 provided anopheline vectors
342 with another source of blood and, together with a decreasing acreage of marshland, con-
343 tributed to a major decline in malaria over the period (Kuhn et al., 2003; Packard, 2007).
344 Indeed, knowledge that malaria once plagued the British Isles has largely been forgotten.
345 This is not without reason - whilst more than 52,000 cases of malaria have been imported
346 since 1953, none of these have lead to secondary cases (Kuhn et al., 2003).

347 Whilst there is evidence that malaria has likely been with us throughout recorded
348 history, the parasite itself, and its association with humans, predates this. It is believed
349 that the ancestors to the malarial parasites first became adapted to the larvae of an aquatic
350 insect of the order Diptera, to which mosquitoes and other blood-sucking flies belong, which
351 first appeared around 150 million to 200 million years ago (Carter and Mendis, 2002).
352 Either during this era or following it, the ancestral malarial parasites became adapted to
353 the two-host life cycle that characterises them now (Carter and Mendis, 2002). Since the
354 emergence of this first malarial parasite, adaptation and evolution of the parasite caused
355 the phylogeny to radiate into many different lines (see Figure 1.4). The malarial parasites
356 came to parasitise many different animal species, of which the greatest number have been
357 recorded amongst primates (Garnham, 1966). Of those that affect humans, there are
358 five, all of the genus *Plasmodium*: *P. falciparum*, *P. vivax*, *P. malariae*, *P. ovale* and *P.*
359 *knowlesi*. The five human malarial parasites have quite different evolutionary histories,
360 and some of these are more related to animal malaria species than they are to one another.
361 The most distinct of the parasites is *P. falciparum*, which today causes the most serious
362 forms of illness. *P. falciparum* is closely related to *P. reichenowi* (Escalante and Ayala,

1994; Escalante, Barrio, and Ayala, 1995; Pacheco et al., 2013), which causes malaria in chimpanzees and, interestingly, each of these are more related to *P. gallinaceum*, the aetiological agent of avian malaria, than they are to many other forms that target mammals (Escalante and Ayala, 1995). The most recent common ancestor of *P. falciparum*, *P. reichenowi* and the avian malarial parasite is estimated to have existed around 130 million years ago (Escalante and Ayala, 1995), with the divergence between *P. falciparum* and *P. reichenowi* having occurred around 4 million to 10 million years ago (Escalante and Ayala, 1994; Escalante, Barrio, and Ayala, 1995). This latter period overlaps with the period, around 5 million years ago, that the human line diverged from the African great apes (Kumar and Hedges, 1998). The other four forms of the parasite exist in clades which include other mammalian parasites (Escalante, Barrio, and Ayala, 1995; Singh et al., 2004; Pacheco et al., 2013). It has been suggested that all forms of human malaria have African sources and that their characteristics (including the illnesses they inflict) have evolved over the past 100,000 years, governed by climatic events and human migration (Carter and Mendis, 2002). Whilst opinions differ on exactly where and how different forms of malaria developed, it is clear that malaria is an ancient human affliction, and that us and the parasite likely co-evolved throughout our joint history.

Malaria is sadly not the only disease borne by mosquitoes that have plagued us through history. The historical force of *yellow fever*, borne by *Ae. aegypti* mosquitoes is, at least, comparable to that of malaria, and the panic it enacted in affected populations likely far exceeded it (Spielman and D'Antonio, 2001). The terror of a yellow fever outbreak is something which we, in the developed world, are thankfully spared. The disease's symptoms are truly horrific in nature: the illness begins with fever, chills, and muscle pain so extreme that people often believe their bones are breaking. The disease commonly causes blinding headaches and stabbing pains in the eyes. In 15-25% of people it infects, the disease moves

on to a deadly second phase: the liver fails, turning the skin yellow; the eyes turn blood red; blood begins to ooze from the nose and mouth, and the internal hemorrhaging causes blood to flow into the stomach, where it later causes a telltale black vomit. At this stage death is inevitable and follows in the next few days (Monath, 2001). Whilst Max Theiler in 1937 formulated an effective vaccine for yellow fever (Norrby, 2007), there remains no effective treatment today.

Before Walter Reed and Jesse Lazear discovered that yellow fever is vectored by *Ae. aegypti* mosquitoes at the turn of twentieth century, there was essentially no way to combat outbreaks of this disease. Like malaria, up until the twentieth century its origins were a complete mystery. The uncertainty surrounding its transmission amplified the fear it installed in communities where yellow fever outbreaks occurred. Philadelphia in the 1790s was the busiest port in the newly independent United States of America. However, in 1793 the city was paralysed by a disastrous outbreak of yellow fever (see Figure 1.5). Those that survived the first few weeks of the epidemic (including President Washington) fled to the countryside (Spielman and D’Antonio, 2001). The arrival of winter eventually caused an end to the epidemic, by killing the *Ae. aegypti* mosquitoes that vector it, but the human toll was enormous - over 5,000 deaths, roughly 10% of its population at the time, perished (Powell, 1949).

Yellow fever also wreaked havoc with the colonial powers by killing expeditionaries sent to colonise parts of Africa and the New World. Indeed, the legend of the “Flying Dutchman”, a ghost vessel fated to haunt the seas, probably had roots in the fate of real expeditions (Vainio, Cutts, and World Health Organisation, 1998). In a landmark paper, Acemoglu, Johnson, and Robinson, 2000 attribute differences in the economic development of nations today to the differing risks of mortality (mainly due to yellow fever and malaria) faced by European settlers between the 17th and 19th centuries. In countries where the risk

of yellow fever was greatest, the colonial powers created extractive states, typified by the Belgian Republic of Congo. The purpose of these states was to expropriate the country's resources with the minimum investment necessary, and so these states lacked many of the key requirements for economic development, for example, property rights. In other, less disease-ridden areas, the Europeans set up "Neo-Europes", with many of the laws and institutions that were commonplace in Europe at the time, for example in Australia and the United States. Acemoglu, Johnson, and Robinson, 2000 argue that these differences in nascent institutions are still felt today in their daughter societies, and have a real effect on their rate of economic development.

Whilst there exist other mosquito borne diseases that have affected the course of history I do not discuss them here, as their influence was likely less extreme than malaria and yellow fever. However, I argue in Section 1.4, that their importance in the developing world today is paramount and is a major contributor to the constellation of sicknesses experienced by those most vulnerable.

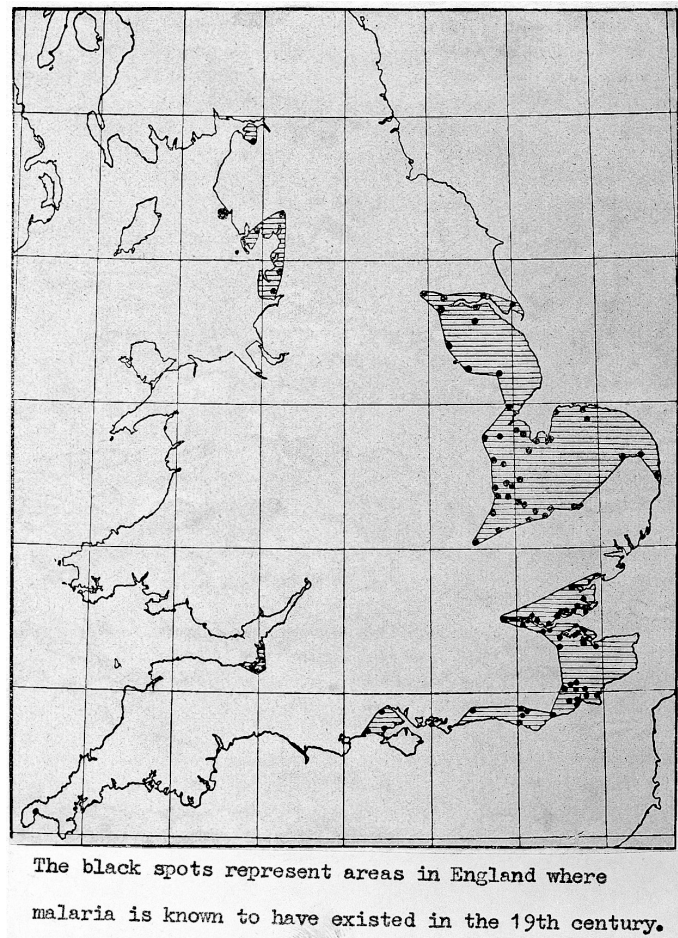


Figure 1.3: The geographical distribution of indigenous malaria (black dots) and marshlands (shaded areas) in England from a survey undertaken in 1864. Image reproduced from the Wellcome collection under the creative commons licence and is based on James, 1929.

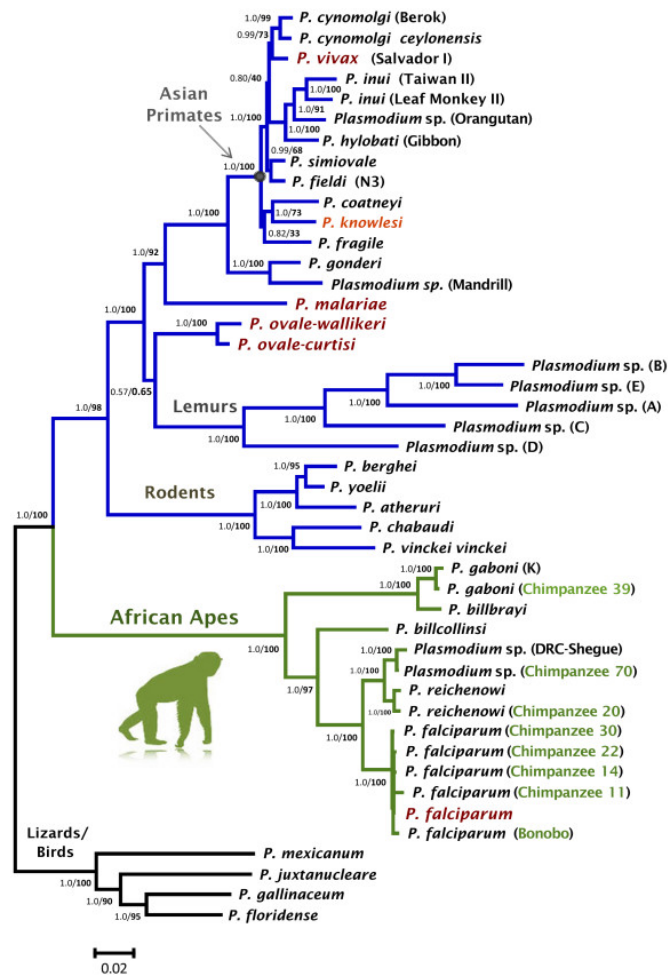


Figure 1.4: A phylogeny of *Plasmodium* spp. based on mitochondrial DNA. The human malarial parasites are indicated in red. Image reproduced from Pacheco et al., 2013.

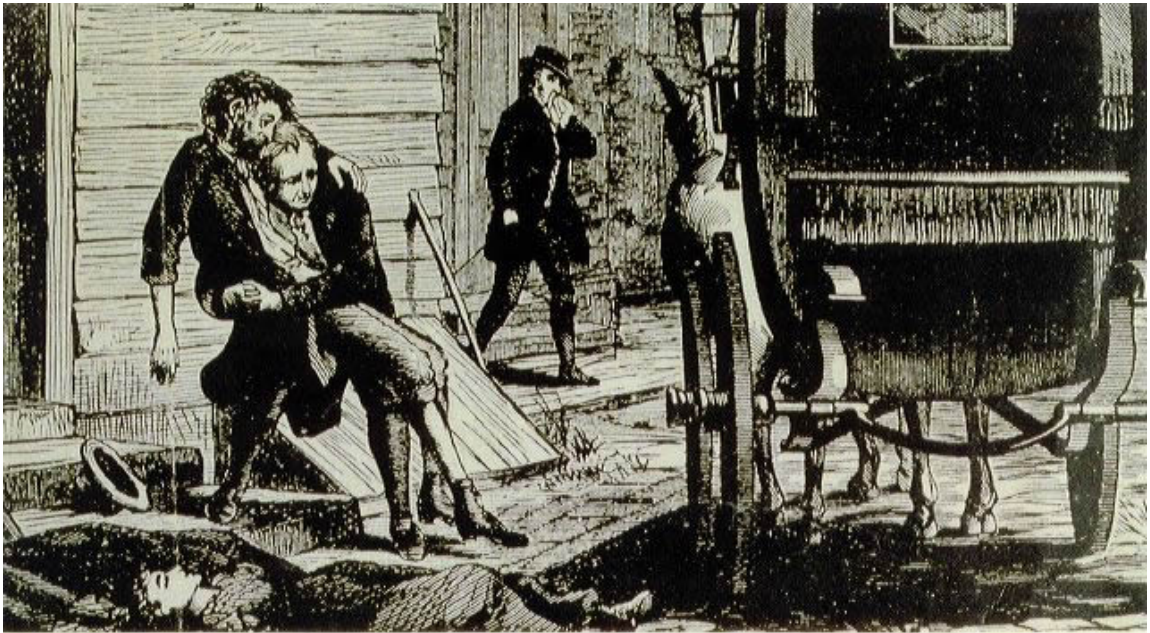


Figure 1.5: A scene depicting life and death in Philadelphia during the 1793 outbreak of yellow fever. Image reproduced from Gumm, 2010.

1.4 The impact of mosquito borne disease in the developing world today

Something is wrong here. War, disease, death, destruction, hunger, filth, poverty, torture, crime, corruption, and the Ice Capades. Something is definitely wrong. This is not good work...I firmly believe, looking at these results, that if there is a God, it has to be a man. No woman could or would ever fuck things up like this.

– The satirist George Carlin, 1937-2008

Before attempting to estimate the costs of malaria to the world today, it is worth remembering that natural selection has already made its own macabre valuation of the disease's devastating power. J.B.S. Haldane in 1948 was the first to propose the "malaria hypothesis" - that certain genetic polymorphisms, particularly those that affect red blood cells, have been selected for because of the protection they provide against malaria, even though they often confer severe fitness costs to the individual (Haldane, 1949). Haldane originally proposed the hypothesis to explain the high prevalence of thalassemia (an illness causing anaemia) along the Mediterranean coast, where malaria used to be endemic, which, he believed, might make individuals more resistant to malaria infection. An individual that is heterozygous for the thalassemia trait may not experience any symptoms, but homozygotes have a more serious version of the illness, whose treatment involves regular blood transfusions, and historically would have almost certainly died before reaching reproductive age. More recently, Haldane's conjecture has been confirmed, with evidence demonstrating that heterozygous and homozygous individuals with certain α^+ thalassemias have a risk of malaria disease that is about one half of normal (Allen et al., 1997). The link between thalassemia and malaria is thought to be so dependable that researchers have used markers on thalassemia genes to trace the migrations of the ancient Greeks, Phoenicians,

451 and Malayo-polynesians (all civilisations likely plagued by malaria; Cavalli-Sforza, 2001).
452 There are, however, polymorphisms that impose an even greater risk of mortality than
453 thalassemia.

454 The sickle cell trait is associated with a point mutation in the gene for the beta chain
455 of haemoglobin S, which leads red blood cells to adopt a rigid ‘S’ shape when oxygen levels
456 in the blood are low (Pauling et al., 1949). The sickle cell trait provides 90% protection
457 against *P. falciparum* malaria when inherited from a single parent, even though the same
458 allele, when inherited from both parents, is fatal (Carter and Mendis, 2002). In parts of
459 Africa where *P. falciparum* malaria is rife, the frequency of the sickle cell allele has been
460 estimated to be from 15% to 30% (Figure 1.6; Livingstone, 2009; Cavalli-Sforza, Menozzi,
461 and Piazza, 1994; Piel et al., 2010), whose upper bound would mean that roughly 1 in 10
462 children are born homozygotes and will likely die soon after birth. Essentially, the risk of
463 death from malaria is so high that naturally selection allows a regularly fatal mutation to
464 circulate in the population (Sachs and Malaney, 2002).

465 The burden of malaria - like for many other infectious diseases - is not distributed
466 equitably around the globe, driven by the particular vector species present, environmental
467 factors (Carter and Mendis, 2002), and differences in economic development (Sachs and
468 Malaney, 2002). Over 90% of malaria deaths occur in Sub-Saharan Africa, the majority
469 of whom are children under five years old (Murray et al., 2012), and estimates suggest
470 that there are 0.2 to 0.6 billion cases of malaria occur globally each year (Murray et al.,
471 2012; World Health Organisation, 2017c). The symptoms of malaria infection take on
472 many different forms according to which of the five species of *Plasmodium* parasite an
473 individual is infected with. Common to all infections are periodic episodes of fevers with
474 chills, sweating, and rigor (Carter and Mendis, 2002). Malaria also has symptoms common
475 to other infectious illnesses, including headaches, body ache, and nausea. Infection with

P. falciparum causes the most lethal form of the disease, which can lead to severe anaemia and major dysfunction of the lungs, kidney, liver, and a much-feared brain condition called *cerebral malaria*. Cerebral malaria is an encephalopathy that causes coma, seizures and life-long neurological problems for those affected (mostly children) and results death in 15% to 50% of patients (Idro, Jenkins, and Newton, 2005; MacPherson et al., 1985). Overall, malaria causes the premature deaths of between 0.5 and 1.2 million people each year (Murray et al., 2012; World Health Organisation, 2015) - about equal to the combined population of Glasgow and Edinburgh or the total population of Estonia.

Malaria weighs so significantly on affected societies that it traps them in poverty. A cross-country regression analysis that controlled for levels of human capital, life expectancy, initial income, as well as a host of geographical factors, concluded that countries where a high proportion of the population lived in regions of *P. falciparum* transmission had an annual GDP growth rate that was 1.3% lower, on average, than other countries between 1965 and 1990 (Gallup and Sachs, 2001). This implies that two countries - one in a malarial zone, another not affected by the disease - with the same income and population size in 1965, which experienced similar population growth rates over the period, would have a minimum difference in standard of living (measured by GDP per capita) of 38% in 1990. This difference of GDP per capita is comparable to that between Sudan and Vietnam today! There are also numerous other, indirect, effects of malaria on those societies it plagues. For example, the inflated risk of childhood mortality means that women tend to have more children (Yamada, 1985), meaning that families invest less in each child's education, and contributes to an already severe gender imbalance in the workplace in the world's poorest countries (Sachs and Malaney, 2002). Since adults in malaria-affected regions develop partial immunity to the disease, children bear the bulk of the disease burden. Malaria causes the death of many pre-school children, but older children also suffer from its effects,

501 resulting in missed days of schooling. The effects of this on children's education are not
502 trivial. For example, a study on absenteeism in Kenyan schools concluded that 11% of
503 school days in primary school are missed because of malaria, and over 4% in secondary
504 school (Leighton and Foster, 1993).

505 Overall, the global footprint of malaria is enormous and inflicts suffering on those
506 most vulnerable. Unfortunately, there are many other mosquito borne diseases today that
507 also inflict suffering whose importance is growing. The dengue virus is a member of the
508 flavivirus family that includes the viruses that cause yellow fever and Japanese encephalitis
509 (Westaway and Blok, 1997). The pandemic today consists of four viral subtypes that co-
510 circulate. Infection with one of these subtypes confers lifetime immunity to that subtype
511 only and, because of this, individuals in hyper-endemic areas (where all subtypes circulate)
512 often become infected with all four subtypes during their lifetime (Gubler, 1998). Later
513 infection with a second strain that is different to the first is more likely to result in more
514 serious illness (Halstead, 1988).

515 Dengue virus infection in humans causes a spectrum of symptoms ranging from no
516 noticeable effect to mild fever to severe and possibly fatal hemorrhagic disease (World
517 Health Organisation, 2009). Dengue fever is primarily an illness of older children and adults
518 (Gubler, 1998). The set of symptoms are similar to other febrile conditions (including
519 malaria) that presumably contribute to its misdiagnosis, and includes headaches, nausea,
520 body aches and vomiting (World Health Organisation, 2009). Dengue fever is rarely fatal
521 and typically lasts less than a week (Gubler, 1998), however, dengue infection can cause a
522 much more serious condition known as *dengue hemorrhagic fever* (DHF). DHF has similar
523 early symptoms as dengue fever but can result in bleeding in the skin, from the gums, and
524 internally, as well as organ impairment (World Health Organisation, 2009). Unfortunately,
525 the serious illnesses caused by dengue infection, including DHF, and another potentially

fatal condition referred to as *dengue shock syndrome*, predominantly prey on children, usually under the age of 15 (Dietz et al., 1996; World Health Organisation, 2009).

The first record of a major outbreak thought to be caused by dengue occurred in the seventeenth century in the Caribbean (McSherry, 1982; Dick et al., 2012), with similar epidemics occurring Asia and Africa probably about a century later (Gubler, 1998). However, accounts of illness similar to that of dengue exist that predate this period. The earliest record found yet is in a Chinese medical encyclopaedia from the Chin Dynasty (265 to 420 AD; Nobuchi, 1979). The disease was referred to as “water poison” by the Chinese and, in an impressive epidemic leap - over one and a half thousand years before Ronald Ross’ discovery that mosquitoes transmit malaria - was thought to connected with flying insects associated with water!

From the emergence of the first likely outbreaks of dengue in the Americas in the 1600s until the 1940s, the pattern of disease were was characterised by relatively infrequent, often large, outbreaks (Gubler, 1998; Dick et al., 2012). The ecological disruption wreaked by the second world war likely created the ideal conditions for the disease to spread. In particular, troop movements probably contributed to expansion of territory of *Ae. aegypti*, which rose in prominence to become, now, the predominant vector of dengue (Gubler, 1998). After the second world war, eradication programmes, fuelled by the recent creation of DDT, resulted in large reductions in the population densities of *Ae. aegypti*, particularly in the Americas (Gubler, 1998), and there was a commensurate fall in cases. The 1970s saw a deterioration of the vector control campaigns, and by the 1990s *Ae. aegypti* had regained much of the territory it held before eradication programmes began (Figure 1.7; Gubler, 1998). Dengue fever inevitably followed on the coat-tails of resurgent *Ae. aegypti* populations, and by the 1980s the Americas were experiencing major epidemics of the disease, in some countries that had been free of the disease for 130 years (Gubler, 1998). Today, the World Health

551 Organisation estimates that over 40% of the world's population are at risk of dengue, with
552 70% of those at risk living in the Asia Pacific countries (World Health Organisation, 2017b).
553 Recent estimates suggest that there may be 100 million symptomatic illnesses caused each
554 year by infection with the dengue virus (Bhatt et al., 2013), resulting in 500,000 cases
555 of DHF (Guha-Sapir and Schimmer, 2005). Dengue fever's meteoric rise in importance,
556 means that it now causes more death and illness in humans than any other virus carried
557 by arthropods. This group of *arboviruses* includes West Nile Virus, Japanese encephalitis,
558 and yellow fever, whose impact I discuss now.

559 It is remarkable that an outbreak of yellow fever in Asia has not yet occurred. The
560 conditions are certainly ripe for the yellow fever virus to emerge, and sustain itself: its
561 predominant vector *Ae. aegypti* is common throughout Asia (Gubler, 2004; Kraemer et
562 al., 2015) and there is evidence that the local population of *Ae. aegypti* can vector yellow
563 fever (Tabachnick et al., 1985). Travellers to many of these countries are not required to
564 provide a yellow fever certificate (Monath, 2001; Jentes et al., 2011), suggesting that the
565 local populace are largely susceptible to the disease. Furthermore, numerous species of
566 monkey are found in Asia, some of which could presumably act as a reservoir for the virus,
567 allowing the disease to become endemic in the region. This has lead to calls for Asian
568 countries to change their vaccination policy for travellers to the region, and heightened
569 vigilance for yellow fever cases (Jentes et al., 2011).

570 Regardless of the potentially disastrous impact of yellow fever on Asia, the disease
571 already forges a lethal path in South America and Africa, where it may infect 200,000
572 people annually (Monath, 2001). In regions of Africa where the population is seasonally
573 exposed to the virus, children, who are yet to acquire natural immunity to the disease, are
574 at highest risk of illness (Monath, 2001). Unfortunately, the disease can persist through
575 the dry season because it is vertically transmitted in mosquitoes, meaning a mosquito's

Disease	Annual number of cases	Annual number of deaths	Annual disability-adjusted life years
Malaria	212,000,000	429,000	NA
Dengue	96,000,000	9,110	1,892,000
Lymphatic filariasis	38,464,000	NA	2,075,000
Chikungunya (Americas)	693,000	NA	NA
Zika virus disease (Americas)	500,000	NA	NA
Yellow fever (Africa)	130,000	500	31,000
Japanese encephalitis	42,500	9,250	431,552
West Nile fever	2,588	111	NA

Table 1.1: The estimated global burden of the major mosquito-borne diseases as of March 2017 (World Health Organisation, 2017a). Note the values in the table represent the central estimates presented by World Health Organisation, 2017a and “NA” indicates data not available.

ova, harbouring the virus, has been shown to be able to lie in tree holes (not the primary aquatic habitat of *Ae. aegypti*) ready to hatch when the rains resume (Monath, 2001; Gubler, 2004). In both South America and Africa, it is the poorest, who work in agriculture in close proximity to the simian hosts of the disease, that are most at risk (Monath, 2001; Gubler, 2004). Indeed, hunters in these regions are known to avoid areas of the forest devoid of the sound of monkeys; this can signal that a recent outbreak of yellow fever has wiped out a local population of monkeys (Spielman and D’Antonio, 2001).

There are many other arboviruses that pose significant threat to swathes of the developing world (see Table 1.1). I do not attempt here to do justice to all of these diseases and, instead, briefly describe two of them: chikungunya and Zika. ‘Chikungunya’ is a word from the language of the Makonde people of southeast Tanzania and northern Mozambique, and translated literally means ‘that which bends up’ (Charrel, 2007) or ‘to walk bent over’ (Pialoux et al., 2007a). The disease is thus called because it is associated with intense pain around the joints, which can incapacitate an affected person (Charrel, 2007; Pialoux et al., 2007a). Chikungunya has been associated with large scale epidemics of illness in the developing world. In 2007, an outbreak of the disease swept through the islands of the Indian Ocean. Reunion has only 770,000 human inhabitants, yet there were 265,000 clinical cases

593 reported during this epidemic (Charrel, 2007). Like dengue and yellow fever, chikungunya
594 is vectored by mosquitoes of the genus *Aedes*, and in different parts of the world, different
595 species are associated with its transmission (although it is largely vectored by *Ae. aegypti*
596 and *Ae. albopictus*). In the Indian Ocean epidemic, the likely vector species was *A. al-*
597 *bopictus* - an aggressive anthropophilic mosquito - whose worldwide prevalence has shifted,
598 and largely, increased over the past 40 years (Charrel, 2007; Kraemer et al., 2015). In an
599 Indian outbreak in 2006, which caused 1.4 million cases of the disease, the main vector was
600 *Ae. aegypti* (Pialoux et al., 2007a), whose increased worldwide distribution has already
601 been discussed.

602 The Zika virus was first discovered in 1947 in the serum of a rhesus monkey caged in
603 the canopy of the Zika forest in Uganda (Dick, Kitchen, and Haddow, 1952). Whilst a
604 handful of Zika cases were recorded in humans (mostly in Africa), the disease was mostly
605 forgotten about, until an outbreak of illness occurred on a small island in the South Pacific
606 in 2007. In April and May of that year, a number of people on Yap island in the Federal
607 States of Micronesia began to suffer from symptoms encompassing rash, fever and arthritis
608 (Duffy et al., 2009). Attending physicians noted that the disease's symptoms were distinct
609 from dengue, and serum samples from the patients suggested that the aetiological agent
610 was the Zika virus (Duffy et al., 2009; Hayes, 2009). No further cases of Zika infection
611 were detected in the South Pacific until 2013, when an explosive outbreak occurred in
612 French Polynesia, with an estimated 28,000 patients seeking medical care (Musso, Nilles,
613 and Cao-Lormeau, 2014). The disease then spread through the rest of the South Pacific
614 (Musso, Nilles, and Cao-Lormeau, 2014) and eventually reached Brazil (Faria et al., 2016),
615 where in 2015 an outbreak began in the city of Feira de Santana in the Bahia region of
616 Brazil (Campos, Bandeira, and Sardi, 2015), where recent modelling work has suggested
617 that the virus infected a large proportion of the city's inhabitants (Lourenco et al., 2017).

By January 2016, at least 30,000 cases of the disease were reported in Brazil (Faria et al., 2016), and tragically the disease is widely believed to be behind the recent rise of microcephaly - a birth defect where the baby's head is smaller than that of a normal child, which can have lifelong neurological impacts - in the region (Mlakar et al., 2016; Oliveira Melo et al., 2016). Zika virus is known to be vectored by mosquitoes of the genus *Aedes* (Hayes, 2009), including *Ae. aegypti* (Campos, Bandeira, and Sardi, 2015; Faria et al., 2016), whose ubiquity in many parts of the world bodes ominously for the future of this epidemic. Further, the Zika virus has been detected in the semen of an infected man (Musso et al., 2015) which supports the earlier suspicion that this disease can be spread sexually (Foy et al., 2011). We are just beginning to understand the magnitude of the Zika epidemic, and the suffering it inflicts on affected individuals. Also, as the disease evolves to avoid our immune responses, it is possible that its virulence will increase (Ewald, 1994) and we could see other serious manifestations of Zika infection.

Mosquitoes also carry a range of other diseases, which affect both humans and animals, and this space cannot do these justice (see Table 1.1). One condition which, I believe, has not received adequate attention, is *filariasis*, and I briefly discuss it now. Although this disease does not regularly cause death, it nonetheless inflicts considerable suffering on those unfortunate to become infected with the aetiological filarial nematodes (Taylor, Hoerauf, and Bockarie, 2010; Murray et al., 2013). Filariasis can manifest in *elephantiasis*, which with little doubt, is one of the most stigmatising of tropical diseases. Elephantiasis occurs when the fluid in the lymphatic system becomes blocked, which frequently causes the legs or genitals to swell to enormous proportions (Farrar et al., 2013). As anyone who has ever seen an individual with these symptoms can testify, it is hard to imagine a disease that could cause more gruesome transfiguration in those it affects. Historically, the scientific study of filariasis played a pivotal role in the discovery that mosquitoes could

643 vector diseases, and it is this historical context which I describe next.

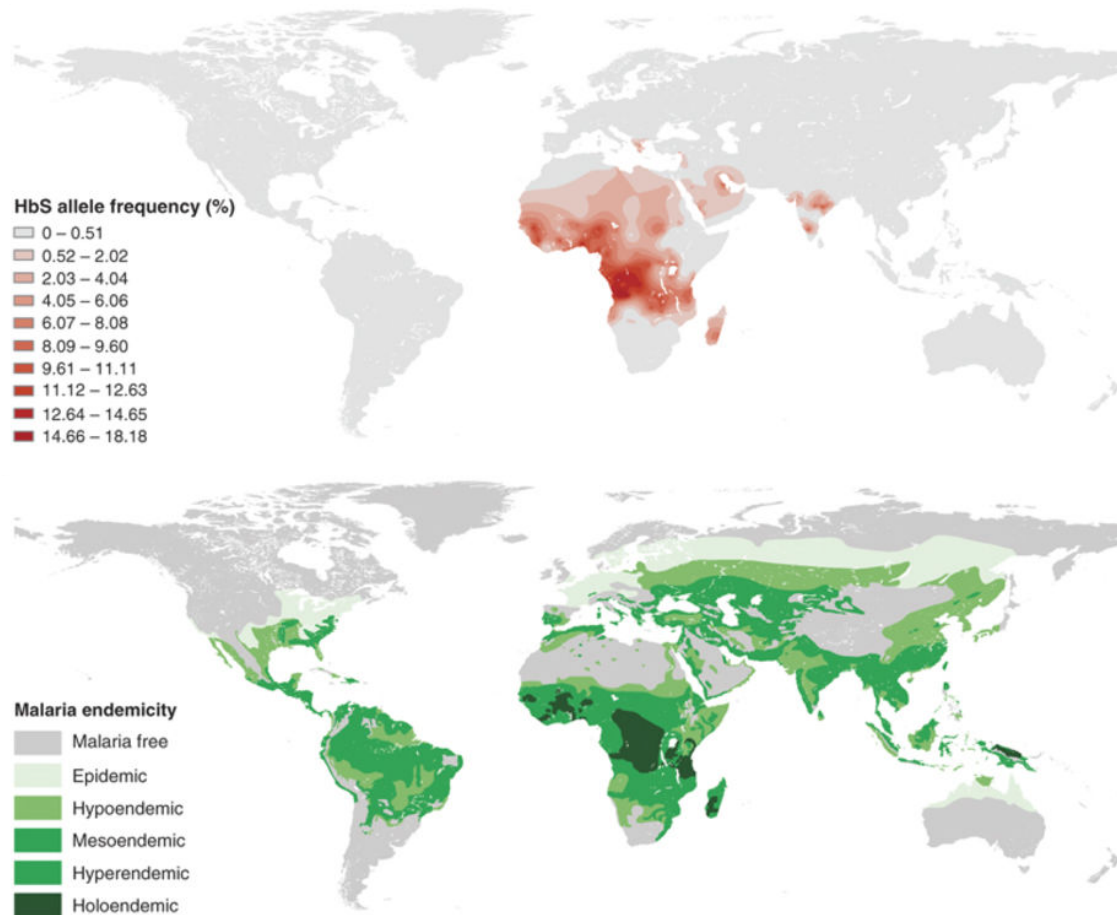


Figure 1.6: **Confirming Haldane's hypothesis: the distribution of sickle haemoglobin (top) and a historical map of malaria endemicity (bottom).** Image reproduced from Piel et al., 2010.

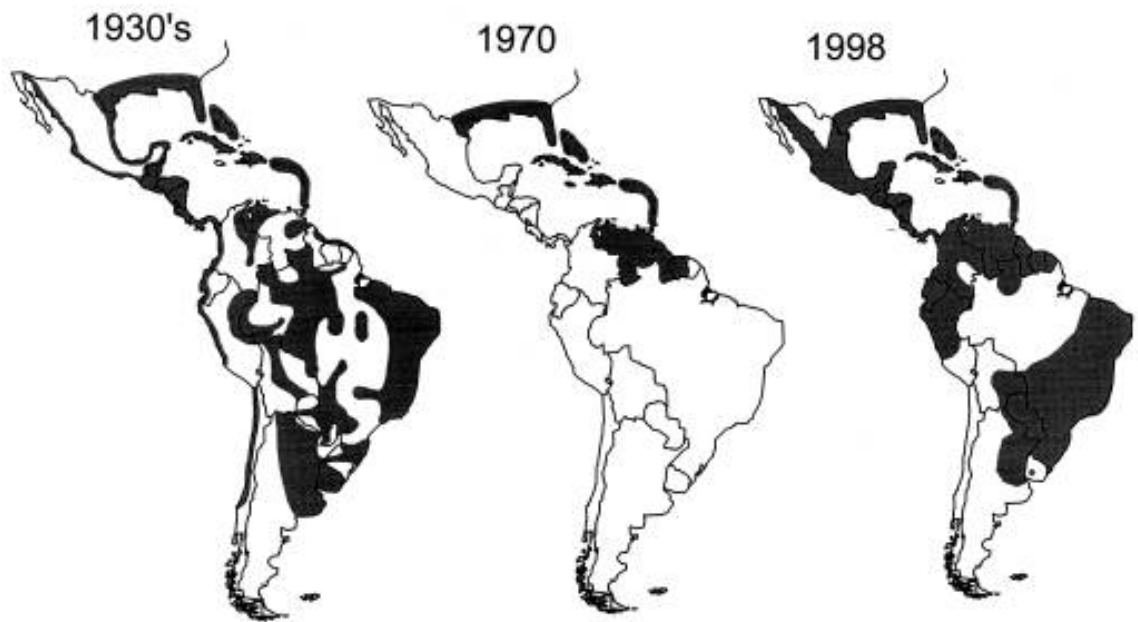


Figure 1.7: The distribution of *Ae. aegypti* mosquitoes in South America before (left), immediately after (middle), and long after (right) the campaign to eradicate this species from the region. Image reproduced from Gubler, 1998.

1.5 Mosquitoes as vectors

The etymological origin of the word “malaria” is rooted in nineteenth century Italy. At the time, it was widely believed that fever was caused by a “miasma”, or poisoning of the air. Indeed, the Italian “mal’aria” literally translated means “bad air”, and this became a popular term for the disease in this period (Snowden, 2008). The path by which mosquitoes came to be implicated in the transmission of malaria is hotly contested, and is a source of national pride for some historians of science. Here, I outline a brief narrative of history described by Spielman and D’Antonio, 2001, although encourage interested readers to consult Ross, 1923 and Grassi, 1924 for two different perspectives on this subject.

The discovery that malaria is transmitted by anopheline mosquitoes actually begins with research into lymphatic filariasis. Elephantiasis made a strong impression on Dr Patrick Manson, a British medical officer stationed in Taiwan (then called Formosa, which was part of China) in the 1870s. In 1875, on a holiday from his service in China, Manson visited London and bought a compound microscope. When he returned to China, he set about investigating the causes of elephantiasis, at first, by conducting autopsies in secret (at the time this practice was outlawed in China) of bodies of patients who had died whilst suffering from this disease. In one patient, Manson found threadlike worms, about the length of a person’s arm, coiled inside within their lymph nodes. In an experiment of questionable ethics, Manson recruited his gardener, Hinlo, whose blood teemed with the filarial parasites, to try to demonstrate that the pathogen could be transmitted from man to mosquito. Hinlo slept within a tent Manson dubbed the “sleeping house” (how much sleep Hinlo got that night is not recorded) whilst *Culex quinquefasciatus* mosquitoes (also known as the southern house mosquito) fed on him. During the next five days, Manson dissected the recaptured mosquitoes and was able to find larval stages of the filarial parasite in some of the specimens. Manson’s achievement was monumental - he was the first to

669 experimentally demonstrate the role of mosquitoes in spreading infectious disease!

670 Around the same time that Manson was performing his filariasis experiments, a French
671 Army doctor named Charles Laveran was stationed in a malaria ridden outpost of the
672 French Foreign Legion in Algeria. Out of concern for the military, not the colonial sub-
673 jects, the French government had sent Laveran to Algeria. Here Laveran was free to perform
674 autopsies and undertake his own experiments. In November of 1880, Laveran drew blood
675 from a patient suffering from extreme fever and placed a droplet of the blood on a mi-
676 croscope slide. Under the magnification, he could see that the blood was alive with tiny
677 protozoan creatures, which he called *Oscillaria malariae* (Cox, 2010). These were the first
678 observations of the malarial parasite, which would later be renamed as *Plasmodium*, and
679 resulted in Laveran being awarded the Nobel Prize in 1907.

680 Laveran's discovery is rightly lauded for its elucidation of the causes of malaria. What
681 remained unclear at the time, however, was how the disease was transmitted to humans.
682 This required the work of another colonial doctor, Ronald Ross, who was serving in India
683 during the 1880s and 1890s. Ross was born in the Himalayas as the eldest of a British
684 general's ten children. He was a true polymath and enjoyed a passion for mathematics and
685 poetry as a child. At the age of twenty-seven, having recently been admitted as a member
686 of the Royal College of Surgeons, Ross was posted to Bangalore.

687 In 1894 in London, Ross was on a break from his service in India and approached
688 Patrick Manson, who was then famous for his work on filariasis. Ross shared with Manson
689 his theory (others, decades before Ross, had also had the same theory) that malaria was
690 also vectored by mosquitoes and Manson agreed with him. Spurred on by Manson, Ross
691 boarded a boat back to India. Ross was so excited about the prospect of finding malarial
692 parasites in humans that he apparently examined the blood of all passengers on the vessel.
693 An excitable graduate student can certainly appreciate his enthusiasm.

694 Interrupted by major cholera outbreaks on the subcontinent, Ross' work continued,
695 without success, for three years. During this time, Ross carried out numerous dissections
696 of a grey mosquito - probably *Culex pipiens* (the common house mosquito) - and a brindled
697 one - probably *Ae. aegypti*, the carrier of yellow fever, dengue and Zika - and in all cases
698 failed to find malarial parasites. It was not until 1897, when his lab assistant raised a
699 number of larvae that matured into "big brown fellows" - a mosquito species which Ross
700 had not seen before (anophelines) - that Ross would find success. Ross fed this mosquito
701 species on malarious patients then set about dissecting the mosquito specimens a few days
702 after the initial infection. On examining the gut of one, Ross saw dark, pigmented cells that
703 were clearly not part of the mosquito itself. These observations were validated by similar
704 dissections the following day. Ross had done it - he had found traces of the malarial
705 parasite in mosquitoes for the first time! In 1897, work by the Italian scientists Bignami,
706 Golgi (who won the Nobel prize for his work on the subcellular compartment which is now
707 known as the "Golgi apparatus"), and Grassi amongst others closed the transmission cycle.
708 They did did this by infecting mosquitoes with malaria by feeding them on febrile human
709 subjects, and then produced an infection in uninfected humans, by feeding the infected
710 mosquitoes on them.

711 Following hotly on the success of Ross and others in demonstrating the anopheline
712 mosquito's role in transmitting malaria, the mystery of the epidemiology of yellow fever
713 was elucidated, again by the work done by an army physician. In 1900, Walter Reed was
714 stationed on Cuba with his assistant Jesse Lazear, following the occupation of the island
715 by their native US at the end of the Spanish-American war. In a series of experiments that
716 would claim the life of Lazear at age 31 (after he purposely was bitten by a mosquito which
717 had fed on the blood of yellow fever patients), the work of the pair demonstrated that *Ae.*
718 *aegypti* was able to transmit the disease. The definitive experiment, performed a month

719 after Lazear's death, involved "volunteers" (this experiment preceded the discovery of the
 720 yellow fever vaccine by 37 years) sleeping in one of two tents: the less lucky volunteers slept
 721 in a tent with mosquitoes which had fed on patients suffering from yellow fever; the slightly
 722 more fortunate individuals slept in another tent, with blankets soiled with the telltale black
 723 vomit of yellow fever victims. Only those in the mosquito tent became ill, confirming that
 724 mosquitoes were typically required for transmission. Whilst a Pyrrhic victory for Lazear
 725 (and the unlucky participants of the experiment), this discovery was crucial in shaping
 726 vector control efforts that would soon be effective in curbing the incidence of this terrible
 727 disease. Despite his crucial part in uncovering the mosquito's role in yellow fever, Reed
 728 did not live much longer than many of his volunteers, and died in 1902, when his appendix
 729 ruptured.

730 1.6 Disease control since Ronald Ross

731 *The head of that statue was made of fine gold, its breast and its arms of silver, its*
 732 *belly and its thighs of bronze, its legs of iron, its feet partly of iron and partly of*
 733 *clay. You continued looking until a stone was cut out without hands, and it struck*
 734 *the statue on its feet of iron and clay and crushed them...Then the iron, the clay,*
 735 *the bronze, the silver and the gold were crushed all at the same time and became*
 736 *like chaff from the summer threshing floors; and the wind carried them away so*
 737 *that not a trace of them was found.*

738 – Daniel's interpretation of Nebuchadnezzar's dream, *Daniel 2:32-35, The Hebrew*
 739 *Bible*

740 Following the discovery of the mosquito's role in malarial transmission, leading Malar-
 741 iologist Giovanni Grassi declared malaria was a "giant with a feet of clay", a biblical

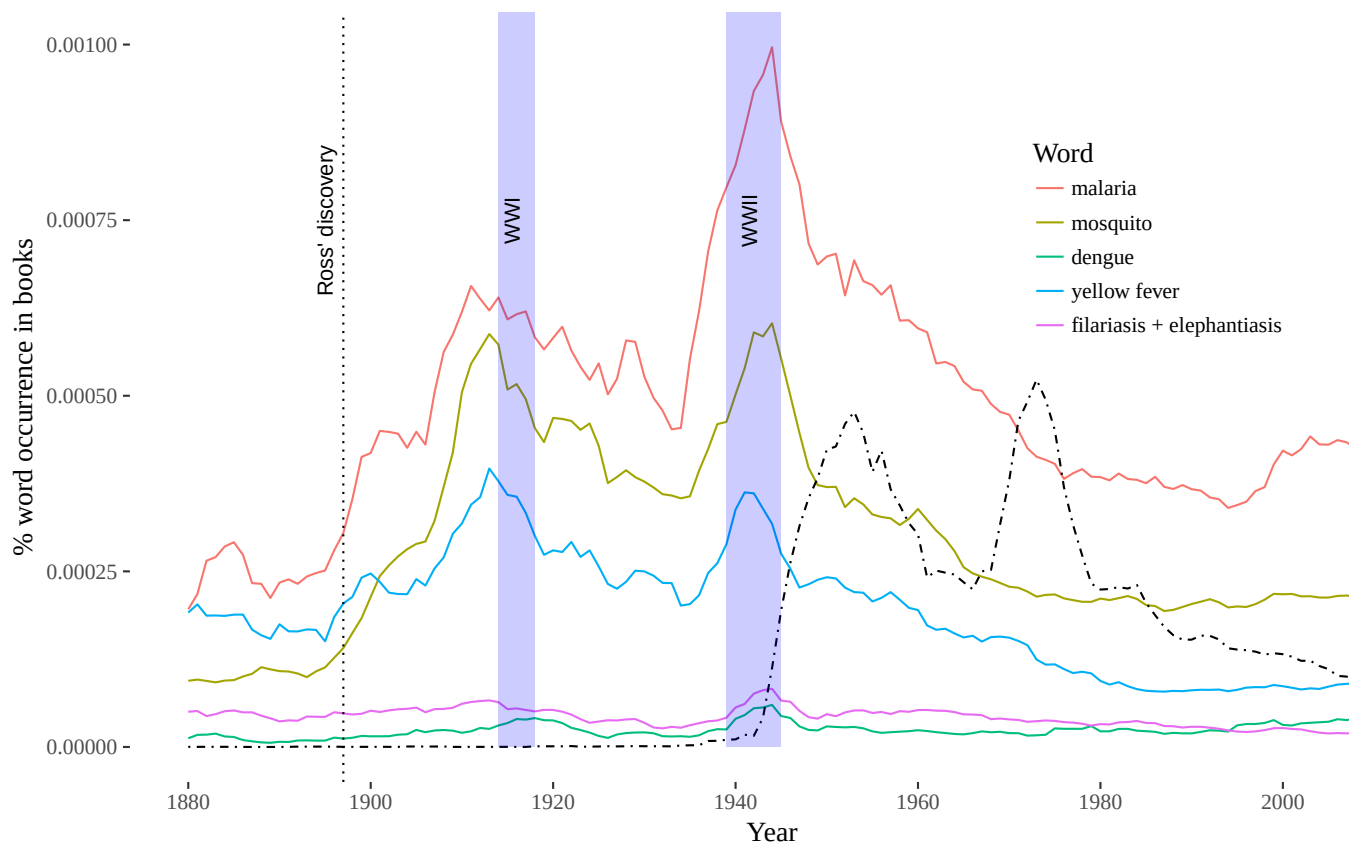


Figure 1.8: **The occurrence of phrases associated with vector borne diseases (the black dashed line represents DDT) in books published in the English language.** The corpus of books from which these quantities are derived is the “English 2012” Google nGram corpus, compiled from digitised books by Google. The requisite data were collected using the *ngramr* R package (Schmidt and Heckendorf, 2017), and represent two period moving average percentages of all n-grams (words for single words; pairs of words for two words) for all books published in that year.

reference (from the book of Daniel, whose original quote appears at the beginning of this section), meaning an obstacle which can be readily overcome (Sachs and Malaney, 2002). Sadly, more than a century on from Grassi’s declaration, malaria still stands firmly (albeit less firmly than 15 years ago (Bhatt et al., 2015) and much less firmly than 100 years ago (Snow et al., 2017)), its feet presumably made of sterner stuff than soil. Over the

747 course of the twentieth century, there have been a few unbridled successes, many partial
748 successes and, unfortunately, a fair number of failed attempts to curtail this disease. What
749 is often lost in historical descriptions of “failed” attempts to control malaria is that, whilst
750 the ultimate goal of globally eradicating the disease may have not been met, these con-
751 siderable efforts were not in vain; many people are now either alive, or live comparatively
752 better off lives, as a result of them. Much of this section considers the global efforts to
753 control malaria but, due to these vector borne diseases’ joint reliance on the mosquito and,
754 importantly, the measures used to control them, their history is, to some degree, shared,
755 which is illustrated in Figure 1.8. This figure shows the annual counts of a few terms
756 relating to vector borne diseases, including malaria, from 1880 to 2008, as a percentage
757 of all words published that year, derived from a corpus of books published in the English
758 language. Whilst this diagram should not be taken too literally, as it is certainly the case
759 that word usage does not directly translate into disease incidence, it nonetheless provides
760 an interesting view of the history of vector borne disease over the period. In this figure,
761 it can be seen that, whilst malaria and yellow fever (the red and turquoise lines) in 1880
762 occupied similar space in the printed pages of the time, in the twenty-first century malaria
763 is more popularly discussed. Also, this figure displays the meteoric rise in discussion of
764 mosquitoes (the gold line) after Ronald Ross’ work in 1897 demonstrated that malaria can
765 be transmitted by mosquitoes. Leading up to each of the world wars (indicated by the
766 blue shaded rectangles), there was increased discussion of the selected diseases, possibly
767 as English-speaking countries discussed ways to control the disease to give their troops
768 advantage. Finally, the black dot-dashed line illustrates how the insecticide DDT became
769 popular following its invention in the 1940s, after which it became the centrepiece of cam-
770 paigns to control yellow fever in the Americas (Packard, 2007) and a global programme to
771 eradicate malaria (Packard, 2007; Spielman and D’Antonio, 2001), which persisted until

the 1970s. As discussed, it is dangerous to ascribe movements in each of these series to the burden of disease or the adoption of certain interventions, however, I cannot resist making one observation: after the introduction of DDT (the black dot-dashed line), there appears to have been a substantial reduction in all the mosquito borne diseases (the red, green, turquoise and mauve lines). This somewhat derivative observation is supported by considerable evidence (see Figure 1.9 and Packard, 2007 for a summary of this evidence), and so I believe that this slight misuse of the data are warranted. Though DDT use is extremely controversial (the second peak in this series around 1970 to 1975 likely reflects discussion of its possible negative impact on humans and ecosystems; see my discussion below) and its use is banned in many countries, 380 scientists, including three Nobel laureates in medicine, across 57 countries signed a letter endorsing continued use of DDT (Taverne, 1999; Roberts, Manguin, and Mouchet, 2000).

Not wishing to venture into the politics of DDT usage here, I instead attempt to provide a history of mosquito borne disease control in the twentieth century, which, I believe is more or less supported by the data from Figure 1.8. Whilst the discovery of Ross concerned malaria, the first major victory against over mosquito borne disease was actually claimed against yellow fever, which I now discuss.

After Reed and Lazear demonstrated that *Ae. aegypti* mosquitoes were responsible for transmitting yellow fever, another US army man, also stationed in Cuba, took up the task of eradicating the disease from the island's capital, Havana (Spielman and D'Antonio, 2001). Major William Gorgas, an uncompromising man whose temperament would later cause him trouble in Panama, organised brigades of soldiers to patrol the streets under orders to search and destroy containers containing standing water, to which the aquatic stages of *Ae. aegypti* are adapted. Ponds were also oiled each week and residents punished if mosquito larvae were found on their premises. One of the first vector control campaigns in history

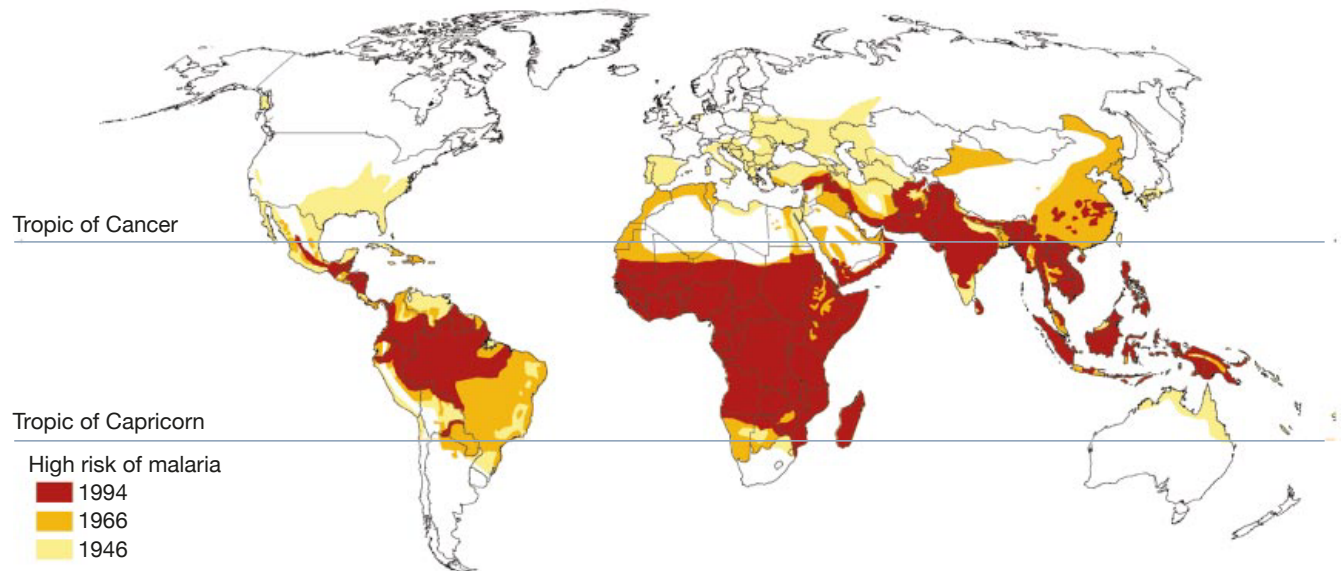


Figure 1.9: **The changing distribution of malaria from 1946 to 1994.** The distribution of disease in 1946 is before the global push for eradication began, whereas the map for 1966 shows prevalence in the midst of the campaign, and for 1994, the same, although long after it ended. Image reproduced from (Sachs and Malaney, 2002).

797 (see World Health Organisation, 2017a Annexe 2 for a list of other such campaigns) was a
 798 great success; after only five months, yellow fever was eradicated from Havana (Spielman
 799 and D’Antonio, 2001).

800 After his victory over yellow fever in Cuba, Gorgas was deployed to Panama, where
 801 America was in midst of attempting to complete the Panama Canal. In their efforts to
 802 construct the Panama canal, the French had lost some 20,000 workers, mainly to yellow
 803 fever and malaria (Packard, 2007). The Americans, who had taken over work on the canal
 804 in 1902, at first fared no better. In 1904, when Gorgas arrived in the country, he began a
 805 militaristic campaign, of a nature similar to that he had used in Cuba. Specifically, Gorgas
 806 divided the canal zone into 25 districts, each of which was placed under the control of

a sanitary inspector who oversaw between 20 and 100 labourers (Gorgas, 1915). Having identified the breeding habits of local malaria and yellow fever carrying mosquitoes, the teams set about destroying their aquatic habitats. The teams cleared the French gardens and water receptacles away from living quarters, and spread lava killing oil on larval aquatic habitats (Spielman and D’Antonio, 2001). He also sent soldiers into homes and barracks to swat adult mosquitoes, one by one (Spielman and D’Antonio, 2001). To kill *anophelines*, the teams drained marshy areas and filled them in with cement (Packard, 2007). Gorgas’ uncompromising nature won him few friends in the commission overseeing the construction of the canal, who even tried to replace him at one point (Spielman and D’Antonio, 2001). Fortunately for Gorgas, his success in Cuba had won him friends in high places, and the then President of the United States, Theodore Roosevelt, decreed that Gorgas should continue his work in Panama (Spielman and D’Antonio, 2001). (It is difficult to conceive of a field entomologist currying such favour today.) By 1906, Gorgas had managed to eradicate yellow fever from the canal zone (Spielman and D’Antonio, 2001), and malaria saw significant declines over the ensuing few years. Malaria was presumably harder to suppress because of the greater difficulty of targeting the anopheline mosquito vectors (Spielman and D’Antonio, 2001). Nonetheless, deaths from malaria declined from a maximum of about 16 per 1000 workers in July 1906, to a minimum of less than 3 per 1000 in December 1909 (Packard, 2007).

During the 1910s and 1920s, malaria control was split between various approaches, including vector control (primarily using the larvicide Paris Green; Packard, 2007), targeting the parasite through quinine and social and economic development programmes that emphasised the human ecological aspect of disease transmission. Schemes aimed at social and economic betterment were pioneered in Italy after the First World War, where so-called *bonifica integrale* or *bonification* policies were instituted. The most notable suc-

cesses of bonification were achieved under Mussolini's fascist government, who managed to drain the Pontine Marshes and eradicate malaria from a province previously gripped by malaria (Spielman and D'Antonio, 2001; Packard, 2007). The scheme began in 1929 and resulted in considerable areas of land being reclaimed for agriculture. This stimulated a large scale immigration to the area, which saw its population rise from 1637 people in 1928 to 60,000 in 1939 (Packard, 2007), largely due to those seeking employment in the newly available agricultural land. Along with land reclamation efforts, vector control (principally by attacking the larval stages of the mosquito using Paris Green) and quinine were integral parts of the project. Bonification schemes also focussed on the societal elements of disease transmission, and included installing high quality housing and setting up regional health clubs. Interestingly, these human aspects of disease control - notably the construction of mosquito-proof houses - have recently proven to be a growing research areas (Lindsay, Emerson, and Charlwood, 2002), with a meta-analysis estimating that individuals living in modern housing had 47% lower odds of contracting malaria compared to traditional housing (although the authors recognised that there was likely sampling bias in the individual studies; Tusting et al., 2015).

The emphasis of human and drug-based approaches to malaria was reflected in the discourse of the time. The League of Nations, founded after the end of the First World War, had a department called the Malaria Commission, whose role was to deal with the increased disease prevalence of this disease during and following the war (Packard, 2007). The first report in 1924 by the British Malariologist, Colonel Sydney James, emphasised the importance in "quininization", housing, education, nutrition and health in disease control. It has been suggested that its aversion to vector control may have stemmed from James' own personal experience in India, where Britain failed to eradicate malaria at Mian Mir, near Lahore (which are both in present day Pakistan; Packard, 2007; Bynum, 1994).

857 Whilst in the early 1930s, the prevailing attitude swung away from vector control, this was
858 about to change, spurred on by the notable success of Fred Soper in Brazil and, later, the
859 discovery of the potent insecticide *dichloro diphenyl trichloroethane*, better known by its
860 acronym, *DDT*.

861 In 1930, Raymond Shannon, an entomologist from the Rockefeller Foundation, was
862 taking a walk from his home in Natal, Brazil, when he noticed strange larvae in a puddle.
863 Shannon immediately recognised the danger of this small organism - he had found the ma-
864 jor African malarial vector *Anopheles gambiae* on Brazilian soil - and alerted the Brazilian
865 authorities. Apart from a few minor outbreaks in the intervening years, the impact of
866 *Anopheles gambiae* on the country was relatively minor until 1938, when a major outbreak
867 of malaria struck in the north-east of Brazil, resulting in over 100,000 disease cases (Spiel-
868 man and D'Antonio, 2001). Whilst *A. gambiae* no doubt played an important role in this
869 outbreak, it was also fuelled by a swollen population of migrant farmers from the interior,
870 where a devastating drought was raging (Packard, 2007). Reacting to this public health
871 disaster, Brazil's president created an antimalarial response unit staffed with 4000 workers,
872 and placed it under sole control of a Rockefeller associated named Fred Soper. Working
873 with the authority of a war General, Soper launched a war on *A. gambiae* in Brazil, and
874 exercised military rigour in his campaign. Cars were fumigated whenever drivers arrived at
875 checkpoints and ships were boarded and investigated for *Anopheles* infestation (Spielman
876 and D'Antonio, 2001). The most powerful weapon in Soper's arsenal was the larvicide
877 *copper acetoarsenite*, better known as *Paris Green* because of its vivid colour which, at the
878 time, was associated with the Impressionist art of the French Riviera. Soper directed the
879 extensive use of Paris Green, which exacted a heavily toll on *A. gambiae* populations. At
880 its peak, this African mosquito held terrain in over 18,000 square miles of territory over
881 three Brazilian states (see Figure 1.10; Spielman and D'Antonio, 2001), yet in a single year,

882 Soper had virtually eliminate the vector from the country!

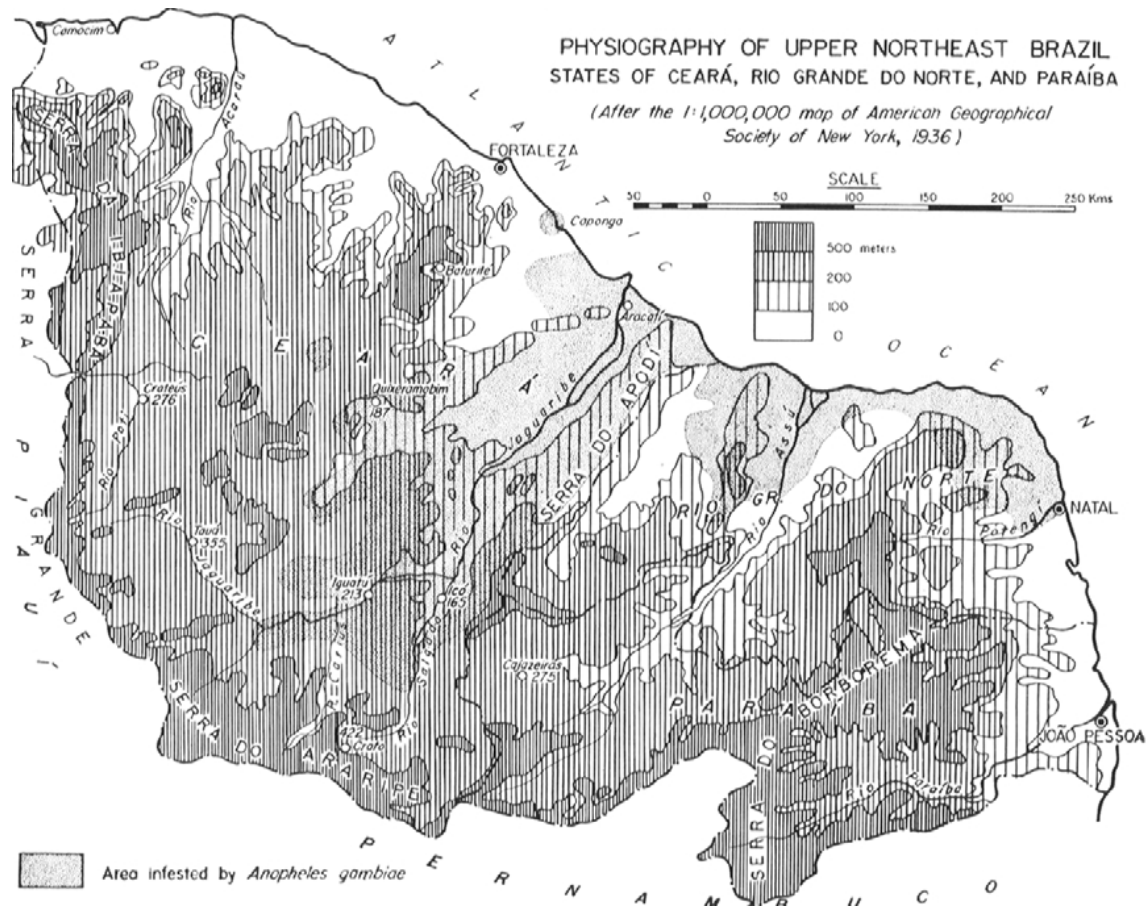


Figure 1.10: A map of the distribution of *Anopheles gambiae* and physical geography in Brazil c.1932. Image reproduced from (Soper and Wilson, 1943).

883 After Fred Soper's Brazilian victory over mosquito, vector control was now back at the
 884 centre of efforts to control mosquito borne disease before the start of the Second World War,
 885 and a wartime discovery was about to cement its position further. In the summer of 1943,
 886 a colleague of Soper's returned from Algeria with news of a new pesticide - then referred
 887 to as Neocide - which had demonstrated great efficacy in vector control (Spielman and
 888 D'Antonio, 2001). Soper subsequently travelled to Algeria and conducted trials on Axis

prisoners, who were successfully deloused with the chemical (Kinkela, 2011). This chemical, hitherto known by its chemical acronym DDT, was to become an integral part of the Allied war effort, with British and American factories manufacturing tonnes of it throughout the war period. Ironically, the chemical was actually derived from a German moth killing pesticide, which originated over fifty years beforehand (Spielman and D'Antonio, 2001). At the end of the war, a field trial of this promising pesticide DDT was conducted on the island of Sardinia, orchestrated by a partnership of the Italian government and the Rockefeller Foundation. Throughout the campaign, an extensive workforce sprayed over 2 million rooms in over 300,000 houses with DDT, and also vast quantities of the chemical were dropped on fields by Italian Air Force helicopters and planes (Spielman and D'Antonio, 2001). The campaign was a success - malaria was eradicated from the island - but, it was not unqualified because, after five years of effort and several million dollars spent, the malaria vector *Anopheles labranchiae* could still be found, which was, unfortunately, a foretelling of what was to follow.

In 1955, member states of the recently established World Health Organisation convened in Mexico City to vote on an ambitious proposal: a plan to globally eradicate malaria (known as the Global Malaria Eradication Programme (GMEP)). The linchpin of the proposed strategy lay in three letters: DDT. The proposal was passed, and over the next 14 years, over \$1.4 billion was spent on the eradication programme, the majority of which was spent on DDT (Spielman and D'Antonio, 2001). So strong was the faith in DDT that following the announcement of the eradication plan, grants for research in malariology dried up virtually overnight, and some universities, Harvard, for example, stopped teaching the subject entirely (Spielman and D'Antonio, 2001). When the eradication programme was passed, most believed that the elimination of disease in sub-Saharan Africa, where the bulk burden of disease lay, would follow after a sustained effort (Packard, 2007). Unfortunately,

914 this proved to be wildly optimistic - the campaign was successful in eliminating malaria
915 from mainland Europe and the southern United States, but in those worst affected parts of
916 the world, the disease stubbornly remained (see Figure 1.9). Before I discuss the reasons
917 behind the failed attempt to eradicate malaria from the face of the globe, I think it is
918 important to stress the campaign's successes. Hundreds of thousands, possibly millions, of
919 people have lived longer, less wretched lives as a direct result of the malaria eradication
920 programme and, today, peoples across many countries likely enjoy a higher living standard
921 because of it. The fact that the campaign did not live up to its ambitious expectations is
922 disappointing, but an argument can nonetheless be made that this programme was one of
923 the most impactful public health campaigns of the twentieth century.

924 Why didn't GMEP meet its goals? There were a complex nexus of factors that con-
925 tributed to this outcome, including waning international support following negative public-
926 ity for the use of DDT (after Rachel Carson published her polemic, *Silent Spring*, against
927 the use of the insecticide in 1962), the failure to recognise local ecological and epidemio-
928 logical variability (Nájera, 2001), the lack of research in mosquito vectors and malariology
929 in general following the campaign's initiation (Nájera, 2001; Nájera, González-Silva, and
930 Alonso, 2011), the loss of the US as a crucial programme donor in 1963 (Nájera, 2001),
931 climatic events (Packard, 2007) and resistance to chloroquine - a major antimalarial drug
932 which is still used today. Also, by design, the programme did not realistically aim to
933 achieve global eradication since it largely ignored Sub-Saharan Africa - a region contribut-
934 ing the majority of the world's disease cases (Nájera, 2001). Indeed, when the programme
935 was approved at the World Health Assembly in 1955 in Mexico, Sub-Saharan Africa was
936 represented by only three independent countries: Ethiopia, Liberia and South Africa, and
937 a handful of metropolitan governments of colonial powers (Nájera, 2001). Rather than dis-
938 cuss in detail all of the causes of GMEP's failure to achieve global eradication (see Nájera,

2001; Nájera, González-Silva, and Alonso, 2011 for a thorough analysis of GMEP), I focus on another leading cause of the programme's dispiriting results - mosquito resistance to DDT. The campaign relied heavily on the indoor spraying of the walls of houses with DDT, based on the assumption that female anophelines predominantly feed indoors and, afterwards, rest on the indoor walls. Unfortunately, whilst, on average, these claims were true, there were considerable exceptions to them, meaning that if the endophilic mosquitoes died because of the chemical, they were likely replaced by other, more exophilic, anophelines (Packard, 2007). Unfortunately, mosquitoes also evolved biochemical resistance to the pesticide, meaning that resistant vectors still fed on humans whilst they slept and rested on the indoor walls afterwards. Because of the experiences in Sardinia, where some mosquitoes remained following the onslaught of DDT, as well as similar experiences in Greece and Saudi Arabia, DDT resistance was actually known about before the start of the campaign and, because of this risk, "TIME IS OF THE ESSENCE" was written in capital letters on the original plan (Spielman and D'Antonio, 2001). Unfortunately, by 1969, when the WHO ostensibly pulled the plug on the eradication programme (and refocussed it towards disease control), this time had run out. Whilst the campaign ended, its legacy likely persists to this day, in some parts of the world where endemic malaria is now lost from memory.

In the 1980s and 1990s, malaria control strategy veered away from vector control and focussed on approaches that target the human stages of the parasite (Godfray, 2013), for example, using the highly effective drug artemisinin (recently discovered by Western science, although it was in use in Chinese medicine for over 2000 years before this; Klayman, 1985; Hsu, 2006) and efforts to develop a vaccine. Over the last twenty or so years, there has been a shift back towards vector control, perhaps driven by the difficulty in developing vaccines against the complex life stages of the malarial parasite (Godfray, 2013), and the growing presence of other mosquito borne diseases, like dengue fever (Gubler,

1998). Whilst these vector control efforts (catalysed by the establishment of the Millennium Development Goals in 2000) caused unprecedented declines in disease cases from 2000-2015 (Bhatt et al., 2015), it seems less certain that this trend will continue. More recently, the downward trend in case numbers appears to have flattened; in 2016, it has been estimated that there were 5 million more disease cases than in 2015 (World Health Organisation, 2017c). The reasons behind this growth in cases remain unclear, although it is feasible that donor fatigue, intervention coverage gaps (for example, fewer insecticide treated nets were delivered in 2016 than in 2015; World Health Organisation, 2017c), climate change, insecticide resistance and parasitic drug resistance could have contributed. Indeed, it seems there is still much we do not know about vector borne disease and its control and, importantly, the mosquito itself.

1.7 Why we need to understand mosquito vectors

...nature is almost infinite in her variety, which means that we can make very few absolute declarations about mosquitoes.

– Mosquito, 2002, Andrew Spielman and Michael D’Antonio

A key characteristic of any infectious disease is its basic reproduction number, R_0 , which measures the ability of a disease to propagate in a population. Specifically, R_0 is the number of new cases generated, on average, from a single case, throughout the period of infection. If $R_0 < 1$, then, on average, the disease does not generate enough subsequent cases to replace those individuals that either recover or die from the infection, and so the prevalence of the disease declines over time. By contrast, if $R_0 > 1$, then the disease will spread through the population. Where $R_0 = 1$, the prevalence should remain constant.

As I have already discussed, mosquitoes act as vectors for a number of diseases that

affect humans. Whilst I focus on malaria here, it is important to note that many of the epidemiological elements for malaria are common to all of these diseases: an uninfected mosquito can become infected itself by biting an infected host - be they human or animal, although in some cases it may receive the disease from its parents (yellow fever, for example, can be vertically transmitted); in all human-affecting mosquito borne diseases, a mosquito transmits the disease by biting an uninfected host, when the pathogen passes from the mosquito into the human. The centrality of the mosquito to both of these processes - its initial infection and the transmission of infection to a subsequent host - underscores why knowledge of the mosquito's ecology and behaviour are so crucial for understanding the epidemiology of these diseases.

The epidemiological details of malaria determined by Ronald Ross were put on a mathematical footing by Macdonald, 1957, who devised a mathematical model of the disease. In doing so, the authors made a number of assumptions, which have come to be known as the “classic assumptions” (Smith and McKenzie, 2004), which are useful idealisations that allow development of a simple model, which nonetheless captures many of the key aspects of the epidemiology of malaria. In particular, it is assumed that the mosquito population is homogeneous and of constant size, that mosquitoes do not experience senescence, that is, age-related death, that mosquitoes bite humans uniformly and at random, and that an infected mosquito never becomes uninfected. The model of Macdonald, 1957 results in an expression for R_0 given by the following,

$$R_0 = \frac{ma^2bce^{-gn}}{rg}, \quad (1.1)$$

where m is the density of mosquitoes per human, a is called the “human feeding rate” and represents the expected number of bites a mosquito will make on humans per day, b is the probability that an uninfected human becomes infected after being bitten by an infectious

1010 mosquito, c is the probability that an uninfected mosquito becomes infected as the result
 1011 of a bite on an infected human, g is the “force of mortality” and represents the per capita
 1012 death rate of a mosquito, n is the duration of the extrinsic incubation period (EIP) - the
 1013 period of time it takes the ingested *Plasmodium* parasite to transmute and migrate from
 1014 the mosquito’s gut to its salivary glands, where it can be passed onto another host when
 1015 the mosquito bites them - and r is the rate of recovery of an infected human from the
 1016 disease. In what follows, I discuss the sensitivity of malaria transmission to each of the
 1017 mosquito-related elements of expression (1.1).

1018 It may appear surprising that R_0 is least sensitive to the number or density of mosquitoes
 1019 in a region, since m enters linearly into expression (1.1). This does not mean that this aspect
 1020 is unimportant for malaria transmission. In many parts of the world, there is significant
 1021 seasonality in malaria cases, driven by changes in the mosquito vector density. However,
 1022 expression (1.1) does indicate that substantial changes in the density of mosquitoes are
 1023 required to significantly affect transmission, and suggests that vector control methods that
 1024 also affect other parts of disease transmission may be advantageous.

1025 Malaria is vectored by mosquitoes of the genus *Anopheles*, and in Sub-Saharan Africa,
 1026 where the majority of deaths occur, the principal vector is the *A. gambiae* mosquito.
 1027 The endemicity of malaria in large parts of Sub-Saharan Africa is likely due mainly to
 1028 the behavioural characteristics of this species of mosquito (Carter and Mendis, 2002): it
 1029 favours biting of humans, and the highly domestic habits of *A. gambiae* mean that there
 1030 is a great deal of contact between them and humans (Coluzzi, 1999). Outside Africa,
 1031 the anopheline mosquito vectors are zoophilic rather than anthropophilic (Bruce-Chwatt,
 1032 Garrett-Jones, and Weitz, 1966), meaning that they prefer to feed on animals rather than
 1033 humans. These vectors of malaria typically have an anthropophilic index (the proportion of
 1034 blood meals taken on humans) of around 50%, but often much lower, between 10% to 20%

(Bruce-Chwatt, Garrett-Jones, and Weitz, 1966). By contrast, *A. gambiae* and *A. funestus* - the two predominant vectors of malaria on the African continent - have anthropophilic indices around 80%-100% and 60%-100%, respectively (Bruce-Chwatt, Garrett-Jones, and Weitz, 1966). The prejudices of a mosquito's specific taste for blood matter, because for malaria to be transmitted, a mosquito must first bite an infected human, then later bite another uninfected person. Because two bites are required for disease transmission to occur, transmission has a strong sensitivity to the specificity of mosquitoes towards biting humans. In particular, the effect size depends on the square of the proportion of meals taken on humans - a^2 in expression (1.1). Consider a mosquito vector with an anthropophilic index of 10%, for example, *A. sinensis* (a vector of malaria in the far east; Bruce-Chwatt, Garrett-Jones, and Weitz, 1966). For *A. sinensis* to generate the same level of disease transmission as *A. gambiae*, its population density would need to be 100 times as great!

