Applied Dynamics for Robust Analytics and Adaptive Malaria Control

David L. Smith and the RAMP Team

2024-05-16

Contents

Foreword

A large fraction of my time over the past 20 years has been devoted to learning about malaria and mathematical models of malaria epidemiology, transmission dynamics, mosquito ecology, vector control, and the evolution of drug and insecticide resistance. All the time I was building and analyzing models, I was looking for a way of organizing, operationalizing, and applying the rich body of theory developed over more than a century of malaria research, starting with Ronald Ross [58, 72]. A new framework would integrate the concepts and models that have influenced malaria through the present day [45, 73], and it might even serve as a platform for recasting a theory of malaria epidemiology, transmission dynamics, and control [71].

After all these years, we have developed a framework that tackles the challenges of using evidence to develop malaria policy. A core part of that framework is the use of malaria simulation models. Policy discussions can benefit from using mathematics, but discussions must focus on the policy questions, the concepts, and evidence rather than on the mathematics *per se*. Ideally, mathematical ideas should give rigor to the analysis help structure policy discussions without becoming the focus of attention, except perhaps when the mathematics draws attention to something that needs to be clarified. When I set out, I knew that building a framework would face some challenges, but I believed that – if we could solve those challenges – then we could make the math easier to use so that discussions would focus on the issues that mattered. Ideally, the framework could be taken up and used by teams of local experts working in their own countries to reduce the burden of malaria and plan for its elimination. The same framework could also support the academic activities taken up on the side.

We wanted a nimble framework for model building that was capable of conforming to the problems at hand. It should be *extensible,* with *plug-and-play* modularity (for major dynamical components), with flexibility in the choice of functional forms describing various relationships. The framework should also be structurally flexible, so that a model could incorporate space, time, and population strata. It should be scalable to handle realism and complexity: it should be possible to scale down spatially for fine-grained spatial simulations [10, 25, 44], to scale down temporally for fine-grained temporal simulations [79], or to scaling up to understand or analyze regional processes and the emerging patterns [74]. To

serve the needs of malaria programs, a framework would need built-in support for exogenous forcing by weather and vector control to model malaria as a changing baseline modified by control. To get integrated vector control right, we went all in on mosquito ecology with an individual-based simulation model (called MBITES) that could handle exquisite biological detail [79], which inspired new ideas about simple models based on mosquito search and behaviors. To get malaria transmission right, we would need to be able to deal with heterogeneity of all sorts.

To serve programmatic needs, we needed algorithms that could address the decay of interventions – coverage could decay over time through loss (*e.g.* bed nets) or waning potency – and the net effect of one unit of vector control in relation to transmission intensity and coverage. To make all the pieces fit together, we needed interfaces that could connect up models in a generic way; in the design phase, we worked with two model families for each major dynamical component – one that was dead simple, and one that had a was highly realistic. In some cases, the interface designs called for development of new algorithms: blood feeding, egg laying, environmental heterogeneity, human mobility, and mosquito dispersal. In making a master list to test the framework's extensibility, we found that some odd cases that needed to pass information among components – endectocides and auto-disseminated larvicides – but it was easy enough to accommodate these. Beyond the algorithms, these models needed the support of mathematical theory. We wanted the software to help understand thresholds, so we wrote the routines that would compute thresholds for malaria transmission in heterogeneous systems, when appropriate.

Sometime in the fall of 2022, the last few pieces came together. We published the first versions of MicroMoB and exDE at CRAN. We submitted a paper to PLoS Computational Biology [78]. A second paper forced us to confront a set of software design questions in relation to exogenous forcing **addCite**. Then it was time to write this book.

The software lowers the costs of building models that are up to the task of guiding malaria policy. An advantage of using this framework and accompanying software is that it has built-in solutions to a large set of problems that arise when building models that combine many factors. These sorts of pitfalls are inevitable and annoying, and they are usually discovered the hard way. We found design solutions that would help others building models to avoid these sorts of problems. The framework took longer than expected, in part, because there were more pitfalls than we had anticipated.

This book has been written to introduce the features of the mathematical framework and the software (see Figure 1.1). The book is accompanied by the RAMP-Model-Library, which was set up during the primary design phase. The RAMP-Model-Library is where we made all our design mistakes: was the software truly plug-and-play, and was the framework truly extensible? As the primary design phase came to a close, the library that was once the laboratory became a classroom and a museum. The library is being transformed into a

resource for any developer who wants to add new base models to the library or add functionality. Most of all, it is being set up for the end user, someone in a malaria program or working with a malaria program who wants to use simulation based analytics to analyze policies. This book is structured into a set of lessons that teach concepts. Some of the concepts build on one another, and others take on new challenges. We combine these lessons into some examples where we show some algorithms to build models fit for purpose. When a topic deserves a deeper dive, we have supplemented this book with vignettes or lessons.

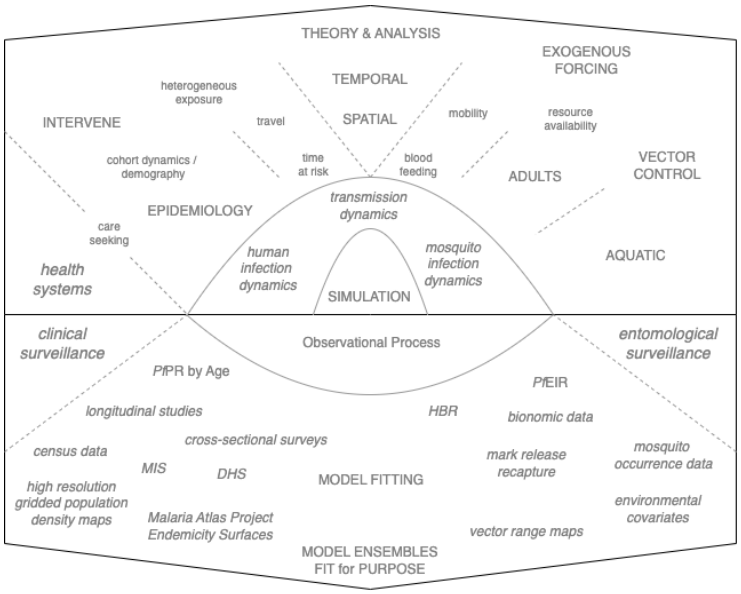


Figure 1: A schematic diagram of the elements in the framework (top half) and the process of model building and model fitting (bottom half)

This book shows, through examples, how to use the software to build malaria models. Even without the software, it fills a gap for students who have taken an introduction to mathematical epidemiology or infectious disease modeling and want to go on in malaria. What are all the special topics that would need to be covered to build models that could be needed in malaria? This book *could* be the basis for such a course, if there were ever enough students. Since there will probably never be enough graduate students at my university who are interested in applied malaria dynamics, the material is being developed for any student anywhere. I will also be writing and recording some short lectures on a subset of topics.

The premise of this book is thus that the reader has started with a solid back-

ground in infectious disease models and malaria. We assume they've seen the Ross-Macdonald model before, that they've taken a class in mathematical epidemiology, and that they know something about how to construct and analyze models. This book emphasizes concepts and teaches through examples. We have left out a lot of the technical and mathemtical details, but we have written some vignettes and lessons to supplement the book. Most of this is found in the documentation for `exDE` or `MicroMoB` or it can be found in the `RAMP-Model-Library`.

While this book should help others build models for malaria policy, but it stops short of applying models to policy. That is covered in another book, ***Robust Analytics for Malaria Policy.***

In doing all this, we are building on an enormous body of work that started with Ronald Ross. While Ross is better known for identifying malaria parasites in a mosquito gut, which proved that malaria is mosquito transmitted, we are more interested in the rest of his career.

After winning the Nobel Prize in 1902, Ross was instrumental in building solid quantitative foundations for malaria transmission and its measurement. Ronald Ross wrote the first models describing malaria transmission. In his writings from 1899 to 1911, it's clear that he was searching for quantitative way of saying something simple – if there are not enough mosquitoes, the malaria transmission can't be sustained. There must be a critical mosquito density, above the cutoff malaria transmission would be sustained, and below it malaria would be eliminated. Ross was looking for a formula that encapsulated his intuition: how were thresholds related to the fact that it took two bites for a mosquito to complete its life cycle? Eventually, Ross wrote down some systems of equations that would describe malaria. The ideas, mathematics, and identification of parameters and processes were extended by other scientists later, most notably Alfred Lotka and George Macdonald.

It seems that the challenge of malaria control was what pushed Ross toward modeling. Ross's first model was a discussion of adult mosquito movement to guide larval source management [47]. The first model describing malaria transmission appeared in a book, *The Prevention of Malaria in Mauritius* [55]. When it came to thinking through control, Ross found it useful to do the math. More than a century later, this book is a look at what his vision has become.

–David L. Smith

Contributors

I've done the primary writing for this book. The framework would not exist without the work of Sean L Wu and a few others. The book borrows from the

work of others, and we have done our best to give credit through citations. It has been a collaborative process (see Contributors). The errors, however, are mostly mine. If you find mistakes or have questions, please drop me a note by email: smitdave@gmail.com.

This is a work in progress, so the list of contributors will change over time.

The software package `MicroMoB` was written by Sean L Wu, Sophie Liebkind, and David L Smith. The software package `exDE` was written by Sean L Wu and David L Smith.

Most of the content so far was written by David L Smith. Contributors from the RAMP Team include:

. . . please consult Dave if you would like a writing role.

Part I

Introduction

Chapter 1

The Prevention of Malaria

Ronald Ross is best known for being the first person to observe malaria parasites in a mosquito [54]. The event, which occurred in a military hospital in Hyderabad, India on August 20^{th}, 1897, is now celebrated as *mosquito day*. In the months that followed, Ross made the case that mosquitoes transmit malaria parasites [60]. Concurrently, the life-cycle of human malaria parasites was described by an Italian team, led by Giovanni Battista Grassi [24, 23]. By 1899, a scientific consensus had formed among malaria experts that malaria was a mosquito-transmitted disease. It would take some time before the "mosquito hypothesis" was accepted more broadly.

The discoveries made by Ross and Battista Grassi closed a loop that Charles Laveran had opened in 1880 when he identified malaria parasites in blood, providing the first clear evidence that human malaria was a parasitic disease [29, 30]. Laveran's observation had spawned almost two decades of research describing malaria infections and disease in humans, all before the mode of transmission was known [22]. The modern view of the parasite life-cycle was not complete until the 1940s, through research describing the parasite's liver stages by Garnham and colleagues [66, 22].

Malaria is a disease that could be treated and measured before anyone knew what caused it or how it was transmitted. Malaria was commonly treated with quinine once the compound had been isolated in 1824. The chemical compound was isolated from the bark of Cinchona tree, which had been used to treat malaria fevers. The spleen rate, a reasonable metric for malaria was described by Dempster in 1948 [12]. All of this occurred before Laveran showed how that malaria was caused by infection with a parasite.

The elements of the story of malaria, including a full account of how the malaria parasite's life cycle was unravelled over time is a topic that worth taking some time to learn [9, 22, 64, 11].

Knowing the mode of transmission, the British military launched vector control efforts to protect British citizens from malaria in its colonial empire. At the time, vector control with bed nets and larval source management, and malaria was treated with quinine. Lest we imagine malaria control in this era was for general humanitarian purposes, Ross emphasizes population segregation as a useful mode of malaria control for the colonizers. Both Ross and Laveran were military doctors supporting colonial governments, a fact that we acknowledge that we can neither ignore nor excuse. We try to learn enough history to avoid repeating the same mistakes. We are sure to make enough mistakes of our own as we blunder forward.

In this first epoch of malaria control, it is Ross who once again draws our attention. After his big discovery, Ross turned his attention to malaria control. In 1910, he published a book, *The Prevention of Malaria.* A second second edition followed in 1911. Ross's work on malaria prevention is of interest because he laid out a vision for a quantitative approach to the prevention of malaria. He clearly articulated how malaria control needed quantitative concepts, including well-developed metrics to measure malaria in populations. Ross himself developed the thick film, a technique that was intended to increase the sensitivity of malaria diagnostics [61, 49]. In the service of understanding malaria for purposes of control, Ross developed the first mathematical models for malaria, in which he addressed basic questions about larval source management [47] and malaria transmission [55, 58, 57].

In the second decade of malaria control, the last phase of Ross's career, he formed an alliance with Hilda Hudson, a mathematics professor from Cambridge University. Between 1915 and 1917, Ross and Hudson developed mathematical theory to support the study of epidemics, which they called *a priori pathometry.* Then, in 1921, Martini developed a model describing the dynamics of infections that immunized their hosts. In 1927, Kermack and McKendrick published the first of a five-part series on the mathematical theory of epidemics that described the dynamics of acute immunizing infections, which overshadowed the work by Ross and Hudson. Nowadays, the field is called *mathematical epidemiology.*

The parasite's life-cycle has been an organizing principle for the study of malaria. The discoveries by Laveran, Ross, Battista Grassi, Garnham and others set the stage for hundreds of thousands of studies over 144 years (at the time of this writing) that have developed a rich and impressive body of knowledge. Over that time, we have learned an enormous amount about malaria and its control. The first comprehensive review of malaria, *Malariology. A Comprehensive Survey of all Aspects of this Group of Diseases from a Global Standpoint,* was published in 1949 in a two volume set edited by Mark F. Boyd with 70 chapters and 1643 pages from 65 contributors [7]. Almost 40 year later, 1988, the second comprehensive review of malaria was published in another two volume set, *Malaria: Principles and Practice of Malariology,* edited by Walter H Wernsdorfer and Sir Ian McGregor; this set had 57 chapters and 1818 pages from 68 contributors,

[77]. Another 40 years has passed, and it is almost time for an update.

This book is a primer on mathematical models to support robust analytics for malaria policy (RAMP). It covers malaria epidemiology, mosquito ecology, human demography, and malaria transmission dynamics and control with a focus on the mathematical theory and models that have been developed over the past century. While this book will cover many of the topics in those two-volume sets, it is focused more narrowly on introducing the mathematical study of malaria that started with Ross. The field made some great leaps forward in the 1950s, when Macdonald wrote a series of papers that reviewed decades of field studies and updated Ross's basic models. Macdonald's paper on the sporozoite transformed the quantitative study of malaria transmission by mosquitoes, drawing attention to the importance of mosquito survival. In his next paper, Macdonald presented a formula for the basic reproductive number for malaria, an idea that almost certainly traces back to Alfred Lotka's work in human demography. Macdonald would have known Lotka from his extensive work on Ross's models of malaria. The simple mathematical models, now called *the Ross-Macdonald model*, thus defined basic theory for malaria transmission dynamics and control [72]. This book starts with the Ross-Macdonald model, and it covers a rich and wonderful set of models that has grown out of it. Around 1970, the mathematical study of malaria entered a period of elaboration through the work of Klaus Dietz, Norman TJ Bailey, Joan Aron, Barbara Hellriegel, and others. Ross was a pioneer in a field that has included important contributions by hundreds of scientists and analysts.

Nowadays, we can describe malaria as a managed, complex adaptive system involving non-linear interactions among mosquitoes, parasites, humans, and the managers. Malaria systems are heterogeneous and locally peculiar in some way: doing the same thing in different places can result in different outcomes. To understand why outcomes of malaria control vary, we need good theory for malaria, including malaria transmission intensity and seasonality, mosquito ecology and behaviors, mosquito and parasite genetics, health systems, and human behaviors. To give advice, we need malaria intelligence – information about these critical factors, and since the systems are peculiar, we will need local information to give tailored advice. We must grapple with questions about malaria that remain poorly understood, and we most offer advice even if there are enormous gaps in data and knowledge. We want to give policy advice that is robust to uncertainty. We will need well organized systems to store and analyze data and to develop intelligence that can help us learn and adapt. In all of these activities, mathematical models of malaria transmission dynamics and control play an important role.

This book takes a deep dive into dynamical systems models for malaria epidemiology, transmission dynamics, and control. It is designed to serve as an introduction and resource for malaria analysts who seek to use evidence and mathematical models to develop advice about malaria policies. We introduce the material using a modular framework for building malaria models designed

to suppport RAMP [78]. Most of the examples use one of two RAMP-branded R packages: one is designed to build and solve systems of ordinary or delay differential equations ramp.xde; and the other supports discrete-time systems ramp.dts. These are further supported by other ramp software packages, including model libraries, and other simulation software that takes a deep dive into mosquito ecology, mosquito dispersal, malaria epidmiology (in the narrow sense). We also hope that the material can be used in scientific research.

To get all this started, we will walk through concepts as they developed over time, which brings us back to Ronald Ross, the study of the prevention of malaria, and the study of epidemics.[1]

1.1 The Quantitative Approach

The goal of malaria control is to prevent malaria. In the long-term, the expectation is that malaria will be eliminated from an ever growing list of places, and that it will eventually be eradicated. All of this revolves around the study of malaria transmission by mosquito populations through blood feeding.

Since malaria is a mosquito-transmitted diseases, malaria control programs need some basic theory to understand how malaria (*i.e.* any disease that is caused by infection with the parasite) is related to exposure risk, malaria transmission intensity, and the prevalence of infection. In particular, programs should should have a basic understanding that relates the factors driving malaria in a place, the metrics that we use to measure malaria, mosquito population densities and bionomics, and the actions taken by malaria control programs.

A vision for doing all this was articulated by Ross in the first decade of studying malaria transmission. Ross was not the only one to grapple with the complex, multi-factorial nature of malaria transmission and vector control, but what set him apart was his attempt to use of mathematics to illustrate his ideas about malaria. We will use Ross's writings throughout this chapter to introduce the nature of the challenges of managing malaria control. With the benefit of decades of research and cheap high-speed computing, we are in a better position to fulfill that vision than Ross ever could have.

At the end of this chapter, we make some attempt to put Ross into a broader historical narrative.

---

[1] The narrative in this book is often historical, which serves two purposes. First, it provides an excuse to cite and discuss old papers that should not get overlooked. Second, a historical narrative allows us to introduce malaria complexity in an ordered way, starting from principles that are simple and abstract, and progressing towards ideas that are non-linear, messy, or subtle. We will often use history as a way of structuring discussions, even as our attention turns increasingly to the mathematics we need to deal with the biological complexity of malaria in populations.

In 1899, Ross announced a new phase of his career in a pair of essays on *exterminating* or *extirpating* malaria through mosquito control [50, 51]. Ross's initial focus was on larval source management:

> *. . . in order to eliminate malaria wholly or partly from a given locality, it is necessary only to exterminate the various species of insect which carry the infection. This will certainly remove the malaria to a large extent and will almost certainly remove it altogether. It remains only to consider whether such a measure is practical [50].*

Over the decade, Ross began to use quantitative concepts to understand some very basic questions. Ross was interested in using mathematics to help him reason through problems. Among Ross's early writings, we find a question posed in 1902, in *Mosquito Brigades and How to Organise Them* [53]:

> It may now be asked, what percentage of diminution in mosquito-borne diseases may be expected to follow a given percentage of reduction in the number of mosquitoes? I regret that I cannot as yet give any actual statistics on the point, but we may perhaps attempt an estimate on *a priori* grounds. We ask, are we to expect a decrease in the same ratio as the decrease in the number of mosquitoes; or in a dupblicate ratio? The disease will probably diminish in a duplicate ratio? —pp. 56 in [53]

Ross was thus trying to develop quantitative intuition to establish reasonable expectations about malaria transmission and the responses to mosquito control. His guess was that the responses should be non-linear because two bites are required for transmission – one to infect a mosquito and another to transmit it back to humans. In paragraph that follows, Ross explains what he meant by a *duplicate ratio.*

> Now, if we reduce the number of mosquitoes in the locality by one-half, the mosquito bites also will be reduced by one half; and, consequently, only half as many people will now become infected as was formerly the case. But, since the mosquitoes themselves are infected by biting previously infected persons, the percentage of infected mosquitoes . . . will also be reduced in its turn, because the insects will now find fewer infected persons to bite. Hence, ultimately, the number of infected persons in the locality will be reduced by much more than one-half. In fact, we may perhaps assume that the number of infected persons will be reduced to one-quarter, that is, in the duplicate ratio of the percentage of reduction of the mosquitoes. —pp. 56 in [53]

While Ross was right that there should be a non-linearity, his quantitative logic failed when he tried to use it without going the rigorous process of developing and analyzing a mathematical model. Using mathematical models, we would now argue that, in fact, the expected reduction varies from place to place. The response would depend on the intensity of transmission before doing any control, which scales linearly with the density of mosquitoes. When transmission is

very intense, halving mosquitoes would scarcely change the fraction of humans infected. In some places, halving mosquitoes might be enough to end local malaria transmission. If malaria transmission were sustained through malaria importation, then Ross's answer would not be too far off.

Despite getting the logic wrong – or at the very least, underestimating the complexity of the question – Ross was asking relevant questions. In this case, we was asking about *scaling relationships* in malaria transmission and their causes. The search for an *a priori* approach would eventually lead to development of his first mathematical model for malaria transmission in 1908, in *Report on the Prevention of Malaria in Mauritius* [55]. A clear basis for giving an answer would finally come when Lotka analyzed Ross's models in 1923 [32, 33, 34, 35], and in Macdonald's analysis of the sporozoite rate [38] and his formula for the basic reproductive number for malaria [36]. The study of these scaling relationships is still evolving.

Ross also discused response timelines. He was aware of differences in duration of infection for yellow fever, lypmhatic filariasis, and malaria. Compared to yellow fever, the responses to vector control should be slow because,

> ...the parasites remain alive for years after the first moment of introduction by the mosquito.... We must not, therefore, expect to see malaria vanish, as if by magic, immediately after our campaign against mosquitoes. – pp. 53 of [53]

In all this, Ross was asking thus questions about what would determine the outcomes of malaria control.

What other factors could affect what was observed? One of the arguments against the "mosquito hypothesis" had been that malaria was sometimes found in places where there was no evidence for malaria transmission. In *The Prevention of Malaria,* Ross spends a great deal of time discussing the importance of imported malaria. In the chapters that follow, we will follow this idea through to its logical conclusion.

1.2 Larval Source Management

The interest in mosquito populations originated from an academic discussion, in 1904, that critiqued a larval source management program in Mian Mar. One fact that clearly irked Ross was the lack of any good method for measuring mosquito populations accurately. The problem at hand:

> Suppose that we have to deal with a country of indefinite extent, every point of which is equally favorable to the propagation of gnats (or of any other animal); and suppose that every point of it is equally attractive to them as regards food supply; and that there is nothing, such for instance as steady winds or local enemies, which tends to

> drive them into certain parts of the country.- Then the density of the gnat population will be uniform all over the country. Of course, such a state of things does not actually exist in nature; but we shall nevertheless find it useful to' consider it as if it does exist, and shall afterwards easily determine the variations from this ideal condition due to definite causes. Let us next select a circumscribed area within this country, and suppose that operations against the insects are undertaken inside it, but not outside it. The question before us is the following: How far will these operations affect the mosquito density within the area and immediately around it?

Ross's question was:

> How large must that radius be in order to render the center entirely mosquito-free?

1.3 Malaria Population Dynamics

He recognized that prevalence could change, and that what set its value was a balance between two processes. The quantitative logic looks like this:

$$\left[\begin{array}{lcl} \text{Fraction Infected Today} & = & \text{Fraction Infected Yesterday} - \\ & & \text{Cleared Infections} + \\ & & \text{Fraction Uninfected that Got Infected} \end{array}\right]$$

This description ignores malaria importation, infected people who died, infected people who emigrated, and infected people who immigrated. The model is neither comprehensive nor perfect. It was a starting point.

For short-lived mosquitoes, we assume they are much more likely to die than to clear infections. So for mosquitoes, the process is slightly different:

$$\left[\begin{array}{lcl} \text{Fraction Infected Today} & = & \text{Fraction Infected that Survived} - \\ & & \text{Fraction Uninfected that Got Infected} \end{array}\right]$$

What's important here is that infectious mosquitoes are infecting people when they bite, and infectious people are infecting mosquitoes when the mosquito blood feeds.

In planning and evaluating malaria control, it was important to develop some kind of expectations about where and when to do malaria control, what types of malaria control are likely to work best, and so on.

Despite all the focus on malaria, transmission Ross's first mathematical model of malaria was about mosquito populations.

In 1911, Ross formulated a model (his second) as a system of two ordinary differential equations. The model appeared in the 2^{nd} edition of *The Prevention of Malaria* [58], and it also appeared in *Nature* [57] in this form:

$$\begin{aligned} dz/dt &= k'z'(p-z)+qz \\ dz'/dt &= kz(p'-z')+q'z' \end{aligned}$$

1.4 Measuring Transmission

Through the first half of the 1900s, Ross promoted malaria control and malaria research. He was a key player in public debates about malaria control and he launched a center to study malaria transmission that is still active today [6, 72]. He engaged in public debates about malaria control hosted by the *British Medical Journal*, and he published a book, *Mosquito Brigades and How to Organize Them* [53]. The end of this phase of his career was marked by publication of the 2^{nd} edition of *The Prevention of Malaria* in 1911 [58].

This phase of Ross's career is highly relevant to our study of **Applied Malaria Dynamics** because it was the beginning of a rigorous approach to malaria control including development of the first three mathematical models describing malaria transmission dynamics or control [47, 55, 58, 57]. Many of the early attempts to control malaria were implemented by the British Military. In the first few years, there were some successes, but there were also some failures [59]. Heterogeneity in the responses to malaria control efforts seem to have turned Ross to mathematics. In 1904, he presented a paper at the International Congress of Arts and Science in St. Louis, Missouri that applied diffusion models to larval source management (published also in *Science*[47]). Ross's transmission models appear in publications that emerged from consulting with national malaria programs in Mauritius and Greece [55, 52]. While he was writing about the nuts and bolts of control, he was also grappling with mathematical formulas that could help him understand malaria control quantitatively. How would reducing mosquito densities change the prevalence of malaria? Was there a critical population density of mosquitoes required to sustain transmission? The result was the first mathematical model to describe malaria transmission dynamics [55]. A short time later, Ross would reformulate the model in the 2^{nd} edition of his book, *The Prevention of Malaria* [58]. He also published it in Nature [57]. For Ross, the mathematical models were a logical next step towards trying to understand malaria control in rigorous quantitative terms.

1.5 The Study of Epidemics

By 1915, when Ross turned his attention to the broader study of epidemics, his intent was to establish a new discipline, which he called *a priori pathometry*. He

set out to establish the mathematical foundations for the study of epidemics [56, 48]. In his 1916 paper, he wrote

> *It is somewhat surprising that so little mathematical work should have been done on the subject of epidemics, and, indeed, on the distribution of diseases in general.*

The case for using mathematics had grown stronger as a result of a scientific revolution caused by invention of the light microscope that established a germ theory of disease, which included Laveran's discovery in 1880. Even without knowing that bacteria, viruses, and parasites were transmitted among hosts causing disease, *contagion* was an impossible concept to ignore. Some mathematical work had already been done on infectious diseases, including a few models. It's not clear whether Ross was aware of Daniel Bernouli's mathematical model of small pox [5, 3], or PD En'ko's discrete time models for measles [13]. Ross may have known about those models, but he certainly didn't cite them. Regardless, the mathematical study of infectious diseases in populations was, at that point in time, underdeveloped. Ross wrote:

> *. . . the principles of epidemiology on which preventive measures largely depend, such as the rate of infection, the frequency of outbreaks, and the loss of immunity, can scarceley ever be resolved by any other methods than those of analysis.*

Ross proposes a rudimentary three-tiered classification system for epidemics based on the patterns of fluctuating incidence with exemplars: 1) leprosy and tuberculosis; 2) measles and malaria; 3) plague and cholera.

> *To what are these differences due? Why indeed should epidemics occur at all, and why should not all infectious diseases belong to the first group and not always remain at an almost flat rate?*

In 1917, he teamed up with Hilda Hudson, a Cambridge mathematician, to finish the first major contribution to mathematical epidemiology, or what he called *a priori pathometry* [62, 63]. Over 18 years (1899-1917), Ross's ideas laid a solid foundations for the modern study of malaria transmission and theory of malaria control [72].

1.6 Mathematical Epidemiology

Despite the heroic efforts of Ross and Hudson on *a priori pathometry*, they were missing a key element in the modern mathematical study of epidemics. In 1920, a new mathematical model was published by Martini, which described epidemics in which recovery from infection was followed by lifelong immunity. Martini's model would be explored in great detail in a series of papers by Kermack, and Mckendrick. In modern language, Ross's models would be studied as SIS compartmental models, while Martini's equations would be called SIR

compartmental models. The two systems of equations have strikingly different features, and so do the kinds of epidemics they mimic.

In Ross's model, a person would become susceptible to infection after recovering, so a person could be infected many times, and prevalence is a naturally meaningful statistic. In fact, Ross's equations can be solved exactly: after Ross published the equations in Nature in 1911 [57], Alfred Lotka published a closed-form solution in 1912 [31].

There is no closed form solution to the SIR model, and prevalence is an ephemerally changing quantity. The natural summary statistic for SIR epidemic is called S_∞, the portion remaining uninfected asymptotically. In simple epidemics – an epidemic in a single population with no replenishment of susceptible hosts – the dynamics differ depending on how many cases that first case would tend to generate, called R_0. If $R_0 < 1$, then there is a closed form solution: $S_\infty = S_0 - I_0(1 - R_0)^{-1}$. If that first case would generate more than one other case, and the number of cases would initially rise to a peak, when each new case would generate exactly one case. Thereafter, the number of cases would decline. The time course of an SIR epidemic does not have a closed form solution.

The differences between SIS and SIR models would provide a mathematical basis for answering Ross's question and point to an important role for the concept of population immunity. In SIS models, there is no immunity, so prevalence tends to approach an endemic equilibrium with very little tendency to cycle. When susceptible populations the SIR models are replenished by birth or migration, there is a natural tendency to cycle.

The SIS model is a useful starting point, but Ross was aware of evidence that immunity to malaria would develop in humans. The role of immunity is more complex, but a critical feature is that infection with one parasite would not prevent reinfection with another. In places where exposure rates were high, reinfection – also called *superinfection* – was quite common. The first model of malaria that was *not* an SIS model – published by George Macdonald but relying on mathematics by P. Armitage – considered a role for *superinfection,* or reinfection of susceptible individuals. In models with superinfection, an interesting statistic is the multiplicity of infection (MoI). A useful discussion of this superinfection model and its mathematical flaws was written by Paul EM Fine [21]. What is, perhaps, more important is that superinfection in malaria facilitates sexual recombination for the parasites. Sexual recombination facilitates development of parasite diversity, which could partly explain population immunity in malaria.

Sticking with Ross's core challenge of how to study epidemics mathematically to understand the tendency to cycle, we must acknowledge the complex natural history of population immunity. In malaria, a model for the mathematical talents of Klaus Dietz combined with the epidemiological skills of Louis Molineaux to produce the first model of population immunity to malaria. That model described infection dynamics in a population with partial immunity. In malaria

epidemiology, a core challenge is the problem of malaria immunity and its relationship to parasite genetic diversity, disease, and infectiousness. In the mathematical study of malaria epidemiology, the formulation of adequate models is among the most pressing and most difficult to address. Measles and a few other pathogens are well-described by the SIR model. For these acute immunizing infections, infection is followed by a life-long immunity. The dengue viruses has four (or maybe five) functionally distinct virues with complex patterns of cross immunity. For other pathogens, population immunity is undermined when immune-escape variantes evolve in the pathogen populations. For the influenza viruses and the coronavirus descendents of the viruses that sparked the global COVID-19 epidemic, escape variants arise at a rate that is high enough to sustain yearly outbreaks.

Adding to the complexity, immunity is not the only factor affecting a tendency to cycle. For arthropod-transmitted pathogens, fluctuations in arthropod population densities and behaviors driven by weather affect trasmission rates. For directly transmitted pathogens, humidity and survival in the environment and the tendency for populations to congregate amplifies transmission. In malaria, fluctuating mosquito one factor that must *a priori* force transmission, and since population immunity probably has a weak effect, seasonal malaria is largely driven by fluctuating mosquito populations. In SIR models, seasonal fluctuations are sustained by an interaction between environmental factors and the depletion of susceptible individuals.

1.7 Approaching Complexity

Malaria is complex and heterogeneous, which makes it difficult to study and manage. A core challenge in both science and policy is the availability of information. Mathematical models can help us understand and analyze all that complexity and make informed decisions, despite the data gaps.

While this book is about applied dynamics, the approach is eclectic. It relies heavily on conventional statistical analysis. An axiom we espouse is that two inferential approaches are better than one. This point of view was prominent in the early study of epidemics, tracing back to both Ross and Browne [20]. After Ross spent a decade studying malaria control, he realized the difficulty of understanding malaria epidemics, so he argued for an integrated approach. About the study of epidemics, he wrote

> *The whole subject is capable of study by two distinct methods which are used in other branches of science, which are complementary of each other, and which should converge towards the same results – the* ***a posteriori*** *and the* ***a priori*** *methods. In the former we commence with observed statistics, endeavor to fit analytical laws to them, and so work backwards to the underlying cause (as done in much statistical*

> *work of the day); and in the latter, we assume a knowledge of the causes, construct our differential equations on that supposition, follow up the logical consequences, and finally test the calculated results by comparing them with the observed statistics.*

More than a century later, it is much easier to use computational tools to fulfill this vision.

While this lays out a useful agenda for *understanding malaria,* another axiom we espouse is that the process of giving advice about a policy must take on a broader set of challenges. Policy and science are different in many ways, so there are good reasons why we might want to use different models and methods for basic research and policy analytics. In policy, decisions *must* be made in a timely way, and they *should* use all available evidence, even if it's weak. Basic research is epistemologically conservative, by design. Studies that are published in peer review must be repeatable, and in anticipation of criticism, the scientists aim to do things so well that they are unassailable. This usually has the effect of narrowing and controlling the conditions under which the study was conducted. When we operate in policy settings, we *should* design our studies in a different way, and we will thus need to deal with the uncertainty differently. The studies that inform policies will need to take an approach that is broader – the models should be realistic enough to address the question of interest. Such studies will often need to make compromises and decisions that are poorly informed by the evidence, which raises the question of how much a policy maker could trust it. One strategy for making policies trustworthy is to repeat the supporting analysis using every reasonable approach, so that we can be reasonably sure our policy recommendations would not change. The idea of fully propagating uncertainty is the essential feature of *robust analytics.* If we make the effort, we can identify key sources of uncertainty, identify priority data needs, and collect new data that could help resolve some of the most important sources of uncertainty. Building models to do this is challenging for practical reasons, and it requires drawing heavily on basic research. In giving advice, we must give different weights to the uncertainty than we would in research.

In basic research, we develop mechanistic models to understand malaria as a biological process. In malaria epidemiology, the states and parameters describe infection, immunity, infectiousness, disease, and drug taking in response to exposure. Scientists focus on basic biological mechanisms in order to understand differences in malaria across spectrum of transmission. Immunity and drug-taking are important factors to consider, but it may be that differences in epidemiology and disease across settings arise from differences in the local parasite populations. The models are a way of summarizing knowledge in a quantitative form – something like a complex hypothesis. A test of a model's adequacy is whether it can describe malaria accurately after accounting for differences in drug taking patterns and pattern of exposure.

We study mosquito ecology and blood feeding to understand malaria transmission and develop theory for malaria control. Transmission models couple

parasite infection dynamics in humans and mosquitoes through blood feeding. Mosquito populations are shaped by the aquatic habitats for immature mosquito populations. These habitats are standing water bodies, and they are shaped by topography, hydrology, land use, and the water chemistry, which is affected by surrounding rocks, soils, vegetation and pollution. These habitats are filled (exogenously forced) by rainfall and after some eggs are laid, the mosquito dynamics are affected by crowding, predation, and other endogenous dynamics. Larval development and parasite development rates are modified by temperature. Adult mosquito activity rates are affected by temperature, relative humidity, and vector control. Indoor residual spraying (IRS) kills mosquitoes when they rest on a sprayed surface, usually after blood feeding or during the process of searching for a host. Insecticide treated nets (ITNs) protect humans from biting and kill some mosquitoes. By reducing the availability of potential blood hosts, nets can slow blood feeding in some contexts. Larval source management (LSM) reduces immature population densities.

By studying mosquito ecology and malaria transmission dynamics, we can start to understand malaria as a changing baseline that has been modified by malaria control. This is the problem confronted daily in malaria programs, but dealing with the evidence requires having the tools available to synthesize data describing different parts of malaria. The models help translate evidence into information that can be used to make decisions, to make strategic plans, and to mark progress against national plans. The models encapsulate information about transmission in context, so it is possible to study how malaria persists in a place over time, and how various factors have modified (or could modify) mosquito population dynamics and blood feeding and thereby suppress transmission. Transmission models help us to set intervention coverage targets based on an understanding of malaria connectivity to surrounding regions and local thresholds.

In policy, we use these models with the expectation that – if we fit the models by adjusting parameters that affect how malaria works in some particular place – they *should* help us understand transmission in some particular context and make good decisions about what to do.

Frustratingly, the heterogeneity and the complexity conspire against us. We would like to be sure about how malaria works across settings before we start using the models to stratify populations, tailor interventions to context, or targeting the interventions. Instead, we must admit that we don't know everything we'd like to, and we probably never will. We must proceed with policy without having satisfactory answers to some basic questions. In policy, we will use the models to evaluate the consequences of having missing information, but we will also use the models to help us prioritize missing data so we can fill in the gaps. What missing data would reduce our uncertainty about what to do about malaria? How do we fill the critical knowledge gaps.

To understand malaria or to give policy advice, we must start simple and then add complexity, layer on layer. To deal with missing information, we start with generic models, and then add details to address concerns about some of

the details that we hope to identify by studying the systems as we intervene. This approach – starting simple and then layering on complexity – makes it possible to learn as we go. A question is when it stops making sense to add realism to a model. A model that it too simple and abstract might help us understand the basic dynamics and give generic advice, but we would question the model's adequacy if it could not reproduce the patterns we cared about in some particular place at some particular time. As a rule of thumb, a model should be just complex enough to *describe* the patterns we care about and *weigh* the relevant options to give advice. Practically speaking, it's hard to know you've gone far enough unless, at some point, it's clear that you've gone a bit too far.

Over the past few years, we developed a new framework for building models that would make it possible to start simple and then build models of malaria transmission at any level of complexity. We wanted to be able to build in realism by adding complexity one feature at a time. Through this process we can create nested, hierarchical models in branching chains. At the ends of the chains, we might find highly realistic models that are, perhaps, overfit. (The cautions against overfitting play out differently in policy given the urgency of acting in a timely way, but it is also possible to go out and collect new data.) We call the framework's ability to do this **scalability** and the resulting swarms have **scalable complexity.** The iterative attempt to make plans, weigh evidence, quantify uncertainty, gather new data to reduce uncertainty, and then restart the annual cycle, is called **adaptive malaria control.**

To make this possible, we needed a way of building models that would keep the focus on the policy questions and on a dialogue between malaria managers and the analytical support team. We thus sought to design modular software with *plug-and-play* functionality and a high degree of structural flexibility. We needed the framework to be extensible. After making a lot of mistakes, the primary design phase is over, and the algorithms have been published in two software packages. We are currently extending the library of *base models*, which includes some simple or classical models that are instructive or of historical interest. We are also fine-tuning the design requirements for models as we develop protocols that streamline fitting models to data. The software avoids the mistakes we made over the past few years, reuses models, and streamlines the model building process. We hope this software has dramatically lowered the costs of building and analyzing these complex, realistic models.

In malaria epidemiology (narrowly defined as a study of infection and disease in humans), the relationship between exposure, infection, immunity, disease, and infectiousness changes in populations as they age, and it is affected by drug taking. This picture grows more complex as we consider intervening with vaccines or monoclonal antibodies, or as we look at interactions with anemia, nutritional status, and human genetics. Our models need to interface with data from clinical settings and research, so they will need to consider diagnostics, parasite counts, detection, and transmission. Combining these factors can give rise to an overwhelming amount of complexity. Later, we will introduce new

models and show how it possible to simplify all this complexity and make sense of malaria.

We are interested in using these models to guide policy, which requires both solid computation and good communication. In this book, we lay a foundation for understanding the complexity by studying some simple compartmental models. We will review classical queuing models for superinfection and the multiplicity of infection (MoI); new models for the age of infection (AoI) or stage of infection (SoI); immunity; parasite densities, fever, disease, and detection; gametocytes and transmission, and drug taking. To end up with models that can handle all the complexity, we build probabilistic models that combine these factors. In doing so, we find that we can do some powerful analysis, and we can map the states in these models onto outcomes that matter for research and policy: test positivity, parasite counts, infectiousness, and disease. With patience, we can combine these factors and develop a framework for understanding malaria in populations that match the features of individual-based simulation models. We end up with a sensible understanding malaria epidemiology as ontogeny – development of immunity as a part of an organisms history. We back this view with some very usable models that capture the changing character of malaria in cohorts of humans as they age.