Why do some mosquito species in Africa (particularly, certain members of the *A. gambiae s.l.* complex) tend to prefer an all-you-can-eat human blood buffet, whereas those outside prefer to feed mainly on the blood of animals? It is possible that this is due to differences in the animals that were available for domestication between Africa and the rest of the world. In Sub-Saharan Africa, there were no animal species that could be readily domesticated, whereas in Europe, Asia, and the Middle East, many animals, including cows, pigs, sheep and goats were domesticated during the Neolithic agrarian revolution around 4,000 to 5,000 years ago (Diamond, 1998). This meant that, whilst in Europe, population densities of animals of a few select species grew enormously, in Sub-Saharan Africa, the animals were more diverse and at lower population densities. This may be why certain anopheline species in Sub-Saharan Africa adapted to become such efficient human blood feeders, since we were the animal species that was available in the highest population densities. In other parts of the world, the mosquito species may have adapted to feed on the

domesticated animals which were available in sufficient numbers and, perhaps, easier prey (Carter and Mendis, 2002). An alternative explanation is that, since humans evolved out of Africa, they have been exposed to the mosquito species there for longer than anywhere else, allowing certain members of the *A. gambiae* complex more opportunity to evolve to exploit humans.

Whilst the a^2 term appears simple, it hides subtle complexities, which are crucial for developing effective interventions against malaria. Because of the sensitivity of R_0 to the rate of human biting, many vector control methods explicitly target reductions in this aspect of disease transmission. All anopheline vector species feed predominantly indoors at night (Gatton et al., 2013), which partly led to the widespread use of indoor residual spraying of insecticides (IRS), as well as the massive increase in the use of long lasting insecticidal nets (LLINs), since the turn of the twenty-first century. Both of these interventions have been credited with causing the first sustained decline in malaria incidence in a generation (Bhatt et al., 2015), however, a consequence of spraying insecticides indoors is that this gives a selective advantage to those mosquitoes that feed on humans earlier in the day, when they are outdoors. Indeed, exophily (the preference of mosquitoes to rest outdoors after they have fed) is believed to be a key reason why indoor residual spraying failed to control malaria in the Garki project in northern Nigeria (Gatton et al., 2013; Molineaux, Gramiccia, and World Health Organisation, 1980). The sensitivity of R_0 to the human biting rate also highlights the importance of understanding the community ecology of vector species (Godfray, 2013). In parts of Africa, the widespread use of LLINs, which provide a barrier between a sleeping person and mosquito, has lead to a shift in vector dominance from the endophilic *A. gambiae sensu stricto* to the more exophilic *A. arabiensis* (Gatton et al., 2013). To better plan effective interventions, it is essential to know the local distribution of species present, and recognise if it is likely that vector

substitution may occur as a result of targeting one of these species.

Disease transmission, as measured by R_0 , is most sensitive to mosquito mortality, whose effect is quantified by the $\frac{e^{-gn}}{g}$ term in expression (1.1). Why is disease transmission so sensitive to this aspect of mosquito ecology? Because a mosquito that lives longer, has more opportunity to bite an infected human and become infected itself, then to survive the extrinsic incubation period (n ; EIP) to allow it to become infectious, and finally has enough time to find and bite an uninfected host. Also, more indirectly, mosquito lifespan influences disease transmission because mosquitoes that live longer produce more offspring and, hence, support larger populations of mosquitoes. The sensitivity of R_0 to mosquito mortality means that it is crucial to determine how long they live in the wild.

The predominant approaches to estimating how long mosquitoes live in the wild are MRR experiments and female mosquito dissection-based approaches. In MRRs, lab-reared or wild-caught mosquitoes are marked, typically with a fluorescent dye, and released into the wild. Over days subsequent to the release, collections are made of mosquitoes and the number of marked mosquitoes recaptured is recorded. In populations of mosquitoes with shorter lifespans, the number recaptured will tail off more quickly, because fewer marked mosquitoes persist in the population. Typically, these experiments do not record information on where the mosquitoes were recaptured, and so the rate at which mosquitoes disperse out of the study area also affects the number of mosquitoes available for recapture. This means that estimates of lifespan determined by this method are more accurately interpreted as “time spent in the study area”. In Chapter 2, I provide a more in depth appraisal of this approach to estimating mosquito lifespan and so do not discuss this further now. In dissection-based studies, wild-caught female mosquitoes are dissected in an attempt to estimate their age. This is possible because it is thought that female mosquitoes undergo predictable changes in their reproductive anatomy throughout their lifespan, which can be

1110 used to estimate the age of a given specimen (Polovodova, 1949; Detinova, 1962). In
1111 Chapter 3, I describe this approach in more detail and so do not discuss it further now.

1112 The most predominant and impactful interventions - LLINs and indoor residual spray-
1113 ing (IRS) - are thought to be so successful because they each simultaneously target multiple
1114 elements of the disease transmission process. Most importantly, it is hoped that the toxic
1115 effects of these chemicals kill mosquitoes and, hence, shorten the mean lifespans of popula-
1116 tions. By shortening the lifespan of mosquitoes, this likely leads to a decline in the density
1117 of mosquitoes that are supported in a given location, again reducing disease transmission.
1118 Finally, LLINs (and probably, to a lesser degree, IRS) reduce the effective contact time
1119 between mosquitoes and humans, leading to lower biting rates.

1120 Whilst there is widespread consensus that LLINs and IRS have been effective means of
1121 vector control, there still exists considerable uncertainty in their mode of action (Gatton et
1122 al., 2013). The insecticides used in LLINs and IRS are thought to have threefold effects on
1123 vector populations: toxic chemical action, spatial deterancy, and contact irritancy (Smith
1124 and Webley, 1969; Lines, Myamba, and Curtis, 1987; Takken, 2002). Whilst there is
1125 considerable evidence in support of each of these factors, there is now widespread prevalence
1126 of mosquito vector resistance to the four main classes of insecticide which are widely used
1127 (Ranson et al., 2011; Ranson and Lissenden, 2016), which threatens to weaken the potency
1128 of these methods of vector control. Indeed, results from experimental hut trials demonstrate
1129 that mosquito mortality is considerably lower in areas with higher levels of insecticide
1130 resistance (N’Guessan et al., 2007). Aside from physiological resistance to the chemicals,
1131 there is reasonable evidence that there has been behavioural adaptation to the widespread
1132 use of bednets and insecticide spraying (see Gatton et al., 2013 for a summary of the
1133 current evidence). This growing resistance means that understanding the mechanism by
1134 which certain interventions are effective is essential to be able to design new interventions

and effectively deploy existing measures.

An element of expression (1.1) that I have not yet discussed is the EIP, to which disease transmission is almost as sensitive as mosquito lifespan. Rather than being a constant, the EIP for malaria has long been known to be temperature sensitive (Boyd, 1949; Macdonald, 1957; Detinova, 1962), and is reasonably well approximated by the downward-sloping “Detinova curve”, $EIP = \frac{111}{T-16}$ (Detinova, 1962; Patz and Olson, 2006). Since the EIP tends to decrease as temperature rises, this supports the argument made by some that the malaria burden may worsen as global temperatures increase (Githeko et al., 2000; Patz and Olson, 2006; Paaajmans, Read, and Thomas, 2009). The interaction between lifespan and EIP in expression (1.1) means that, unsurprisingly, any intervention that targets killing adult mosquitoes will necessarily lower the fraction of individuals who survive the EIP, and can significantly impact disease transmission.

A key assumption underlying the model from which expression (1.1) results, is that mosquito age does not affect the rate at which they die. Since work by Clements and Paterson, 1981 indicated that mosquito populations may experience the effects of old age some have questioned whether this is likely to be the case in reality, and has lead to attempts to detect mosquito senescence. In Chapters 2 and 3, I use meta-analyses of MRR experiments and dissection-based studies, respectively, to determine whether mosquitoes experience the effects of old age. To motivate this aim, I now summarise the current modelling approaches that have been used to assess the importance of mosquito senescence for malaria transmission, and also the current experimental evidence for this aspect of mosquito ecology.

1157 1.8 The importance of mosquito senescence

1158 In the Appendix to this thesis, I include a brief toy model of form similar to that presented
 1159 by Bellan, 2010. This modelling illustrates that age-dependent mortality can affect the
 1160 impact of vector control programmes, however, unsurprisingly, the importance of this facet
 1161 of mosquito ecology depends on the degree of the effect strength in wild populations. The
 1162 results that I present in Chapters 2 and 3 suggest that the magnitude of age-dependent
 1163 mortality faced by wild populations is relatively minor and, hence, I argue that the impact
 1164 of mosquito senescence is more likely to be marginal rather than revelatory. I also believe
 1165 that this assertion is supported by the results of other attempts to model this effect.

1166 Novoseltsev et al., 2012 model the impact of different mortality curves on vectorial
 1167 capacity and R_0 , finding that the impact of age-dependent mortality is most important
 1168 when vector densities and/or population growth rates are sufficiently low, so that $R_0 \sim 1$.
 1169 In this knife-edge case, the form of age-dependent mortality assumed can determine whether
 1170 or not enough mosquitoes survive past the EIP to allow a mosquito borne pathogen to
 1171 persist.

1172 Hancock, Thomas, and Godfray, 2009 model the use of fungal biopesticides, which
 1173 kill malaria-carrying mosquitoes before they are old enough to transmit the disease, using
 1174 an age-structured susceptible-exposed-infectious model of adult mosquito lifetime, with a
 1175 constant rate of larval recruitment. They investigate how different pesticides, each of which
 1176 inflicts a different age-related mortality risk (modelled using a Weibull distribution) on a
 1177 mosquito population, affect the transmission of disease. This modelling suggests that the
 1178 shape of the curve of mosquito mortality is an important factor for determining the impact
 1179 of spraying on disease transmission. In their simulations, disease transmission was most
 1180 sensitive to the shape of the additional mortality hazard when considering highly effective
 1181 pesticides. The age-dependent action of these pesticides is, however, likely more extreme

than the natural mortality faced by mosquitoes (those populations in the absence of vector control), meaning that it is possible that the effect of senescence on disease transmission for most populations is more subtle.

A range of physiological mechanisms have been suggested may lead to age-related death, including changes in mosquito flight (Nayar and Sauerman, 1973), immune functioning (Christensen, LaFond, and Christensen, 1986), and detoxification mechanisms (Hazelton and Lang, 1984; Lines and Nassor, 1991). To my knowledge, however, experiments have not been done that demonstrate which, if any, of these factors influence wild populations.

A number of studies have also found evidence for mosquito senescence, in both lab populations (Styer et al., 2007; Dawes et al., 2009), and also for wild mosquitoes (Clements and Paterson, 1981; Harrington et al., 2008; Hugo et al., 2014). In large cage experiments consisting of over 100,000 *Ae. aegypti* mosquitoes, Styer et al., 2007 found that the mortality rate increased with age across both mosquito sexes, where females were found to live about twice as long as males (females had a mean lifespan of c.32 days whereas for males it was c.16 days). Similarly, Dawes et al., 2009 conducted experiments on 1,000 *A. stephensi* (an important vector of malaria in India) and concluded that mortality depended on age. In both cases, the mortality rate was low and relatively constant, however, for mosquitoes younger than 10 days old. Since estimates of wild mosquito lifespan are often less than 10 days (see Chapters 2 and 3), however, it is less likely that senescence will be as important a factor for wild populations since most mosquitoes will die due to extrinsic causes before intrinsic mortality affects them. Additionally, Dawes et al., 2009 found evidence that infection with *P. berghei* reduced the lifespan of mosquitoes. Whilst this result may hold for wild populations, it is unlikely given the millions of years that the parasite has probably had to adapt to minimise its impact on the mosquito host (Carter and Mendis, 2002) and, instead, may reflect the inevitable inbred nature of laboratory colonies. More generally,

1207 this suggests that age-related mortality may be a more significant factor in laboratory
1208 colonies than it is for wild mosquitoes.

1209 In a meta-analysis of MRR experiments and mosquito dissection studies conducted
1210 on eleven different species of mosquito, Clements and Paterson, 1981 compared the fit of
1211 an exponential distribution for lifespan with that of a Gompertz model. In a majority
1212 of experiments (mainly corresponding to dissection based studies), the Gompertz model -
1213 representing an increase in mortality with age - provided a better fit to the data. Harring-
1214 ton et al., 2008 performed MRR experiments on cohorts of laboratory-reared *Ae. aegypti*
1215 mosquitoes, consisting of ages between 0 and 20 days, and found that the number of
1216 recaptured individuals decreased with the age of the cohort; suggesting an elevated mor-
1217 tality for older mosquitoes. Because mosquitoes in the laboratory probably live artificially
1218 long, however, it is possible that the older cohorts were already experiencing the effects
1219 of age-related decay before release and, therefore, these results may not generalise to wild
1220 populations. Hugo et al., 2014 used a novel technique of mosquito ageing, based on tran-
1221 scriptional profiling of individual mosquito specimens, to estimate *Ae. aegypti* mosquito
1222 mortality across different times of the year in Vietnam. In the wetter and cooler of the
1223 seasons, mosquitoes were estimated to live shorter lives, and the exponential model - which
1224 assumes an age-independent mortality risk - was a better fit to the data. In the dry and
1225 hot season, the mosquitoes lived longer, and the Gompertz distribution provided a better
1226 fit to the data. This mixed result suggests that mosquito mortality may be more complex
1227 than is currently believed, and outside factors, such as the weather and the availability of
1228 hosts, may affect the pattern of mortality.

1229 Overall, I argue that on the basis of previously-published work, it is unclear whether age-
1230 related mosquito mortality matters for wild populations. This, in part, probably reflects
1231 the subtlety of the effect, and the cost of carrying out experiments with the large number

of mosquitoes needed to detect it. It also motivates the approach I take in Chapters 2 and 3, where I pool data from a number of studies to estimate mosquito mortality.

1.9 Future vector control by genetic drive

So-called ‘selfish genes’ (Dawkins, 1976) can spread through a population even when they confer a fitness cost to the individual. The eminent evolutionary biologist Bill Hamilton (Hamilton, 1967) recognised the significance of these types of genes for controlling pests. In a landmark paper, Hamilton argued that a gene on the Y chromosome in a heterogametic male that causes this chromosome to beat the X chromosome in a race to fertilise the egg, would have a selective advantage. As the mutant gene spreads, it would drive sex ratios away from the 1:1 equilibrium argued for by Fisher, 1930, resulting in smaller and smaller populations that are increasingly male-biased. Eventually, stochastic loss of the few remaining females would then lead to eradication of the local population. The idea of using gene drive for vector control (Hickey and Craig Jr, 1966; Hamilton, 1967) and pest management (Curtis, 1968) was first proposed in the 1960s, yet serious attempts to introduce gene drive systems into organisms did not occur until the twenty first century.

Homing endonuclease genes (HEGs) code for nonessential genes which occur in a variety of unicellular organisms, including yeast, bacteria, fungi and viruses (Mueller et al., 1993; Belfort and Roberts, 1997). HEGs code for a protein that targets a c.24 base pair sequence in the DNA and causes a double-strand cut there (Stoddard, 2005). For heterozygous individuals, the cell typically repairs the break using the chromosome containing the HEG as a template and, hence, the HEG is copied over to the other chromosome, in a process known as *homing*, converting a heterozygote into a homozygote (Stoddard, 2005). This trick allows HEGs to spread from rare to fixation in populations. Given the selective advantage that these genetic elements possess, it appears surprising that they do not occur

1256 in larger multicellular organisms. This is because, after reaching fixation, the absence of
1257 further selection means that these elements can accumulate mutations and become non-
1258 functional (Godfray, North, and Burt, 2017). Indeed, Goddard and Burt, 1999 found
1259 evidence in yeast that horizontal transmissions of HEGs between species are required for
1260 long-term persistence of these genes.

1261 Burt, 2003 was the first to propose that HEGs could be used to control mosquito
1262 vector populations. Burt imagined a number of different ways a HEG could be used to
1263 impose a fitness load on either the pathogen population and/or the mosquito population.
1264 First, a HEG could be engineered to recognise a site which is essential for the pathogen
1265 yet non-essential for the mosquito. The gene would spread to fixation and, in doing so,
1266 replace the target gene, meaning that the pathogen is unable to be transmitted by the
1267 mosquito. Second, a HEG could target a gene that is essential for the fitness of the vector,
1268 for example, a gene associated with female fecundity or survival. If the HEG spread to
1269 fixation, it would impose a large a genetic load on the mosquito population. This would
1270 likely lead to population suppression, potentially to a level where the pathogen could be
1271 lost from circulation. Third, Burt imagined that a HEG could be inserted onto the Y
1272 chromosome but recognise a site on the X chromosome, under a promoter that caused it
1273 to be expressed only at male meiosis. When the gene was expressed, it would cause the X
1274 chromosome to be cut, and attempts to repair it would result in a non-viable X gametes,
1275 meaning all of the offspring would be sons (Godfray, 2013).

1276 The recent discovery of the Crispr-Cas9 system in bacteria (Jinek et al., 2012; Doudna
1277 and Charpentier, 2014), and the idea that it might be used to introduce artificial homing
1278 mechanisms into wild populations (Esvelt et al., 2014; Simoni et al., 2014) has provided
1279 further stimulus for research into the use of driving endonuclease gene systems (DEGs)
1280 - a term proposed by Godfray, North, and Burt, 2017 for the collection of constructs

which have characteristics similar to HEGs - to control vector populations. Mathematical modelling of DEGs has played a crucial role in shaping the research into this area, and the field is expansive (see Godfray, North, and Burt, 2017 for a summary of the work done, thus far). Rather than attempt to summarise all models that have been proposed, I discuss a few indicative cases.

Many current proposals for DEGs aim to cause harm to the mosquito vector population. Whether a driving endonuclease gene (DEG) will spread is a function of the fitness cost (s) it imposes on a homozygous individual (for this analysis, I assume that there is no cost for carrying a DEG on a single chromosome), and the rate at which homing occurs (e), in other words, the rate at which a heterozygote is converted into a homozygote. Following the derivation in Godfray, North, and Burt, 2017, I use a non-spatial model to determine the equilibrium frequency reached by the DEG q , where equal numbers of copies produced for randomly chosen wild-type and DEG alleles should be equal. A randomly chosen wild-type allele will be present in a wild-type homozygote with probability $(1 - q)$ and be transmitted to half the offspring, and will be present in a heterozygote with probability q , where the probability it is transmitted to its offspring is $\frac{1}{2}(1 - e)$. A randomly chosen DEG allele will be present in a DEG homozygote with probability q , where it will be present in half of the offspring that survive $(1 - s)$, and in a heterozygote with probability $(1 - q)$, where it be present in $\frac{1}{2}(1 + e)$ of the offspring. Equating the numbers of offspring for wild-type and DEGs, we obtain,

$$\frac{1}{2}(1 - q) + \frac{1}{2}q(1 - e) = \frac{1}{2}q(1 - s) + \frac{1}{2}(1 - q)(1 + e), \quad (1.2)$$

which, when rearranged yields,

$$-qe = (1 - q)e - qs, \quad (1.3)$$

which can be interpreted intuitively as the net effect of a DEG on a wild-type allele (left-hand side) must equal its net effect on another DEG allele (right-hand side). I now solve for the equilibrium frequency of a DEG using expression (1.3), to obtain,

$$q = \frac{e}{s}. \quad (1.4)$$

If $e < s$, a stable polymorphism with both alleles present will exist in the population. If $e > s$, meaning that the rate at which homing occurs exceeds the fitness cost of a homozygous DEG, then $q > 1$ (which is impossible), indicating that the DEG will spread to fixation and the wild-type will disappear.

The load that a DEG imposes on a population at equilibrium is straightforward to calculate. At equilibrium, the frequency of homozygous DEGs equals q^2 , meaning that the relative population fitness decreases from 1 to $1 - q^2s$. Using expression (1.4), the relative fitness of a population at equilibrium hence equals $1 - e^2/s$. As $e \rightarrow 1$, then DEGs targeting genes essential for survival ($s = 1$) would spread and, therefore, the overall fitness decreases to 0, and population elimination is likely.

Because the action of a DEG interacts with other processes that regulate the population, it is difficult to determine exactly what its impact will be on population size. Of particular importance is to consider how a DEG might interact with processes which are density dependent (Alphey and Bonsall, 2014; Godfray, North, and Burt, 2017), because these are the processes that typically regulate the population size. In the mosquito vectors responsible for the bulk burden of disease, it is generally believed that density dependence acts only in the larval stage (Gimnig et al., 2002; Muriu et al., 2013). This means that it matters when during their life the fitness cost resulting from the DEG's action is paid. If the DEG acts during the egg stage, for example, this could lower competition for resources amongst larvae, which might result in similar numbers of individuals surviving the

larval stage, with little change to the adult population. Ideally, for mosquitoes, a DEG would, hence, act at the pupal stage or on newly emergent adults, to avoid the interaction with density dependence (Godfray, North, and Burt, 2017). It has also been shown that the precise shape of the function dictating density dependent mortality can affect population dynamics (May, 2001). For example, were density dependence to act in an over-compensating way - where low population densities in one generation lead to much larger densities in the next, and *vice versa* - then this would certainly be of concern to a DEG release (Alphey and Bonsall, 2014). However, Godfray, North, and Burt, 2017 argue that this sort of population regulation probably does not occur in mosquito populations and, hence, is not likely to influence population dynamics of an introduced DEG.

The DEG that I have considered, thus far, is assumed to knock out a gene that affects the fitness of an individual. As I have already discussed, however, other varieties of DEGs are also conceivable. Deredec, Burt, and Godfray, 2008 considered a HEG inserted on the Y chromosome which causes cuts in the X chromosome, rendering it non-viable, sometimes called an *X-shredding* or *Y drive* construct. Using a non-spatial population genetic model, they demonstrate that the equilibrium sex ratio (the proportion of males) is given by $1/[1 + (1 - e)^k]$, where e is the rate of homing and k is the number of HEGs (each with a unique target site on the X chromosome) inserted onto the Y chromosome. As the rate of homing is increased, therefore, the sex ratio increases towards 1, meaning a fully male population is obtained. The more independent HEGs that are inserted onto the Y chromosome, the greater the effective rate at which homing occurs, meaning that a male-biased population can be achieved at lower homing rates for a single construct.

As I have already discussed, the malaria eradication campaign initiated in 1955 largely overlooked the difficulty of eradicating malaria in sub-Saharan Africa, where access can be problematic and the public health infrastructure is generally poor. Today, the condition of a

number of countries in the region (for example, the Democratic Republic of Congo, Burkina Faso and Somalia) means that the gains from the recent declines in disease transmission have been distributed unequally throughout the continent (Bhatt et al., 2015). A benefit of a gene drive campaign is the ability of a DEG to spread across a region (being carried by the insects, themselves), meaning that it likely needs less on the ground support than some other interventions. The spatial spread of DEGs has been considered in a number of mathematical models, one of which I extend in Chapter 5. In the same Chapter, I discuss implications of spatial models describing the population dynamics of a Y drive HEG and, hence, do not discuss this aspect further now.

Until now, all the vector control measures that have been widely deployed against mosquitoes and the malarial pathogens have suffered as a result of evolved resistance. It is thus pertinent to consider how resistance might arise in response to the release of a DEG, and how this may influence the resultant spread of disease. The most likely form of resistance allele to arise is one that does not contain the DEG recognition site (Godfray, North, and Burt, 2017). Whilst considerable efforts are under way to locate recognition sites that are highly conserved, it is possible that a few wild-type forms not containing the recognition site may naturally exist. Alternatively, the homing action itself could cause a mutation in the recognition site (Godfray, North, and Burt, 2017). Homing typically results when a double-strand break in the DNA is repaired using the homologous chromosome as a template. There are, however, non-homologous ways in which the DNA can be repaired. One particular mechanism is called *end joining* (EJ) and involves the ligation of the loose chromosome ends (Weterings and Chen, 2008). If the DEG recognition site is repaired during this process, EJ simply results in a reduction of homing efficiency. However, if EJ destroys or significantly mutates the recognition site, then a mutant allele will arise which confers resistance to the homing. If the mutant allele has a fitness similar to the one

with the recognition site, then modelling suggests it will spread throughout the population (Deredec, Burt, and Godfray, 2008). Perhaps the most straightforward way to mitigate against this risk is to engineer a DEG to cut at a number of sites in the target gene (Alphey, 2016; Godfray, North, and Burt, 2017), or to combine multiple HEGs in a single intervention (Godfray, North, and Burt, 2017).

It is now an exciting time to work in vector biology since a number of laboratories have successfully introduced DEGs into mosquitoes (Windbichler et al., 2011; Gantz et al., 2015; Hammond et al., 2016), and work on the technology is continuing at a fevered pace. There is also a research consortium called Target Malaria (of which I am part), whose aim is to consider these sorts of technology for use in malaria control and, when the crucial approval from stakeholders, regulators and fellow scientists is obtained, hopefully to use it to help control the disease.

1.10 Motivating this thesis

The impact of mosquito-borne disease on historical cultures can be read about in books ranging from literature to archaeology. The current impact of mosquito-borne disease on some developing countries can, however, be seen first hand. History teaches us that a single strategy for controlling, and later, eradicating these diseases, will likely fail, and that multi-pronged approaches are necessary. Vector control has been a critical factor in those successful past attempts to eradicate vector borne disease from a given location, as well as the recent declining incidence of malaria. Yet, mosquito resistance to insecticides, climate change, and donor fatigue, amongst other factors, threatens to weaken the potency of these methods. In this context, it becomes even more crucial to better understand mosquito ecology, in order to assess existing interventions, and develop new policies which effectively eliminate disease. A crucial element of mosquito ecology, to which the majority

1399 of this thesis is devoted, is the length of time that mosquitoes live. In Chapters 2 and 3, I
1400 use data from the two prominent approaches to estimate this quantity - MRR studies and
1401 dissection-based experiments - to derive estimates of mosquito longevity which reflect the
1402 weight of data that has been collected. In Chapter 4, I analyse the results of experiments
1403 from a new technology, known as *near-infrared spectroscopy*, which can age individual
1404 mosquitoes, and show that it may be worth pursuing as a tool to estimate mosquito lifespan
1405 in the wild.

1406 The worrying recent parasite resistance to insecticide, which has been recorded in Asia,
1407 means that now, more than ever, it is essential for us to develop new methods for disease
1408 control. In Chapter 5, I model a recently-proposed method for vector control, using a HEG
1409 which biases the sex ratio towards males. If such a gene drive mechanism were released into
1410 the wild, it will be important to monitor its effects on the mosquito population, particularly
1411 to search for any signs of resistance. It is conceivable that the three types of method which
1412 I have mentioned, MRRs, mosquito dissection, and near-infrared spectroscopy, could be
1413 used for this purpose.

Chapter 2

A meta-analysis of mark-release-recapture experiments of mosquitoes

2.1 Abstract

Mosquito-borne disease transmission is sensitive to the duration of adult mosquito lifespan. An important method to estimate lifespan in the field is mark-release-recapture (MRR) experiments. These experiments, however, are costly to undertake and can resort in estimates associated with considerable uncertainty. Also, there is considerable heterogeneity in the methods used to analyse such experiments, meaning that it is not straightforward to compare estimates. In this chapter, I apply a common statistical methodology to the analysis of a dataset of 232 historical MRR experiments, which allows direct comparison between the estimates. Further, I use a Bayesian hierarchical model to produce estimates of lifespan at the species and genus levels and, also, estimates of lifespan by sex. This

1428 analysis indicated that there were significant differences in the lifespan of mosquitoes be-
1429 tween species and genera. The analysis indicates, also, that female mosquitoes live on
1430 average for 0.8 days longer than males. By fitting a range of different survival models, each
1431 of which embodies different assumptions about age-dependent mortality, I test whether
1432 mosquitoes experience senescence. In all cases, the exponential survival model, which as-
1433 sumes a mortality risk which is independent of age, fit the data as well as or better than
1434 each of the five other models which incorporated age-dependent mortality, suggesting that
1435 wild populations of mosquitoes do not undergo senescence.

1436 2.2 Introduction

1437 Of particular importance to the epidemiology of mosquito-borne disease is the length of
1438 time that adult mosquitoes live in the wild, which has a fourfold effect on disease transmis-
1439 sion. All else being equal, mosquitoes with longer lifespans (i) form larger populations, (ii)
1440 are more likely to become infected, (iii) are more likely to survive the extrinsic incubation
1441 period (EIP) of the disease to become infectious and (iv), after becoming infectious, are
1442 more likely to bite and infect a human Ross, 1910; Macdonald, 1957; Anderson and May,
1443 1992; Smith and McKenzie, 2004. The sensitivity of disease transmission to these factors
1444 means that information on mosquito longevity is crucial to assess disease control strategies.

1445 There are currently two methods that have been used to determine mosquito lifespan in
1446 the wild: MRR studies and dissection-based approaches that estimate the age of captured
1447 samples using a knowledge of the physiological effects of ageing on mosquitoes. The first
1448 MRR experiments were used to study the flying distance of anopheline mosquitoes in the
1449 Panama during construction of the canal (Zetek, 1915) and in South Carolina, USA (Le
1450 Prince and Griffitts, 1917) but, subsequently, these experiments have been employed to
1451 determine lifespan (see Guerra et al., 2014 for a summary of MRR experiments). In MRR

experiments mosquitoes are marked (typically with a fluorescent die) and released into the wild, and subsequent captures of wild mosquitoes are carried out over a number of days. Over time the released mosquitoes die and thus the number of marked mosquitoes that are recaptured declines. The faster the rate of decline the higher the mortality hazard, and the shorter the lifespan. Hence by monitoring the decline in recaptures over time the lifespan of the marked mosquitoes can be estimated by statistical methods. Dissection-based methods have mostly been only applicable to female mosquitoes, and relies on the assumption that irreversible changes occur to female reproduction organs throughout their lifetime (Polovodova, 1949; Detinova, 1962). These methods either examine the proportion of mosquitoes that have laid eggs, or estimate the number of gonotrophic cycles a female mosquito has experienced by counting the number of beadlike dilations in the ovarioles (Polovodova, 1949; Gillies and Wilkes, 1965; Samarawickrema, 1968).

Dissection-based methods are used less frequently than MRR experiments to determine mosquito lifespan. We speculate on the reasons for the lower popularity of this method: MRR experiments can be used to estimate other characteristics of wild mosquito populations, such as mosquito dispersal and population sizes; dissecting individual mosquitoes requires specialised knowledge and equipment; and dissection-based approaches require the experimenter to make assumptions about the effect of the gonotrophic cycle on female mosquito physiology. MRR experiments require fewer assumptions to be made about mosquito biology to arrive at estimates of lifespan (Reiter, 2007): behaviour and survival are not affected by marking or handling mosquitoes; the probability of recapture does not change with age; and the rate of mosquito dispersal is small compared to the risk of mortality. The assumption that MRR estimates of lifespan are unaffected by mosquito dispersal is likely violated in practice, and we believe strengthens the need for other methods (such as dissection-based approaches) that do not depend on mosquito dispersal.

1477 Unfortunately, individual MRR studies are costly and can result in poor quality esti-
1478 mates of mosquito lifespan, due to the difficulty of capturing mosquitoes in natural environ-
1479 ments. Furthermore, the variety of experimental and statistical methods used across dif-
1480 ferent studies makes it difficult to reach a consensus about the lifespan of wild mosquitoes.

1481 In this chapter, we overcome some of these issues by combining data from 232 separate
1482 MRR time-series covering 33 separate species, and three genera, from a recently-published
1483 database (Guerra et al., 2014). We apply a hierarchical Bayesian model to these data,
1484 allowing us to estimate average lifespans under a range of assumptions about how the
1485 individual time-series are related. In particular, we estimate lifespan at both the species
1486 and genus levels. We also examine the effect of exogenous factors such as air temperature,
1487 mosquito sex, and blood-feeding.

1488 In addition to our investigations of average mosquito lifespan, we further ask if there is
1489 evidence that mosquito mortality increases with age in wild populations. This is of interest
1490 because a mosquito population whose members are subject to senescence will contain a
1491 smaller fraction of very old mosquitoes, which are particularly effective in transmitting
1492 disease, than a similar population where senescence does not occur (Styer et al., 2007;
1493 Bellan, 2010; Novoseltsev et al., 2012). To address this question we use six different
1494 statistical models of mosquito mortality taken from mosquito studies. Each of these models
1495 embodies different assumptions about senescence and we fit the parameters within our
1496 Bayesian framework.

2.3 Materials and methods

2.3.1 Data

The MRR database in Guerra et al., 2014 contains 393 individual time-series, along with meta-data for a range of factors for each experiment (for example, species, geography, and date of study). We then transformed the data into a form amenable to our statistical model. In particular, we removed time-series with fewer than six separate recapture observations, as well as species with fewer than two separate MRR time-series. The final data encompassed 232 time-series (see Section 7.2 for a list of the studies included in our final dataset), with a range of characteristics (Tables 2.1 and 2.2). The data encompassed time-series from MRR experiments across 33 different species and three genera: *Aedes* (91 separate time-series), *Anopheles* (94), and *Culex* (47), and spanning a wide geographical range (Figure 2.1). The data comprise 177 female-only, 35 male-only and 18 mixed-sex release time-series. Where the age of the adult mosquitoes at release-time was available (102 time-series) we include this information in our time variable (that represents mosquito age) in the statistical models. Where age was unavailable (130 time-series), we assume that the age at release is zero. We recognise this assumption will likely underestimate mosquito age for some studies and so we interpret our estimates as lower bounds on mosquito lifespan.

2.3.2 Statistical model

Data for a typical MRR experiment consists of a single release of N_R marked mosquitoes followed by a series of marked mosquito recaptures on subsequent days (Figure 2.2 shows three such example series.) We model the number $y(t)$ of marked mosquitoes recaptured on day t using a negative binomial sampling model,

$$y(t) \sim \text{NB}((N_R - Y(t-1))S(t)\psi, \kappa), \quad (2.1)$$

Data characteristic	Total
Number of species	33
Number of genera	3
Number of <i>Anopheles</i>	94
Number of <i>Aedes</i>	91
Number of <i>Culex</i>	47
Number of female-only MRR	179
Number of male-only MRR	35
Number of mixed sex MRR	18
Number of pre-release blood-feeding only MRR	71
Number of pre-release sugar-feeding only MRR	41
Number of pre-release both blood- and sugar-feeding MRR	4
Number of pre-release neither blood- and sugar-feeding MRR	116
Number of MRR time-series	232

Table 2.1: Summary of variables across all MRR time-series.

Data characteristic	Min	Mean	Median	Max	Standard deviation
Study duration, days	6.0	11.8	10.0	71.0	4.7
Number of days on which collections took place	6.0	10.1	9.0	47	9.1
Number of separate release days	1.0	1.9	1.0	23	3.0
Number released	66	4,929	1,297	86,200	12,043
Number recaptured	2	163	63	4,090	399
Recapture percentage	0%	8.6%	5.2%	57.1%	10.04%
Age at release, days	0	1.8	0	13	2.8

Table 2.2: Summary of data from individual MRR time-series.

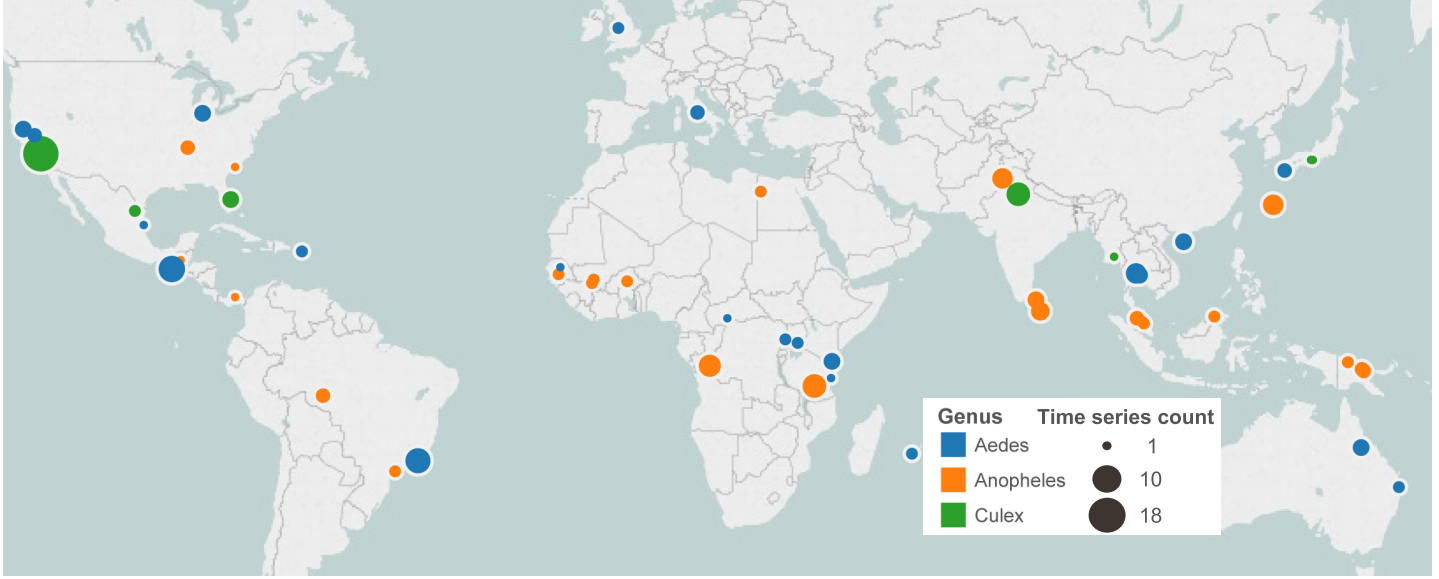


Figure 2.1: **The location and number of separate time-series for the MRR studies included in this analysis.** The area of each circle indicates the number of time-series at each study site. The colour shows the genus of the mosquitoes in each study.

where $Y(t-1)$ is the cumulative number of mosquitoes caught on all days before t , $S(t)$ is the probability that an individual mosquito survives and remains in the study area until time t , ψ is the daily recapture probability for an individual mosquito which is assumed constant through time, and κ is the time-independent shape parameter of the negative binomial distribution that controls the extent of variance in recapture rate likely due mostly to environmental heterogeneity. We chose a parameterisation of the negative binomial such that its mean is given by $\mu(t) = (N_R - Y(t-1))S(t)\psi$, and its variance by $\sigma(t)^2 = \mu(t) + \frac{\mu(t)^2}{\kappa}$. We use the negative binomial sampling distribution for its ability to account for temporal aggregation in captures due to environmental heterogeneity (Bliss and Fisher, 1953), and has previously been used to model mosquito MRR experiments (Service, 1971; Nedelman, 1983).

For some MRR experiments, there were a number of releases ($q \geq 2$) of marked

1531 mosquitoes throughout the duration of the study. In contrast to single releases, with
 1532 multiple releases over time, it is not in general possible to determine the particular release
 1533 to which a recaptured mosquito belongs. To avoid the complication of directly inferring
 1534 this quantity, we choose to represent previous recaptures probabilistically. This results in
 1535 a slightly different mean to that of the single release model,

$$\mu(t) = N_{Released}(1 - \psi)^{t-1}S(t)\psi, \quad (2.2)$$

1536 where the factor of $(1 - \psi)^{t-1}$ represents the probability that a mosquito is *not* recaptured
 1537 on a previous day. Where experiments consisted of two or more releases occurring at
 1538 distinct points in time we assumed that recaptures of individual mosquitoes from either
 1539 batch were independent of one another, although with the same sampling parameters (ψ
 1540 and κ). This results in an overall mean composed of the sum of those from all q releases,

$$\mu(t) = \mu_1(t) + \dots + \mu_q(t). \quad (2.3)$$

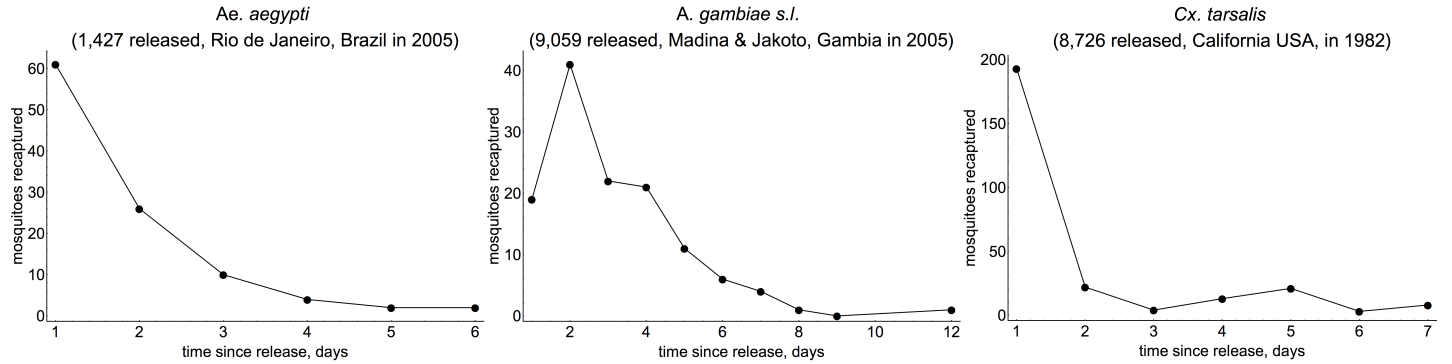


Figure 2.2: **Example MRR time-series for three different studies from the Guerra et al., 2014 database.** In all the above experiments there was a single release of a batch of marked mosquitoes.

1541 The simplest assumption for mosquito survival is that death and dispersal rates are

independent of mosquito age, giving an exponential ‘survival’ function,

$$S(t) = e^{-\lambda t}, \quad (2.4)$$

where λ is the sum of the rates of death and dispersal from the study area. To estimate the lifespan of mosquitoes we use this model because we find no evidence in support of age-dependence across all species (see below for the method used to determine this, and later the ‘Results’).

We assume that the number of mosquitoes recaptured on one day is independent of the number recaptured on any other day, apart from their joint dependence on λ , ψ and κ . This results in a likelihood of the data of the form,

$$\mathcal{L}(y(t_1), y(t_2), \dots, y(t_R) | \lambda, \psi, \kappa) = \prod_{i=1}^R p(y(t_i) | \psi, \lambda, \kappa), \quad (2.5)$$

where R is the number of individual days where collections took place, and $p(y(t_i) | \psi, \lambda, \kappa)$ is the probability of recapturing $y(t_i)$ mosquitoes on day t_i determined from a negative binomial distribution with parameters $\mu = e^{-\lambda t} \psi$, and κ .

2.3.3 Individual time-series estimates

We first treat each time-series separately, and estimate individual (λ, ψ, κ) parameters of the statistical model (described in the previous section) for each time-series. To use a Bayesian methodology we must specify priors on this set of parameters. Here we choose to specify independent priors on each parameter of the form: $\lambda \sim \mathcal{N}(-1.8, 1)$, $\psi \sim \exp(20)$ and $\kappa \sim \log\text{-normal}(1, 1)$. This prior on the rate parameter of the exponential distribution (λ) corresponds to a wide range of possible lifespans (Figure 2.3A), with a mean of 10 days. It was necessary to use an informative prior on ψ to allow it to be estimated; the prior on

1561 κ is fairly uninformative and allows a wide range of values (Figure 2.3B,C).

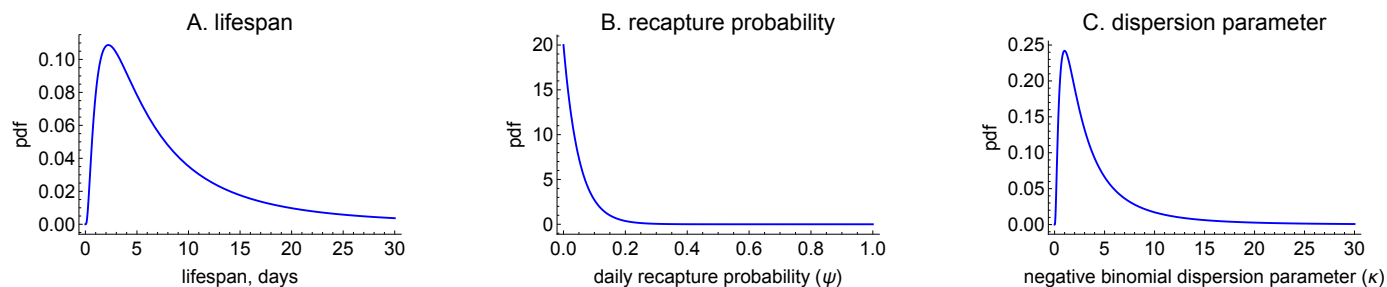


Figure 2.3: The prior probability distributions used for lifespan (A), recapture probability (B) and dispersal parameter (C) in the individual time-series analysis for the MRR studies.

1562 2.3.4 Estimating lifespan at the species, genus and overall groupings

1563 We want to synthesise information from across all MRR experiments to produce more
 1564 robust estimates of mosquito lifespan than can be obtained from considering the indi-
 1565 vidual time-series separately. However there exists considerable heterogeneity across the
 1566 experiments. This heterogeneity has two sources. There is that arising from variability in
 1567 experimental methodology. However there is also variability arising from actual differences
 1568 in lifespan across the different mosquito cohorts; for example due to genetic differences
 1569 between mosquito populations or due to climatic differences. Because of this heterogeneity
 1570 we argue that it is not appropriate to assume the parameters of our models are the same
 1571 across all time-series. We therefore choose to use a Bayesian hierarchical model which is
 1572 akin to a ‘random effects’ model in classical statistics. This type of model assumes that
 1573 there is random variation in parameters at the individual time-series level, although each
 1574 of the parameters is drawn from a common ‘population-level’ distribution. In our case we
 1575 separately estimate three different model groupings where the population-level distribu-
 1576 tions correspond to the species, genus and overall (across all studies) levels respectively

(Figure 2.4). This allows us to estimate average mosquito lifespan for a given species or genus, or alternatively across all time-series, by independent sampling from the posterior predictive distribution from which the individual λ_i are drawn.

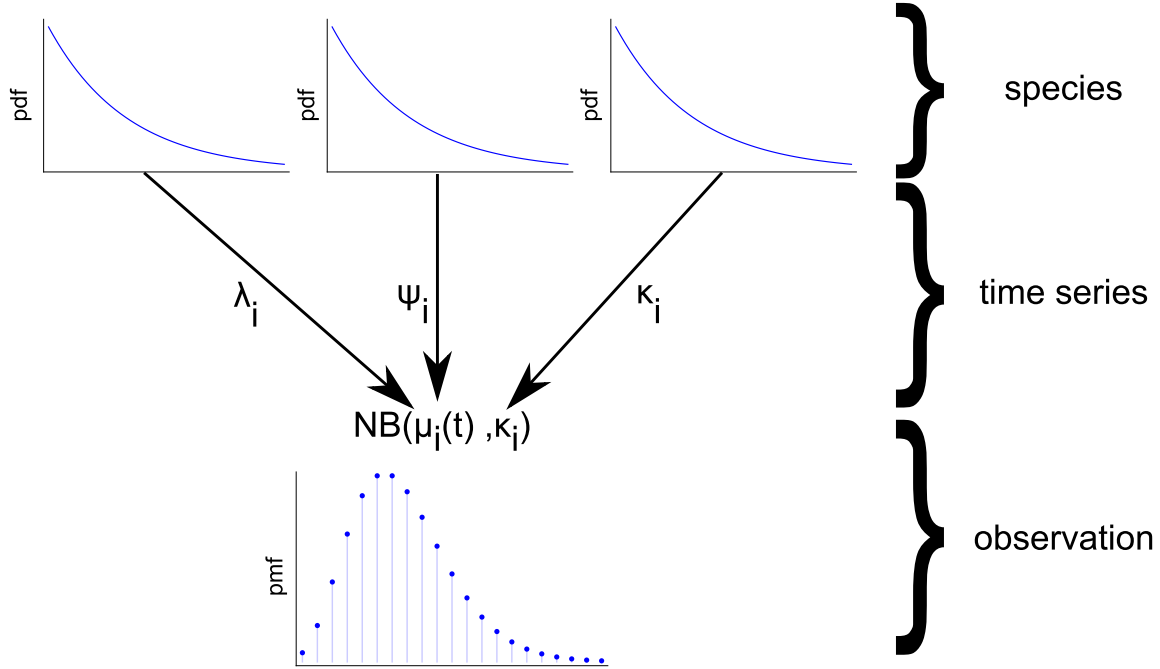


Figure 2.4: **A representation of the hierarchical Bayesian model, for the constant mortality model case when considering the species-level grouping.** Here ‘pdf’ indicates ‘probability density function’, and λ_i , ψ_i , and κ_i represent the force of mortality for the exponential distribution, daily recapture probability and over-dispersion parameter of the negative binomial distribution for time-series i in the dataset; $\mu_i(t) = \psi_i e^{-\lambda_i t}$ is the mean number of mosquitoes recaptured on day t .

Whilst we assume that the variation in mosquito mortality across experiments is random in nature, we recognise that there may exist systematic sources of variation that we do not include in our model. To account for this potential source of bias we examined how the individual estimates of lifespan correlated with experimental covariates (number of traps, range of traps, trapping method), but found no such evidence of systematic variation in

lifespan (data not shown).

A Bayesian framework requires that we incorporate our pre-analysis beliefs into our estimates through the use of priors for all parameters in a particular model. In our hierarchical model we are required to specify prior distributions at two levels of the analysis. The first of these links the individual time-series estimates with the overarching group-level distributions. The parameters of these group-level distributions are then set prior distributions (Figure 2.4). As an example, we suppose that the rate parameter λ_i of our exponential survival model in experiment i is drawn from a group-level distribution which we assume to be log-normal,

$$\lambda_i \sim \text{log-normal}(\mu_\lambda, \sigma) \quad (2.6)$$

where μ_λ and σ are the location and scale parameters of the log-normal distribution that summarise group-level characteristics. We then specify priors on these group-level parameters, $\mu_\lambda \sim N(-1.8, 1)$ and $\sigma \sim \text{log-normal}(-2, 1)$. The relatively complex nature of priors in hierarchical models make it important to determine their influence on inferences. We chose the above – somewhat uninformative – priors to allow a range of mosquito lifespans in order to minimise their effect on the estimates we report (Figure 2.5). We also chose to set hierarchical priors on the remaining parameters in our models – ψ the probability of daily recapture, and κ the ‘over-dispersion’ parameter of a negative binomial distribution. For ψ we chose a fairly informative prior that placed most weight where $\psi \leq 10\%$, since in most experiments the overall recapture percentage was well beneath this value. For κ we chose a fairly wide prior that had most of its probability mass below $\kappa = 20$ (Figure 2.6). In all cases the hierarchical priors were chosen to have similar implications on lifespan, recapture probability and over-dispersion at the individual time-series level as for the non-hierarchical analysis described previously.

2.3.5 Testing for age-dependent mortality in wild mosquitoes

Previous work has found evidence of age-dependent mortality in lab populations (Styer et al., 2007; Dawes et al., 2009), and less-commonly in wild mosquitoes (Clements and Paterson, 1981). Furthermore, recent modelling work has examined the implications of departures from a constant risk of mortality (Styer et al., 2007; Hancock, Thomas, and Godfray, 2009; Novoseltsev et al., 2012). To determine whether senescence occurs in wild mosquitoes we re-estimate our model using survival functions, $S(t)$, that allow for a rate of mortality that can vary with age. Specifically, we re-estimate our model using five other models that were previously used in the literature (see Table 2.3 for a description of these). Each of these models makes different assumptions about how the rate of death and dispersal varies with age, but all can be represented by a general form,

$$S(t) = e^{-\int_0^t \lambda(\tau) d\tau}. \quad (2.7)$$

where we constrain the parameters of our models to preclude the possibility of a hazard that decreases with age ($\frac{d\lambda(t)}{dt} \geq 0$). Whilst this could occur if older mosquitoes disperse less than younger mosquitoes, we do not believe this effect would outweigh any declines in survival associated with old age.

The parameters of the survival function of each model were assigned hierarchical priors that allowed considerable variation in mosquito lifespan (see Figure 2.5). Otherwise the statistical model we used remained the same as for the constant mortality case.

2.3.6 Model estimation by MCMC

The likelihood and priors we use result in posterior distributions whose analytic form cannot be calculated with existent computational methods. Instead we use Markov Chain Monte

Survival function	Hazard rate	Interpretation	Mean time spent in study area, $\bar{T}$	Papers assuming this function
Exponential	λ	Constant mortality risk	$\frac{1}{\lambda}$	Ross, 1910; Anderson and May, 1992; Smith and McKenzie, 2004
Gompertz	$\alpha e^{\beta t}$	Mortality increases with age at an ever-increasing rate	$\frac{e^{\frac{\alpha}{\beta}} \Gamma(0, \frac{\alpha}{\beta})}{\beta}$	Clements and Paterson, 1981; Styer et al., 2007; Novoseltsev et al., 2012
Weibull	$\alpha t^{\beta}; \beta \geq 1$	Mortality increases with age at an ever-increasing rate	$\alpha^{-\frac{1}{\beta}} \Gamma(1 + \frac{1}{\beta})$	Hancock, Thomas, and Godfray, 2009; and Carey, 2001 considering general insect demography
Gompertz-Makeham	$\alpha e^{\beta t} + c$	Two additive mortality risks: one that increases with age, and another that is age-independent	No simple analytic form	Styer et al., 2007
Logistic	$\frac{\alpha e^{\beta t}}{1 + \frac{\alpha s}{b}(e^{\beta t} - 1)}$	Mortality risk increases with age, although at a declining rate	No simple analytic form	Styer et al., 2007; Novoseltsev et al., 2012
Logistic-Makeham	$\frac{\alpha e^{\beta t}}{1 + \frac{\alpha s}{b}(e^{\beta t} - 1)} + c$	Two separate additive mortality risks: an age-dependent risk of the same form as the logistic model, and a constant hazard	No simple analytic form	Styer et al., 2007; Bellan, 2010

Table 2.3: **A description of the survival functions used in this study, arranged in rough order of model complexity (simple-complex from top-bottom.)** All parameters are defined to be non-negative. $\Gamma(\theta)$ and $\Gamma(\theta_1, \theta_2)$ refer to the Euler gamma function, and incomplete gamma function respectively. The ‘mean time spent in study area’ is an estimate of the combined effects of mosquito mortality and dispersal from the area of the study where collections take place, since our data does not provide spatially-resolved data.

Survival function	time-series-level priors	Group-level priors
Exponential	$\lambda \sim \text{log-normal}(\mu_\lambda, \sigma)$	$\mu_\lambda \sim N(-1.8, 1), \sigma \sim \text{log-normal}(-2, 1)$
Gompertz	$\alpha, \beta \sim \text{log-normal}(\mu_{\alpha \beta}, 0.2)$	$\mu_\alpha \sim N(-3, 1), \mu_\beta \sim N(-3, 0.5)$
Weibull	$\alpha, (\beta - 1) \sim \text{log-normal}(\mu_{\alpha \beta}, 0.2)$	$\mu_\alpha \sim N(-4.8, 1), \mu_\beta \sim N(-4, 1)$
Gompertz-Makeham	$\alpha, \beta, c \sim \text{log-normal}(\mu_{\alpha \beta c}, 0.2)$	$\mu_\alpha \sim N(-3, 1), \mu_\beta \sim N(-3, 0.4), \mu_c \sim N(-4.5, 0.5)$
Logistic	$\alpha, \beta, s \sim \text{log-normal}(\mu_{\alpha \beta s}, 0.2)$	$\mu_\alpha \sim N(-2.8, 1), \mu_\beta \sim N(-3, 1), \mu_s \sim N(-3, 1)$
Logistic-Makeham	$\alpha, \beta, s, c \sim \text{log-normal}(\mu_{\alpha \beta s c}, 0.2)$	$\mu_\alpha \sim N(-3.5, 0.5), \mu_\beta \sim N(-3.2, 1), \mu_s \sim N(-3, 0.5), \mu_c \sim N(-4, 0.5)$

Table 2.4: **The priors used on parameters of each different survival model.** For the exponential model the ‘group-level’ priors were the same for the genus and ‘overall’ models that were also estimated. The exponential model was the only model that was simple enough to allow the scale parameter of the log-normal (σ) to be estimated by the data. The notation $\alpha, \beta \sim \text{log-normal}(\mu_{\alpha|\beta}, 0.2)$ means that α and β were assigned independent log-normal priors with location parameters μ_α and μ_β respectively, and a scale parameter of 0.2 in both cases.

Carlo (MCMC) methods to sample from each posterior distribution. In particular we used *Stan* software (Carpenter et al., 2016) that implements an efficient MCMC algorithm known as NUTS (Hoffman and Gelman, 2014).

To judge convergence of the sampling algorithm to the posterior density, we calculated $\hat{R}$ across all Markov chains – a ratio that compares the between-chain variation to that within each chain that is commonly used to measure convergence in MCMC (Gelman and Rubin, 1992). For each model that we estimated we conducted an independent MCMC run consisting of 12 independent Markov chains with 1,000 iterations per chain, discarding the first half of these iterations as warm-up (Gelman et al., 2014), and continued to run the algorithm until $\hat{R} < 1.1$ for all model parameters. We also ensured that across each MCMC run the number of divergent iterations (that can bias the MCMC away from the true posterior density) was minimal. In the majority of cases, the number of divergent iterations was far fewer than 1% of the total number of samples.

1642 2.3.7 K-Fold cross validation

1643 One way to estimate a model’s out-of-sample predictive capability is to use Akaike In-
1644 formation Criterion, Bayesian Information Criterion, Deviance Information Criterion or
1645 Watanabe Akaike Information Criterion that explicitly penalise a model in accordance to
1646 its complexity (typically indexed by its number of free parameters) to correct for model
1647 over-fit. The way in which this correction takes places however, is fairly heuristic, and
1648 less appropriate for hierarchical models. An alternative method, common in the machine
1649 learning literature, is known as ‘cross-validation’ (see for example, Kohavi, 1995). Here
1650 to estimate out of sample predictive capability, the original data set is partitioned into
1651 training and test sets. The model is then fitted to the training set, and its predictive
1652 performance measured on the independent test set.

1653 We use cross-validation to compare the predictive power of the species, genus and overall
1654 models, and later to determine whether age-dependent mortality occurs. In particular,
1655 we use K-Fold cross-validation where we repeatedly partition our dataset into a test set
1656 composed of a number of individual time-series, and a training set of the remainder (see
1657 for example, Marsland, 2015). We then use the fitted model to measure the predictive
1658 performance on the test set. The particular measure we use is the expected log point-wise
1659 predictive density (elpd)(Vehtari, Gelman, and Gabry, 2015) which sums the predictive
1660 performance for each data point averaged across all posterior samples.

Since the models are hierarchical we must specify how to draw parameters for a given
test sample time-series. Here we chose to draw values of the individual time-series pa-
rameters as samples from the population distribution that corresponds to that particular
sample’s respective grouping. So if the particular test sample time-series pertains to *Ae.*
aegypti (and we are using the species-level model) then we draw independent samples for

the statistical model's parameters from the estimated *Ae. aegypti* population distribution,

$$\theta_i \sim p(\theta^{\text{Ae. aegypti}} | \text{data}), \quad (2.8)$$

1661 where θ_i is the value of the parameters used in test time-series i , and $p(\theta^{\text{Ae. aegypti}} | \text{data})$ is
 1662 the overall posterior distribution across all *Ae. aegypti* experiments.

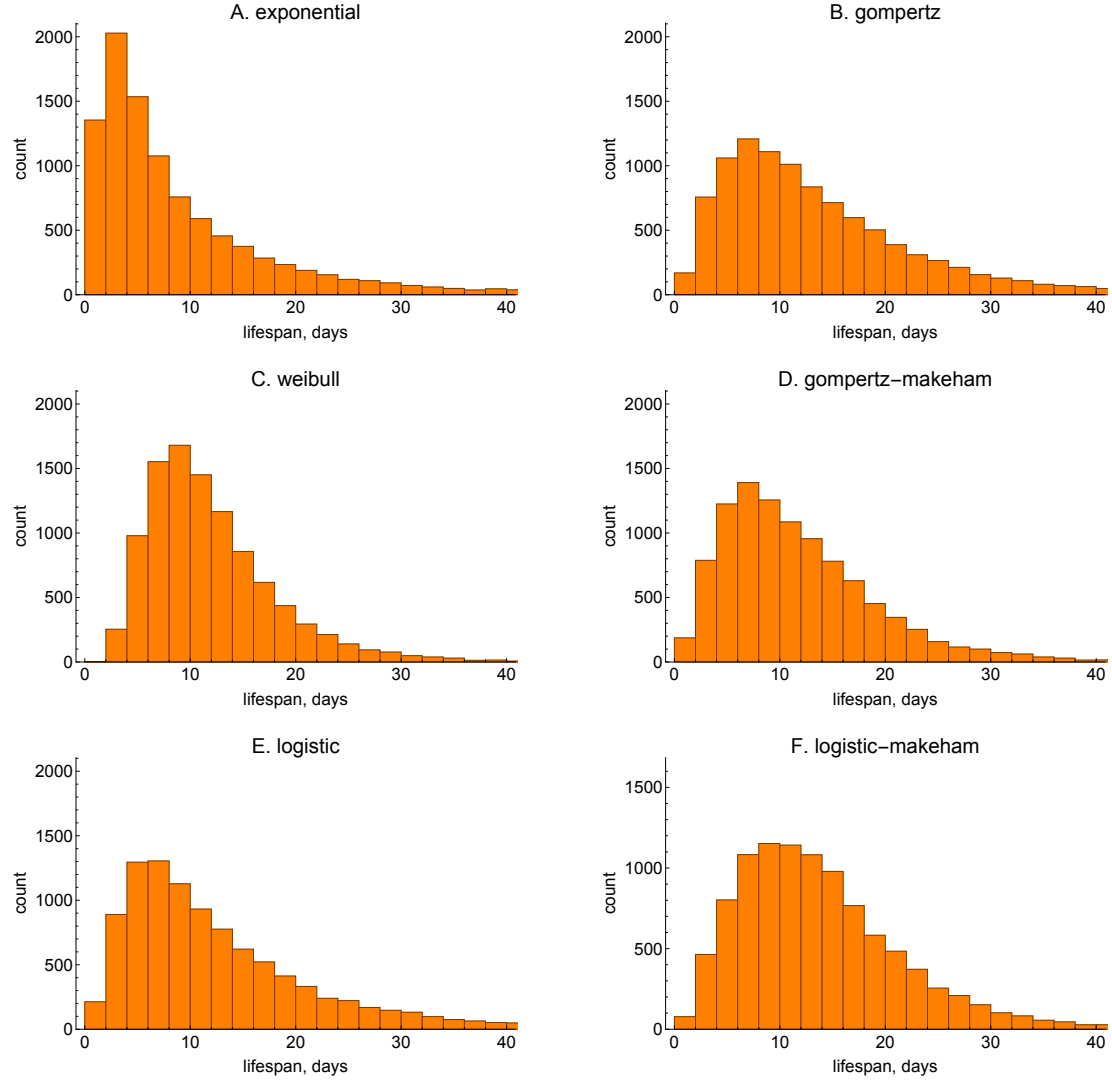


Figure 2.5: **Samples from the prior distribution of mean lifespan across the six models of mosquito mortality introduced in Table 2.3.** In all cases 10,000 samples were generated using the priors indicated in Table 2.4.

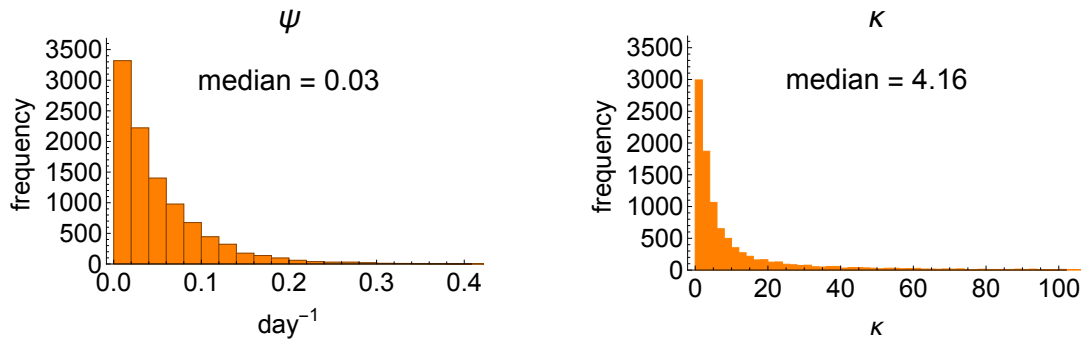


Figure 2.6: **Samples from the priors for daily recapture probability (ψ) and over-dispersion parameter (κ).** In all cases 10,000 samples were generated using the following priors: $\psi \sim \exp(\theta_\psi)$ and $\kappa \sim \exp(\theta_\kappa)$.

2.4 Results

2.4.1 Mosquito lifespan varies across species although variation is less marked at the genus level

We first present estimates of mean mosquito lifespan obtained by separate analysis of each individual time-series (Figure 2.7). Whilst there is variability in these estimates, in 211 of the 230 cases we report the median lifespan was below 10 days. (We do not present estimates for the two time-series that corresponded to experiments done on *Anopheles balabacensis* since this data was highly unusual – recaptures increased over time – and resulted in unrealistic estimates.) However, from this separate treatment of individual MRR time-series it is not immediately clear whether differences exist at the species or genus level. To determine this we estimated hierarchical Bayesian models where we grouped by species, genus or across all time-series (‘overall’.) These estimates show considerable variation at the species level (Figure 2.8 main), with the lowest estimated mean lifespan corresponding to *Anopheles (A.) subpictus s.l.*, a major vector of malaria in Asia (Sinka et al., 2012), with a median lifespan of 1.0 days. It is unlikely that this estimate is representative of the true lifespan of this mosquito species, since with such a short lifespan there would be negligible transmission of malaria. Since we cannot disentangle the effects of mosquito mortality from dispersal in our estimates, the estimates we present represent lower bounds on mosquito lifespan, which in some cases are very conservative. The highest mean lifespan was estimated for *Aedes (Ae.) cantans*, a mosquito species found in Western Europe (Lakhani, 1974; Hubálek, 2008), with a median lifespan of 14.5 days. ANOVA tests using the posterior medians from each time-series suggested that differences in mean lifespan exist at the species level ($p < 0.01$, $F = 10.76$, $df = 32$), and was robust to non-inclusion of *Anopheles balabacensis* ($p < 0.01$, $F = 2.74$, $df = 31$). The Kruskal-Wallis test also indicated that

this result was robust to non-normal residuals ($p < 0.01$, $\chi^2 = 99.83$, $df = 32$). Lifespan estimates were less variable at the genus level (Figure 2.8 inset): *Culex* mosquitoes had the lowest median estimate at 2.4 days; *Anopheles* were second highest at 3.6 days; and *Aedes* were estimated to have the highest mean lifespan at 4.5 days. ANOVA tests suggested that there were differences in mean lifespan at the genus level ($p < 0.01$, $F = 4.74$, $df = 2$). Again this result was robust to the assumption of non-normal residuals with the Kruskal-Wallis test results indicating statistically different lifespans ($p < 0.01$, $\chi^2 = 25.48$, $df = 2$). Across all genera the median posterior lifespan was estimated to be 3.7 days (Figure 2.8 inset).

To determine whether grouping by species explained more of the variation in the data than genus we conducted K-Fold cross-validation (see ‘Methods’) on each of the models. These suggested that the species and genus groupings performed better than the model where we simply pooled all data together ($p_{species} = 0.02$, $p_{genus} = 0.04$). However the species model did not lead to better prediction than the one where we grouped by genus, suggesting that the majority of variation in lifespan is explained by genus (Figure 2.9).

Why do we see such variation at the species level, and yet this model is not significantly better at prediction than the genus one? Whilst for some species grouping at this level leads to significantly better predictions, in other cases the model is over-fit to the data and so does not perform as well (Figure 2.10). In biological terms the species-grouping model performs better when the data are representative of real underlying differences in lifespan, but less well when the experimental data are influenced by significant idiosyncratic ‘random’ factors that lead to differences in lifespan estimates.

In an MRR experiment there are two mechanisms that can cause a reduction in mosquito numbers after release: 1. mosquitoes die before they are recaptured, or 2. mosquitoes disperse out of the study area. Since our analysis is based on spatially-averaged

MRR data our estimates actually represent the combined effect of mortality and dispersal ('the time in study area'). However, if mosquito dispersal causes a considerable reduction in mosquito numbers after release then studies with a wider recapture radius should have a higher estimated longevity than studies with a smaller radius. The spatial extent of the study should therefore be *positively* correlated with our estimates if dispersal is an important omitted variable in our models. However there is no such correlation (Figure 2.11). This lack of correlation suggests that our estimates of lifespan may not be strongly biased downwards. However, it is possible that such a correlation exists due to low recapture success of outer traps in studies with a larger recapture range.

2.4.2 A small proportion of mosquitoes are likely responsible for the bulk burden of disease for malaria and Zika

Whilst our principal estimates of adult mosquito lifespan (Figure 2.7 and 2.8) represent *means*, it is the distribution of lifespans at the individual level that dictates vectorial capacity. We therefore used the posterior parameter samples for the rate parameter in our exponential distribution to sample quantiles from the survival curves. These curves show the proportion of individual mosquitoes that live up until a given age. In particular, we estimated the proportion of mosquitoes that survived the extrinsic incubation period (EIP) for those species that are identified as vectors for malaria, chikungunya, dengue fever and Zika.

For the *Anopheles* species included in this analysis there was a large range in the percentage of mosquitoes surviving the EIP for malaria (Figure 2.12) which we assumed to be 10 days (Blanford et al., 2013): our posterior median estimates indicate that <0.1% of *An. subpictus s.l.* and 46% *An. sergenti* survived long enough to become infectious with malaria. Further ANOVA tests on the posterior median proportions for the individual

time-series suggested that differences exist in the fraction of mosquitoes that survive the EIP for malaria ($p < 0.01$, $F = 32.91$, $df = 16$, $n = 88$ time-series). The Kruskal-Wallis test confirmed this result was robust to non-normal residuals ($p < 0.01$, $\chi^2 = 62.43$, $df = 16$, $n = 88$ time-series). Importantly, for the major African vectors of malaria, *An. gambiae s.l.* and *An. funestus s.l.*, we obtained posterior median estimates that indicate that only 2% and 4% of individual mosquitoes survive the EIP.

For the three other diseases we consider: chikungunya, dengue fever and Zika, the two species *Ae. aegypti* and *Ae. albopictus* are the vectors responsible for causing the majority of disease cases (Pialoux et al., 2007b; World Health Organisation, 2002; Bogoch et al., 2016). The EIP for Chikungunya is the shortest of the diseases we consider, at around 2 days (Dubrulle et al., 2009). The proportion of mosquitoes surviving this period is considerably higher than for malaria (Figure 2.12, bottom): we estimate that 59% *Ae. aegypti* and 76% *A. albopictus* mosquitoes live long enough to transmit chikungunya. The EIP for dengue fever is thought to depend particularly strongly on temperature, and so we consider the EIP at a fixed temperature of 30° C, estimated to be approximately 6.5 days in a recent meta-analysis (Chan and Johansson, 2012). Assuming this EIP we estimate that 22% *Ae. aegypti* and 42% *Ae. albopictus* mosquitoes live long enough to pass on dengue fever. Whilst the research into Zika is at a fairly early stage, the existent estimates of the EIP for the disease suggest that its EIP is longer than for the other diseases we consider. Here we assume an EIP of 12.5 days, which represents an average of two recently-published estimates (Chouin-Carneiro et al., 2016; Di Luca et al., 2016). For this EIP we estimate that 8% *Ae. aegypti* and 19% *A. albopictus* survive to be able to transmit Zika.

2.4.3 The longevity of mosquitoes does not vary systematically over the range of temperatures of MRR study locations

The Guerra et al., 2014 database included the latitude and longitude of each study, along with the date when the study began. We used this information to find estimates of the air temperature for each study using the European Centre for Medium Range Weather Forecasts' historical database. For each study we calculated the mean monthly temperature across a spatial area of ($\text{lat} \pm 1$ degree, $\text{long} \pm 1$ degree), for the month at which each study was carried out. Unfortunately, the records for this database begin in 1979, which pre-dates the study date for 65 of our 232 MRR time-series. For these time-series, we chose to estimate the air temperature by an average of monthly temperatures over the years 1979-89.

We carried out a regression of the posterior median lifespan estimated for each time-series (using the hierarchical model where we grouped by species) against air temperature including both linear and quadratic temperature terms (Figure 2.13). The effect of temperature was not statistically significant at the 5% level.

2.4.4 Male mosquitoes have shorter lives than females

The Guerra et al., 2014 database contains information on whether the releases consisted only of females, males or a mixture of mosquitoes. For some species in the database there are not releases of all of these sex groupings. This meant that we could not produce species-level estimates of lifespan for each sex. Fortunately for each genus, as well as overall, we were able to produce estimates of lifespan for female-only and male-only releases (Figure 2.14).

Overall we estimate that female mosquitoes live for approximately 0.8 days longer than males (difference in posterior median estimates); female mosquitoes live for approximately

3.8 days and males for 3.0 days. An ANOVA test using the posterior median estimates from the individual series indicated that there are no significant differences between males and female lifespans. However, this pattern repeated across all three genera in our sample. For *Culex* mosquitoes females lived for 0.5 days longer, on average, than males; for *Anopheles* females lived for 1.5 days longer; and for *Aedes* the difference in lifespan was 1.6 days. Since male mosquitoes have previously been observed to live shorter lives than females in the lab (Styer et al., 2007) and field (Silver, 2007), we believe that our results (whilst not statistically significant) may be indicative of this trend.

We do not know of any previous work that has modelled the effect of a mosquito lifespan that varies by sex. However further modelling work is needed to determine the implications of such differences in lifespan by sex to vector control efforts.

2.4.5 Blood- and sugar-fed mosquitoes do not live longer than those that were not fed before release

The Guerra et al., 2014 database contains data on whether the released mosquitoes were blood- and/or sugar-fed prior to release. To ensure that we had sufficient statistical power we estimated the lifespans of each category across all time-series in our dataset (Figure 2.15). Whilst there is a slight increase from no-feeding to sugar-only then blood-only then sugar-and-blood (also seen in the genus-level estimates – data not shown), ANOVA tests found that this effect was not statistically significant.