We are interested in understanding malaria control in context, which requires delving into mosquito ecology and behavior. In this book, we start with a simple model for mosquito ecology and parasite infection dynamics in mosquitoes. We add aquatic population dynamics, mosquito population regulation, and exogenous forcing by weather. Later, we worry about adult mosquito behavioral states such as mating, sugar feeding, and egg laying. We introduce the concept of resource availability, and we develop an understanding of mosquito search and movement in response to resource availability. We take some deep dives to understand how mosquito spatial dynamics work at a fine spatial grain, and then we scale up to understand mosquito populations on landscapes.

At first, we describe mosquito blood feeding and transmission with a few simple parameters. Later, we develop a new model for mosquito blood feeding in a dynamically changing host population with parameters that allow host strata to be more or less available. We also modify our understanding of heterogeneous exposure to biting. We develop a methods for modeling environmental heterogeneity, heterogeneous exposure by age, and a generalized way of handling failty– other sources of heterogeneous biting – through stratification.

We must take a detour to understand how to handle the effects of temperature on the parasite's extrinsic incubation period (EIP). We need a way of dealing with mosquito survival and dispersal through the EIP. This problem has been effectively solved.

To round out this picture, we need a way of dealing with other aspects of human ecology that affect malaria transmission dynamics, including human mobility, human demography, bed net usage, adherance to drugs, and care

seeking. Differences among humans call for a synthesis of studies that have identified traits that affect malaria, stratification, and simulation to identify useful ways of propagating the heterogeneity through analyses.

To go along with a theory of transmission, we need a theory of control. We compute effect sizes and evaluate area effects. We develop a generalized concept of effect modification that considers the total effect of a single unit of control. We modify basic processes by including the effects of vector control and mass medical interventions (*e.g.* seasonal malaria chemoprotection, mass drug administration, vaccines, and monoclonal antibodies). Relying on behavioral state models and the concept of resource availability, we develop a models for integrated vector control.

This is a book about how to do the math that is required for robust analytics for malaria programs. The goal is to use all the data available, but especially the data generated by malaria programs, to paint a clear picture of malaria transmission as a changing baseline that has been modified by control.

1.8 Historical Notes

We note that contemporary scholars of Ross, including Brownlee and Lotka, were also applying mathematics in human demography and health. Ross's mathematical models of malaria transmission were also not the first models for the transmission of an infectious disease. First was Bernoulli's model of smallpox, and P'Enko had published a model for flu a few years before. Ross was

Our philosophy has been to design a framework for model building that can be used by programs. In this context, *model building* means designing ensembles of models. The next step involves applying a set of tools to computational tasks that are beyond what our brains could do. To accomplish our goals, we need more than the mathematical framework. We need to be able to implement and compute models. This requires new software.

The software is structured around three major dynamical components and two interfaces. The dynamical components are: 1) the humans and malaria epidemiology, including the effects of treating malaria with drugs; 2) adult mosquito ecology, behavior and infection dynamics; and 3) aquatic mosquito ecology. Malaria transmission by mosquito populations, including the 2^{nd} and 3^{rd} dynamical component are set up to consider the effects of weather and vector control. The first interface links humans, adult mosquitoes, and parasites to describe parasite transmission through mosquito blood feeding and human exposure to infective mosquito populations. The second interface links adult and mosquito populations through egg laying and emergence. Within each component and interface, there are multiple sub-domains, and there are built in design features to deal with heterogeneity and other features for malaria

control. After a 140 years of studying malaria, there's a lot of detail that could be important in some way.

The software we have developed is meant to lower the costs of building and using models. We want programs to be focused on the decisions, the data, the concepts, and the analysis. As a metaphor, some students learn a numerical method for approximating $\sqrt{2}$ in school, but after learning it once, they stop worrying about *how* it is computed and they punch buttons into a calculator. Knowing how to compute something is sometimes useful, but worrying about how to compute it each time would interrupt the process that called for computing it. Instead, we punch the formula into a scientific calculator or any software that does computation confident that the machine knows how to do it. In applying models, the same kind of logic applies. People need to understand the concepts, but like a calculator, the tools should hide the technical details that don't add to a discussion. The software we have developed is a reliable interface for calculations designed to support policy.

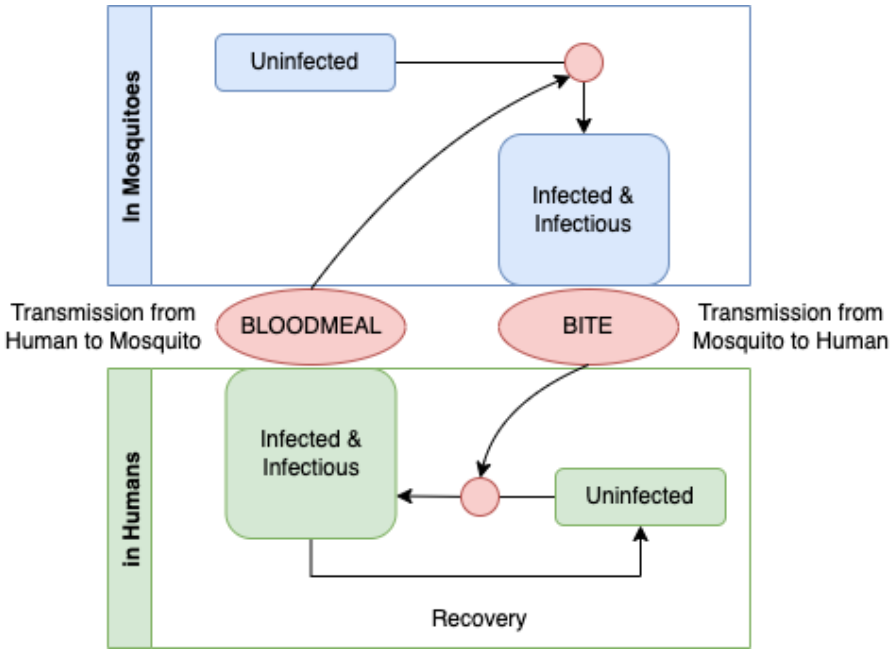
To learn how to use that software, we need to get through a lot of material. The background material in the following presentation is fairly sparse. We are trying to introduce just *enough* mathematics to teach users the critical concepts so they know what the software can do. We assume that the work will be done by teams that include a few people who understand the mathematics, who can guide others through the process.

All of this sounds very complex, but we must start with something simple and layer on complexity in an ordered way. The first model we present is a Ross-Macdonald model.

Chapter 2

Basic Malaria Models

Ross was interested in understanding malaria so he could make plans about malaria control. Given his interests in measuring malaria, he immediately noticed the importance of the fact that malaria prevalence differed from place to place. He was interested in understanding what caused differences in malaria prevalence. Ross developed mathematical models of malaria as a way of understanding malaria prevalence in humans in a population. His first models focused on *local* transmission, ignoring exposure elsewhere (*e.g.* through travel). How does malaria persist in a population? It persists through the transmission of parasites among hosts during blood feeding. Malaria infections in the two populations are connected through blood feeding, so malaria transmission works as an integrated system. As a diagram, the process looks something like this:



To develop a mathematical model, we must make some assumptions about the process.

We imagine a population comprised of individual mosquitoes and individual humans, and we are interested in computing the fraction of humans and mosquitoes that is infected. The number of infected mosquitoes and humans is constantly changing. Since infections don't last forever, the prevalence of infection in mosquitoes and mosquitoes was a balance between infections acquired through parasite transmission and the natural loss of infection and host mortality.

Mosquitoes lived short lives, so the main issue was survival. This formulation of the model thus ignores the loss of infection in mosquitoes and mosquito migration. Critically, this equation does not explicitly say anything about a change in the number of uninfected mosquitoes through *emergence* from aquatic habitats. This model also ignores the lag between the point in time when a mosquito gets infected and when it becomes infectious. The basic mathematical model we will describe in the next section effectively assumes that the mosquito population does not change, so every mosquito that dies is replaced with an uninfected mosquito.

$$\text{MOSQUITOES} \left[\begin{array}{lcl} \text{Infected Tomorrow} & = & \text{Infected Today} \\ & - & \text{Infected: Died} \\ & + & \text{Uninfected: Got Infected} \end{array} \right]$$

Humans live long lives, so the main issue was clearing infections. This formulation

of the model thus ignores human demography, including births, deaths, and migration. The model also has an abstract concept of place, that we can call *here.* The model imagines one mosquito population and one human population in one place. This model ignores malaria acquired during travel away from here.

$$\text{HUMANS} \left[\begin{array}{lcl} \text{Infected Tomorrow} & = & \text{Infected Today} \\ & - & \text{Infected: Cleared Infection} \\ & + & \text{Uninfected: Got Infected} \end{array} \right]$$

We can translate this basic description of a process into a mathematical model. There are at least four ways of doing this. The model can be formulated in either discrete or continuous time, and we can describe a deterministic process or we can develop a stochastic process. Regardless of the type of model we choose to formulate, the first step is to start using mathematical notation.

To describe the factors that affect malaria transmission and that determine the prevalence of malaria in human populations over time, we will need to use mathematical models. In the following, we will write down four distinct mathematical equations that follow these basic rules:

1. As a deterministic system in discrete time;
2. As a deterministic system in continuous time;
3. As a stochastic system in discrete time;
4. A simulation model in continuous time.

In all of these models, we will adhere rigidly to Ross's simplifying assumptions. At this point, we're not trying to understand how the malaria models work. Later, we'll worry about modifying the assumptions to ensure that we have identified and evaluated all the factors that affect malaria in real populations.

2.1 Discrete Time, Deterministic

We can translate Ross's basic description of a process into a model, but first, we need to start using mathematical symbols and notation to represent the process. The model itself is formulated as a set of coupled difference equations, but we will write down the equations last. To get there, we must learn about the parts of a model: variables, initial conditions, and parameters.

In the following, we also want to write R code to implement and solve the model. We will be adopting some conventions that end up being very useful. So really, this is a lesson about two things.

2.1.1 Variables

Variables are quantities that we compute and that change over time. In this model, the variables we compute are the fraction of humans and mosquitoes that are infected at each point in time. Since it is a discrete time system, the values of the variables are defined only at integer values of t.

- Let x_t be the fraction of people who are infected at time t, and $0 \leq x_t \leq 1$.
- Let y_t be the fraction of mosquitoes who are infected at time t, $0 \leq y_t \leq 1$.

2.1.2 Initial Conditions

Since the values of our variables in the next time step (at time $t+1$) depend on their values now (at time t), we can't really compute anything unless we can say how the process gets started. What are the values of our variables at the beginning of our simulations (usually, at time $t = 0$)?

We set these initial values to be small:

```
# Initial Conditions, as a Named Vector
x0 = .01
y0 = .001
xy = c(t=0, x=x0, y=y0)
```

2.1.3 Parameters

The parameters describe the processes that are occurring in the population, the fraction of humans that clear an infection each day, and the fraction of infected mosquitoes that die. Unlike variables, parameters are passed to the model.

- Let s denote the fraction of people who clear infections after one day; $0 < s < 1$.
- Let u denote the fraction of mosquitoes who die in one day; $0 < u < 1$.
- Let a denote the fraction of mosquitoes who blood feed on a human in a day; $0 < a < 1$.
- Let m denote the number of mosquitoes per human; $m \geq 0$.

```
# The Parameters, as a Named Vector
par = list(
  s = 1/200, # The fraction of infections that clear each day
  u = 1/12,  # The fraction of mosquitoes that die each day
  a = 1/4,   # The fraction of mosquitoes that blood feed on a human each day
  m = 2      # The number of mosquitoes per human
)
```

2.1.4 Equations

Finally, we put all this together into a mathematical statement that has translated the description of a process. There are four terms:

- The fraction of humans who are infected is x_t; a fraction s clear infections.
- The fraction of mosquitoes who are infected is y_t; a fraction u die.
- The fraction of humans who are infected is $1 - x_t$; a fraction may_t will become infected.
- The fraction of mosquitoes who are not infected is $1 - y_t$; a fraction ax_t become infected.

$$\begin{aligned} x_{t+1} &= x_t - sx_t + may_t(1 - x_t) \\ y_{t+1} &= y_t - uy_t + ax_t(1 - y_t) \end{aligned}$$

We write a function that computes and returns the updated values of the variables. We adopt a simple naming convention: since this is a discrete time system, we will call the function `dts_xy_1`. It accepts the variables, which describe the state of the system at the current time (as a named vector) and the parameter values (as a list) and it returns the state of the system (as a named vector):

```
dts_xy_1 = function(xy, p){with(as.list(xy), with(p,{
  xn = x + m*a*y*(1-x) - s*x
  yn = y + a*x*(1-y) - u*y
  return(c(t=t+1, x=xn, y=yn))
}))}
```

2.1.5 Simulation

With the R code we developed, we can now compute the values of the variables after one time step. We initialize `xy` to hold the computed variables over time.

```
# xy_t stores the values of the variables
xy_t = xy
```

We name the variables so that later we can call them by name.

```
# Compute and store the values of the variables
xy = dts_xy_1(xy, par)
xy_t = rbind(xy_t, xy)
```

We can take a peak at the values we computed:

```
# Print
print(xy)
```

```
##           t           x           y
## 1.000000000 0.010445000 0.003414167
```

We can iterate over many time steps, storing the values in columns:

```
# Iterate to compute the values as they change over time
for(t in 2:40){
  xy = dts_xy_1(xy, par)
  xy_t = rbind(xy_t, xy)
}
```

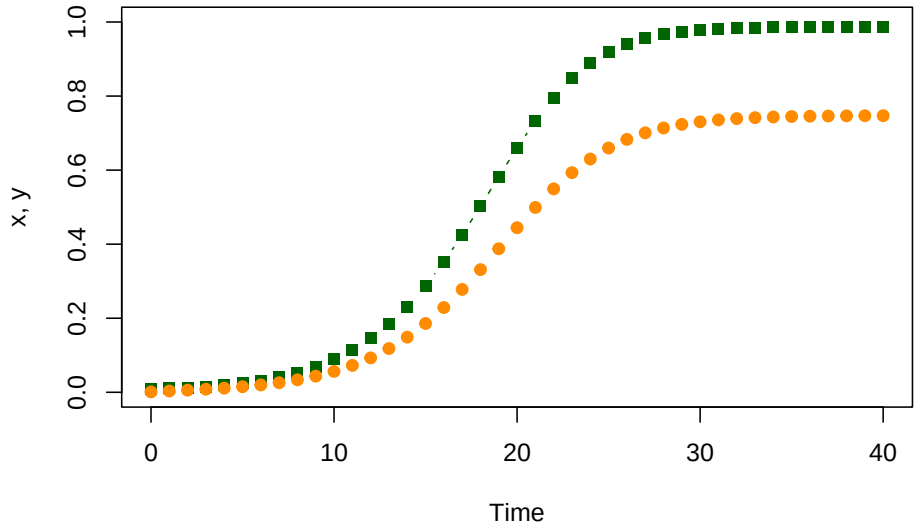
The way we implemented this, the values of x and y are stored in columns:

```
print(tail(xy_t, 3))
```

```
##     t         x         y
## xy 38 0.9865573 0.7465869
## xy 39 0.9866426 0.7468730
## xy 40 0.9866975 0.7470700
```

We can now plot out the values over the first 40 days

```
with(data.frame(xy_t),{
  plot(t, x, type = "b", ylim = c(0,1), col = "darkgreen", ylab = "x, y", xlab = "Time"
  lines(t, y, type = "b", col = "darkorange", pch =19)
})
```



After 40 days, the values of x and y are not changing very much:

```
# The last value is still stored as xy; the [-1] omits t
xy[-1] - dts_xy_1(xy, par)[-1]
```

```
##             x             y
```

```
## -3.545225e-05 -1.355019e-04
```

If we iterate another hundred days and check again, the differences are effectively zero.

```
for(i in 41:140){
  xy = dts_xy_1(xy, par)
  xy_t = rbind(xy_t, xy)
}
xy[-1] - dts_xy_1(xy, par)[-1]
```

```
## x y
## 0 0
```

After simulating, the variables reach a steady state, where asymptotically $x_{t+1} = x_t$ and $y_{t+1} = y_t$.

2.1.6 Analysis

We can figure out the steady state values are by substituting $x_{t+1} = x_t = x$ and $y_{t+1} = y_t = y$ into the equations above, and then solving for x and y.

$$\begin{aligned} x &= x - sx + may(1-x) \\ y &= y - uy + ax(1-y) \end{aligned}$$

or even easier:

$$\begin{aligned} may(1-x) &= sx \\ ax(1-y) &= uy \end{aligned}$$

The most obvious solution to these equations is $x = y = 0$, which in these models means that there is no malaria. We call it the disease-free steady state. The equilibrium makes some sense: if there are no infected mosquitoes or infected humans in a deterministic model, there can never be any.

There is another solution where malaria is present. We solve the second equation first:

$$y = ax/(u + ax).$$

We can substitute this for x into the first equation to get:

$$ma^2(1-x) = s(u + ax)$$

and now we solve for x

$$x = \frac{ma^2 - su}{ma^2 + sa}$$

We can write a function to compute this steady state:

```
# Compute the steady state
find_eq_dts_1 = function(par){with(par,{
  xx = (m*a^2 - s*u)/ (m*a^2 + s*a)
  yy = a*xx/(u+a*xx)
  c(x=xx,y=yy)
})}
```

2.1.7 Verification

We want to get used to double checking everything to avoid inserting mistakes. One way to do this is to find two or more ways of computing the same thing, for verification.

If we've done everything right, we ought to get the same values for the steady states through our analysis and simulation.

```
xy[-1]
```

```
##         x         y
## 0.9867987 0.7475000
```

```
find_eq_dts_1(par)
```

```
##         x         y
## 0.9867987 0.7475000
```

It's tempting to look at the printout and assume these two numbers are exactly equal. This is the land of computation, so things won't be exact. We can simply sum up the absolute values of the differences:

```
verify_dts_1 = function(xy, pars){
  sum(abs(xy[-1] - find_eq_dts_1(pars)))
}
verify_dts_1(xy, par)
```

```
## [1] 3.330669e-16
```

If we wanted to reduce this to a simple error check, we should pick a tolerance level – say 10^{-9} – and then just ask if we are closer than that:

```
check_it_dts_1 = function(xy, pars, tol=1e-9)
{
  verify_dts_1(xy, pars) < tol
```

```
}
check_it_dts_1(xy, par)
```

```
## [1] TRUE
```

2.1.8 Packaging Workflows

If we want to repeat tasks, we can simply write a wrapper around other functions that expedite the work. In this case, we want to write a function that *solves* the equations (*i.e.* that iteratively computes and stores the values) over some time interval:

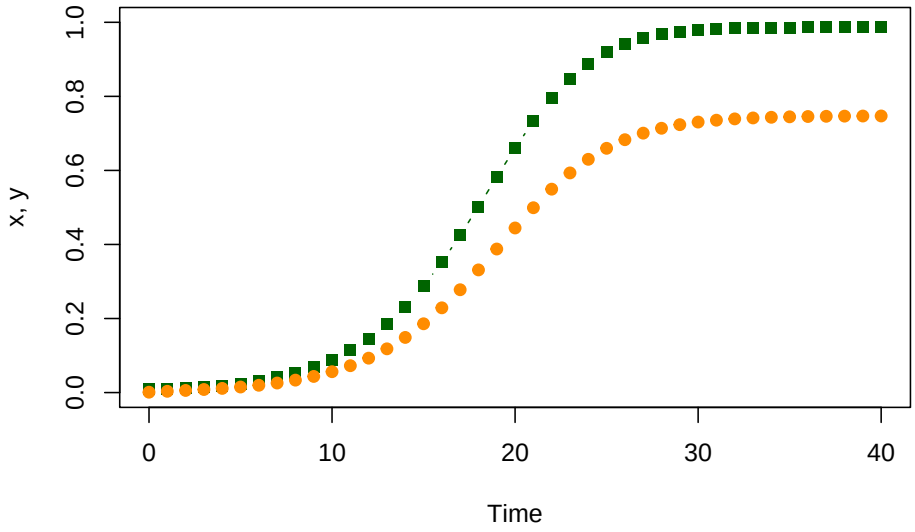
```
solve_dts_xy_1 = function(pars, x0=.01, y0 = 0.001, tmax=100){
   xy = c(t=0, x=x0, y=y0)
   xy_t = xy
   for(t in 1:tmax){
     xy = dts_xy_1(xy, pars)
     xy_t = rbind(xy_t, xy)
   }
   return(list(t=xy_t[,1], x=xy_t[,2], y = xy_t[,3], last = xy))
}
```

We can write another function that *plots* the equations (*i.e.* that iteratively computes and stores the values) over some time interval:

```
plot_xy = function(xy_t, add_points = FALSE){with(xy_t,{
  llty = "l"
  if(add_points == TRUE) llty = "b"
  plot(t, x, type = llty, ylim = c(0,1), col = "darkgreen", ylab = "x, y", xlab = "Time", pch=15)
  lines(t, y, type = llty, col = "darkorange", pch =19)
})}
```

Now, all the work we did above can be plotted using a single function call:

```
plot_xy(solve_dts_xy_1(par, tmax=40), add_points=TRUE)
```



Now, it is easier to use the models in various ways.

2.1.9 Thresholds

All our analysis worked out well for the parameter values that we chose, but what if we had picked different parameters?

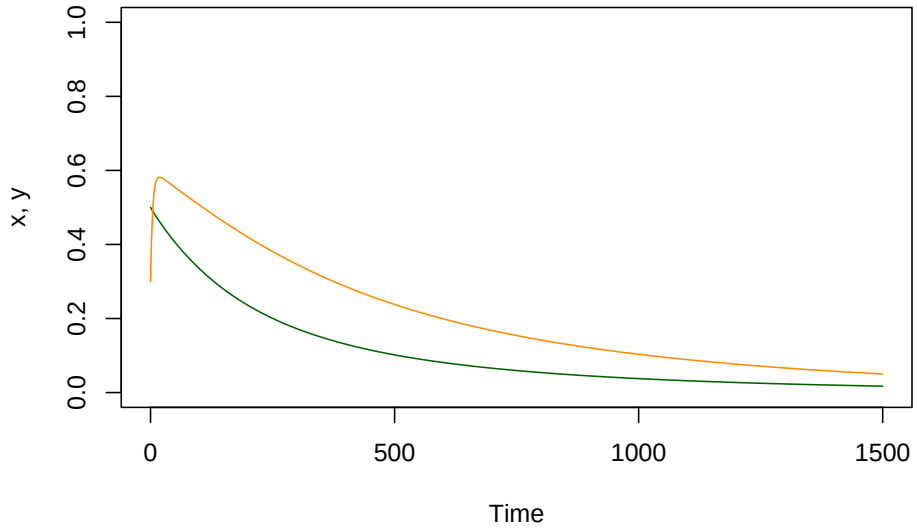
There must be some very low level of mosquitoes, for example, where malaria can't reproduce. If we reduce m to 0.005 and evaluate the expression at the steady state, we get negative values for x and y.

```
par1 = par
par1$m = 0.005
find_eq_dts_1(par1)
```

```
##           x           y
## -0.06666667 -0.25000000
```

What happens if we simulate this? (Let's set the initial conditions to reasonably high values)

```
plot_xy(solve_dts_xy_1(par1, x0 = .5, y0=.3, tmax=1500))
```



If we look at the equations, it's easy enough to spot the problem. Since x and y must be positive, then it must be true that

$$ma^2 > su.$$

We call this a threshold condition.

```
m_crit = with(par, s*u/a^2)
m_crit
```

```
## [1] 0.006666667
```

If we check, we find that this gives us the disease free equilibrium.

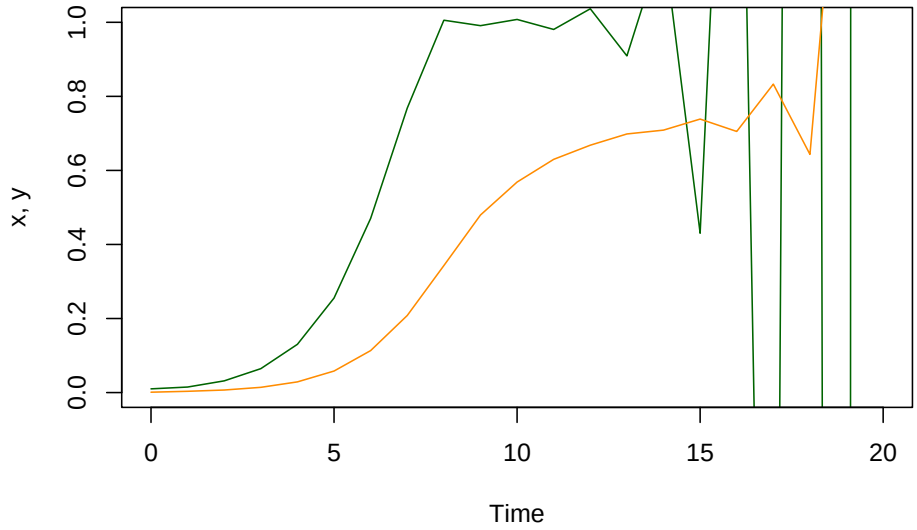
```
par2 = par
par2$m = m_crit
find_eq_dts_1(par2)
```

```
## x y
## 0 0
```

2.1.10 Numerical Stability

There is another problem with the equations. If we set m too high, such that at some point $may_t > 1$, then the whole system eventually crashes:

```
par3 = par
par3$m = 20
plot_xy(solve_dts_xy_1(par3, x0 = .01, y0=.001, tmax=20))
```



In discrete time formulations, we must be very careful to ensure that we have formulated a proper model. How can we fix this problem?

2.1.11 Improved Model

We have to go back to the assumption that the fraction getting infected is linearly proportional to the number of infective bites. One way to do this is to say that the *expected* number of bites would follow a Poisson distribution with mean may_t, so that the fraction getting infected is the zero term from a Poisson:

$$1 - e^{-may_t}$$

Now, our equations are the following:

$$\begin{aligned} x_{t+1} &= x_t - sx_t + (1 - e^{may_t})(1 - x_t) \\ y_{t+1} &= y_t - uy_t + ax_t(1 - y_t) \end{aligned}$$

Now, if we want to compute the steady state, we're stuck with the problem of solving this:

$$(1 - e^{-ma^2x/(u+ax)})(1 - x) = sx$$

It's surprisingly easy to write down equations, like this one, that we can't solve with pencil and paper. We can still find a way of computing the steady state, but we have to write R code that solves for x numerically.

```
find_eq_dts_2 = function(pars){with(pars,{
   f_xx = function(x, pp){with(pp,{
     xx = (1 - exp(-m*a^2*x/(u+a*x)))*(1-x) - s*x
     y = a*xx/(u+a*xx)
     yy = u*y + a*xx*(1-y)
     return(xx^2 + yy^2)
   })}
   xx = optimize(f_xx, c(0,1), pp=pars)$min
   yy = a*xx/(u+a*xx)
   c(xx, yy)
})}
```

The new equilibrium is at:

```
find_eq_dts_2(par3)
```

```
## [1] 0.9949026 0.7490406
```

We can just as easily write the code to numerically solve the discrete time system:

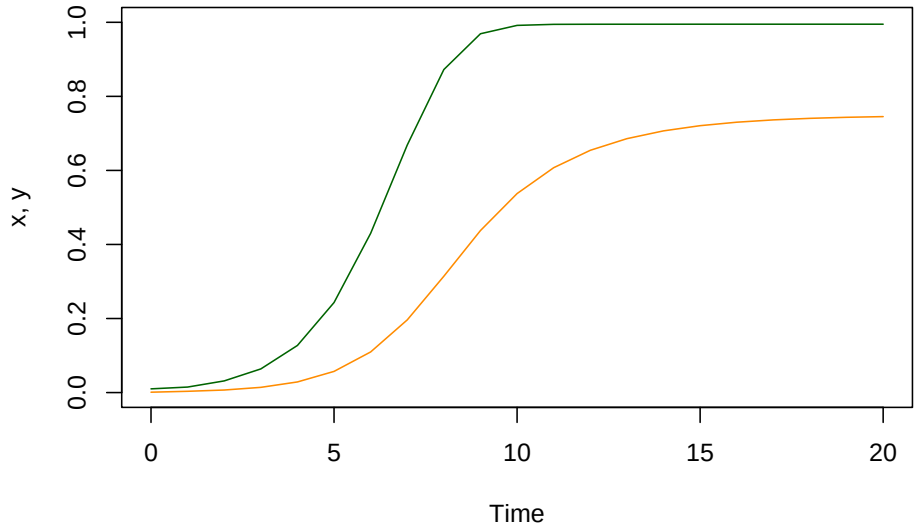
```
dts_xy_2 = function(xy, p){with(as.list(xy), with(p,{
  xn = x - s*x + (1-exp(-m*a*y))*(1-x)
  yn = y - u*y + a*x*(1-y)
  t=t+1
  return(c(t=t, x=xn, y=yn))
}))}
```

Once again, we can wrap a function around the solver so that it's easier to use the code:

```
solve_dts_xy_2 = function(pars, x0=.01, y0 = 0.001, tmax=100){
   xy = c(t=0, x=x0, y=y0)
   xy_t = xy
   for(t in 1:tmax){
     xy = dts_xy_2(xy, pars)
     xy_t = rbind(xy_t, xy)
   }
   return(list(t=xy_t[,1], x=xy_t[,2], y = xy_t[,3], last = xy))
}
```

Now, we can visualize the output and see that we have fixed our stability problem.

```
plot_xy(solve_dts_xy_2(par3, x0 = .01, y0=.001, tmax=20))
```




```
find_eq_dts_2(par3)
```

```
## [1] 0.9949026 0.7490406
```

```
solve_dts_xy_2(par3, x0=.1, y0=.05,tmax=300)$last[-1]
```

```
##         x         y
## 0.9949051 0.7490410
```

Once again, we want to ensure that our code does not have any mistakes, so we write a function to verify our results.

```
verify_dts_2 = function(pars, x0 = .1, y0=.05, tmax=200){
  xyt = solve_dts_xy_2(pars, x0=x0, y0=y0, tmax=tmax)
  sum(abs(xyt$last[-1] - find_eq_dts_2(pars)))
}
```

```
verify_dts_2(par3, tmax=300)
```

```
## [1] 2.924525e-06
```

We can write the function to check if the approximation is accurate to some level of tolerance:

```
check_it_dts_2 = function(pars, x0 = .1, y0=.05, tmax=200, tol=1e-5)
{verify_dts_2(pars, x0, y0, tmax) < tol}
```

```
check_it_dts_2(par3, tmax=200)
```

```
## [1] TRUE
```

2.1.12 Notes

Some useful notes:

1. The two models we just presented above can be implemented in a spreadsheet.
2.

2.2 Discrete Time, Stochastic

The world is stochastic.

2.3 Continuous Time, Deterministic

Here, we present a continuous time, deterministic model. It is homologous to Ross's 2nd model [57], after renaming the parameters. To deal with these equations, you need a background in calculus.

2.3.1 Variables and Initial Conditions

Once again, variables are quantities that we compute and that change over time. In this model, the variables we compute are the fraction of humans and mosquitoes that are infected at each point in time. Since it is a continuous time system, the values of the variables are defined for all values of t.

- Let $x(t)$ be the fraction of people who are infected at time t, and $0 \leq x(t) \leq 1$.
- Let $y(t)$ be the fraction of mosquitoes who are infected at time t, $0 \leq y(t) \leq 1$.

As before, we will need to define initial conditions.

2.3.2 Parameters

The parameters describe the processes that are occurring in the population, the fraction of humans that clear an infection each day, and the fraction of infected mosquitoes that die. Unlike variables, parameters are passed to the model.

- Let r denote the fraction of people who clear infections after one day; $0 < r < 1$.
- Let g denote the fraction of mosquitoes who die in one day; $0 < g < 1$.

- Let a denote the fraction of mosquitoes who blood feed on a human in a day; $0 < a < 1$.
- Let m denote the number of mosquitoes per human; $m \geq 0$.

```
# The Parameters, as a Named Vector
ross_xde_par = list(
  r = 1/200, # The fraction of infections that clear each day
  g = 1/12,  # The fraction of mosquitoes that die each day
  a = 1/4,   # The fraction of mosquitoes that blood feed on a human each day
  m = 2      # The number of mosquitoes per human
)
```

2.3.3 Equations

Finally, we put all this together into a mathematical statement that has translated the description of a process. There are four terms:

- The fraction of humans who are infected is x; infections clear at the rate r.
- The fraction of mosquitoes who are infected is y; mosquitoes die at the rate g.
- The fraction of humans who are infected is $1 - x$; infections occur at the rate may.
- The fraction of mosquitoes who are not infected is $1 - y$; infections occur at the rate ax.

$$\begin{aligned} dx/dt &= may(1-z) - rx \\ dy/dt &= ax(1-y) - gy \end{aligned}$$

```
ross_xde = function(t, y, params){
  with(params, with(as.list(y),{

  dxdt = m*a*y*(1-x) - r*x
  dydt = a*x*(1-y) - g*y

  return(list(c(dxdt, dydt)))
}))}
```

2.3.4 Solving

We use the R package ***deSolve***

```
library(deSolve)
```

The wrapper:

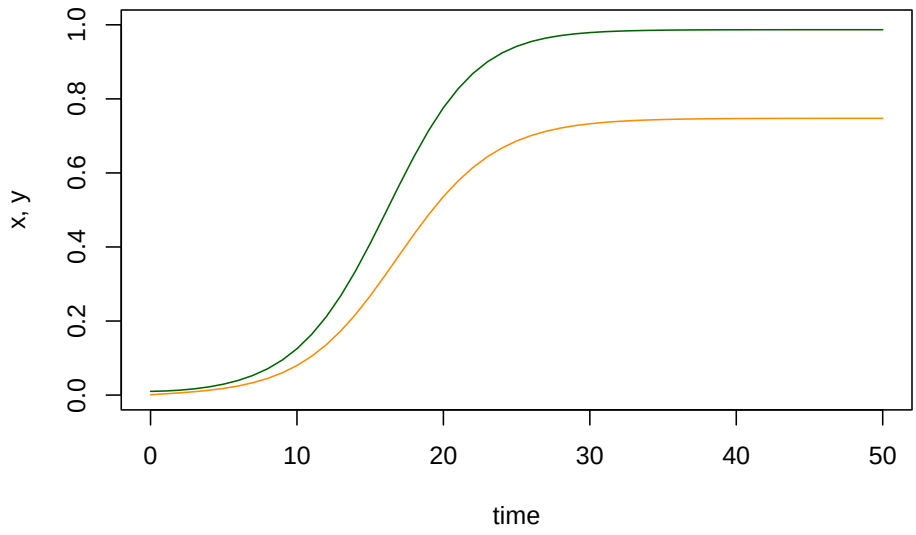
```
solve_ross_xde = function(pars, x0=.01, y0=0.001, tmax=50, dt=1){
  tms = seq(0, tmax, by = dt)
  xy0 = c(x=x0, y=y0)
  data.frame(ode(xy0, tms, ross_xde, pars))
}
```

It works.

```
deout <- solve_ross_xde(ross_xde_par)
```

Plot

```
with(deout, {
  plot(time, x, type = "l", col = "darkgreen", ylim = c(0,1), ylab = c("x, y"))
  lines(time, y, col = "darkorange")
})
```



These differential equations require

2.3.5 Verification

We set $dx/dt = dy/dt = 0$

$$\begin{aligned} 0 &= may(1-x) - rx \\ 0 &= ax(1-y) - gy \end{aligned}$$

Solve the second equation for y

$$y = ax/(g + ax)$$

Substitute back into the first equation:

$$ma^2(1 - x) = r(g + ax)$$

and solve.

$$x = \frac{ma^2 - rg}{ma^2 + ra}$$

Noting that $ma^2 > rg$ must be a threshold condition, we write a function to compute the steady state using the formula we just derived:

```
find_eq_xde_1 = function(pars){with(pars,{
  xx = ifelse(m*a^2 > r*g, (m*a^2 - r*g)/(m*a^2 + r*a), 0)
  yy = a*xx/(g+a*xx)
  c(x=xx, y=yy)
})}
```

```
find_eq_xde_1(ross_xde_par)
```

```
##         x         y
## 0.9867987 0.7475000
```

We check it against the other way, which involves running the system for a very long time.

```
de_eq <- solve_ross_xde(ross_xde_par, tmax=500, dt=5)
with(de_eq, c(x=tail(x,1), y=tail(y,1)))
```

```
##         x         y
## 0.9867987 0.7475000
```

2.4 Continuous Time, Stochastic

The world is stochastic.

2.4.1 Deterministic

2.4.2 Stochastic

We want to to introduce a mathematical model that is a good starting point. None of the models presented in the previous chapter are suitable, largely because

the variables in those models were proportions. The models we will use are almost identical, but the variables are densities.

The Ross-Macdonald model is useful as a concise way of understanding malaria transmsission through the parasite's life-cycle. There is, however, a great deal of complexity that the Ross-Macdonald model does can not address, and some of it matters a great deal in policy. We must extend the model using new mathematical constructs to deal with several features of malaria:

- malaria immunity and disease in cohorts of humans as they age;
- human demography;
- anti-malaria drugs and other pharmaceutical interventions;
- heterogeneous exposure
- spatial dynamics and malaria importation;
- mosquito ecology and behavior;
- exogenous forcing by weather, giving rise to seasonal and unstable dynamics;
- and vector control.

In the following chapters, we will introduce and critically review the models that have been developed to address these topics in some detail. As we take on these new features and add complexity, we will (in effect) be going around the parasite and mosquito life-cycles over and over again, each time pursuing a new theme.

We want to pursue the idea of scaling complexity in dynamic models ranging from simple and abstract to highly realistic. We want to arrive at a happy compromise – a model that is "as simple as possible, but no simpler." (This is often attributed to Einstein, who probably never said it that way exactly [46]). It might be hard to convince anyone that a model of intermediate complexity is good enough unless at least *some* of the models have clearly gone too far. We don't want to add new features to our models unless they would affect some important aspect of malaria transmission dynamics and control.

Ideally, we would let the data tell us about the right level of complexity, but how would we ever convince someone to collect the data we need? If we don't expect to be able to collect all the data, we can do some of the hard work up front to set some prioritize. One way to pursue questions about what details are likely to be biologically relevant is through model building and model-model comparison. If we can't identify a case where knowing something would change our policies using a mathematical model, it's probably not worth collecting the data for policy. We think of all these models as crude approximations to a messy reality – the question we want to address is whether the approximations are good enough to use as a basis for giving policy advice. Building very complex

models that are also mathematically and biologically sound and meaningful is a daunting task, but we have developed a modular system for building models that streamlines the model building process.

To get started, we will reintroduce the Ross-Macdonald model in a form that is easier to extend than the one Macdonald made famous [72]. We will start with a model that first appeared in a 1982 book chapter written by Joan Aron and Robert May [2]. It is a good starting point because it is *extensible.*

We chose it because the variables represent population densities, which are used to compute proportions, like *prevalence.* In the version of the Ross-Macdonald we presented before (in [Rebuilding Macdonald's Model]), the variables are proportions. In some models, we would like the *total* number of hosts to change over space or time, but if our state variables are proportions, then these densities appear in the denominators, which would make them much more difficult to modify. By writing down equations using densities, it is far easier to add complexity.

Why use densities?

To show why we use densities, we present a simple example. If we write down an equation describing changes in the density of infected humans, X, in a population with total human population density H. We let V denote vectorial capacity, and b the fraction of infective bites that cause an infection, and we assume the force of infection is bVX/H. The dynamics of infection are described by this simple equation:

$$\frac{dX}{dt} = bV\frac{X}{H}(H - X) - rX$$

In this equation, prevalence is $x = X/H$. Following through with the change in variables, we can write down the equation for the change in prevalence:

$$\frac{dx}{dt} = \frac{1}{H^2}\left(H\frac{dX}{dt} - X\frac{dH}{dt}\right)$$

and with some rearranging, we get:

$$\frac{dx}{dt} = bVx(1 - x) - rx - x\frac{dH}{dt}$$

The second equation is as simple as the first only if $dH/dt = 0$. Since we will want to deal with dynamical changes in host populations, we will avoid formulating base models that have proportions.