2.4.6 Wild mosquitoes do not appear to experience strong age-dependent mortality

We estimated six different models for mosquito mortality; each representing a previously-used model in the literature (see ‘Methods’). Five of these models allowed for the possibility

1805 of a mortality hazard that increases with age; the exponential model is the only model which
1806 does not allow a variable age-dependent mortality. Here we compare the predictive perfor-
1807 mance of each of these models using K-Fold cross-validation (see ‘Methods’). In particular
1808 we estimate the average log probability for a single data point across our independent test
1809 sets, with higher values for this quantity indicating a better fit of the model to the data.

1810 There was some variability in the fit of each model across all the species (Figure 2.16).
1811 However, in all cases the exponential model either performed statistically equivalent to, or
1812 better than, the models that allowed for an age-dependent mortality.

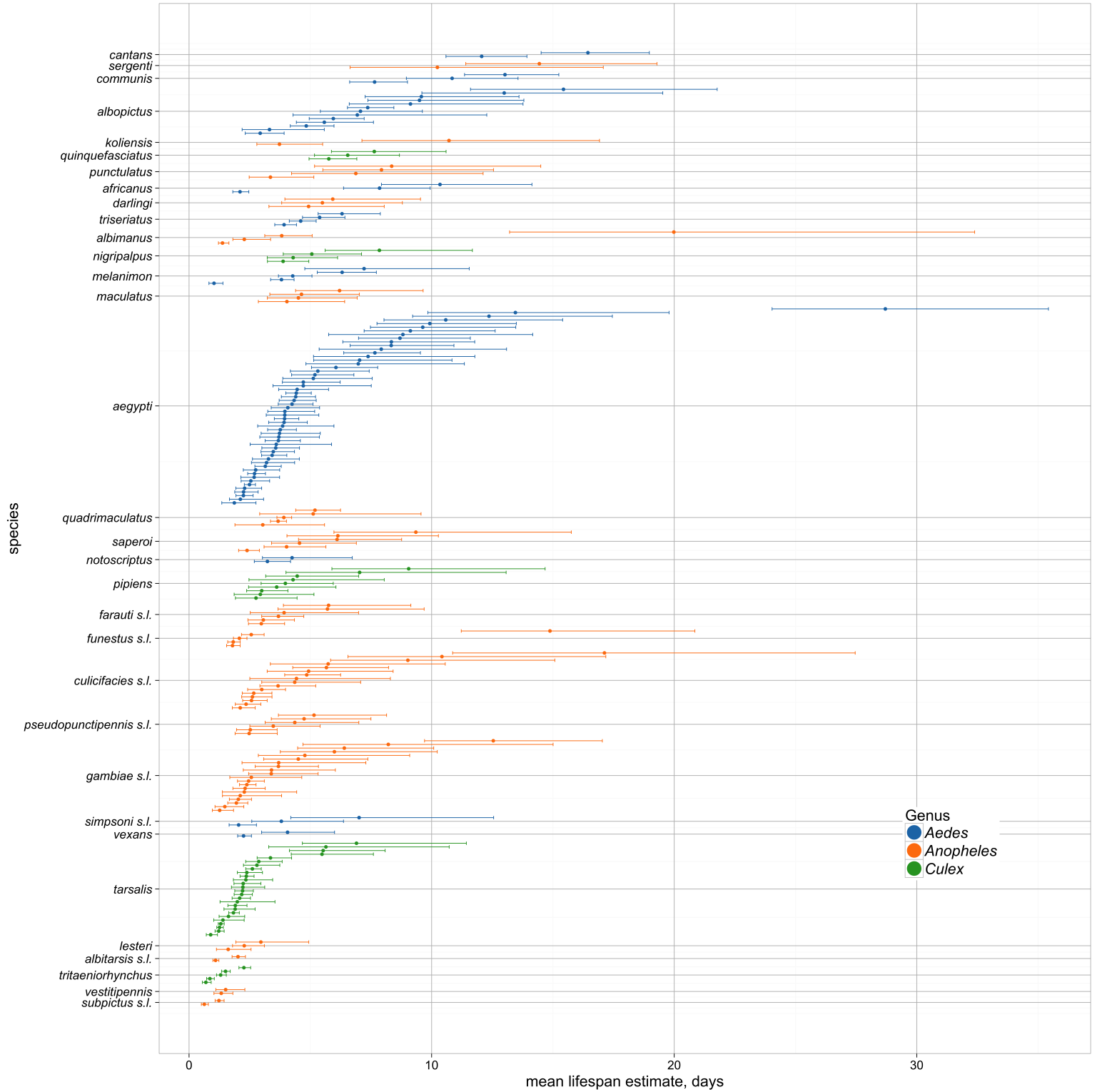


Figure 2.7: **Individual time-series estimates of adult mosquito mean lifespan from the MRR analysis ordered by species.** The middle line in each box shows the median estimates. The left and right box whiskers show the 25%, and 75% posterior quantiles respectively. All estimates were obtained using the non-hierarchical exponential survival model as described in text.

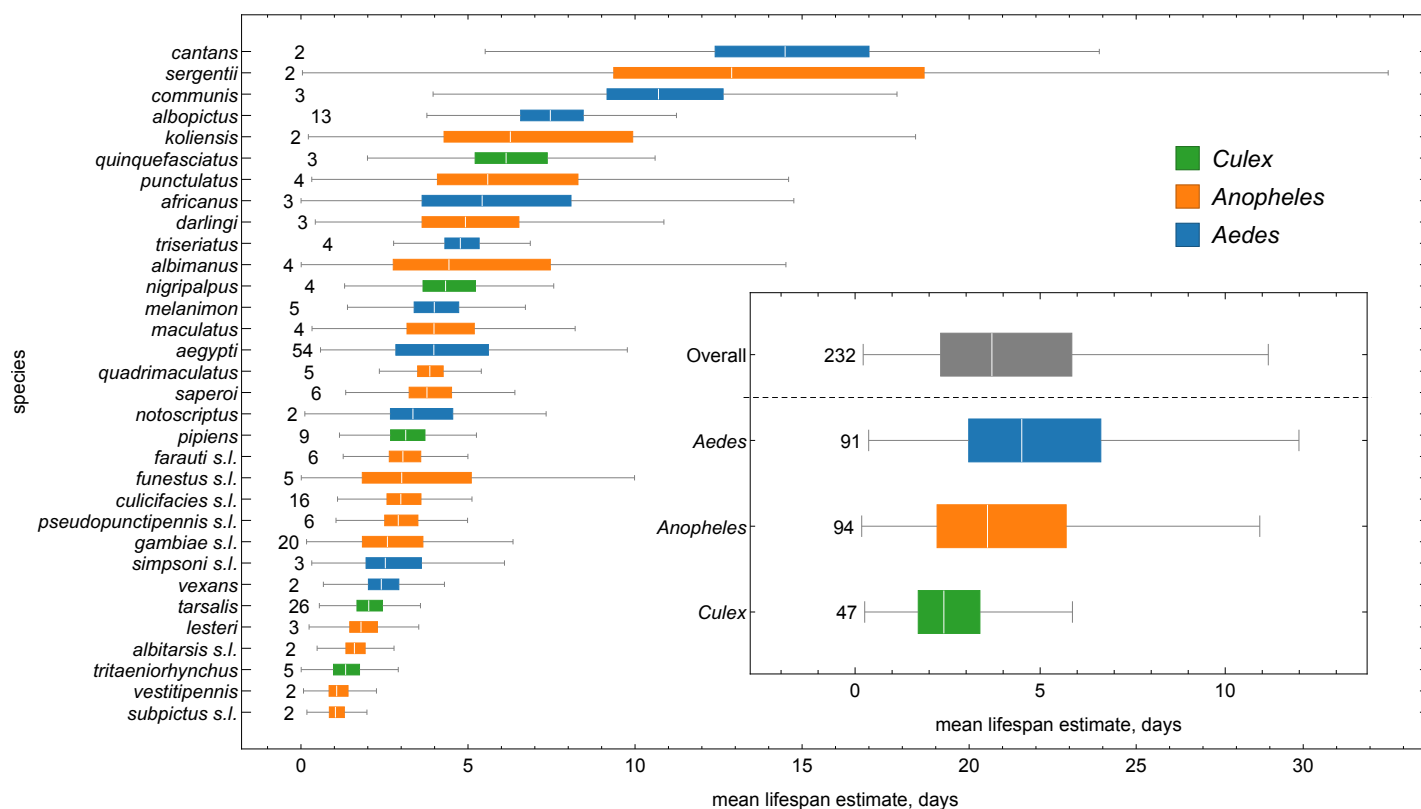


Figure 2.8: **Posterior estimates of adult mosquito mean lifespan across species, genus and overall groupings as determined from the MRR data.** The middle line in each box shows the median estimates. The left and right box edges show the 25%, and 75% posterior quantiles respectively. The whiskers show the range of the data, excluding points lying more than 1.5 times the interquartile range away from each edge of the box. The numbers before the start of the left whisker indicate the number of individual time-series within each species. All estimates were obtained using the exponential survival model.

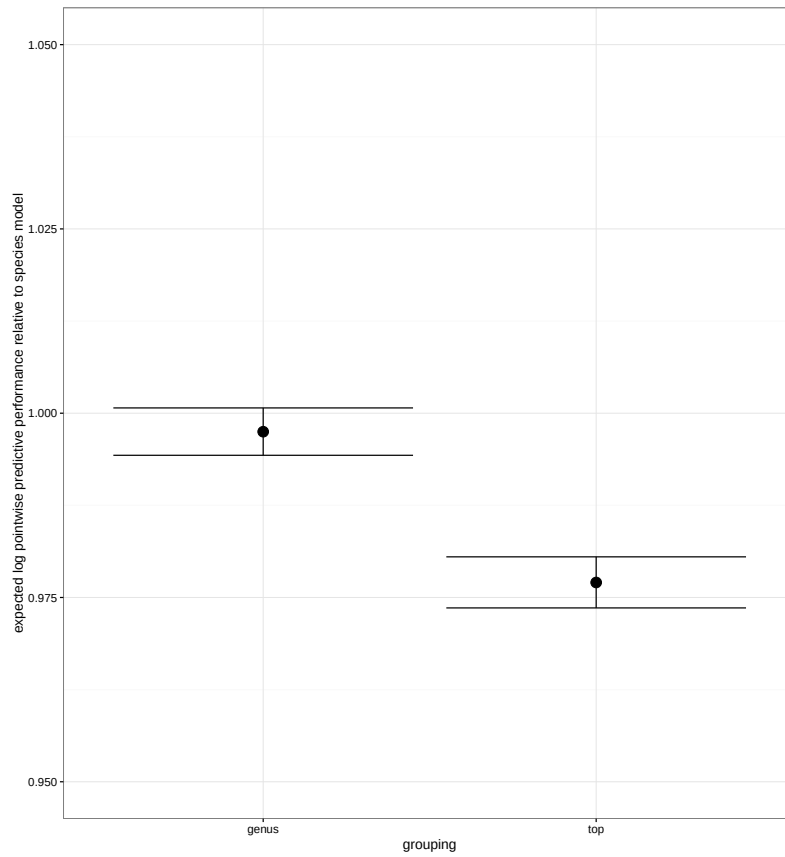


Figure 2.9: **The estimated out-of-sample average log pointwise predictive performance of the genus- and overall-level models relative to the species-level model.** Values higher up indicate a more predictive model. In all cases the predictive performance of each model was measured via K-Fold cross-validation. The upper and lower tails show the approximate 95% confidence intervals on the predictive performance relative to the species-level model.

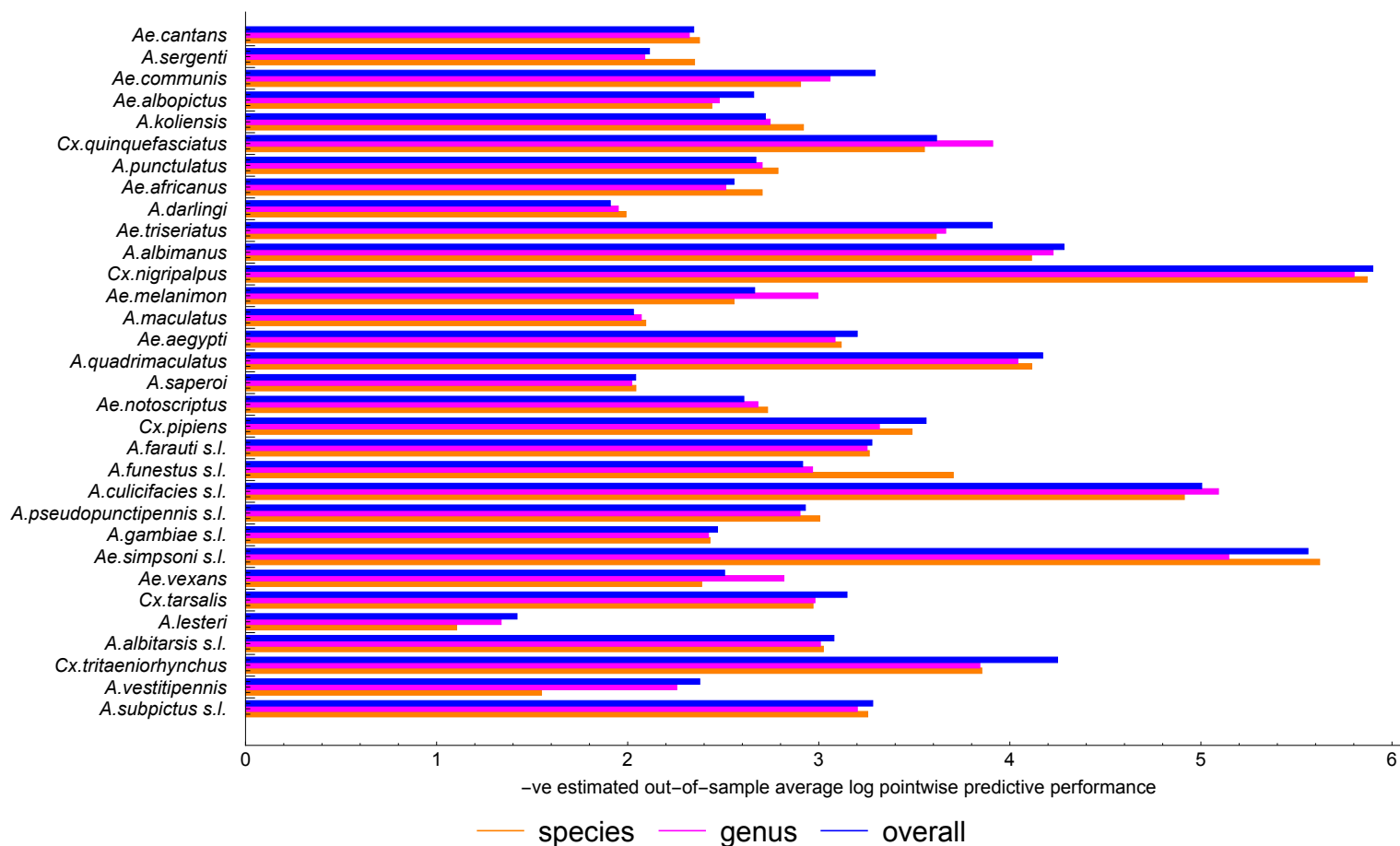


Figure 2.10: The estimated out-of-sample average log pointwise predictive performance of the species-, genus- and overall-level models, across each species. Note the scale of the horizontal axis is -1 times the predictive performance, and hence values closest to zero indicate the best performing model. In all cases the out-of-sample predictive performance was estimated by K-Fold cross-validation.

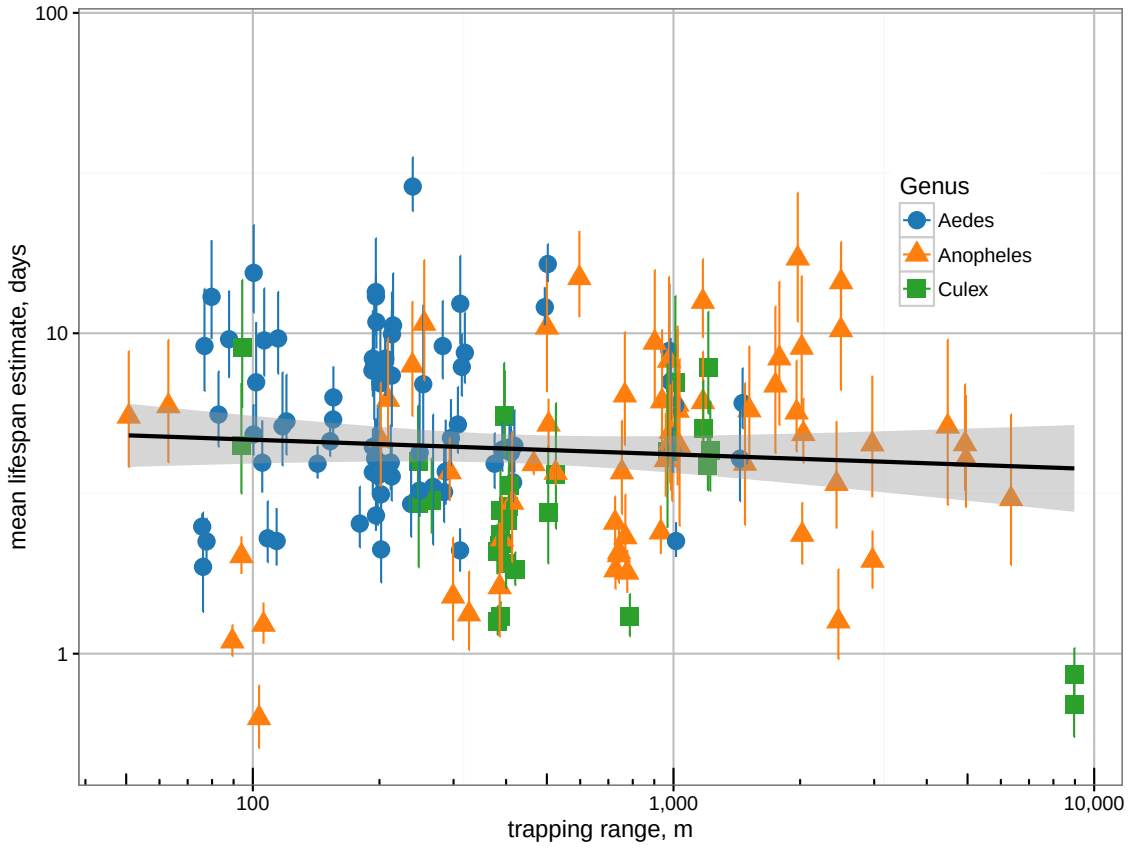


Figure 2.11: **Posterior estimates of mean mosquito lifespan for each time-series, versus the trapping range.** The markers show the median posterior estimates, with the lower and upper bounds indicating the 25% and 75% quantiles, respectively. The black line shows a linear regression line estimated using the median posterior lifetimes, with the grey shading indicating 95% confidence intervals. All estimates were obtained using the exponential survival model.

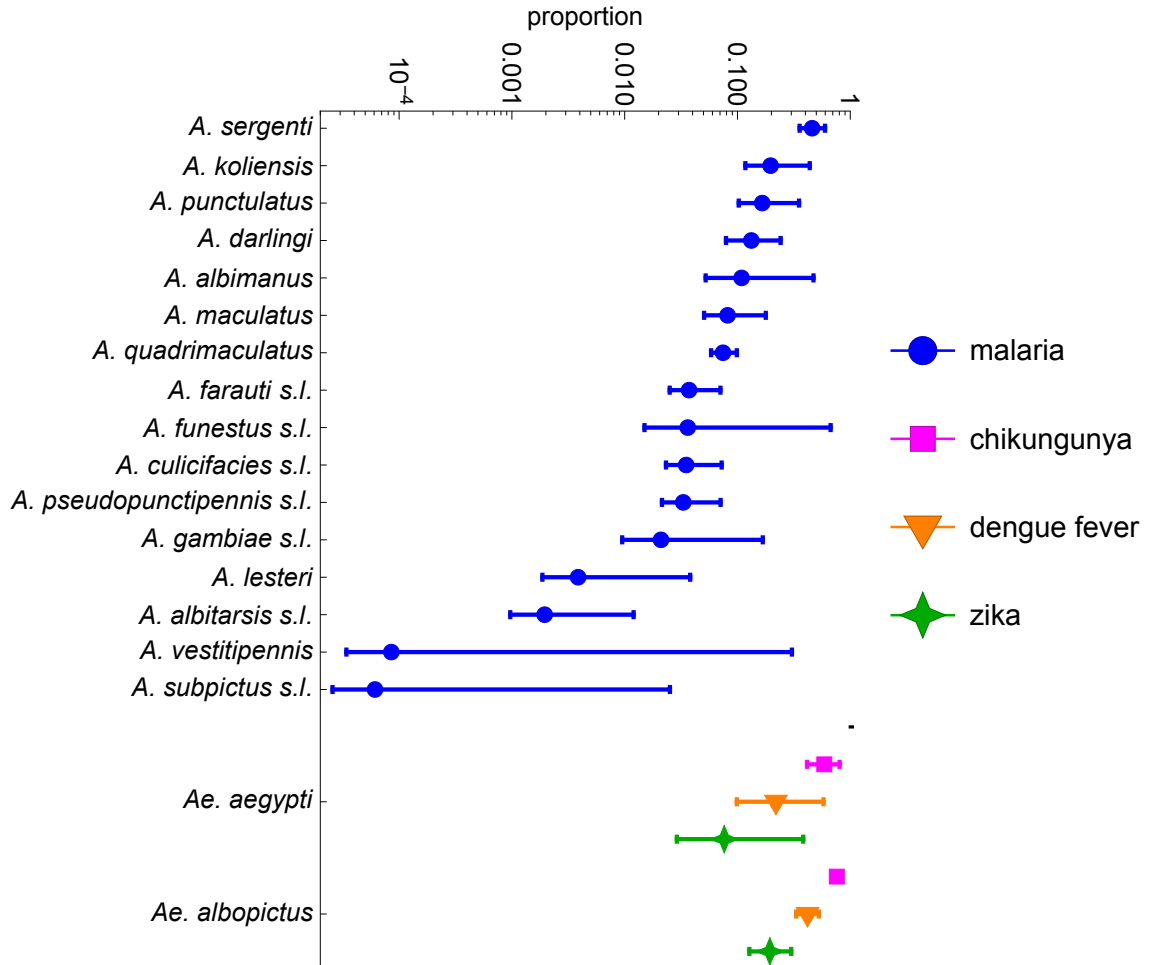


Figure 2.12: The estimated proportion of mosquitoes surviving the EIP of malaria, chikungunya, dengue fever and Zika for the main disease vectors. Here we assumed that the EIPs were: malaria-10 days, chikungunya-2 days, dengue fever-6.5 days and Zika-12.5 days. The left and right whiskers indicate the 25% and 75% posterior quantiles, and the points indicate the mean. All estimates were obtained using the exponential survival model.

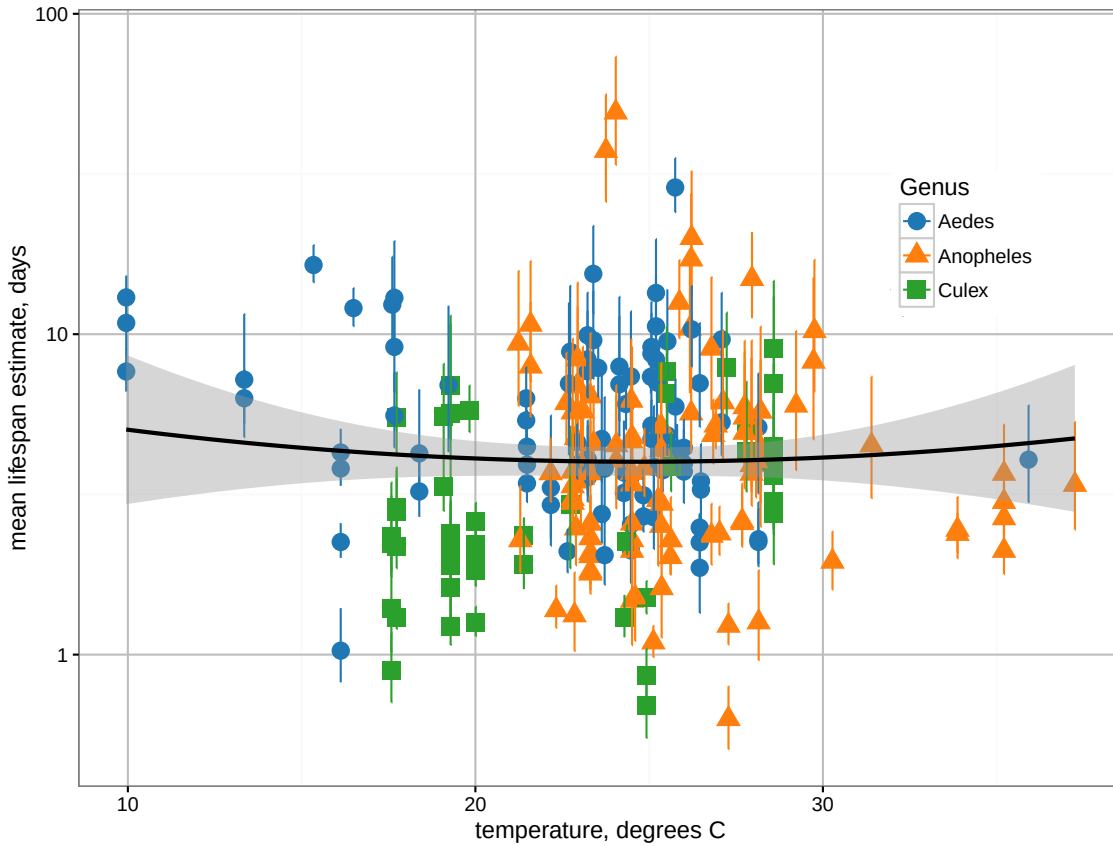


Figure 2.13: **Posterior estimates of mean mosquito lifespan for each time-series versus the average monthly temperature for each study location and date.** The markers show the median posterior estimates, with the lower and upper bounds indicating the 25% and 75% quantiles, respectively. The black line shows a regression line with linear and quadratic terms estimated using the median posterior lifetimes, with the grey shading indicating 95% confidence intervals. All estimates of lifespan were obtained using the hierarchical exponential survival model.

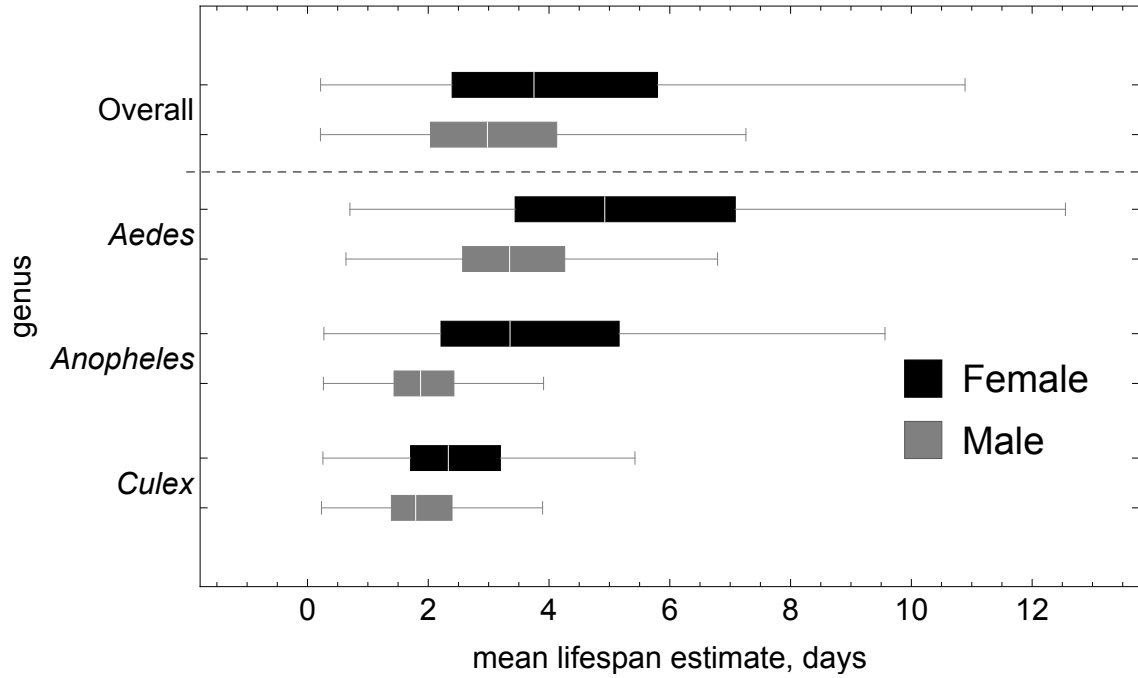


Figure 2.14: **Posterior estimates of mean mosquito lifespan across female-only and male-only releases, across all studies (“overall”) and by genus.** The markers show the median posterior estimates, with the lower and upper bounds indicating the 25% and 75% quantiles, respectively. The numbers before the start of the left whisker indicate the number of individual time-series within each category. All estimates were obtained using the exponential survival model. **Add numbers involved in each case.**

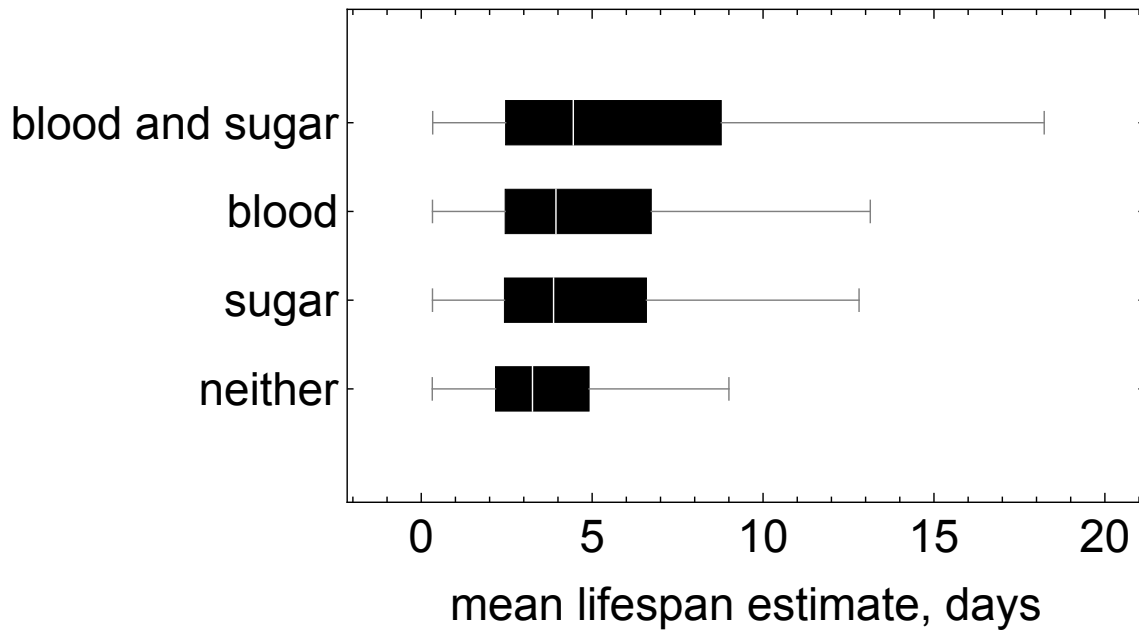


Figure 2.15: **Posterior estimates of mean mosquito lifespan by pre-release feeding method across all studies.** The markers show the median posterior estimates, with the lower and upper bounds indicating the 25% and 75% quantiles, respectively. The numbers before the start of the left whisker indicate the number of individual time-series within each category. All estimates were obtained using the exponential survival model. **Add numbers involved in each case.**

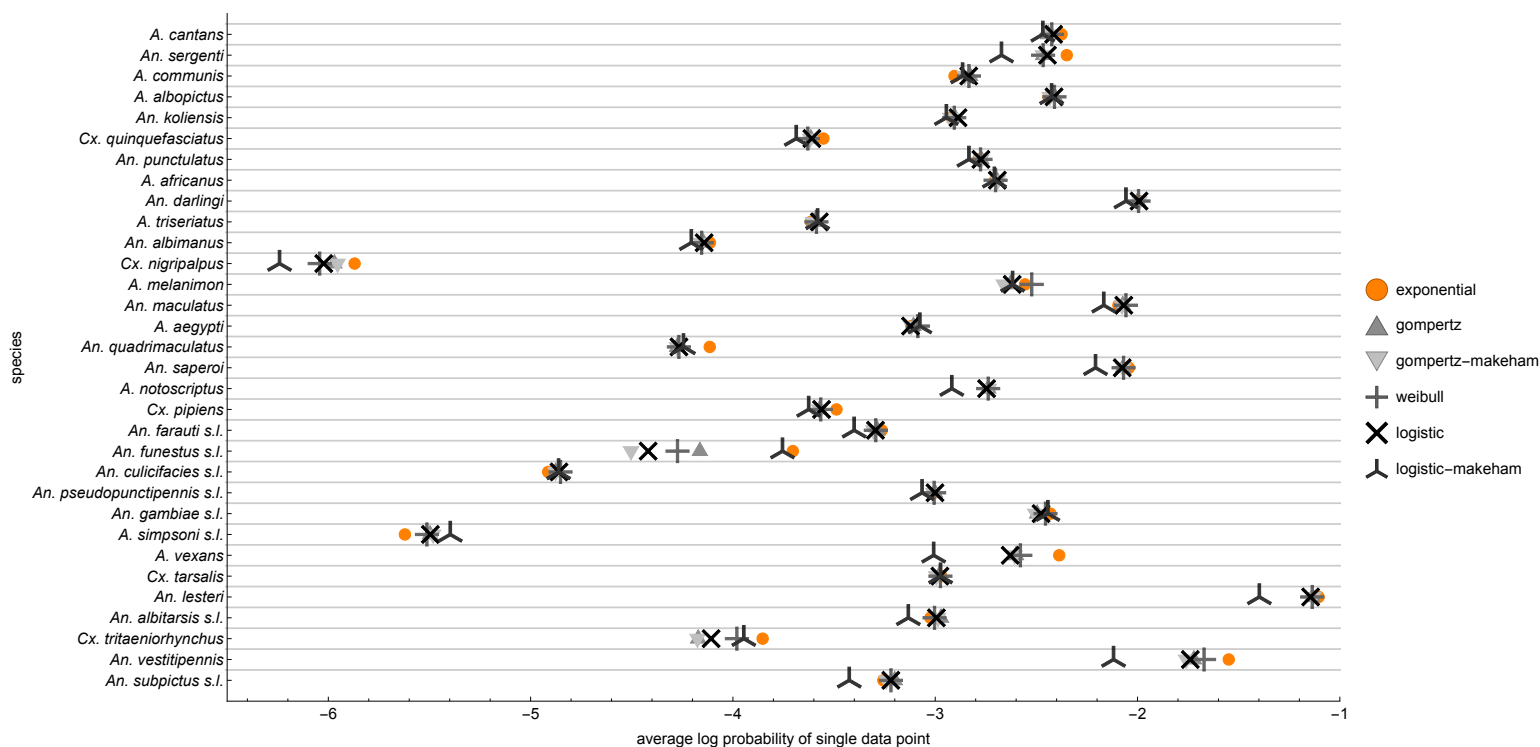


Figure 2.16: **Average log probability for each data point from each of the estimated models of mosquito mortality.** Values towards the right indicate a better predictive fit for a given model. Across all species there were no cases where the exponential model performed significantly worse than the model with the best fit. The average log probability was calculated on independent test sets using K-Fold cross-validation.

2.5 Discussion

Mosquito longevity is an important epidemiological parameter, yet one that cannot be easily inferred with MRR. The MRR experiment has for many decades been a widely used method to estimate longevity (Silver, 2007), yet the approach is often beset by low recapture rates leading to large confidence bounds on results. Furthermore, the methods used to estimate longevity from MRR data often differ, while several unavoidable assumptions may bias MRR-based results systematically. In this study we have applied common estimation procedures to data from a large number of MRR studies to reduce, as far as is possible, the sampling noise inherent in individual experiments, and to overcome the difficulty of interpreting results based on different analytical procedures. We now discuss these findings although delay comparison with dissection-based methods (Polovodova, 1949; Detinova, 1962) until Chapter 3 where we carry out a meta-analysis of such experiments. Finally, we draw on our analysis to present an appraisal of the MRR experimental procedure itself, as a means to estimating mosquito longevity.

The main results of this analysis are as follows:

1. Most mosquitoes live for less than ten days, which implies that only a small fraction of adult females can act as vectors for disease (depending on the disease and mosquito species). Although a few species live longer than this, the main vectors of malaria, dengue fever, Zika and chikungunya are all found to exhibit relatively short average lifespans (Figure 2.8).
2. The data does not reject the exponential survival model, which assumes a risk of mortality that is independent of age. However, age-dependent mortality has been detected in lab populations (Styer et al., 2007; Harrington et al., 2008; Dawes et al., 2009) and in field populations based on dissections of wild-caught adult fe-

1837 males Clements and Paterson, 1981. This result may be due to the rarity of marked
1838 mosquitoes that live long enough to experience physiological decline, which implies
1839 that senescence may not be an important effect in the field.

1840 3. We found no evidence of a systematic effect of temperature on mosquito lifespan
1841 (Figure 2.13). However, field observations and lab experiments have suggested the
1842 adult lifespan of *Ae. aegypti* mosquitoes is affected by temperature Jepson, Moutia,
1843 and Courtois, 1947; Yang et al., 2009; Lambrechts et al., 2011. It is unclear whether
1844 our result is because our sample has relatively few MRR experiments at extreme
1845 temperatures, or because the data lacks the statistical power to identify an effect.
1846 Nevertheless, our results suggest that the effect of temperature on adult mosquito
1847 lifespan may be relatively small in magnitude.

1848 4. The data indicates that wild female mosquitoes live longer than males (Figure 2.14),
1849 which has been previously found in laboratory (Styer et al., 2007) and wild (Silver,
1850 2007) populations. There is increasing interest in the biology of male mosquitoes due
1851 to the development of male-based population control methods (Ferguson et al., 2005),
1852 such as sterile insect technique (Benedict and Robinson, 2003; Alphey et al., 2010),
1853 and transgenic methods that aim to bias the sex-ratio of wild populations towards
1854 males (Lees et al., 2014; Galizi et al., 2014; Adelman and Tu, 2016). This increased
1855 focus on male mosquitoes as well as ethical concerns have likely lead to an increased
1856 frequency of MRR experiments that release only males (Figure 2.17). However we
1857 believe that experiments that elucidate aspects of the ecology of female mosquitoes
1858 continue to be of value so long as adequate measures are taken to mitigate risks to
1859 humans.

1860 The key assumption of MRR-based methods are: (i) the behaviour and survival of

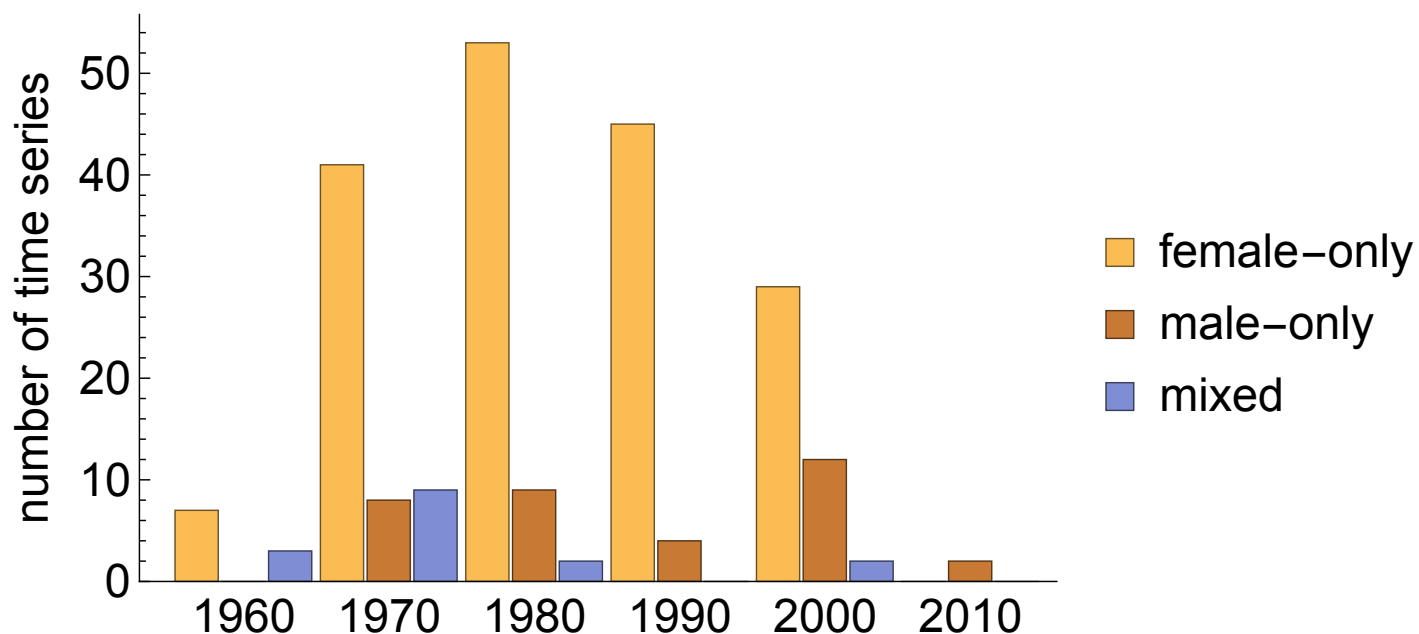


Figure 2.17: **The number of MRR time-series by date and sex of released mosquitoes.** The numbers represent by-decade totals of the 232 time-series we used in the main analysis.

mosquitoes are not affected by their marking or handling, and (ii) the rate of mosquito emigration from the study area is small compared to the rate of mortality. Mosquitoes are delicate insects and it is plausible that many marking methods will impair fitness to some degree, yet unfortunately there are seldom data to quantify this. A few studies have measured the survival of *Ae. aegypti* and *A. gambiae* mosquitoes in laboratory conditions after marking with a number of the fluorescent dusts commonly used in MRR studies (Muir and Kay, 1998; Verhulst, Loonen, and Takken, 2013; Dickens and Brant, 2014). Most experiments found that marking leads to reductions in survival (depending on the particular dye), indicating that survival estimates based on MRR experiments will underestimate longevity. If the assumption of limited emigration is incorrect, then the results from the MRR analysis will also underestimate the longevity of wild mosquitoes.

1872 There was no correlation between estimated lifespan and trapping range (fig. 2.11),
1873 suggesting that in most MRR experiments, the study area is sufficient to preclude a major
1874 influence of emigration. However, the apparent absence of an effect of trapping range may
1875 be due to a low recapture success of outer traps in studies with a larger recapture range.
1876 Indeed we believe that our estimates are likely downwardly biased due to the effects of
1877 mosquito dispersal. This is supported by our estimates of lifespan, which suggest there
1878 are insufficient mosquitoes that survive long enough to support mosquito-borne diseases in
1879 geographies where these diseases are endemic.

1880 Taken together, both assumptions indicate that our estimates of longevity are down-
1881 wardly biased, and that our results represent lower bounds on mosquito longevity. Addi-
1882 tional data that addresses these assumptions will clearly be very helpful in clarifying the
1883 possible extent of systematic bias in MRR based longevity estimates. Laboratory based
1884 studies are a useful approach to address the effects of marking, and we encourage en-
1885 tomologists to consider performing similar experiments in conjunction with future MRR
1886 experiments.

1887 **2.5.1 The statistical power of MRR studies**

1888 Consider an ideal MRR experiment, where the assumptions of no emigration and harmless
1889 marking are fully satisfied. With these data, how accurately can we estimate mosquito lifes-
1890 pan? To address this question we generated artificial MRR data and attempted to estimate
1891 the (known) parameters by maximum likelihood (a ‘Monte Carlo’ analysis). Specifically,
1892 we simulated releases of N mosquitoes that are monitored for m days in each case, and
1893 determined how statistical power to estimate lifespan depends on these two parameters
1894 (Figure 2.18). In order to focus on estimating lifespan, we assumed the mosquitoes ex-
1895 perience constant mortality, and recaptures follow the negative binomial sampling model

(eqn. 2.1) where the recapture probability parameters (ψ and κ) were the averages of those estimated for the actual data.

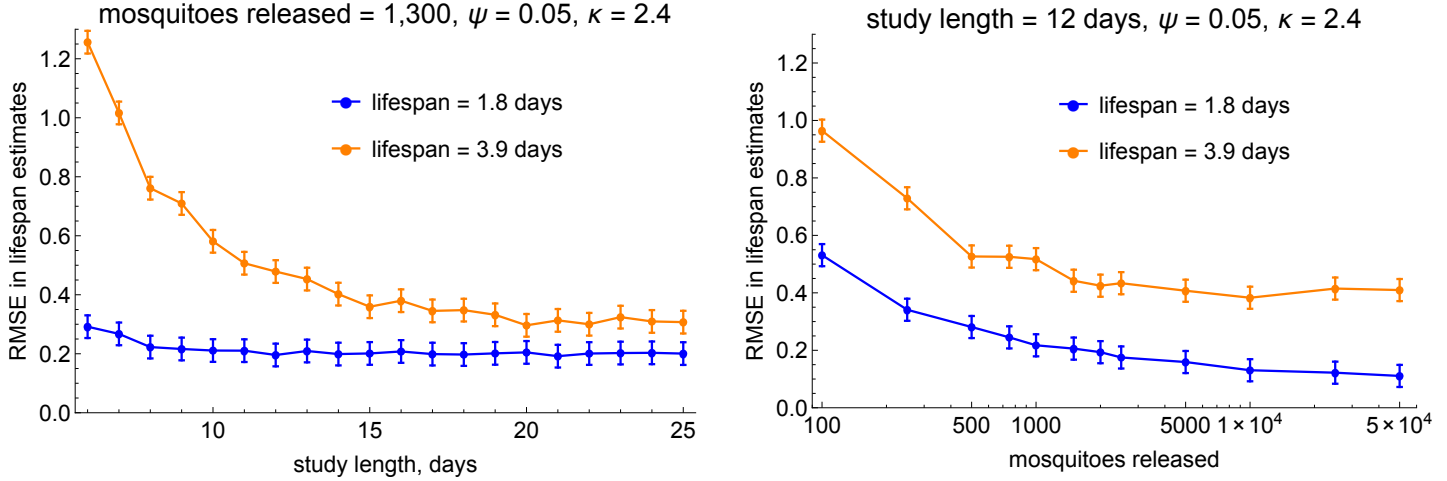


Figure 2.18: **The average error in predicting mean lifespan as a function of study length (left) and number of released mosquitoes (right) in a Monte Carlo analysis.** For each parameter set we generated 400 simulated data series using the negative binomial sampling model with an exponential survival function as described in the text, and estimated the mean lifespan using maximum likelihood. The error bars show the standard deviation in the prediction error.

Unsurprisingly, the error in predicting lifespan declines as the duration of study is increased (Figure 2.18A). However once a study length is much longer than the lifespan of the majority of mosquitoes, the predictive power cannot be improved by extending the study duration. The effect of increasing release size (while holding study length constant) is similarly a case of diminishing returns (Figure 2.18B). For the parameters we use, there

are significant gains in accuracy from releasing 1,000 rather than 100 mosquitoes, but very modest gains from releasing 10,000 rather than 1000 mosquitoes.

We also conducted a power analysis for the detection of senescence in MRR experiments (Figure 2.19). Here we calculated the power of a maximum likelihood estimator of the ‘senescence parameter’ β of the Gompertz survival function (see Table 2.3) for case study populations with three different levels of senescence (Figure 2.19A.). These indicate that the power to detect senescence is strongly dependent on study length (Figure 2.19B.), but is insensitive to release size (Figure 2.19C.). Clements and Paterson, 1981 conducted a meta-analysis of MRR and dissection-based experiments and found evidence of senescence that is, at least, qualitatively similar in magnitude to that of the ‘mild’ case considered above. For this case detecting senescence with a power of 80% requires a study length of at least 18 days. Since the median study length for experiments included in our analysis was 10 days (Table 2.1) this could partly explain our failure to detect senescence at the species level.

2.5.2 MRR design

To help inform future MRR experimental design, we investigated the influence of MRR experimental set up on recapture rate. Specifically we regressed the recapture percentage from each individual MRR time-series against a range of experimental design covariates, including both procedural factors (duration of the study, release size, collection method, trapping range etc.) and biological factors (mosquito diet prior to release, genus, sex; Table 2.5).

Surprisingly, this analysis suggests that the only reliable means of improving recapture rate is to increase the duration of a study. The number of traps was found to negatively affect the recapture rate, in contrast to our expectations. A possible explanation for this

Independent variable	(1)	(2)	(3)
experimental conditions			
duration of study	0.08** (0.03)	0.19** (0.05)	0.14** (0.05)
duration of study, squared	-	-0.03** (0.001)	-0.03** (0.001)
number released, '000s	-0.03** (0.01)	-0.03** (0.01)	-0.03** (0.01)
number of traps	-0.003** (0.001)	-	-
trapping range	-0.0003** (0.000)	-0.0002** (0.000)	-0.0003** (0.000)
human landing catches	-	0.31 (0.37)	0.22 (0.36)
CO ₂ /light traps	-	0.35 (0.30)	0.29 (0.85)
aspirators	-	-0.10 (0.33)	-0.11 (0.36)
animal bait	-	0.51 (0.40)	0.77 (0.44)
mosquito ecology			
<i>Anopheles</i>	-	-	0.30 (0.33)
<i>Aedes</i>	-	-	0.41 (0.36)
male	-	-	-0.76** (0.27)
mix	-	-	0.20 (0.31)
blood-fed	-	-	0.11 (0.21)
sugar-fed	-	-	0.24 (0.22)
other			
constant	-4.43** (0.45)	-4.51** (0.47)	-4.40** (0.53)
R-squared	0.23	0.32	0.37

Table 2.5: **Results from a regression analysis of determinants of recapture success rate.** Columns (1), (2), and (3) indicate different regressions. In all three cases the dependent variable was the log of the percentage of marked mosquitoes that were recaptured over the course of each MRR experiment. The standard errors for each parameter are shown in parentheses, and */** indicate significance at the 5%/1% levels. The models were estimated using ordinary least squares.

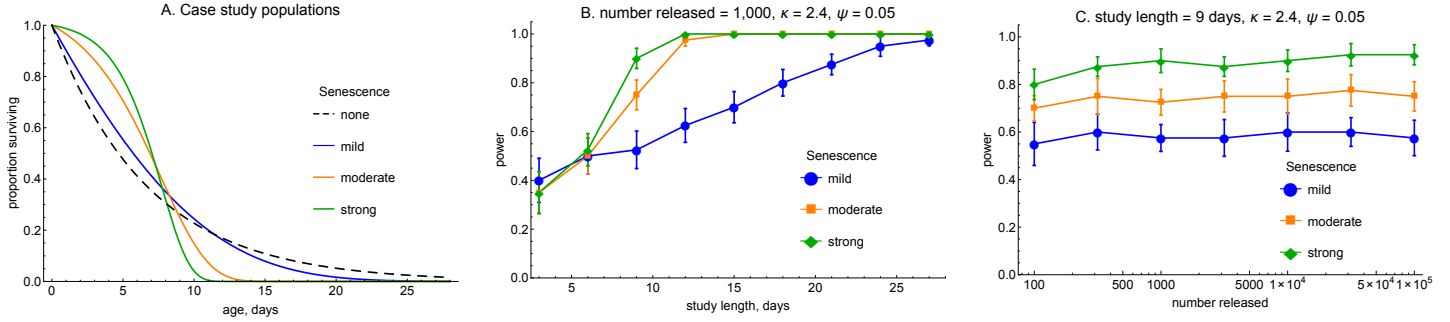


Figure 2.19: **The statistical power to detect senescence for A. three case study populations as a function of B. study length and C. number released.** For each parameter set we generated 500 simulated data series using the negative binomial sampling model with a gompertz hazard function ($\lambda(t) = \alpha e^{\beta t}$), and estimated the β parameter using maximum likelihood and tested the null hypothesis $\beta = 0$ (constant mortality risk) against the alternative $\beta > 0$ using an likelihood ratio test with a 5% test size. The error bars show the standard deviation in the power for each parameter set. The recapture probability parameters (ψ and κ) used were the averages of those estimated for the actual data.

1927 result is that we are failing to adequately capture the spatial distribution of traps, by
 1928 simply using the distance of the furthest trap from the release point. So as the number of
 1929 traps increases, they may be distributed less densely in space, resulting in a lower recapture
 1930 rate. The trapping range and release size were similarly found to negatively impact the
 1931 recapture rate.

1932 We believe that the quality of the data was not sufficient to rank the four collection
 1933 methods (human landing catches, light traps, aspirators and animal bait), where based
 1934 on previous observations it is expected that human landing catches likely produce the
 1935 highest recapture rates (see for example Mathenge et al., 2004; Kroeckel et al., 2006).
 1936 However, it is possible that differences between recapture methods are masked by not

accounting for species-level variation the efficiency of different methods are known to vary by species Mathenge et al., 2005. Male-only releases suffer from a lower recapture rate than female-only releases, which is unsurprising since most trapping methods are based on female-specific behaviour, but may also reflect their shorter lifespan.

2.6 Conclusion

In this chapter, I presented estimates of mosquito lifespan from an analysis of a database of MRR experiments. This analysis allows us, for the first time, to estimate lifespan by species and genus, and to investigate how environmental factors affect mosquito mortality. This work indicates that mosquito lifespan varies by species and genus and that female mosquitoes live longer than males. There was no evidence to suggest any systematic variation in mosquito mortality according to the average outdoor temperature of study location, however, we cannot discount the possibility that other environmental covariates, such as distance from human habitation, suitable water bodies or rainfall may affect lifespan. By fitting a range of different survival models to the data – each embodying different assumptions about mosquito mortality – this allowed us to test whether wild mosquitoes experience senescence. Across all species, we did not find any evidence that age-dependent mortality could better explain the data than the simpler exponential survival model.

In estimating mosquito lifespan from spatially-averaged MRR experiments, two assumptions are made: mosquito dispersal out of the study area is negligible compared to the rate of death; and that handling of mosquitoes and their marking does not damage them enough to affect their survival. In reality, both of these assumptions may be violated – mosquitoes have been found to migrate long distances (Dao et al., 2014) and dye has been demonstrated to significantly affect survival (Verhulst, Loonen, and Takken, 2013) – leading to estimates of lifespan that are downwardly biased. Without a gold standard

1961 approach for estimating mosquito survival, however, it is unclear how significant these ef-
1962 fects bias estimates of lifespan obtained from MRR experiments. Fortunately, alternative
1963 approaches do exist for determining lifespan, although admittedly, each also requires us to
1964 make assumptions about mosquito ecology. In the next chapter, I consider an approach
1965 which is based on dissection of female mosquito experiments, which allows, for some species
1966 in the dataset we collect, to compare estimates from this approach with those from MRR
1967 experiments, allowing a critical appraisal of each method.

1968 Whilst MRR experiments provide imperfect measures of mosquito lifespan, it is worth
1969 noting that they are also often used to simultaneously determine a range of other bio-
1970 nomic parameters, such as dispersal distance and population densities (see Silver, 2007).
1971 Furthermore, many of these parameters cannot be determined as directly by other ap-
1972 proaches. This means that the MRR experiment, whilst costly and variable in results, is
1973 likely to remain a mainstay approach in a field entomologist's toolbox.

Chapter 3

A meta-analysis of mosquito dissection experiments

3.1 Abstract

An alternative method to MRR experiments for estimating lifespan of wild mosquitoes is by capturing samples of mosquitoes and dissecting the specimens. Female mosquitoes undergo predictable changes in their reproductive anatomy throughout their lives, providing a biological clock which can be used to estimate the physiological age of individual specimens when dissected. In this chapter, I describe how a dataset of female mosquito dissection experiments was compiled by a literature survey, then apply a Bayesian hierarchical model (of similar nature to that used in Chapter 2) to estimate the physiological lifespan of mosquitoes by species and genus. This analysis indicated that there were significant differences in the lifespan of mosquitoes between species and genera. To convert physiological age - a measure of the number of gonotrophic cycles a female mosquito has undergone - to calendar age, requires estimates of the length of the gonotrophic cycle. In

1989 this chapter, I describe how a further search of the literature was undertaken to compile
 1990 a database of estimates of gonotrophic cycle duration. I then apply a statistical approach
 1991 to this data to produce estimates of gonotrophic cycle duration. These estimates are then
 1992 used to convert the physiological lifespan estimates into lifespan in terms of calendar time,
 1993 which show good correspondence with the estimates from the MRR studies. As for the
 1994 MRR analysis, I also fitted six different survival models to the data, each of which embod-
 1995 ies different assumptions about age-dependent mortality. Unlike the MRR analysis, there
 1996 was a lack of consensus as to whether age-dependent mortality affects wild populations:
 1997 across 5 species, the exponential distribution, which assumes no increase in mortality with
 1998 age, fit the data best; in 14 others, one of the five models incorporating age-dependent
 1999 mortality improved the fit over the exponential; in the remaining 6 species, both classes
 2000 of model provided an equally good fit. Overall, I interpret these results as indicating that
 2001 the importance of age-dependent mortality varies according to a range of factors, including
 2002 climate, geography, and genetics, and is likely a subtle effect.

2003 3.2 Introduction

2004 Dissection-based methods to determine the physiological age of adult female mosquitoes
 2005 are the principal alternative to MRR for estimating survival and longevity (Detinova,
 2006 1962; Dye, 1992). Perhaps the simplest such method is to sample the parous rate in a
 2007 population (the proportion of adult females that have laid eggs) by separating parous from
 2008 nulliparous females (Detinova, 1962). Parous rate gives an estimate of adult female survival
 2009 by the simple formula $\text{daily survival} = \sqrt[T]{\text{parous rate}}$ where T is the duration of the first
 2010 gonotrophic cycle (Davidson, 1954). An alternative approach to determine the lifespan of
 2011 wild mosquitoes is to use changes in the insects' morphology that occur at regular intervals
 2012 throughout life to estimate the physiological age of specimens. The most common approach

to estimate physiological age was introduced by Polovodova, 1949, and involves counting the number of dilations in female mosquito ovarioles. In this method the maximum number of dilations found in the ovarioles is said to represent the parity status of an individual, defined as the number of times a female has laid eggs. Egg-laying typically occurs once during each gonotrophic cycle - the process of seeking a host to allow blood-feeding, followed by digestion of the blood and maturation of eggs, and finally oviposition in a suitable site (Beklemishev, 1940). If estimates exist for the duration of each gonotrophic cycle, then parity status of an individual mosquito can be converted into an estimate of its calendar age. By collecting a sufficient number of wild mosquitoes and estimating their age it is hoped that a survival curve can be reconstructed from which we can estimate population quantities of interest, for example, the mean lifespan.

The female reproductive system contains two paired ovaries, each of which comprises a number of ovarioles which attach to a central tube, the calyx, which is continuous with the oviduct (Figure 3.1) (Fox and Brust, 1994). Each ovariole consists of a growth zone or germarium, and typically two follicles in gonoactive females. A follicle contains one oocyte and seven nurse cells, although in follicles recently emerged from the germarium these two cell types are phenotypically indistinguishable (Figure 3.2) (Detinova, 1962). The developing follicle goes through a range of stages, first described by Christophers, 1911¹, who classified these as stages I-V.

¹Sir Rickard Christophers was a giant amongst mosquito people, who worked until well after his one hundredth birthday. As well as his significant contribution to many aspects of mosquito biology, he is known for writing a letter to the *American Journal of Tropical Medicine and Hygiene* objecting to their publication of his obituary (Spielman and D'Antonio, 2001).

Chapter glossary

- **Gonotrophic cycle** - the time interval between two consecutive blood meals for female mosquitoes, that encompasses (i) the search for a host and blood-feeding, (ii) the digestion of blood and egg maturation, and (iii) the search for a suitable oviposition site and oviposition (Lardeux et al., 2008).
- **Physiological age** - the number of gonotrophic cycles a female mosquito has undergone.
- **Nulliparous** - describes female mosquitoes that have yet to complete a gonotrophic cycle.
- **Parous** - describes female mosquitoes that have completed at least one gonotrophic cycle.
- **Ovariole** - One of the tubes which composes a ovary (of which mosquitoes have two) which connects to an oviduct, where immature oocytes are deposited. Ovarioles are composed of a germarium and a series of follicles.

- I - One oocyte in the distal portion of follicle (the end nearest the calyx) lies below seven nurse cells.
- II - Yolk granules appear in the protoplasm of the egg, which itself grows to become significantly bigger than the nurse cells. At the end of this stage the egg accounts for roughly half the space of the follicle.
- III - The share of the follicle taken up by the egg increases to roughly three-quarters the space, and the follicle elongates. The nucleus can no longer be seen through the dense yolk.

- IV - The follicle becomes longer still and the nurse cells only occupy the uppermost tenth of the whole, with the rest of the follicle consisting of yolk.
- V - The chorion (the outermost membrane) surrounds the entire egg, and the floats appear on each side of the egg, ready to be laid.

Whilst there still exists some controversy regarding the details of ovulation (see Section 3.5), we now describe the process as it was originally summarised by Detinova, 1962, since this motivates the use of this method as a way to age wild female mosquitoes. During ovulation contraction of the ovarian muscle causes the egg to move towards the distal end of the ovariole (Figure 3.3). Here it passes through the pedicle into the lateral oviduct, where it forms a queue with the other eggs in the common oviduct, ready to be laid. After the eggs leave the ovariole a large empty sack remains at the point where the egg developed. Over time the intima of the sack begins to contract, and after about a day there remains a dilation of similar size to a follicle that has just left the germarium. Above the dilation sits a follicle that is ready to develop into the second egg during the next gonotrophic cycle. The whole process is then repeated again during ovulation, with the result that after two ovulations there now exist two distinct dilations in the ovariole. After each subsequent ovulation another dilation in the ovarioles forms. Hence by dissecting the ovaries and counting the number of such dilations it is possible to determine the number of gonotrophic cycles that a mosquito has undergone.

In this chapter, we describe the collection and analysis of data from a number of experiments in which wild female mosquito populations are sampled and dissected to estimate the parity status of each individual (the number of gonotrophic cycles undergone). By using a hierarchical Bayesian model where individual experiments from within a species or genus are grouped together this allows estimation of mean mosquito lifespan at these levels. These lifespan estimates are in terms of numbers of gonotrophic cycles. To convert

2066 these into calendar ages we require estimates of the duration of each cycle. Accordingly in
2067 Section 3.3.3 we describe the collection and analysis of data for studies that estimate the
2068 duration of gonotrophic cycles, which is used to convert mean lifespan estimates in terms
2069 of gonotrophic cycles to those in terms of calendar age.

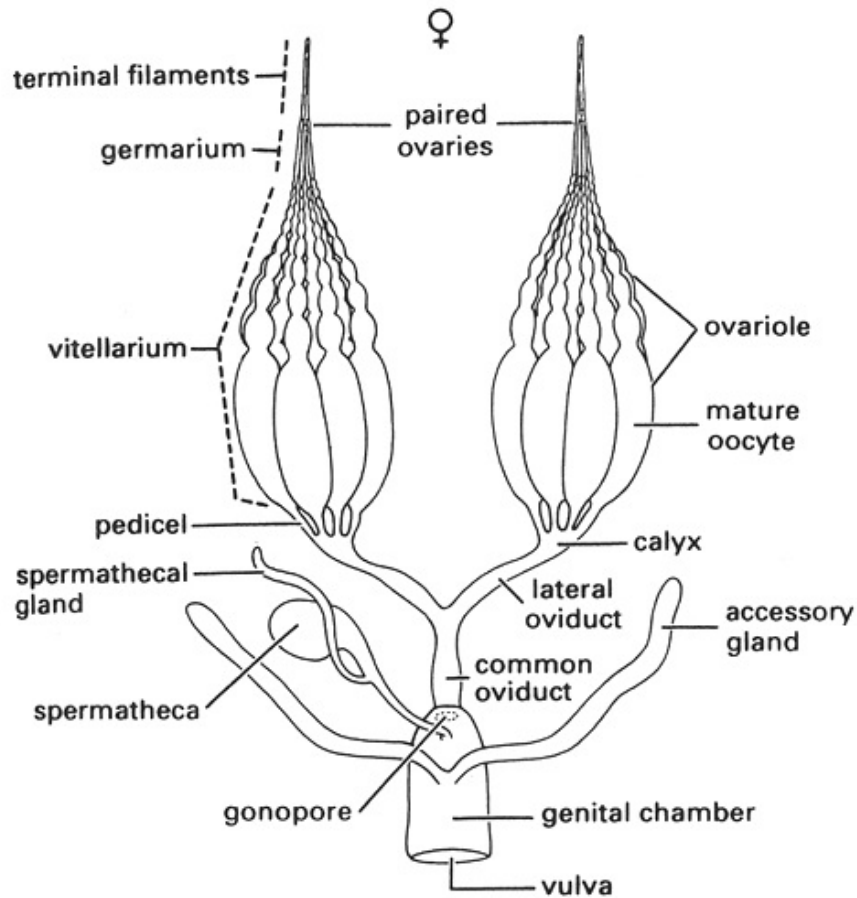


Figure 3.1: **The female mosquito reproductive system.** Figure reproduced from University of Sydney, 2004.

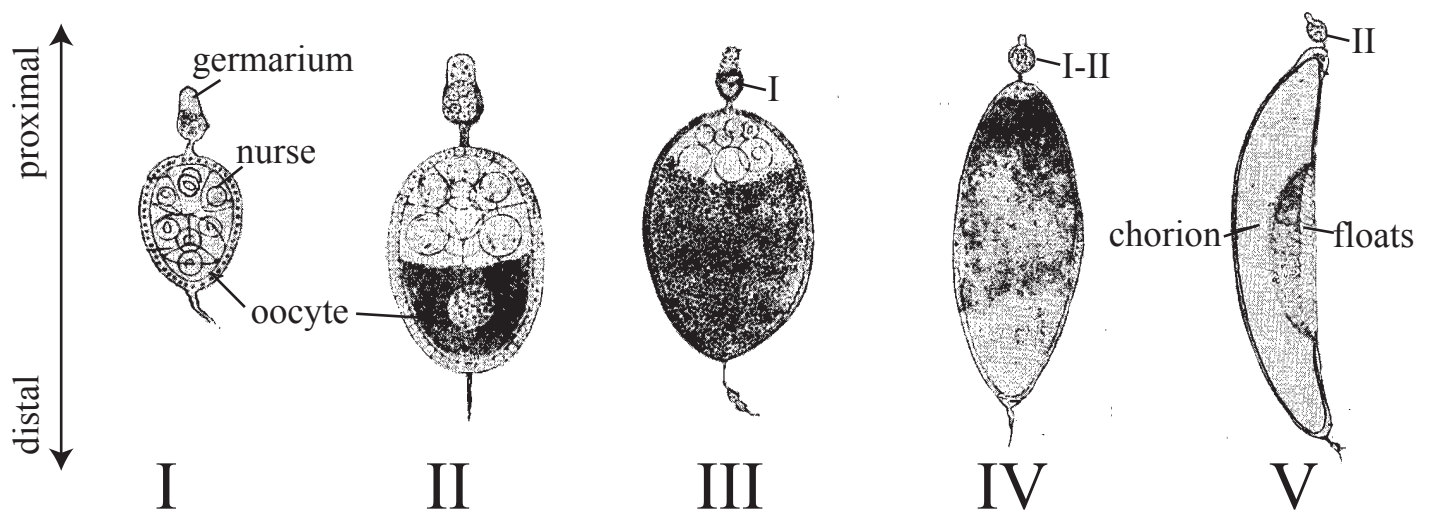


Figure 3.2: **The Christophers stages of ovariole development.** The distal portion connects with the lateral oviduct, through which the egg passes to be laid. Figure adapted from Detinova, 1962.

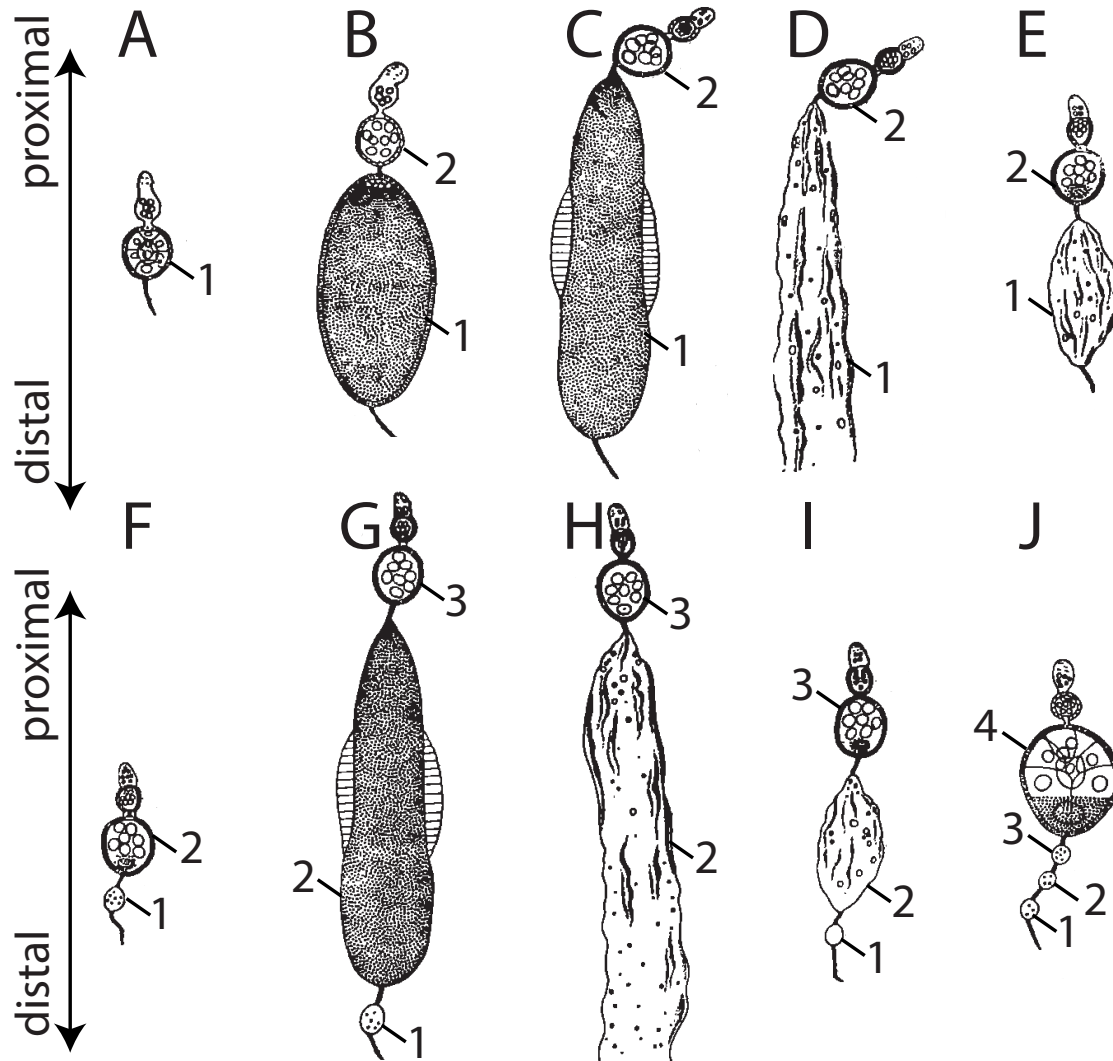


Figure 3.3: **The formation of dilations in ovariole development (A-J temporally).** The numbers indicate the order in which the ovarioles (and the resultant dilations) develop. Figure adapted from Detinova, 1962.

3.3 Method

3.3.1 Collection of dissection data

A comprehensive search of the literature using Google Scholar (scholar.google.co.uk) was performed using various combinations of the following keywords: dissection, mosquito, parity, parous, age, and physiological age. No constraints were placed on publication date, location or type. This list was supplemented with a number of author-specific searches for those individuals most prevalent in the literature. In particular we searched for all articles authored by: Charlwood, Muller, Schlein, Samarawickrema, Reisen, Detinova, Polodova, Gillies, and Wilkes. An additional list of potential articles was provided by doing a forward article search on some of the most highly-cited articles in the database: Polovodova, 1949; Detinova, 1962; Gillies and Wilkes, 1965; Clements and Paterson, 1981. The list of results was then filtered manually by examining the titles and abstracts to produce a candidate list of the articles most likely to contain data on the physiological age of dissected mosquitoes caught in the wild, as determined by the Polovodova, 1949 method.

A relational database was used to store the raw data from the actual experiments, along with the meta-data associated with each of the experiments. In many of the published articles wild mosquitoes were caught and dissected over a period of time, and the raw data thus consists of snapshots of the age structure of the population at regular intervals in time. In the cases where the data was more sparse (fewer than ten individuals, on average, per date), we aggregated across dates and recorded this as a single entry; otherwise we recorded the snapshots of the population at each different date. We record separate series for each species that was captured, or for those that were recorded at separate capture locations (potentially with an alternative collection method), and do not aggregate over these datasets.

For each individual series we recorded the following meta-data: genus, species, collection method, whether or not insecticide was mentioned as being used in or around the time of mosquito collections, and the start and end dates of the experiment. At the article level we recorded the following meta-data: author, year, title, country, location (within country), start and end date, whether insecticide was used during any of the experimental replicates, whether a MRR experiment occurred alongside the dissection study, and the collection location (indoor or outdoor). At either the individual series or article levels we record additional meta-data describing the nature of data collection, for example explaining where the data was contained within the article, whether it was obtained by digitising graphs, and the number of separate dated series. For those few cases ($n = 16$ series) where the data was obtained by digitising graphs, we used the WebPlotDigitizer online tool (Rohatgi, 2017).

The data collection method resulted in 568 separate dissection series, across 72 published articles (see Section 7.3 for a list of the studies included in our final dataset). The published datasets cover the period from 1960-2015 with comparable numbers of studies across each decade (Figure 3.4). The statistical approach applied to the data relies on the assumption that there is a constant rate of recruitment into the adult population (see Section 3.3.2). If populations fluctuate strongly from month-to-month this assumption will likely be violated. To try to mitigate against such a risk, we aggregate the data across all dates to obtain a single series for each identifier, resulting in 201 such series. To obtain a reasonable level of accuracy on estimates we remove all those individual (aggregated) series where there are fewer than 100 mosquitoes in total in the series. Finally we remove data for any species where there was only one series resulting in 131 series across four genera (*Anopheles*, *Culex*, *Aedes*, and *Mansonia*) and 25 species (Table 3.1). These studies are distributed across a wide range of geographies (Figure 3.5).