The model is extensible because, as we will see, mosquito population density is a variable, not a term. A parameter in the equation describing changes in mosquito density provides a link to aquatic mosquito ecology, which makes it possible to develop models for mosquito population dynamics, when that is an important consideration. In this version, we find that the formulas describing R_0 and vectorial capacity tell the story of parasite transmission through two blood meals. We will thus update the classical formulas. While Macdonald's analysis and formulas are familiar to many, they were incomplete [69, 70]. We develop a formula for vectorial capacity that is consistent with the intent of the original, but our analysis of sensitivity to parameters includes effects on mosquito ecology [8]. (A lengthy and philosophical discussion of the history and its failings is planned.)

We chose it because it is *realistic.* Time does not appear in most versions of the Ross-Macdonald model: the equations are *autonomous.* These equations use time to drive a seasonal pattern: they are *non-autonomous.* Since we know we are interested in dealing with exogenous forcing, we start out with a model that is forced.

While the following model is basic, we recommend reading it, if only because we introduce concepts and conventions that are important for the software design.

2.5 What is a Model?

2.6 Historical Notes

If we were simply learning the math, ...

When we analyze these equations to determine their *stability* and to identify *threshold conditions,* we focus on threshold conditions and the behavior of these systems when malaria is rare. In most places, malaria is endemic so we need to be concerned about malaria immunity and its effects on transmission; malaria is under some level of control; and because of weather and other factors, the baseline conditions change from year to year.

Chapter 3

Macdonald's Model

A few years ago, I went searching for *The Ross-Macdonald Model*. I was looking for a single publication where I could find a model written down in a concise form as a system of differential equations. Instead, I found the history of an idea that was spread across several publications [72]. What I sought was first published in 1982 [2, 4].

Macdonald never wrote down his model for malaria transmission as a system of equations in a single place. In 1950, he published a review of malaria malaria epidemiology [40], and a new model for superinfection [37]. In 1952, Macdonald published his analysis of the sporozoite rate [38], and then he wrote a paper about endemic malaria that included a formula for the basic reproductive number, R_0 [36]. Macdonald reports that the mathematical analysis of superinfection and the sporozoite rate had been done by Armitage; a paper appeared in 1953 [1]. The formulas have enough information to write down the equations he must have been working from. We present it in the next section.

After reading old papers, it became clear that Alfred Lotka played an important yet overlooked role in development of basic theory for malaria. Ross had actually published three mathematical models. Two of these described malaria transmission. Lotka analyzed both in 1923, and in a collaboration with Sharpe, he had extended the models to consider the effects of a delay. Lotka had also introduced the idea of a basic reproductive number for human demography. Macdonald adapted the idea to malaria, largely without attribution. It should, perhaps, be called the Ross-Lotka-Macdonald model.

All this history, spanning the period from roughly 1899 to 1969, has been reviewed before [72], and some historical notes are recounted at the end of this chapter. The codification of the Ross-Macdonald model in the 1980s coincided with a rising interest in disease ecology and mathematical epidemiology in academic departments of ecology.

3.1 The Model

3.1.1 Superinfection

Macdonald's model for superinfection...

3.1.2 The Sporozoite Rate

Macdonald made several important contributions to the mathematical study of malaria. Ross's original formulation remained the standard until Macdonald's paper on the sporozoite rate [38]. Because of it's historical interest, we'll discuss Ross's formulation in the discussion of Macdonald's Model, in part, to make it clear why Macdonald's contribution was significant enough to get his name attached.

In the 1950s, George Macdonald published a set of papers describing malaria transmission dynamics and control, culminating in publicataion of a book *The Epidemiology and Control of Malaria* [41].

In 1950, Macdonald published a new model with human malaria superinfection [37] and a synthetic review of malaria epidemiology [40]. In 1952, Macdonald published a synthetic review of medical entomology [38] and introduced the concept of a basic reproductive number, R_0 [36].

As a side note, it's highly likely that Macdonald was aware of Lotka's demographic concept of R_0, as well as his work with Sharpe on delay differential equations [65], though we have been unable to find a single line where Macdonald gives Lotka the credit. For these reasons, and others, the equations ought to be described as the Ross-Lotka-Macdonald model.

3.1.3 Variables

The model has three variables:

- x is the fraction of humans who are infected;
- y is the fraction of adult female mosquitoes who are infected.
- z is the fraction of adult female mosquitoes who are infective.

3.1.4 Parameters

The model has several parameters:

- g: mosquitoes die at a constant rate, g. This is equivalent to assuming that the mosquito lifespan is exponentially distributed with a mean $1/g$. Macdonald's equation use p, the fraction surviving one day, $p = e^{-g}$.
- τ: it takes τ days for parasites to mature and reach the salivary glands, called the EIP. The fraction surviving the EIP is $p^\tau = e^{-g\tau}$.
- m: there are m mosquitoes per human;
- a: mosquitoes blood feed on humans at the rate a;
- b: a fraction of bites by *infectious* mosquitoes causes an infection;
- r: human malaria infections clear at the rate r;
- c: a fraction of bites on infected humans infect a mosquito, denoted c

3.1.5 Equations

In this formulation of the model, we ignore the delay for the EIP but we count the mortality:

$$\begin{aligned} dx/dt &= bmaz(1-x) - rx \\ dy/dt &= cax(1-y) - gy \\ dz/dt &= e^{-g\tau}cax_\tau(1-y_\tau) - gz \end{aligned}$$

this.

3.2 The Basic Reproductive Number

3.3 Sensitivity to Parameters

3.4 The Classification of Transmission

3.5 Historical Notes

Ross's first model described malaria transmission dynamics using difference equations [55]. That model was reviewed, analyzed, and critiqued first by H. Waite in 1910 [75].

Ross's second model.

The Ross-Macdonald model and its development have been discussed elsewhere [2, 4, 72]; it was a model developed by Macdonald that was based on Ross's

earlier work and that was supported by the mathematical talents of Armitage. There is not a single paper where Macdonald described the system of differential equations. Instead, the model appeared in parts of several different publications that presented equations and formulas describing malaria transmission, and that reviewed existing data.

The model that Maconald started with was first developed by Ronald Ross (who published two models of malaria transmission), but we also owe a lot to Alfred Lotka, who analyzed both models.

The formulation is a bit hard to follow, so we will present a system of difference equations in place of the equations Ross actually wrote down.

Lotka

Ross's second model was thoroughly analyzed by Lotka, who had taken an active interest in Ross's malaria models. In 1912, he published a set of solutions to Ross's equations [31]. In 1923, Lotka published an analysis of both of Ross's models in five parts. The first two parts reformulate Ross's models [32, 33]. The third part tackles numerical issues, which includes a photograph of a clay model of the phase plane as a surface [34], and the fifth part is a concise summary [35]. In the 4^{th}, which was led by Sharpe, a new model was introduced that included delays for the latent periods [65].
Alfred J. Lotka

- While Alfred J. Lotka is more famous for his work in demography and ecology, he took an interest in Ross's work on malaria and he made some important contributions to mathematical malaria epidemiology:
- Most importantly, Lotka developed the concept of the basic reproductive number in his work in human demography, which was defined as the expected number of females that would be born to a newborn female.

3.5.1 George Macdonald

we'll discuss Ross's actual formulation,

Chapter 4

Measuring Malaria

4.1 Infection

4.2 Malaria Incidence

4.3 Exposure Risk

4.4 Mosquito Bionomics

Chapter 5

Malaria Control

5.1 Sensitivity to Parameters

Part II

Malaria Modeling

Chapter 6

Model Building

> *There must have been a time, in the beginning, when we could have said – no. But somehow we missed it.* (Guildenstern)
> — from *Rosencranz and Guildenstern are Dead*, Tom Stoppard

The Ross-Macdonald model is a good way to convey of the basic features of malaria epidemiology. The model is useful in the classroom, and it can help to inform our basic understanding of malaria. If we want to use the model to answer some policy questions, we will soon find that the model has some limitations.

but this book is about how we want to use the models for other purposes.

There are many reasons why we would not want to rely on the Ross-Macdonald model to simulate malaria for policies.

If we want to scale up simulation-based analytics to support malaria policy through eradication, it will inevitably involve capacity building, but what sort of training is needed? Certainly, we hope that an analytics team would include someone who has the skills required to critically evaluate models, but this is not as easy as you think. What makes a model good? What makes an analysis useful? It is easy to imagine training up a generation who could do some kind of analysis, but we don't want to train automata – people who would keep doing some analysis, over and over, just because it is what has been done before. We ought to have a reason for doing the analysis. It ought to serve some purpose. Success in policy analytics requires people who have skills in thinking creatively and critically to solve problems. We need people who can build and evaluate models.

If we want to discuss the models we might use to analyze malaria policies, and if we want to have a discussion about that process, it will probably help us to be more careful about how we use terms like *model* and *model building.* In the service of defining the skill sets required to be a good analyst, this chapter discusses the

life-cycle of a model, including the process of building and evaluating models from the beginning to the end (or in math from α to ω).

6.1 Compartment Models

6.2

6.3 Primary Model Development

Most models started with a question: something like *Why is the sky blue?* or *How many mosquitoes does it take to sustain malaria transmission?* Blue sky questions don't go anywhere unless there's a basis for understanding the underlying process. If there is some kind of basis, the next step is to try represent the process in some logically compelling form, such as a diagram with boxes and arrows.

6.4 Implementation and Verification

6.4.1 Solving

- **Fitted Parameters:**
- **Initial Conditions:**

6.4.2 Mathematical Analysis

- **steady states**
- **stability analysis**
- **algorithms**
- **primary analysis**

6.4.3 Simulation

6.5 What is a Model?

6.6 Accuracy & Validation

- Model features
- Functional range
- Heterogeneity

6.6.1 Validation Points

6.7 The Measure of a Model

If we want to evaluate whether a model is good

6.7.1 Limitations

- **Functional Range** the maximum granularity

6.7.2 Parsimony

6.8 Secondary Model Development

- **elaboration**

Model building is a core part of applied dynamics, but if we are to othere are many different activities

What does it mean to build a model?

6.9 Conclusions

Chapter 7

Adult Mosquito Dynamics

If we are to do anything in malaria, we must move past Macdonald's assumption that mosquito population density does not fluctuate.

The goal of this book is to present mathematical models for malaria transmission dynamics and control that are developed well enough to support malaria policies.

7.1 Mosquito Ecology

7.1.1 Variables

Aron & May, 1982

We define the following variables:

- M is the density of mosquitoes.
- Y is the density of infected mosquitoes.
- Z is the density of infectious mosquitoes.
- X is the density of infected humans.

In dynamical systems, we ask how the variables (*i.e.* M, Y, Z, and X) change over time. In the following, we describe the changes on variable about a time.

For our first equation, we start with adult, female mosquito populations. (It is tiresome to repeat *adult, female* each time, and we're ignoring male mosquitoes at this point anyway, so *mosquito* hereafter means *adult, female mosquito*, unless we say otherwise.) The number of mosquitoes is changing as new adults emerge from aquatic habitats or die.

To model changes in M, we assume the following:

- mosquitoes emerge from aquatic habitats at the rate of $\Lambda(t)$ adults, per day;
- mosquitoes die at a constant rate, g. This is equivalent to assuming that the mosquito lifespan is exponentially distributed with a mean $1/g$. The fraction surviving one day is e^{-g}.

Our first equation describes changes in the number of mosquitoes:

$$\frac{dM}{dt} = \Lambda(t) - gM \tag{7.1}$$

7.1.2 Blood Fed Mosquitoes

At this point, we will take a detour and define a variable describing the density of mosquitoes that have blood fed at least once, V. After blood feeding, a mosquito is either gravid or *parous*, meaning its ovaries are distended from laying an egg batch. We do this, in part, because the fraction of mosquitoes that are parous is routinely collected, and because it gives us a chance to focus on blood feeding.

To describe *blood feeding*, we assume the following:

- mosquitoes blood feed at the rate f, per mosquito, per day; in this model, this implies that the waiting time to a blood meal is $1/f$ days.
- a fraction of all mosquito blood meals, q, is taken on humans; we call this the *human fraction*
- the human blood feeding rate is the product of these two parameters, fq, which is defined as the number of human blood meals, per mosquito, per day.

The number of human blood meals by a population of vector mosquitoes, per person, per day is called the human biting rate (HBR). In this model, HBR is given by a formula:

$$\text{HBR} = \frac{fqM}{H}$$

Later, we discuss the correspondence between the HBR in models and data.

$$\frac{dY}{dt} = fq(M - Y) - gV \tag{7.2}$$

We won't use V to describe the dynamics of infection, but we might find it useful to understand how parity changes in mosquito populations.

7.1.3 Infected Mosquitoes

Mosquitoes become infected after blood feeding on an infectious human. To model changes in Y, we extend the model of blood feeding to include infection. We need to know what fraction of blood meals end up infecting a mosquito that has not already been infected.

To model changes in Y, we need to describe *infection rates*. We assume the following:

- a fraction of human blood meals, infects mosquitoes. We call this quantity *net infectiousness* (NI) and (for reasons that we will discuss in a moment), we give it a name, κ:

$$\kappa(t) = c\frac{X(t)}{H} \tag{7.3}$$

- infected mosquitoes die at the same rate as uninfected mosquitoes.

We can now write down our second equation describing changes in the number of infected mosquitoes:

$$\frac{dY}{dt} = fq\kappa(M - Y) - gY \tag{7.4}$$

7.1.4 Infectious Mosquitoes

To become infectious, a mosquito has to become infected and then survive through the extrinsic incubation period (EIP). We assume:

- mosquitoes become infectious after a fixed delay, τ days, called the EIP. The fraction of mosquitoes that survive through the EIP is $e^{-g\tau}$.
- infectious mosquitoes die at the same rate as other mosquitoes.

For a mosquito to become infectious, it must have become infected τ days ago and survived through τ days with probability $e^{-g\tau}$. To write this in equations, we use a subscripted τ to denote the value of a variable (M, Y or X) or term (κ) at time $t - \tau$. For example X_τ is the number of people who were infected and infecious at time $t - \tau$, and M_τ is the number of mosquitoes at time $t - \tau$.

The number of infectious mosquitoes that are added to the population at a point in time includes all the mosquitoes that became infected at time $t - \tau$ and survived the EIP. This is our third equation describing changes in the number of infectious mosquitoes:

$$\frac{dZ}{dt} = fq\kappa_\tau(M_\tau - Y_\tau)e^{-g\tau} - gZ \tag{7.5}$$

Here, Z represents the number of mosquitoes with *sporozoites* in their salivary glands. The *fraction* of mosquitoes with sporozoites in their salivary glands has been called the *sporozoite rate* (SR), which in our notation is

$$z = \frac{Z}{M}$$

The number of bites by vector mosquitoes, per person, per day is called the entomological inoculation rate (EIR). It is defined as the product of the HBR and the SR:

$$\text{EIR} = \text{SR} \times \text{HBR}$$

In our notation, the EIR is:

$$\text{EIR} = z\frac{fqM}{H} = \frac{fqZ}{H}$$

As with the HBR, we would like to know how to connect estiamted values of the EIR to our formulas. Since that's *really* complicated, we've spent a lot of time in the following sections discussing it.

7.1.5 Infected Humans

Humans become infected after being bitten by an infectious mosquito. We assume the following:

- A fraction b of all bites by infectious mosquitoes cause an infection.
- The hazard rate for infection, also called the *force of infection* (FoI) and denoted h is $b\times$ EIR:
$$h = fqb\frac{Z}{H}$$
- Infections clear at the rate r, per infection, per day (the average time to clear is $1/r$ days), and after clearing an infection a person becomes susceptible to infection again.

We can now write down our fourth equation describing changes in the number of infected humans:

$$\frac{dX}{dt} = h(H - X) - rX \tag{7.6}$$

7.1.6 . . . as a System

While we presented these equations one at a time, they work as a system. To see it all at once, we write it here as a system with four equations and two terms:

$$\begin{array}{rl} \frac{dM}{dt} & = \Lambda(t) - gM \\ \frac{dY}{dt} & = fq\kappa(M - Y) - gY \\ \frac{dZ}{dt} & = fq\kappa_\tau(M_\tau - Y_\tau)e^{-g\tau} - gZ \\ \frac{dX}{dt} & = h(H - X) - rX \end{array} \qquad (7.7)$$

$$\begin{array}{rl} \kappa & = c\frac{X(t)}{H} \\ h & = bfq\frac{Z(t)}{H} \end{array}$$

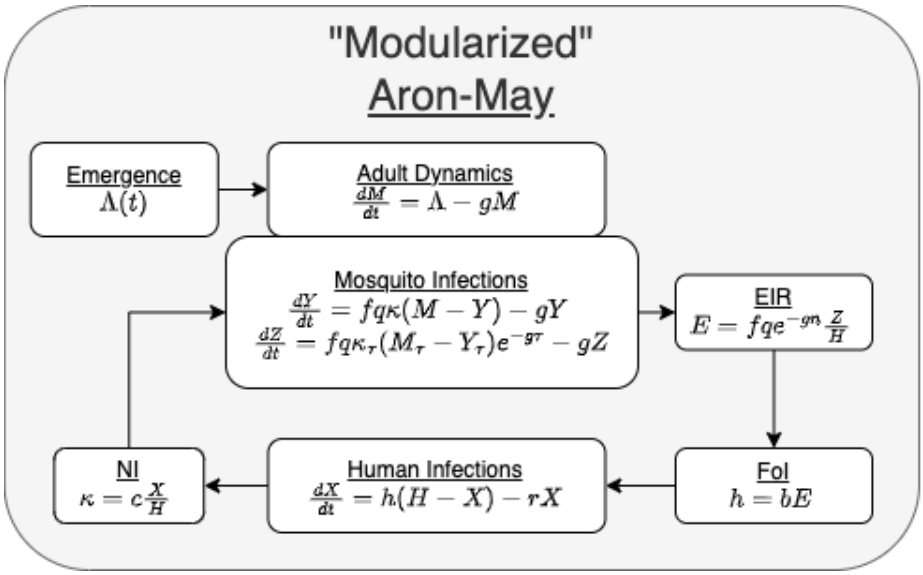


Figure 7.1: A diagram of the a version of the Ross-Macdonald model, using equations from Aron and May [2]

These equations describe processes in three domains (Figure 2.1):

- adult mosquito ecology (M, and perhaps V);
- parasite infection dynamics in mosquito populations (Y and Z);
- parasite infection dynamics in human populations (X).

The equations describing parasite infections in mosquito populations also include the variable M, so the mosquito infection dynamics are coupled to the mosquito population dynamics. The way we've written the equations, each compartment

has an input term (*i.e.*, Λ, κ, or h) that depends on something else. We've passed Λ as a parameter. For the infection dynamics, the terms κ and h couple two separate systems. For adult mosquito dynamics, emergence is passed to the model as a parameters.

There are, of course, more compact ways of writing these equations. We have written the equations this way to emphasize a few things. First, the terms make it clear exactly how the equations in one domain are connected to another. Second, if we wanted to start *changing* some of the assumptions, these terms help to isolate the parts we might like to change. By writing the equations in this modularized form, we can start to understand how we might be able to write software that would allow us to represent mosquito infection dynamics with different systems of equations.

The next step is to find solutions.

NOTE: We don't introduce `exDE` or `MicroMoB` until Modularity and Software.

7.1.7 Solutions

What does a **solution** to these equations look like?

Solutions to these equations are values of the variables over time $(M(t), Y(t), Z(t), X(t))$ that satisfy the system of four equations described above. We call these solutions *orbits.* To put it another way, if we took the derivatives of the orbits for any variable at any point in time using the basic definition

$$\lim_{h \to 0} \frac{x(t+h) - x(t)}{h},$$

and then we used the values of the variables at time t to compute dM/dt, dY/dt, dZ/dt, and dX/dt (*i.e.*, using the formulas), we would get the same values.

It is important that these orbits are unique: after specifying the *initial values* of the variables, there is one and only one set of orbits that solves the equations. When we solve the equations, we usually produce solutions from a starting point into a future, but the orbits are defined for all time – *i.e.* the process implies the existence of solutions far back into the past. These are deterministic equations, after all.