Genus	Species	Frequency
<i>Anopheles</i>	<i>gambiae s.l.</i>	19
<i>Culex</i>	<i>quinquefasciatus</i>	12
<i>Anopheles</i>	<i>maculipennis</i>	11
<i>Anopheles</i>	<i>farauti s.l.</i>	10
<i>Anopheles</i>	<i>sergentii</i>	8
<i>Aedes</i>	<i>polynesiensis</i>	8
<i>Culex</i>	<i>pipiens</i>	6
<i>Anopheles</i>	<i>culicifacies</i>	5
<i>Anopheles</i>	<i>darlingi</i>	5
<i>Anopheles</i>	<i>quadrimalatus</i>	5
<i>Anopheles</i>	<i>stephensi</i>	4
<i>Anopheles</i>	<i>melas</i>	4
<i>Culex</i>	<i>annulirostris</i>	4
<i>Aedes</i>	<i>aegypti</i>	3
<i>Aedes</i>	<i>samoanus</i>	3
<i>Anopheles</i>	<i>minimus</i>	3
<i>Anopheles</i>	<i>rivulorum</i>	3
<i>Culex</i>	<i>thalassius</i>	3
<i>Mansonia</i>	<i>uniformis</i>	3
<i>Anopheles</i>	<i>subpictus</i>	2
<i>Aedes</i>	<i>sollicitans</i>	2
<i>Aedes</i>	<i>vexans</i>	2
<i>Anopheles</i>	<i>bellator</i>	2
<i>Anopheles</i>	<i>cruzi</i>	2
<i>Culex</i>	<i>tritaeniorhynchus</i>	2
<i>Anopheles</i>		83
<i>Culex</i>		27
<i>Aedes</i>		18
<i>Mansonia</i>		3
Total		131

Table 3.1: The numbers of dissection series for each species or genus included in the overall dataset.

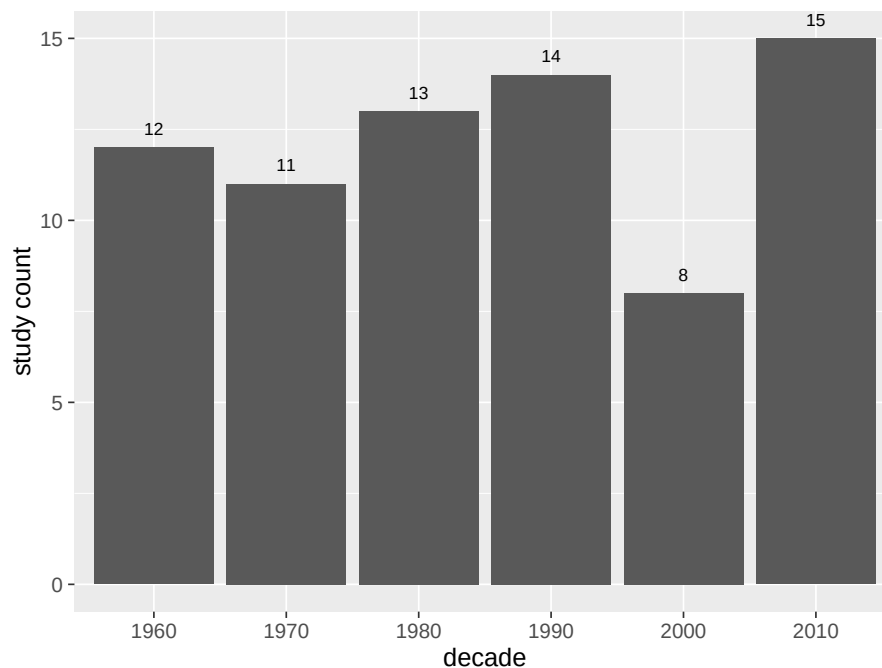


Figure 3.4: The numbers of published studies that estimate gonotrophic age by dissection across the study period.

2119 3.3.2 Statistical analysis of dissection data

We suppose that the number of individuals recruited into the adult female population is constant over time, meaning that the age-structure of the population is stable. Whilst environmental heterogeneity will naturally lead to variation in adult recruitment over time, we hope that by aggregating data across studies undertaken over a range of dates we may reduce the impact of this effect on our results. Suppose that the probability an individual female mosquito survives until age a is given by the survival function $S(a)$, then the number of individuals in the population surviving to this age is given by,

$$A(a) = A(0)S(a), \quad (3.1)$$

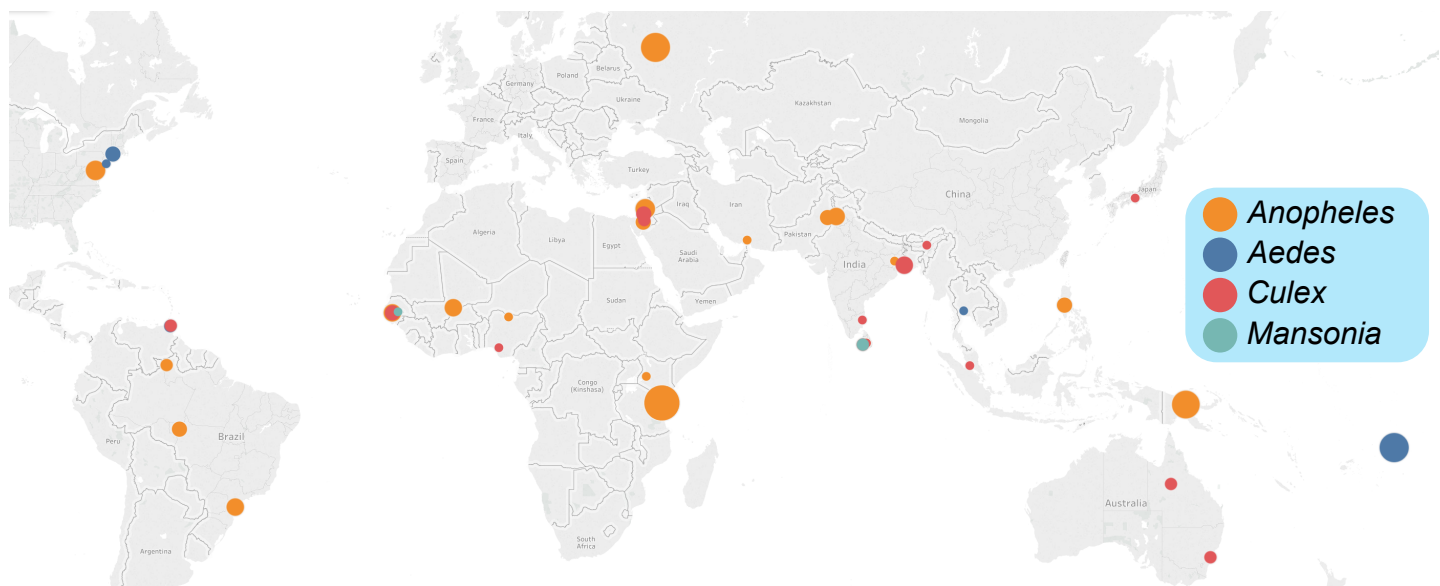


Figure 3.5: **The geographic location of each of the dissection databases included in the meta-analysis.** The area of the bubbles indicates the number of unique data series available.

where $A(0)$ is the number of female mosquitoes recruited into the population per unit time. Consider one individual experiment where we randomly sample from a wild population that is structured as per eq. (3.1). In this case the number of individuals sampled at each age is binomially-distributed,

$$X(a) \sim \mathcal{B}(A(a), p(a)), \quad (3.2)$$

where $p(a)$ is the probability of recapturing a single mosquito of age a . In what follows we typically assume that the probability of capturing a given mosquito is independent of their age, so that $p(a) = p = \text{const}$. Since in general $A(a)$ is large and $p(a)$ is small, we

Variable	Min	Median	Mean	Max	Standard deviation
Min age of captures, gonotrophic cycles	0	0	0.02	1.00	0.12
Median age of captures, gonotrophic cycles	0	1.00	0.63	2.00	0.62
Mean age of captures, gonotrophic cycles	0.17	0.90	1.00	3.03	0.58
Max age of captures, gonotrophic cycles	1.00	5.00	5.77	13.00	2.94
Number captured	100	565	1,317	14,012	1,885

Table 3.2: **Summaries of the characteristics of the 131 individual dissection data series that are included in this analysis.** Note when series are censored (see Section 3.3.2) we assume that all mosquitoes dissected over the threshold are of this age when calculating summary statistics. **Add number of censored series.**

can approximate the above using a poisson distribution,

$$X(a) \sim \text{poisson}(A(a)p). \quad (3.3)$$

However we believe that the assumption of *independent* captures of individual mosquitoes, which underlies the binomial and poisson models, is likely suspect for the same reasons as for the MRR analysis (see Section 2.3.2). As before we choose to specify a negative binomial sampling distribution that allows for non-independent captures,

$$X(a) \sim \text{NB}(A(a)p, \kappa), \quad (3.4)$$

2120 where we use the parameterisation such that the mean is given by $A(a)p$, and the over-
2121 dispersion parameter, κ , where as $\kappa \rightarrow \infty$ the above sampling distribution approaches a
2122 poisson.

In the field, unfortunately, we do not know the number of mosquitoes recruited into the adult population, $A(a)$, nor the probability of capturing an individual at a given point in time, p . Instead we model their product $\Psi = A(0)p$ (the population of adult female

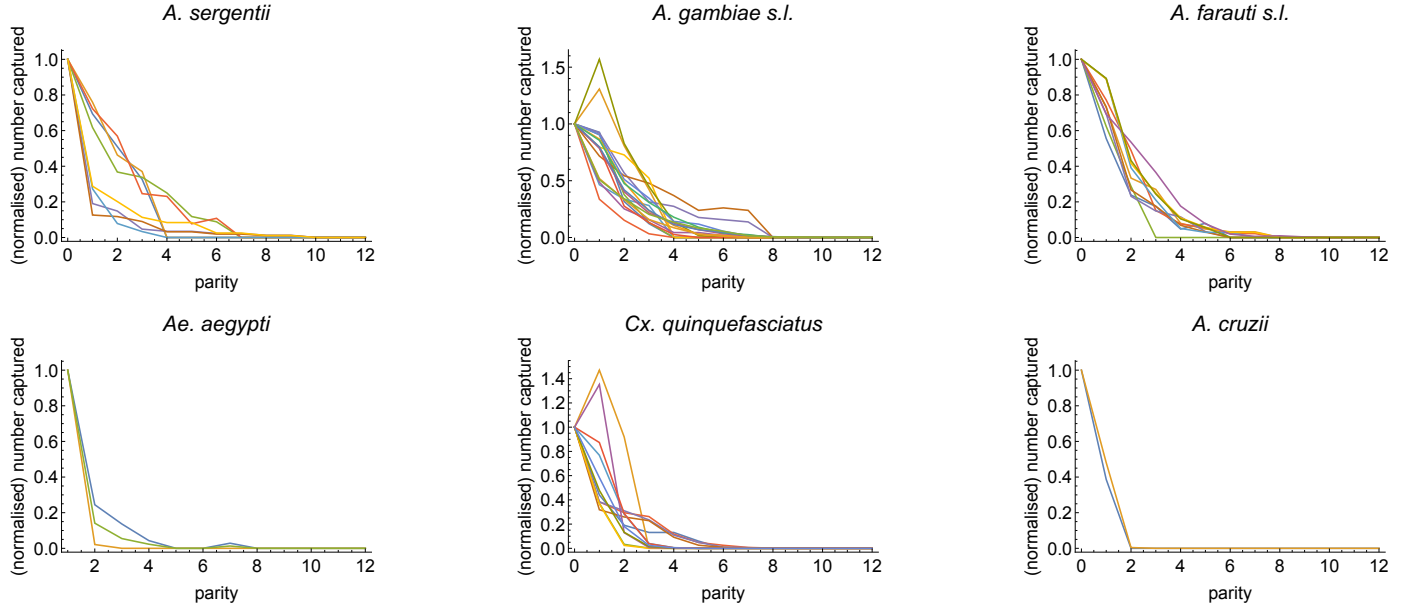


Figure 3.6: **The normalised physiological age series for six species in the database.** Each different coloured line represents an individual series. In each case the count for all ages has been normalised by the nulliparous (or for *Ae. aegypti* uniparous) count. For *Ae. aegypti* we start the series at uniparous individuals since there is a deficiency of nullipars for all series for this species (see Section 3.3.2). In all cases we do not include any data for censored observations (see Section 3.3.2).

mosquitoes of age zero that can be captured) probabilistically resulting in a model,

$$X(a) \sim \text{NB}(\Psi S(a), \kappa). \quad (3.5)$$

2123 The resultant likelihood of a data series consisting of counts: $(y(a_1), y(a_2), \dots, y(a_R))$, is
 2124 then calculated by assuming (conditional) independence of the observations,

$$\mathcal{L}(y(a_1), y(a_2), \dots, y(a_R) | S(\cdot), \Psi, \kappa) = \prod_{a=a_1}^{a_R} p(y(a) | S(a), \Psi, \kappa), \quad (3.6)$$

where $p(y(a) | S(a), \Psi, \kappa)$ corresponds to the negative binomial probability mass function

for a count of $y(a)$ mosquitoes aged a as specified in eqn. (3.5), and R is the number of separate physiological age classes in which the count was non-zero. Since we do not know Ψ we must learn it from the data. One approach to estimate this parameter could be use the number of captures of nulliparous mosquitoes, $X(0)$ in place of Ψ . We prefer to allow for some uncertainty in this parameter and use the data to estimate its value. However, unfortunately, the negative binomial likelihood allows too much variation in this parameter (because the data are over-dispersed), and instead we specify a likelihood of the form,

$$X(0) \sim \mathcal{N}(\Psi, \sqrt{\Psi}), \quad (3.7)$$

solely for the first data point $X(0)$. The above allows for some freedom in Ψ whilst ensuring that the parameter's probability mass lies near enough to $X(0)$ to allow useful model estimates. Ψ is set a uniform prior over the range $[0, \frac{3}{2}X(0)]$.

Since $S(a)$ is monotonically-decreasing we know that, if recruitment to the adult population is constant, then the numbers of nulliparous individuals should exceed the count in subsequent parity states. However in a number of data series there is a relative dearth of nulliparous mosquitoes versus uniparous individuals. This deficiency has been noted in a previously-published study where the authors hypothesise that it is due to the issue of sampling the nulliparous population, since they are more likely to rest outside, compared with parous individuals (Gillies and Wilkes, 1965). However there is evidence to suggest that, on average, the first gonotrophic cycle is longer than subsequent cycles (see Section 3.3.3), meaning that a relative surplus of nulliparous mosquitoes may exist in captured samples (Clements and Paterson, 1981).

In our analysis, we chose to remove the counts of nulliparous individuals from the series where the count was too low relative to uniparous or higher parity individuals since this may indicate issues with sampling the nulliparous population. Specifically we stipulated

2141 that the nulliparous count should exceed 90% of the count for the uniparous individuals.
 2142 For those cases where this condition is not met we remove the nulliparous observation
 2143 and analyse the series of counts for all subsequent pars (uniparous and subsequent parous
 2144 states).

2145 Here we estimate our model with one of six different survival functions (the same as for
 2146 the MRR case; Table 2.3), each of which makes different assumptions regarding how the
 2147 force of mortality is affected by mosquito age. To estimate lifespan at the species, genus
 2148 and overall groupings we then use a hierarchical Bayesian model of the same mathematical
 2149 form as for the analysis of MRR experiments (see Section 2.3.4). However the priors for the
 2150 analysis of dissection data were modified so that they represented a mean prior lifespan of
 2151 three gonotrophic cycles, although allowed considerable variation in mean lifespan (Figure
 2152 3.7). The priors used for the over-dispersion parameter (κ) for the negative binomial
 2153 likelihood were the same as for the MRR analysis.

2154 To determine whether mosquitoes experience age-dependent mortality we compared the
 2155 predictive power of each of the models that incorporate a hazard function that depends on
 2156 age with that from the exponential model. As for the MRR analysis we also used K-Fold
 2157 cross-validation to perform this comparison, where the data are randomly partitioned into
 2158 training and test sets (see Section 2.3.7). The model is then fitted to each training set
 2159 and used to predict the data in the independent test set. Since there are fewer series than
 2160 for the MRR dataset we used two partitions for each species, where each partition was of
 2161 roughly the same number of individual series.

2162 The estimates of mean mosquito lifespan that we present here (Figure 3.12) resulted
 2163 from the use of the exponential survival model (i.e. no age dependence). This choice was
 2164 made because we found limited evidence in support of age-dependent mortality (see Section
 2165 3.4.4). However, since the priors for each of the survival functions were specified to allow

Survival function	Physiological age series-level priors	Group-level priors
Exponential	$\lambda \sim \text{log-normal}(\mu_\lambda, \sigma)$	$\mu_\lambda \sim N(-1.2, 1), \sigma \sim \text{log-normal}(-2, 1)$
Gompertz	$\alpha, \beta \sim \text{log-normal}(\mu_{\alpha \beta}, 0.2)$	$\mu_\alpha \sim N(-1.5, 1), \mu_\beta \sim N(-2.5, 0.5)$
Weibull	$\alpha, (\beta - 1) \sim \text{log-normal}(\mu_{\alpha \beta}, 0.2)$	$\mu_\alpha \sim N(-2.5, 1), \mu_\beta \sim N(-4, 0.5)$
Gompertz-Makeham	$\alpha, \beta, c \sim \text{log-normal}(\mu_{\alpha \beta c}, 0.2)$	$\mu_\alpha \sim N(-1.4, 1), \mu_\beta \sim N(-3, 0.4), \mu_c \sim N(-4.5, 0.5)$
Logistic	$\alpha, \beta, s \sim \text{log-normal}(\mu_{\alpha \beta s}, 0.2)$	$\mu_\alpha \sim N(-1.4, 1), \mu_\beta \sim N(-2, 1), \mu_s \sim N(-3, 1)$
Logistic-Makeham	$\alpha, \beta, s, c \sim \text{log-normal}(\mu_{\alpha \beta s c}, 0.2)$	$\mu_\alpha \sim N(-3, 1), \mu_\beta \sim N(-1.3, 1), \mu_s \sim N(-2, 0.5), \mu_c \sim N(-4, 0.5)$

Table 3.3: **The priors used on parameters of each different survival model for the dissection data analysis.** For the exponential model the ‘group-level’ priors were the same for the genus and ‘overall’ models that were also estimated. The exponential model was the only model that was simple enough to allow the scale parameter of the log-normal (σ) to be estimated by the data. The notation $\alpha, \beta \sim \text{log-normal}(\mu_{\alpha|\beta}, 0.2)$ means that α and β were assigned independent log-normal priors with location parameters μ_α and μ_β respectively, and a scale parameter of 0.2 in both cases.

for a wide variety of lifespans (Figure 3.7) the specific choice of survival function made little difference to resultant estimates (data not shown).

To demonstrate that the results we obtain are not sensitive to the particular form of hierarchical prior structure chosen we also provide estimates of the lifespan for each series analysed on its own. To do so we assume priors on the rate parameter of the exponential distribution that provides support over a wide range of possible lifespans (Figure 3.8A). We also specified a prior on the over-dispersion parameter κ that was comparable to the hierarchical case (Figure 3.8B).

As for the MRR analysis we use *Stan* that implements a form of MCMC algorithm known as NUTS (Hoffman and Gelman, 2014) to sample from the posterior. The protocol followed for the MCMC sampling was the same as described in Section 2.3.6.

2177 **Analysis of censored data series**

2178 In some of the published papers the data are censored above some threshold age. In these
 2179 cases we only know the number, $y(a_T)$, of individuals who were captured and dissected with
 2180 an estimated physiological age that is greater than or equal to some threshold, a_T . Since
 2181 we do not know the estimated physiological age of individual specimens i , we represent
 2182 this as a parameter, $\aleph_i$, in our statistical model. This means that the joint distribution
 2183 is a function of $y(a_T)$ different $\aleph_i$ parameters. Since these parameters are not directly
 2184 of interest, and our chosen MCMC engine, *Stan* (Stan Development Team, 2014), does
 2185 not directly allow discrete parameters in models, we marginalise these out of the joint
 2186 distribution. To do this exactly would require an infinite sum over all the possible ages for
 2187 all $y(a_T)$ mosquitoes that have been caught whose age exceeds the threshold. Rather than
 2188 carry out this intractable number of summations we instead make the approximation that
 2189 all mosquitoes in this group are of the same age $\aleph_i = \aleph$, $\forall i \in (1, \dots, y(a_T))$. We believe this
 2190 assumption is justifiable, particularly since the numbers of mosquitoes in each subsequent
 2191 age category is a strongly-decreasing function of age. This means that whilst we do not
 2192 know with certainty individuals' ages, it is likely that many will be of the threshold age
 2193 a_T .

This approximation means that to marginalise the parameter $\aleph$ out of the joint distribution, we are only required to do a single summation. Specifically if we have $y(a_T)$ captured individuals of age equal to, or exceeding, some threshold age, a_T , the probability of these observations is given by,

$$q(y(a_T)|S(\cdot), \Psi, \kappa) = \sum_{\aleph=a_T}^{\infty} p(y(a_T), \aleph|S(\aleph), \Psi, \kappa) \quad (3.8)$$

$$= \sum_{\aleph=a_T}^{\infty} p(y(a_T)|S(\aleph), \Psi, \kappa) \times p(\aleph) \quad (3.9)$$

where $q(y(a_T)|S(\aleph), \Psi, \kappa)$ is the probability of observing $y(a_T)$ counts for mosquitoes of an age $\aleph$. In practice, since we do not believe mosquitoes live for longer than 20 gonotrophic cycles (the maximum observed in the data was 13), we cut-off the summation at this point, and assume that the discrete prior probability distribution $p(\aleph)$ is uniform over this range. The overall likelihood for the cases where the series are censored therefore has the form,

$$\mathcal{L}(y(a_1), y(a_2), \dots, y(a_{T-1}), y(a_T)|S(\cdot), \Psi, \kappa) = \left(\prod_{a=a_1}^{a_T-1} p(y(a)|S(a), \Psi, \kappa) \right) q(y(a_T)|S(\aleph), \Psi, \kappa). \quad (3.10)$$

In practice, the number of mosquitoes captured in those series that are censored represents a small percentage of total captures (the median is approximately 2%), and so the effect of the $q(\cdot)$ term above is likely minimal on resultant inferences.

3.3.3 Data collection of gonotrophic cycle duration

To convert the estimates of lifespan in physiological age into chronological age we require estimates of the duration of the gonotrophic cycle. To determine this characteristic we conducted a meta-analysis of previously-published studies that estimate the duration of the gonotrophic cycle. A search of the literature using Google Scholar (`scholar.google.co.uk`) was performed using the search term: ‘gonotrophic cycle duration’. The list of articles was then supplemented with a list of references discussed by Silver, 2007. Based on the abstracts of the resultant list of published studies we then decided whether to search each article for estimates of the duration of the gonotrophic cycle. Overall 79 separate estimates of this parameter were found across 42 published articles (see Section 7.4 for a list of the studies included in our dataset). Along with information about the estimates we also recorded study and series meta-data, including the location of the study, the method used for estimation, the species and genus, as well as the temperatures and/or seasons in

which the experiments were carried out. The methods used to estimate gonotrophic cycle duration in the literature largely fell into two distinct categories: those based on MRR studies ($n = 29$); and those based on observations of mosquitoes in a laboratory setting ($n = 42$). In the MRR studies estimates are made of the duration of gonotrophic cycles by dissecting recaptured mosquitoes at each time point using the method of Polovodova, 1949 to count ovariolar dilations. For the laboratory studies the duration of gonotrophic cycles is determined more directly by observing the time taken for mosquitoes to mate then blood-feed and finally oviposit.

Along with point estimates of the parameter we also collected information about the uncertainty in the estimates (if available). In many articles the duration of the gonotrophic cycle was estimated separately for the first versus subsequent cycles, and these estimates were recorded separately. Raw estimates of gonotrophic cycle duration were obtained for species across four different genera (*Aedes*, *Anopheles*, *Culex* and *Masonia*), although there was a bias towards *Anopheles* with $n = 47$ estimates out of the total of 79 (Figure 3.9).

Since we collected data on the method used to estimate the duration of the gonotrophic cycle we show the raw estimates for the two most common approaches: MRR and laboratory (Figure 3.10). From these results it is evident that the estimates from the laboratory studies are, on average, higher than those using the MRR method, (although we wait until Section 3.4.2 for a quantitative comparison).

3.3.4 Statistical analysis of gonotrophic cycle data

As discussed in Section 3.3.3, there was considerable study-level heterogeneity in the information provided for the estimates of the gonotrophic cycle duration. A number of studies ($n = 24$) provided no estimates of uncertainty in gonotrophic cycle duration, whereas the rest gave some indication of confidence or alternatively a range of possible estimates. How-

ever the types of uncertainty intervals that were specified varied considerably from study to study, from the vague (but common) ‘4-6 days’ to the more helpful ‘ 5 ± 1 day (95% confidence interval)’.

The heterogeneous nature of the gonotrophic cycle estimates requires a method that explicitly accounts for this characteristic of the data. We decided to model the estimates as representing quantiles from an underlying normal distribution that represents uncertainty over possible durations of the gonotrophic cycle. In those circumstances where lower, central and upper bounds were given explicitly the data was fairly symmetric and so we believe assuming the observations come from an unskewed normal distribution is not unreasonable. Explicitly we treated each type of uncertainty interval as follows,

- Simple range, for example, ‘4-6 days’: treat the lower and upper bounds as the 2.5% and 97.5% quantiles of a normal distribution.
- Confidence intervals, ‘ 5 ± 1 day (95% confidence interval)’: treat the lower and upper bounds as the relevant quantiles of a normal distribution.
- Point estimates, for example ‘5 days’: treat this as the median of a normal distribution.

The benefit of this approach is that we can convert the quantiles from our normal distribution parameterised by a mean (μ) and standard deviation (σ) to equivalent quantiles from a standard normal,

$$Z_{50} = \frac{G_{50} - \mu}{\sigma}, \quad (3.11)$$

where Z_{50} and G_{50} indicate 50% quantiles from a standard normal distribution and a normal distribution respectively. By manipulating the above expression we obtain the

following,

$$G_{50} = \sigma Z_{50} + \mu, \quad (3.12)$$

2255 which forms a straight line in (Z, G) space with slope σ and y-intercept μ . Therefore by
 2256 plotting the raw observations of (Z, G) quantiles then estimating a linear regression line
 2257 we can characterise the underlying $\mathcal{N}(\mu, \sigma)$ distribution from which we assume they are
 2258 drawn.

2259 Due to the relative unavailability of estimates for gonotrophic cycle duration across
 2260 the different genera we generated a single distribution to represent the uncertainty in this
 2261 parameter by pooling all the data. We believe this is justified because ANOVA tests
 2262 indicated that there was no significant variation in gonotrophic cycle duration across the
 2263 different genera (Figure 3.9).

2264 3.3.5 Conversion of lifespan from physiological to calendar age

2265 The estimates of lifespan produced from analysing the dissection data are in terms of
 2266 physiological age (the number of gonotrophic cycles undertaken). To allow comparison
 2267 with the estimates from the MRR studies (Figure 2.8) as well as produce more useful
 2268 estimates to inform disease transmission dynamical models we convert these estimates to
 2269 calendar days. To do this we use the estimated parameters of the normal densities that
 2270 we have assumed represent uncertainty in gonotrophic cycle duration (see Section 3.3.4)
 2271 and use them to convert our posterior samples of mean physiological lifespan $(L_1^p, \dots, L_S^p)$ to
 2272 lifespan in calendar ages $(L_1^c, \dots, L_S^c)$. To do this we iterate the following for all $i \in (1, \dots, S)$
 2273 posterior samples,

- 2274 1. Sample $G_{1i} \sim \mathcal{N}(\mu_{1i}, \sigma_{1i})$, to obtain a duration for the 1st gonotrophic cycle.

2275 2. Sample $G_{2i} \sim \mathcal{N}(\mu_{2i}, \sigma_{2i})$, to obtain a duration for subsequent gonotrophic cycles.

2276 3. If $L_i^p > 1$, the mean lifespan is longer than one gonotrophic cycle:

2277 then $L_i^c = G_{1i} + G_{2i} \times (L_i^p - 1)$.

2278 4. Else:

2279 then $L_i^c = G_{1i} \times L_i^p$.

2280 In the using the above methodology to convert from physiological to calendar age we
 2281 implicitly assume that for each gonotrophic cycle after the first that the cycles are of the
 2282 same length (we choose one G_{2i} per sample). However since we allow a different G_{2i} for
 2283 each sample we nonetheless believe that the above approach allows for sufficient uncertainty
 2284 in the estimates of subsequent gonotrophic cycle durations. If instead we believed that
 2285 significant variation in gonotrophic cycle occurred within a particular mosquito's life (as
 2286 well as between mosquitoes) then we could draw a new value of G_{2i} for each subsequent
 2287 cycle. This produces results with a little more uncertainty than previously (data not
 2288 shown).

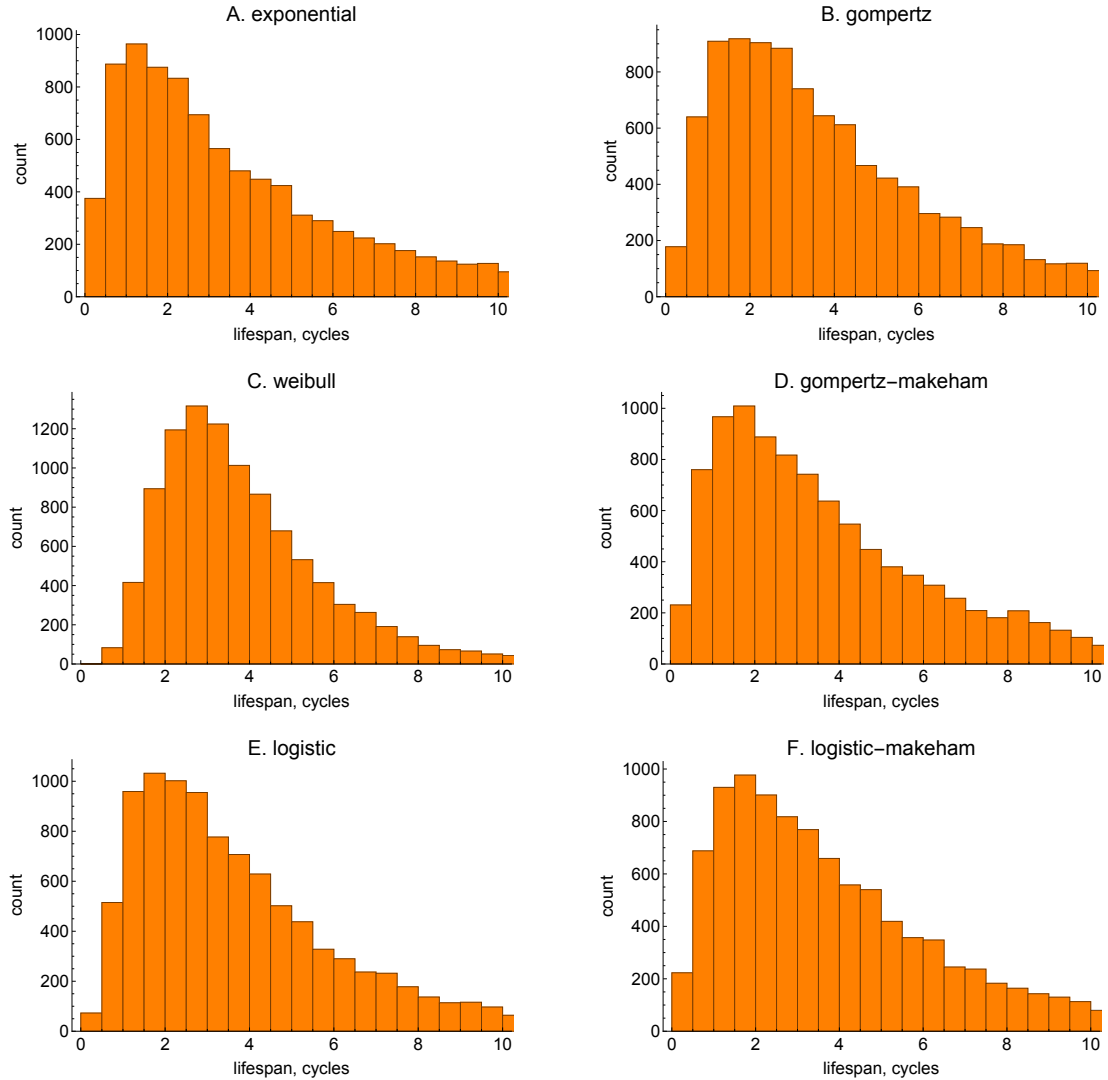


Figure 3.7: **The priors over mean lifespan used to analyse the data for each of the hierarchical models.** The distributional form of the priors is given in Table 3.3. In all cases the plots show data for 10,000 samples from the relevant prior distribution.

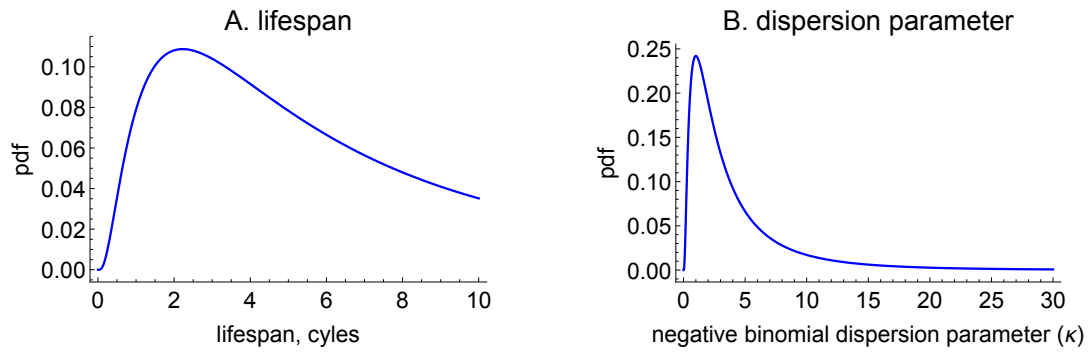


Figure 3.8: **The priors over A. mean lifespan and B. over-dispersion parameter used to analyse the individual dissection series.** The prior on the rate parameter of the exponential distribution was $\lambda \sim \text{log-normal}(-1.8, 1)$. The prior on the dispersion parameter was $\kappa \sim \text{log-normal}(1, 1)$.

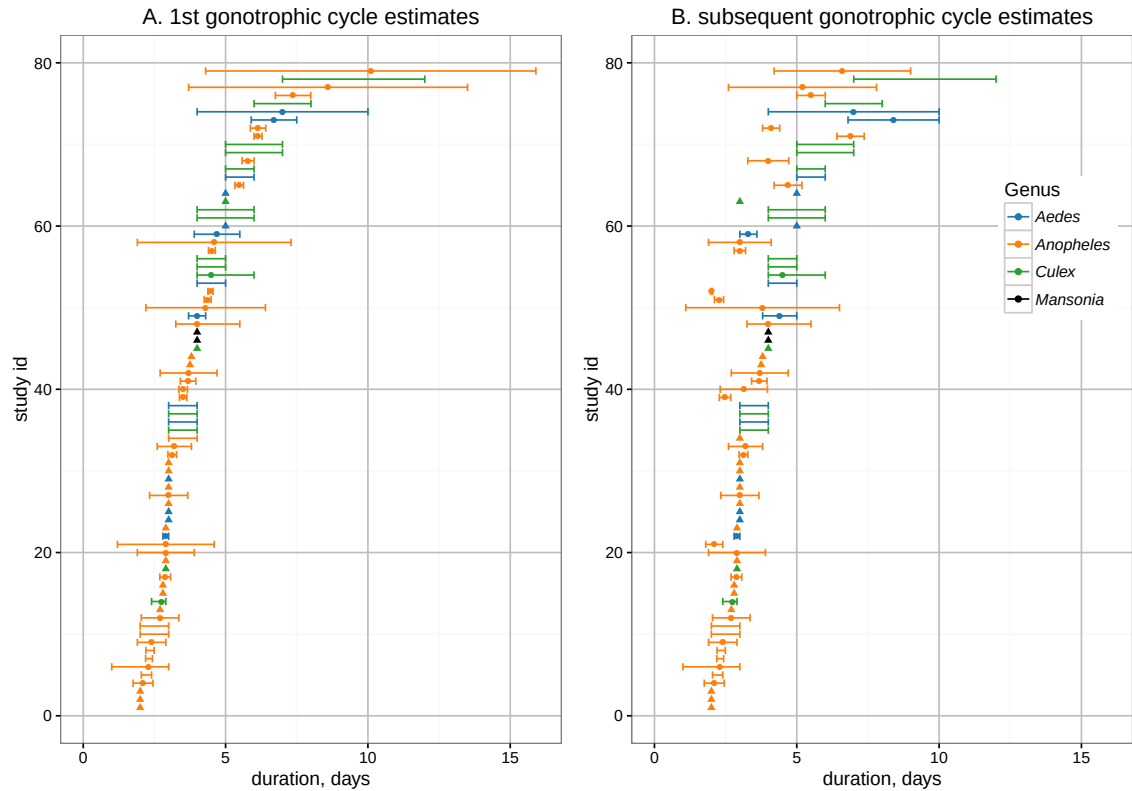


Figure 3.9: **Raw gonotrophic cycle duration estimates for A. the 1st cycle and B. subsequent cycles.** The confidence intervals represent a range of uncertainties since they were not available in a standardised form (see Section 3.3.4 for a discussion of how we handle this issue). If no statement of uncertainty was given in the estimates then we indicate this by a triangle marker. The studies are ordered according to the central estimate of gonotrophic duration.

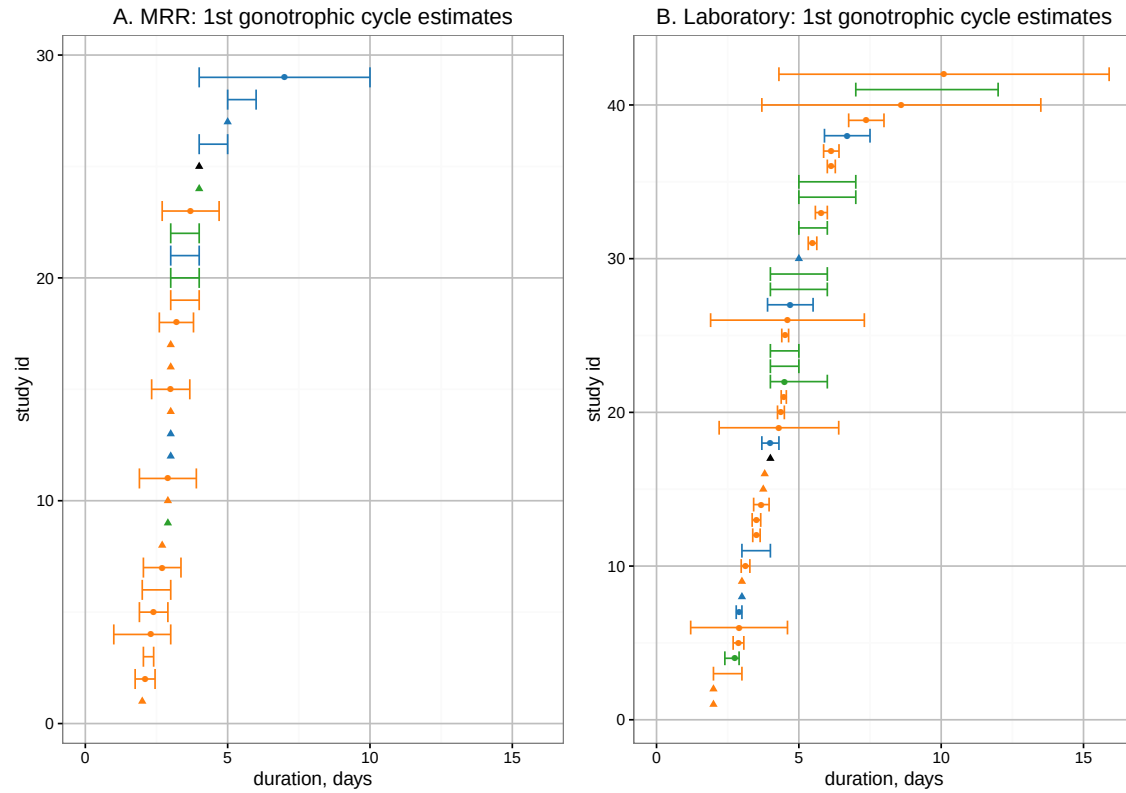


Figure 3.10: **Raw gonotrophic cycle duration estimates from A. the MRR and B. from laboratory-based studies.** The confidence intervals represent heterogeneous measures since they were not available in a standardised form (see Section 3.3.4 for a discussion of we handle this issue). If no statement of uncertainty was given in the estimates then we indicate this by a triangle marker. The studies are ordered according to the central estimate of gonotrophic duration.

3.4 Results

3.4.1 Wild mosquito lifespan in gonotrophic cycles

We first present estimates of the mean mosquito lifespan where we estimate a separate model for each data series (Figure 3.11). Whilst there is considerable variability in these estimates, in 125 of the 131 cases we estimate the median lifespan was below three gonotrophic cycles. However from this separate treatment of individual MRR time-series it is not immediately clear whether differences exist at the species or genus level. To investigate whether such group-level differences exist we estimated a hierarchical model to yield estimates at the species, genus and overall groupings (Figure 3.12). The species with the highest mean lifespan estimate was *A. sergentii*, with a posterior median of 2.5 gonotrophic cycles. This species is known as the ‘oasis vector’ or ‘desert malaria vector’ for its ability to persevere in the harsh conditions throughout the region (Sinka et al., 2010), and so it is perhaps unsurprising that both this analysis concludes that this species endures longer than most other species. At the other end of the scale is *Anopheles bellator* (median posterior estimate 0.5 cycles), which is a major vector of malaria in the Brazil’s Atlantic forests (Forattini et al., 1999; Lorenz et al., 2012). An ANOVA test using the posterior medians from each of the 131 series (from the hierarchical model) concluded that there were significant differences in mean lifespan between the species ($p < 0.01$, $F = 3.53$, $df = 24$). The result of this test remained the same if instead we used the median estimates compared to when we estimated a separate model for each series ($p < 0.01$, $F = 2.24$, $df = 24$). The Kruskal-Wallis test also concluded that this result was robust to non-normal residuals within each species ($p < 0.01$, $\chi^2 = 89.70$, $df = 24$).

The shortest-lived genus was *Aedes*, with a mean lifespan of 0.8 gonotrophic cycles²;

²This was a, somewhat, surprising result, indicating that most aedines did not survive long enough to reproduce and hinting that the effective sample size may be small. This is at odds with population genetic studies which estimate a large effective population size (Lehmann et al., 1998), although these results

followed by *Culex* (posterior median of 1.0 cycles); then *Mansonia* (1.1 cycles); and finally *Anopheles* was found to be longest with a mean lifespan of 1.4 cycles. This result was a different ordering compared with the estimates from the MRR data (Figure 2.8) which estimated that *Culex* was the shortest-lived, followed by *Anopheles* and *Aedes*. However some of the differences in estimates can probably be attributed to the different assumptions that underpin MRR and dissection-based studies (see Section 3.5.1). An ANOVA test using the posterior medians from each of the 131 series (from the hierarchical model) concluded that there were significant differences in mean lifespan between the genera ($p < 0.01$, $F = 5.23$, $df = 3$), which was robust to the non-inclusion of *Mansonia* mosquitoes, for which there were only three separate data series ($p < 0.01$, $F = 7.59$, $df = 2$). The results were again robust to the use of the data from the estimations of separate models for each of the series ($p < 0.05$, $F = 3.41$, $df = 3$). Again the Kruskal-Wallis test concluded that this result was robust to the assumption of normality for data within each species ($p < 0.01$, $\chi^2 = 25.52$, $df = 3$).

3.4.2 Gonotrophic cycle duration

We next present estimates of the duration of the 1st and subsequent gonotrophic cycles which were obtained by pooling data from all genera and across all experimental methods (Figure 3.13). The uncertainty in gonotrophic cycle duration was estimated as being described (by assumption) by $\mathcal{N}(4.3, 0.4)$ and $\mathcal{N}(3.8, 0.4)$ distributions for the 1st and subsequent cycles respectively. Using these distributions we estimate that the 1st gonotrophic cycle is, on average, longer by 0.5 days, although this difference is not statistically significant ($p = 0.16$).

For each of the studies we collected information on the method used to estimate

may be reflective of large migratory flows between local populations rather than a high survival rate for individuals.

gonotrophic cycle duration. This allows us to estimate the duration of the gonotrophic cycle for the two predominant methods: those studies using MRR versus those using laboratory-based approaches (Figure 3.14). We determined that the uncertainty in the 1st gonotrophic cycle duration for MRR studies as a $\mathcal{N}(3.3, 0.2)$ distribution versus a $\mathcal{N}(4.8, 0.5)$ distribution for laboratory-based studies. The difference in estimated cycle length was therefore, on average, 1.5 days, and is statistically-significant ($p < 0.01$). It is not clear why the estimated gonotrophic cycle should be longer for laboratory studies, particularly since mosquitoes in the wild must locate hosts to blood-feed, and sites in which to oviposit.

3.4.3 Wild mosquito lifespan in calendar age

The estimates of the gonotrophic cycle duration across all species and methods (Figure 3.13) were used to convert estimates of lifespan in terms of cycles to that in terms of calendar days (Figure 3.15; see Section 3.3.5). As for the MRR analysis we find that the majority of species have a mean lifespan whose posterior median is less than 10 days (Figure 3.15). However we delay a direct comparison between the MRR and dissection-based estimates until Section 3.5.

3.4.4 The evidence for age-dependent mortality

We estimated models that assumed six different forms for mosquito mortality (see Table 2.3). These included the exponential model that assumes that the risk of mortality is independent of mosquito age, and other forms where we constrained the parameters in estimation so that they only allowed for a mortality that *increases* with age. To investigate evidence for age-dependent mortality we compare the predictive power of each of the five models that incorporated age-dependent mortality with the exponential model. As for the

2358 MRR analysis we used K-Fold cross-validation to do this (see Section 2.3.7), and then
2359 obtained the average log pointwise predictive density for each of the models (Figure 3.16).

2360 In 14 cases we conclude that the best performing model was one that incorporated age-
2361 dependent mortality. However in the 11 other cases we either conclude that the exponential
2362 model performed best ($n = 5$) or neither the exponential nor one of the five other survival
2363 function models performed statistically better ($n = 6$). Overall we take this to mean
2364 that there is mixed evidence in favour of age-dependent mortality, particularly since we
2365 pitted the exponential model against a pool of five other models with age-dependence.
2366 And in many of the cases where the exponential model performed worse than a model with
2367 age-dependent mortality it nonetheless performed better than one or more models with
2368 age-dependence.

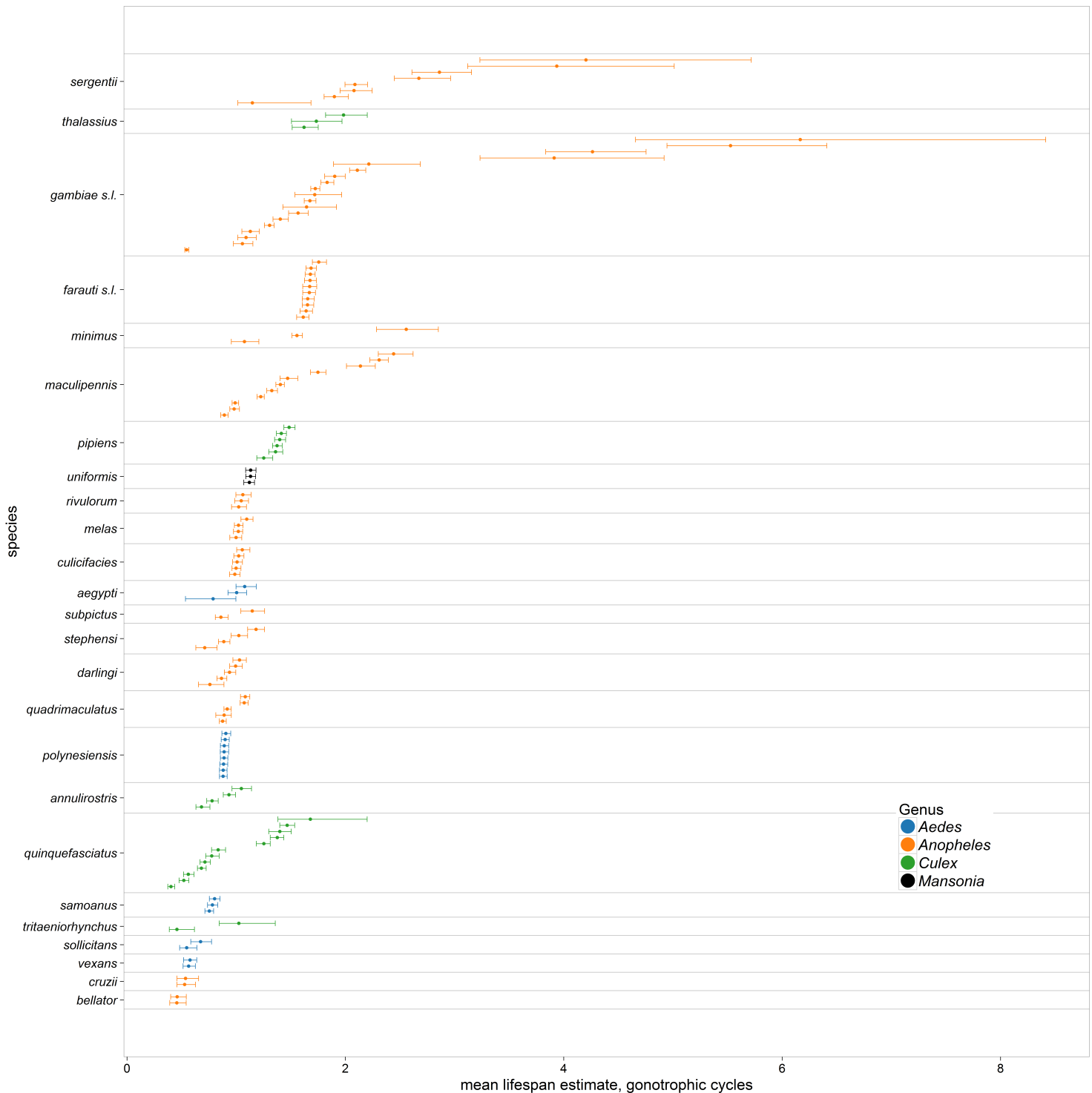


Figure 3.11: Individual time-series estimates of adult mosquito mean lifespan ordered by species median as determined from the dissection data. The middle line in each box shows the median estimates. The left and right box whiskers show the 25%, and 75% posterior quantiles respectively. All estimates were obtained using the non-hierarchical exponential survival model.

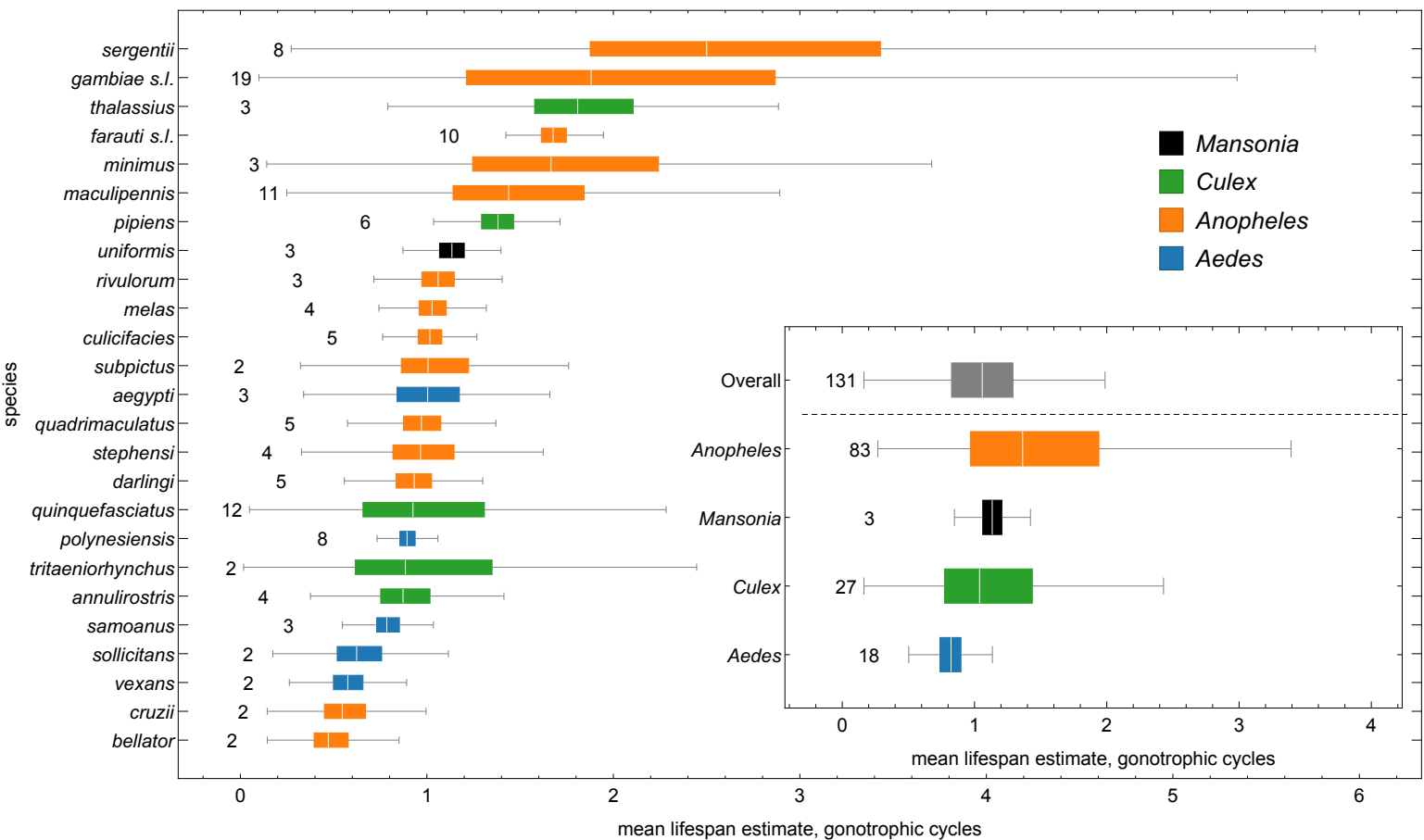


Figure 3.12: **Estimates of mean mosquito lifespan in terms of physiological age at the species, genus and overall levels as determined from the dissection data.** The middle line in each box shows the median estimates. The left and right box edges show the 25%, and 75% posterior quantiles respectively. The whiskers show the range of the data, excluding points lying more than 1.5 times the interquartile range away from each edge of the box. The numbers before the start of the left whisker indicate the number of individual time-series within each species. All estimates were obtained using the exponential survival model.

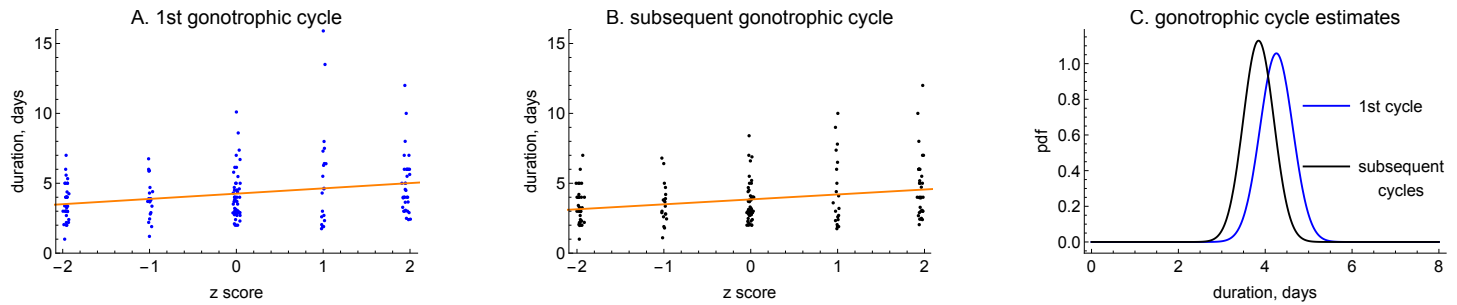


Figure 3.13: The estimation process (A. & B.) and C. the resultant densities for gonotrophic cycle duration for the 1st (blue line) and subsequent cycles (black line). A./B. shows the quantiles of the data for the 1st/subsequent gonotrophic cycle duration versus the equivalent quantiles of a standard normal (z scores). The vertical axis intercepts and slopes from linear regression lines (orange) are then used to estimate the mean and standard deviation of the normal distribution over gonotrophic cycle duration, which are shown in C. Note that jitter in the horizontal axis has been added to plots A. and B.

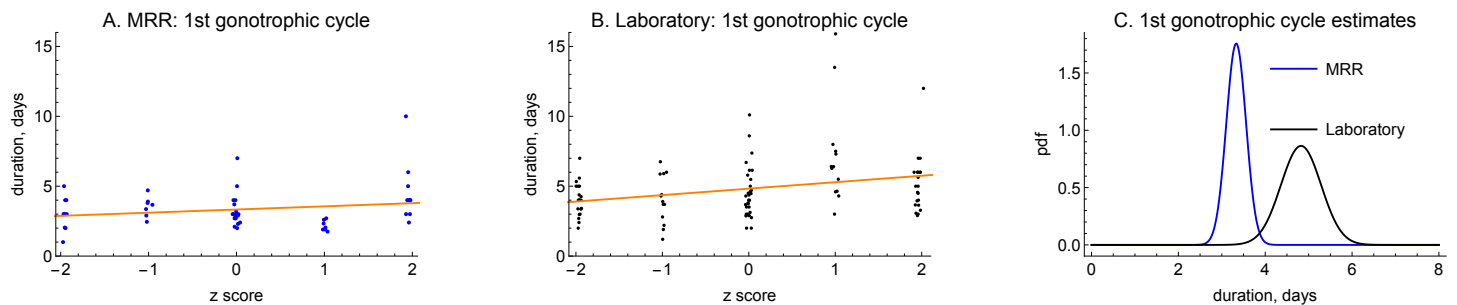


Figure 3.14: The estimation process (A. & B.) and C. the resultant densities for the 1st gonotrophic cycle duration from those studies using a 'MRR' method (blue line) versus those with a 'laboratory' method (black line). A./B. shows the quantiles of the data for the gonotrophic cycle duration from A. the MRR studies and B. the laboratory studies versus the equivalent quantiles of a standard normal (z scores). The vertical axis intercepts and slopes from linear regression lines (orange) are then used to estimate the mean and standard deviation of the normal distribution over gonotrophic cycle duration, which are shown in C. Note that jitter in the horizontal axis has been added to plots A. and B.

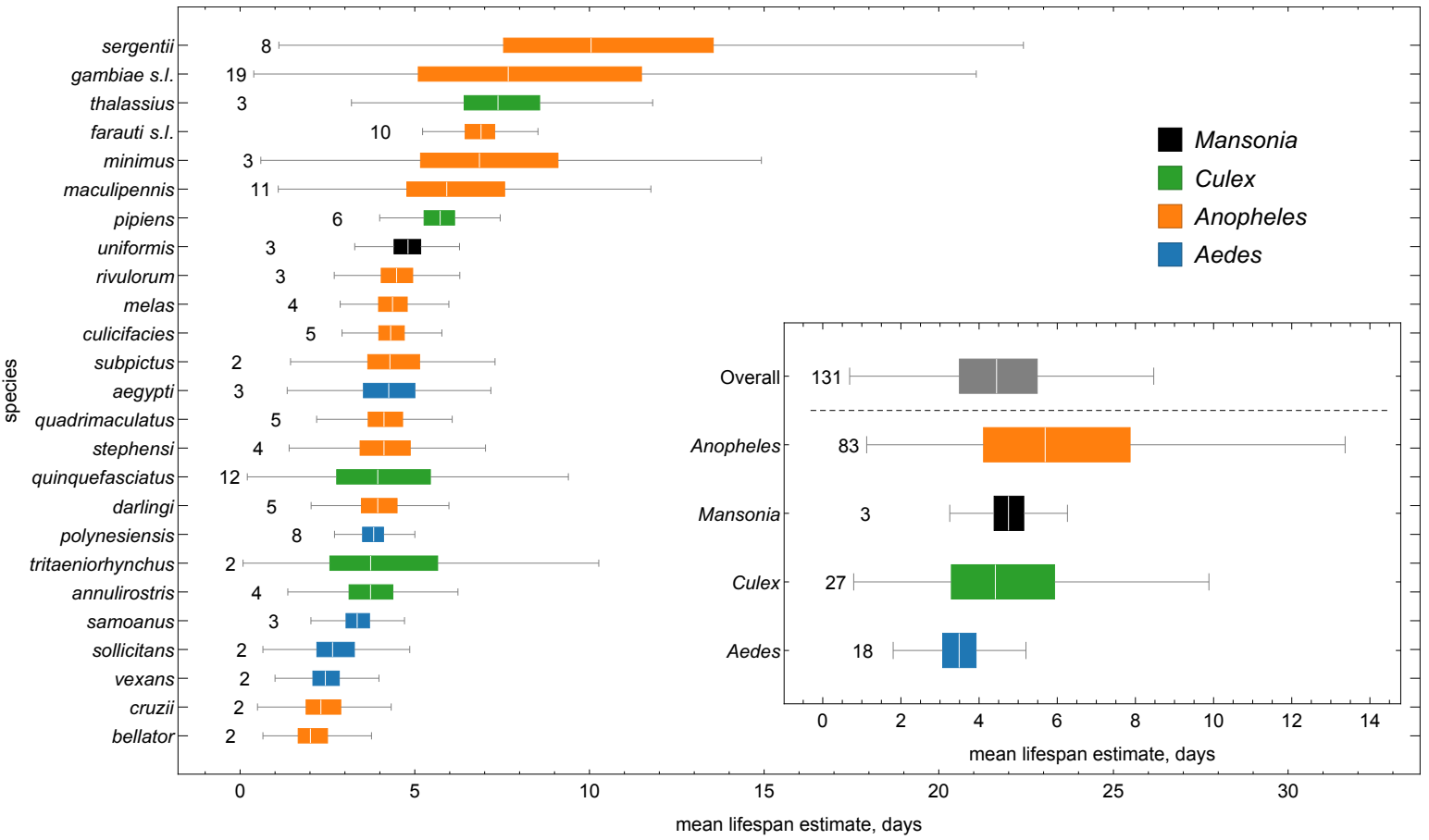


Figure 3.15: **Estimates of mean mosquito lifespan in terms of calendar age at the species, genus and overall levels as determined from the dissection data.** The middle line in each box shows the median estimates. The left and right box edges show the 25%, and 75% posterior quantiles respectively. The whiskers show the range of the data, excluding points lying more than 1.5 times the interquartile range away from each edge of the box. The numbers before the start of the left whisker indicate the number of individual time-series within each species. All estimates were obtained using the exponential survival model. Samples from $\mathcal{N}(4.3, 0.4)$ and $\mathcal{N}(3.8, 0.4)$ distributions for the duration of the 1st and subsequent gonotrophic cycles were used to convert the estimates of physiological lifespan (Figure 3.12) into calendar lifespan.

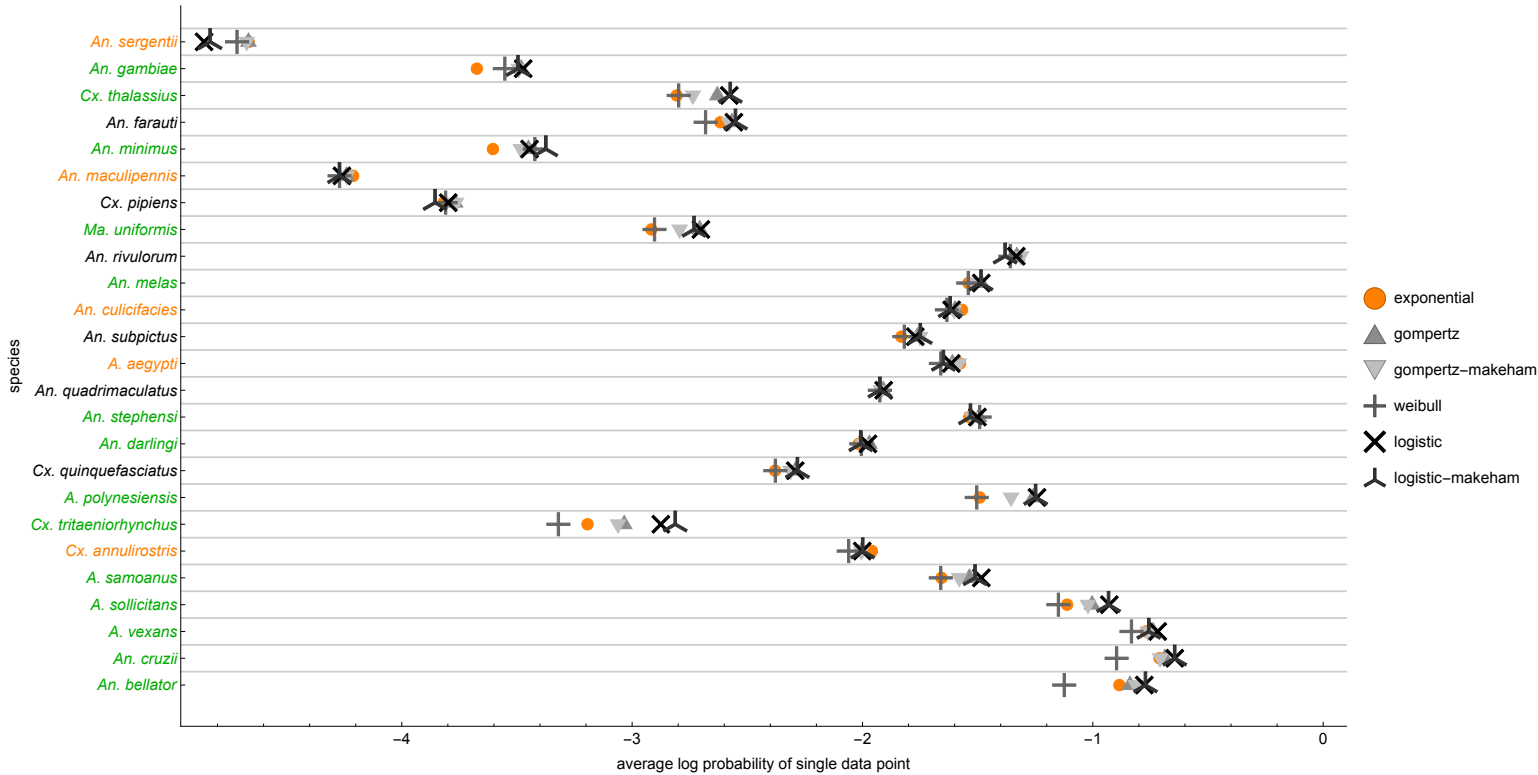


Figure 3.16: **Average log probability for each data point from each of the estimated models of mosquito mortality for the dissection data.** We colour those species names according to whether the exponential model provided the best fit (orange) or if a model with age-dependent mortality fit best (green). Finally if neither the exponential nor any other model out-performed (i.e. if $p > 0.05$ for the pairwise comparison of average log-likelihood assuming underlying data are normally-distributed) then we coloured the name black. The average log probability was calculated on independent test sets using K-Fold cross-validation with two folds for each species.

3.5 Discussion

3.5.1 Comparison with MRR estimates

To facilitate comparison between estimates of lifespan derived from dissection-based methods and those from MRR we combine the estimates from those species and genera involved in both studies (Figures 3.15 & 2.8) into a single plot (Figure 3.17). In comparing estimates from the two approaches it is reassuring that there is considerable agreement between the estimated posteriors. In both cases we estimate that *Anopheles sergentii* was the longest lived of the anopheline species with a posterior median of 12.9 days from the MRR meta-analysis and 10.1 days from the analysis of dissection studies. This species is a major vector of malaria in the Sahara (Sinka et al., 2010), where to act as a disease vector it must persevere through these hard conditions. It is reasonable to hypothesise that this species should live longer than those in environments where the potential for blood-feeding and oviposition is greater. The species with the greatest discrepancy in the posterior estimates was the major African malaria vector *Anopheles gambiae* (Sinka et al., 2012), where we estimate a (posterior median) mean lifespan of 2.6 days from the MRR analysis and 7.7 days from the laboratory-based approaches. For the individual genera we found the greatest discrepancy for *Anopheles* (a posterior median of 3.6 days from the MRR versus 5.7 days from the dissection analysis), followed by *Culex* (2.4 versus 4.4 days). By contrast for *Aedes* the estimates from the MRR studies (posterior median of 4.5 days) exceed those of dissection-based studies (3.5 days). Across all studies we estimate from the MRR analysis that mean mosquito lifespan is 3.7 days versus 4.4 days from the dissection-based studies.

Some of the differences in these group-level estimates between the two approaches is likely due to environmental and genetic differences between mosquitoes in the experiments that were analysed in each meta-analysis. However, we believe that part of the discrepancy

can be explained by the methodological differences in approaches. In particular we speculate that differences in dispersal rate can explain some of the discrepancy in estimates between the MRR and dissection-based approaches. Both *Anopheles* and *Culex* mosquitoes are generally thought to fly farther during their lifetimes than *Aedes*³, meaning that the estimates from MRR-based approaches will be most downwardly-biased for these genera. This is supported by our results since the dissection-based estimates (themselves not reliant on dispersal) exceed the MRR estimates for *Anopheles* and *Culex* mosquitoes, but not for *Aedes*.

The key assumptions of dissection based methods are: (i) physiological age can be accurately determined by dissection, (ii) the relationship between physiological and chronological age is known, and (iii) the population being sampled is in equilibrium (recruitment matches mortality) and (iv) individual mosquitoes can be randomly sampled from the population.

The number of gonotrophic cycles can be estimated by dissecting female mosquitoes and counting the number of dilations in the ovarioles (Polovodova, 1949; Detinova, 1962). However the reliability and accuracy of this method has been questioned by practitioners. The objections include the difficulty of the dissections (Kay, 1979; Hoc and Wilkes, 1995), the impracticality of dissecting more than a small proportion of ovarioles (Hoc and Wilkes, 1995) and the related issue of locating ovarioles that are deemed ‘diagnostic’ (Fox and Brust, 1994), as well as the considerable variation in numbers of ovariole dilations for mosquitoes known to be of the same physiological age (Kay, 1979; Russell, 1986b; Hugo et al., 2008). Indeed there is considerable uncertainty concerning the fundamental question of how dilations in the ovarioles form. Whilst the ‘Old School’ of thought (term coined by

³The stylised fact that aedines fly smaller distances throughout their lifetimes is, perhaps, biased towards *Ae. aegypti* - a primarily urban mosquito - because others, such as *Ae. albopictus* originate from rural areas (Bonizzoni et al., 2013), and have been shown to fly further in population genetic studies (Schmidt et al., 2017).

2416 Fox and Brust, 1994) headed by Polovodana (Polovodova, 1949) and Detinova (Detinova,
2417 1962) considers dilations to result from normal oogenesis, a ‘New School’ headed by Lange
2418 and Hoc (Lange and Hoc, 1981) has challenged this assertion. The New School believe that
2419 only abortive oogenesis results in follicular dilations because normal oogenesis destroys the
2420 sack-like structures (Fox and Brust, 1994). This means that the success of Polovodana’s
2421 method hinges on dissecting sufficient numbers of ovarioles to discover those with the high-
2422 est numbers of dilations, where abortive oogenesis has occurred in each gonotrophic cycle.
2423 They deem these ovarioles ‘diagnostic’ since only in these cases the number of dilations
2424 equals the number of gonotrophic cycles that have occurred. However as a mosquito ages
2425 the numbers of diagnostic ovarioles diminishes, since the random occurrence of normal oo-
2426 genesis in a particular ovariole means it no longer has a number of dilations that equals the
2427 number of gonotrophic cycles undertaken. This increased difficulty of finding diagnostic
2428 ovarioles as a mosquito ages would mean that there is an elevated chance of age ‘hypodi-
2429 agnosis’ for older specimens (Fox and Brust, 1994), and may bias estimates of population
2430 mean lifespan downwards. The difficulty of locating diagnostic ovarioles has recently been
2431 investigated using lab populations of *Culex* and *Aedes* mosquitoes by Hugo et al., 2008, who
2432 conclude that only a small percentage of ovarioles have a number of dilations equal to the
2433 number of gonotrophic cycles. Overall the bulk of evidence suggests that dissection-based
2434 physiological age determination likely understates true physiological age.

2435 Two common methods are used to estimate the duration of gonotrophic cycles (required
2436 to convert physiological lifespan to chronological lifespan): MRR studies (see for example,
2437 Gillies and Wilkes, 1965) where marked mosquitoes are recaptured and dissected to deter-
2438 mine the number of gonotrophic cycles that have occurred since release; and laboratory-
2439 based experiments using colonies of (typically) wild-caught females (or their progeny),
2440 whose biting and oviposition behaviour is observed over time (see for example, Afrane

et al., 2005). Whilst it is unclear how each of these methods should bias estimates of gonotrophic cycle duration, in our analysis we find that laboratory-based studies indicate a longer gonotrophic cycle (Figure 3.14). The distributions we ultimately used to convert physiological age into calendar age were calculated by pooling data across both of these two approaches, and so it is hoped our estimates incorporate this uncertainty over gonotrophic cycle duration. Since it is unclear what biases in estimation should result from MRR or laboratory-based studies, it is not straightforward to comment on this pooling affects our estimates of lifespan.

If a population of mosquitoes is shrinking this leads to a relative under-abundance of young mosquitoes, and a flattening of the survival curve, resulting in over-estimates of lifespan, and *vice versa* for growing populations. For stable populations, periods when shrinking occurs must result in equal changes in the population size compared to those when it expands. If mosquito collections occur equally-prevalently in each of these two modes then aggregating the data across all sampling times, and estimating a single model, should yield an approximately unbiased estimate of lifespan. As such, to mitigate against the risk of dynamical population structure (where possible) we aggregated series that were gathered at different points in time. However, the additional uncertainty of a fluctuating population size mean the credible intervals on the results are likely overly narrow.

The assumption of random sampling of the wild mosquito population has been questioned by field entomologists, although there are conflicting opinions as to whether this results in a relative paucity (Gillies and Wilkes, 1965) or abundance (Clements and Paterson, 1981) of nulliparous individuals. In our database we found a number of cases where there was a deficit of nulliparous individuals, some of which has previously been ascribed to the differing distribution of resting females between indoor and outdoor traps (Detinova, 1962; Gillies and Wilkes, 1965). Indeed, differences in parity have been found for mosquitoes

collected at different times of day (Gruchet, 1962; Gillies and Wilkes, 1963; Charlwood and Wilkes, 1979), and if the mosquitoes were caught indoor or outdoors (Gruchet, 1962; Gillies and Wilkes, 1963; Gillies and Wilkes, 1965). The combined bias induced by both of these effects is unclear since individually their results are of conflicting signs. Rather than a deficit of nulliparous individuals Clements and Paterson, 1981 suppose that the longer duration of the 1st gonotrophic cycle (not supported by our analysis; Figure 3.13) should mean a relative abundance of younger individuals, and hence apply a correction factor to inflate their count of nulliparous individuals. If the population size is stationary we know that the numbers of nulliparous individuals should exceed the uniparous group. Assuming this to be the case we chose to not include those counts of nulliparous individuals in our analysis where their number was less than 90% of the uniparous. Whilst we see no obvious differences in lifespan (data not shown) according to collection method or location we cannot discount the possibility that the assumption of random sampling is violated, although the directionality of the bias induced by this is unclear. However it is likely that our uncertainty intervals for lifespan are overly narrow.

Overall the evidence suggests that the estimates of mosquito lifespan derived from dissection-based methods will likely understate true lifespan⁴. This is driven mainly by the practical difficulty of dissecting sufficient numbers of ovarioles to locate those that are diagnostic of the underlying physiological state. However, the often higher estimates of lifespan from dissection-based analyses compared with those from MRR suggests that MRR estimates are more downwardly biased. Nonetheless, the different nature of the assumptions behind each of the two methods means they may offer complimentary information on mosquito survival. In addition to the differences of assumption, we note that dissection based studies require a highly specialised expertise which will often be unavailable, whereas

⁴At least those using the Polovodova, 1949 method compared to the Detinova, 1962 approach, which merely determines whether a mosquito is parous or not, and requires fewer assumptions.

2490 MRR methods are relatively simpler to use. Furthermore, most if not all dissection methods
 2491 that have been used previously are only applicable to female mosquitoes, whereas MRR can
 2492 be applied to either sex. Although dissection data gives detailed and, in comparison with
 2493 MRR, relatively unbiased estimates of age-structure, we thus foresee a continued reliance
 2494 on MRR experiments in many studies on particular mosquito populations. Efforts to use
 2495 both approaches concurrently will be particularly useful, and will allow us to quantify the
 2496 biases induced by the assumptions of each.

2497 3.5.2 Comparison with laboratory-based methods

2498 It is widely believed mosquitoes live artificially long under the comparatively benign con-
 2499 ditions of the laboratory. We find it informative to consider estimates of lifespan derived
 2500 from observations of such populations as they constitute an upper bound on the lifespan
 2501 of wild populations. Also, since the numbers of mosquitoes involved in large cage experi-
 2502 ments often numbers in the thousands, these estimates have lower uncertainty than those
 2503 obtained from field experiments. Styer et al., 2007 carried out such an experiment using
 2504 45,054 female and 55,997 male *Ae. aegypti*, and found that females lived nearly twice as
 2505 long as males; the median lifespan was estimated as 31.69 ± 0.06 days for females and 16.39
 2506 ± 0.03 days for males. A similar study by Dawes et al., 2009 using a lab colony of over 1000
 2507 female *Anopheles stephensi* found similar estimates for median lifespan (31-42 days). It is
 2508 clear that these estimates are typically many multiples of the estimates that result from the
 2509 analysis of field data. As discussed previously it is likely that both of the methods, MRR
 2510 and dissection studies, which we analyse in Chapters 2 and 3, represent lower bounds on
 2511 lifespan. However, without a unbiased method to measure mosquito lifespan it is difficult
 2512 to quantify the gap that exists between lifespan in the field and that in the laboratory.
 2513 The development of additional methods to estimate mosquito age, such as “Near-Infrared

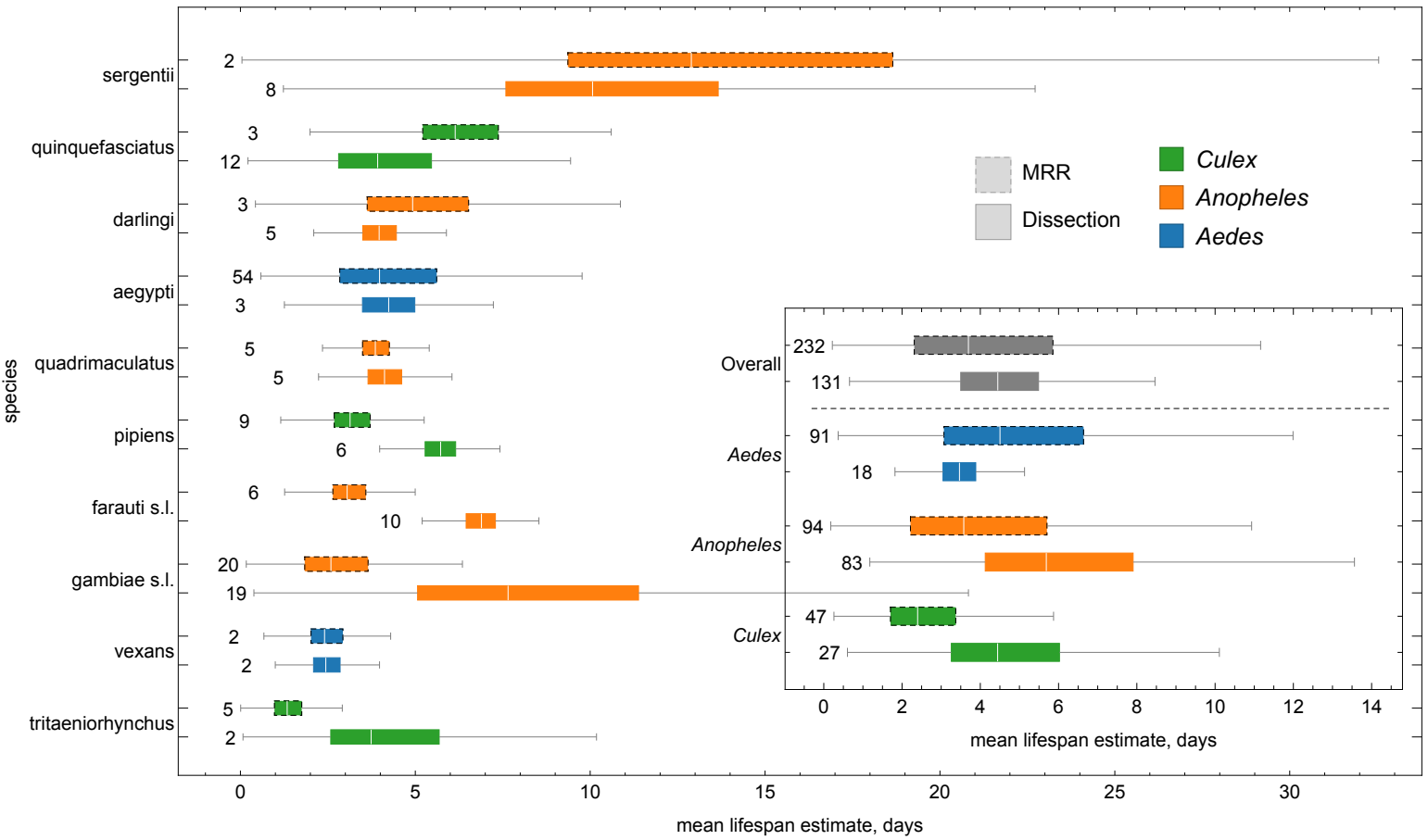


Figure 3.17: **A comparison of the estimates of mean lifespan from the meta-analysis of MRR studies (dashed box outlines) with those from dissection-based experiments (solid box outlines).** The middle line in each box shows the median estimates. The left and right box edges show the 25%, and 75% posterior quantiles respectively. The whiskers show the range of the data, excluding points lying more than 1.5 times the interquartile range away from each edge of the box. The numbers before the start of the left whisker indicate the number of individual time-series within each species. All estimates were obtained using the exponential survival model. Samples from $\mathcal{N}(4.3, 0.4)$ and $\mathcal{N}(3.8, 0.4)$ distributions for the duration of the 1st and subsequent gonotrophic cycles were used to convert the estimates of physiological lifespan from the dissection-based studies (Figure 3.12) into calendar lifespan.

2514 Spectroscopy” (NIR) which has been piloted in laboratory populations (Mayagaya et al.,
 2515 2009; Sikulu et al., 2011) are likely of considerable worth here (see Chapter 4).

2516 A better understanding of the causes of the elevated mosquito mortality hazard in the
 2517 field would be beneficial in attempts to control wild populations. Presumably this is due to
 2518 a range of factors that depend on the local ecology: predation, lack of food, the difficulty
 2519 in finding a mate, the additional risk of mortality from blood-feeding (animals and/or
 2520 humans), the relative sparsity of sites in which to oviposit, as well as variation in weather
 2521 conditions. Understanding the importance of each of these factors to mosquito mortality
 2522 is crucial to develop effective vector control, and we argue that field experiments to these
 2523 ends are important future research directions.

2524 **3.5.3 Does age-dependence matter?**

2525 In the meta-analysis of MRR experiments we found no evidence that mosquito mortality
 2526 increases with age, whereas in the dissection analysis the evidence was more mixed. We
 2527 argue that this is not a fair test of the exponential model since it was pitted against a pool
 2528 of five other models that incorporate age dependence. Indeed in some of the cases where
 2529 age-dependent survival functions fit the data better we found that the exponential model
 2530 nonetheless performed better than the worst-performing of the other models.

2531 The absence of a finding of age-dependence for the MRR analysis could be partly ex-
 2532 plained by the shorter estimated lifespan found by this approach relative to those dependent
 2533 on dissection. If marked mosquitoes live for shorter times it may be difficult to detect the
 2534 relatively weak age-dependence (see below) that we find in some dissection studies. Also,
 2535 it is possible that a reduction in mosquito flying ability as they age may mask the effects
 2536 of age-dependent mortality as detected by MRR studies (since older mosquitoes may die
 2537 at a higher rate but be more likely to remain in the study area). However, we cannot
 2538 discount the possibility that the failure to detect senescence using MRR studies is due to
 2539 the relatively low power of these experiments, where the experimental durations are often

short (see Section 2.5.1).

A number of experiments have found evidence of age-dependence in laboratory populations (Styer et al., 2007; Dawes et al., 2009). However, the artificially benign environment of the laboratory means mosquitoes may live considerably longer than in the wild, where they may die because of exogenous factors, before the effects of physiological decline have had time to manifest. Field experiments have also found evidence for age-dependent mortality. Clements and Paterson, 1981 presents a meta-analysis of survival data from both MRR and dissection experiments. On the basis of log-log plots of the counts versus time they conclude that there is evidence of age-dependent mortality in many of their studies. Here we used a Bayesian hierarchical model to combine information across a large pool of MRR and dissection experiments, and found mixed evidence for a mortality risk that increases with age. However, our results allow the possibility that age-dependent mortality can affect individual populations of mosquitoes. Harrington et al., 2008 conduct a field experiment where mosquitoes reared under laboratory conditions are marked and released at different ages. Analysis of the resultant MRR time-series suggest that mosquito mortality increases with age at release. However, it is possible that this type of field experiment may suffer from the same biases as laboratory-based approaches, because the released mosquitoes were often of ages considerably higher (up to 20 days) than typical estimates of wild mosquito lifespan.

3.6 Conclusion

In this chapter, I presented estimates of mosquito lifespan from a meta-analysis of mosquito dissection experiments which used the Polovodana approach of age determination (Polovodova, 1949). This method assumes that ovariole dilations appear each time a mosquito oviposits, providing a determination of mosquito lifespan that is measured in terms of physiological

2564 age. Since oviposition requires a mosquito to blood feed, the results from this approach
 2565 have direct epidemiological implications. If a mosquito does not blood feed at least twice
 2566 in its life, then disease transmission cannot occur. This means that, rather than calendar
 2567 age, mosquito physiological age, in the sense of gauging biting frequency, provides a more
 2568 direct measure of disease transmission potential; particularly if mosquito gonotrophic cycle
 2569 duration varies according to environmental conditions (Rúa et al., 2005; Lardeux et al.,
 2570 2008). In our results, we determined that, for the majority of mosquito species, an individ-
 2571 ual female mosquito, on average, undergoes a single gonotrophic cycle during its lifetime,
 2572 meaning relatively few individuals survive long enough to transmit disease. There is an
 2573 ambiguity with this assertion, however. Mosquitoes that survive only a single gonotrophic
 2574 cycle may still transmit disease if they are killed during or shortly after their second blood
 2575 meal, before completing a full gonotrophic cycle. If the increase in mortality associated
 2576 with blood feeding is substantial, then this effect may be particularly marked. Indeed, it
 2577 has also been demonstrated that mosquitoes laden with *Plasmodium spp.* may have an
 2578 elevated risk of mortality whilst blood feeding (Anderson, Knols, and Koella, 2000), since
 2579 at this point in the parasite's lifespan, there is a conflict of interests between those of the
 2580 mosquito and parasite (Ewald, 1994). A mosquito would like to take a meal of sufficient
 2581 size to allow egg development and fly away as rapidly as possible. Whereas, for the para-
 2582 site, it is preferable for the mosquito to take a large blood meal, to provide sporozoites the
 2583 greatest opportunity to be transmitted. Even if a mosquito does feed twice in its lifetime,
 2584 however, it may still not be able to transmit disease. If the blood meals are taken in too
 2585 rapid succession then the parasite has not the opportunity to transmute and migrate from
 2586 the gut to the salivary glands (i.e. survive the extrinsic incubation period). In this sense, it
 2587 is calendar age, rather than reproductive age, that matters most for disease transmission.

2588 Calendar, rather than physiological time, is also typically the measure of time used in

mathematical or computational models of disease epidemiology. It is hoped that this work will inform future modelling work and so, we convert our estimates of mosquito lifespan in physiological age into calendar age using estimates of the duration of the gonotrophic cycle. These estimates were obtained from a meta-analysis of laboratory and MRR experiments that determined this property of mosquito populations. The results we obtained indicated that laboratory mosquitoes, on average, had longer gonotrophic cycle durations than wild individuals. This could conceivably occur if mosquitoes adjust their feeding habits according to their lifespan because, in the wild, mosquitoes survive for less time than in the laboratory. The estimates of lifespan in calendar age indicated a similar story to the results from the MRR meta-analysis. As per the MRR analysis, in the majority of species, the average lifespan was estimated as shorter than 10 days, which is generally insufficient time to transmit malaria, which has an EIP of around 10 days. Taken together, these results indicate that likely a minority of mosquitoes are responsible for the bulk burden of disease. This suggests that disease transmission should be sensitive to interventions which reduce adult lifespan.

There was a good correspondence in the estimates of lifespan obtained from each method for those species which appeared in both datasets. This similarity is encouraging as it lends support to each methodology. It is interesting, however, to comment on the differences in lifespan estimates that exist for particular cases, as this may stem from the different assumptions that underpin each approach. In *Anopheles* and *Culex* mosquitoes, the lifespan estimated from dissection experiments exceed those obtained from MRR studies, whereas for aedine mosquitoes the MRR estimates exceed those from dissections. Anopheline and culicine mosquitoes are believed to fly further in their lives than aedine mosquitoes, which may be because many of their species are generally more adapted to rural environments. Since MRR estimates assume that mosquito dispersal is negligible compared to the

2614 effects of mortality, these differences in flight distances could explain these results. Further
2615 experimental data, however, is required to confirm this result.

2616 Models incorporating age-dependent mortality provided a better fit to the data in most
2617 species, in contrast to the MRR experimental results. There were, however, comparable
2618 numbers of species where exponential model performed equally well or better than the
2619 collection of age-dependent models. Overall, the results from the MRR and dissection meta-
2620 analyses indicate relative weak evidence in favour of a risk of mortality that increases with
2621 age, since in most cases the exponential model fit the data as well as or better than models
2622 which incorporated senescence. This suggests that the Ross-Macdonald model, created in
2623 1957, need not be changed to incorporate senescence, which considerably complicates the
2624 maths and likely, in reality, has only marginal repercussions on disease transmission.

Chapter 4

Using near-infrared spectroscopy to estimate mosquito population mean lifespan

4.1 Abstract

In recent laboratory experiments, it has been demonstrated that changes in cuticular hydrocarbons in mosquitoes which occur throughout their lifetime can be determined by shining near-infrared light through the mosquito specimens and examining the resultant spectrogram. The changes in mosquitoes' body chemistry are consistent across specimens and it has been shown that machine learning methods applied to the near-infrared spectra can predict mosquito age with reasonable accuracy. In this chapter, I apply a recently developed machine learning algorithm to predict the age of mosquito specimens across a database of 13 laboratory experiments conducted on anopheline and aedine species, and demonstrate that our approach considerably bolsters individual specimen predictive per-

2639 formance versus previously published studies. This analysis also indicates that for near-
2640 infrared spectroscopy (NIRS) to be useful for field entomological studies, considerable care
2641 to standardise experimental protocol will be necessary, since the between-study predictive
2642 performance of the machine learning algorithm is poor. Previous work has emphasised
2643 the importance of predicting the age of individual mosquito specimens, whereas in field
2644 epidemiology it is generally of more interest to estimate population mean lifespan. In this
2645 chapter, I undertake an *in silico* Monte Carlo experiment incorporating collection of wild
2646 mosquitoes and their subsequent analysis using NIRS, and use it to demonstrate that this
2647 technique could be used to monitor changes in mosquito lifespan that may occur during
2648 vector control campaigns.

2649 4.2 Introduction

2650 Mosquito borne diseases remain a major cause of suffering and death. Malaria is thought to
2651 have killed 438,000 people in 2015 (World Health Organisation, 2015) and it is estimated
2652 that the global cost of dengue exceeded \$39 billion in 2011 (Selck, Adalja, and Boddie,
2653 2014). Killing adult mosquitoes is the primary method of control for malaria, dengue,
2654 chikungunya and Zika amongst others and therefore one of the most important global public
2655 health interventions. Regular trapping of wild mosquitoes allows investigators to determine
2656 whether changes to population sizes have occurred, but mosquito abundance fluctuates
2657 substantially from day to day and according to where traps are set, rendering population
2658 size estimates imprecise. More importantly, the number of mosquitoes in itself is a poor
2659 predictor of disease transmission because many vector borne diseases have relatively long
2660 extrinsic incubation periods (see Section 2.4.2). Since many adult mosquito vectors live for
2661 less than 10 days (see Chapters 2 and 3), even in the absence of vector control it is only the
2662 rarer, older female mosquitoes that transmit the infection. Simple measures of abundance

are therefore insufficient to determine transmission risks or characterise the impact of vector control programmes. In the past, vector control was assessed by the percentage of the human population protected by interventions, such as the numbers sleeping under bednets or living in houses sprayed by insecticide. However, resistance to the most commonly-used insecticide used to control mosquito vectors of malaria is now ubiquitous throughout Africa (Ranson et al., 2011; Ranson and Lissenden, 2016), weakening the link between intervention coverage and protection.

It has long been recognised that the epidemiology of mosquito-borne disease is sensitive to wild mosquito lifespan (see Chapter 1), though there are no simple to use methods to reliably measure it in the field. Recently a number of different age-specific interventions have been proposed, including fungal biopesticides (Thomas and Read, 2007) and strains of *Wolbachia* that induce population mortality in older mosquitoes (Rasgon, Styer, and Scott, 2003), meaning that monitoring the age of wild populations is crucial to monitor the impact of these interventions. A key method for estimating mosquito survival in the wild is MRR studies (Chapter 2). A benefit of these experiments is that as well as estimating wild mosquito lifespan, other important ecological parameters can be estimated, such as mosquito dispersal rate or population size. However, these experiments are time-consuming, and difficult to do (Silver, 2007) and it can be costly to release the large numbers of marked mosquitoes required to achieve a reasonable level of accuracy (Desena et al., 1999). Furthermore, there are ethical concerns of releasing lab-reared or wild-caught mosquitoes into the wild due to the risk that they may contribute to the burden of disease. A more direct way of estimating mosquito age is to dissect the ovaries and assess whether (and how many times) a female has laid eggs (Chapter 3). This approach provides estimates of mosquito lifespan that are complimentary to those obtained from MRR, since the estimates do not depend on mosquito dispersal. However, this method is laborious,

2688 meaning it is difficult to scale and doesn't directly indicate biological age (just the number
2689 of feeding cycles, the length of which might vary with vector control).

2690 A number of novel laboratory methods have been developed to age mosquitoes including
2691 the analysis of cuticular hydrocarbons (Cook et al., 2006), pteridine fluorescence (Wu and
2692 Lehane, 1999), and near-infrared spectroscopy (NIRS) (Mayagaya et al., 2009; Liebman
2693 et al., 2015). Unlike the other techniques, NIRS is a rapid, non-destructive method which
2694 requires little technical training and no reagents. Large numbers of mosquitoes can be
2695 processed relatively economically making it feasible to use NIRS to routinely assess the
2696 efficacy of vector control interventions.

2697 NIRS works by measuring the change in absorbance of different wavelengths of light
2698 by organic compounds in specimens. Specifically, the stretching and bending of mainly
2699 C-H, N-H, and O-H functional groups may cause different wavelengths of light to be ad-
2700 sorbed according to the particular species or insect age group (Mayagaya et al., 2009).
2701 Prior to its use in mosquitoes, NIRS has been used in a range of agricultural and pest
2702 management applications: to identify the species present in grain stores (Dowell et al.,
2703 1999), the sub-species of termites (Aldrich et al., 2007), to differentiate between male and
2704 female tsetse fly pupae (Dowell et al., 2005), and to age-grade house flies (Perez-Mendoza
2705 et al., 2002). In applying NIRS to mosquitoes, the head and thorax of specimens of known
2706 ages are scanned (see Figure 4.1 for a picture of an near-infrared (NIR) machine) and a
2707 simple machine learning algorithm is then used to convert spectral data (see Figure 4.2)
2708 into estimates of biological age. Like all the above methods, NIRS has considerable mea-
2709 surement error making predictions of individual mosquito age relatively poor. Therefore
2710 researchers have resorted to using NIRS to define mosquitoes as young (<7 days) or old
2711 (≥ 7 days) (Sikulu et al., 2010; Sikulu et al., 2011; Dowell, Noutcha, and Michel, 2011;
2712 Sikulu et al., 2014; Liebman et al., 2015). Despite this simplification, NIRS is still only



Figure 4.1: **A near-infrared spectroscopy machine being used to scan individual mosquito samples.** Image adapted from Sikulu et al., 2011.

able to predict the percentage of young or old mosquitoes with 78% to 90% accuracy across both male and female specimens (Sikulu et al., 2010; Sikulu et al., 2011; Dowell, Noutcha, and Michel, 2011; Sikulu et al., 2014; Liebman et al., 2015). From a practical perspective NIRS predictions of mosquito age appear to be relatively stable, with only a slight loss of precision for mosquitoes killed earlier and kept in RNAlater or other preservation methods (Sikulu et al., 2011; Ntamatungiro et al., 2013). Equally, in the laboratory mosquito age was not noticeably impacted by the physiological status of the mosquito (Ntamatungiro et al., 2013) (i.e. whether it had mated, blood fed and laid eggs) or whether or not it had been exposed to pyrethroid insecticide (Sikulu et al., 2014).

To date, a relatively simple machine learning method called Partial Least Squares (PLS; in software provided by the NIRS manufacturer) has been used to convert spectral data into

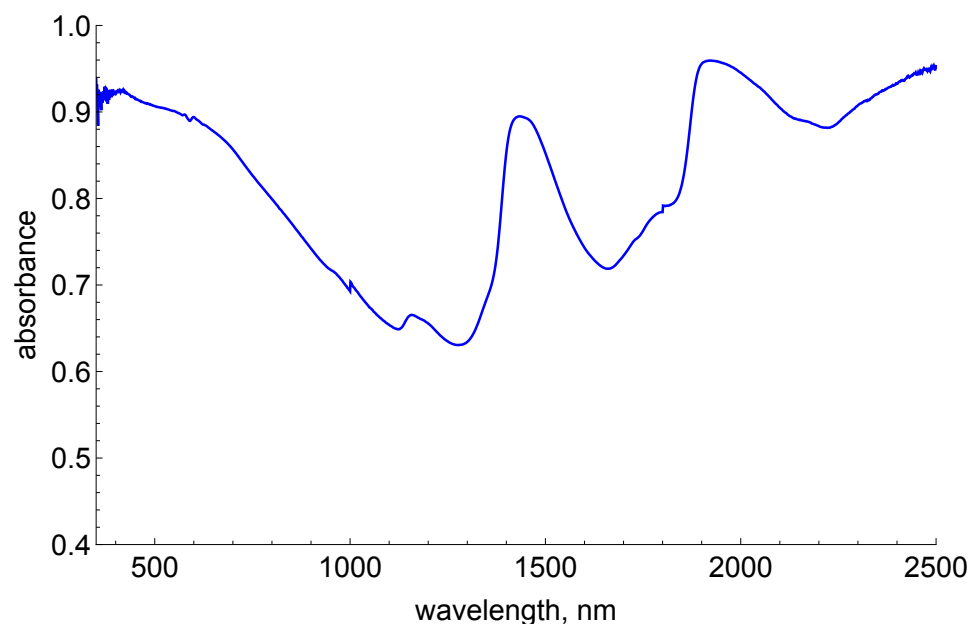


Figure 4.2: **An example NIR spectrum for a *Anopheles gambiae s.s.* mosquito taken from our database (Table 4.1).**

2724 predictions of individual mosquito age. Here we collate data from several published studies
 2725 that assessed the ability of NIRS to age *Anopheles gambiae s.l.* and *Ae. aegypti* mosquitoes
 2726 and use it to show how a more modern method known as interval PLS (iPLS) (Norgaard
 2727 et al., 2000) boosts the precision of individual mosquito age estimates. Furthermore we
 2728 show that the accuracy of iPLS is consistent across all 12 datasets in our database and
 2729 robust to variation in mosquito species. In the past mosquito aging techniques have been
 2730 evaluated on their ability to predict individual mosquito age. However, when monitoring
 2731 vector control intervention in the field, it is the overall mean and age distribution across
 2732 the mosquito population that is necessary to gauge intervention efforts. To ensure unbiased
 2733 estimation of population lifespan it is necessary for NIRS to produce unbiased estimates
 2734 of individual age for mosquitoes of any age. In line with the results of previous studies
 2735 (Sikulu et al., 2011; Dowell, Noutcha, and Michel, 2011) that used PLS we also find that

iPLS results in some biases in individual age estimates; overestimating the age of young mosquitoes and underestimating the age of old samples. Since this bias is consistent across test datasets we fit a general additive model to the errors in prediction and use this to produce unbiased estimates of mosquito age across each of our datasets. We use the bias-corrected iPLS model to demonstrate that by sampling modest numbers of mosquitoes even relatively imprecise metrics on the individual-level can generate highly accurate population-level estimates. The accuracy of predictions however is very sensitive to the nature of the calibration and testing sets. When both of these datasets were composed of mosquitoes from the same experiment, we could produce accurate predictions of mosquito lifespan. However we show that when samples from one experiment were used to calibrate our machine learning algorithm its performance on a separate test set is poor. To our knowledge this intra-experiment heterogeneity has not been discussed previously, and suggests the importance of standardising the methods for NIRS (mosquito killing, preservation, as well as instrument calibration) for field applications of NIRS.

4.3 Method

4.3.1 Data

The paired age-spectra data we analyse comes from five studies of NIRS experiments on mosquitoes produced by our experimental collaborators, Floyd Dowell and Maggy Sikulu-Lord, in experiments carried out between 2011-2015 (see Table 4.1 caption). Prior to scanning, mosquitoes were either anaesthetised and scanned fresh, or killed and preserved (see Table 1 for the preservation method used in each case). In studies A, C, D, E, F, G, H, and I, a QualitySpec Pro spectrometer (3502500 nm; ASD Inc, Boulder, CO) was used to scan the mosquitoes to generate the spectra; in studies, B, J, K and L were generated

2759 using a LabSpec 5000 NIR spectrometer (ASD Inc, Boulder, CO). In all experiments, the
2760 mosquitoes were scanned using the method described in Mayagaya et al., 2009, which we
2761 briefly summarise: the individual mosquitoes were laid on their backs 2mm underneath a 3
2762 mm-diameter bifurcated fiber-optic probe, containing 4 collection fibers and 33 illumination
2763 fibers. The viewing area spot size was set to approximately 3mm and was focused on the
2764 thorax and head. Each individual spectrum that we used in our analysis was the result
2765 of the average of 20 repeat spectra collected by the instrument. We note that, whilst
2766 previous studies use the term ‘near-infrared’ to describe the region of the spectrum used
2767 in mosquito studies, the wavelengths considered also include parts of the spectrum in the
2768 visible light range. Our resultant database contains 4549 unique NIR-spectra measurements
2769 obtained from 13 separate experiments using lab-reared populations of *Anopheles* and
2770 *Aedes* mosquitoes of known age (Table 4.1). Whilst not knowable, the unique number
2771 of mosquitoes in the data are considerably fewer than the total suggests, since repeated
2772 observations were made from the same mosquito throughout their lifetime. The majority
2773 of these samples are for *Anopheles arabiensis* (n=1898), followed by *Anopheles gambiae*
2774 *s.s.* (n=1904), *Anopheles gambiae s.l.* (n=295), and *Ae. aegypti* (n=452); mostly at ages
2775 less than 15 days (see Figure 4.3).

2776 4.3.2 Machine learning method: existent work

Previous studies have relied upon software produced by the NIRS manufacturer to build models to estimate mosquito age from individual spectra (Mayagaya et al., 2009; Sikulu et al., 2011). This software uses a statistical method known as Partial Least Squares (PLS) to estimate the age of individual mosquito samples from their respective spectra. PLS is a commonly used method in chemometrics for categorising data based on spectra. (For example, Norgaard et al., 2000 apply such a method to determine the concentration of

Study	Species	Country of origin	Sex	Preservation method	Number of mosquitoes in each age group					Total
					< 5 days	$5 \leq \text{age} < 10$	$10 \leq \text{age} < 15$	$15 \leq \text{age} < 20$	20+ days	
A	<i>A. arabiensis</i>	Tanzania	female	fresh	196	302	100	101	172	871
B	<i>A. arabiensis</i>	Tanzania	female	Silica gel	103	168	103	51	0	425
C	<i>A. arabiensis</i>	Tanzania	female	fresh	99	50	95	69	0	313
D	<i>A. arabiensis</i>	Tanzania	male	fresh	100	50	72	67	0	289
E	<i>A. gambiae s.s.</i>	Tanzania	female	fresh	208	305	0	102	93	708
F	<i>A. gambiae s.s.</i>	Tanzania	female	fresh	100	50	93	90	0	333
G	<i>A. gambiae s.s.</i>	Tanzania	female	fresh	100	50	86	94	0	330
H	<i>A. gambiae s.s.</i>	Tanzania	male	fresh	100	50	67	77	0	294
I	<i>A. gambiae s.s.</i>	Tanzania	male	fresh	100	50	76	13	0	239
J	<i>A. gambiae s.l.</i>	Tanzania	female	RNALater	91	162	42	0	0	295
K	<i>Ae. aegypti</i>	Australia	male	RNALater	46	43	46	50	42	227
L	<i>Ae. aegypti</i>	Australia	female	RNALater	41	45	46	43	50	225
Total	all				1284	1325	826	757	357	4549

Table 4.1: **A summary of the mosquito-age data for each study in our dataset.**

The numbers represent the binned counts of unique mosquito-age samples within each age class. The data for experiment B are from Mayagaya et al., 2015; the data for J are from Sikulu et al., 2014; the data for K and L are from Sikulu-Lord et al., 2016; the data for experiments C, D, F, G, H, I is from Mayagaya et al., 2009; the data for experiments A and E are not yet published although the experimental protocol followed Mayagaya et al., 2009.

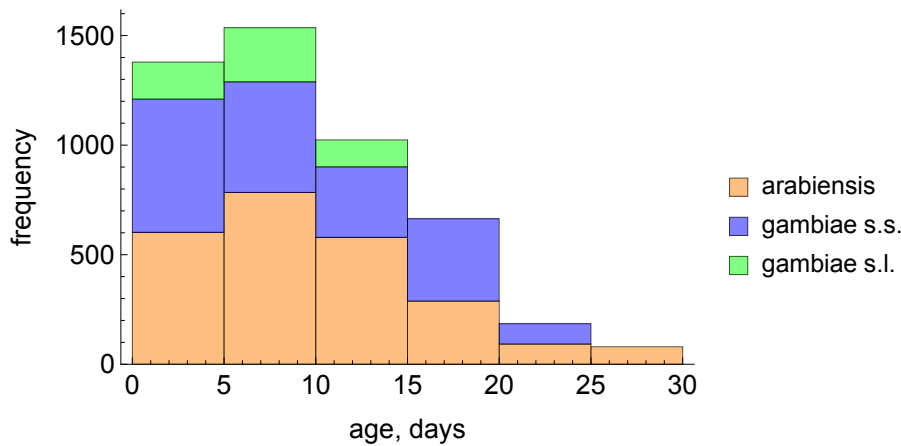


Figure 4.3: **The number of *Anopheles* spectra/age samples, coloured by species.** See Table 4.1 for the values used to produce this graph.

key chemical components in samples of beer!) It is a dimension reduction technique that is similar in nature to principal component analysis (PCA). PCA attempts to reduce the dimensionality of the data by decomposing the variance-covariance matrix for the independent factors (in our case, the NIR absorbance at each wavelength) into a few linear combinations of the independent variables which explain the majority of the variation. These *principal components* are then frequently used as features in machine learning algorithms to categorise or predict values of a dependent variable (in our case, the age of mosquitoes). In this sense, PCA is an *unsupervised* technique since it does not take into account the variable we seek to predict when forming the components. PLS by contrast is a *supervised* technique, since it explicitly takes into account the dependent variable when forming components (Maitra and Yan, 2008; Robinson, 1967). It does this by estimating latent factors in the independent *and* dependent variables which have the highest correlation (Figure 4.4). Whilst PLS can work with multidimensional outputs, in our case the output is a single variable (the age), although the algorithm's mechanism remain the same. Mathematically PLS amounts to the following decomposition,

$$\mathbf{X} = \mathbf{T}\mathbf{P}' + \mathbf{E} \quad (4.1)$$

$$\mathbf{Y} = \mathbf{U}\mathbf{Q}' + \mathbf{F} \quad (4.2)$$

2777 where $\mathbf{X}$ and $\mathbf{Y}$ are matrices of the independent and dependent variables respectively
 2778 (although in our case the dependent variable is univariate, and hence a vector); $\mathbf{T}$ and $\mathbf{U}$
 2779 are the projections of the independent and dependent variables into a lower-dimensional
 2780 subspace; $\mathbf{P}$ and $\mathbf{Q}$ are loading matrices that give the weighting of each variable on the
 2781 latent factors; $\mathbf{E}$ and $\mathbf{F}$ are error terms since the prediction of the dependent from the
 2782 independent variables is, in general, not exact. A number of PLS algorithms exist to
 2783 estimate the latent factors that have the highest covariance (see for example, Abdi, 2003

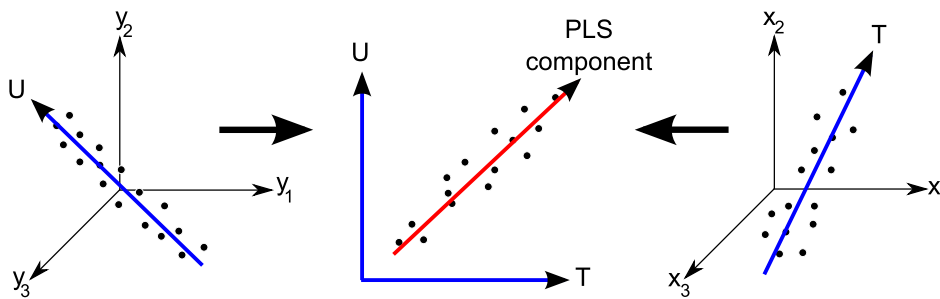


Figure 4.4: **A visual depiction of the method underlying Partial Least Squares.** In the left-hand plot a three dimensional dependent variable is decomposed into a component $\mathbf{U}$, which has a high correlation with the component $\mathbf{T}$ (middle plot), that is formed from the independent variables (right hand plot). In our case the output is unidimensional, and is the age of the mosquito samples. The independent variables are measurements of NIR absorbance at each wavelength.

for an introduction).

4.3.3 Machine learning method: our approach

We used a range of recently-developed methods to boost the predictive performance of the PLS algorithms that were used in previous mosquito age prediction experiments. The method we use is based on interval Partial Least Squares (iPLS), Norgaard et al., 2000. In this method each individual spectrum is split up into non-overlapping windows of separate wavelength intervals. The collection of individual wavelength windows are then used to train independent PLS models. The benefits of this approach are threefold: first, the individual tuning of the algorithm to each segment increases predictive performance; second, it aids understanding of the algorithms (since we are able to see how different parts of the spectra perform in predicting different characteristics of the data); lastly, this method allows us to throw away those parts of the spectra that are less useful in prediction, reducing the amount of noise in eventual predictions. After finding the most predictive parts of the spectra, it is important to look for the optimal combination of these segments. We

used the Matlab (MATLAB, 2010) Toolbox introduced by Leardi and Nørgaard, 2004 to determine the most predictive wavelength windows, and the combinations of these that had the highest accuracy. This resulted in different combinations of wavelength windows being optimal for each genus,

• *Anopheles*: $\lambda \in ([350nm, 708nm], [709nm, 1066nm], [1067nm, 1424nm])$,

• *Aedes*: $\lambda \in ([709nm, 1066nm], [1067nm, 1424nm])$.

We combined the models trained on the individual wavelength segments, with a PLS model trained on a wider spectrum [350nm,1850nm], since the inclusion of the latter was found to improve predictive performance for older mosquitoes. The predictions from the separate models were then combined using a generalised additive model with splines to result in a final estimate $\widehat{age}$,

$$\widehat{age} = \sum_{i=1}^K f_i(\widehat{age}_i) \quad (4.3)$$

where $\widehat{age}_i$ is the estimated age from wavelength interval $i = 1, 2, \dots, K$ and the individual $f_i()$ are functions estimated semi-parametrically, using splines. Other machine learning methods have been suggested as improvements on PLS in the literature, (in particular neural networks when the data are considerably non-linear Blanco et al., 2000) however, all other trialled methods fell short of the performance of the above methodology.

4.3.4 Cross validation and bias correction

We gauge the predictive performance of a model by training it on one dataset (the “training” set) and using this to predict individual mosquito ages on an independent dataset (the “test” set). The allocation of individual samples into either training or test sets was done in one of two ways: first, we generate training and testing sets by randomly selecting samples produced from a single experiment; second, we trained a model on data from one

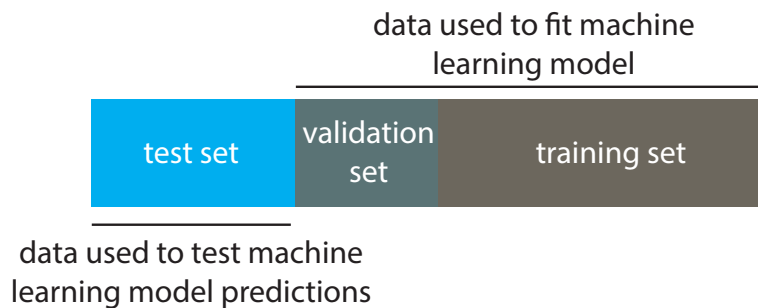


Figure 4.5: **The partitioning of the data used to train and test our machine learning model.**

experiment and used it to predict mosquito age in a test set composed of an independent experiment. The first approach allows us to gauge the within-study predictive capability of NIR, whereas the latter approach estimates between-study predictive performance.

For the within-study experiments we selected 70% of overall data to form the training set, and the other 30% comprised the test set. To avoid the risk of overfitting, we then split the training set further into a validation set (30%), and a separate *smaller* training set (the other 70%; see Figure 4.5). This final training set was used to train PLS models on each of the spectral windows, with the individual model predictions on the validation set used as independent variables in training the generalised additive model on the same set (and also for bias correction; see below). The test set was then used to get an independent evaluation the performance of the final model.

To implement the iPLS model we used the Caret package for R (Kuhn, 2008). We used the default settings of the Caret package which chose a number of PLS components to minimise the RMSE across 25 bootstrapped samples of the training set. We also tried K-Fold cross-validation (using $K = 10$ folds) but the performance of the model was indistinguishable from the default bootstrapping method. For the individual wavelength interval models, the maximum allowable number of PLS components was 20 (for short

intervals) or 40 (for the 350-1850nm interval).

The predictive performance of the models on the test set exhibited bias for certain age ranges; overestimation for young mosquitoes and underestimation for the oldest mosquitoes (Figure 4.6A.). Since any biases in individual mosquito age predictions result in biased estimates of the population mean lifespan we sought to correct these. To do this we regressed errors in age prediction on the true age of samples in the independent validation set using a general additive model of the form,

$$error_i = \sum_{i=1}^K f_i(age_i) \quad (4.4)$$

where there were $K = 5$ smoothing components $f_i()$, which were estimated semi-parametrically using splines. The estimated error from the above model was then used to correct the iPLS predictions on the independent test set. This process resulted in predictions that were unbiased estimates of the true age across the range of ages we surveyed (Figure 4.6B.). The cost of this bias correction was that there was a slight increase in RMSE versus the uncorrected iPLS by approximately 10%. This increase is because we trade-off a lower bias for a slightly higher variance.

The predictions of individual mosquito age that result from our machine learning approach are not restricted to be positive (Figure 4.6B.). This could be rectified by a post-prediction step where each negative mosquito age is made zero however, for our purposes this was undesirable. To produce unbiased estimates of the population lifespan we require that the estimates of mosquito age are unbiased across the range of ages encountered. For our case this means that the mean of the predicted age should be the true age across all age groups. With negative ages (whilst nonsensical biologically) this condition is met, and our estimator of mean mosquito lifespan is unbiased.

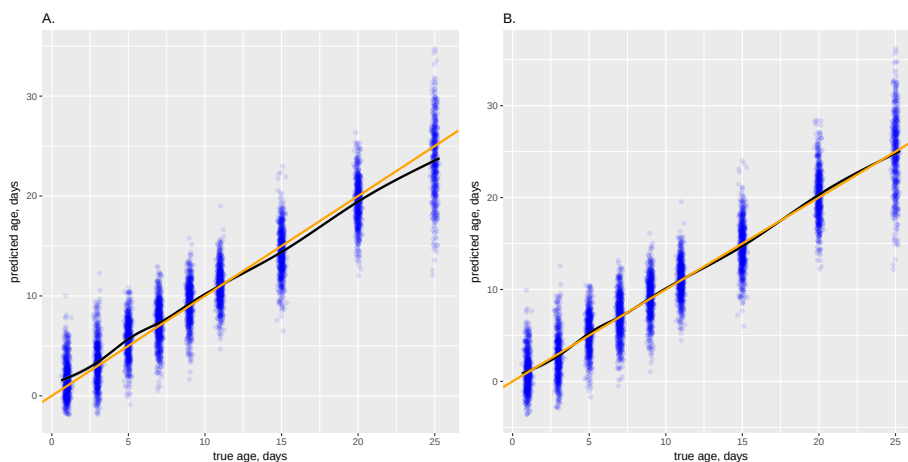


Figure 4.6: **The true versus predicted age of samples using iPLS (A.) and iPLS with bias-correction (B.).** The orange line is the *true = predicted* line and the black is a loess curve fitted to the data. The data come from study A. (Table 4.1), which used *Anopheles arabiensis*, and represent 5,821 test samples.

4.3.5 Surrogate Bayesian model

Following the results of Chapters 2 and 3 we suppose that wild mosquitoes face a constant mortality hazard (i.e. no senescence), and further assume that the population is of constant size, meaning that mosquito lifetime is exponentially distributed. We suppose we can sample randomly from this population, subject these samples to NIRS, and then use our iPLS model to predict the age of specimen. We use these NIRS-predicted ages to infer the *mean* mosquito lifetime of the population from which they came. Here we conduct a Monte Carlo analysis of the use of our machine learning algorithm as an estimator of this underlying wild-type population characteristic (see Figure 4.8 for a visual depiction of this process).

Since we cannot actually subject our *in silico* population of mosquitoes to NIRS we create a *surrogate* model that mimics the combined act of undertaking the NIRS together with prediction from our iPLS algorithm. This surrogate model takes a true mosquito age

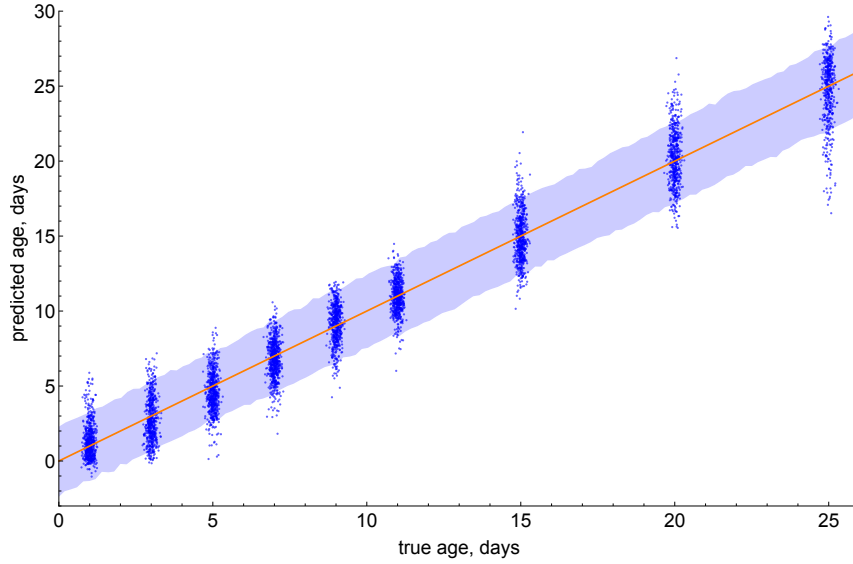


Figure 4.7: **The true versus predicted age of samples for an example study along with the 95% credible interval (blue shading) for the surrogate model predictions.** The orange line represents the true=predicted line. The data come from study A. (Table 4.1), which used *Anopheles arabiensis*, and represent 5,821 test samples.

as an input and outputs a predicted age. For the surrogate model we fit a Bayesian linear regression model with heteroscedastic errors to the combined (true, predicted) ages of the dataset resulting from NIRS predictions on the test set. Specifically we fit the following regression model to this data,

$$predicted_i \sim \mathcal{N}(true_i, \sigma_0 + \sigma_1 true_i^2), \quad (4.5)$$

where $\sigma_0, \sigma_1 > 0$, and represented a good fit to the data (Figure 4.7).

4.3.6 Population lifespan inference

in silico populations of mosquitoes together with the aforementioned surrogate model were used to assess the performance of using NIRS to predict population life expectancy. The

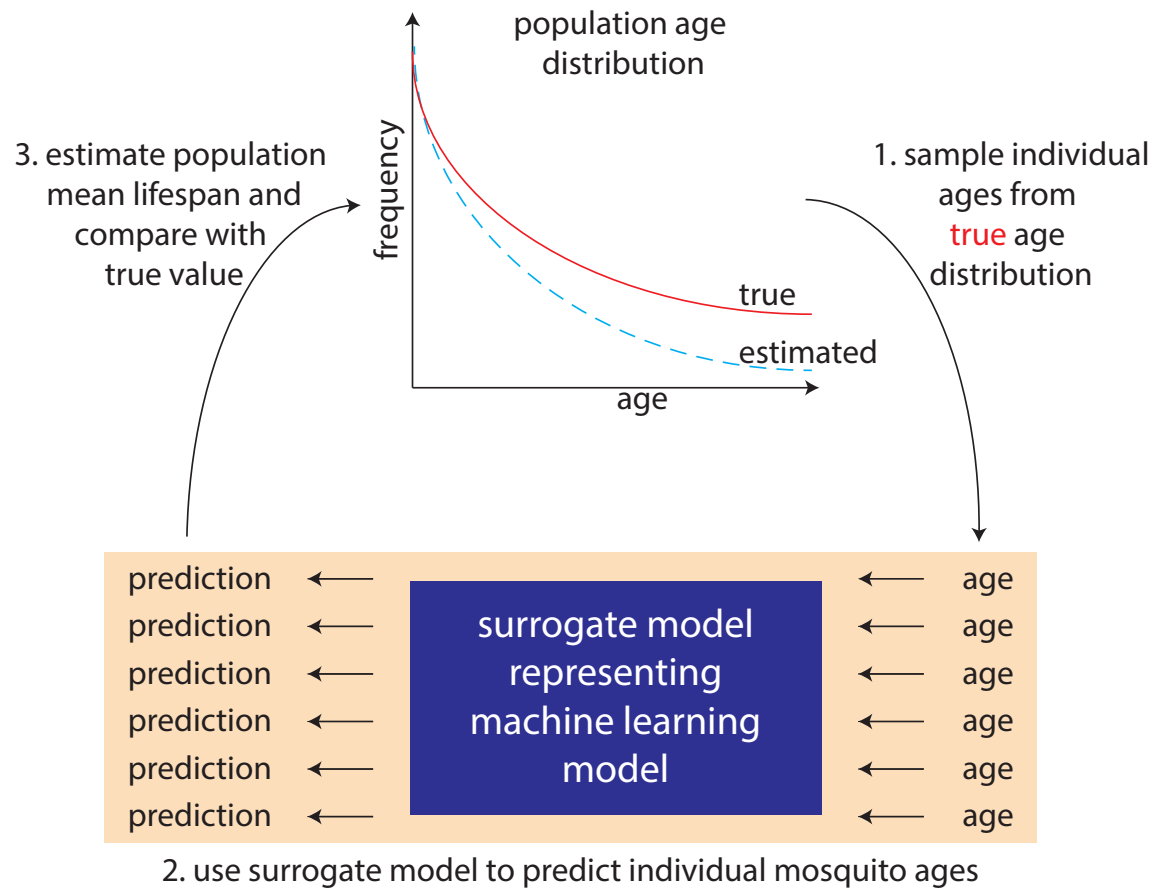


Figure 4.8: **The Monte Carlo process used to evaluate the use of NIRS & machine learning to infer population mean age.** Across each iterate we record the difference between the true population mean lifespan and the estimated.

2880 procedure for carrying out this analysis for the population mean lifetime, $\bar{L}$ is shown
 2881 in algorithm 1 (see also Figure 4.8 for further explanation of this process). Since the
 2882 population age structure has an exponential age distribution, the maximum likelihood
 2883 estimator of the mean lifespan is simply the sample mean age.

2884 Here we choose $n_{iteration} = 1000$ which resulted in a well-characterised distribution of
 2885 $\hat{L}_i$ being obtained. In what follows we investigate how the distribution of mean lifespan
 2886 estimates changes as the number of wild mosquito sampled varies.

Algorithm 1 A pseudo-algorithm illustrating the method behind Monte Carlo testing of the iPLS machine learning algorithm.

```

1: procedure DATA GENERATION
2:   for  $i = 1$  to  $n_{iteration}$  do
3:     sample  $(\sigma_{0i}, \sigma_{1i})$  from posteriors
4:     for  $j = 1$  to  $n_{sample}$  do
5:       sample  $true_{ij} \sim \exp(\lambda)$ 
6:       sample  $predicted_{ij} \sim N(true_{ij}, \sigma_{0i} + \sigma_{1i}true_{ij}^2)$ 
7: procedure INFERENCE
8:   for  $i = 1$  to  $n_{iteration}$  do
9:     estimate  $\hat{L}_i = \frac{1}{n_{sample}} \sum_{j=1}^{n_{sample}} predicted_{ij}$  (via MLE)
10: calculate bias, variance and RMSE

```

4.4 Results

4.4.1 Accurate prediction of individual mosquito age within a given experiment

The machine learning algorithm (iPLS & bias-correction) accurately predicts the ages of individual samples in an independent test set which was composed of mosquitoes from the same experimental study as the training set (Figure 4.9). The predictive performance of the model was negatively correlated with the sample size, although this was not statistically significant ($\rho = -0.52$, $t = -1.9$, $p = 0.08$). We expect that the lack of statistical significance is in part related to the small sample size ($n = 12$) that we use here, since further results suggest a fairly strong dependence of machine learning performance on sample size (Section 4.4.2).

To isolate the heterogeneity in predictive performance due to idiosyncratic experimental conditions we next fixed the sample size at 200 across all experimental datasets. We then compared the performance of our machine learning algorithm with the method used by previous studies across 100 replicates (Figure 4.10). In all cases the performance of our

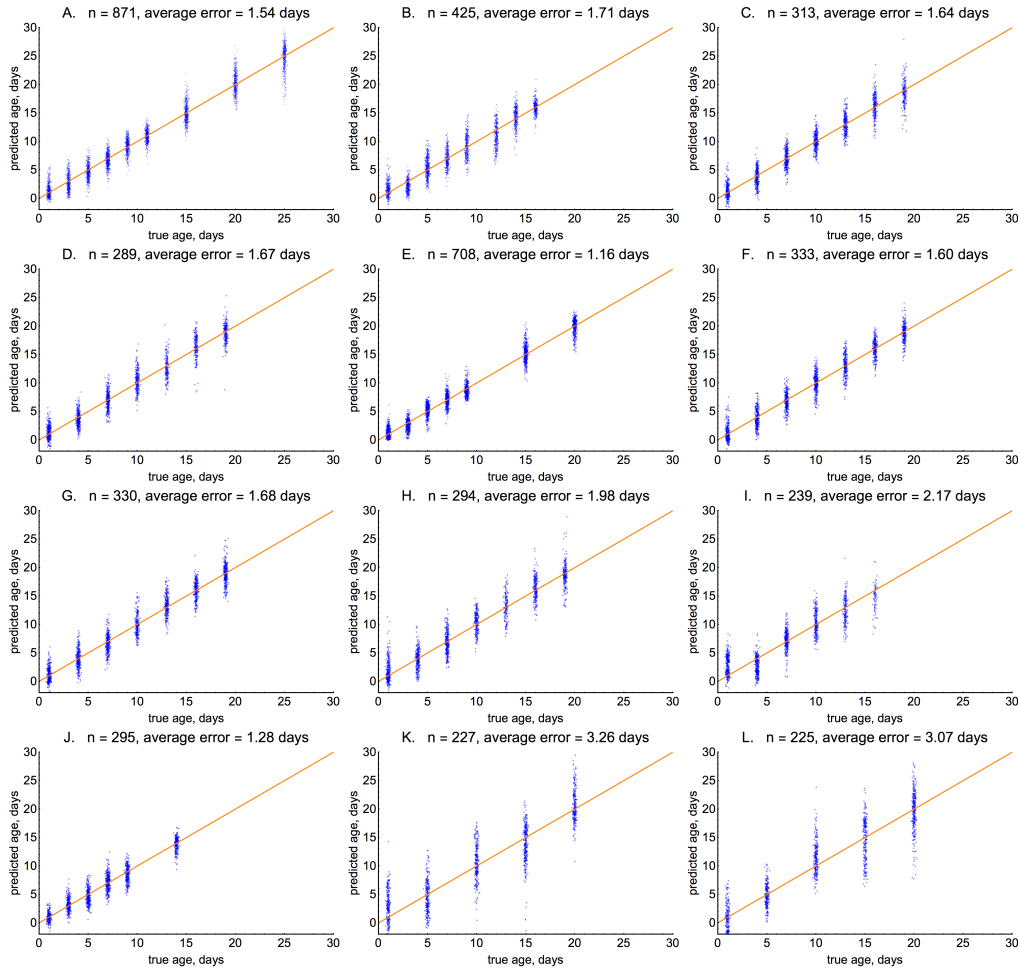


Figure 4.9: The true versus predicted age for each of the studies (A. to L.; Table 4.1) using test data composed of within-experiment samples. The sample size and RMSE ('average error') is as indicated above each panel. Studies A. to D. are were carried out with *Anopheles arabiensis* mosquitoes; E. to I. with *Anopheles gambiae* s.s.; J with *Anopheles gambiae* s.l., and K. & L. with *Ae. aegypti*. The test data was generated by randomly partitioning the sample into training (70% of overall sample) and testing (the remaining 30%) datasets across 20 repetitions.

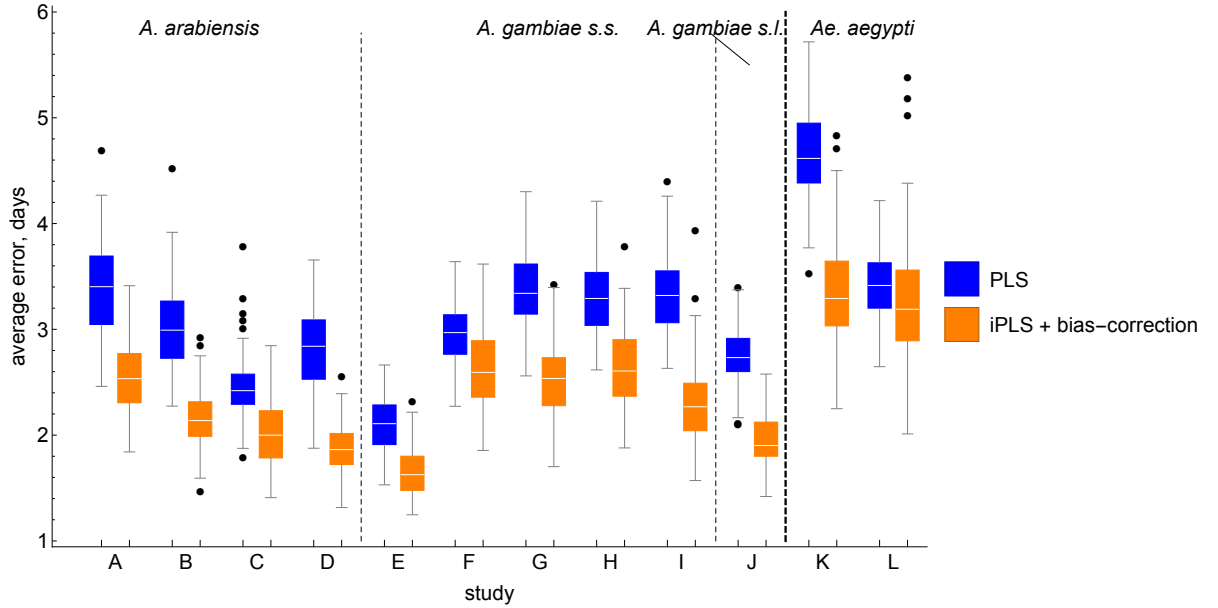


Figure 4.10: **Comparing the predictive performance of the previously-used machine learning method (‘simple PLS’) with our approach (‘iPLS + bias-correction’) on independent within-experiment test datasets.** In all cases a fixed sample size of 200 was used, and we randomly partition the data into training (70% of the sample) and testing (the remaining 30%) datasets. The predictive performance was estimated across 100 iterations of the random partitioning.

algorithm performed significantly better at the 5% level (Table 4.2). Overall the mean RMSE for our machine learning method was 0.72 days less than that of the previous method, which was consistent across species (*A. arabiensis*, 0.77 days; *A. gambiae s.s.*, 0.67 days; *A. gambiae s.l.*, 0.79 days; *Ae. aegypti*, 0.73 days).

4.4.2 Individual prediction errors decrease with sample size

To evaluate the improvement in prediction due to increased sample size we took the study with the greatest number of samples (study A., $n=871$) and estimated predictive accuracy at a range of sample sizes (Figure 4.11). As the number of samples increased the average

Study	t	p
A	15.54	<0.01
B	18.52	<0.01
C	10.43	<0.01
D	22.72	<0.01
E	13.07	<0.01
F	7.41	<0.01
G	17.97	<0.01
H	13.72	<0.01
I	20.86	<0.01
J	22.56	<0.01
K	19.69	<0.01
L	2.32	0.01

Table 4.2: **The results of a two sample t test comparing the mean predictive performance of our machine learning algorithm with the method used in previous studies.** In all cases the alternative hypothesis was that our test performs better than the previous method. The test assumes that the data in each sample are independent, although allows for heterogeneity in variances between the samples.

error decreased although with decreasing marginal returns. In particular, we found that the predictive error was inversely dependent on the square root of the sample size. These results suggest that there are benefits to collecting larger sample sizes, although these are most significant when increasing sample sizes from a few hundred to a thousand mosquitoes.

4.4.3 Inferring population mean lifespan is possible with modest sample sizes

Since our machine learning algorithm requires individual spectra to produce estimates we created a Bayesian surrogate model that took as an input the true age of a mosquito and outputted an estimate (see Section 4.3.5). This allowed us to undertake a Monte Carlo evaluation of the use of NIRS to infer population lifespan (Figure 4.8). In this analysis we assume that there is no age-dependent mortality and that there is constant recruitment into

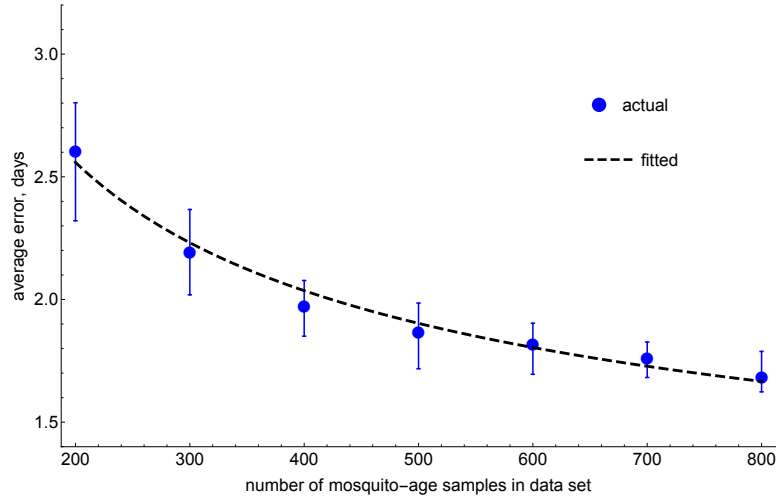


Figure 4.11: **Increasing the number of mosquitoes used in the training dataset substantially improves the predictive accuracy of our algorithm.** The solid points represent median average error seen in Study A. (Table 4.1), and the upper and lower fences represent the 75% and 25% quantiles, each of which was calculated across 100 replicates at each sample size. The regression line here was estimated as: $y = 0.44 + 34.81x^{-0.5}$, where y is the median RMSE and x is the number of mosquitoes in the training dataset.

the adult class, meaning that the population structure was exponentially distributed with respect to age. We assume that we can randomly sample individuals from this population and subject them to NIRS (here by inputting their true age into our surrogate model) which results in an estimated age. We also assume that alongside the wild-caught specimens we use a lab population of mosquitoes reared from wild-caught larvae as a training dataset (which hopefully mitigates against risk of poor predictive performance between experiments that we discuss in Section 4.4.5).

We assess the use of NIRS to infer mean lifespan for three different populations with mean lifespans of 3, 5 and 7 days as we vary the number of wild mosquitoes that are sampled (Figure 4.12I.). Assuming predictive accuracy of NIRS on the wild-caught mosquitoes matched that of study A, we conclude that sample sizes of 150 are sufficient to resolve between these three populations. However, it is likely that in the wild mosquitoes will possess greater heterogeneity in their physiology as they age compared to lab populations (mainly due to environmental heterogeneity). To account for this we inflate the estimated errors from study A. by 100% and 200% (panels II. and III. of Figure 4.12). Even if the errors are inflated by 200% inflation NIRS can still resolve between the three populations if at least 200 wild mosquitoes are caught.

4.4.4 The bulk of uncertainty in population lifespan estimates is due to sampling variability

To assess the relative effect of sampling variation and NIRS predictive errors on the inference of the population mean we considered a hypothetical scenario of applying NIRS in the wild. In particular we assume we have a calibration dataset composed of 500 mosquitoes whose error distribution is assumed to mimic that of study A. We then suppose that we can randomly sample 100 mosquitoes from a wild population whose ages were distributed

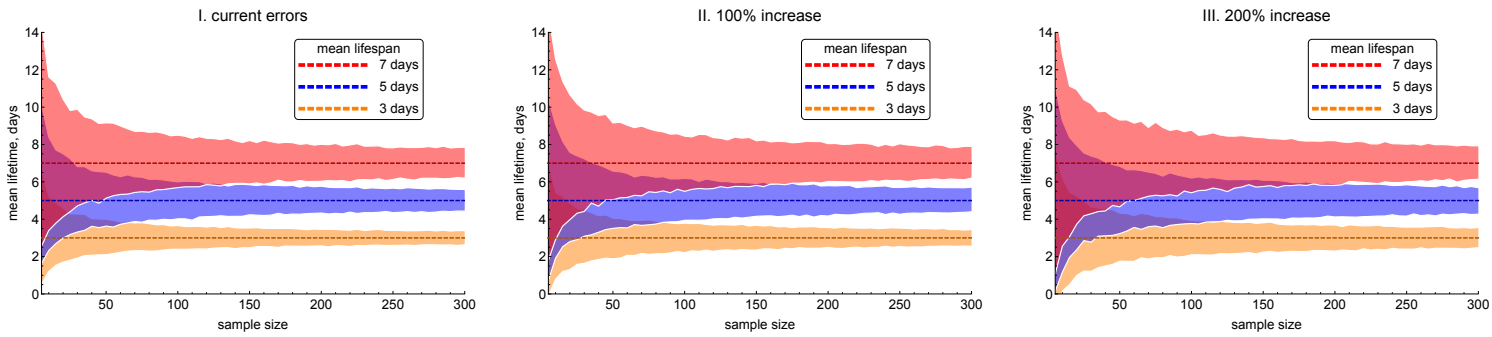


Figure 4.12: **The effect on sample size on the use of NIRS to infer mean lifespan for populations with life expectancies of 3 days (orange), 5 days (blue) and 7 days (red) assuming predictive errors from study A. (panel I.) versus inflated errors (panels II. and III.).** Dashed lines indicate the true value whilst shaded area shows the 95% quantile estimates across 1000 replicates as the number of mosquitoes sampled from the population increases. In all cases we assume that the NIRS errors followed a surrogate model fit to the data from Study A. To inflate the errors we inflated (σ_0, σ_1) in eqn. (4.5) by the requisite amount.

exponentially with a given mean lifespan. We then record the true age, and generate a prediction from NIRS using our surrogate model. We next computed the RMSE in prediction for the population mean lifespan from first using the true age of each sample and second using the NIRS-predicted age. We repeated this exercise across a range of population mean lifespans to obtain a decomposition of the mean lifespan prediction error into that due to sampling variation and the incremental effect of NIRS predictive error, as a function of life expectancy (Figure 4.13I.). Only for the very youngest populations (mean lifespan < 2 days) was the bulk of uncertainty in prediction due to NIRS measurement error. Whereas for the rest (mean lifespan 2+ days), the dominant source of noise was sampling variation. This is because the youngest populations have the least variation in lifespan meaning that there is little sampling variation and NIRS measurement error dominates. If instead we only sample 10 wild mosquitoes, both the sampling variation and NIRS errors become inflated, and the overall errors are disproportionately larger for populations with higher life expectancies (Figure 4.13II.).

In the wild NIRS measurement error will likely exceed that which we obtained in the laboratory. To investigate the performance of NIRS in this setting we carried out the same analysis except assuming that the overall error was 100% greater than we have found for study A. (Figure 4.13III.). In this context the relative role of sampling variation is diminished, although for populations of mean lifespans exceeding 3 days it nonetheless represented the dominant contribution to predictive uncertainty.

4.4.5 Between-study prediction is poor

The between study heterogeneity is evident from observing the individual spectra (Figure 4.14). Within a single experiment there are clear commonalities in the spectra collected, even across mosquitoes of different ages (Figure 4.14I.). However, within a single age cohort

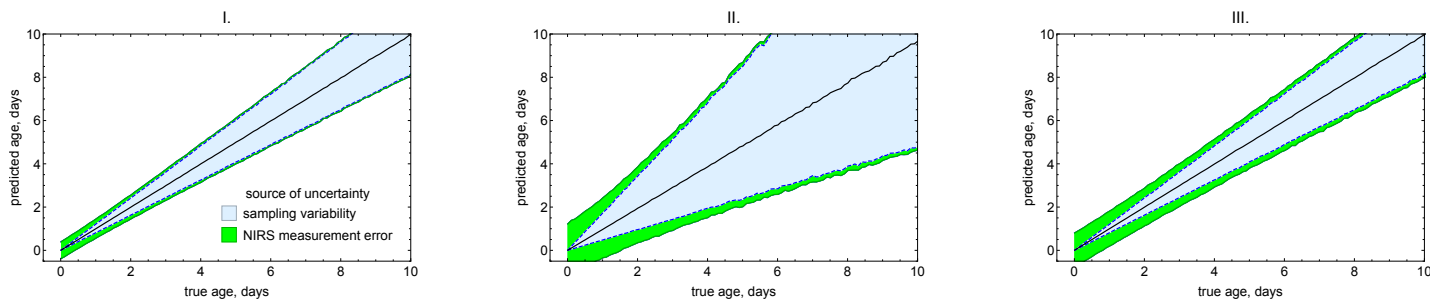


Figure 4.13: **The contribution of sampling variability (light blue shading) and NIRS measurement error (green shading) to the overall error in prediction of mean mosquito lifespan (solid dark green lines) for current NIRS errors with a sample size of (I.) 100 and (II.) 10 and (III.) a sample size of 100 with a 100% increase in errors.** The data shown represent the averages across 10,000 simulations. The NIRS errors were assumed to follow those estimated for Study A.

the spectra demonstrate considerable heterogeneity (Figure 4.14II.).

To assess how experimental heterogeneity affected the predictive power of NIRS we trained a model using samples from one experiment and used this to predict ages using spectra from another. We were not exhaustive here and instead picked three indicative training and testing experimental pairs to illustrate the performance of NIRS (Figure 4.15). In all cases the average error and bias were considerably worse than the within-experiment cases (see Figure 4.9). Furthermore this poor performance did not abate when we used a training set composed of the same species as the test (Figure 4.15 left panel).

It is unclear why out of the three cases considered the performance was best when using data from study A. to predict ages in study E. Since these studies had the highest sample sizes in our dataset we investigated how sample size affects prediction for this case by varying the number of samples from study A. we include in our training set (Figure 4.16). The effect of increasing sample size was relatively modest, although it was evident that there was lower uncertainty in the predictions for each age class. The reduction in bias with increasing sample size was also not a strong effect, although for the largest

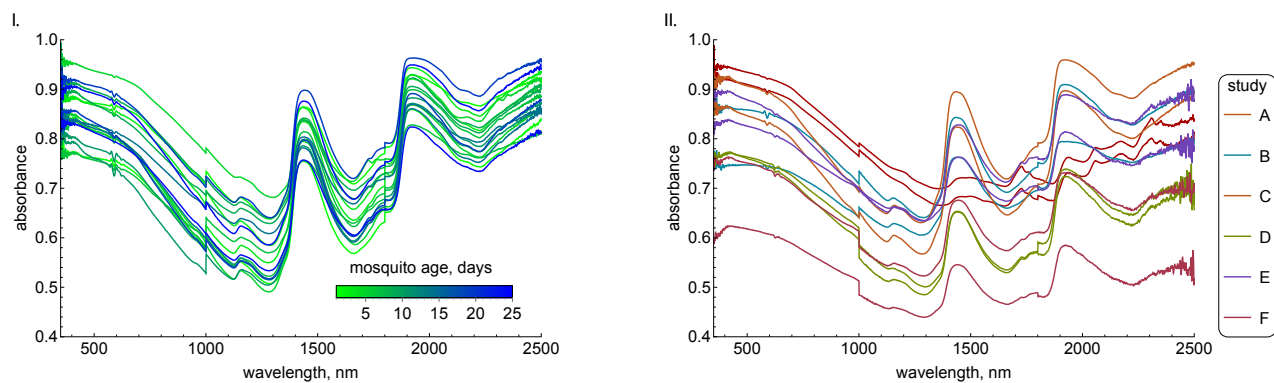


Figure 4.14: (I.) Example spectra for a variety of mosquito ages from within a single experiment, and (II.) example spectra for mosquitoes of age 7 days across six experiments (A. $\rightarrow$ F.). In both cases the datasets chosen represent a randomly chosen subset of the overall series.

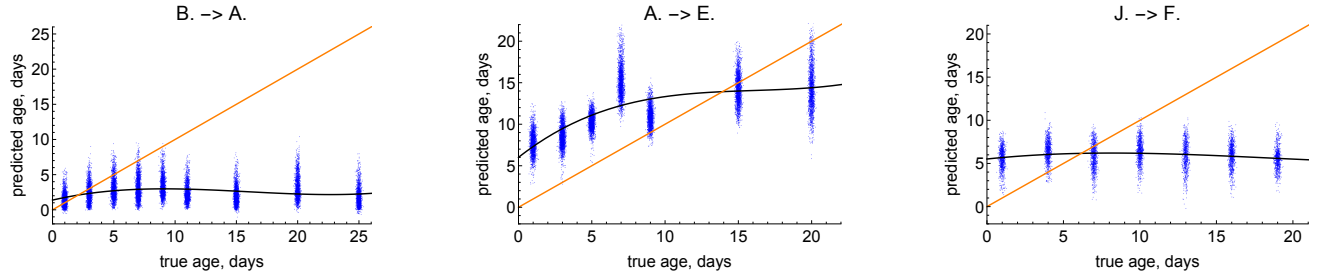


Figure 4.15: **The between-study predictive power of NIRS: using data from one study (B., A., & J.) to predict mosquito age in another (A., E., & F.).** The data (blue dots) represent the results across 20 iterations, and the orange line shows the *predicted = true* age ideal. In all cases the entire dataset from one study was used as a training set to predict all the ages from the test experiment.

sample size (Figure 4.16III.) there was a slight improvement in prediction for the oldest mosquitoes. Since sample size does not strongly affect our results we hypothesise that the better prediction in this case was due to a similar experimental protocol used across the two studies. Unfortunately since we do not have access to a detailed description of experimental protocols we are unable to test this hypothesis, although we hope that future work will further elucidate the impact of experimental heterogeneity on NIRS performance.

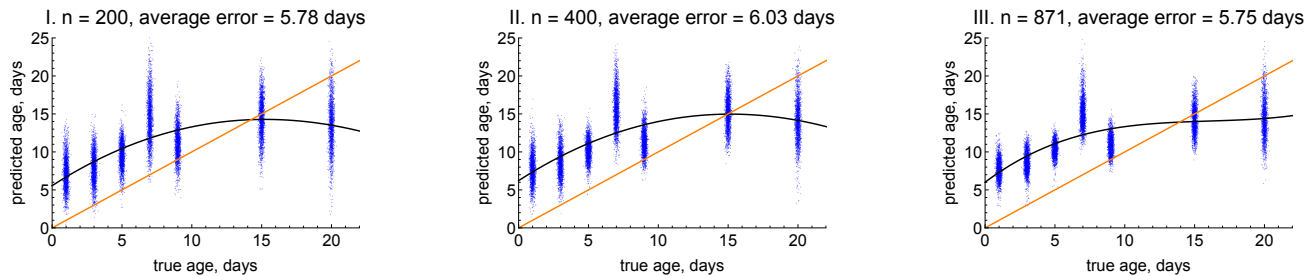


Figure 4.16: **The predictive power of NIRS using data from study A. to predict ages of samples in study E. as a function of sample size (I. to III.)** The data (blue dots) represent the results across 20 iterations, and the orange line shows the *predicted* = *true* age ideal. In cases I. and II. in each iteration a random subset of the data from study A. was chosen as training data.