As written, the equations do not define a *model.* Instead, the equations define a process or a **model family.** A model is something that *can* produce orbits. A model is defined only after assigning specific values to the parameters. Informally, we will often slip and use the "model" to describe a model family. It's easy to slip up, and sometimes we can get by with being sloppy, but we need to remember the distinction. When we say that the software is *modular,* we mean that it is easy to swap out one *model family* for another.

To find solutions of equations we use an R software package called `deSolve.` Because of the delay for the EIP, these are called *delay differential equations,* which are handled using a function called `dede`. An important step in solving delay differential equations is a function `lagvalue()` that computes and returns the values of variables at a time lag, ℓ. In these equations, the lag is set by the EIP, τ, so we must evaluate `lagvalue(t-tau).`

In solving *ordinary differential equations,* we must pass initial conditions. To solve a delay differential equations with a maximum lag ℓ, we must specify the initial conditions for the interval $[-\ell, t_0)$, where t_0 is the point in time when we start computing solutions. In these equations, since the equation for dZ/dt looks back τ units, we must specify values of $M(t)$, $Y(t)$, and $X(t)$ for all values of $t \in [-\tau, t_0)$. This forces an awkward choice, since we don't know the solutions backwards in time, but would need to know those solutions to use them. What is typically done – and we've done it here – is to specify a constant set of initial values and moving on.

Doing this introduces a little *numerical slop.* By slop, we mean that these values are *not* what we would get if we ran the equations backwards in time. In these equations, it won't affect our analysis most of the time, so we're happy to acknowledge this little problem and find ways around it. It's a little thing, but we should never forget it, because we might find that it *is* affecting our analysis at some point.

With `deSolve,` solving differential equations is not difficult – it just involves following a few steps. In the following, we walk through these steps:

- Write a function that computes the derivatives;
- Define initial conditions;
- Define the values of the parameters;
- Define a mesh on time;
- Call a function that solves the equations, such as `dede` for delay differential equations.

Many users will find that reading this code is like learning how to compute $\sqrt{2}$. If so, feel free to learn it once and then skip it.

7.1.7.1 Derivatives

The first step is to write down the equations to compute the derivatives. The solver expects a function with three required arguments (in this order):

- `t` is time
- `y` is the list of variables
- `params` is a set of parameters

The derivatives are computed and returned in the same order as 'y' in a `list`. To make code that is easy to read, we make `params` as a `list` with parameter names (see below), so that inside the function `with(params,{...})`, the parameter names are visible.

```
dAronMay = function(t, y, params){with(params,{

  # Variables
  if(t<=tau) ylag<-y0 else ylag <- lagvalue(t-tau)
  M=y[1]; M_tau = ylag[1]
  Y=y[2]; Y_tau = ylag[2];
  Z=y[3];
  X=y[4]; X_tau = ylag[4]

  # Terms
  kappa = c*X/H; kappa_tau = c*X_tau/H
  h = b*f*q*Z/H

  # Dynamics
  dM = Lambda(t) - g*M
  dY = f*q*kappa*(M-Y) -g*Y
  dZ = f*q*kappa_tau*(M_tau-Y_tau)*exp(-g*tau) -g*Z
  dX = h*(H-X)-r*X

  return(list(c(dM, dY, dZ, dX)))
})}
```

7.1.8 Initial Values

To run the model, we must supply initial values. If you were writing code yourself, it would be important to remember that the initial values and the return value for the derivatives must occur in the same order.

A useful convention in {R} is to pass the initial values as a named list. Later, we can turn the outputs into a data frame, and then we can retrieve the variables by name.

```
y0= c(M=60, Y=0, Z=0, X=1)
```

The object `y0` is a named list – the names are attached but invisible.

```
y0
```

```
##  M  Y  Z  X
## 60  0  0  1
```

When we turn it into a list, with `as.list,` the names are attached to the values:

```
as.list(y0)$M
```

```
## [1] 60
```

If we use `with`, we create an environment where we can simply use the names:

```
with(as.list(y0), {
  M
})
```

```
## [1] 60
```

7.1.9 Parameter Values

We pass the parameters as a list. It might seem like overkill, but we have written a function `makeParams()` that takes default values and generates a list. This makes it easy to generate a new set of parameter values with alternative values, and it also helps us to write and pass function $\Lambda(t)$ with parameters we like. By passing the parameter as a list, the parameter values are available to the function `dAronMay` when we use `with(params, {})`.

Note that we have also attached the initial values of the variables as a parameter set, which are the return values for `lagvalue(t)` when `t<0`.

```
makeParams = function(y0,
                      g=1/12, f=1/2.5, q=0.95,
                      c=0.15,
                      b=0.55, r=1/200, H=1000,
                      m=.05, ss=1,
                      tau=10
                      ){
  ss = min(1,max(0, ss))
  Lambda = function(t){m*H*(1 + ss*sin(2*pi*t/365))}
  return(list(y0=y0,g=g,f=f,q=q,c=c,
              H=H,m=m,tau=tau,b=b,r=r,Lambda=Lambda))
}
params = makeParams(y0)
```

To make it absolutely clear, we are assuming:

- $g = 1/12$: mosquitoes live about 12 days, on average
- $f = 1/2.5$: mosquitoes feed every 2.5 days, on average
- $q = 0.95$: the human fraction is 95%; mosquitoes feed on humans 95% of the time
- $c = 0.15$: about 15% of bites on infectious humans infect a mosquito

- $b = 0.55$: about 55% of bites by infective mosquitoes cause an infection
- $r = 1/200$: human infections last about 200 days, on average
- $H = 1000$: we're simulating transmission in a population of a thousand humans
- $\tau = 10$: the extrinsic incubation period is about 10 days
- For emergence, we tune the average value using m and it is scaled to H:
  - The parameter m in the function above has been set to 0.05 by default.
  - The parameter ss affects the amplitude of the fluctuations. We force it to take on values between 0 and 1.
  - Emergence is modeled as a sinusoidal function with a yearly cycle.

$$\Lambda(t) = mH\left(1 + \sin\left(\frac{2\pi t}{365}\right)\right)$$

7.1.10 Solving

We define a mesh over time – the points in time when we would like to know the values of the variables:

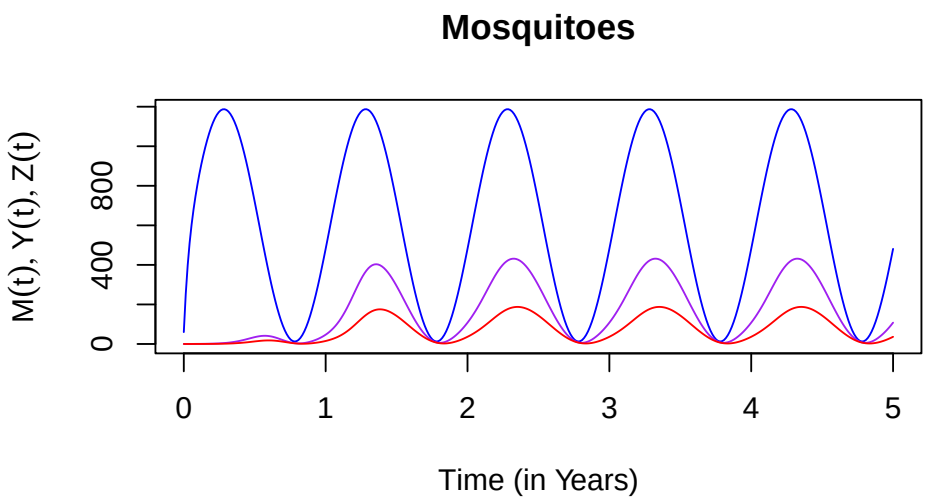
```
tt = seq(0,5*365, by=5)
```

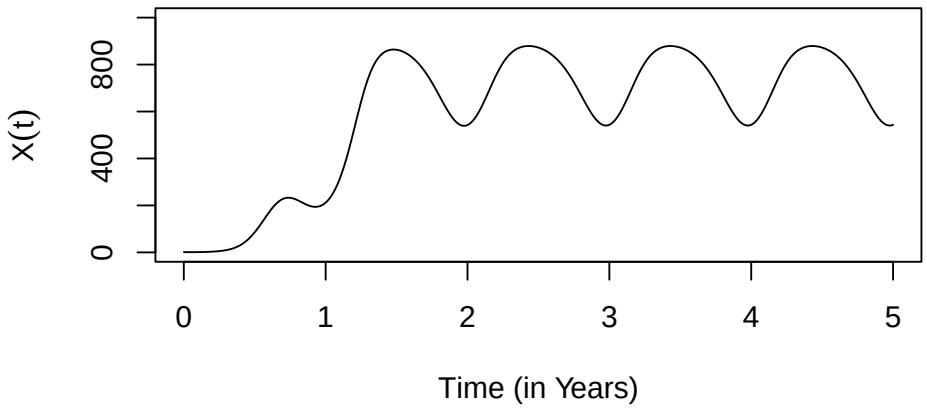
This code solves the equations:

```
require(deSolve)
yout <- dede(y=y0, times=tt, func=dAronMay, parms=params)
```

7.1.11 Visualizing

When we plot out the solutions, they look like this.


Mosquitoes
M(t), Y(t), Z(t)
0
400
800
0
1
2
3
4
5
Time (in Years)


Humans
X(t)
0
400
800
0
1
2
3
4
5
Time (in Years)

7.2 Steady States

Here, we analyze the system of equations in a narrow case when there is *no seasonality,* and the system reaches a *steady state.* To do so, we set the parameter `ss=1`, so that $\Lambda(t)$ is a constant; the resulting system is *autonomous.* We do this, in part, because the resulting system is easier to understand. We can develop intuition that can be applied (albeit with caution) to more complex systems. To be clear, we are dealing with this system:

$$\begin{array}{ll} \frac{dM}{dt} & = \Lambda - gM \\ \frac{dY}{dt} & = fq\kappa(M - Y) - gY \\ \frac{dZ}{dt} & = fq\kappa_\tau(M_\tau - Y_\tau)e^{-g\tau} - gZ \\ \frac{dX}{dt} & = h(H - X) - rX \end{array} \qquad (7.8)$$

$$\begin{array}{ll} \kappa & = c\frac{X(t)}{H} \\ h & = bfq\frac{Z(t)}{H} \end{array}$$

As before, we have put the equations in their modularized form above, and the connecting terms below.

The first thing to note is that M affects Y and Z, which affect X; but M is not affected by Y or Z. Mosquito population density is *exogenous* to malaria dynamics.

7.2.1 Mosquito Density

We can thus treat it separately in the analysis:

$$\frac{dM}{dt} = \Lambda - gM \qquad (7.9)$$

Since emergence rates are steady, mosquito population density reaches a steady state when $dM/dt = 0$, which occurs at:

$$\bar{M} = \frac{\Lambda}{g} \qquad (7.10)$$

7.2.2 EIR

Next, we note that at a steady state, the delayed values of variables and terms don't change, so from dY/dt, we get:

$$g\bar{Y} = fq\kappa(\bar{M} - \bar{Y}) \tag{7.11}$$

If we substitute the formula for $\bar{M}$ and solve for $\bar{Y}$, we get:

$$\bar{Y} = \frac{fq\kappa}{g + fq\kappa}\frac{\Lambda}{g} \tag{7.12}$$

and since at the steady state, any variable at time $t + \tau$ is equal to its value at time t, we substitute the formula for $g\bar{Y}$ into dZ/dt to get:

$$g\bar{Y}e^{-g\tau} = g\bar{Z} \tag{7.13}$$

Solving for $\bar{Z}$ we get:

$$\bar{Z} = \frac{fq\kappa}{g + fq\kappa}\frac{\Lambda}{g}e^{-g\tau} \tag{7.14}$$

At the steady state,

$$\text{EIR} = fq\frac{\bar{Z}}{H}.$$

In field studies, the EIR is the product of the HBR and the sporozoite rate (SR). The sporozoite rate (SR, z) is given by:

$$\bar{z} = \frac{Z}{M} = \frac{fq\kappa}{g + fq\kappa}e^{-g\tau} \tag{7.15}$$

So we can understand the EIR as having two parts:

$$\text{EIR} = \text{HBR} \times \text{SR} \tag{7.16}$$

or equivalently

$$\text{EIR} = \frac{fq\Lambda}{H} \times \frac{fq\kappa}{g + fq\kappa}e^{-g\tau} \tag{7.17}$$

This formula for the SR (albeit with slightly different notation) was originally derived as part of the Ross-Macdonald model [38, 1]. Also, Smith and McKenzie (2004) have written a useful discussion of mosquito demography [69].

7.2.3 Vectorial Capacity

Here, we pause to define a term that describes the number of human blood meals each mosquito would take over its whole life:

$$S = \frac{fq}{g}.$$

Since $1/g$ is the mosquito lifespan in days, and fq is the human blood feeding rate, S is the number of human bloodmeals a mosquito would take over its lifespan. Intuitively, it makes sense that this *should* be what drives transmission, since it takes two human blood meals to transmit malaria parasites.

If we rearrange the terms a bit, we can rewrite out the expression for the EIR:

$$\text{EIR}(\kappa) = fq\frac{\bar{Z}}{H} = \frac{\Lambda}{H}S^2 e^{-g\tau}\frac{\kappa}{1+S\kappa} \tag{7.18}$$

This formula for the EIR has two parts. We call the first part *vectorial capacity* (V):

$$V = \frac{\Lambda}{H}S^2 e^{-g\tau} \tag{7.19}$$

The second part is an expression that involves mainly κ.

$$\frac{\kappa}{1+S\kappa} \tag{7.20}$$

The relationship between VC and EIR at a steady state is a product:

$$\text{EIR}(\kappa) = V\frac{\kappa}{1+S\kappa} \tag{7.21}$$

Vectorial capacity describes the slope of the EIR when κ is small:

$$d\frac{\text{EIR}(\kappa)}{d\kappa}\bigg|_{\kappa=0} = V \tag{7.22}$$

We say that VC describes *potential* transmission, even if the parasites are absent. Another way to say the same thing is that when κ is small, then:

$$\text{EIR}(\kappa) \approx V\kappa \tag{7.23}$$

We can interpret vectorial capacity (V) in simple terms. It describes *the number of infective bites that would arise from all the mosquitoes biting a single human*

on a single day but only *if all those mosquitoes became infected.* Vectorial capacity tells the story of potential parasite transmission by mosquitoes in four steps, which highlights the fact that two human blood meals are required for the parasite to be transmitted and complete its life-cycle.

$$\left|\begin{array}{c} \Lambda/H \\ \\ \text{Mosquito} \\ \text{Emerges} \end{array}\right| \rightarrow \left|\begin{array}{c} S\kappa \\ \\ \text{Parasite} \\ \text{Infects} \\ \text{Mosquito} \end{array}\right| \rightarrow \left|\begin{array}{c} e^{-g\tau} \\ \\ \text{Mosquito} \\ \text{Survives} \\ \text{EIP} \end{array}\right| \rightarrow \left|\begin{array}{c} S \\ \\ \text{Parasite} \\ \text{Infects} \\ \text{Human} \end{array}\right| \tag{7.24}$$

As a reminder, while Eq. (7.24) includes κ, the formula for VC, in Eq. (7.19), assumes that $\kappa = 1$: the VC describes transmission as if humans were perfectly infectious. It was defined this way on purpose: it was meant to include mosquito parameters and exclude human factors. We can think of VC as defining something like a conditional expectation, a maximum, or (as we have already said) a measure of *potential transmission by mosquitoes* that is independent of human factors.

While κ (the numerator in Eq.(7.20) accounts for *most* of the difference between the EIR and the VC, the rest of the difference is due to the denominator in Eq. (7.20), $1 + S\kappa$, which traces back to the formula from dY/dt, which assumes that mosquitoes are either infected or not. The denominator is a measure of saturation – the fraction of mosquitoes that get *superinfected* with parasites. The main point here is that as κ increases, saturation increases. If we set S to the values in the previous plots, we can isolate the relationship:

In these formulas, the measure of saturation is exactly $1 + S\kappa$. We could rewrite the relationship between the EIR and VC in a way that tells us something about how we might be underestimating a parasite's reproductive success:

$$\text{EIR} \times (1 + S\kappa) = V\kappa$$

which suggests that each infectious bite is passing along an excess $S\kappa$ bites.

7.2.4 Malaria Prevalence & Thresholds

We let x denote infection prevalence:

$$x = \frac{X}{H} \tag{7.25}$$

so $\kappa = cx$, and

$$\frac{dx}{dt} = \frac{1}{H}\frac{dX}{dt} = h(1 - x) - rx \tag{7.26}$$

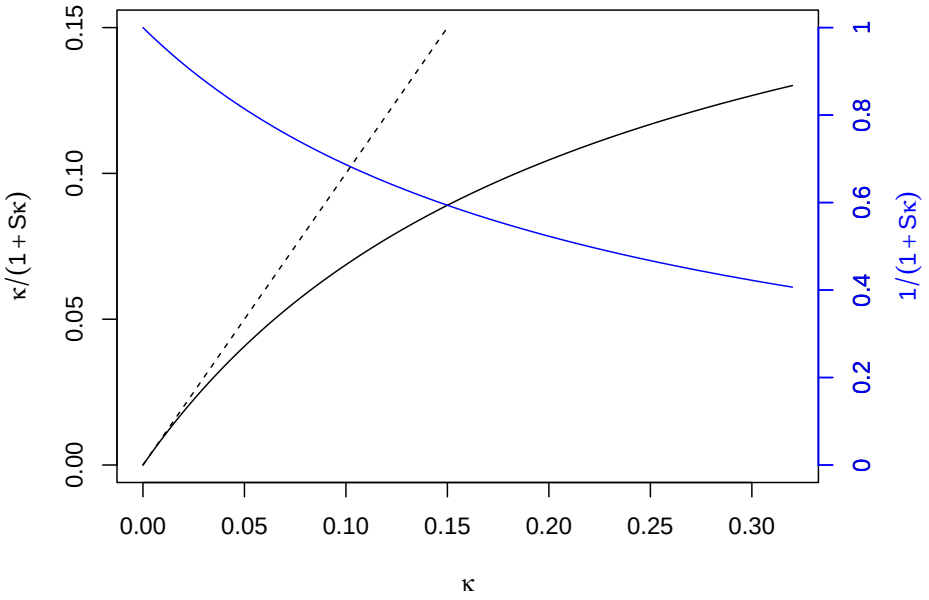


Figure 7.2: The effect (compare the solid and dashed black lines) and effect size of saturation (blue), graphically.

We can also define the basic reproductive number:

$$R_0 = \frac{bcV}{r}. \tag{7.27}$$

It is the product of four terms:

- Vectorial capacity, V, the number of infective bites, per person, per day;
- The number of days a person would remain infectious, $1/r$;
- The fraction of infectious bites that would infect a human, b;
- The fraction of blood meals on infectious humans that would infect a mosquito, c

After taking their product, we can interpret R_0 as a measure of the parasite's reproductive success after a single generation. It only depends on where we start counting. It could be one of the following:

- the number of infected mosquitoes that would arise from a single infected mosquito;
- the number of infectious mosquitoes that would arise from a single infectious mosquito;

- the number of infected and infectious humans that would arise from a single infected and infectious human.

Here, R_0 plays an important role in these equations if we start with dX/dt; then transform it to dx/dt; then replace h with bE; then replace κ with cx; then divide by r; and rearrange:

$$\frac{1}{r}\frac{dx}{dt} = x\left[R_0\left(\frac{1-x}{1+cSx}\right) - 1\right] \tag{7.28}$$

Since x is the prevalence, it is always in the interval $[0, 1]$. When x is very close to 0, then

$$\frac{1-x}{1+cSx} \lesssim 1. \tag{7.29}$$

and as x grows very small:

$$\lim_{x\to 0}\frac{1-x}{1+cSx} = 1. \tag{7.30}$$

It follows that when x is small, $dx/dt > 0$ if and only if $R_0 > 1$. Depending on R_0, only one of two possibilities can hold:

- either $R_0 < 1$, so that $x = 0$ is the steady state;
- or $R_0 > 1$, and the steady state is:

$$\bar{x} = \frac{R_0 - 1}{R_0 + cS} \tag{7.31}$$

Since at the steady state, $\kappa = c\bar{x}$, we can plug this back into the formulas above to get $\bar{Y}$ and $\bar{Z}$.