4.5 Discussion

NIRS can generate accurate estimates of the average age of mosquito populations despite predictions of individual mosquitoes being relatively uncertain. This makes NIRS a promising tool for directly monitoring the efficacy of vector control interventions that targets adult mosquitoes. The ability of an insect vector to transmit disease (the vectorial capacity) is highly sensitive to the death rate of the mosquito (Macdonald, 1957; Smith and McKenzie, 2004; Brady et al., 2016). This makes average mosquito age an important metric for assessing community-level protection and the likelihood of a disease outbreak.

Enthusiasm for the use of NIRS (and other aging methods) has been tempered by the relatively high measurement error seen in published studies. All previous work has used laboratory reared mosquitoes which are likely to be more uniform (and give more consistent spectra) than those caught in the wild due to differences in diet, environment and diversity of hazards. It is therefore assumed that NIRS precision will further diminish once these real-life factors are taken into consideration. Here we show that NIRS can generate unbiased estimates of mosquito age which means that it can generate accurate

population-level estimates of average mosquito life-expectancy. Shifting the scale at which mosquito age-grading methods are evaluated from the individual-level to the population makes epidemiological sense as this is likely to be most important for most uses of the technology. This work therefore shows that NIRS is very well suited to monitor wild mosquito populations. More accurate methods of aging individual mosquitoes (with lower measurement error) will make relative little difference to the precision of population-level estimates. This is especially the case in mosquito populations with an average life-expectancy of 3-7 days, which is the range of age distributions mathematical models predict are going to be the most informative for the control of malaria (White et al., 2011) and likely for other mosquito-borne diseases also.

New methods for aging vectors should therefore concentrate on minimising bias to ensure the technology can be used at a population-level. It is also imperative to evaluate the NIRS in more natural field settings before its use as a tool for routinely monitoring vector populations can be assessed. It is reassuring that simulations which artificially increase the individual error of mosquito age estimates did not substantially reduce overall mean population precision. The average age of wild caught mosquitoes will also be sensitive to biases in trapping techniques, as some methods are thought to preferentially catch certain aged mosquitoes (see for example, Gillies and Wilkes, 1965). These biases need to be quantified and adjusted for in any mean estimate of population age. This work also assumed that mosquitoes die at a constant rate in the wild and the models would need to be adjusted to account for senescence should evidence arrive that they senesce.

Future studies using NIRS should carefully consider the structure of the calibration dataset to improve predictive accuracy. Increasing the number of mosquitoes improves precision though care should be taken to ensure a broad range of ages, each with the same number of mosquitoes. In the studies analysed here each had fewer older mosquitoes.

It is therefore not surprising that the naïve model better fitted young mosquitoes and underestimates the age of older vectors. Given the epidemiological importance of older mosquitoes a more balanced calibration dataset would be advisable.

Further work is also needed to improve the machine learning methods used to convert spectral data into mosquito age. There are a multitude of different predictive learning methods currently being developed, many of which may be better suited to analysing these data than the relatively simple techniques employed here. For example, it is possible that using deep learning methods (see for example, LeCun, Bengio, and Hinton, 2015) could generate significant improvements to individual predictions.

This study is the first-time data from different laboratory reared mosquito populations are analysed at the same time. Unfortunately it is clear that a calibration dataset generated with one population of mosquitoes cannot be used to predict the age of a different mosquito population. Further work is needed to test whether this is true as the data analysed here were from a disparate selection of studies which span multiple years, locations, different spectrograms, mosquito killing techniques and different mosquito sub-species. Further work that standardises the procedures or refining the machine learning methods may enable this intra-study variability to be reduced, possibly allowing a universal calibration dataset. If not then every study may require their own group of mosquitoes with known age in order to train the machine learning algorithm. These calibration datasets could be generated in the field using the F1 generation of wild caught mosquitoes which would encompass the genetic heterogeneity seen in the wild. Nevertheless the increased investment required to do this would diminish the utility of NIRS for routine vector monitoring, though it would be still invaluable as a research tool.

4.6 Conclusion

In this chapter, I applied recently developed machine learning methods to predict the age of individual mosquitoes in near-infrared spectroscopy experiments from a database of previously published studies, which was compiled by my collaborators and me. This statistical approach allowed a significant improvement in the predictive error for individual specimens over previously published methods. Unfortunately, however, whilst the NIR spectra were predictive of mosquito ages within a given experiment, models trained on data from one experiment could not be used to predict the ages of mosquitoes in another. The source of poor between-experimental predictive power is not currently knowable since the experiments used differing spectroscopy machines, preparation methods and were conducted on different laboratory strains of mosquitoes. If the experimental methodology drives the poor predictive power then, standardising experimental protocols would make NIRS an attractive method for entomological monitoring. By contrast, if genetic differences in the mosquitoes or environmental covariates are the source of low predictive power between experiments, then NIRS would require a separate calibration set for each application. Whilst this may be possible in some circumstances, this would diminish the field applicability of NIRS. Further experiments, therefore, to determine the source of between experiment heterogeneity in NIR spectra will be crucial.

Epidemiologists are primarily concerned with estimates of population lifespan. To prioritise minimising bias in this quantity we modified the machine learning method to incorporate a bias-correction step. Mean estimates can then be improved by increasing the number of mosquitoes sampled. This trade-off between precision and bias therefore depends on the ease with which mosquitoes can be caught in the wild. This work indicates that the number of mosquitoes required to generate useful estimates are well within the range that can be feasibly collected by routine vector monitoring programmes (see Tables

2.2 and 3.2).

NIRS has already proved its ability to differentiate between morphologically indistinguishable *An. gambiae s.s.* from *An. arabiensis* in wild caught mosquitoes from Tanzania (Sikulu et al., 2010) with relatively good accuracy. Further work is needed to test other vector species in different locations and to determine whether its predictive accuracy can be improved by refining the machine learning methods. At the moment PCR is expensive and require specialized training, restricting their use for routine monitoring. If so then NIRS would be ideally placed to be a single method capable of estimating mosquito age, species and infection status in mosquito populations.

Chapter 5

Using a homing endonuclease to control vector populations in a landscape with temporal heterogeneity

5.1 Abstract

The use of gene drive systems to manipulate populations of malaria vectors is currently being investigated as a method of malaria control. One potential system uses driving endonuclease genes (DEGs) to spread genes that impose a genetic load. Previously, models have shown that the introduction of DEG-bearing mosquitoes could suppress or even extinguish vector populations in spatially-heterogeneous environments which were constant over time. In this chapter, I combine a stochastic spatially-explicit model of mosquito ecology with a rainfall model which enables the generation of a variety of daily precipitation

3100 patterns. I then use the model to investigate how releases of a DEG that cause a bias in
3101 population sex ratios towards males are affected by seasonal or random rainfall patterns.
3102 The parameters of the rainfall model are then fitted using data from Bamako, Mali, and
3103 Mbita, Kenya, to evaluate release strategies in similar climatic conditions.

3104 5.2 Introduction

3105 Vector control lies at the forefront of efforts to combat malaria, and is expected to have con-
3106 tinued importance for disease control in the foreseeable future (World Health Organisation,
3107 2015; World Health Organisation, 2016a; World Health Organisation, 2017a). Insecticides,
3108 delivered via bednets and indoor residual spraying, are the current mainstay of vector
3109 control and their widespread application has helped bring about dramatic reductions in
3110 malaria incidence over recent years (World Health Organisation, 2015). Insecticide-based
3111 programmes alone, however, are not expected to be sufficient to eliminate malaria from the
3112 worst affected parts of sub-Saharan Africa, where climatic and socio-economic conditions
3113 are particularly conducive to endemic transmission (Le Menach et al., 2007; Ferguson et
3114 al., 2010b). Moreover, the emergence of insecticide resistance in vector populations threat-
3115 ens the continuing efficacy of these measures (Read, Lynch, and Thomas, 2009; Ranson
3116 et al., 2011; Ranson and Lissenden, 2016). Thus there is considerable need for innovative
3117 complimentary methods of vector control (Tanner et al., 2015).

3118 Homing endonuclease genes (HEGs) (and more generally, other driving genetic con-
3119 structs - DEGs - such as those involving Crispr-Cas9; Godfray, North, and Burt, 2017)
3120 offer potential for vector control by enabling the genetic manipulation of vector popu-
3121 lations (Burt, 2003). HEGs spread rapidly in populations by using a process known as
3122 homing, where heterozygous cells are converted to homozygotes (Mueller et al., 1993).
3123 They are found in many single-celled organisms, and are typically associated with self-

3124 splicing introns or inteins, that catalyse their own removal from mRNA, tRNA and rRNA,
3125 meaning that there is low toxicity to the host organism (Chevalier and Stoddard, 2001).
3126 For use in vector control, a HEG could be inserted into a functional mosquito gene in order
3127 to induce a population-wide knockout of that gene (the so called classic-HEG approach).
3128 Alternatively, a HEG could be inserted onto the male-specific Y-chromosome so that it
3129 disrupts the female-specific X-chromosome and thus biases the sex-ratio towards male off-
3130 spring (Y-drive). Both approaches aim to either suppress or eliminate the target vector
3131 population.

3132 The potential of these approaches has been demonstrated in laboratory conditions
3133 (Windbichler et al., 2011; Galizi et al., 2014), and investigated using mathematical models
3134 (Burt, 2003; Deredec, Burt, and Godfray, 2008; Deredec, Godfray, and Burt, 2011; North,
3135 Burt, and Godfray, 2013). The earlier models considered HEG releases in large panmictic
3136 populations to help understand the population load a HEG could impose in idealised
3137 conditions (Burt, 2003), how this depends on the specific nature of the HEG deployed
3138 (Deredec, Burt, and Godfray, 2008), and the implications for malaria reduction (Deredec,
3139 Godfray, and Burt, 2011). More recently, North, Burt, and Godfray, 2013 investigated
3140 how the spread of a HEG may be affected by the spatial structure of natural populations
3141 in heterogeneous landscapes. This work suggests that spatial variation in habitat quality
3142 does not affect the likelihood of spread in landscapes with abundant mosquito resources.
3143 However, in sparse landscapes a HEG may become extinct before reaching all parts of the
3144 population, meriting the use of multiple release site.

3145 These models provide a foundation for exploring how HEGs might influence vector
3146 populations, yet have not considered the role of temporal variance in rainfall. This may
3147 be important because members of the *Anopheles gambiae* complex, which include many of
3148 the major malaria vectors in Africa, tend to develop as larvae in either semi-permanent or

ephemeral water bodies created by rain (Gimmig et al., 2001; Shililu et al., 2003; Fillinger et al., 2004). The population dynamics of these species are therefore, to varying degrees, sensitive to fluctuations in rainfall (Shililu et al., 2003; Fillinger et al., 2004; Koenraadt, Githeko, and Takken, 2004; Fillinger and Lindsay, 2011).

The aim of this paper is to investigate how fluctuations in rainfall will influence the performance of a Y-drive HEG. We introduce a flexible model of daily precipitation which enables the generation of artificial rainfall data with specified characteristics (Bárdossy and Plate, 1992). This is combined with the mosquito population model of North, Burt, and Godfray, 2013 to explore HEG releases in different climates. In particular, we address the roles of seasonality and random fluctuations in rainfall and how these forms of variability should be accounted for when planning releases. We use two case-studies, based on Bamako in Mali and Mbita in Kenya, to investigate the combined influence of seasonal and random variability in these locations. These locations were chosen because they exemplify two distinct climatic regions of Africa where malaria incidence is high and HEGs may eventually be deployed (the Sahel and tropical east Africa respectively), and in each case we use historical data to fit our rainfall model. In summary, our analysis addresses the following questions. 1. When is the best time to release a HEG in a strongly seasonal environment? 2. How will shorter-term random variability in rainfall affect the performance of a HEG? 3. What is the optimal release strategy in the two case-study locations?

5.3 Method

5.3.1 Mosquito population model

Our model extends the spatially explicit model of *Anopheles* population dynamics developed by North, Burt, and Godfray, 2013, by incorporating temporal variability in environ-

mental conditions. Since the demographic component of both the current model and its predecessor are the same, we give only a brief review of the assumptions here. The model considers a population of mosquitoes that are each characterised by their sex, life-stage (juvenile or adult), genotype (whether or not they possess HEGs), and location in two-dimensional space. The landscape consists of small water bodies (aquatic habitats), and houses (feeding sites). The simulation is continuous time, where time steps are decided by the Gillespie algorithm (Gillespie, 1977). The values of demographic and environmental parameters used in the simulations are as detailed in Table 5.1.

Juvenile mortality can occur at any point within the simulation, and is increasing in the number of larvae at a particular aquatic habitat. We assume that density-dependent competition for food only acts on the instantaneous juvenile mortality rate, but does not affect the developmental rate, nor adult mortality. Whilst the latter two effects have been observed (Gimnig et al., 2002; White et al., 2011; Muriu et al., 2013), for simplicity we do not model these effects, although this could be considered in future work. The rate of mortality from juvenile competition is given by $\alpha n_{J_i}(t)$, where $n_{J_i}(t)$ is the number of juveniles at aquatic habitat i at time t . We also assume that the length of the juvenile stage is fixed at 9 days. Further work might relax this assumption by considering the individual stages of juvenile development: eggs, larvae (within this stage into instars) and pupae (Hancock and Godfray, 2007).

Emergent adult females undergo gonotrophic cycles, and are categorised according to whether they are searching for feeding sites in order to blood-feed, or aquatic habitats, in which to oviposit. Adult female mosquitoes move through the landscape at rates governed by their proximity to objects of search, representing area-restricted search behaviour. Dispersal is controlled by three parameters: a jumping rate r , the strength of the reduction in movement near an object of search, and the maximum jump distance s_G . A host-seeking

female jumps at rate $r/\beta^{C_H(t)}$, where $C_H(t)$ is the number of houses within the detection distance s_H (when there are no houses within a distance, the jumping rate is r). When a jump is made, a new location is drawn uniformly at random within a radius s_G centred on the mosquito's current location. Oviposition occurs in the same manner, except that a count is made of the number of aquatic habitats $C_O(t)$ in the local vicinity of the mosquito.

Mating occurs only between a male and unmated female which are within a given distance s_M of each other. The mating rate of a given female is given by $mC_M(t)$, where $C_M(t)$ is the number of males within radius s_M of her current location at time t . Blood feeding occurs at a rate $\gamma_H C_H(t)$ and, afterwards, the individual female becomes endowed with q eggs, where q is drawn from a Poisson distribution with mean κ . Oviposition occurs at a rate $\nu C_O(t)$, and during an oviposition, a female deposits p_X female and p_Y male eggs into a given aquatic habitat, where $p_X \sim \text{Poisson}(\omega/2)$ and $p_Y \sim \text{Poisson}(\omega/2)$, if the female carries wild-type sperm and $p_X \sim \text{Poisson}(\omega/2(1-e))$ and $p_Y \sim \text{Poisson}(\omega/2(1+e))$, if the female carries transgenic sperm, where e is the cleavage rate.

Following North, Burt, and Godfray, 2013, we assume that adult males do not disperse, and are confined to areas near aquatic habitats. This assumption was made because the behaviour of males, for example, where swarms form, and how individuals move between them, is not well understood. We are currently, however, working towards a model where both sexes of mosquito move. The males are categorised by whether their Y-chromosome carries the HEG, and mated females by whether the sperm they carry is X, Y or Y-HEG. The HEG biases the offspring towards males, by shredding 85% of the X-chromosomes in randomly-selected gametes, reducing the proportion of viable females.

The landscapes, which consisted of toroidal two-dimensional space, were populated at the start of each simulation with a stochastically-generated distribution of aquatic habitats and houses. The houses occurred at static locations in the domain, whereas the aquatic

habitats were created dynamically according to a spatial Poisson process (see Section 5.3.2). North, Burt, and Godfray, 2013 investigated the effect of spatial covariance between the location of aquatic habitats and houses, finding that the effects of low aquatic habitat density are amplified or mitigated if this covariance is negative or positive, respectively. Here, since we are most interested in the effects of temporal covariance, we assume that the spatial covariance between aquatic habitats and houses is zero.

Parameter	Symbol	Default value	Units
Demographic parameters			
Egg laying rate	ν	0.12	$day^{-1}C_O^{-1*}$
Feeding rate	γ_H	0.015	$day^{-1}C_H^{-1**}$
Egg load after feeding	κ	120	
Rate of mortality from competition	α	0.05	$day^{-1}n_J^{-1}$
Mortality rate in adult stages	μ	0.1	day^{-1}
Mating rate	m	0.01	$day^{-1}C_M^{-1***}$
Rate of HEG cleavage	e	0.6	$generation^{-1}$
Mean number of eggs laid per oviposition	ω	40	
Maximum mating distance	s_M	0.3	km
Feeding site detection distance	s_H	0.5	km
Breeding site detection distance	s_O	0.15	km
Basic jump rate	r	10	day^{-1}
Maximum jump distance		0.3	km
Stepping reduction near search object		10	
Environmental parameters			
Simulation domain area		16	km^{-2}
Mean density of feeding sites		$2^{1.5}$	km^{-2}
Mean density of aquatic habitats	κ	varies between 100-240	km^{-2}
Breeding site turnover rate	χ	0.1	$day^{-1}site^{-1}$
Mean density of sample points		100	km^{-2}
Sample point turnover rate		0.02	day^{-1}

Table 5.1: **The parameter values used in the rainfall simulations.** The local densities $*C_O(x, y)$, $**C_H(x, y)$ and $***C_M(x, y)$ at position (x, y) are the number of aquatic habitats, feeding sites and mates within distances s_O , s_H and s_M of that point respectively. The references for using these parameter values are given in North, Burt, and Godfray, 2013.

5.3.2 Linking rainfall with aquatic habitat density

The model was previously used to explore how landscape structure will influence the use of HEGs, by simulating releases across a wide range of feeding and aquatic habitat distributions (North, Burt, and Godfray, 2013). The analysis considered only static feeding site distributions, yet the aquatic habitats were allowed to dynamically appear and disappear to reflect the observation that *Anopheles* larvae often develop in small water bodies subject to rainfall and evaporation (Gimnig et al., 2001; Shililu et al., 2003; Fillinger et al., 2004). North, Burt, and Godfray, 2013 assumed that the rates of site creation and destruction were unchanging and equal in any given landscape, so that the average number of sites per unit area remained constant through time.

To investigate the role of climate, we suppose that variation in rainfall influences the creation of aquatic habitats. We denote by θ_d the rainfall amount on day d and we will refer to $\{\theta_d\}$, the time-series of daily rainfall, as the weather. The population model is linked to weather by supposing that on a given day d , aquatic habitats are randomly created at a rate $\chi \times \kappa \times \theta_d$ per unit area, which means that the number of sites created on day d is poisson distributed with expectation also equal to $\chi \times \kappa \times \theta_d$ per unit area. Sites are destroyed at rate χ per site, which means the longevity of a given site is exponentially distributed with mean $\frac{1}{\chi}$. Note that a large value of χ means aquatic habitats appear quickly during rain and disappear quickly during drought, and we thus use this parameter to control the lag between rainfall and aquatic habitat density. The parameter κ will be used to control the long-run average density of aquatic habitats, allowing us to investigate HEG releases across landscapes that share similar climates yet differ in their overall extent of breeding habitat.

5.3.3 A general model for rainfall

We model the rainfall dynamics using the procedure first described by Bárdossy and Plate, 1992. We consider a time-series $\{W_d\}$ that represents the net sum of meteorological processes influencing rainfall that we will (subsequently) transform to yield the rainfall time-series $\{\theta_d\}$. Rather than attempting to model the meteorological processes explicitly, we assume that on each day d , W_d is a random variable that depends on three factors, (i) the conditions of the previous day (W_{d-1}), (ii) the trend in meteorological conditions ($\{\mu_d\}$), and (iii) random noise ($\epsilon_d \sim \mathcal{N}(0, \sigma)$). Specifically, $\{W_d\}$ is a first-order autoregressive (AR-1) process given by the equation,

$$W_d = \rho W_{d-1} + (1 - \rho)\mu_d + \epsilon_d, \quad (5.1)$$

where the parameter ρ controls the degree to which conditions are dictated by recent history rather than the underlying trend $\{\mu_d\}$. We will use the standard deviation of the noise term σ to control the extent of randomness in conditions. Since the time-series $\{W_d\}$ may shift between positive and negative values, it is necessary to transform the series to derive the rainfall time-series $\{\theta_d\}$. Following Bárdossy and Plate, 1992, we use the transformation,

$$\theta_d = \begin{cases} W_d^\beta, & \text{if } W_d \geq 0 \\ 0, & \text{otherwise} \end{cases} \quad (5.2)$$

where the exponent β is used to control the variance in the rainfall distribution at the particular location. The flexibility of this model lies in our ability to manipulate the autocorrelation, noise, and trend of the simulated rainfall series. In the analyses that follow, we first explore two limiting cases before fitting the model using data from Bamako in Mali and Mbita in Kenya. In the first limiting case, rainfall is described by a sinusoidal

3257 trend with annual periodicity ($\mu_d = \mu_0 \sin(2\pi d/365)$) and no noise ($\sigma = 0$). In the second,
 3258 rainfall is subject to noisy and possibly autocorrelated variability ($\sigma > 0, \rho \geq 0$) yet with
 3259 no underlying trend ($\mu_d = \mu = \text{constant}$).

We fitted the model to the two case-study locations using 20 years of daily precipitation data in each case (1st January 1995 to 31st December 2014, obtained from the European Centre for Medium-Range Weather Forecasts, ECMWF). The parameter β was fixed *a priori* to give a good correspondence between simulated and actual rainfall series (Table 5.2, $\beta = 4$ for Bamako and $\beta = 5$ for Mbita). These values of β were used to compute $\{W_d\}$ by transforming the rainfall series (eqn. 5.2), where we assumed $W_d = 0$ on days with no rainfall. To estimate the trend $\{\mu_d\}$, we next computed the mean value of W_d for each day of the year, averaged over the 20 year period. The resulting series $\{\mu_d\}$, was then approximated by a Fourier series with K harmonics,

$$\mu_d = \sum_{k=0}^{K-1} a_k \cos\left(\frac{2\pi dk}{365}\right). \quad (5.3)$$

3260 $K = 4$ was chosen, which generated smoothed approximations of the $\{\mu_d\}$ series, which
 3261 produced rainfall series that had the most similar characteristics to the actual data (Table
 3262 5.2). Finally, ρ and σ were estimated by maximum likelihood.

3263 5.3.4 HEG releases

3264 The simulation environment is a 16 km^2 square area with periodic boundary conditions.
 3265 In all cases the landscape was run for a period of 50 days before the introduction of
 3266 mosquitoes, in order to allow the aquatic habitat dynamics to reach equilibrium. Similarly,
 3267 for all replicates the wild-type population was run for a period of at least 2 years before
 3268 the HEGs were introduced, in order to allow the population to reach equilibrium.

3269 For the purely deterministic sinusoidal model, the releases occurred at four equally-

characteristic	Bamako		Mbita	
	observed	simulated	observed	simulated
Mean wet day amount (mm)	4.90	4.65	9.13	9.76
Standard deviation of wet day amount (mm)	4.8	4.61	13.30	13.90
Median amount (mm)	0.00	0.04	3.35	1.82
P(dry tomorrow dry today)	0.87	0.90	0.65	0.68
P(wet tomorrow wet today)	0.67	0.66	0.85	0.79
P(wet)	0.28	0.23	0.70	0.60
Mean wet spell length (days)	7.65	9.58	2.87	3.10
Mean dry spell length (days)	2.99	2.91	6.61	4.68
Standard deviation of monthly totals (mm)	56.78	50.88	103.14	102.47
Maximum daily precipitation (mm)	110.75	42.26	158.95	222.83
Lag 1 autocorrelation	0.41	0.46	0.29	0.51

Table 5.2: **A comparison between rainfall data and model simulated data for Bamako, Mali, and Mbita, Kenya.** For simulated series, the values represent averages across 100 replicates.

spaced times of the year, corresponding to the start of the wet season, the middle of the wet, the start of the dry season, and the middle of the dry season. The default release strategy was 500 HEG males released uniformly at random across all the aquatic habitats in the landscape. After the introduction, the simulations were allowed to run for a further 6 years, to allow sufficient time for a HEG to affect the population dynamics of the wild-types. Finally, for the models for rainfall fit to Bamako, Mali and Mbita, Kenya, we considered releases in each of the different months of the year. In all cases, the releases of HEGs occurred at the midpoint of each month.

5.4 Results

5.4.1 Constant environment

In a constant environment, a released HEG will either spread to fixation or fail to establish (North, Burt, and Godfray, 2013). If fixation occurs, the load will either suppress or

extinguish the population, depending largely on the intrinsic growth rate of the target population (Deredec, Godfray, and Burt, 2011). Consider, for example, a release where rainfall is constant so that the aquatic habitat density is $142km^{-2}$ at all times. Fixation occurred in half the simulations of this scenario, while the HEG became extinct with no lasting consequences in the remaining half. Population extinction occurred in 80% of the cases where the HEG became fixed, suggesting that the population intrinsic growth rate is relatively small for the parameters we assume.

5.4.2 HEG dynamics are more variable in a seasonal environment

In an environment with an average aquatic habitat density of $142km^{-2}$, yet with seasonal variation in the density, a wider range of dynamics is possible (Figure 5.1). Now, fixation occurred in 34% of simulations and invariably led to population extinction (Figure 5.1a). The HEG became extinct with no lasting consequence to the population in 49% of simulations (Figure 5.1b). In the remaining simulations, the HEG induced extinction despite first becoming lost from the population (17% of simulations, Figure 5.1c) or, occasionally, the HEG and wildtype males both remained in the population throughout the simulated period of 6 years post release (0.4%, Figure 5.1d). A common feature of these dynamics is a tendency of the HEG frequency to fluctuate in an annual yet unpredictable manner (Figure 5.1 a-d grey lines). If a HEG fixes in a wet season, the population is likely to become extinct in the following dry season because the scarcity of aquatic habitats exacerbates the genetic load. If a HEG fails to fix in a given wet season, it may become lost from the population in the following dry season. In most simulations, the population will recover after the HEG is lost, although this may take several years. In some simulations, however, the remaining population is so impaired from the invasion that it too becomes extinct during a subsequent dry season.

5.4.3 The start of the wet season is the optimal time for release

The relative probabilities of these outcomes are influenced by both the time of year of the HEG release, and by the average density of aquatic habitats through the year (Figure 5.2). Extinction is generally most likely if the HEG is released at the start of the wet season (blue lines) and least likely if the release is at the start of the dry season (mauve lines). Release time is important in seasonal environments because of differences in the probability that the HEG becomes fixed and, in particular, that this occurs in the year of release. Due to the tendency of HEG frequency to fluctuate, it is most effective to release a HEG when it is likely to start increasing in frequency. We find this coincides with the start of the wet season, when the target population is growing. An additional benefit of releasing at this time is a relatively high initial HEG frequency due to the target population being small. If the initial frequency is high, fixation may occur during the first wet season leading to population extinction in the following dry season. We thus find that the release size is generally more important in seasonal than constant environments, and large releases may compensate for sub-optimal release timing (Figure 5.3).

5.4.4 In seasonal environments eradication is easier to achieve if the average aquatic habitat density is low or high rather than intermediate

Irrespective of release time, we found the HEG is more likely to induce eradication in environments with a low rather than intermediate average aquatic habitat density, in contrast to the trend when conditions are constant (Figure 5.2a). This partly reflects a higher initial HEG frequency in the case of low average aquatic habitat density, because the target population is smaller, and thus a greater chance of fixation in the year of release. In addition, the possibility of extinction occurring after the HEG is lost from the population is more likely if the average aquatic habitat density is low (Figure 5.2b). If aquatic habitats

are generally rare, the mosquito intrinsic population growth rate will be particularly small during the dry season; in these conditions a population that has been suppressed by the impact of the HEG may randomly become extinct.

In contrast to this trend, we found the probability of extinction generally increases with aquatic habitat density in environments where the average density is already high. A higher density of aquatic habitats reduces the risk of a HEG becoming lost from the target population, and thus promotes HEG establishment.

Our results suggest that the roles of initial HEG frequency and underlying vulnerability of the target population are more important when aquatic habitats are generally rare, while HEG loss in the dry season is important when aquatic habitats are more abundant. However, the balance of these factors depends on release timing, and we expect that detailed models will be required to discern their relative importance in specific scenarios.

5.4.5 Random variation in rainfall generally increases the probability that a HEG will induce population extinction

We next investigate the role of random weather variation, by allowing variance in the distribution of daily rainfall (controlled by σ) and autocorrelation in the rainfall time-series (controlled by ρ ; Figure 5.4). In cases of low variance and low autocorrelation, we observed only minor effects of the variability on the outcome of a HEG release (Figure 5.4a, blue versus black). The probability that a HEG induces population extinction increases, however, if either the variance in rainfall is increased (purple versus blue) or if the rainfall time-series becomes more autocorrelated (magenta versus blue). We found the probability of extinction is highest if the rainfall distribution has both high variance and high autocorrelation (red).

Random variability in rainfall increases the probability of extinction because it increases

variability in the dynamics of the target population. If a HEG becomes fixed in a variable climate, the population is likely to become extinct in a subsequent drought (Figure 5.4d). Moreover, if the variability is great, the impact of a HEG may lead to population extinction even after the HEG becomes itself lost from the population (Figure 5.4b). These effects of variability are to some extent lessened by a reduced probability of fixation (Figure 5.4c). Rainfall variability makes HEG establishment more difficult since a release may occur during drought, when there are few available aquatic habitats. For the parameter ranges investigated, we found the greater potential impact of a HEG in a variable environment almost always outweighed the greater difficulty in establishment.

Note that our analysis assumes the releases occur on random dates with respect to weather conditions, whereas a release programme may have some capacity to delay a release if conditions are particularly adverse. A more systematic strategy to dealing with an unpredictable climate may be to release multiple time-delayed batches of HEG-bearing mosquitoes rather than a single large batch, and the probability of fixation increases if this bet-hedging approach is taken (results not shown).

5.4.6 Releases during the first half of the wet season result in the highest chance of population extinction, although in some cases the time taken for extinction to occur may be lower by releasing HEGs during the dry season

We investigated HEG releases in Bamako, Mali and Mbita, Kenya, using simulated rainfall series after fitting the rainfall model to 20 years of data in each case. For each location a typical model-generated rainfall series gave a good visual correspondence to the actual series, and resulted in similar estimated aquatic habitat dynamics (Figure 5.5). A more formal comparison is given in Table 5.2, showing that the model is able to replicate a range

3378 of important rainfall characteristics, as described by Bárdossy, Stehlik, and Caspary, 2001.

3379 Simulations of a mosquito population in each of these locations typically end in pop-
3380 ulation extinction during the dry season (Bamako) or a relative dry spell (Mbita) in the
3381 absence of permanent aquatic habitats. The issue of how mosquito populations survive
3382 dry periods remains controversial, with aestivation and long-range migration possible ex-
3383 planations (Dao et al., 2014). Since neither of these possibilities occur in our model, it is
3384 unsurprising that the simulations are unable to replicate dry season survival; particularly
3385 for the case of Bamako which experiences more annual variability. In order to investigate
3386 HEG releases (in Bamako and Mbita) without modelling the uncertain dry season ecol-
3387 ogy, we focus on environments with permanent aquatic habitats (representing riverside or
3388 lakeside pools) as well as ephemeral aquatic habitats subject to rainfall (Figure 5.6). To
3389 construct Figure 5.6, we assumed that aquatic habitats were 65% from variable sources for
3390 Bamako, and 80% variable for Mbita. In both cases these were the maximum percentage
3391 of variable aquatic habitats that could sustain wild-type populations at least 99% of the
3392 time, and still allowed a reasonable running time.

3393 For Bamako, there is a clear benefit of releasing during the first half of the wet season,
3394 for landscapes with low, or high aquatic habitat densities (red and purple lines in top-left
3395 panel of Figure 5.6). The relatively large effect for the lowest aquatic habitat density land-
3396 scape (purple line), was due to a combination of increased occurrence of extinction due
3397 to a HEG partially-invading the population, together with increased fixation probability
3398 (data not shown). This contrasted with the high aquatic habitat case (red line), where
3399 the incremental probability of extinction during June-August was due to increased fixation
3400 probability, and a higher chance that a fixed HEG induces extinction. The different mech-
3401 anisms driving the extinction dynamics account for the slight difference in optimal time of
3402 year of release for the low and high aquatic habitat landscapes. Interestingly, landscapes

with moderate aquatic habitat densities did not show a noticeable effect of release time on the probability of population extinction, although compared to the other cases there was more variation in the time taken for extinction to occur (data not shown). By releasing in March, before the wet season begun, the average time for extinction occur was c.1.5 years, whereas releasing in the previous month resulted in the HEG taking roughly twice as long to drive population extinction. This sensitivity of time of extinction to release month is due to the importance of a HEG establishing in its first wet season. March releases tended to produce a HEG that established during the coming wet season, whereas for February this establishment was delayed until the next years wet season; in both cases extinction often occurring during the first dry season following the wet.

The effect of time of year of HEG release was much less marked for Mbita, due to the absence of a marked dry season for this location (Figure 5.6 top-right panel). In contrast to Bamako, the release time which resulted in the shortest time to population extinction was the months before the short rains; March-April for medium-high aquatic habitat densities, and January for the landscape with the lowest aquatic habitat density considered. Although, again this effect was weaker than for Bamako, where a long dry season causes a more pronounced effect.

In the event of planning a HEG release, an important consideration will be the speed of effect, as well as the effect itself. It is preferable for the HEG to act as quickly as possible, both for operational reasons (the sooner to reduce the prevalence of malaria) and practical reasons (a faster acting HEG gives less opportunity for HEG resistance to evolve). Across the landscape qualities investigated here, there was a lack of consensus over the release time which resulted in the fastest-acting HEG. However, in the case of Bamako across some of the parameter range, this optimum did not coincide with the release time that resulted in the greatest chance of population extinction. For example, in the case of the

lowest aquatic habitat density landscapes, releases at the start of the wet season (in June) resulted in the highest likelihood of population extinction, whereas releases in the middle of the dry season (December-January) resulted in the quickest eradication. This effect can be explained by the fact that releases during the dry season are usually unsuccessful, but if the HEG does manage to fix, an extinction may follow soon after.

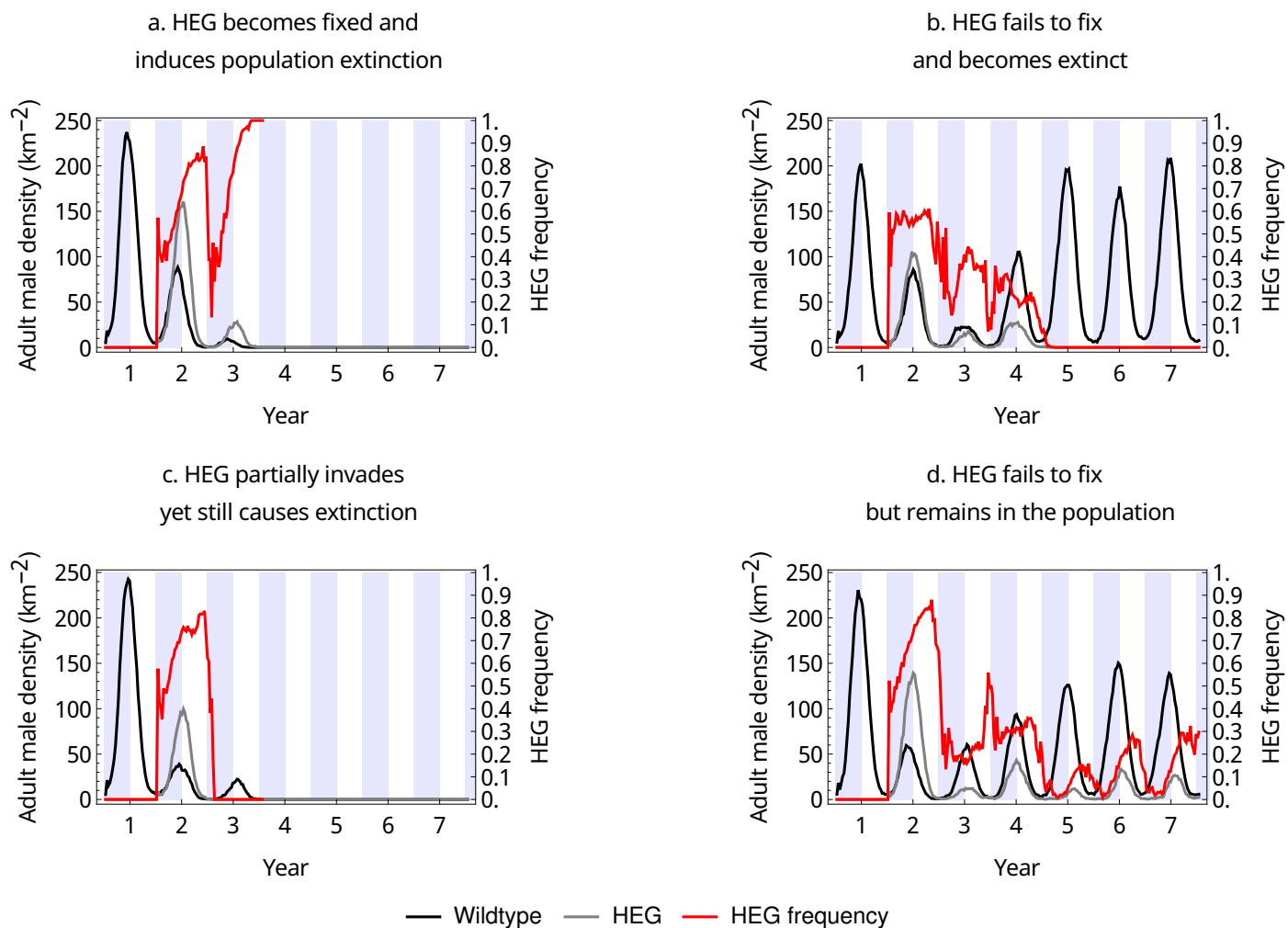


Figure 5.1: **Four examples of the outcomes following a HEG release, for adult population densities of wild-type and HEGs (black and grey lines), and frequencies (red lines).** Blue shading indicates wet seasons. Demographic parameters are set at values as described in Table 5.1. The mean aquatic habitat density was set at $\kappa = 142 \text{ km}^{-2}$ with seasonal variation about this level as described in the text.

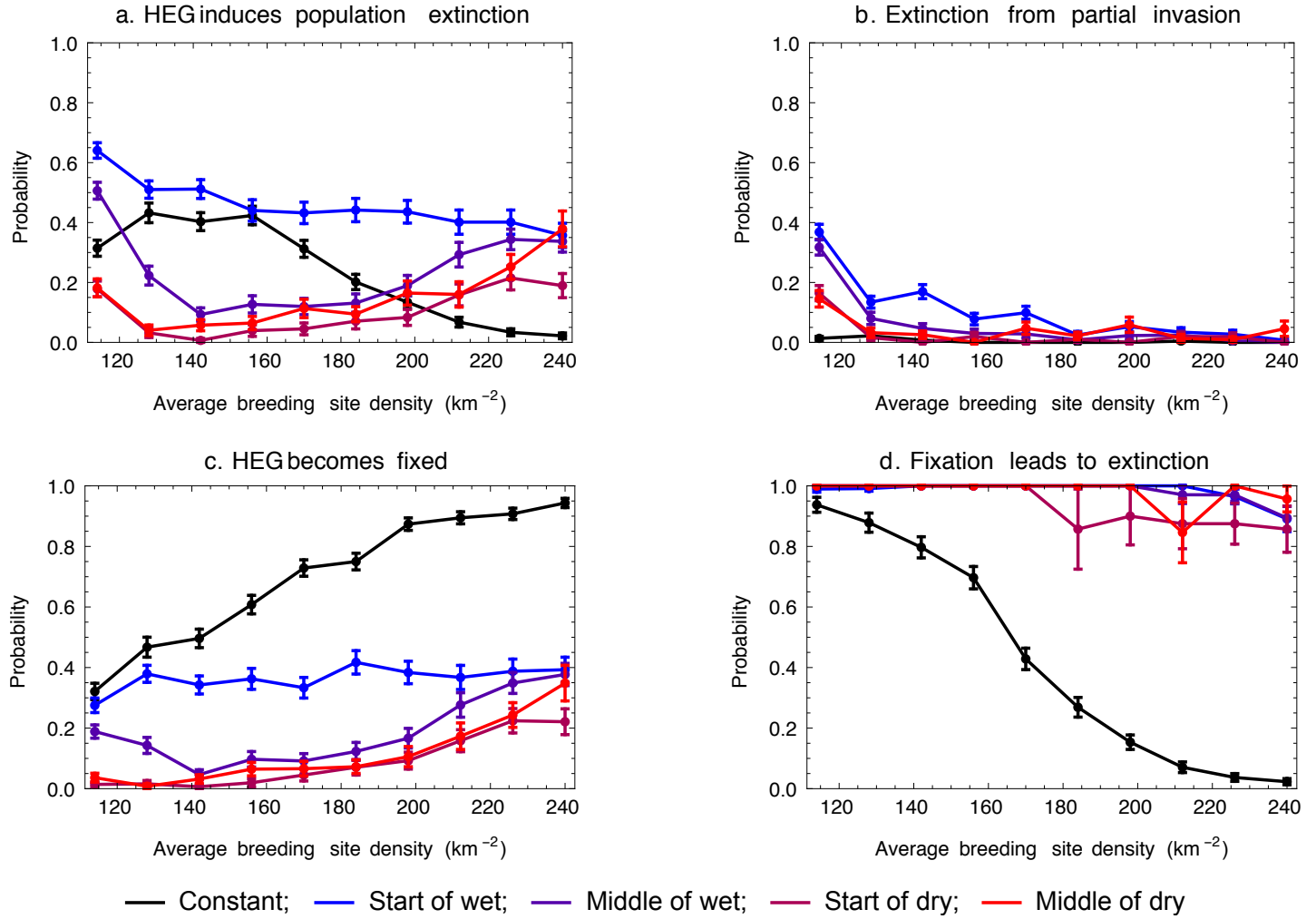


Figure 5.2: The impact of release timing (determined by line colouration) in an environment with an annual rainfall cycle, and in a constant environment (black), as a function of average aquatic habitat density. a. The unconditional probability that the HEG causes population extinction, which is decomposed ($a = b + c \times d$) into, b. the probability that a HEG that does not spread to fixation induces a population extinction, and the probability that a fixed HEG induces extinction ($c \times d$). The error bars indicate 95% confidence intervals based on a normal approximation with c.200 simulation runs. Demographic parameters are set at values as described in Table 5.1.

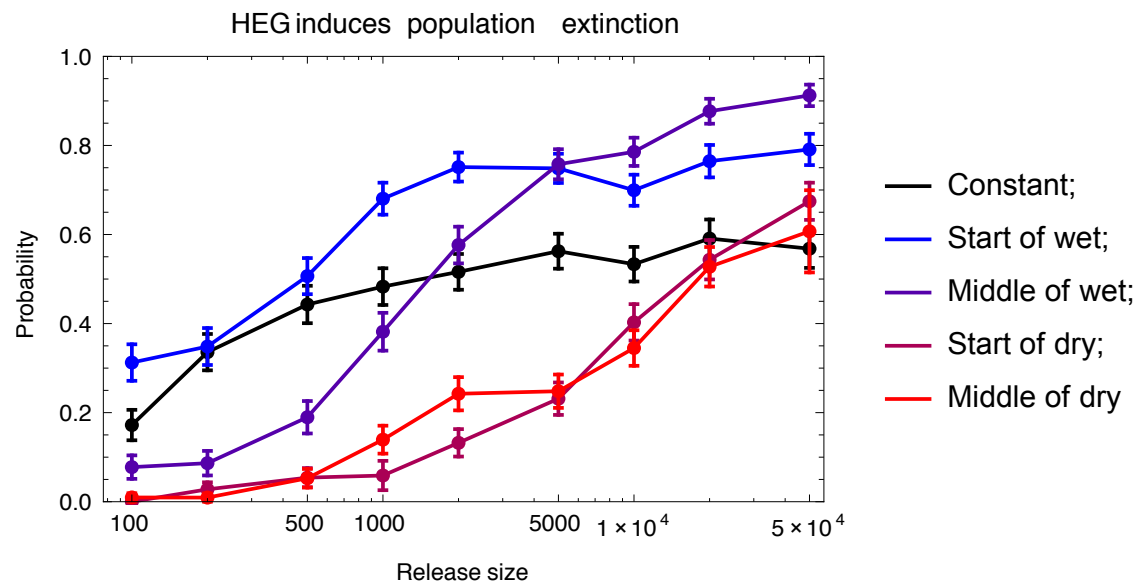


Figure 5.3: **The effect of release size (the number of HEG bearing males introduced) on the probability of population extinction.** The parameters and colouration follows Figure 5.2, with average aquatic habitat density = 128 km^{-2} .

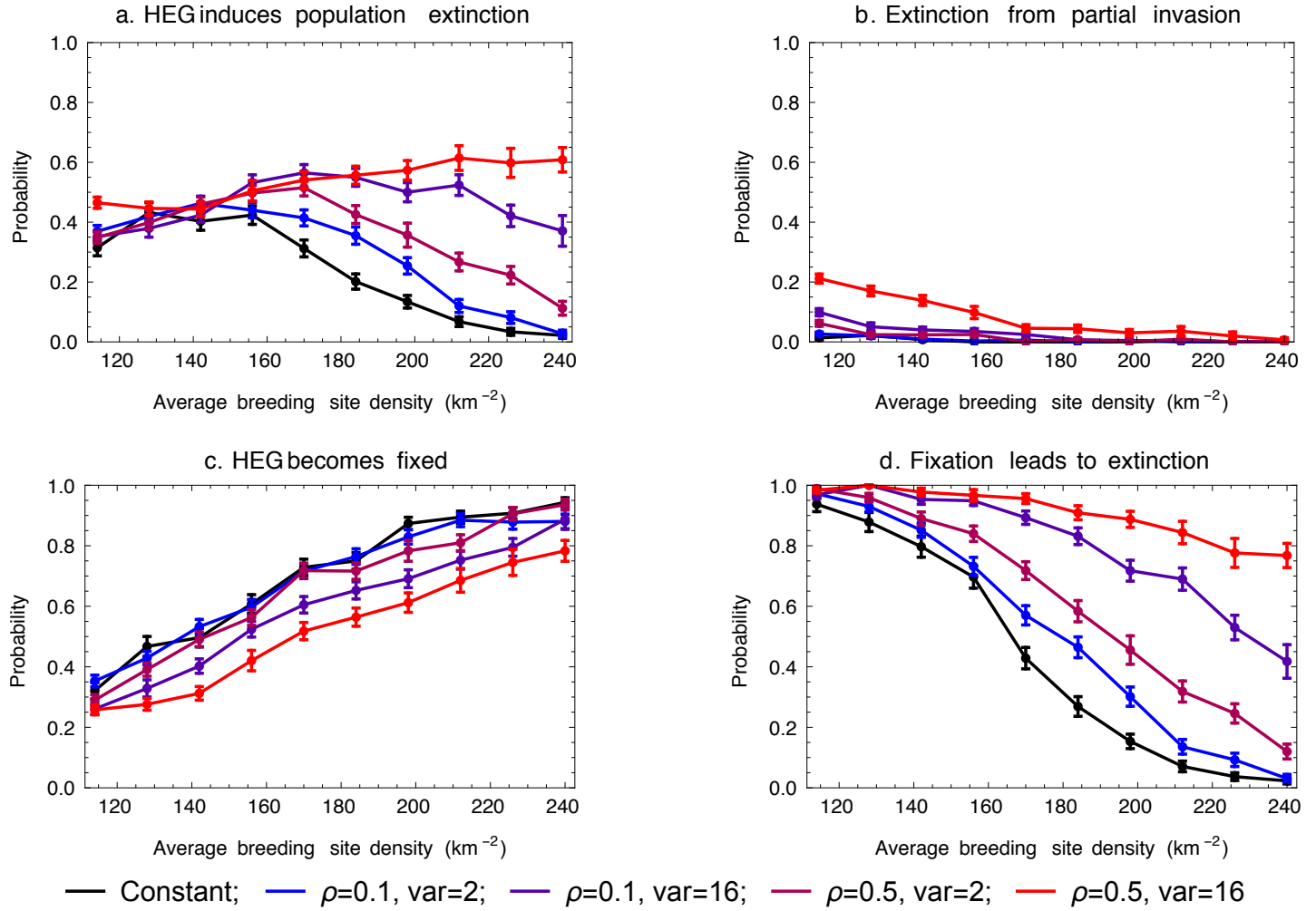


Figure 5.4: The impact of HEG releases in environments with stochastic variability in rainfall (coloured lines), and in a constant environment (black), as a function of average aquatic habitat density. The simulated rainfall series were constructed with low (blue and purple) or high (magenta and red) levels of autocorrelation, and with low (blue and magenta) or high (purple and red) levels of variance. The error bars indicate 95% confidence intervals based on a normal approximation with c.200 simulation runs. Demographic parameters are set at values as described in Table 5.1.

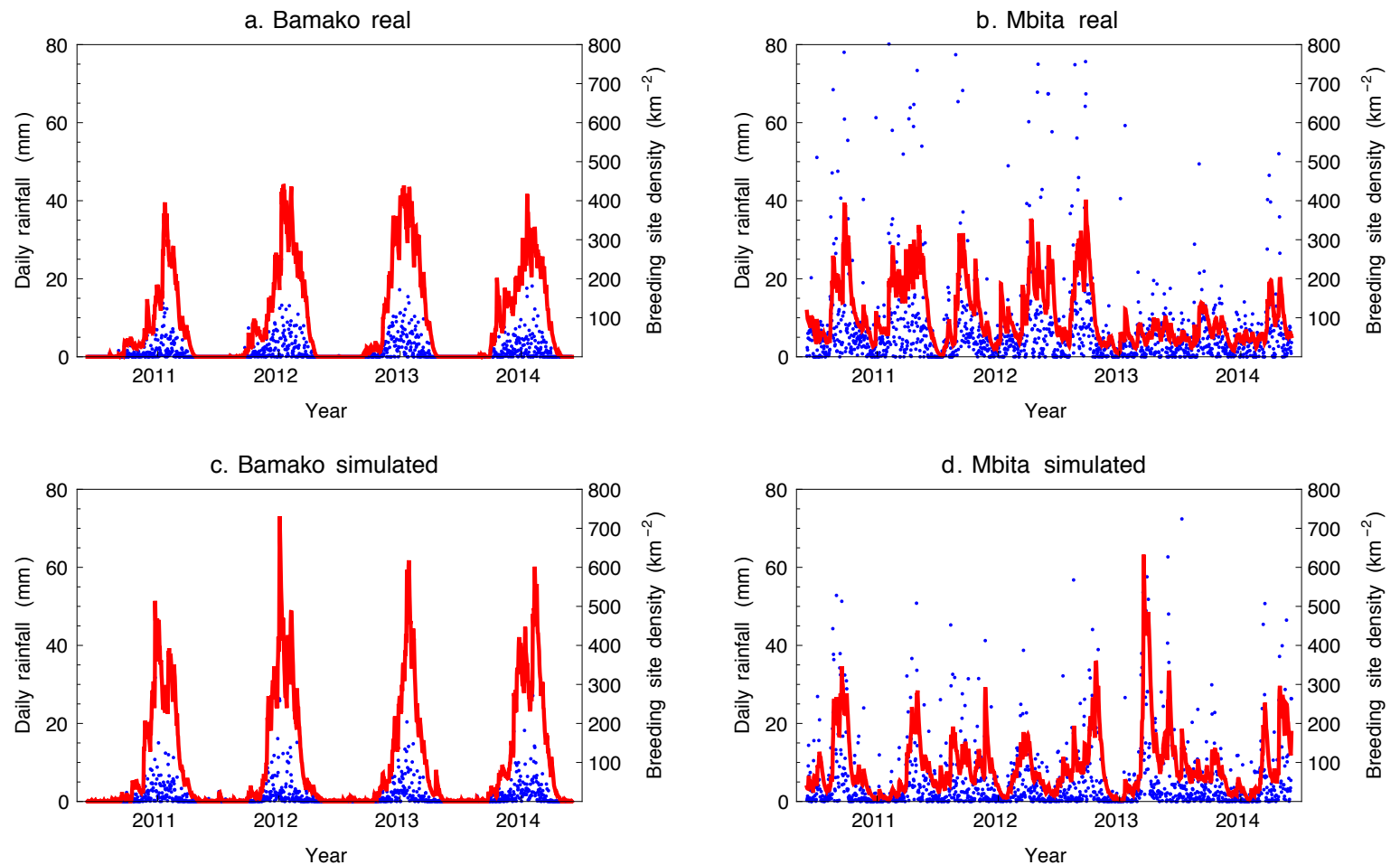


Figure 5.5: Rainfall data from Bamako, Mali and Mbita, Kenya (blue dots a. and c.), and corresponding time-series simulated from the rainfall model with fitted parameters (b. and d.) The red lines plot the aquatic habitat densities computed from the rainfall time-series by the method outlined in the text, with $\chi = 0.1$ and $\kappa = 100$.

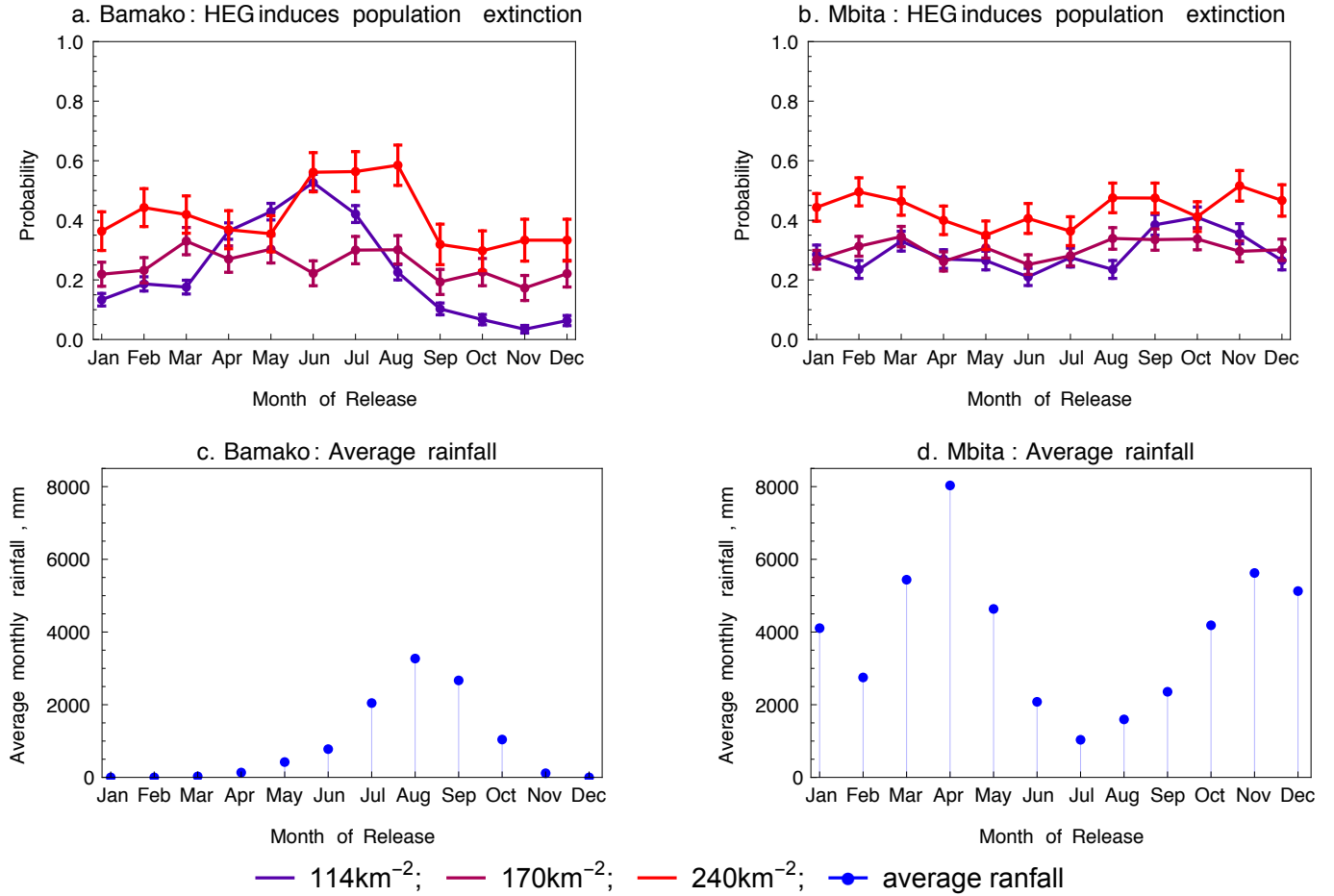


Figure 5.6: **Simulated HEG releases in Bamako and Mbita (a. and b.)** The probability of extinction depends on the average density of aquatic habitats in the environment (determined by line colouration) and the month of release. Graphs c. and d. show the average monthly rainfall over the 20 years of data we collected for each location. Demographic parameters are set at values as described in Table 5.1. The error bars indicate 95% confidence intervals based on a normal approximation with c.200 simulation runs.

5.5 Discussion

The use of HEGs to suppress vector populations holds clear potential as a tool for suppressing malaria. Mathematical modelling plays a key role in developing this technology by enabling investigation of how HEGs will perform in mosquito populations that are shaped by complex factors such as genetic diversity and environmental heterogeneity. In this paper we have focussed on rainfall because this is a key factor underpinning the population dynamics of malaria vectors, yet one that has previously received little attention. Our results suggest that HEG-based programmes can be successful in a wide range of climates, although the timing of HEG-release may be important in environments where rainfall is highly variable.

In general, HEGs should be released at the start of wet periods when the target species is growing at the fastest rate, since this maximises the probability that the genetic construct will successfully establish to become fixed in the population. In highly seasonal environments, where the wet season is more predictable, applying this rule may be marginally easier, although often considerable uncertainty remains over the start of the rains. In environments where wet weather is less predictable, it may be harder to plan releases to coincide with wet periods. For environments with less predictable weather, a bet-hedging strategy of few and often releases will be preferable to occasional large releases. Furthermore, releasing small batches of mosquitoes relatively often may be easier from an operational standpoint, since it may be harder to rear large populations of modified insects in individual insectories. In addition, releases could be focussed on locations where there is permanent breeding habitat and so the target population is more stable. Although variability in rainfall will increase the difficulty of establishing a HEG, it may also increase the impact of HEGs that do establish. Our results show an established HEG will be more likely to extinguish a target population if rainfall is variable, because such a population

3458 must persist through dry periods when the load imposed by the HEG is more likely to be
3459 fatal.

3460 The strong selection pressure for the mosquitoes to evolve a defence against the HEG's
3461 action needs to be taken into account when designing a release. This also depends on
3462 the specific nature of the mutation that could arise; to avoid spontaneous mutations, it
3463 is important to minimise the number of mosquitoes alive between release and extinction,
3464 whereas for HEG-induced mutations more importance should be given to the total number
3465 of transgenic mosquitoes. Time taken for eradication to occur is also of practical impor-
3466 tance, since shorter periods would result in a faster decline in vectorial capacity, more
3467 rapidly reducing the malaria burden within a region. Here we used the time taken between
3468 release and extinction both as a proxy for the probability that resistance arises, and as a
3469 measure of suppression of malaria. However, future analyses should extend the model to
3470 explicitly include mutations (see Deredec, Burt, and Godfray, 2008 for a spatially-averaged
3471 model of resistance for HEGs). It is also possible that even before a HEG is released, a
3472 minority of the wild population may have a natural immunity to the HEGs action. In the
3473 event of a HEG release, the (strong) selective advantage given to these mosquitoes would
3474 cause their population to grow rapidly. Whether population extinction would result in this
3475 circumstance, however, would depend on the degree of immunity conferred to an individual,
3476 the heritability of the defence, as well as its prevalence and spread in the population.

3477 The rainfall model developed in this paper has proved a useful tool for investigating
3478 effects of temporal variation on mosquito populations, and how this will influence HEG-
3479 based control . The strength of the model lies in its flexibility, whereby we have been
3480 able to reproduce a wide range of realistic patterns of rainfall with only a small number of
3481 parameters. The model may provide a useful framework for other studies that investigate
3482 the effects of temporal variation on vector biology, or more generally the relationships

between malaria and climatic variability.

In the Sahel region of Africa, as exemplified here by Bamako, seasonality is marked (Xue et al., 2004). Our model thus predicts that HEG releases occurring at the start of the wet season will be particularly effective in this region, resulting in extinction of the target population. It is beyond the scope of the current model to quantify the difference in impact between wet and dry season releases at a particular location given that many complex facets of mosquito behaviour are ignored. In particular, by assuming that mosquito behaviour is the same all year round, the model predicts that a population in Bamako will become extinct each dry season unless there are permanent aquatic habitats (even before HEGs are introduced). Although it is well documented that *Anopheles gambiae s.l.* mosquitoes become hard to find during a Sahelian dry season (Touré et al., 1998; Taylor et al., 2001; Dao et al., 2014), they invariably become common soon after each wet season begins, suggesting that total collapses do not occur.

The suggested early wet season release for HEGs can be contrasted with the results of modelling other biological methods of disease control. *Wolbachia* are symbiotic bacteria that are found in many insect species, and spread through mosquito populations via a mechanism known as cytoplasmic incompatibility (Yen and Barr, 1971), which is currently being explored as a way of controlling dengue fever (Moreira et al., 2009; Bian et al., 2010), carried primarily by the *Ae. aegypti* mosquito. The introduced bacteria has been demonstrated to cause a reduction in vectorial capacity, either by lowering mosquito mean lifetime (McMeniman et al., 2009), or by inhibiting disease transmission (Moreira et al., 2009; Bian et al., 2010). In order to overcome the deleterious effects of carrying the bacteria and spread to the wider population, it is necessary that the released frequency of infected females exceeds a threshold (Hancock, Sinkins, and Godfray, 2011). Hence, it could be optimal to release at times when population densities are lowest, although Hancock,

3508 Sinkins, and Godfray, 2011 suggest that it may be preferable to release a larger number of
3509 mosquitoes during the wet season, in order to avoid the high mortality in the dry season.
3510 Similarly, in SIT (sterile insect technique) the load imposed on the population would be
3511 greatest during times of low population densities, although practically release timings are
3512 typically chosen to coincide with particular crop-growing seasons (Dyck, Hendrichs, and
3513 Robinson, 2006).

3514 The question of how *Anopheles* populations survive the dry season remains controver-
3515 sial, and is currently dominated by two competing hypotheses. The first contends that
3516 adult mosquitoes undergo a dry season diapause (aestivation) by hiding in (currently un-
3517 known) shelters while the second contends that large numbers of adults migrate, possibly
3518 over large distances, soon after the onset of each wet season to repopulate suitable habitats
3519 (Lehmann et al., 2010; Adamou et al., 2011; Dao et al., 2014). The former hypothesis is
3520 supported by mark-recapture observations (remarkably, Lehmann et al., 2010 discovered
3521 a female *A. gambiae* s.s. still alive over seven months after being marked in the previous
3522 wet season) and experimental evidence (Adamou et al., 2011). It is possible that different
3523 members of the *Anopheles s.l.* complex use different strategies; recent Mali-based studies
3524 have suggested that *A. coluzzii* aestivates while the closely related *A. gambiae* s.s. persists
3525 by long distance migration (Adamou et al., 2011; Dao et al., 2014).

3526 We chose not to incorporate either survival mechanism into our model due to the lack
3527 of consensus over which, if either, mechanism predominates, yet it is worth speculating
3528 how these behaviours might alter our conclusions. A target population that aestivates
3529 will essentially skip the dry season, and the population at the onset of a wet season will
3530 be a subset of the previous wet season population (those that survive aestivation). A
3531 HEG that becomes established during a wet season will therefore remain present in the
3532 following year, and so will continue to suppress the target population. The overall extent

of suppression may be less than predicted by the current model because the role of the dry season is reduced, yet the general dynamics will be similar. If dry season survival is due to long distance migration, on the other hand, the population will be replaced each year and a HEG that establishes in one season will be unlikely to persist. In this case, it may be necessary to release HEGs annually to guarantee persistent population suppression. However, it may also be possible to disseminate HEGs across a landscape by their release in a population that becomes a source of migration, thus knowledge of *Anopheles gambiae s.l.* migration patterns will be invaluable.

The clearly important role of dry season behaviour highlights the value of research in mosquito ecology to assessing HEGs as a vector-control intervention. Before HEGs might be deployed, it will be important to develop detailed site-specific models to predict the outcome of initial releases as accurately as possible. This work will clearly benefit from further research on the dry season behaviour of *Anopheles gambiae s.l.* in Sahelian Africa, and also from further study into the effects of weather in other locations where releases may take place. Field studies that shed light on the link between rainfall and breeding habitat will likewise be extremely helpful. Whilst in our modelling here rainfall is taken always to increase mosquito population sizes, it has been suggested that flood waters may actually wash away larvae, reducing the availability of viable aquatic habitats (Reiner et al., 2015). Temperature is known also to affect survival in a nonlinear way; if too hot or too cold, there is increased mortality (Craig, Snow, and Sauer, 1999), and has led to temperature being used a predictor variable in generating maps of malaria prevalence as well as its inclusion in the many studies which investigate the climatic determinants of vector population dynamics (Reiner et al., 2015). Local geology and species-specific preferences for habitat may also be important considerations for planning releases, and further research into these areas would be beneficial.

5.6 Conclusion

I have demonstrated that mosquito population control using a driving homing endonuclease gene (DEG) that causes severe sex ratio distortion can be a viable strategy in a highly seasonal environment, and that the precise timing of releases during the year can be important in maximising success. In particular, our simulation results indicate that releases at the start of the wet season are preferable for maximising the probability that the wild-type mosquito population is extinguished. In practice, however, operational aspects of rearing and releasing batches of mosquitoes will influence the release timing which can be achieved. In harder to reach locations, it may not be possible to locally release DEGs and, therefore, it will be important to develop models for larger regions of the environment - which account for environmental and human ecological aspects of geography - which include models of weather, to determine optimal release strategies. How informative the result of these simulations are depends on the assumptions made about local mosquito ecology, however, and so emphasises the continued importance of field entomology.

I finish by noting that though the use of DEGs and other gene-editing strategies is a novel technology with great promise, it is essential that a rigorous governance framework is set up to regulate its use, both to avoid any negative health or environmental impacts and to assure a licence to operate by civil society in the countries where it may be employed.

Chapter 6

Conclusion

From Macdonald's mathematical model for the transmission of malaria, we know that mosquito borne diseases are most sensitive to the duration of the lifespan of mosquitoes. This is because mosquitoes that live longer (1) have more opportunity to bite an infected host and become infected themselves, (2) have a greater chance of surviving the time it takes for the pathogenic organism to develop inside the mosquito, such that the mosquito becomes infectious, and (3) have more time to bite an uninfected human host and transmit the disease. In this thesis, I presented the results from meta-analyses of previously-published data from the two predominant approaches for estimating wild mosquito longevity: mark-release-recapture (MRR) experiments (in Chapter 2), and studies that involve the dissection of female mosquitoes (in Chapter 3). The evidence from both approaches suggests that the majority of mosquitoes live for less than 10 days. Whilst these estimates likely represent lower bounds on the true lifespan of wild mosquitoes (see below), their correspondence indicates that mosquitoes may live shorter lives than is generally believed. If we take the estimates as they stand, this implies that only a small minority of mosquitoes live long enough to vector the major mosquito borne diseases. If this is the case, then disease

transmission should be particularly sensitive to interventions that kill adult mosquitoes, since small changes in adult mortality would drastically alter the number of mosquitoes surviving the EIP to become infectious. Indeed, this was the main conclusion drawn from the Ross-Macdonald model used to justify indoor residual spraying during the Global Malaria Eradication Programme (Smith et al., 2012).

There are reasons to believe that both MRR and dissection-based experiments likely underestimate the lifespan of mosquitoes, however. In MRR experiments, the handling of mosquitoes may damage them, and since the experimental data I analysed did not include information about where the recaptures occurred, the estimates represent the combined effects of mortality and dispersal of mosquitoes out of the study area, resulting in conservative estimates of lifespan. In dissection-based studies, it is assumed that the count of dilations in a particular ovariole faithfully represents the number of gonotrophic cycles that a female mosquito has undergone; an ovariole where this is the case is termed *diagnostic*. Previous work indicates that the proportion of diagnostic ovarioles decreases as a mosquito gets older, due to a range of physiological processes that destroy the ovariole dilations required to age mosquitoes by this method. This would mean that older mosquitoes require dissections of a large number of ovarioles to locate those few with the greatest number of dilations, which faithfully represent the gonotrophic age of a mosquito. Whilst the various entomologists who carried out the dissection experiments were probably aware of this issue, it is possible, and has been shown to frequently be the case (Hugo et al., 2008), that the parous age of many mosquito specimens is less than the true gonotrophic age. Overall, the assumptions of the two methods suggest that the estimates I present are best viewed as lower bounds on mosquito lifespan.

Despite the different assumptions made by each method, it is encouraging that those species and genera where observations were available from both sources had comparable

estimates of lifespan. Where discrepancies exist, however, I tentatively suggest that some of the differences in lifespan may be attributable to the variation in dispersal across mosquito species or genera. Specifically, because the estimates from MRR studies effectively include the rate of dispersal whereas dissection-based estimates do not, it is possible MRR study estimates may be relatively downward-biased. Indeed, across the majority of species and genera, the estimates from the dissection studies were greater. At the genus level, the greatest difference was obtained for *Culex* mosquitoes followed by *Anopheles* then *Aedes* (where the estimates from the MRR studies were actually higher). It is thought that aedine mosquitoes, which live in close proximity to humans often in urban settings, disperse the least during their lifetimes, explaining why this genus was associated with the smallest difference in estimates from both approaches. Further work is required, however, to confirm the variation in dispersal of species and genera.

Across both classes of experiment, there was evidence that mosquito lifespan varies between species and genera. This is perhaps not a surprising result given that many mosquito species are found in relatively confined geographic areas, and that mosquitoes will inevitably adapt to their local environment. It should be noted, however, that within most species, the variability in lifespan equalled, and in many cases, exceeded, the variability between species. This was particularly pronounced for those species, for example, *Ae. aegypti* and *A. gambiae*, which are found in a variety of geographies. On this basis, I argue that it is likely adaptation to local climate and other environmental pressures that drives the observed differences in lifespan. Whilst I investigated air temperature as a potential source of this variability (and found no systematic pattern), I believe that other factors, such as rainfall, may be more predictive of lifespan.

Laboratory colonies of mosquitoes are essential for studying mosquito borne diseases, yet we know that these mosquitoes and their environmental conditions are considerably

different to those in the wild. This is typified by the differences in mosquito lifespan for laboratory and wild mosquitoes, where those in captivity regularly live for many multiples of the estimated lifespan of their wild counterparts. Some of these differences are likely due to biases in the methods used to estimate wild mosquito lifespan, yet these do not explain the bulk of the discrepancy. Predation, the risk of mortality from blood-feeding and a lack of food, amongst other factors present in wild environments, are likely major contributors to these differences, yet our current knowledge of mosquito ecology is insufficient to conclude which of these predominates. These differences in lifespan, I believe, highlight the issues inherent with extrapolating conclusions drawn from laboratory colonies to the wild, and I argue that a number of such published results (including those concerning age-dependent mortality) may not generalise well to the field. Whilst more expensive, and constrained in their use to those geographies that naturally support mosquito populations, I argue that semi-field experiments - where many of the conditions are similar to wild environments - will be of increasing importance for the study of mosquito borne disease. Just how representative these semi-field environments are of the wild is currently unclear, however, and further work is needed to justify their added expense. Despite this, it is encouraging to see that these studies appear to be becoming more prevalent in recent years (see Hugo et al., 2014 for an example of an experiment used to estimate mosquito mortality in Vietnam). Additionally, it is important to have methods for determining characteristics of mosquito ecology from wild caught specimens, and I believe that MRRs and dissection studies will continue to play an important role in this area. Yet, it is still desirable to develop new methodologies which are simpler to implement than these traditional approaches, and currently there are a number of promising approaches which are subject to current research.

In Chapter 4, I discussed one such approach to estimating mosquito lifespan known as *near-infrared spectroscopy* (NIRS). In this approach, light is shined through the head and

thorax of mosquitoes, and the adsorption across a range of wavelengths is measured. As a mosquito gets older, its biochemistry changes, causing changes to its infrared spectrum. Previous work has demonstrated that these changes to the spectrum are regular enough that machine learning algorithms can be trained on them and used to predict the ages of individual mosquitoes. In this chapter, I applied a more recently developed machine learning algorithm than has been used previously, to this task. The initial results from using this approach showed a significant improvement in the average errors at the individual level, but (like previous work) produced bias estimates of lifespan, with over-prediction for younger mosquitoes and under-prediction for older individuals. Since this bias was systematic, by modifying the machine learning algorithm, I was able to obtain an unbiased estimator of individual mosquito lifespan.

A potential use of NIRS in the field could be to detect changes in the average lifespan of the local mosquito population in response to vector control measures, such as the deployment of insecticide treated bednets and the indoor spraying of insecticide. In Chapter 4, I demonstrated that NIRS applied to a modest number of caught mosquitoes can reliably detect differences in the average lifespan of a population, along those age ranges thought to be realistic. Further, because the bulk of the errors in population lifespan were due to sampling error (the random variation in mosquitoes caught by collections), NIRS was found to be relatively insensitive to inflations in the individual prediction errors, so long as these estimates remain unbiased.

There is reason to believe, however, that the errors obtained from applying NIRS in practice will not be of qualitatively different form to those I have assumed here. In Chapter 4, I demonstrated that NIRS calibrated on data from one experiment cannot predict the ages of individual mosquito specimens scanned in another. There is insufficient data to determine the source of the inter-experiment heterogeneity, but it is possible that this is

3693 due to differences in the spectrometer, the method used to preserve the caught mosquitoes,
3694 and the genetics of the specimens, amongst other factors. If the bulk of this variation is
3695 due to methodological differences, then it should be possible to establish protocols which
3696 minimise this effect, and NIRS could be straightforwardly used to survey wild populations.
3697 If the source of the variation is due to non-experimental heterogeneities, such as species and
3698 geography, then each application of NIRS will require its own calibration dataset composed
3699 of local mosquitoes (possibly from wild larvae raised in a laboratory) of known ages. Whilst
3700 this necessity would diminish the worth of NIRS, it could, nevertheless, be a potentially
3701 useful tool for field entomology.