What we've learned about these equations is that if mosquito population densities are constant, then malaria reaches a steady state: if $R_0 > 1$, then there is a positive endemic equilibrium, and if $R_0 < 1$, then malaria is absent from the system. The system is said to be stable – in fact, is is globally asymptotically stable, which means that all the orbits end up converging to the steady state. This statement has been proved many times in many papers, and since this book is focused on policy, we'll let others worry about proofs.

7.2.5 Checking our Work

An advantage of working in this environment is that we can check our work. One way we could solve these equations would be to run them for a very long time:

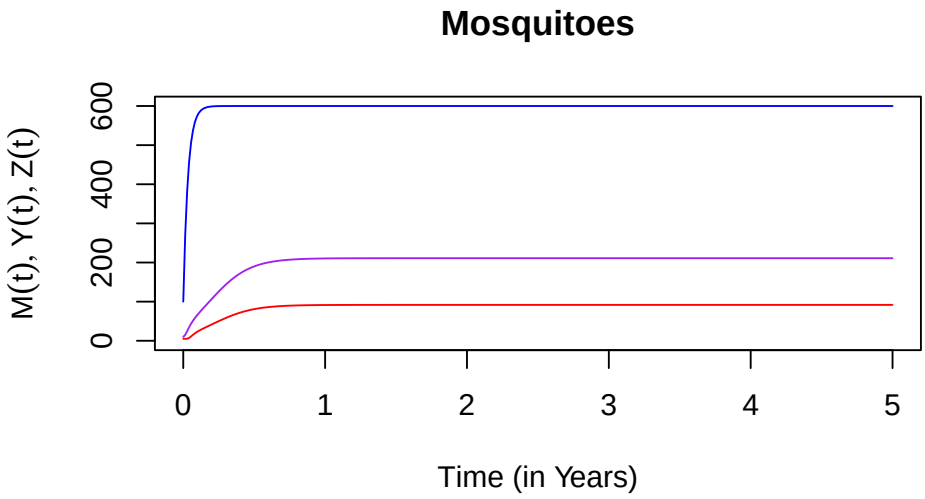
We make a parameter set that defines the model:

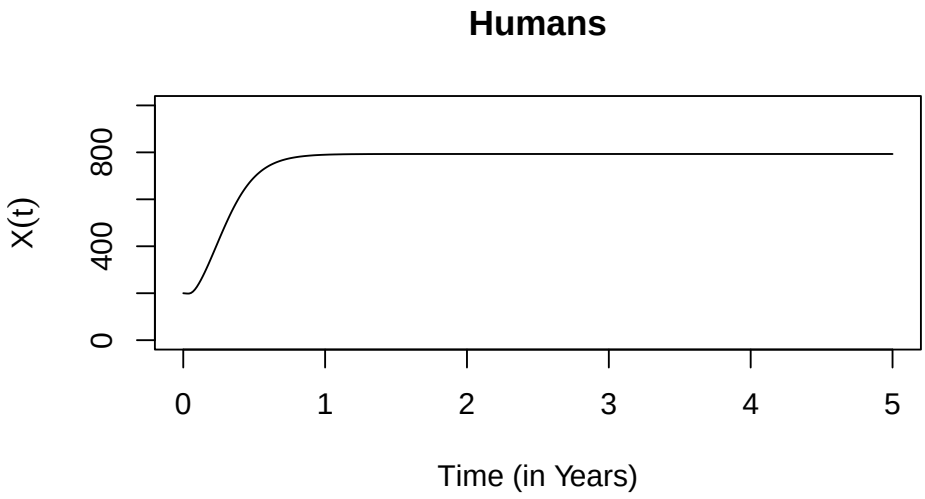
```
y0=c(M=100, Y=10, Z=5, X=200)
paramsSteady = makeParams(y0, ss=0)
```

```
dede(y0, times=tt, func=dAronMay, parms=paramsSteady) -> yout
tail(yout, 1)[-1] -> eq1
eq1
```

```
## [1] 600.00000 211.01706  91.70764 793.10541
```

By plotting it out, we can check to see if we've run it for long enough:





We can check our algebra by computing the same quantities, and R_0 and other quantities we care about:

```
steadyStates_AronMay = function(params){with(params,{
  Lambda = m*H
  Meq = Lambda/g
  S = with(paramsSteady, f*q/g)
  V = m*S^2*exp(-g*tau)
  R0 = b*c*V/r
  x = ifelse(R0>1,(R0-1)/(R0+c*S), 0)
  Xeq = x*H
  kappa = c*x
  Yeq = S*kappa/(1+S*kappa)*Meq
  Zeq = Yeq*exp(-g*tau)
  EIR = f*q*Zeq/H
  FoI = b*EIR
  aEIR = 365*EIR
  aFoI = 365*FoI
  extra=c(S=S, V=V, R0=R0, x=x, kappa=kappa,
          EIR=EIR, FoI=FoI)
  annual =c(aEIR = aEIR, aFoI=aFoI)
  list(std=c(M=Meq, Y=Yeq, Z=Zeq, X=Xeq),
       extra=signif(extra, 3),
       annual = signif(annual,3))
})}
steadyStates_AronMay(paramsSteady) -> eq2
eq2
```

```
## $std
##         M         Y         Z         X
## 600.00000 211.01706  91.70764 793.10541
##
## $extra
##      S      V     R0      x  kappa    EIR    FoI
## 4.5600 0.4520 7.4600 0.7930 0.1190 0.0348 0.0192
##
## $annual
## aEIR aFoI
## 12.7  7.0
```

Now, we can compare directly:

```
rbind(eq1=eq1, eq2=eq2$std)
```

```
##       M        Y        Z        X
## eq1 600 211.0171 91.70764 793.1054
## eq2 600 211.0171 91.70764 793.1054
```

7.3 Stable Orbits

If emergence rates vary seasonally, how much of the analysis that we did to understand *steady states* still holds? Obviously, if conditions are changing seasonally, the model does not reach a steady state. In fact, after modification to suit the context, many of the same principles translate. The steady state analysis provides a good qualitative guide, but that the answers will look different. Here, we illustrate by solving systems to illustrate some basic points, which is easy enough. Analysis of the resulting dynamics can be quite difficult; it is covered in Temporal Dynamics.

7.3.1 Thresholds

There is a threshold condition $R_0 > 1$ that determines whether malaria is endemic, but the formula for R_0 depends on the form of $\Lambda(t)$. If we set $R_0 = 1$, we can show that the threshold for persistence in a seasonal environment is $R_0 > \sigma > 1$ (see Figure 1.1). The math to compute threshold conditions in seasonal environments is in Temporal Dynamics.

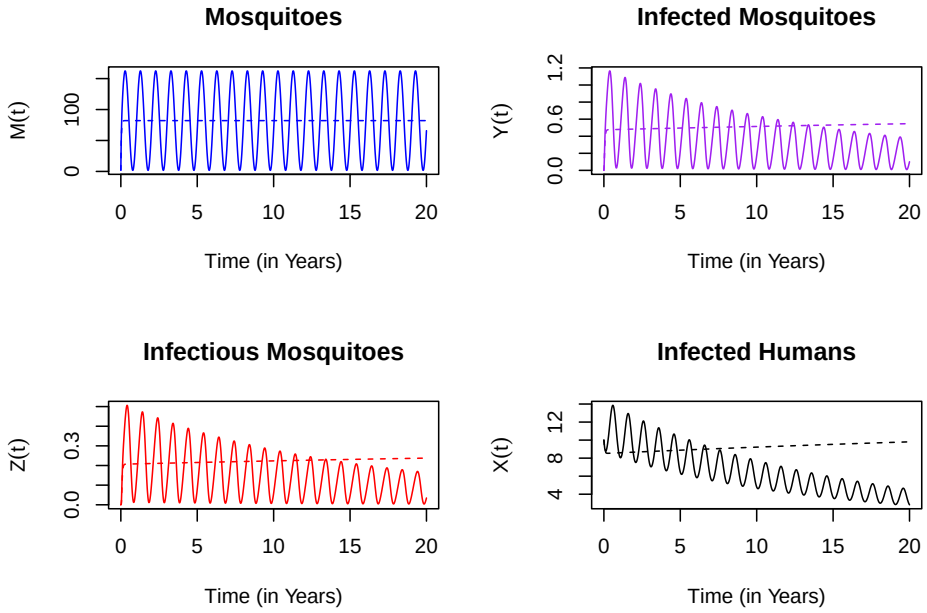


Figure 7.3: Here, we set $R_0 = 1.02$ for the model with constant emergence, and we show that malaria persists. For the same parameters and for the same *average* emergence rate, malaria declines with seasonality.

7.3.2 Orbits

If $R_0 > 1$, then all orbits converge to a set of *stable orbits* (See Figure 1.1). If $\Lambda(t)$ has an annual cycle, then after the orbits converge:

- $M(t+365) = M(t)$;
- $Y(t+365) = Y(t)$ and $Z(t+365) = Z(t)$;
- $X(t+365) = X(t)$.

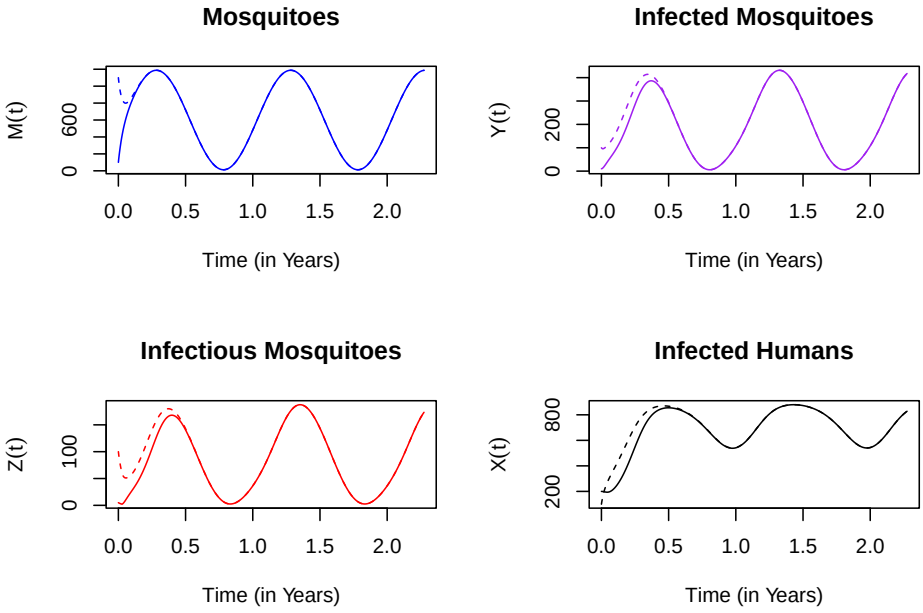


Figure 7.4: With different initial values, the orbits converge and eventually lie on top of one another.

7.3.3 Average Dynamics

If $R_0 > 1$ and malaria is endemic, the *average* prevalence of malaria infection is variable in a seasonal environment. While the prevalence is higher at the peak, the average for the whole year tends to be lower.

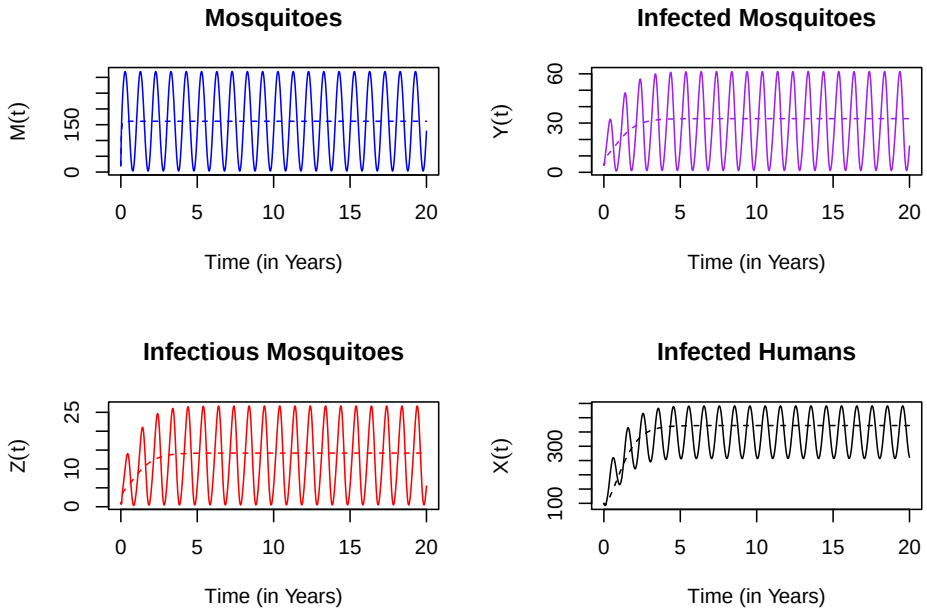


Figure 7.5: Here, we set $R_0 = 2$ for the model with constant emergence, and we show that the prevalence of malaria is similar in the seasonal environment, but it's higher as transmission peaks, lower in the off-season, and lower overall.

Chapter 8

Blood Feeding

8.1

Chapter 9

Mosquito Ecology

9.1 Egg Laying

Chapter 10

Human Age & Demography

A core limitation of most malaria models is that they need to have the capability of handling human demography and age. Here, we present the Ross-Macdonald model for exposure and infection in a population with age.

Human Demography: The problem, in a nutshell, is that our human population is changing dynamically as it ages over time. We let $H(a,t)$ denote the density of a human population of age a at time t, and we can describe age-specific survival as an age- and time-specific rate $\mu(a,t)$. Demographic changes in cohorts are described as follows:

$$\frac{\partial H(a,t)}{da} + \frac{\partial H(a,t)}{dt} = -\mu(a,t)H(a,t)$$

In this case, births are specified as a boundary condtion $H(0,t) = B(t)$.

Malaria Dynamics: To model malaria infection dynamics in this population, we let $h(a,t)$ denote the force of infection for the cohort of age a at time t:

$$\frac{\partial X(a,t)}{da} + \frac{\partial X(a,t)}{dt} = h(a,t)\left(H(a,t) - X(a,t)\right) - rX(a,t)$$

The problem for malaria in humans is exacerbated by the development of immunity in populations. We would like to have some methods that make it possible to investigate data describing malaria in populations as they age.

10.1 Cohort Dynamics

We now

10.2 Gallerkin

We now

Chapter 11

Realism & Complexity

There are several reasons why models for policy will tend to be more complex than models developed for scientific research. In science, the questions are narrow and the study designs are focused enough to form tests of ideas that have clear outcomes. When models are used to guide policy, the same level of scientific rigor should be applied, but weighing all the evidence requires a broder synthesis.
To build models that can guide malaria policies, the models must be *realistic* enough to be compelling, and they *ought* to reflect the knowledge and experience accumulated over years of studying and controlling malaria. Policy advice should be checked for consistency across studies. The models must be complex enough to serve many purposes all at once. To weigh tradeoffs, policy requires broad, synthetic models that allow for comparisons across subject matter domains.

To carry a conversation forward, the models used to guide discussions will need to retain a memory of what has been learned already, so they will tend to add features and grow more complex. Given the uncertainty, policy should be based on model swarms that propagate the uncertainty. The predictions of those models must be specific enough to be proven wrong, so that over time *some* of the models can be trusted over others. The same models can be used to identify which missing data would have the greatest impact on a policy, and ideally, studies can be conducted to gather this data. Over time, the advice should shift from *generic* advice to *specific* advice as more evidence is gathered. This is, in a nutshell, how daptive management works.

It might take a lot of work to build a model that has been fit to all the evidence describing malaria in a management unit over the recent past, and it might cut against the instincts we have as scientists to add all that realism, but it's worth it to make the effort if it helps communicate with malaria managers.

In designing a software solution to the problem of building realistic models, we

designed a framework for building models and a toolbox to build model swarms that would address the concerns of malaria programs. In the chapters that follow, we'll show the features of this framework by constructing examples. Even if we're principled about adding complexity, a cost of doing so is *computational complexity.* That is something the software was designed to manage. For the moment, we thus want to set aside concerns about *realism* vs. *abstraction*, about *parsimony*, and about *error propagation,* and we want to simply ask the question of how to build models with the features we want.

This chapter is an overview of the historical development of malaria models and an introduction to the toolbox. We'll cover the same material in much greater detail in the chapters that follow, and we'll construct examples using `exDE` and `MicroMoB`.

11.1 Malaria Epidemiology

11.2 Transmission

The second topic we must tackle is blood feeding, which is an interaction between mosquitoes and humans. It is an asymmetric relationship – mosquitoes search for blood hosts, select a host, and blood feed. Humans, for their part, attract mosquitoes from a distance, move around, and spend time in places when mosquitoes are biting. Humans can wear protective clothing (or not), use bed nets (or not), or do other things that make them more or less available to humans. Despite all this, humans are often unaware that they have been bitten.

Transmission occurs during blood feeding, and models of blood feeding *should* be able to take all this heterogeneity into account. If the models do a proper accounting, then the total number of human blood meals taken by mosquitoes would equal the number of blood meals received by humans. In doing so, we find no inspiration from Macdonald, whose description of human blood feeding was simple and phenomenological: a single parameter described the human blood feeding rates. After Garrett-Jones described the human blood index, drawing on decades of work, the one parameter was split into an overall blood feeding rate (f) and a human fraction (q). The question left unaddressed by Macdonald was how these rates vary by context, and the consequences for exposure. To do this, we reformulated the algorithm describing blood feeding [78].

Over the past two decades, several papers have drawn attention to the way blood feeding behaviors are or ought to be constrained by the availability of vertebrate hosts. It may be fine to assume that the density of vertebrate hosts doesn't change, but *something* should change when a large fraction of people are using bednets. Even with static parameters, we should think through the limiting cases: if there are no vertebrate hosts, then there blood feeding should not occur (*i.e.*, $f = 0$); if there are no human hosts, then there should be no human blood

meals ($q = 0$); and if there are no alternatives to humans, all blood meals should be on humans ($q = 1$).

The concepts we devised for blood feeding must, therefore, integrate the notion of frailty with the process of mosquito search. On the one hand, the mosquitoes should blood feed at a slower rate if hosts are unavailable. On the other hand, human biting should become heterogeneous. To arrive at an adequate description, we need to formalize this notion of host availability. The logic is that mosquitoes *search* for humans. Differences among humans in their attractiveness are represented by a *search weight.* Mosquito search in a place depends on the amount of time spent by humans, but also by daily mosquito activity patterns; from these, we develop a notion of *time at risk* that characterize the way human activities expose them to mosquitoes. The mosquitoes add up all the time at risk spent by all the humans, which gives a measure of their *availability.* Availability describes humans as well as other vertebrate hosts, which are modified by mosquito preferences. The overall feeding rates and the human fraction are computed from availability using *functional responses.*

To complete the picture, we consider how the expected rate of exposure could have a distribution in the population, which we call environmental heterogeneity.

11.3 Mosquito Ecology

11.3.1 Resource Availability

11.3.2 Egg Laying

11.3.3 Regulation

11.3.4 Exogenous Forcing

11.4 Heterogeneity & Stratification

Human populations are heterogeneous. Some kinds of heterogeneity affect how we understand malaria and what we should do, including who to target. To deal with heterogeneity in models, we will often need to *segment* a human population into sub-populations, or *strata.* When we talk about *stratification,* we mean it the narrow sense of segmenting a human population (*i.e.* not subdividing landscapes spatially[1]), because the model predictions made by creating strata that are more homogeneous should be more accurate. The guiding principle is

[1] In a broader sense, stratification is also about subdividing landscapes into a set of spatial domains that share relevant features in order to *tailor interventions to context.* That is a topic we take up in a separate book, (**Robust Analytics for Malaria Policy.**).

that our analytics will should strive to be more accurate, and that we should thus identify and remove those sources of heterogeneity that would affect policy advice, whether it affects estimating the impact of interventions in the past or projecting those impacts into the future. We acknowledge that models are approximations, and that our approximations don't have to be perfect. The goal is to find ways of propagating uncertainty that are *good enough* for the task at hand.

In malaria epidemiology, *some* kinds of endogenous heterogeneity *could* be built into the *epidemiological state space*. Other kinds of heterogeneity, including consistent differences in exposure, differences in care seeking and drug taking, and differences created by malaria control (*e.g.* net ownership or vaccination), usually require stratification. The decision about how to strike the right balance depends on the model and the purpose of a study.