3702 In Chapters 2 and 3, I compared the fit of a number of different models for mosquito
3703 mortality to data. One of these models - the exponential distribution - assumes a constant
3704 (i.e. age independent) risk of mortality, whereas the other five models I used allowed for a
3705 mortality risk that increased with age. For the MRR experiments, the exponential model
3706 fitted the data as well as the other five models across each of the species in the dataset,
3707 indicating that we cannot reject the hypothesis that wild mosquitoes do not experience an
3708 elevated mortality risk as they age. For the dissection experiments, the results were more
3709 mixed. In nearly half of the species, the exponential model fit the data as well as or better
3710 than the best performing of the five models which incorporate senescence. In the rest of the
3711 species, the exponential did not provide as good a fit to the data as the best of the other
3712 five, however, although in some of these cases, the exponential model still performed better
3713 than one or more of the models with age-dependence. Collectively, I believe these results
3714 indicate that mosquito senescence, should it occur in the wild, is likely a subtle effect;
3715 considerably smaller than suggested from experiments on laboratory colonies (for example,
3716 see Styer et al., 2007). Presumably, in the wild, environmental factors (as discussed above)
3717 result in the premature death of mosquitoes before they get to ages where they experience

senility. Again, experiments which indicate the relative importance of these factors would allow a better understanding of these differences and, perhaps, suggest novel methods for vector control. Due to the subtlety of mosquito senescence, I argue that differences in the modelled impact of vector control using the Ross-Macdonald model and those incorporating senescence may be less major than is sometimes suggested.

Widespread use of existing vector control methods, including most notably, insecticide treated bednets and the spraying of insecticides indoors, has likely been a major contributor to the recent gains against malaria (Bhatt et al., 2015). The growing prevalence of insecticide resistance amongst the major disease vectors, however, threatens to halt and, perhaps, reverse these trends (Ranson et al., 2011). Similarly, the spread of parasites resistant to artemisinin - an important drug used to treat human malaria - in Asia (Ashley et al., 2014; Imwong et al., 2017), is extremely concerning and suggests a greater emphasis on finding new ways to control the disease. A recently proposed approach involves using genetic drive technologies to spread traits which impose a fitness load on a mosquito vector population, leading to declines in its numbers and, potentially, its elimination. In Chapter 5, I investigate the use of one such technology, known as a homing endonuclease (HEG), which has been engineered to bias the sex ratio of mosquito progeny towards males. As a so-called Y-drive HEG spreads through the population, it should become increasingly composed of males, leading to an eventual population collapse when the few remaining females die. Previous modelling work has investigated the release of these HEGs in both panmictic (Burt, 2003; Deredec, Burt, and Godfray, 2008; Deredec, Godfray, and Burt, 2011) and spatially heterogeneous (North, Burt, and Godfray, 2013) environments. In this chapter, I build on this work by incorporating a rainfall model which is linked to the availability of aquatic habitats in a previously-published spatial model of mosquito ecology (North, Burt, and Godfray, 2013), and using it to investigate the release of HEGs in environments

with temporal heterogeneity. This work demonstrates that increased heterogeneity can detrimentally impact the ability of a HEG to establish in a population, but that this effect is more often than not compensated for by the population's elevated vulnerability. This means that, across much of parameter space, temporal variation in the environment can actually boost the performance of a HEG release, meaning it more often leads to population elimination. This suggests that HEGs, or more generally driving endonuclease mechanisms (DEGs), show great potential for vector control in the field. Encouragingly, a number of DEGs have been successfully inserted into laboratory populations of mosquitoes (Windbichler et al., 2011; Gantz et al., 2015; Hammond et al., 2016) and a consortium called Target Malaria lead by Austin Burt (the first to propose using HEGs for vector control) has been actively involved in evaluating the potential use of such technologies for vector control.

The model I used in Chapter 5 assumes that resistance to DEGs does not arise over the time scales it takes for a DEG to suppress mosquito population densities. We know, however, that there exist mechanisms that could give rise to resistant strains of mosquitoes. For example, it is possible that the DNA sequence that a DEG targets may not be conserved across all individuals in a wild population. After the release of a DEG, any individuals lacking this recognition sequence would, hence, have a selective advantage, and likely spread through the population. Similarly, chromosomal repair can occur by non-homologous mechanisms such as end joining, which could generate alleles which did not contain the requisite recognition sequence. Regardless of the source of resistant alleles, its spread will be dictated by the other fitness costs associated with the mutation. Modelling work has determined the characteristics of the engineered DEG which determine the population dynamics associated with resistance in non-spatial environments (Deredec, Burt, and Godfray, 2008; Godfray, North, and Burt, 2017), and attempts to model the spread of

a DEG with resistance in spatially explicit environments are now being actively pursued. It is hoped that the modelling can inform the work of biological teams who are responsible for engineering the DEGs and inserting them into the mosquito genome, leading to a technology that avoids a number of sources of resistance.

Others have argued that the gaps in our knowledge of mosquito ecology are significant barriers to disease eradication (Godfray, 2013; Ferguson et al., 2010a), and it is my hope that this thesis helps to slightly narrow the gap associated with mosquito survival. It is clear, however, that there are important areas of mosquito ecology and behaviour about which we know little, and I now speculate on those areas which I believe may prove fruitful in the future.

Whilst it is generally accepted that there is substantial variation in the ecology of different species of mosquito, I believe that there is less recognition of the variation within a given species. For example, the use of insecticide treated bednets and the indoor spraying of insecticides for malaria control is based on the stereotypical perspective that vectors bite indoors (at night) and rest on the inside walls of houses (Ferguson et al., 2010a). It is known, however, that endophilic species may sometimes feed outside, and may do so in response to interventions which target domestic behaviours (Pates and Curtis, 2005; Reddy et al., 2011). Whether this behaviour represents phenotypic plasticity or has underlying genetic causes, is currently not well known. Our view of putting mosquitoes in ecological and behavioural ‘boxes’, whilst historically useful, is now, perhaps, detrimental to the cause. In order to build mathematical models which can be used to accurately predict disease transmission following interventions, I believe that we need to move away from regarding a species’ behaviour as a discrete point in ecological space towards a distribution, whose spread is informed by experimental work. Characterising the covariances in mosquito ecological space between vector species is also of crucial importance. For

example, when interventions suppress endophilic species, this gives a selective advantage to exophilic vectors, which in some cases, may be equally capable of vectoring disease (a good example being the major malaria vector *A. arabiensis*). Understanding how vector species interact with predators and prey throughout their various life stages is also, of importance (Godfray, 2013). An intervention which effects one part of an ecosystem could cause an ecological disturbance resulting in another vector species arising in prominence (Ferguson et al., 2010a). Unless we have a reasonable understanding of these intra- and inter-species interactions, I believe, that computational models of mosquito ecology will likely overestimate the effects of interventions.

An element of mosquito ecology that I argue is currently poorly understood is the dispersive behaviour of mosquitoes. As well as being an important element of models of disease transmission, spatial dispersal is of particular importance to the design of DEG programmes (Godfray, North, and Burt, 2017), where it is hoped that modified mosquitoes in one area will seed the invasion of DEGs into other locations. Classic metapopulation theory suggests that the population dynamics which result from a DEG release are likely to be highly sensitive to the assumptions made about the spatial dispersal kernels of mosquitoes (Hanski, 1998). The spatial kernels determine the connectivity of patches (idealised discrete areas of the landscape) and, hence, determine the rate at which wild-type populations are invaded by DEGs, and also the rate at which empty patches are colonised by either DEGs or wild-type individuals. The meta-population dynamics also depends on any differences in dispersive fitness which exist between wild-types and DEG individuals. Clearly, if wild-type individuals are considerably fitter than the DEGs then it will be more difficult to eliminate wild-type populations. The spatial dispersal kernel not only affects the rate of DEG spread but also how resistant alleles, should they arise, will spread through wild populations. Again, the outcome of the invasion of the resistant allele will depend on its

relative fitness versus wild-type and DEG individuals.

Mark-release-recapture experiments remain our best source of information on the dispersal of mosquitoes (see, for example, Guerra et al., 2014). Yet, it remains unclear how to design these experiments to capture the requisite information, and the numbers of individuals that would need to be released to best estimate the tails of the kernels; likely to be of particular importance to the design of interventions. I am currently working with Ace North on a spatial model of MRR experiments which, we believe, may help to answer some of these questions. Of course, it may prove to be the case that the numbers of individuals required to probe the tails of the spatial kernels is beyond the remit of traditional experimental approaches. Godfray, 2013 has suggested that an alternative marking methodology would be to spray a large area of a landscape with harmless chemical elements, such as Rubidium, which have been demonstrated to be taken up by larvae and are retained into the adult stage (Wilkins et al., 2007).

Another part of mosquito ecology that is crucial to the spread of a DEG is the dry season behaviour of anopheline mosquitoes in those climates (for example, the Sahel region of Africa) where there are distinct wet and dry seasons. There are a number of hypothesis concerning how these mosquitoes seem to appear soon after the first rains fall, including aestivation (Lehmann et al., 2010), long-range migration (Dao et al., 2014) and ovipositing in the few water bodies that persist through the dry season. Yet, to date, there is not strong consensus on which of these, if any, can explain the observed population dynamics.

As well as the existing gaps in our knowledge of mosquito ecology, our manipulation of our environment brings new challenges. We are currently in the midst of the greatest period of urbanisation in human history. In 1910, 10% of the world's population resided in cities, whereas today there are more urbanites than non-urbanites, and the percentage of the population living in cities is expected to continue growing in the near future (UN,

2002). This unprecedented, and widespread change to our way of life dramatically changes the ecology of an environment (Grimm et al., 2008), and likely exerts potent evolutionary forces on species that dwell within cities (Partecke, Schwabl, and Gwinner, 2006). Others have argued that urbanisation of cities changes the ecological landscape enough to affect the transmission of mosquito borne disease (although differ in their proposed directionality of the effect on disease transmission; Keiser et al., 2004; Hay et al., 2005), but it is possible that the selective pressure exerted on mosquitoes may have deleterious consequences for us. In particular, it is possible that hitherto zoophilic mosquitoes might adapt their biting behaviour towards human sources - a ubiquitous source of food in cities. Alternatively, anthropophilic species that are accustomed to living in natural environments might adapt to oviposit and breed in close proximity to humans (for example, typical of the behaviour of *Ae. aegypti*). If an African anopheline species that liked urban environments and favoured biting humans arose, then projections that suggest malaria transmission will decrease (see, for example, (Hay et al., 2005)) would be wildly incorrect, and a new, urban, era of malaria might begin.

Malaria and many other vector borne diseases have plagued humans from a time pre-dating civilisation. Since the work of Manson, Ross, Lazear, and Reed, amongst others at the turn of the twentieth century put mosquitoes at the centre of transmission for these horrific illnesses, we have had considerable successes and, unfortunately, notable failures against these flying emissaries. The recent declines in malaria cases in those parts of the world most affected have, finally, given us some cause for optimism. It is important, however, that we do not now rest our laurels on the clay appendages of the opposition. There are many reasons to believe that malaria and other mosquito borne diseases will continue to affect those vulnerable many throughout the course of this century. A sinister recent development is the discovery in Cambodia of a single dominant artemisinin-resistant form

3868 of *P. falciparum*, which has now developed resistance to piperazine - the drug the country
3869 uses as a first line treatment for malaria (Imwong et al., 2017). The tragedy of the recent
3870 Zika epidemic for those worst affected parts of the world is unarguable and is, perhaps, the
3871 most serious infectious disease to threaten the Western world since HIV/AIDS invaded the
3872 United States in the 1980s. Even though this disease doesn't kill, we are just beginning
3873 to understand its impact and it hints at the real risk posed by emergent *Aedes*-borne dis-
3874 eases. If any good is to come from all this human suffering, it is through our - the West's
3875 - re-acquaintance with the ills carried by our mosquito foe which, I hope, leads to greater
3876 funds being allocated to its study.

Chapter 7

Appendix

7.1 A mathematical model of senescence

I now present a simple model to provide insight into the effect of mosquito senescence - that is age-related death - on vector control efforts. The model is similar in nature to that presented in Bellan, 2010, although here I use a continuous age structured model and explicitly include a larval population (incorporating density dependence). I use this model only to undertake an analysis of the statics of the system, in other words, when the populations are at their steady states, as this provides ample insight into the impact of an age-dependent mortality on vector control efforts. I leave the analysis of the dynamic behaviour of the system for future work.

Vectorial capacity is a measure of the transmission potential of a mosquito population and is often used to assess the impact of vector control efforts (Smith and McKenzie, 2004). Vectorial capacity is similar in nature to R_0 and represents the expected additional human infections that will be generated each day from a single infected human. This quantity is derived from the product of four terms: (1) the emergence rate of mosquitoes ϵ , (2) the

3893 average number of human bites for a single mosquito throughout its lifetime, S , representing
 3894 the initial biting of an infected human (for vectorial capacity calculations we assume the
 3895 mosquito always becomes infected), (3) the probability that a mosquito survives through
 3896 the entomological inoculation period (EIP), $Pr(A > n)$, so that it can transmit sporozoites
 3897 to an uninfected human in a subsequent bite, that (4) is again represented by S ,

$$C = \epsilon \times S \times Pr(A > n) \times S = \epsilon S^2 Pr(A > n). \quad (7.1)$$

3898 If mosquitoes do not experience the effects of old age then individual lifespan is exponen-
 3899 tially distributed and, hence, $Pr(A > n) = e^{-ng}$, where g is the (age-independent) force of
 3900 mortality, and n is the duration of the EIP. Further, for an exponential lifespan distribu-
 3901 tion, it is straightforward to calculate S , the average number of human bites a mosquito
 3902 takes in its lifetime; it is given by the product of the average lifespan of a mosquito, equal to
 3903 $\frac{1}{g}$, and the human biting rate, a , which measures the average number of bites per mosquito
 3904 per day,

$$S = \frac{a}{g}. \quad (7.2)$$

3905 If we assume that the mosquito population is at its steady state,

$$\frac{dM}{dt} = \epsilon - g\tilde{M} = 0, \quad (7.3)$$

3906 meaning $\epsilon = \tilde{M}g$, where $\tilde{M}$ is the steady state population density. Using this result we
 3907 obtain an alternative expression for vectorial capacity, which is the same as that derived
 3908 from the Ross-Macdonald model of malaria transmission (Macdonald, 1957),

$$C = \frac{\tilde{M}a^2e^{-ng}}{g}. \quad (7.4)$$

To derive an equivalent quantity for a population that experiences an age-dependent mortality risk, I consider a continuous time age-structured model of a population of adult mosquitoes (see, for example, Murray, 2002, for an introduction to these types of model). To enable an assessment of the effects of vector control approaches employing larvicide, I also include a larval population in the model. To keep things simple, I assume that the larval population does not experience age-dependent mortality and, hence, can be well-described by an age-averaged model. I assume that the larval population compete for resources and represent this density-dependence by a logistic term. The model system is of the form,

$$\frac{dL(t)}{dt} = fM(t) \left(1 - \frac{L(t)}{\kappa}\right) - \eta L(t) \quad (7.5)$$

$$\frac{\partial m(A, t)}{\partial A} + \frac{\partial m(A, t)}{\partial t} = -\mu(A)m(A, t) \quad (7.6)$$

where $L(t)$ is the density of the larval population, $m(A, t)$ is the density of the adult population of age A at time t , f is the fecundity of an adult mosquito (here I consider only a population of females and implicitly assume the presence of a male population), which is assumed to be independent of age, $M(t) = \int_0^\infty m(A, t)dA$ is the overall density of the adult population, κ is the carrying capacity of the larval population, η is the rate of recruitment of the larval population into the adult stages, and $\mu(A)$ is the force of adult mortality, that can depend on age. To close this system, I assume that those larvae recruited into the adult population enter it at age $A = 0$ (here A measures adult age),

$$m(0, t) = \eta L(t). \quad (7.7)$$

Considering the populations at their steady states yields the following relations,

$$f\tilde{M}\left(1 - \frac{\tilde{L}}{\kappa}\right) = \eta\tilde{L} \quad (7.8)$$

$$\frac{d\tilde{m}(A)}{dA} = -\mu(A)\tilde{m}(A), \quad (7.9)$$

where $\tilde{L}$, $\tilde{m}$ and $\tilde{M}$ represent steady state quantities. The second of these expressions can be solved by separation of variables to yield the steady state age structure of the adult population,

$$\tilde{m}(A) = \tilde{m}(0)e^{-\int_0^A \mu(\tau)d\tau} \quad (7.10)$$

$$= \eta\tilde{L}\underbrace{e^{-\int_0^A \mu(\tau)d\tau}}_{S(A)} \quad (7.11)$$

where $S(A)$ is the survival function, which indicates the proportion of the adult population surviving till age A . The steady state adult population density is calculated by integrating expression (7.11),

$$\tilde{M} = \eta\tilde{L} \int_0^\infty S(A)dA \quad (7.12)$$

$$= \eta\tilde{L}\bar{D} \quad (7.13)$$

where the second line follows because the integral of the survival function across all ages is the average lifespan of the population (Cox and Oakes, 1984), $\bar{D}$. By substituting expression (7.13) into expression (7.8), we can solve for the steady state larval population density,

$$\tilde{L} = \kappa \left[1 - \frac{1}{\bar{D}f} \right], \quad (7.14)$$

where we obtain a non-zero steady state larval density if $\bar{D}f > 1$, in other words, the average number of offspring produced by a mosquito during its lifespan exceeds one - sufficient to replace itself. The steady state number of larvae recruited into the adult population per day is, hence, given by,

$$\tilde{m}(0) = \eta\kappa \left[1 - \frac{1}{\bar{D}f} \right]. \quad (7.15)$$

I now follow Styer et al., 2007 and calculate the vectorial capacity of each age cohort (a collection of individuals of the same age). Its derivation follows the same approach that we used to calculate the overall vectorial capacity, and results in the following expression,

$$C(A) = a^2 \tilde{m}(A) e(A) Pr(t > A + n | t > A) \quad (7.16)$$

3917 where $e(A)$ is the expected remaining lifespan of a mosquito that has reached age A , and
 3918 $Pr(t > A + n | t > A)$ is the probability that this mosquito survives for a subsequent duration
 3919 of time equal to the EIP (n).

3920 In what follows, I suppose that adult mosquito lifespans can be described by the Gompertz distribution (see Chapter 2), which has an age-specific force of mortality equal to,

$$\mu(A) = de^{A\lambda} \quad (7.17)$$

3922 where d and λ are constant parameters that affect the shape of this distribution. When
 3923 $\lambda = 0$, the force of mortality is a constant (equal to d), independent of age, and we
 3924 obtain an exponentially-distributed lifespan distribution (blue line in Figure 7.1A). For an
 3925 exponential distribution for lifespans, the distribution of mosquito ages is also exponentially
 3926 distributed (with the same rate parameter), and the survival curve is also an exponential

function (blue line in Figure 7.1B). As λ is increased, the mosquitoes are more sensitive to the effect of old age, and the lifespan distribution becomes increasingly truncated towards older mosquitoes (orange and green lines in Figure 7.1A), and few mosquitoes survive much longer than the mean lifespan (orange and green lines in Figure 7.1B).

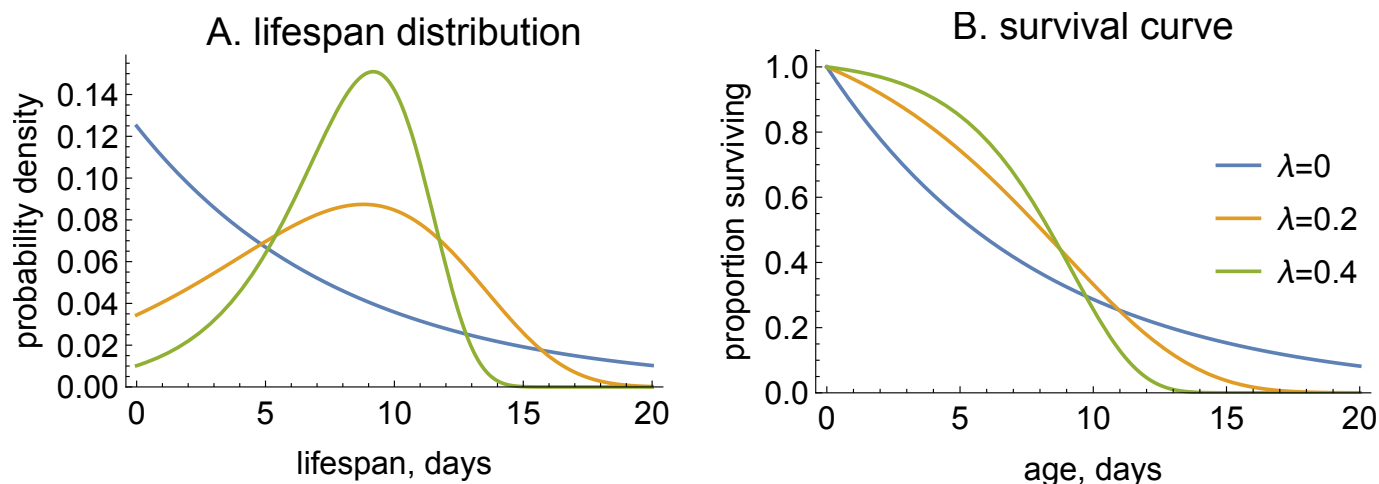


Figure 7.1: **The probability distribution for mosquito lifespan (left) and the survival curves for these mosquito populations (right) as a function of the Gompertz distribution's λ parameter.** In all cases, the populations have a mean lifespan of 8 days.

I now use expression (7.16) to illustrate how the vector capacity is affected by age-dependent mortality. To do this, I compare a population that experiences no age-related mortality (left column of Figure 7.2) with one that experiences significant age-related decay (right column). In particular, for the senescent population, I specify $\lambda = 0.2$, which corresponds to the lifespan distribution and survival curves shown by the orange lines in the left and right panels of Figure 7.1, respectively. I assume the mean lifespan (8 days) and EIP (10 days) are the same in both cases. Expression (7.16) shows that the age-specific

vector capacity is the product of three things (in what follows, I assume $a = 1$ since this does not affect the reasoning): (1) the density of mosquitoes at a given age (top row of Figure 7.2), (2) the residual life expectancy for a mosquito of that age (second row from top) and, (3) the probability that the mosquito subsequently survives for a duration equal to the EIP (second row from bottom). The bottom row in Figure 7.2 shows the age-specific vectorial capacity (solid lines) and the overall vectorial capacity (shaded areas) for both populations.

The most significant differences between the elements of vectorial capacity for the two populations are the residual life expectancy (second row from top of Figure 7.2) and the probability of surviving the EIP (second row from bottom). For the population that does not experience senescence, these two quantities are independent of age (left column), due to the memoryless property of the exponential distribution. Whereas for the senescent population (right column), both of these quantities decrease with age, with the probability of surviving the EIP being particularly sensitive. The result of this age-dependence is that the vectorial capacity is considerably lower for the senescent population: for the parameters I use here, the population experiencing age-independent mortality has a transmission intensity over three times that of the senescent case.

I now derive expressions for the overall vector capacity for a population experiencing age-related mortality. To do this, we note that $Pr(t > A + n | t > A)$ is a conditional probability and, hence, can be derived from the law of conditional probability,

$$Pr(t > A + n | t > A) = \frac{Pr(t > A + n)}{Pr(t > A)} \quad (7.18)$$

$$= \frac{S(A + n)}{S(A)}. \quad (7.19)$$

Using expression (7.16), this allows us to derive another expression for the age-specific

vectorial capacity,

$$C(A) = a^2 \tilde{m}(A) e(A) \frac{S(A+n)}{S(A)} \quad (7.20)$$

$$= a^2 \tilde{m}(0) S(A+n) e(A). \quad (7.21)$$

This result can be interpreted intuitively: $S(A+n)e(A)$ is the average number of days a mosquito of age A will last after it has survived the EIP, meaning that the above expression represents the average number of infections generated per day from mosquitoes of age A . The overall vectorial capacity is then just the integral of this expression,

$$C = a^2 \tilde{m}(0) \int_0^\infty S(A+n) e(A) dA. \quad (7.22)$$

If the population experiences no effect of old age, then the above reduces to,

$$C = \frac{\tilde{m}(0) a^2 e^{-gn}}{g} \int_0^\infty e^{-gA} dA \quad (7.23)$$

$$= \frac{a^2 \tilde{m}(0) e^{-gn}}{g^2} \quad (7.24)$$

$$= \frac{\tilde{M} a^2 e^{-gn}}{g}, \quad (7.25)$$

3955 and we recapitulate the Ross-Macdonald result given in expression (7.4). (To obtain the
3956 final line I used $\tilde{m}(0) = \tilde{M}g$.)

I now use the model to compare how interventions that reduce adult lifespan (for example, insecticide treated nets or pyrethroid spraying) or larval survivorship (for example, the destruction of larval habitat or chemical larvicide) affect the vectorial capacity. I assume that both classes of interventions act in age-independent fashion (see Hancock, Thomas, and Godfray, 2009 for an analysis of the use of fungal pesticides that target older

mosquitoes), with an elevated risk of mortality given by ζ , for adults, and χ , for the larval stage. For the adult population, this amounts to using a Gompertz-Makeham model (see Chapter 2) for the lifespan distribution, which has an age-specific force of mortality equal to,

$$\mu(A) = de^{A\lambda} + \zeta. \quad (7.26)$$

To consider the effect of larval interventions, I change the larval population dynamic equation to,

$$\frac{dL(t)}{dt} = fM(t) \left(1 - \frac{L(t)}{\kappa}\right) - \eta L(t) - \chi L(t), \quad (7.27)$$

where η represents the rate at which larvae emerge into adults and χ is the rate at which they die due to the treatment.

I consider the effect of these interventions on the steady state values of the vectorial capacity. In Figure 7.3, I illustrate how the vectorial capacities of a population that does not experience senescence (left column) and a senescent population of the same mean lifespan ($\lambda = 0.2$, right column) are affected by an intervention that increases the adult mosquito mortality rate by $\zeta = 0.1 \text{ day}^{-1}$. In both cases, the elevated adult death rate causes a decline in the mosquito population density (top row of Figure 7.3). For the population where age does not affect mortality risk, the increased adult mortality significantly reduces the residual life expectancy of mosquitoes of all ages (left panel, second from top of Figure 7.3), whereas for the senescent case, the reduction in life expectancy is greatest for the youngest mosquitoes although not as large in magnitude (right panel, second from top). The effects on the life expectancy are mirrored by the probability that a mosquito of a given age survives the EIP (second row from bottom of Figure 7.3) where, again, the

3971 impact is greater on the non-senescent population. Overall, this means that the impact on
 3972 the disease transmission is greatest for the population where mortality is age-independent
 3973 (bottom row of Figure 7.3): for the parameters I use here, the vectorial capacity is reduced
 3974 by about 80% for the population that does not experience age-related death, and 68% for
 3975 the senescent population.

3976 I next compare the sensitivity of the vectorial capacity for the same two populations
 3977 as the intervention effectiveness is varied. Figure 7.4A shows the percentage reduction
 3978 in vectorial capacity for the population without senescence (blue line) and the one with
 3979 age-dependent mortality (orange line). In all cases, the proportional reduction in disease
 3980 transmission is lower for the senescent population. The greatest difference between the two
 3981 populations in the magnitude of the effect occurs when $\zeta \approx 0.06$, when the intervention
 3982 reduces the vectorial capacity for the population without senescence by about 16% more
 3983 than for the other case. However, I also believe it is worth noting that, in both cases,
 3984 the reduction in disease transmission that can be achieved by killing adult mosquitoes is
 3985 considerable for even modestly-effective interventions.

To calculate effects of larvicide on the vectorial capacity, I need only consider its effect on the larval population density (which affects vectorial capacity through expression (7.7)), because it does not impact adult survivorship. For the system described by expressions (7.6) and (7.27), the steady state larval density is given by,

$$\tilde{L} = \kappa \left[1 - \frac{\eta + \chi}{f\eta\bar{D}} \right], \quad (7.28)$$

where the population can sustain itself so long as $\eta(f\bar{D} - 1) > \chi$, in other words, the average density of newly-emergent adults exceeds the rate at which larvae are killed. Expression

(7.28) results in a steady state density of newly-emergent adults equal to,

$$\tilde{m}(0) = \eta\kappa \left[1 - \frac{\eta + \chi}{f\eta\bar{D}} \right]. \quad (7.29)$$

Using this expression, we can determine the percentage reduction in vectorial capacity by the following expression (obtained from the ratio of the steady state density of newly-emergent adults after the intervention to that before it),

$$r = \frac{\chi}{\eta(f\bar{D} - 1)}. \quad (7.30)$$

3986 I illustrate how this function varies as a function of the strength of the intervention χ in
 3987 Figure 7.4B. Since this relationship is independent of the specific form of mortality for the
 3988 adult population, I show only a single line in this plot. For the parameter values that I
 3989 use here, this intervention has relatively minor effects on disease transmission compared
 3990 to those that impact adult mortality (Figure 7.4A). This echoes the result of Macdonald,
 3991 1957, whose equation for R_0 partly motivated the programme initiated in the mid-twentieth
 3992 century to kill adult mosquitoes by DDT spraying (Packard, 2007).

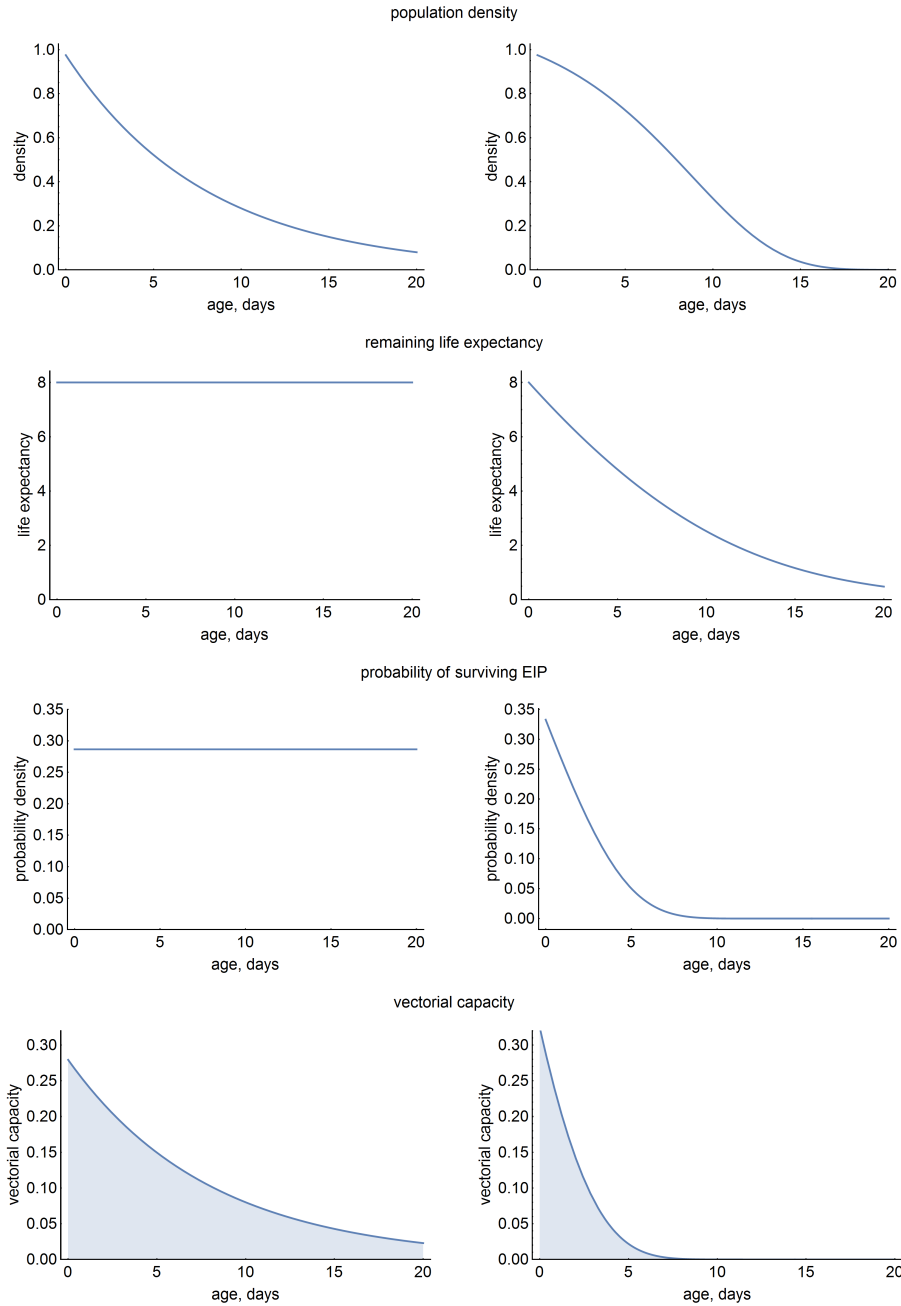


Figure 7.2: The vector capacity at a given age (bottom row) for a population without senescence (left column) and another with age-dependence ($\lambda = 0.2$, right column) is calculated by taking the product of the density of mosquitoes of that age (top row) multiplied by the remaining life expectancy of that mosquito (second row from top) and the probability it survives the EIP (second row from bottom). In both cases, I assume that the life expectancy of a mosquito is 8 days, the EIP is 10 days, $\kappa = 1$, $f = 5$, and $\eta = 1$.

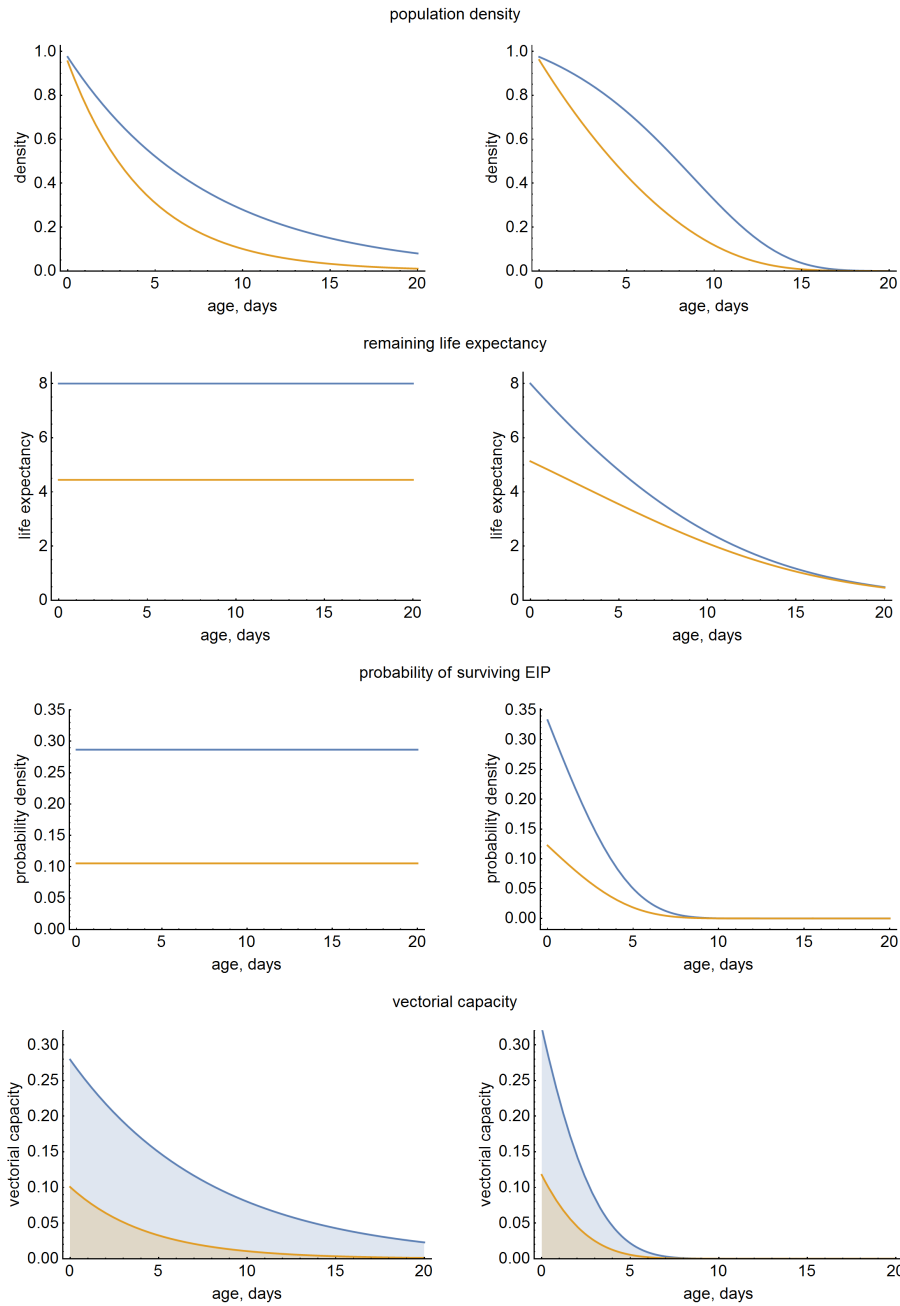


Figure 7.3: The impact of an intervention that kills adult mosquitoes ($\zeta = 0.1$, orange lines) versus the pre-intervention quantities (blue lines) on the components of vectorial capacity (top three rows) and the vectorial capacity itself (bottom row) for a population without senescence (left column) and another with age-dependent mortality (right column). In both cases, I assume that the life expectancy of a mosquito is 8 days, the EIP is 10 days, $\kappa = 1$, $f = 5$, and $\eta = 1$.

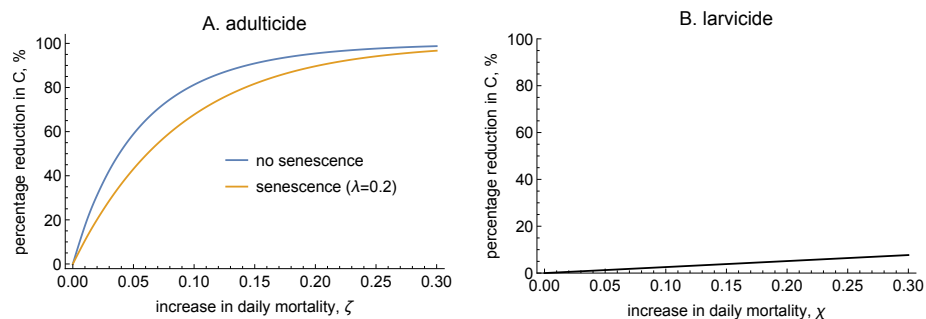


Figure 7.4: **The sensitivity of vectorial capacity to elevated adult mortality (A) for populations without senescence (blue line) and with senescence (orange line) and the same sensitivity for elevated larval mortality (B, black line).** In all cases, I assume that the life expectancy of a mosquito is 8 days, the EIP is 10 days, $\kappa = 1$, $f = 5$, and $\eta = 0.1$

7.2 Studies included in the MRR meta-analysis

The following is the subset of studies from the original Guerra et al., 2014 database which were used in the MRR meta-analysis of Chapter 2: Marini et al., 2010; Baber et al., 2010; Lacroix et al., 2009; Freitas, Eiras, and Oliveira, 2008; Midega et al., 2007; Maciel-De-Freitas, Codeco, and Lourenco-De-Oliveira, 2007; Elizondo-Quiroga et al., 2006; Ba et al., 2005; Fabian et al., 2005; La Corte Dos Santos, Forattini, and Burattini, 2004; Watson, Saul, and Kay, 2000; Harrington et al., 2001; Tsuda et al., 2001; Muir and Kay, 1998; Touré et al., 1998; Quinones et al., 1997; Costantini et al., 1996; Trpis, Häusermann, and Craig Jr, 1995; Jensen and Washino, 1994; Fernandez-Salas, Rodriguez, and Roberts, 1994; Jaal and MacDonald, 1992; Rodriguez et al., 1992; Chiang et al., 1991; Jensen and Washino, 1991; Eldridge and Reeves, 1990; Macdonald, Sebastian, and Tun, 1968; Pumpuni and Walker, 1989; Charlwood and Bryan, 1987; Charlwood and Alecrim, 1989; Birley and Charlwood, 1989; Arredondo-Jiménez, Rodríguez, and Washino, 1998; Hii, Birley, and Sang, 1990; Renshaw, Service, and Birley, 1994; Milby and Reisen, 1989; Freitas, Codeco, and Oliveira, 2007; Loong et al., 1990; Nelson and Milby, 1980; Freitas et al., 2006; McDonald, 1977; Curtis and Rawlings, 1980; Rawlings and Davidson, 1982; Conway, Trpis, and McClelland, 1974; Reisen, Mahmood, and Parveen, 1979; Nelson et al., 1978; Rawlings et al., 1981; Sempala, 1981; Takagi et al., 1995; Buei et al., 1980; Eyles and Bishop, 1943; Ordóñez González et al., 2001; Pant and Yasuno, 1973; Charlwood, Graves, and Marshall, 1988; Reisen et al., 1982; Nayar, Provost, and Hansen, 1980; Carnevale et al., 1979; Eyles, Sabrosky, and Russell, 1946; Reisen et al., 1984; Charlwood, Graves, and Birley, 1986; Trpis and Hausermann, 1975; Lutwama and Mukwaya, 1994; Wada et al., 1969; Takken et al., 1998; Abdel-Malek and Abdel-Aal, 1966; Valerio et al., 2012; Zetek, 1915; Takagi et al., 1995; Yasuno, Rajagopalan, and Laürequé, 1975; Eyles and Cox, 1943; Germain, Hervé, and Geoffroy, 1974.

7.3 Studies included in the dissection study meta-analysis

The following is a list of studies included in the dissection study meta-analysis of Chapter 3: Catangui, 1971; Charlwood and Wilkes, 1979; Chang et al., 1991; Charlwood, Vij, and Billingsley, 2000; Edalat et al., 2015; Barros, Honorio, and Arruda, 2011; Lines, Wilkes, and Lyimo, 1991; Magesa et al., 1991; Lines et al., 1991; Forattini et al., 1996; Samarawickrema, 1967; Samarawickrema, 1968; Barros et al., 2007; Charlwood et al., 1985; Chandra, Seal, and Hati, 1996; Hoc and Wilkes, 1995; Vythilingam et al., 1997; Russell, 1986a; Chadee and Ritchie, 2010; Ebsary and Crans, 1977; Charlwood, 1980; Shriram and Krishnamoorthy, 2011; Mahmood and Reisen, 1981; Samarawickrema, Sone, and Cummings, 1987; Uttah et al., 2013; Ramaiah and Das, 1992; Buei and Ito, 1982; Chadee et al., 1995; Chandra, 2008; Foll and Pant, 1966; Kanda, Joo, and Choi, 1975; Mala et al., 2014; Shalaby and World Health Organisation, 1962; World Health Organisation, 1960; Reisen, Mahmood, and Parveen, 1980; Detinova, 1962; Smith and Kurtz, 1994; Jayanetti, Wijesundera, and Amaresinghe, 1987; Ch et al., 2013; Pant and World Health Organisation, 1962; Ghosh, Mandal, and Chandra, 2010; Mahanta et al., 1999; Schlein and Müller, 2012; Penilla et al., 2002; Schlein and Müller, 2015; Snow and Boreham, 1978; Detinova, 1968; Nathan, 1981; Charlwood, Dagoro, and Paru, 1985; Schlein and Müller, 2008; Beier et al., 2012; Müller and Schlein, 2005; Surendran et al., 2006; Müller et al., 2010; Mendis et al., 1998; Wilkes, Matola, and Charlwood, 1996; Chatterjee and Chandra, 2000; Müller et al., 2010; Tuchinda et al., 1969; Chan, 1971; Aigbodion et al., 2011; Qualls et al., 2015; Reisen, Mahmood, and Azra, 1981; Gillies and Wilkes, 1972; Gillies and Wilkes, 1965; Kay, 1979; Hitchcock Jr, 1968.

7.4 Studies included in the gonotrophic cycle duration meta-analysis

The following is a list of studies included in the gonotrophic cycle study meta-analysis of Chapter 3: Gillies and Wilkes, 1965; Sheppard et al., 1969; Meillon, Sebastian, and Khan, 1967; Pant and Yasuno, 1973; Germain, Hervé, and Geoffroy, 1974; Lowe et al., 1973; Fernandez-Salas, Rodriguez, and Roberts, 1994; Buei et al., 1980; Rawlings and Curtis, 1982; Mori and Wada, 1977; Sempala, 1981; Reisen et al., 1983; Birley and Charlwood, 1989; Suzuki, 1978; Charlwood and Bryan, 1987; Charlwood, Graves, and Birley, 1986; Charlwood and Wilkes, 1979; Charlwood et al., 1995; Chang et al., 1991; Edalat et al., 2015; Barros, Honorio, and Arruda, 2011; Samarawickrema, 1968; Samarawickrema, 1967; Charlwood et al., 1985; Chandra, Seal, and Hati, 1996; Russell, 1986a; Chadee and Ritchie, 2010; Mahmood and Reisen, 1981; Ahumada, Lapointe, and Samuel, 2004; Kenawy, 1991; Rajagopalan, 1980; Chandra, 2008; Scholl, Porter, and Defoliart, 1979; Mala et al., 2014; Afrane et al., 2005; Gillies, 1953; Quinones et al., 1997; Rúa et al., 2005; Delatte et al., 2009; Arredondo-Jiménez, Rodríguez, and Washino, 1998; Wong et al., 2014; Ijumba, Mosha, and Lindsay, 2002.

Bibliography

- [1] A. A. Abdel-Malek, M. Abdel-Aal, et al. “Study of the dispersion and flight range of *Anopheles sergentii* Theo in Siwa Oasis using radioactive isotopes as markers”. In: *Bulletin of the World Health Organisation* 35 (1966), pp. 968–973.
- [2] H. Abdi. “Partial least square regression (PLS regression)”. In: *Encyclopedia for Research Methods for the Social Sciences* 6.4 (2003), pp. 792–795.
- [3] D. Acemoglu, S. Johnson, and J. A. Robinson. “The colonial origins of comparative development: An empirical investigation”. In: *National Bureau of Economic Research* (2000).
- [4] A. Adamou et al. “The contribution of aestivating mosquitoes to the persistence of *Anopheles gambiae* in the Sahel”. In: *Malaria Journal* 10.1 (2011), p. 151.
- [5] Z. N. Adelman and Z. Tu. “Control of mosquito-borne infectious diseases: sex and gene drive”. In: *Trends in Parasitology* 32.3 (2016), pp. 219–229.
- [6] Y. A. Afrane et al. “Effects of microclimatic changes caused by land use and land cover on duration of gonotrophic cycles of *Anopheles gambiae* (Diptera: Culicidae) in western Kenya highlands”. In: *Journal of Medical Entomology* 42.6 (2005), pp. 974–980.

- 4073 [7] J. A. Ahumada, D. Lapointe, and M. D. Samuel. “Modeling the population dynam-
4074 ics of *Culex quinquefasciatus* (Diptera: Culicidae), along an elevational gradient in
4075 Hawaii”. In: *Journal of Medical Entomology* 41.6 (2004), pp. 1157–1170.
- 4076 [8] F. Aigbodion et al. “Pelagia Research Library”. In: *European Journal of Experi-*
4077 *mental Biology* 1.4 (2011), pp. 173–180.
- 4078 [9] B. T. Aldrich et al. “Identification of termite species and subspecies of the genus
4079 *Zootermopsis* using near-infrared reflectance spectroscopy”. In: *Journal of Insect*
4080 *Science* 7.1 (2007).
- 4081 [10] S. Allen et al. “ α +-Thalassemia protects children against disease caused by other
4082 infections as well as malaria”. In: *Proceedings of the National Academy of Sciences*
4083 94.26 (1997), pp. 14736–14741.
- 4084 [11] L. Alphey. “Can CRISPR-Cas9 gene drives curb malaria?” In: *Nature Biotechnology*
4085 34.2 (2016), p. 149.
- 4086 [12] L. Alphey et al. “Sterile-insect methods for control of mosquito-borne diseases: an
4087 analysis”. In: *Vector-Borne and Zoonotic Diseases* 10.3 (2010), pp. 295–311.
- 4088 [13] N. Alphey and M. B. Bonsall. “Interplay of population genetics and dynamics in
4089 the genetic control of mosquitoes”. In: *Journal of the Royal Society Interface* 11.93
4090 (2014), p. 20131071.
- 4091 [14] R. Anderson, B. Knols, and J. Koella. “*Plasmodium falciparum* sporozoites increase
4092 feeding-associated mortality of their mosquito hosts *Anopheles gambiae* sl”. In: *Par-*
4093 *asitology* 120.4 (2000), pp. 329–333.
- 4094 [15] R. M. Anderson and R. M. May. *Infectious diseases of humans: dynamics and con-*
4095 *trol*. Vol. 28. Wiley Online Library, 1992.

- [16] J. I. Arredondo-Jiménez, M. H. Rodríguez, and R. K. Washino. “Gonotrophic cycle and survivorship of *Anopheles vestitipennis* (Diptera: Culicidae) in two different ecological areas of southern Mexico”. In: *Journal of Medical Entomology* 35.6 (1998), pp. 937–942.
- [17] E. A. Ashley et al. “Spread of artemisinin resistance in *Plasmodium falciparum* malaria”. In: *New England Journal of Medicine* 371.5 (2014), pp. 411–423.
- [18] Y. Ba et al. “Aspects of bioecology of two Rift Valley fever virus vectors in Senegal (West Africa): *Aedes vexans* and *Culex poicilipes* (Diptera: Culicidae)”. In: *Journal of Medical Entomology* 42.5 (2005), pp. 739–750.
- [19] I. Baber et al. “Population size and migration of *Anopheles gambiae* in the Bancoumana Region of Mali and their significance for efficient vector control”. In: *PLoS One* 5.4 (2010), e10270.
- [20] A Bárdossy, J Stehlik, and H.-J. Caspary. “Generating of areal precipitation series in the upper Neckar catchment”. In: *Physics and Chemistry of the Earth, Part B: Hydrology, Oceans and Atmosphere* 26.9 (2001), pp. 683–687.
- [21] A. Bárdossy and E. J. Plate. “Space-time model for daily rainfall using atmospheric circulation patterns”. In: *Water Resources Research* 28.5 (1992), pp. 1247–1259.
- [22] F. S. M. de Barros, N. A. Honorio, and M. E. Arruda. “Survivorship of *Anopheles darlingi* (Diptera: Culicidae) in relation with malaria incidence in the Brazilian Amazon”. In: *PLoS One* 6.8 (2011), e22388.
- [23] F. S. M. de Barros et al. “Parity and age composition for *Anopheles darlingi* Root (Diptera: Culicidae) and *Anopheles albittarsis* Lynch-Arribálzaga (Diptera: Culicidae) of the northern Amazon Basin, Brazil”. In: *Journal of Vector Ecology* 32.1 (2007), pp. 54–68.

- 4120 [24] J. P. Basilua Kanza et al. “Pyrethroid, DDT and malathion resistance in the malaria
4121 vector *Anopheles gambiae* from the Democratic Republic of Congo”. In: *Transac-*
4122 *tions of the Royal Society of Tropical Medicine and Hygiene* 107.1 (2012), pp. 8–
4123 14.
- 4124 [25] J. C. Beier et al. “Attractive toxic sugar bait (ATSB) methods decimate populations
4125 of *Anopheles* malaria vectors in arid environments regardless of the local availability
4126 of favoured sugar-source blossoms”. In: *Malaria Journal* 11.1 (2012), p. 31.
- 4127 [26] W. Beklemishev. “Gonotrophic rhythm as a basic principle of the biology of *Anophe-*
4128 *les*”. In: *Vop Fiziol Ekol Malyar Konara* 1 (1940), pp. 3–22.
- 4129 [27] M. Belfort and R. J. Roberts. “Homing endonucleases: keeping the house in order”.
4130 In: *Nucleic Acids Research* 25.17 (1997), pp. 3379–3388.
- 4131 [28] S. E. Bellan. “The importance of age dependent mortality and the extrinsic incu-
4132 bation period in models of mosquito-borne disease transmission and control”. In:
4133 *PLoS One* 5.4 (2010), e10165.
- 4134 [29] M. Q. Benedict and A. S. Robinson. “The first releases of transgenic mosquitoes: an
4135 argument for the sterile insect technique”. In: *Trends in Parasitology* 19.8 (2003),
4136 pp. 349–355.
- 4137 [30] S. Bhatt et al. “The effect of malaria control on *Plasmodium falciparum* in Africa
4138 between 2000 and 2015”. In: *Nature* 526.7572 (2015), pp. 207–211.
- 4139 [31] S. Bhatt et al. “The global distribution and burden of dengue”. In: *Nature* 496.7446
4140 (2013), pp. 504–507.
- 4141 [32] G. Bian et al. “The endosymbiotic bacterium *Wolbachia* induces resistance to dengue
4142 virus in *Aedes aegypti*”. In: *PLoS Pathogens* 6.4 (2010), e1000833.

- [33] M. Birley and J. Charlwood. “The effect of moonlight and other factors on the oviposition cycle of malaria vectors in Madang, Papua New Guinea”. In: *Annals of Tropical Medicine & Parasitology* 83.4 (1989), pp. 415–422.
- [34] M Blanco et al. “NIR calibration in non-linear systems: different PLS approaches and artificial neural networks”. In: *Chemometrics and Intelligent Laboratory Systems* 50.1 (2000), pp. 75–82.
- [35] J. I. Blanford et al. “Implications of temperature variation for malaria parasite development across Africa”. In: *Scientific Reports* 3 (2013).
- [36] C. I. Bliss and R. A. Fisher. “Fitting the negative binomial distribution to biological data”. In: *Biometrics* 9.2 (1953), pp. 176–200.
- [37] I. I. Bogoch et al. “Anticipating the international spread of Zika virus from Brazil”. In: *The Lancet* 387.10016 (2016), pp. 335–336.
- [38] M. Bonizzoni et al. “The invasive mosquito species *Aedes albopictus*: current knowledge and future perspectives”. In: *Trends in Parasitology* 29.9 (2013), pp. 460–468.
- [39] M. F. Boyd et al. *Malariology. A Comprehensive Survey of all Aspects of this Group of Diseases from a Global Standpoint*. Vol. 2. W.B. Saunders Company, 1949.
- [40] O. J. Brady et al. “Refining the global spatial limits of dengue virus transmission by evidence-based consensus”. In: *PLoS Neglected Tropical Diseases* 6.8 (2012), e1760.
- [41] O. J. Brady et al. “Vectorial capacity and vector control: reconsidering sensitivity to parameters for malaria elimination”. In: *Transactions of the Royal Society of Tropical Medicine and Hygiene* 110.2 (2016), pp. 107–117.
- [42] Britannica. *Encyclopaedia of Britannica website*. <https://www.britannica.com/event/World-War-I/Killed-wounded-and-missing>. Accessed: 2017-09-18. 2017.

- 4166 [43] Britannica. *Encyclopaedia of Britannica website*. [https://www.britannica.com/](https://www.britannica.com/event/World-War-II)
4167 [event/World-War-II](https://www.britannica.com/event/World-War-II). Accessed: 2017-09-18. 2017.
- 4168 [44] L. J. Bruce-Chwatt, C Garrett-Jones, and B Weitz. “Ten years’ study (1955-64) of
4169 host selection by anopheline mosquitos”. In: *Bulletin of the World Health Organi-*
4170 *sation* 35.3 (1966), p. 405.
- 4171 [45] L. Bruce-Chwatt et al. “Cinchona and its alkaloids: 350 years.” In: *New York State*
4172 *Journal of Medicine* 88.6 (1988), pp. 318–322.
- 4173 [46] K Buei et al. “Field studies on the gonotrophic cycle of *Culex tritaeniorhynchus*”.
4174 In: *Medical Entomology and Zoology* 31.1 (1980), pp. 57–62.
- 4175 [47] K. Buei and S. Ito. “The age-composition of field populations and the survival rates
4176 in *Culex tritaeniorhynchus* Giles”. In: *Medical Entomology and Zoology* 33.1 (1982),
4177 pp. 21–25.
- 4178 [48] A. Burt. “Site-specific selfish genes as tools for the control and genetic engineering
4179 of natural populations”. In: *Proceedings of the Royal Society of London. Series B:*
4180 *Biological Sciences* 270.1518 (2003), pp. 921–928.
- 4181 [49] W. F. Bynum. “An experiment that failed: malaria control at Mian Mir.” In: *Paras-*
4182 *itologia* 36.1-2 (1994), pp. 107–120.
- 4183 [50] G. S. Campos, A. C. Bandeira, and S. I. Sardi. “Zika virus outbreak, Bahia, Brazil”.
4184 In: *Emerging Infectious Diseases* 21.10 (2015), p. 1885.
- 4185 [51] J. R. Carey. “Insect biodemography”. In: *Annual Review of Entomology* 46.1 (2001),
4186 pp. 79–110.
- 4187 [52] P. Carnevale et al. “Etude du cycle gonotrophique d’*Anopheles gambiae* (Diptera,
4188 Culicidae)(Giles, 1902) en zone de forêt dégradée d’Afrique Centrale”. In: *Cahiers*
4189 *Orstom. Série Entomologie Médicale et Parasitologie* 17 (1979), pp. 55–75.

- 4190 [53] B. Carpenter et al. “Stan: A probabilistic programming language”. In: *Journal of*
4191 *Statistical Software* (2016).
- 4192 [54] R. Carter and K. N. Mendis. “Evolutionary and historical aspects of the burden of
4193 malaria”. In: *Clinical Microbiology Reviews* 15.4 (2002), pp. 564–594.
- 4194 [55] F. Catangui et al. “Studies on the gonotrophic cycle of *Anopheles minimus flavi-*
4195 *rostris* and the application of physiological age grading technique on the same
4196 species.” In: *Southeast Asian Journal of Tropical Medicine and Public Health* 2.3
4197 (1971), pp. 384–92.
- 4198 [56] L. L. Cavalli-Sforza. *Genes, peoples, and languages*. University of California Press,
4199 2001.
- 4200 [57] L. L. Cavalli-Sforza, P. Menozzi, and A. Piazza. *The history and geography of human*
4201 *genes*. Princeton University Press, 1994.
- 4202 [58] A. Celli et al. “The history of Malaria in the Roman Campagna.” In: *Memorie*
4203 *Accademia dei Lincei* 1.Series 6 (1925).
- 4204 [59] G. Ch et al. “Comparative Studies of Different Indices Related to Filarial Vector of
4205 a Rural and an Urban Area of West Bengal, India”. In: *Tropical Medicine & Surgery*
4206 (2013).
- 4207 [60] D. D. Chadee and S. A. Ritchie. “Oviposition behaviour and parity rates of *Aedes*
4208 *aegypti* collected in sticky traps in Trinidad, West Indies”. In: *Acta Tropica* 116.3
4209 (2010), pp. 212–216.
- 4210 [61] D. D. Chadee et al. “Seasonal abundance, biting cycle and parity of the mosquito
4211 *Haemagogus leucocelaenus* in Trinidad, west indies”. In: *Medical and Veterinary*
4212 *Entomology* 9.4 (1995), pp. 372–376.

- 4213 [62] K.-L. Chan et al. "Life table studies of *Aedes albopictus* (Skuse)". In: *Sterility prin-*
4214 *ciples for insect control or eradication* (1971), pp. 131–144.
- 4215 [63] M. Chan and M. A. Johansson. "The incubation periods of dengue viruses". In:
4216 *PLoS One* 7.11 (2012), e50972.
- 4217 [64] G Chandra, B Seal, and A. Hati. "Age composition of the filarial vector *Culex quin-*
4218 *quefasciatus* (Diptera: Culicidae) in Calcutta, India". In: *Bulletin of Entomological*
4219 *Research* 86.3 (1996), pp. 223–226.
- 4220 [65] G. Chandra. "Age composition of incriminated malaria vector in a rural foothills in
4221 West Bengal, India". In: *Indian Journal of Medical Research* 127.6 (2008), p. 607.
- 4222 [66] M.-S. Chang et al. "Comparative transmission potential of three *Mansonia spp.*
4223 (Diptera: Culicidae) for filariasis in Sarawak, Malaysia". In: *Bulletin of Entomolog-*
4224 *ical Research* 81.4 (1991), pp. 437–444.
- 4225 [67] J. D. Charlwood and W. Alecrim. "Capture-recapture studies with the South Amer-
4226 ican malaria vector *Anopheles darlingi*, Root". In: *Annals of Tropical Medicine &*
4227 *Parasitology* 83.6 (1989), pp. 569–576.
- 4228 [68] J. Charlwood. "Observations on the bionomics of *Anopheles darlingi* Root (Diptera:
4229 Culicidae) from Brazil". In: *Bulletin of Entomological Research* 70.4 (1980), pp. 685–
4230 692.
- 4231 [69] J. Charlwood and J. Bryan. "A mark-recapture experiment with the filariasis vector
4232 *Anopheles punctulatus* in Papua New Guinea". In: *Annals of Tropical Medicine &*
4233 *Parasitology* 81.4 (1987), pp. 429–436.
- 4234 [70] J. Charlwood, H Dagoro, and R Paru. "Blood-feeding and resting behaviour in
4235 the *Anopheles punctulatus* Dönitz complex (Diptera: Culicidae) from coastal Papua
4236 New Guinea". In: *Bulletin of Entomological Research* 75.3 (1985), pp. 463–476.

- 4237 [71] J. Charlwood, P. Graves, and M. Birley. "Capture-recapture studies with mosquitoes
4238 of the group of *Anopheles punctulatus* Dönitz (Diptera: Culicidae) from Papua New
4239 Guinea". In: *Bulletin of Entomological Research* 76.2 (1986), pp. 211–227.
- 4240 [72] J. Charlwood, P. Graves, and T. d. C. Marshall. "Evidence for a memorized home
4241 range in *Anopheles farauti* females from Papua New Guinea". In: *Medical and Vet-
4242 erinary Entomology* 2.2 (1988), pp. 101–108.
- 4243 [73] J. Charlwood, R. Vij, and P. Billingsley. "Dry season refugia of malaria-transmitting
4244 mosquitoes in a dry savannah zone of east Africa." In: *The American Journal of
4245 Tropical Medicine and Hygiene* 62.6 (2000), pp. 726–732.
- 4246 [74] J. Charlwood and T. Wilkes. "Studies on the age-composition of samples of *Anophe-
4247 les darlingi* Root (Diptera: Culicidae) in Brazil". In: *Bulletin of Entomological Re-
4248 search* 69.02 (1979), pp. 337–342.
- 4249 [75] J. Charlwood et al. "Assessing survival rates of *Anopheles farauti* (diptera: Culici-
4250 dae) from Papua New Guinea". In: *The Journal of Animal Ecology* (1985), pp. 1003–
4251 1016.
- 4252 [76] J. Charlwood et al. "The rise and fall of *Anopheles arabiensis* (Diptera: Culicidae)
4253 in a Tanzanian village". In: *Bulletin of Entomological Research* 85.1 (1995), pp. 37–
4254 44.
- 4255 [77] R. N. Charrel. "Chikungunya outbreaks—the globalization of vectorborne diseases".
4256 In: *The New England Journal of Medicine* 356.8 (2007), p. 769.
- 4257 [78] S. Chatterjee and G. Chandra. "Role of *Anopheles subpictus* as a primary vector of
4258 malaria in an area in India". In: *Japanese Journal of Tropical Medicine and Hygiene*
4259 28.3 (2000), pp. 177–181.

- [79] B. S. Chevalier and B. L. Stoddard. “Homing endonucleases: structural and functional insight into the catalysts of intron/intein mobility”. In: *Nucleic Acids Research* 29.18 (2001), pp. 3757–3774.
- [80] G. Chiang et al. “Capture-recapture studies with *Anopheles maculatus* Theobald (Diptera: Culicidae) the vector of malaria in peninsular Malaysia”. In: *Southeast Asian Journal of Tropical Medicine and Public Health* 22 (1991), pp. 643–647.
- [81] T. Chouin-Carneiro et al. “Differential Susceptibilities of *Aedes aegypti* and *Aedes albopictus* from the Americas to Zika Virus”. In: *PLoS Neglected Tropical Diseases* 10.3 (2016), e0004543.
- [82] B. M. Christensen, M. M. LaFond, and L. A. Christensen. “Defense reactions of mosquitoes to filarial worms: Effect of host age on the immune response to *Dirofilaria immitis* microfilariae”. In: *The Journal of Parasitology* (1986), pp. 212–215.
- [83] S. R. Christophers. *Malaria in the Punjab*. Superintendent Government Printing, India, 1911.
- [84] A. Clements and G. Paterson. “The analysis of mortality and survival rates in wild populations of mosquitoes”. In: *Journal of Applied Ecology* (1981), pp. 373–399.
- [85] M. Coluzzi. “The clay feet of the malaria giant and its African roots: hypotheses and inferences about origin, spread and control of *Plasmodium falciparum*.” In: *Parassitologia* 41.1-3 (1999), pp. 277–283.
- [86] G. Conway, M. Trpis, and G. McClelland. “Population parameters of the mosquito *Aedes aegypti* (L.) estimated by mark-release-recapture in a suburban habitat in Tanzania”. In: *The Journal of Animal Ecology* (1974), pp. 289–304.

- [87] P. E. Cook et al. “The use of transcriptional profiles to predict adult mosquito age under field conditions”. In: *Proceedings of the National Academy of Sciences* 103.48 (2006), pp. 18060–18065.
- [88] C. Costantini et al. “Density, survival and dispersal of *Anopheles gambiae* complex mosquitoes in a West African Sudan savanna village”. In: *Medical and Veterinary Entomology* 10.3 (1996), pp. 203–219.
- [89] D. R. Cox and D. Oakes. *Analysis of survival data*. Vol. 21. CRC Press, 1984.
- [90] F. E. Cox. “History of the discovery of the malaria parasites and their vectors”. In: *Parasites & Vectors* 3.1 (2010), p. 5.
- [91] M. H. Craig, R. Snow, and D. le Sueur. “A climate-based distribution model of malaria transmission in sub-Saharan Africa”. In: *Parasitology today* 15.3 (1999), pp. 105–111.
- [92] C. Curtis. “Possible use of translocations to fix desirable genes in insect pest populations”. In: *Nature* 218.5139 (1968), pp. 368–369.
- [93] C. Curtis and P Rawlings. “A preliminary study of dispersal and survival of *Anopheles culicifacies* in relation to the possibility of inhibiting the spread of insecticide resistance”. In: *Ecological Entomology* 5.1 (1980), pp. 11–17.
- [94] A Dao et al. “Signatures of aestivation and migration in Sahelian malaria mosquito populations”. In: *Nature* 516.7531 (2014), pp. 387–390.
- [95] H. C. Darby. *The draining of the Fens*. Cambridge University Press, 2011.
- [96] G Davidson. “Estimation of the survival-rate of anopheline mosquitoes in nature”. In: *Nature* (1954).

- 4304 [97] E. J. Dawes et al. “*Anopheles* mortality is both age-and *Plasmodium*-density de-
4305 pendent: implications for malaria transmission”. In: *Malaria Journal* 8.1 (2009),
4306 p. 228.
- 4307 [98] R. Dawkins. *The selfish gene*. Oxford University Press, 1976.
- 4308 [99] H. Delatte et al. “Influence of temperature on immature development, survival,
4309 longevity, fecundity, and gonotrophic cycles of *Aedes albopictus*, vector of chikun-
4310 gunya and dengue in the Indian Ocean”. In: *Journal of Medical Entomology* 46.1
4311 (2009), pp. 33–41.
- 4312 [100] A. Deredec, A. Burt, and H. C. J. Godfray. “The population genetics of using
4313 homing endonuclease genes in vector and pest management”. In: *Genetics* 179.4
4314 (2008), pp. 2013–2026.
- 4315 [101] A. Deredec, H. C. J. Godfray, and A. Burt. “Requirements for effective malaria
4316 control with homing endonuclease genes”. In: *Proceedings of the National Academy*
4317 *of Sciences* 108.43 (2011), E874–E880.
- 4318 [102] M. Desena et al. “Potential for aging female *Aedes aegypti* (Diptera: Culicidae) by
4319 gas chromatographic analysis of cuticular hydrocarbons, including a field evalua-
4320 tion”. In: *Journal of Medical Entomology* 36.6 (1999), pp. 811–823.
- 4321 [103] T. S. Detinova et al. “Age grouping methods in Diptera of medical importance with
4322 special reference to some vectors of malaria”. In: *Monograph series World Health*
4323 *Organisation* (1962).
- 4324 [104] T. Detinova. “Age structure of insect populations of medical importance”. In: *An-*
4325 *ual Review of Entomology* 13.1 (1968), pp. 427–450.
- 4326 [105] M Di Luca et al. “Experimental studies of susceptibility of Italian *Aedes albopictus*
4327 to Zika virus”. In: *Eurosurveillance* 21.18 (2016), p. 30223.

- 4328 [106] J. M. Diamond. *Guns, germs and steel: a short history of everybody for the last*
4329 *13,000 years*. Random House, 1998.
- 4330 [107] G. Dick, S. Kitchen, and A. Haddow. “Zika virus (I). Isolations and serological
4331 specificity”. In: *Transactions of the Royal Society of Tropical Medicine and Hygiene*
4332 46.5 (1952), pp. 509–520.
- 4333 [108] O. B. Dick et al. “The history of dengue outbreaks in the Americas”. In: *The*
4334 *American Journal of Tropical Medicine and Hygiene* 87.4 (2012), pp. 584–593.
- 4335 [109] B. L. Dickens and H. L. Brant. “Effects of marking methods and fluorescent dusts
4336 on *Aedes aegypti* survival”. In: *Parasites & Vectors* 7.1 (2014), p. 1.
- 4337 [110] V Dietz et al. “The 1986 dengue and dengue hemorrhagic fever epidemic in Puerto
4338 Rico: epidemiologic and clinical observations.” In: *Puerto Rico Health Sciences Jour-*
4339 *nal* 15.3 (1996), pp. 201–210.
- 4340 [111] M. Dobson. ““Marsh fever” - The geography of malaria in England”. In: *Journal of*
4341 *Historical Geography* 6.4 (1980), pp. 357–389.
- 4342 [112] M. J. Dobson. *Contours of death and disease in early modern England*. Vol. 29.
4343 Cambridge University Press, 2003.
- 4344 [113] A. M. Dondorp et al. “Artemisinin resistance in *Plasmodium falciparum* malaria”.
4345 In: *New England Journal of Medicine* 361.5 (2009), pp. 455–467.
- 4346 [114] J. A. Doudna and E. Charpentier. “The new frontier of genome engineering with
4347 CRISPR-Cas9”. In: *Science* 346.6213 (2014), p. 1258096.
- 4348 [115] F. Dowell et al. “Sex separation of tsetse fly pupae using near-infrared spectroscopy”.
4349 In: *Bulletin of Entomological Research* 95.3 (2005), pp. 249–257.

- 4350 [116] F. E. Dowell, A. E. Noutcha, and K. Michel. “The effect of preservation methods on
4351 predicting mosquito age by near Infrared spectroscopy”. In: *The American Journal*
4352 *of Tropical Medicine and Hygiene* 85.6 (2011), pp. 1093–1096.
- 4353 [117] F. E. Dowell et al. “Identifying stored-grain insects using near-infrared spectroscopy”.
4354 In: *Journal of Economic Entomology* 92.1 (1999), pp. 165–169.
- 4355 [118] M. Dubrulle et al. “Chikungunya virus and *Aedes* mosquitoes: saliva is infectious
4356 as soon as two days after oral infection”. In: *PLoS one* 4.6 (2009), e5895.
- 4357 [119] M. R. Duffy et al. “Zika virus outbreak on Yap Island, federated states of Microne-
4358 sia”. In: *New England Journal of Medicine* 360.24 (2009), pp. 2536–2543.
- 4359 [120] V. A. Dyck, J. Hendrichs, and A. S. Robinson. *Sterile insect technique: principles*
4360 *and practice in area-wide integrated pest management*. Springer, 2006.
- 4361 [121] C. Dye. “The analysis of parasite transmission by bloodsucking insects”. In: *Annual*
4362 *review of entomology* 37.1 (1992), pp. 1–19.
- 4363 [122] B. Ebsary and W. Crans. “The biting activity of *Aedes sollicitans* in New Jersey.”
4364 In: *Mosquito News* (1977).
- 4365 [123] H. Edalat et al. “Vectorial capacity and age determination of *Anopheles stephens*
4366 *Liston* (Diptera: Culicidae), during the malaria transmission in Southern Iran”. In:
4367 *Journal of Entomology and Zoology Studies* 3.1 (2015), pp. 256–263.
- 4368 [124] B. Eldridge and W. Reeves. “Daily survivorship of adult *Aedes communis* in a high
4369 mountain environment in California.” In: *Journal of the American Mosquito Control*
4370 *Association* 6.4 (1990), pp. 662–666.
- 4371 [125] A. Elizondo-Quiroga et al. “Gonotrophic cycle and survivorship of *Culex quinque-*
4372 *fasciatus* (Diptera: Culicidae) using sticky ovitraps in Monterrey, northeastern Mex-

- 4373 ico". In: *Journal of the American Mosquito Control Association* 22.1 (2006), pp. 10–
4374 14.
- 4375 [126] A. A. Escalante and F. J. Ayala. "Evolutionary origin of *Plasmodium* and other
4376 *Apicomplexa* based on rRNA genes". In: *Proceedings of the National Academy of*
4377 *Sciences* 92.13 (1995), pp. 5793–5797.
- 4378 [127] A. A. Escalante and F. J. Ayala. "Phylogeny of the malarial genus *Plasmodium*,
4379 derived from rRNA gene sequences". In: *Proceedings of the National Academy of*
4380 *Sciences* 91.24 (1994), pp. 11373–11377.
- 4381 [128] A. A. Escalante, E. Barrio, and F. J. Ayala. "Evolutionary origin of human and
4382 primate malarias: evidence from the circumsporozoite protein gene." In: *Molecular*
4383 *Biology and Evolution* 12.4 (1995), pp. 616–626.
- 4384 [129] K. M. Esvelt et al. "Concerning RNA-guided gene drives for the alteration of wild
4385 populations". In: *eLife* 3 (2014), e03401.
- 4386 [130] P. W. Ewald. *Evolution of infectious disease*. Oxford University Press on Demand,
4387 1994.
- 4388 [131] D. Eyles, L. Bishop, et al. "An Experiment on the Range of Dispersion of *Anopheles*
4389 *quadrimaculatus*". In: *American Journal of Hygiene* 37.3 (1943).
- 4390 [132] D. E. Eyles, W. W. Cox, et al. "The Measurement of a Population of *Anopheles*
4391 *quadrimaculatus* Say." In: *Journal of the National Malaria Society* 2.2 (1943).
- 4392 [133] D. E. Eyles, C. W. Sabrosky, J. C. Russell, et al. *Long-range dispersal of Anopheles*
4393 *quadrimaculatus*. US Government Printing Office, 1946.
- 4394 [134] M. M. Fabian et al. "Mark-release-recapture experiments with *Anopheles saperoi*
4395 (Diptera: Culicidae) in the Yona Forest, northern Okinawa, Japan". In: *Southeast*
4396 *Asian Journal of Tropical Medicine and Public Health* 36.1 (2005), p. 54.

- 4397 [135] N. R. Faria et al. “Zika virus in the Americas: early epidemiological and genetic
4398 findings”. In: *Science* 352.6283 (2016), pp. 345–349.
- 4399 [136] J. Farrar et al. *Manson’s Tropical Diseases E-Book*. Elsevier Health Sciences, 2013.
- 4400 [137] H. M. Ferguson et al. “Ecology: a prerequisite for malaria elimination and eradica-
4401 tion”. In: *PLoS Medicine* 7.8 (2010), e1000303.
- 4402 [138] H. M. Ferguson et al. “Redressing the sex imbalance in knowledge of vector biology”.
4403 In: *Trends in Ecology & Evolution* 20.4 (2005), pp. 202–209.
- 4404 [139] H. Ferguson et al. “Ecology: a prerequisite for malaria elimination and eradication”.
4405 In: *PLoS Medicine* 7 (2010), e1000303.
- 4406 [140] I. Fernandez-Salas, M. H. Rodriguez, and D. R. Roberts. “Gonotrophic cycle and
4407 survivorship of *Anopheles pseudopunctipennis* (Diptera: Culicidae) in the Tapachula
4408 foothills of southern Mexico”. In: *Journal of Medical Entomology* 31.3 (1994), pp. 340–
4409 347.
- 4410 [141] U. Fillinger and S. W. Lindsay. “Larval source management for malaria control in
4411 Africa: myths and reality”. In: *Malaria Journal* 10.1 (2011), p. 353.
- 4412 [142] U. Fillinger et al. “The practical importance of permanent and semipermanent habi-
4413 tats for controlling aquatic stages of *Anopheles gambiae* sensu lato mosquitoes: op-
4414 erational observations from a rural town in western Kenya”. In: *Tropical Medicine*
4415 *& International Health* 9.12 (2004), pp. 1274–1289.
- 4416 [143] R. A. Fisher. *The genetical theory of natural selection*. Oxford University Press,
4417 1930.
- 4418 [144] C. Foll and C. Pant. “The conditions of malaria transmission in Katsina Province,
4419 Northern Nigeria, and a discussion of the effects of dichlorvos application”. In:
4420 *Bulletin of the World Health Organisation* 34.3 (1966), p. 395.

- 4421 [145] O. P. Forattini et al. "Role of *Anopheles (Kerteszia) bellator* as malaria vector in
4422 southeastern Brazil (Diptera: Culicidae)". In: *Memórias do Instituto Oswaldo Cruz*
4423 94.6 (1999), pp. 715–718.
- 4424 [146] O. P. Forattini et al. "Studies on mosquitoes (Diptera: Culicidae) and anthropic
4425 environment: 11-Biting activity and blood-seeking parity of *Anopheles (Kerteszia)*
4426 in South-Eastern Brazil". In: *Revista de Saude Publica* 30 (1996), pp. 107–114.
- 4427 [147] A. Fox and R. Brust. "How do dilatations form in mosquito ovarioles?" In: *Para-
4428 sitology Today* 10.1 (1994), pp. 19–23.
- 4429 [148] B. D. Foy et al. "Probable non-vector-borne transmission of Zika virus, Colorado,
4430 USA". In: *Emerging Infectious Diseases* 17.5 (2011), p. 880.
- 4431 [149] R. Maciel-de Freitas, C. Codeco, and R. Lourenço-de Oliveira. "Body size-associated
4432 survival and dispersal rates of *Aedes aegypti* in Rio de Janeiro". In: *Medical and
4433 Veterinary Entomology* 21.3 (2007), pp. 284–292.
- 4434 [150] R. Maciel-de Freitas, Á. E. Eiras, and R. Lourenço-de Oliveira. "Calculating the
4435 survival rate and estimated population density of gravid *Aedes aegypti* (Diptera,
4436 Culicidae) in Rio de Janeiro, Brazil". In: *Cadernos de Saúde Pública* 24.12 (2008),
4437 pp. 2747–2754.
- 4438 [151] R. Maciel-de Freitas et al. "Movement of dengue vectors between the human mod-
4439 ified environment and an urban forest in Rio de Janeiro". In: *Journal of Medical
4440 Entomology* 43.6 (2006), pp. 1112–1120.
- 4441 [152] R. Galizi et al. "A synthetic sex ratio distortion system for the control of the human
4442 malaria mosquito". In: *Nature Communications* 5 (2014).
- 4443 [153] J. L. Gallup and J. D. Sachs. "The economic burden of malaria". In: *The American
4444 Journal of Tropical Medicine and Hygiene* 64.1 (2001), pp. 85–96.

- 4445 [154] V. M. Gantz et al. “Highly efficient Cas9-mediated gene drive for population modi-
4446 fication of the malaria vector mosquito *Anopheles stephensi*”. In: *Proceedings of the*
4447 *National Academy of Sciences* 112.49 (2015), E6736–E6743.
- 4448 [155] P. C. Garnham. *Malaria parasites and other haemosporidia*. Blackwell Scientific
4449 Publications Ltd., 1966.
- 4450 [156] B Gates. *The Deadliest Animal in the World*. [https://www.gatesnotes.com/Health/Most-](https://www.gatesnotes.com/Health/Most-Lethal-Animal-Mosquito-Week)
4451 *Lethal-Animal-Mosquito-Week*. 2014.
- 4452 [157] M. L. Gatton et al. “The importance of mosquito behavioural adaptations to malaria
4453 control in Africa”. In: *Evolution* 67.4 (2013), pp. 1218–1230.
- 4454 [158] A. Gelman and D. B. Rubin. “Inference from iterative simulation using multiple
4455 sequences”. In: *Statistical Science* (1992), pp. 457–472.
- 4456 [159] A. Gelman et al. *Bayesian data analysis*. Vol. 2. Taylor & Francis, 2014.
- 4457 [160] M. Germain, J.-P. Hervé, and B. Geoffroy. “Evaluation de la durée du cycle trophogo-
4458 nique d’*Aedes africanus* (Théobald), vecteur potentiel de fièvre jaune, dans une
4459 galerie forestière du sud de la République Centrafricaine”. In: *Cahiers ORSTOM.*
4460 *Série Entomologie Médicale et Parasitologie* 12.2 (1974), pp. 127–133.
- 4461 [161] A. Ghosh, S. Mandal, and G. Chandra. “Seasonal distribution, parity, resting, host-
4462 seeking behavior and association of malarial parasites of *Anopheles stephensi* Liston
4463 in Kolkata, West Bengal”. In: *Entomological Research* 40.1 (2010), pp. 46–54.
- 4464 [162] D. T. Gillespie. “Exact stochastic simulation of coupled chemical reactions”. In: *The*
4465 *Journal of Physical Chemistry* 81.25 (1977), pp. 2340–2361.
- 4466 [163] M. Gillies. “The duration of the gonotrophic cycle in *Anopheles gambiae* and *Anophe-*
4467 *les funestus*, with a note on the efficiency of hand catching”. In: *East African Medical*
4468 *Journal* 30.4 (1953).

- 4469 [164] M. Gillies and T. Wilkes. “A study of the age-composition of populations of *Anophe-*
4470 *les gambiae* Giles and *A. funestus* Giles in North-Eastern Tanzania”. In: *Bulletin of*
4471 *Entomological Research* 56.02 (1965), pp. 237–262.
- 4472 [165] M. Gillies and T. Wilkes. “Observations on nulliparous and parous rates in a pop-
4473 ulation of *Anopheles funestus* in East Africa”. In: *Annals of Tropical Medicine &*
4474 *Parasitology* 57.2 (1963), pp. 204–213.
- 4475 [166] M. Gillies and T. Wilkes. “The range of attraction of animal baits and carbon
4476 dioxide for mosquitoes. Studies in a freshwater area of West Africa”. In: *Bulletin of*
4477 *Entomological Research* 61.3 (1972), pp. 389–404.
- 4478 [167] J. E. Gimnig et al. “Characteristics of larval anopheline (Diptera: Culicidae) habi-
4479 tats in Western Kenya”. In: *Journal of Medical Entomology* 38.2 (2001), pp. 282–
4480 288.
- 4481 [168] J. E. Gimnig et al. “Density-dependent development of *Anopheles gambiae* (Diptera:
4482 Culicidae) larvae in artificial habitats”. In: *Journal of Medical Entomology* 39.1
4483 (2002), pp. 162–172.
- 4484 [169] A. K. Githeko et al. “Climate change and vector-borne diseases: a regional analysis”.
4485 In: *Bulletin of the World Health Organisation* 78.9 (2000), pp. 1136–1147.
- 4486 [170] M. R. Goddard and A. Burt. “Recurrent invasion and extinction of a selfish gene”.
4487 In: *Proceedings of the National Academy of Sciences* 96.24 (1999), pp. 13880–13885.
- 4488 [171] H. C. J. Godfray. “Mosquito ecology and control of malaria”. In: *Journal of Animal*
4489 *Ecology* 82.1 (2013), pp. 15–25.
- 4490 [172] H. C. J. Godfray, A. North, and A. Burt. “How driving endonuclease genes can be
4491 used to combat pests and disease vectors”. In: *BMC Biology* 15.1 (2017), p. 81.
- 4492 [173] W. C. Gorgas. *Sanitation in Panama*. Appleton, 1915.

- 4493 [174] B. Grassi. “Twenty-five Years After: A Chronicle of the discoveries relating to the
4494 mode of transmission of Human Malaria”. In: *Parasitology* 16.4 (1924), pp. 355–364.
- 4495 [175] B. M. Greenwood et al. “Malaria: progress, perils, and prospects for eradication”.
4496 In: *The Journal of Clinical Investigation* 118.4 (2008), p. 1266.
- 4497 [176] N. B. Grimm et al. “Global change and the ecology of cities”. In: *Science* 319.5864
4498 (2008), pp. 756–760.
- 4499 [177] H. Gruchet. “Etude de l’âge physiologique des femelles d’*Anopheles funestus* Giles
4500 dans la région de Miandrivaso, Madagascar”. In: *Bulletin de la Société de Pathologie*
4501 *Exotique* 55.1 (1962), pp. 165–174.
- 4502 [178] D. J. Gubler. “Dengue and dengue hemorrhagic fever”. In: *Clinical Microbiology*
4503 *Reviews* 11.3 (1998), pp. 480–496.
- 4504 [179] D. J. Gubler. “Epidemic dengue/dengue hemorrhagic fever as a public health, social
4505 and economic problem in the 21st century”. In: *Trends in Microbiology* 10.2 (2002),
4506 pp. 100–103.
- 4507 [180] D. J. Gubler. “The changing epidemiology of yellow fever and dengue, 1900 to 2003:
4508 full circle?” In: *Comparative Immunology, Microbiology and Infectious Diseases* 27.5
4509 (2004), pp. 319–330.
- 4510 [181] C. A. Guerra et al. “A global assembly of adult female mosquito mark-release-
4511 recapture data to inform the control of mosquito-borne pathogens”. In: *Parasite &*
4512 *Vectors* 7.1 (2014), p. 276.
- 4513 [182] D. Guha-Sapir and B. Schimmer. “Dengue fever: new paradigms for a changing
4514 epidemiology”. In: *Emerging Themes in Epidemiology* 2.1 (2005), p. 1.
- 4515 [183] S. A. Gumm. *Philadelphia Under Siege: The Yellow Fever of 1793*. <http://pabook2.libraries.psu.edu/>
4516 2010.

- 4517 [184] J. S. Haldane. “The rate of mutation of human genes”. In: *Hereditas* 35.S1 (1949),
4518 pp. 267–273.
- 4519 [185] S. B. Halstead. “Pathogenesis of dengue: challenges to molecular biology”. In: *Sci-*
4520 *ence* (1988), pp. 476–481.
- 4521 [186] W. D. Hamilton. “Extraordinary sex ratios”. In: *Science* 156.3774 (1967), pp. 477–
4522 488.
- 4523 [187] A. Hammond et al. “A CRISPR-Cas9 gene drive system targeting female reproduc-
4524 tion in the malaria mosquito vector *Anopheles gambiae*”. In: *Nature Biotechnology*
4525 34.1 (2016), p. 78.
- 4526 [188] P. Hancock, M. Thomas, and H. Godfray. “An age-structured model to evaluate
4527 the potential of novel malaria-control interventions: a case study of fungal biopesti-
4528 cide sprays”. In: *Proceedings of the Royal Society of London B: Biological Sciences*
4529 276.1654 (2009), pp. 71–80.
- 4530 [189] P. A. Hancock, S. P. Sinkins, and H. C. J. Godfray. “Population dynamic models of
4531 the spread of *Wolbachia*”. In: *The American Naturalist* 177.3 (2011), pp. 323–333.
- 4532 [190] P. A. Hancock and H. C. J. Godfray. “Application of the lumped age-class technique
4533 to studying the dynamics of malaria-mosquito-human interactions”. In: *Malaria*
4534 *Journal* 6.1 (2007), p. 98.
- 4535 [191] I. Hanski. “Metapopulation dynamics”. In: *Nature* 396.6706 (1998), p. 41.
- 4536 [192] L. C. Harrington et al. “Age-dependent survival of the dengue vector *Aedes aegypti*
4537 (Diptera: Culicidae) demonstrated by simultaneous release–recapture of different
4538 age cohorts”. In: *Journal of Medical Entomology* 45.2 (2008), pp. 307–313.

- 4539 [193] L. C. Harrington et al. "Analysis of survival of young and old *Aedes aegypti* (Diptera:
4540 Culicidae) from Puerto Rico and Thailand". In: *Journal of Medical Entomology* 38.4
4541 (2001), pp. 537–547.
- 4542 [194] S. I. Hay et al. "Urbanization, malaria transmission and disease burden in Africa".
4543 In: *Nature reviews. Microbiology* 3.1 (2005), p. 81.
- 4544 [195] E. B. Hayes. "Zika virus outside Africa". In: *Emerging Infectious Diseases* 15.9
4545 (2009), p. 1347.
- 4546 [196] G. A. Hazelton and C. A. Lang. "Glutathione levels during the mosquito life span
4547 with emphasis on senescence". In: *Experimental Biology and Medicine* 176.3 (1984),
4548 pp. 249–256.
- 4549 [197] E. Hempelmann and K. Krafts. "Bad air, amulets and mosquitoes: 2,000 years of
4550 changing perspectives on malaria". In: *Malaria Journal* 12.1 (2013), p. 232.
- 4551 [198] W. A. Hickey and G. B. Craig Jr. "Genetic distortion of sex ratio in a mosquito,
4552 *Aedes aegypti*". In: *Genetics* 53.6 (1966), p. 1177.
- 4553 [199] J. Hii, M. Birley, and V. Sang. "Estimation of survival rate and oviposition interval
4554 of *Anopheles balabacensis* mosquitoes from mark-recapture experiments in Sabah,
4555 Malaysia". In: *Medical and Veterinary Entomology* 4.2 (1990), pp. 135–140.
- 4556 [200] J. G. Hitchcock Jr. "Age composition of a natural population of *Anopheles quadri-*
4557 *maculatus* Say (Diptera: Culicidae) in Maryland, USA". In: *Journal of Medical*
4558 *Entomology* 5.1 (1968), pp. 125–134.
- 4559 [201] T. Hoc and T. Wilkes. "The ovariole structure of *Anopheles gambiae* (Diptera: Culi-
4560 cidae) and its use in determining physiological age". In: *Bulletin of Entomological*
4561 *Research* 85.01 (1995), pp. 59–69.

- 4562 [202] M. D. Hoffman and A. Gelman. “The No-U-turn sampler: adaptively setting path
4563 lengths in Hamiltonian Monte Carlo.” In: *Journal of Machine Learning Research*
4564 15.1 (2014), pp. 1593–1623.
- 4565 [203] E. Hsu. “The history of qing hao in the Chinese materia medica”. In: *Transactions*
4566 *of the Royal Society of Tropical Medicine and Hygiene* 100.6 (2006), pp. 505–508.
- 4567 [204] Z. Hubálek. “Mosquito-borne viruses in Europe”. In: *Parasitology research* 103.1
4568 (2008), pp. 29–43.
- 4569 [205] L. E. Hugo et al. “Adult survivorship of the dengue mosquito *Aedes aegypti* varies
4570 seasonally in central Vietnam”. In: *PLoS Neglected Tropical Diseases* 8.2 (2014),
4571 e2669.
- 4572 [206] L. E. Hugo et al. “Evaluations of mosquito age grading techniques based on mor-
4573 phological changes”. In: *Journal of Medical Entomology* 45.3 (2008), pp. 353–369.
- 4574 [207] R. A. Hutchinson and S. W. Lindsay. “Malaria and deaths in the English marshes”.
4575 In: *The Lancet* 367.9526 (2006), pp. 1947–1951.
- 4576 [208] R. Idro, N. E. Jenkins, and C. R. Newton. “Pathogenesis, clinical features, and
4577 neurological outcome of cerebral malaria”. In: *The Lancet Neurology* 4.12 (2005),
4578 pp. 827–840.
- 4579 [209] J. Ijumba, F. Mosha, and S. Lindsay. “Malaria transmission risk variations derived
4580 from different agricultural practices in an irrigated area of northern Tanzania”. In:
4581 *Medical and Veterinary Entomology* 16.1 (2002), pp. 28–38.
- 4582 [210] M. Imwong et al. “Spread of a single multidrug resistant malaria parasite lineage
4583 (PfPailin) to Vietnam”. In: *The Lancet Infectious Diseases* 17.10 (2017), pp. 1022–
4584 1023.

- 4585 [211] Z Jaal and W. MacDonald. “A mark-release-recapture experiment with *Anopheles*
4586 *lesteri paraliae* in northwest Peninsular Malaysia”. In: *Annals of Tropical Medicine*
4587 *& Parasitology* 86.4 (1992), pp. 419–424.
- 4588 [212] S. P. James. “The disappearance of malaria from England”. In: *Proceedings of the*
4589 *Royal Society of Medicine* 23 (1929), pp. 70–87.
- 4590 [213] S. Jayanetti, M. d. S. Wijesundera, and F. Amaresinghe. “A Study on the Bionomics
4591 of Indoor Resting Mosquitoes in Kandy”. In: (1987).
- 4592 [214] T. Jensen and R. K. Washino. “An assessment of the biological capacity of a Sacra-
4593 mento Valley population of *Aedes melanimon* to vector arboviruses”. In: *The Amer-*
4594 *ican Journal of Tropical Medicine and Hygiene* 44.4 (1991), pp. 355–363.
- 4595 [215] T. Jensen and R. K. Washino. “Comparison of recapture patterns of marked and
4596 released *Aedes vexans* and *Ae. melanimon* (Diptera: Culicidae) in the Sacramento
4597 Valley of California”. In: *Journal of Medical Entomology* 31.4 (1994), pp. 607–610.
- 4598 [216] E. S. Jentes et al. “The revised global yellow fever risk map and recommendations for
4599 vaccination, 2010: consensus of the Informal WHO Working Group on Geographic
4600 Risk for Yellow Fever”. In: *The Lancet Infectious Diseases* 11.8 (2011), pp. 622–632.
- 4601 [217] W. Jepson, A Moutia, and C Courtois. “The malaria problem in Mauritius: the
4602 bionomics of Mauritian anophelines”. In: *Bulletin of Entomological Research* 38.01
4603 (1947), pp. 177–208.
- 4604 [218] M. Jinek et al. “A programmable dual-RNA-guided DNA endonuclease in adaptive
4605 bacterial immunity”. In: *Science* 337.6096 (2012), pp. 816–821.
- 4606 [219] T Kanda, C. Joo, and D. Choi. “Epidemiological studies on Malayan filariasis in an
4607 inland area in Kyungpook, Korea, 2: The periodicity of the microfilariae and the

- 4608 bionomics of the vector [*Brugia malayi*, *Anopheles sinensis*].” In: *Mosquito News*
4609 (1975).
- 4610 [220] B. Kay. “Age structure of populations of *Culex annulirostris* (Diptera: Culicidae)
4611 at Kowanyama and Charleville, Queensland”. In: *Journal of Medical Entomology*
4612 16.4 (1979), pp. 309–316.
- 4613 [221] J. Keiser et al. “Urbanization in sub-saharan Africa and implication for malaria
4614 control”. In: *The American Journal of Tropical Medicine and Hygiene* 71.2_suppl
4615 (2004), pp. 118–127.
- 4616 [222] M. A. Kenawy. “Development and survival of *Anopheles pharoensis* and An. mul-
4617 ticolor from Faiyum, Egypt”. In: *J Am Mosq Control Assoc* 7.4 (1991), pp. 551–
4618 5.
- 4619 [223] D. Kinkela. *DDT and the American century: global health, environmental politics,*
4620 *and the pesticide that changed the world*. University of North Carolina Press, 2011.
- 4621 [224] D. L. Klayman. “Qinghaosu (artemisinin): an antimalarial drug from China”. In:
4622 *Science* 228 (1985), pp. 1049–1056.
- 4623 [225] C. Koenraadt, A. Githeko, and W Takken. “The effects of rainfall and evapotranspi-
4624 ration on the temporal dynamics of *Anopheles gambiae* ss and *Anopheles arabiensis*
4625 in a Kenyan village”. In: *Acta Tropica* 90.2 (2004), pp. 141–153.
- 4626 [226] R. Kohavi et al. “A study of cross-validation and bootstrap for accuracy estimation
4627 and model selection”. In: *Ijcai*. Vol. 14. 2. 1995, pp. 1137–1145.
- 4628 [227] M. U. Kraemer et al. “The global distribution of the arbovirus vectors *Aedes aegypti*
4629 and Ae. albopictus”. In: *eLife* 4 (2015), e08347.

- 4630 [228] U. Kroeckel et al. “New tools for surveillance of adult yellow fever mosquitoes:
4631 comparison of trap catches with human landing rates in an urban environment”. In:
4632 *Journal of the American Mosquito Control Association* 22.2 (2006), pp. 229–238.
- 4633 [229] K. G. Kuhn et al. “Malaria in Britain: past, present, and future”. In: *Proceedings*
4634 *of the National Academy of Sciences* 100.17 (2003), pp. 9997–10001.
- 4635 [230] M. Kuhn. “Caret package”. In: *Journal of Statistical Software* 28.5 (2008), pp. 1–26.
- 4636 [231] S. Kumar and S. B. Hedges. “A molecular timescale for vertebrate evolution”. In:
4637 *Nature* 392.6679 (1998), p. 917.
- 4638 [232] R. La Corte Dos Santos, O. P. Forattini, and M. N. Burattini. “*Anopheles albitarsis*
4639 *sl* (Diptera: Culicidae) survivorship and density in a rice irrigation area of the State
4640 of São Paulo, Brazil”. In: *Journal of Medical Entomology* 41.5 (2004), pp. 997–1000.
- 4641 [233] R. Lacroix et al. “Dispersal and survival of male and female *Aedes albopictus* (Diptera:
4642 Culicidae) on Reunion Island”. In: *Journal of Medical Entomology* 46.5 (2009),
4643 pp. 1117–1124.
- 4644 [234] K. Lakhani et al. “Estimated mortalities of the immature stages of *Aedes cantans*
4645 (Mg.) (Diptera, Culicidae) in a natural habitat”. In: *Bulletin of Entomological Re-*
4646 *search* 64.02 (1974), pp. 265–276.
- 4647 [235] L. Lambrechts et al. “Impact of daily temperature fluctuations on dengue virus
4648 transmission by *Aedes aegypti*”. In: *Proceedings of the National Academy of Sciences*
4649 108.18 (2011), pp. 7460–7465.
- 4650 [236] A. Lange and T. Hoc. *Abortive oogenesis and physiological age in blood-sucking*
4651 *mosquitoes (Diptera: Culicidae). Meditsinskaya Parazitologiya i Parazitarnye Bolezni*
4652 *50, 48–56. 1981.*

- 4653 [237] F. J. Lardeux et al. “A physiological time analysis of the duration of the gonotrophic
4654 cycle of *Anopheles pseudopunctipennis* and its implications for malaria transmission
4655 in Bolivia”. In: *Malaria Journal* 7.1 (2008), p. 141.
- 4656 [238] A. Le Menach et al. “An elaborated feeding cycle model for reductions in vectorial
4657 capacity of night-biting mosquitoes by insecticide-treated nets”. In: *Malaria Journal*
4658 6.1 (2007), p. 10.
- 4659 [239] J. A. A. Le Prince and T. H. D. Griffitts. “Flight of mosquitoes: studies on the
4660 distance of flight of *Anopheles quadrimaculatus*”. In: *Public Health Reports (1896-
4661 1970)* (1917), pp. 656–659.
- 4662 [240] R. Leardi and L. Nørgaard. “Sequential application of backward interval partial
4663 least squares and genetic algorithms for the selection of relevant spectral regions”.
4664 In: *Journal of chemometrics* 18.11 (2004), pp. 486–497.
- 4665 [241] Y. LeCun, Y. Bengio, and G. Hinton. “Deep learning”. In: *Nature* 521.7553 (2015),
4666 pp. 436–444.
- 4667 [242] R. S. Lees et al. “Review: Improving our knowledge of male mosquito biology in
4668 relation to genetic control programmes”. In: *Acta Tropica* 132 (2014), S2–S11.
- 4669 [243] T. Lehmann et al. “Aestivation of the African malaria mosquito, *Anopheles gambiae*
4670 in the Sahel”. In: *The American Journal of Tropical Medicine and Hygiene* 83.3
4671 (2010), pp. 601–606.
- 4672 [244] T. Lehmann et al. “The effective population size of *Anopheles gambiae* in Kenya:
4673 implications for population structure.” In: *Molecular Biology and Evolution* 15.3
4674 (1998), pp. 264–276.
- 4675 [245] C. Leighton and R. Foster. *Economic impacts of malaria in Kenya and Nigeria*. Abt
4676 Associates Bethesda, 1993.

- 4677 [246] K. Liebman et al. "The Influence of diet on the use of Near-infrared spectroscopy to
4678 determine the age of female *Aedes aegypti* mosquitoes". In: *The American Journal*
4679 *of Tropical Medicine and Hygiene* 92.5 (2015), pp. 1070–1075.
- 4680 [247] S. W. Lindsay, P. M. Emerson, and J. D. Charlwood. "Reducing malaria by mosquito-
4681 proofing houses". In: *Trends in Parasitology* 18.11 (2002), pp. 510–514.
- 4682 [248] J. Lines, J Myamba, and C. Curtis. "Experimental hut trials of permethrin-impregnated
4683 mosquito nets and eave curtains against malaria vectors in Tanzania". In: *Medical*
4684 *and Veterinary Entomology* 1.1 (1987), pp. 37–51.
- 4685 [249] J. Lines and N. Nassor. "DDT resistance in *Anopheles gambiae* declines with mosquito
4686 age". In: *Medical and Veterinary Entomology* 5.3 (1991), pp. 261–265.
- 4687 [250] J. Lines, T. Wilkes, and E. Lyimo. "Human malaria infectiousness measured by
4688 age-specific sporozoite rates in *Anopheles gambiae* in Tanzania". In: *Parasitology*
4689 102.2 (1991), pp. 167–177.
- 4690 [251] J. Lines et al. "Monitoring human-biting mosquitoes (Diptera: Culicidae) in Tan-
4691 zania with light-traps hung beside mosquito nets". In: *Bulletin of Entomological*
4692 *Research* 81.1 (1991), pp. 77–84.
- 4693 [252] F. B. Livingstone. *Abnormal hemoglobins in human populations*. Transaction Pub-
4694 lishers, 2009.
- 4695 [253] K. Loong et al. "Survival and feeding behaviour of Malaysian strain of *Anopheles*
4696 *maculatus* Theobald (Diptera: Culicidae) and their role in malaria transmission."
4697 In: *Tropical Biomedicine* 7.1 (1990), pp. 71–76.
- 4698 [254] C. Lorenz et al. "Morphometrical diagnosis of the malaria vectors *Anopheles cruzii*,
4699 *An. homunculus* and *An. bellator*". In: *Parasites & Vectors* 5.1 (2012), p. 257.

- 4700 [255] J. Lourenco et al. “Epidemiological and ecological determinants of Zika virus trans-
4701 mission in an urban setting.” In: *eLife* 6:e29820 (2017).
- 4702 [256] R. Lowe et al. “Reproductive behaviour of *Culex pipiens quinquefasciatus* released
4703 into a natural population.” In: *Mosquito News* 33.2 (1973), pp. 221–7.
- 4704 [257] J. Lutwama and L. Mukwaya. “Mark-release-recapture studies on three anthro-
4705 pophilic populations of *Aedes (Stegomyia) simpsoni* comple (Diptera: Culicidae) in
4706 Uganda”. In: *Bulletin of Entomological Research* 84.4 (1994), pp. 521–527.
- 4707 [258] G. Macdonald et al. *The epidemiology and control of malaria*. London, Oxford Uni-
4708 versity Press, 1957.
- 4709 [259] W. Macdonald, A Sebastian, and M. M. Tun. “A mark-release-recapture experiment
4710 with *Culex pipiens fatigans* in the village of Okpo, Burma”. In: *Annals of Tropical*
4711 *Medicine & Parasitology* 62.2 (1968), pp. 200–209.
- 4712 [260] R. Maciel-De-Freitas, C. T. Codeco, and R. Lourenco-De-Oliveira. “Daily survival
4713 rates and dispersal of *Aedes aegypti* females in Rio de Janeiro, Brazil”. In: *The*
4714 *American Journal of Tropical Medicine and Hygiene* 76.4 (2007), pp. 659–665.
- 4715 [261] G. MacPherson et al. “Human cerebral malaria. A quantitative ultrastructural anal-
4716 ysis of parasitized erythrocyte sequestration.” In: *The American Journal of Pathol-*
4717 *ogy* 119.3 (1985), p. 385.
- 4718 [262] S. Magesa et al. “Trial of pyrethroid impregnated bednets in an area of Tanzania
4719 holoendemic for malaria Part 2. Effects on the malaria vector population”. In: *Acta*
4720 *Tropica* 49.2 (1991), pp. 97–108.
- 4721 [263] B Mahanta et al. “Temporal variations in biting density and rhythm of *Culex quin-*
4722 *quefasciatus* in tea agro-ecosystem of Assam, India.” In: (1999).

- 4723 [264] F. Mahmood and W. K. Reisen. “Duration of the gonotrophic cycles of *Anopheles*
4724 *culicifacies* Giles and *Anopheles stephensi* Liston, with observations on reproductive
4725 activity and survivorship during winter in Punjab province, Pakistan.” In: *Mosquito*
4726 *News* (1981).
- 4727 [265] S. Maitra and J. Yan. “Principle component analysis and partial least squares: Two
4728 dimension reduction techniques for regression”. In: *Applying Multivariate Statistical*
4729 *Models* 79 (2008).
- 4730 [266] A. O. Mala et al. “Gonotrophic cycle duration, fecundity and parity of *Anopheles*
4731 *gambiae* complex mosquitoes during an extended period of dry weather in a semi
4732 arid area in Baringo County, Kenya”. In: *International Journal of Mosquito Research*
4733 1.2 (2014), pp. 28–34.
- 4734 [267] F Marini et al. “Study of *Aedes albopictus* dispersal in Rome, Italy, using sticky traps
4735 in mark–release–recapture experiments”. In: *Medical and Veterinary Entomology*
4736 24.4 (2010), pp. 361–368.
- 4737 [268] S. Marsland. *Machine learning: an algorithmic perspective*. CRC press, 2015.
- 4738 [269] E. M. Mathenge et al. “Comparative field evaluation of the Mbita trap, the Centers
4739 for Disease Control light trap, and the human landing catch for sampling of malaria
4740 vectors in western Kenya”. In: *The American Journal of Tropical Medicine and*
4741 *Hygiene* 70.1 (2004), pp. 33–37.
- 4742 [270] E. M. Mathenge et al. “Comparative performance of the Mbita trap, CDC light trap
4743 and the human landing catch in the sampling of *Anopheles arabiensis*, *An. funestus*
4744 and culicine species in a rice irrigation in western Kenya”. In: *Malaria Journal* 4.1
4745 (2005), p. 1.

- 4746 [271] MATLAB. *version 7.10.0 (R2010a)*. Natick, Massachusetts: The MathWorks Inc.,
4747 2010.
- 4748 [272] V. Maximus and J. Valerius. *Factorum et dictorum memorabilium libri novem: Cum*
4749 *Iulii Paridis et Ianuarii Nepotiani epitomis*. BG Teubner, 1888.
- 4750 [273] R. M. May. *Stability and complexity in model ecosystems*. Vol. 6. Princeton Univer-
4751 sity Press, 2001.
- 4752 [274] V. S. Mayagaya et al. “Non-destructive determination of age and species of *Anophe-*
4753 *les gambiae* sl using near-infrared spectroscopy”. In: *The American Journal of Trop-*
4754 *ical Medicine and Hygiene* 81.4 (2009), pp. 622–630.
- 4755 [275] V. S. Mayagaya et al. “Evaluating preservation methods for identifying *Anopheles*
4756 *gambiae* ss and *Anopheles arabiensis* complex mosquitoes species using near infra-
4757 red spectroscopy”. In: *Parasites & Vectors* 8.1 (2015), p. 60.
- 4758 [276] P. McDonald. “Population characteristics of domestic *Aedes aegypti* (Diptera: Culi-
4759 cidae) in villages on the Kenya Coast I. Adult survivorship and population size”.
4760 In: *Journal of Medical Entomology* 14.1 (1977), pp. 42–48.
- 4761 [277] C. J. McMeniman et al. “Stable introduction of a life-shortening *Wolbachia* infection
4762 into the mosquito *Aedes aegypti*”. In: *Science* 323.5910 (2009), pp. 141–144.
- 4763 [278] J. McSherry. “Some medical aspects of the Darien scheme: was it dengue?” In:
4764 *Scottish Medical Journal* 27.2 (1982), pp. 183–184.
- 4765 [279] B. de Meillon, A. Sebastian, and Z. Khan. “Time of arrival of gravid *Culex pipiens*
4766 *fatigans* at an oviposition site, the oviposition cycle and the relationship between
4767 time of feeding and time of oviposition”. In: *Bulletin of the World Health Organi-*
4768 *sation* 36.1 (1967), p. 39.

- 4769 [280] C Mendis et al. "Cordon sanitaire or laissez faire: differential dispersal of young and
4770 old females of the malaria vector *Anopheles funestus* Giles (Diptera: Culicidae) in
4771 southern Mozambique". In: *African Entomology* 6.1 (1998).
- 4772 [281] S. R. Meshnick and M. J. Dobson. "The history of antimalarial drugs". In: *Anti-*
4773 *malarial Chemotherapy* (2001), pp. 15–25.
- 4774 [282] J. T. Midega et al. "Estimating dispersal and survival of *Anopheles gambiae* and
4775 *Anopheles funestus* along the Kenyan coast by using mark–release–recapture meth-
4776 ods". In: *Journal of Medical Entomology* 44.6 (2007), pp. 923–929.
- 4777 [283] M. Milby and W. Reisen. "Estimation of vectorial capacity: vector survivorship."
4778 In: *Bulletin of the Society of Vector Ecology* 14.1 (1989), pp. 47–54.
- 4779 [284] J. Mlakar et al. "Zika virus associated with microcephaly". In: *The New England*
4780 *Journal of Medicine* 2016.374 (2016), pp. 951–958.
- 4781 [285] L. Molineaux, G. Gramiccia, World Health Organisation, et al. "The Garki project:
4782 research on the epidemiology and control of malaria in the Sudan savanna of West
4783 Africa". In: (1980).
- 4784 [286] T. P. Monath. "Yellow fever: an update". In: *The Lancet Infectious Diseases* 1.1
4785 (2001), pp. 11–20.
- 4786 [287] L. A. Moreira et al. "A *Wolbachia* symbiont in *Aedes aegypti* limits infection with
4787 dengue, Chikungunya, and *Plasmodium*". In: *Cell* 139.7 (2009), pp. 1268–1278.
- 4788 [288] A Mori, Y Wada, et al. "The gonotrophic cycle of *Aedes albopictus* in the field." In:
4789 *Tropical Medicine* 19.3/4 (1977), pp. 141–146.
- 4790 [289] J. E. Mueller et al. "Homing endonucleases". In: *Cold Spring Harbor Monograph*
4791 *Series* 25 (1993), pp. 111–111.

- 4792 [290] L. E. Muir and B. H. Kay. “*Aedes aegypti* survival and dispersal estimated by
4793 mark-release-recapture in northern Australia.” In: *The American Journal of Trop-*
4794 *ical Medicine and Hygiene* 58.3 (1998), pp. 277–282.
- 4795 [291] C. Mulamba et al. “Widespread pyrethroid and DDT resistance in the major malaria
4796 vector *Anopheles funestus* in East Africa is driven by metabolic resistance mecha-
- 4797 nisms”. In: *PLoS One* 9.10 (2014), e110058.
- 4798 [292] G Müller and Y Schlein. “Plant tissues: the frugal diet of mosquitoes in adverse
4799 conditions”. In: *Medical and Veterinary Entomology* 19.4 (2005), pp. 413–422.
- 4800 [293] G. C. Müller et al. “Successful field trial of attractive toxic sugar bait (ATSB)
4801 plant-spraying methods against malaria vectors in the *Anopheles gambiae* complex
4802 in Mali, West Africa”. In: *Malaria Journal* 9.1 (2010), p. 210.
- 4803 [294] S. M. Muriu et al. “Larval density dependence in *Anopheles gambiae* ss, the major
4804 African vector of malaria”. In: *Journal of Animal Ecology* 82.1 (2013), pp. 166–174.
- 4805 [295] C. J. Murray et al. “Disability-adjusted life years (DALYs) for 291 diseases and
4806 injuries in 21 regions, 1990–2010: a systematic analysis for the Global Burden of
4807 Disease Study 2010”. In: *The Lancet* 380.9859 (2013), pp. 2197–2223.
- 4808 [296] C. J. Murray et al. “Global malaria mortality between 1980 and 2010: a systematic
4809 analysis”. In: *The Lancet* 379.9814 (2012), pp. 413–431.
- 4810 [297] J. D. Murray. *Mathematical biology I*. Springer-Verlag, New York, 2002.
- 4811 [298] D Musso, E. Nilles, and V.-M. Cao-Lormeau. “Rapid spread of emerging Zika virus
4812 in the Pacific area”. In: *Clinical Microbiology and Infection* 20.10 (2014).
- 4813 [299] D. Musso et al. “Potential sexual transmission of Zika virus”. In: *Emerging Infectious*
4814 *Diseases* 21.2 (2015), p. 359.

- 4815 [300] J. A. Nájera. “Malaria control: achievements, problems and strategies.” In: *Paras-*
4816 *sitologia* 43.1-2 (2001), pp. 1–89.
- 4817 [301] J. A. Nájera, M. González-Silva, and P. L. Alonso. “Some lessons for the future
4818 from the Global Malaria Eradication Programme (1955–1969)”. In: *PLoS Medicine*
4819 8.1 (2011), e1000412.
- 4820 [302] M. B. Nathan. “Bancroftian filariasis in coastal North Trinidad, West Indies: in-
4821 tensity of transmission by *Culex quinquefasciatus*”. In: *Transactions of the Royal*
4822 *Society of Tropical Medicine and Hygiene* 75.5 (1981), pp. 721–730.
- 4823 [303] J. Nayar, M. Provost, and C. Hansen. “Quantitative bionomics of *Culex nigripalpus*
4824 (Diptera: Culicidae) populations in Florida: 2. Distribution, dispersal and survival
4825 patterns”. In: *Journal of Medical Entomology* 17.1 (1980), pp. 40–50.
- 4826 [304] J. Nayar and D. Sauerman. “A comparative study of flight performance and fuel
4827 utilization as a function of age in females of Florida mosquitoes”. In: *Journal of*
4828 *Insect Physiology* 19.10 (1973), pp. 1977–1988.
- 4829 [305] J. Nedelman. “A negative binomial model for sampling mosquitoes in a malaria
4830 survey”. In: *Biometrics* (1983), pp. 1009–1020.
- 4831 [306] R. Nelson and M. Milby. “Dispersal and survival of field and laboratory strains of
4832 *Culex tarsalis* (Diptera: Culicidae)”. In: *Journal of Medical Entomology* 17.2 (1980),
4833 pp. 146–150.
- 4834 [307] R. Nelson et al. “Estimates of survival, population size, and emergence of *Culex*
4835 *tarsalis* at an isolated site”. In: *Annals of the Entomological Society of America*
4836 71.5 (1978), pp. 801–808.

- 4837 [308] R. N’Guessan et al. “Reduced efficacy of insecticide-treated nets and indoor residual
4838 spraying for malaria control in pyrethroid resistance area, Benin”. In: *Emerging*
4839 *Infectious Diseases* 13.2 (2007), p. 199.
- 4840 [309] H Nobuchi. “The symptoms of a dengue-like illness recorded in a Chinese medical
4841 encyclopedia”. In: *Kanpo Rinsho* 26 (1979), pp. 422–425.
- 4842 [310] L Norgaard et al. “Interval partial least-squares regression (iPLS): a comparative
4843 chemometric study with an example from near-infrared spectroscopy”. In: *Applied*
4844 *Spectroscopy* 54.3 (2000), pp. 413–419.
- 4845 [311] E. Norrby. “Yellow fever and Max Theiler: the only Nobel Prize for a virus vaccine”.
4846 In: *Journal of Experimental Medicine* 204.12 (2007), pp. 2779–2784.
- 4847 [312] A. North, A. Burt, and H. C. J. Godfray. “Modelling the spatial spread of a homing
4848 endonuclease gene in a mosquito population”. In: *Journal of Applied Ecology* 50.5
4849 (2013), pp. 1216–1225.
- 4850 [313] V. N. Novoseltsev et al. “An age-structured extension to the vectorial capacity
4851 model”. In: *PLoS One* 7.6 (2012).
- 4852 [314] A. J. Ntamatungiro et al. “The influence of physiological status on age prediction
4853 of *Anopheles arabiensis* using near infra-red spectroscopy”. In: *Parasites & Vectors*
4854 6.1 (2013), p. 298.
- 4855 [315] A. Oliveira Melo et al. “Zika virus intrauterine infection causes fetal brain abnormal-
4856 ity and microcephaly: tip of the iceberg?” In: *Ultrasound in Obstetrics & Gynecology*
4857 47.1 (2016), pp. 6–7.
- 4858 [316] J. G. Ordóñez González et al. “The use of sticky ovitraps to estimate dispersal
4859 of *Aedes aegypti* in northeastern Mexico”. In: *Journal of the American Mosquito*
4860 *Control Association* 17.2 (2001), pp. 93–97.

- 4861 [317] K. P. Paaijmans, A. F. Read, and M. B. Thomas. “Understanding the link between
4862 malaria risk and climate”. In: *Proceedings of the National Academy of Sciences*
4863 106.33 (2009), pp. 13844–13849.
- 4864 [318] M. A. Pacheco et al. “Malarial parasite diversity in chimpanzees: the value of com-
4865 parative approaches to ascertain the evolution of *Plasmodium falciparum* antigens”.
4866 In: *Malaria Journal* 12.1 (2013), p. 328.
- 4867 [319] R. M. Packard. *The making of a tropical disease: a short history of malaria*. Johns
4868 Hopkins University Press, 2007.
- 4869 [320] C. P. Pant, World Health Organisation, et al. “Distribution of anophelines in relation
4870 to altitude in Nepal”. In: (1962).
- 4871 [321] C. Pant and M Yasuno. “Field studies on the gonotrophic cycle of *Aedes aegypti* in
4872 Bangkok, Thailand”. In: *Journal of Medical Entomology* 10.2 (1973), pp. 219–223.
- 4873 [322] J. Partecke, I. Schwabl, and E. Gwinner. “Stress and the city: urbanization and its
4874 effects on the stress physiology in European blackbirds”. In: *Ecology* 87.8 (2006),
4875 pp. 1945–1952.
- 4876 [323] H. Pates and C. Curtis. “Mosquito behavior and vector control”. In: *Annual Review*
4877 *of Entomology* 50 (2005), pp. 53–70.
- 4878 [324] J. A. Patz and S. H. Olson. “Malaria risk and temperature: influences from global cli-
4879 mate change and local land use practices”. In: *Proceedings of the National Academy*
4880 *of Sciences* 103.15 (2006), pp. 5635–5636.
- 4881 [325] L. Pauling et al. “Sickle cell anemia, a molecular disease”. In: (1949).
- 4882 [326] R. Penilla et al. “Pteridine concentrations differ between insectary-reared and field-
4883 collected *Anopheles albimanus* mosquitoes of the same physiological age”. In: *Med-*
4884 *ical and Veterinary Entomology* 16.3 (2002), pp. 225–234.

- 4885 [327] J. Perez-Mendoza et al. “Chronological age-grading of house flies by using near-
4886 infrared spectroscopy”. In: *Journal of Medical Entomology* 39.3 (2002), pp. 499–
4887 508.
- 4888 [328] G. Pialoux et al. “Chikungunya, an epidemic arbovirosis”. In: *The Lancet Infectious*
4889 *Diseases* 7.5 (2007), pp. 319–327.
- 4890 [329] G. Pialoux et al. “Chikungunya, an epidemic arbovirosis”. In: *The Lancet Infectious*
4891 *Diseases* 7.5 (2007), pp. 319–327.
- 4892 [330] F. B. Piel et al. “Global distribution of the sickle cell gene and geographical confir-
4893 mation of the malaria hypothesis”. In: *Nature Communications* 1 (2010), p. 104.
- 4894 [331] V. Polovodova. “The determination of the physiological age of female *Anopheles*
4895 by the number of gonotrophic cycles completed”. In: *Meditinskaja Parazitologija*
4896 *Parazitar Bolezni* 18 (1949), pp. 352–355.
- 4897 [332] J. H. Powell. *Bring out your dead: the great plague of yellow fever in Philadelphia*
4898 *in 1793*. University of Pennsylvania Press, 1949.
- 4899 [333] C. Pumpuni and E. Walker. “Population size and survivorship of adult *Aedes trise-*
4900 *riatus* in a scrap tireyard in northern Indiana.” In: *Journal of the American Mosquito*
4901 *Control Association* 5.2 (1989), pp. 166–172.
- 4902 [334] W. A. Qualls et al. “Indoor use of attractive toxic sugar bait (ATSB) to effectively
4903 control malaria vectors in Mali, West Africa”. In: *Malaria Journal* 14.1 (2015),
4904 p. 301.
- 4905 [335] M. Quinones et al. “*Anopheles gambiae* gonotrophic cycle duration, biting and exit-
4906 ing behaviour unaffected by permethrin-impregnated bednets in The Gambia”. In:
4907 *Medical and Veterinary Entomology* 11.1 (1997), pp. 71–78.

- 4908 [336] P. Rajagopalan. "Population dynamics of *Culex pipiens fatigans*, the filariasis vector,
4909 in pondicherry: influence of climate and environment". In: *Proceedings of the Indian*
4910 *National Sciences Academy*. Vol. 46. 1980, pp. 745–52.
- 4911 [337] K. Ramaiah and P. Das. "Non-involvement of nulliparous females in the transmission
4912 of bancroftian filariasis". In: *Acta Tropica* 52.2-3 (1992), pp. 149–153.
- 4913 [338] H. Ranson and N. Lissenden. "Insecticide resistance in African *Anopheles* mosquitoes:
4914 a worsening situation that needs urgent action to maintain malaria control". In:
4915 *Trends in Parasitology* 32.3 (2016), pp. 187–196.
- 4916 [339] H. Ranson et al. "Pyrethroid resistance in African anopheline mosquitoes: what
4917 are the implications for malaria control?" In: *Trends in Parasitology* 27.2 (2011),
4918 pp. 91–98.
- 4919 [340] J. L. Rasgon, L. M. Styer, and T. W. Scott. "*Wolbachia*-induced mortality as a
4920 mechanism to modulate pathogen transmission by vector arthropods". In: *Journal*
4921 *of Medical Entomology* 40.2 (2003), pp. 125–132.
- 4922 [341] P. Rawlings and C. F. Curtis. "Tests for the existence of genetic variability in the
4923 tendency of *Anopheles culicifacies* species B to rest in houses and to bite man". In:
4924 *Bulletin of the World Health Organisation* 60.3 (1982), p. 427.
- 4925 [342] P. Rawlings and G. Davidson. "The dispersal and survival of *Anopheles culicifacies*
4926 *Giles* (Diptera: Culicidae) in a Sri Lankan village under malathion spraying". In:
4927 *Bulletin of Entomological Research* 72.1 (1982), pp. 139–144.
- 4928 [343] P. Rawlings et al. "The influence of age and season on dispersal and recapture of
4929 *Anopheles culicifacies* in Sri Lanka". In: *Ecological Entomology* 6.3 (1981), pp. 307–
4930 319.

- 4931 [344] A. F. Read, P. A. Lynch, M. B. Thomas, et al. “How to make evolution-proof
4932 insecticides for malaria control”. In: *PLoS Biol* 7.4 (2009), e1000058.
- 4933 [345] M. R. Reddy et al. “Outdoor host seeking behaviour of *Anopheles gambiae* mosquitoes
4934 following initiation of malaria vector control on Bioko Island, Equatorial Guinea”.
4935 In: *Malaria Journal* 10.1 (2011), p. 184.
- 4936 [346] R. C. Reiner et al. “Seasonality of *Plasmodium falciparum* transmission: a system-
4937 atic review”. In: *Malaria Journal* 14.1 (2015), p. 343.
- 4938 [347] W. K. Reisen, F. Mahmood, and K. Azra. “*Anopheles culicifacies giles*: Adult eco-
4939 logical parameters measured in rural punjab province, pakistan using capture-mark-
4940 release-recapture and dissection methods, with comparative observations on *An.*
4941 *stephensi liston* and *An. subpictus grassi*”. In: *Researches on Population Ecology*
4942 23.1 (1981), pp. 39–60.
- 4943 [348] W. K. Reisen, F. Mahmood, and T. Parveen. “*Anopheles culicifacies Giles*: a release-
4944 recapture experiment with cohorts of known age with implications for malaria epi-
4945 demiology and genetical control in Pakistan”. In: *Transactions of the Royal Society*
4946 *of Tropical Medicine and Hygiene* 74.3 (1980), pp. 307–317.
- 4947 [349] W. K. Reisen, F. Mahmood, and T. Parveen. “*Anopheles subpictus Grassi*: ob-
4948 servations on survivorship and population size using mark-release-recapture and
4949 dissection methods”. In: *Researches on Population Ecology* 21.1 (1979), pp. 12–29.
- 4950 [350] W. K. Reisen et al. “Population ecology of *Culex tarsalis* (Diptera: Culicidae) in a
4951 foothill environment of Kern County, California: temporal changes in female relative
4952 abundance, reproductive status, and survivorship”. In: *Annals of the Entomological*
4953 *Society of America* 76.4 (1983), pp. 800–808.

- 4954 [351] W. Reisen et al. “*Anopheles culicifacies*: observations on population ecology and
4955 reproductive behavior.” In: *Mosquito News* 42.1 (1982), pp. 93–101.
- 4956 [352] W. Reisen et al. “The impact of aerial applications of ultra-low volume adulticides
4957 on *Culex tarsalis* populations (Diptera: Culicidae) in Kern County, California, USA,
4958 1982”. In: *Journal of Medical Entomology* 21.5 (1984), pp. 573–585.
- 4959 [353] P. Reiter. “From Shakespeare to Defoe: malaria in England in the Little Ice Age.”
4960 In: *Emerging Infectious Diseases* 6.1 (2000), p. 1.
- 4961 [354] P. Reiter. “Oviposition, dispersal, and survival in *Aedes aegypti*: implications for the
4962 efficacy of control strategies”. In: *Vector-Borne and Zoonotic Diseases* 7.2 (2007),
4963 pp. 261–273.
- 4964 [355] M Renshaw, M. Service, and M. Birley. “Host finding, feeding patterns and evi-
4965 dence for a memorized home range of the mosquito *Aedes cantans*”. In: *Medical and
4966 Veterinary Entomology* 8.2 (1994), pp. 187–193.
- 4967 [356] G Rezza et al. “Infection with chikungunya virus in Italy: an outbreak in a temperate
4968 region”. In: *The Lancet* 370.9602 (2007), pp. 1840–1846.
- 4969 [357] D. R. Roberts, S. Manguin, and J. Mouchet. “DDT house spraying and re-emerging
4970 malaria”. In: *The Lancet* 356.9226 (2000), pp. 330–332.
- 4971 [358] D. Robinson. “Particle Motion and Segregation During the Mixing of Solid Partic-
4972 ulate Materials”. PhD thesis. King’s College London, 1967.
- 4973 [359] M. H. Rodriguez et al. “Gonotrophic cycle and survivorship of *Anopheles albimanus*
4974 (Diptera: Culicidae) in southern Mexico”. In: *Journal of Medical Entomology* 29.3
4975 (1992), pp. 395–399.
- 4976 [360] A. Rohatgi. *WebPlotDigitizer*. Version 3.11. Austin, Texas, USA, 2017.
- 4977 [361] R. Ross. *The prevention of malaria*. Dutton, 1910.

- 4978 [362] R. Ross et al. *Memoirs: with a Full Account of the Great Malaria Problem and its*
4979 *Solution*. John Murray, 1923.
- 4980 [363] G. L. Rúa et al. “Laboratory estimation of the effects of increasing temperatures
4981 on the duration of gonotrophic cycle of *Anopheles albimanus* (Diptera: Culicidae)”.
4982 In: *Memorias Do Instituto Oswaldo Cruz* 100.5 (2005), pp. 515–520.
- 4983 [364] R. C. Russell. “*Culex annulirostris* Skuse (Diptera: Culicidae) at Appin, NSW? bionomics
4984 and behaviour”. In: *Austral Entomology* 25.2 (1986), pp. 103–109.
- 4985 [365] R. C. Russell. “Population age composition and female longevity of the arbovirus
4986 vector *Culex annulirostris skuse* near Echua, Victoria, in the Murray Valley of south-
4987 eastern Australia 1979-1985”. In: *Australian Journal of Experimental Biology & Med-*
4988 *ical Science* 64.6 (1986).
- 4989 [366] J. Sachs and P. Malaney. “The economic and social burden of malaria”. In: *Nature*
4990 415.6872 (2002), pp. 680–685.
- 4991 [367] R. Sallares. *Malaria and Rome: a history of malaria in ancient Italy*. Oxford Uni-
4992 versity Press, 2002.
- 4993 [368] W. Samarawickrema. “A study of the age-composition of natural populations of
4994 *Culex pipiens fatigans* Wiedemann in relation to the transmission of filariasis due
4995 to *Wuchereria bancrofti* (Cobbold) in Ceylon”. In: *Bulletin of the World Health*
4996 *Organisation* 37.1 (1967), p. 117.
- 4997 [369] W. Samarawickrema. “Biting cycles and parity of the mosquito *Mansonia* (Manso-
4998 nioides) uniformis (Theo.) in Ceylon”. In: *Bulletin of Entomological research* 58.02
4999 (1968), pp. 299–314.
- 5000 [370] W. Samarawickrema, F. Sone, and R. Cummings. “Seasonal abundance, diel bit-
5001 ing activity and parity of *Aedes polynesiensis marks* and *A. samoanus* (Grünberg)

- (Diptera: Culicidae) in Samoa”. In: *Bulletin of Entomological Research* 77.2 (1987), pp. 191–200.
- [371] Y. Schlein and G. C. Müller. “An approach to mosquito control: Using the dominant attraction of flowering *Tamarix jordanis* trees against *Culex pipiens*”. In: *Journal of Medical Entomology* 45.3 (2008), pp. 384–390.
- [372] Y. Schlein and G. C. Müller. “Decrease of larval and subsequent adult *Anopheles sergentii* populations following feeding of adult mosquitoes from *Bacillus sphaericus*-containing attractive sugar baits”. In: *Parasites & Vectors* 8.1 (2015), p. 244.
- [373] Y. Schlein and G. C. Müller. “Diurnal resting behavior of adult *Culex pipiens* in an arid habitat in Israel and possible control measurements with toxic sugar baits”. In: *Acta Tropica* 124.1 (2012), pp. 48–53.
- [374] D. Schmidt and C. Heckendorf. *Guide to the ngram Package: Fast n-gram Tokenization*. R Vignette. 2017. URL: <https://cran.r-project.org/package=ngram>.
- [375] T. L. Schmidt et al. “Genome-wide SNPs reveal the drivers of gene flow in an urban population of the Asian Tiger Mosquito, *Aedes albopictus*”. In: *PLoS Neglected Tropical Diseases* 11.10 (2017), e0006009.
- [376] P. Scholl, C. H. Porter, and G. R. Defoliart. “*Aedes triseriatus*: persistence of nulliparous females under field conditions”. In: *Mosquito News* 39 (1979), pp. 368–371.
- [377] F. W. Selck, A. A. Adalja, and C. R. Boddie. “An estimate of the global health care and lost productivity costs of dengue”. In: *Vector-Borne and Zoonotic Diseases* 14.11 (2014), pp. 824–826.

- 5024 [378] S. Sempala. “The ecology of *Aedes* (*Stegomyia*) *africanus* (Theobald) in a tropical
5025 forest in Uganda: mark-release-recapture studies on a female adult population”. In:
5026 *International Journal of Tropical Insect Science* 1.3 (1981), pp. 211–224.
- 5027 [379] Q. Serenus and R. Pépin. *Liber medicinalis*. Presses Universitaires de France, 1950.
- 5028 [380] M. Service. “Studies on sampling larval populations of the *Anopheles gambiae* com-
5029 plex”. In: *Bulletin of the World Health Organisation* 45.2 (1971), p. 169.
- 5030 [381] A. Shalaby, World Health Organisation, et al. “Studies on the age composition of
5031 *Anopheles Culicifacies* Giles at different phases of the development of resistance to
5032 DDT in Panchmahals district of Gujarat State, India”. In: (1962).
- 5033 [382] M. A. Shampo and R. A. Kyle. “Nei Ching - oldest known medical book”. In: *Mayo*
5034 *Clinic Proceedings*. Vol. 64. 1. Elsevier. 1989, p. 134.
- 5035 [383] P. Sheppard et al. “The dynamics of an adult population of *Aedes aegypti* in relation
5036 to dengue haemorrhagic fever in Bangkok”. In: *The Journal of Animal Ecology*
5037 (1969), pp. 661–702.
- 5038 [384] J. Shililu et al. “Larval habitat diversity and ecology of anopheline larvae in Eritrea”.
5039 In: *Journal of Medical Entomology* 40.6 (2003), pp. 921–929.
- 5040 [385] A. Shriram and K Krishnamoorthy. “Population dynamics, age composition and
5041 survival of *Downsiomyia nivea* in relation to transmission of diurnally subperiodic
5042 filariasis”. In: *Journal of Asia-Pacific Entomology* 14.1 (2011), pp. 34–40.
- 5043 [386] M. Sikulu et al. “Evaluating RNAlater® as a preservative for using near-infrared
5044 spectroscopy to predict *Anopheles gambiae* age and species”. In: *Malaria Journal*
5045 10.1 (2011), p. 186.

- 5046 [387] M. Sikulu et al. “Near-infrared spectroscopy as a complementary age grading and
5047 species identification tool for African malaria vectors”. In: *Parasites & Vectors*
5048 (2010).
- 5049 [388] M. T. Sikulu et al. “Using a near-infrared spectrometer to estimate the age of
5050 anopheles mosquitoes exposed to pyrethroids”. In: *PLoS One* 9.3 (2014), e90657.
- 5051 [389] M. T. Sikulu-Lord et al. “Near-infrared spectroscopy, a rapid method for predicting
5052 the age of male and female wild-type and *Wolbachia* infected *Aedes aegypti*”. In:
5053 *PLoS Neglected Tropical Diseases* 10.10 (2016), e0005040.
- 5054 [390] J. B. Silver. *Mosquito ecology: field sampling methods*. Springer Science & Business
5055 Media, 2007.
- 5056 [391] A. Simoni et al. “Development of synthetic selfish elements based on modular nucle-
5057 ases in *Drosophila melanogaster*”. In: *Nucleic Acids Research* 42.11 (2014), pp. 7461–
5058 7472.
- 5059 [392] B. Singh et al. “A large focus of naturally acquired *Plasmodium knowlesi* infections
5060 in human beings”. In: *The Lancet* 363.9414 (2004), pp. 1017–1024.
- 5061 [393] M. E. Sinka et al. “A global map of dominant malaria vectors”. In: *Parasites &*
5062 *Vectors* 5.1 (2012), p. 1.
- 5063 [394] M. E. Sinka et al. “The dominant *Anopheles* vectors of human malaria in Africa,
5064 Europe and the Middle East: occurrence data, distribution maps and bionomic
5065 précis”. In: *Parasites & Vectors* 3.1 (2010), p. 117.
- 5066 [395] A Smith and D. Webley. “A veranda-trap hut for studying the house-frequenting
5067 habits of mosquitoes and for assessing insecticides. III. The effect of DDT on be-
5068 haviour and mortality”. In: *Bulletin of Entomological Research* 59.1 (1969), pp. 33–
5069 46.

- [396] D. L. Smith, J. Dushoff, and F. E. McKenzie. “The risk of a mosquito-borne infection in a heterogeneous environment”. In: *PLoS Biology* 2.11 (2004), e368.
- [397] D. L. Smith and F. E. McKenzie. “Statics and dynamics of malaria infection in *Anopheles* mosquitoes”. In: *Malaria Journal* 3.1 (2004), p. 13.
- [398] D. L. Smith et al. “Ross, Macdonald, and a theory for the dynamics and control of mosquito-transmitted pathogens”. In: *PLoS Pathogens* 8.4 (2012), e1002588.
- [399] S. M. Smith and R. Kurtz. “The age structure of a population of *Aedes provocans* (Diptera: Culicidae) in southwestern Ontario”. In: *Great Lakes Entomologist* 27.2 (1994), p. 113.
- [400] R. W. Snow et al. “The prevalence of *Plasmodium falciparum* in sub-Saharan Africa since 1900”. In: *Nature* 550.7677 (2017), p. 515.
- [401] W. Snow and P. Boreham. “The host-feeding patterns of some culicine mosquitoes (Diptera: Culicidae) in The Gambia”. In: *Bulletin of Entomological Research* 68.4 (1978), pp. 695–706.
- [402] F. Snowden. *The conquest of malaria: Italy, 1900-1962*. Yale University Press, 2008.
- [403] F. L. Soper, D. B. Wilson, et al. *Anopheles gambiae in Brazil 1930 to 1940*. New York City: The Rockefeller Foundation., 1943.
- [404] A. Spielman and M. D’Antonio. *Mosquito: The story of man’s deadliest foe*. Hachette Books, 2001.
- [405] Stan Development Team. *Stan: A C++ Library for Probability and Sampling, Version 2.5.0*. 2014. URL: <http://mc-stan.org/>.
- [406] B. L. Stoddard. “Homing endonuclease structure and function”. In: *Quarterly Reviews of Biophysics* 38.1 (2005), pp. 49–95.

- 5093 [407] L. M. Styer et al. "Mosquitoes do senesce: departure from the paradigm of constant
5094 mortality". In: *The American Journal of Tropical Medicine and Hygiene* 76.1 (2007),
5095 pp. 111–117.
- 5096 [408] S. Surendran et al. "*Anopheles culicifacies* sibling species B and E in Sri Lanka differ
5097 in longevity and in their susceptibility to malaria parasite infection and common
5098 insecticides". In: *Medical and Veterinary Entomology* 20.1 (2006), pp. 153–156.
- 5099 [409] T. Suzuki. "Preliminary studies on blood meal interval of *Aedes polynesiensis* in the
5100 field". In: *Medical Entomology and Zoology* 29.2 (1978), pp. 169–174.
- 5101 [410] W. Tabachnick et al. "Oral infection of *Aedes aegypti* with yellow fever virus: geo-
5102 graphic variation and genetic considerations". In: *The American Journal of Tropical*
5103 *Medicine and Hygiene* 34.6 (1985), pp. 1219–1224.
- 5104 [411] M. Takagi et al. "Movement of individually marked *Aedes albopictus* females in
5105 Nagasaki, Japan." In: *Tropical Medicine* 37.2 (1995), pp. 79–85.
- 5106 [412] W. Takken et al. "Dispersal and survival of *Anopheles funestus* and *A. gambiae* sl
5107 (Diptera: Culicidae) during the rainy season in southeast Tanzania". In: *Bulletin of*
5108 *Entomological Research* 88.5 (1998), pp. 561–566.
- 5109 [413] W. Takken. "Do insecticide-treated bednets have an effect on malaria vectors?" In:
5110 *Tropical Medicine & International Health* 7.12 (2002), pp. 1022–1030.
- 5111 [414] M. Tanner et al. "Malaria eradication and elimination: views on how to translate a
5112 vision into reality". In: *BMC Medicine* 13.1 (2015), p. 167.
- 5113 [415] Tanogabo. *Dea Febris*. [https://tanogaboblog.it/portalino/il-giorno-della-candelora/dea-](https://tanogaboblog.it/portalino/il-giorno-della-candelora/dea-febris/)
5114 [febris/](https://tanogaboblog.it/portalino/il-giorno-della-candelora/dea-febris/). 2015.
- 5115 [416] J. Taverne. "DDT—to ban or not to ban?" In: *Parasitology Today* 15.5 (1999),
5116 pp. 180–181.

- 5117 [417] C. Taylor et al. "Gene flow among populations of the malaria vector, *Anopheles*
5118 *gambiae*, in Mali, West Africa". In: *Genetics* 157.2 (2001), pp. 743–750.
- 5119 [418] M. J. Taylor, A. Hoerauf, and M. Bockarie. "Lymphatic filariasis and onchocercia-
5120 sis". In: *The Lancet* 376.9747 (2010), pp. 1175–1185.
- 5121 [419] M. B. Thomas and A. F. Read. "Can fungal biopesticides control malaria?" In:
5122 *Nature Reviews Microbiology* 5.5 (2007), pp. 377–383.
- 5123 [420] Y. T. Touré et al. "Mark–release–recapture experiments with *Anopheles gambiae*
5124 *sl* in Banambani Village, Mali, to determine population size and structure". In:
5125 *Medical and Veterinary Entomology* 12.1 (1998), pp. 74–83.
- 5126 [421] M. Trpis and W. Hausermann. "Demonstration of differential domesticity of *Aedes*
5127 *aegypti* (L.)(Diptera, Culicidae) in Africa by mark-release-recapture". In: *Bulletin*
5128 *of Entomological Research* 65.2 (1975), pp. 199–208.
- 5129 [422] M. Trpis, W. Häusermann, and G. B. Craig Jr. "Estimates of population size, dis-
5130 persal, and longevity of domestic *Aedes aegypti aegypti* (Diptera: Culicidae) by
5131 mark–release–recapture in the village of Shauri Moyo in eastern Kenya". In: *Jour-*
5132 *nal of Medical Entomology* 32.1 (1995), pp. 27–33.
- 5133 [423] Y. Tsuda et al. "Movement of *Aedes aegypti* (Diptera: Culicidae) released in a small
5134 isolated village on Hainan Island, China". In: *Journal of Medical Entomology* 38.1
5135 (2001), pp. 93–98.
- 5136 [424] P. Tuchinda et al. "On the diurnal rhythmus of biting behavior of *Aëdes aegypti* in
5137 relation to the age and to the hemorrhagic fever in Bangkok, 1964". In: *Japanese*
5138 *Journal of Tropical Medicine* 10.1 (1969), pp. 1–6.
- 5139 [425] L. S. Tusting et al. "The evidence for improving housing to reduce malaria: a sys-
5140 tematic review and meta-analysis". In: *Malaria Journal* 14.1 (2015), p. 209.

- 5141 [426] UN. *World urbanization prospects: the 2001 revision*. Vol. 216. United Nations Pub-
5142 lications, 2002.
- 5143 [427] University of Sydney. *BIOL2007 Introductory Entomology course website*. 2004.
- 5144 [428] E. C. Uttah et al. “Physiological age composition of female anopheline mosquitoes
5145 in an area endemic for malaria and filariasis”. In: *International Journal of Scientific
5146 and Research Publications* 3.7 (2013), pp. 1–4.
- 5147 [429] J. Vainio, F. Cutts, World Health Organisation, et al. “Yellow fever”. In: (1998).
- 5148 [430] L. Valerio et al. “Dispersal of male *Aedes aegypti* in a coastal village in southern
5149 Mexico”. In: *The American Journal of Tropical Medicine and Hygiene* 86.4 (2012),
5150 pp. 665–676.
- 5151 [431] A. Vehtari, A. Gelman, and J. Gabry. “Efficient implementation of leave-one-out
5152 cross-validation and WAIC for evaluating fitted Bayesian models”. In: *arXiv preprint
5153 arXiv:1507.04544* (2015).
- 5154 [432] N. O. Verhulst, J. A. Loonen, and W. Takken. “Advances in methods for colour
5155 marking of mosquitoes”. In: *Parasites & Vectors* 6.1 (2013), p. 1.
- 5156 [433] I. Vythilingam et al. “Abundance, parity, and Japanese encephalitis virus infection
5157 of mosquitoes (Diptera: Culicidae) in Sepang District, Malaysia”. In: *Journal of
5158 Medical Entomology* 34.3 (1997), pp. 257–262.
- 5159 [434] Y. Wada et al. “Dispersal experiment of *Culex tritaeniorhynchus* in Nagasaki area
5160 (Preliminary report)”. In: *Tropical Medicine* 11.1 (1969), pp. 37–44.
- 5161 [435] T. M. Watson, A. Saul, and B. H. Kay. “*Aedes notoscriptus* (Diptera: Culicidae)
5162 survival and dispersal estimated by mark-release-recapture in Brisbane, Queensland,
5163 Australia”. In: *Journal of Medical Entomology* 37.3 (2000), pp. 380–384.

- 5164 [436] E. Westaway and J Blok. “Taxonomy and evolutionary relationships of flaviviruses”.
5165 In: *Dengue and dengue hemorrhagic fever* (1997), pp. 147–173.
- 5166 [437] E. Weterings and D. J. Chen. “The endless tale of non-homologous end-joining”.
5167 In: *Cell research* 18.1 (2008), p. 114.
- 5168 [438] M. T. White et al. “Modelling the impact of vector control interventions on *Anophe-*
5169 *les gambiae* population dynamics”. In: *Parasites & Vectors* 4 (2011), p. 153.
- 5170 [439] N. J. White. “Artemisinin resistance - the clock is ticking”. In: *The Lancet* 376.9758
5171 (2010), pp. 2051–2052.
- 5172 [440] J. Whitfield. “Portrait of a serial killer”. In: *Nature news* (2002).
- 5173 [441] T. Wilkes, Y. Matola, and J. Charlwood. “*Anopheles rivulorum*, a vector of human
5174 malaria in Africa”. In: *Medical and Veterinary Entomology* 10.1 (1996), pp. 108–
5175 110.
- 5176 [442] E. Wilkins et al. “Rubidium marking of *Anopheles* mosquitoes detectable by field-
5177 capable X-ray spectrometry”. In: *Medical and Veterinary Entomology* 21.2 (2007),
5178 pp. 196–203.
- 5179 [443] N. Windbichler et al. “A synthetic homing endonuclease-based gene drive system in
5180 the human malaria mosquito”. In: *Nature* 473.7346 (2011), pp. 212–215.
- 5181 [444] J. Wong et al. “Sampling considerations for designing *Aedes aegypti* (Diptera: Culi-
5182 cidae) oviposition studies in Iquitos, Peru: substrate preference, diurnal periodic-
5183 ity, and gonotrophic cycle length”. In: *Journal of Medical Entomology* 48.1 (2014),
5184 pp. 45–52.
- 5185 [445] A. C. Wootton. *Chronicles of pharmacy*. Vol. 1. Macmillan and Company, limited,
5186 1910.

- 5187 [446] World Health Organisation. “Comprehensive guidelines for prevention and control
5188 of dengue and dengue haemorrhagic fever”. In: (2002).
- 5189 [447] World Health Organisation. “Dengue: guidelines for diagnosis, treatment, prevention
5190 and control”. In: (2009).
- 5191 [448] World Health Organisation. “Global vector control response 2017-2030”. In: (2017).
- 5192 [449] World Health Organisation. *World Health Organisation: Western Pacific Region*.
5193 [http://www.wpro.who.int/mediacentre/factsheets/fs_09032012_Dengue/](http://www.wpro.who.int/mediacentre/factsheets/fs_09032012_Dengue/en/)
5194 [en/](http://www.wpro.who.int/mediacentre/factsheets/fs_09032012_Dengue/en/). Accessed: 2017-09-18. 2017.
- 5195 [450] World Health Organisation. *World malaria report 2008*. World Health Organisation,
5196 2015.
- 5197 [451] World Health Organisation. “World Malaria Report 2017”. In: *Google Scholar* (2017).
- 5198 [452] World Health Organisation et al. “Global Technical Strategy for Malaria 2016–2030.
5199 2015”. In: *Geneva. 35pp* (2016).
- 5200 [453] World Health Organisation et al. “Preliminary appraisal on the use of age-grouping
5201 methods in anopheline mosquitos”. In: (1960).
- 5202 [454] World Health Organisation et al. “WHO statement on the first meeting of the
5203 International Health Regulations (2005) - Emergency Committee on Zika virus and
5204 observed increase in neurological disorders and neonatal malformations”. In: 37.3
5205 (2016), pp. 332–333.
- 5206 [455] D Wu and M. Lehane. “Pteridine fluorescence for age determination of *Anopheles*
5207 mosquitoes”. In: *Medical and Veterinary Entomology* 13.1 (1999), pp. 48–52.
- 5208 [456] Y. Xue et al. “The Sahelian climate”. In: *Vegetation, Water, Humans and the Cli-*
5209 *mate*. Springer, 2004, pp. 59–77.

- 5210 [457] T. Yamada. “Causal relationships between infant mortality and fertility in developed
5211 and less developed countries”. In: *Southern Economic Journal* (1985), pp. 364–370.
- 5212 [458] H. Yang et al. “Assessing the effects of temperature on the population of *Aedes ae-*
5213 *gypti*, the vector of dengue”. In: *Epidemiology and Infection* 137.08 (2009), pp. 1188–
5214 1202.
- 5215 [459] M Yasuno, P. Rajagopalan, G. Laürique, et al. “Migration patterns of *Culex fatigans*
5216 around Delhi, India.” In: *Tropical Medicine* 17.2 (1975), pp. 91–96.
- 5217 [460] J. H. Yen and A. R. Barr. “New hypothesis of the cause of cytoplasmic incompati-
5218 bility in *Culex pipiens L.*” In: *Nature* 232.5313 (1971), pp. 657–658.
- 5219 [461] J. Zetek. “Behavior of *Anopheles albimanus Wiede* and *tarsimaculata Goeldi*”. In:
5220 *Annals of the Entomological Society of America* 8.3 (1915), pp. 221–271.