The framework and supporting software offer a toolbox for stratification. It is designed to stratify populations in a principled way, so that we can *understand* how the heterogeneity affects transmission or outcomes that we care about, but we can also *combine* effects. We want to stratify populations by applying rules that *split* populations when the differences are large enough. (If we started with complex models, we might choose to *join* populations if the differences were small.) By so doing, we can *compare* the behaviors of models that differ from each other in only one way. If the differences are not too large, or if the differences in dynamical behaviors we care about are not too large, we might decide not to split the strata, and use the average. Because of the way models are encoded, it's easy to build models that split the strata in multiple, independent ways.

11.4.1 Strata in the Ross model

As a simple example, consider a simple Ross-style model for infection with exposure and recovery (described in Section 7.1.5):

$$\frac{dX}{dt} = h(H - X) - rX$$

If exposure is heterogeneous, we could split this population into two strata and add subscripts (*i.e.*, indexed by $i \in \{1, 2, \ldots\}$):

$$\frac{dX_i}{dt} = k_i h(H_i - X_i) - rX_i$$

We hold the *average* FoI constant by constraining the values of k_i:

$$\frac{\sum_i k_i H_i}{H} = 1$$

Stratification is important if the differences are large. With two strata, it would not make sense to stratify if $k_1 \approx k_2$, but if $k_2 \gg k_1$ then it might change our expectations, or it might change what we recommend.

11.4.2 Frailties

We will introduce segmentation first through models of [Heterogeneous Exposure] to malaria, where we consider various sources of *frailty* – proportional differences in the average hazard rate for infection (k_i, in the example above). These differences in exposure can arise because of age, house type, risky behaviors, other factors. Frailty that is attributable to location (*e.g.* proximity of home to aquatic habitats) can be dealt with by sub-dividing space into *patches,* a topic that is taken up in [Space] below and Spatial Dynamics. Depending on the size of the patches, some differences in average rates of exposure due to location can persist, and these could be dealt with by generic stratification into high *vs.* low exposure strata.

Some of the heterogeneous traits that we care about change dynamically, so we will also need to consider population *flows* among strata, which change the sizes of the strata. We would like to deal with these flows in a principled way. Bed net ownership and use are among the human behaviors that matters most for programs. In some cases, we will want to understand dynamic changes in bed net ownership, the patterns of use among those who own a net, personal protection, and community effects. Later, we show how to construct an example that *describes* all of these aspects of bednets.

Segmentation is what we need to build models of pharmaceutical interventions with waning effectiveness, such as mass vaccination. Among the most important factors in malaria is age. We have defined algorithms to model [Aging] and other demographic change, the loss of bednets, waning protection or changing housing quality.

11.4.3 Age

Immunity to malaria develops with age and exposure. The development of immunity is probably changing throughout life, so it makes sense to think of malaria epidemiology as ontogeny.

For systems described generically by the state space, $\mathscr{X}$, the dynamics we care about have the form:

$$\frac{\partial \mathscr{X}(a,t)}{\partial a} + \frac{\partial \mathscr{X}(a,t)}{\partial t}$$

We might want to deal with malaria differently if we are studying malaria in

cohorts. In a population where the FoI over time is $h(t)$, we might want to follow a birth cohort, so we define $h_d(a) = h(t - a)$ for all $t > d$. We can then solve:

$$\frac{d\mathscr{X}}{da}$$

which produces states in cohorts as they age, $\mathscr{X}(a|h)$.

When we simulate malaria transmission dynamics in populations for policy, we will want to put a mesh on age and segment the population. The dynamics are define for age strata, where the FoI is defined differently for each age stratum:

$$\frac{d\mathscr{X}_a}{dt}$$

which produces age-dependent states over time, $\mathscr{X}_a(t|h)$.

Our algorithms should guarantee that the epidemiological states over time provide an accurate match for the epidemiological states over age.

11.5 Spatial Dynamics

11.5.1 Human Migration, Mobility & Travel

The notion of a spatially distributed risk for humans and the modalities of human travel.

- Humans move around, so we develop a model of *time spent*. Time spent is sub-divided into three parts:
  - time spent at home;
  - time spent traveling, when a night is spent away from home;
  - human mobility, which describes time spent around home when not traveling.
- For travel, we estimate a travel FoI.
- For time at home and mobility, after weighing time spent and mosquito diurnal activity patterns by time of day, we modify time spent to get a notion of *time at risk*
- After modifying time at risk by search weights, mosquito blood meals are distributed among all hosts according to their availability.

11.5.2 Mosquito Search & Dispersal

11.5.3 The Mixing Matrix

11.5.4 Pathogen Dispersal by Humans

11.5.5 Pathogen Dispersal by Mosquitoes

To describe mosquito spatial dynamics, we

11.6 Exogenous Forcing

11.6.1 Weather

11.6.2 Habitat Dynamics

Space is big, so we start by drawing boundaries around a part of the world we want to study, that we call the *spatial domain.*

11.6.3 The EIP

11.6.4 Seasonality

11.7 Integrated Vector Control

11.8 Integrated Malaria Control

11.9 Context

Chapter 12

Modularity and Software

Hundreds of publications have described new models of malaria [45, 73]. The challenge we have taken on is to find a new way of building models for malaria that draws from all those good ideas to build models at any level of complexity. We want to do this with reusable, professional quality software. Ideally, the models that we develop would be sufficiently complex to address policy questions, yet remain amenable to analysis. To get there, we took a step back to try and understand *malaria models*, and to put this into a birds-eye view of the process of model building.

From Ross's first published model in 1905 to the first draft of this book, 117 years have passed. The story of malaria models can be summarized in three epochs.

Ross's models, and contributions to mathematical study of malaria made by Alfred J Lotka (1912-1923), George Macdonald (1950-1968), and Garrett-Jones (1964-1970) take us to the end first epoch, which is marked by the end of the Global Malaria Eradication Programme (GMEP, 1955-1969). As part of the GMEP, Macdonald's formulas were extended by Garrett-Jones into the concept of *vectorial capacity* and a rudimentary theory of vector control. By 1970, the *Ross-Macdonald* model was more than just a set of equations. It was a theory for malaria dynamics and control supported by a well-developed set of concepts, parameters and metrics [72].

Over that same period of time, mathematical theory for directly transmitted diseases took a parallel path, with important mathematical contributions from Kermack and McKendrick, NTJ Bailey, and Bartlett. Sometime around 1980, mathematical epidemiology began a period of innovation and synthesis, particularly after the publications of Robert May and Roy Anderson made it a mainstream activity in departments of ecology.

In malaria and mosquito-borne diseases, Klaus Dietz publications span the second epoch (1971-2006), including development of a mathematical model with immunity for the Garki Project [16], work on the dynamics of malaria under treatment by drugs [17], seasonality [14], and heterogeneous biting [18, 15]. During this time, theory developed for malaria borrowed concepts and methods. In spatial dynamics, the patch models of Yorke and ** were modified to by Dye and Hasibeder to describe mosquito-borne pathogens [19, 26].

The last epoch of malaria, which starts around 2006, is marked by two major developments: a maturing theory of malaria control; and the rise of branded, individual-based models.

The publication of *OpenMalaria* in 2006 marks the beginning of the last epoch of malaria. Some important antecedents were Dana Fochs models for *Aedes* dynamics and dengue virus transmission, as *CIMSiM* and *DENSiM*. In malaria, several within-host models had been developed [42, 43]. *OpenMalaria* traces its history back to an intrahost model developed by Dietz and Louis Molineaux [43]. After *OpenMalaria,* two other branded individual-based models were developed. One was developed by a team at Imperial College called *Malaria Tools.* Another was developed by a team at the Institute for Disease Modeling called *eMod.* The fact that the models were named and branded was significant – the authors had developed software that they would maintain and that they were willing to stand behind. The models had finally dealt with *disease* in a serious way, and through publications, the fitted models demonstrated a fidelity to evidence. The branding signaled continuity and consistency.

Around 2007, new models of vector control began to appear that related inteven-tion coverage levels to effect sizes. Macdonald's work had focused on sensitivity to parameters, and the GMEP emphasized technical efficiency to achieve very high coverage (with IRS). Garrett-Jones developed vectorial capacity as a way of understanding vector control and effect modification by insecticide resistance. The new models extended Garrett-Jones ideas. The need for new models was motivated, at least in part, by the goal of achieving universal coverage with ITNs. What were reasonable coverage targets? The new generation of vector control models introduced the concept of an effect size on transmission as a function of intervention coverage levels, where coverage had one definition for operations (*e.g.* something like ownership) and another for effect sizes (*e.g.* related to vector contact rates with interventions). The goal of achieving very high coverage with ITNs bumped into the reality that nets are not durable, so new models have been devised to look at intervention coverage in relation to distribution schemes and product durability. While these concepts had been considered during the GMEP design phase, they did not appear in Macdonald's models.

If we want to take advantage of all the research that has been done, we need a way of understanding malaria models and the whole business of model building.

12.1 Model Building

Model building is a fairly involved process that includes several unavoidable steps:

- There must be some motivation for building a model, which usually starts with a conversation, boxes and arrows drawn on paper or a chalkboard or whiteboard. The process involves refining the questions, until there's a well-formed idea – a reason for building a model.
- The idea gets translated into mathematics. The boxes get translated into variables, the arrows are rate parameters, a mathematical formalism is selected.
- The model gets analyzed. In some cases, when the model complexity exceeds a very low threshold on complexity, this is done with pencil and paper. It is only possible to analyze individual parts of the model this way.
- The model gets translated into pseudo-code, and then it gets implemented as software that can produce output. This is followed by a long and painful process of verifying that the software does what the pseudo-code says it *should* do. After awhile, the software is trusted, and it's time to use it.
- Some thought is given to the correspondence between the variables in a model, observable quantities, and the observational process itself. This process can be a part of what happens above, but at some point, the models need to be fitted to data.
- The software produces output and then: the outputs are visualized; models are fitted to data; graphs are made; papers or reports are published; and careers advance.

That's the simple story of model development. What happens next is could be one of the following:

- Someone re-examines an existing model and notices it is inadequate in some way: it is missing some features, or it might make an assumption that ought to be modified. Simple models become spatial models, single populations are structured.
- Someone decides to implement the model in a different way, perhaps with a different mathematical formalism. Continuous time models are translated into discrete time models. Deterministic models become stochastic. Autonomous processes become non-autonomous.

Through this process, hundreds of malaria models were published.

A problem with this process has been that the software is often developed for bespoke tasks (*i.e.* to publish a paper). The software is often poorly documented and difficult to reuse. The costs of building a model for one task limited the complexity of the model. It was difficult to combine elements of one model

developed for one purpose, with someone else's model developed for another purpose.

In malaria, this *ad hoc* process of writing new models was found to be inadequate to serve the broad range of policy questions. One way of dealing with the complexity was to build individual-based models, but individual-based models have some of the same limitations as reality.

12.2 Modular Computation

Before *OpenMalaria*, most models of malaria modified the Ross-Macdonald model in one way [45]. The innovation was focused on specific themes or questions: how long would an infection last in models with superinfection?

12.2.1 `exDE`

We have written the software that solves these equations in a package called `exDE`.

Part III

Epidemiology

Chapter 13

Malaria Epidemiology

13.1 Parasite

13.2 Epidemiology

A major challenge for malaria dynamics is how to define an state space describing malaria infection and immunity in human populations that captures the essential elements of malaria dynamics well enough to trust for making policies. There are features of malaria infections that have been identified and studied in the past: superinfection; the complex time course of an infection – including fluctuating parasite densities – and the problem of detection; gametogenesis, gametocyte maturation and gametocyte dynamics; fever and disease; development of immunity with exposure including its effects on infection, disease, and infectiousness; treatment, adherence to drug regimens, chemoprotection and infection curing. Over time, these issues have been addressed in various models. We need model that is good enough for policy, but this also means developing a commmon understanding of malaria that can serve as a basis for discussion.

To get to that point, we must start simple and add complexity.

The model for malaria infection that we presented in [Malaria Dynamics] was developed by Ross. In today's vernacular, it would be called an SIS compartment model. The model is very simple, and it is probably inadequate for every task, but it is useful and it has been used. The model assumes that malaria infections clear at a constant rate regardless of the age of infection or other factors. The persistence of malaria infections over decades tells us that this assumption is clearly false, but it is good enough for some programmatic needs. During the GMEP, Ross's model was used to characterized the response timelines for the *Pf*PR after the interruption of transmission. Drawing on multiple sources, Macdonald estimated that the duration of infection was around 200 days, which

was good enough to use as a basis for monitoring and evaluating the interruption of transmission [39]. The simpler model was used even though Macdonald had already proposed an alternative model that considered superinfection [37]. Despite the simplicity, the model was adequate to the task [67]. An important lesson is that the simplicity has some advantages, and the models that get used in policy tend to be very simple.

The question is how to develop models that are simple and yet are up to policy tasks, which means that the models must (at some point) get validated against research data. Doing so means having sufficient complexity to deal with exposure, infection, detection, immunity, disease, infectiousness, care seeking, and drug taking. Whatever model is selected as a basis for policy, it should be simple enough to understand and yet complex enough to capture the *gist* of malaria epidemiology. The models, however chosen, must get it right. Sorting through all the complexity to get a model that is good enough is a daunting task. This introduction is mainly historical, but we use it to preview some of the themes. In the following history, we discuss some of the important innovations.

13.3 Infection

13.3.1 Superinfection

From early on in malaria epidemiology it was clear that exposure to malaria differed among populations, and that in some places, the rate of exposure was far higher than the rate of clearance. Ross emphasized a need to measure exposure both entomologically, through metrics that are known today as the EIR and the FoI, and parasitologically, through the prevalence of infection by light microscopy (or more commonly today, through RDTs), which was called the malaria *parasite rate* . There was no good reason to believe that people in highly malarious areas would be exposed faster than they would clear infections, so they would carry infections that could be traced back to many infectious mosquitoes [76]. This phenomenon was called superinfection.

Macdonald was the first to develop a mechanistic model of superinfection [37], but the mathematical formulation was at odds with his description [21]. It is an interesting bit of history for a different time.

A mathematical basis for understanding superinfection was worked out as a problem in the study of stochastic processes as part of *queueing theory.* This may seem strange, but understanding how many people are queueing involves understanding how people come in and how fast they are processed. One of these queuing models has become a mainstay of malaria epidemiology; in queuing theory, it is called $M/M/\infty$.

The model tracks the **multiplicity of infection** (MoI). It assumes that infections arrive through exposure at a rate h (the FoI), and that they clear

independently. Without clearance, the MoI, denoted ζ, would just go up. The model assumes that each infection clears at the rate r; if the MoI were 3 then infections would clear at the rate $3r$. Regardless of how fast infections arrive, the fact that the pressure for the MOI to go down increases with MoI means that the MoI will reach a stable state. The mean MoI is h/r. The following diagram illustrates and provides the equations:

$$\zeta_0 \underset{r}{\overset{h}{\rightleftarrows}} \zeta_1 \underset{2r}{\overset{h}{\rightleftarrows}} \zeta_2 \underset{3r}{\overset{h}{\rightleftarrows}} \zeta_3 \underset{4r}{\overset{h}{\rightleftarrows}} \cdots$$

$$\begin{aligned} d\zeta_0/dt &= -h\zeta_0 + r\zeta_1 \\ d\zeta_i/dt &= -(h+r)\zeta_i + h\zeta_{i-1} + r(i+1)\zeta_{i+1} \end{aligned}$$

If one is willing to abandon compartment models, then it is possible to formulate more elegant solution using hybrid models. The mean MoI, m changes according to the equation:

$$\frac{dm}{dt} = h - rm.$$

Using queuing models, it is easy to show that the distribution of the MoI is Poisson, and in these hybrid models, if the initial distribution is not Poisson, then it will converge to the Poisson distribution asymptotically. The complex dynamics of superinfection can thus be reduced to this simple equation.

Unfortunately, things become more complex if we add simple features such as treatment with drugs, or heterogeneous exposure. The distribution of the MoI is no longer Poisson [28]. Superinfection is an important part of malaria epidemiology, and we will use these models for superinfection in developing some adequate models for infection and immunity.

In the Garki Model (see below), the waiting time to clear an infection used these queuing models to formulate an approximate clearance rate:

$$\frac{h}{e^{h/r} - 1}$$

13.4 Disease & Immunity

The biggest failing of Ross's model, perhaps, was that it did not make any attempt to grapple with acquired immunity to malaria. It had always been clear that immunity to malaria was important because the prevalence of infection declined throughout adolescence and was consistently lower in adults, and because disease and severe disease were common in young children. The data accumulated through years of studying malaria, done as part of malaria therapy, provided supporting evidence for immunity. There was a difference in outcomes from being exposed to the same parasite (homologous challenge) compared with a

different parasite (a heterologous challenge). Immunity had something to do with the number of different parasites that a person had seen.

The first model to grapple with immunity was the Garki Model [16]. The main idea in the Garki Model was that it would be possible to understand malaria dynamics by expanding the number of compartments: the population was sub-divided into two non-immmune or semi-immune. Infection dynamics were tracked separately within each immune category: the infections would clear faster from semi-immune individuals, they were are not infectious, and they are less likely to test positive if they were infectious. Some features of the Garki model seem odd in retrospect: there were two infected states for non-immunes (y_1 and y_2), but only one for semi-immunes; there was no way to lose immunity; and the assumption that semi-immunes are not infectious.

The Garki Model has had a poweful influence on malaria modeling. Several models since then have expanded on various themes. Several compartment models have been developed that replicate infection dynamics across immune stages: we call this *stage-structured immunity.*

In the Garki Model, we can simulate the immuno-epidemiology of cohorts as they age. Eventually, the cohort *would* settle to an equilibrium. At that point, everyone is semi-immune, a sizable fraction remains non-immune after a century. By the time the cohort reaches the steady state, everyone in the cohort has died. If we focus on the dynamics in the first two decades of life, prevalence rises as people become infected, and then it falls as people become semi-immune. The changing epidemiology as cohorts age is an important feature of malaria. In models like this, the concept of a *steady state* teaches us something, but the models draw attention to the sharp changes in malaria that occur throught the first 20 years of life. We can adapt the idea of steady state to suit our needs – under constant exposure, cohorts trace out *stable orbits* as they age. These stable orbits are a basis for understanding malaria dynamics *vs.* age.

One application of these stable orbits is to understand the the relationship between age and infection prevalence as a function of exposure. Curiously, the Garki Model captures the basic shape of age-*Pf*PR curves, but it does not get the details right. When we start to look at the factors affecting the *Pf*PR by age in populations, we must acknowledge the need to add other features: drug taking and chemoprotection; differences in exposure that arise for a number of reasons; anemia, perhaps; seasonality. Not everything is about immuno-epidemiology.

13.5 Gametocytes and Infectiousness

In Ross's models, everyone who is infected is also infectious. This is clearly wrong, but it may not be a terrible assumption under most circumstances. The Garki Model made the extreme assumption that semi-immune individuals are not infectious at all. There is now copious evidence that adults *do* transmit parasites

to mosquitoes, but they are not as infectious. This decline in infectiousness occurs for two reasons: first, the densities of asexual-stage parasites in adults are controlled by immunity, so they are lower. Since a fraction of asexual parasites becomes gametocytes, the densities of gametocytes are also lower in adults. Second, gametocyte densities are modulated by an immune response that affects malaria parasites in mosquitoes, which is called *gametocyte-stage transmission blocking* immunity. The dynamics of gametocyte-stage immunity change with age and exposure, and we will need to understand how this form of immunity waxes and wanes.

There are some other important details about malaria infections that might be relevant in some contexts. First, *P. falciparum* gametocytes take 8-12 days to mature. When combined with the 6 days in the liver, we must acknowledge that the latent period is at least 2 weeks. Because gametocytes must reach densities high enough to be transmitted, the effective latent period for humans is probably closer to 20 days. Since the parasites also need 10 days or more to mature in mosquitoes, the shortest parasite generations are probably at least a month long.

Another feature of gametocytes that matters is that gametocyte populations are not always affected by anti-malarial drugs, so after taking drugs that clear all the asexual-stages, some people will remain infectious to mosquitoes for quite a while after being treated with some drugs.

Ross's assumption may serve most needs, but the models must be good enough to guide policies, such as MDA or malaria elimination, when details about gametocytes and infectiousness can affect the outcomes of policies.

13.6 Treatment and Chemoprotection

It is impossible to understand malaria infection dynamics without accounting for treatment with anti-malarial drugs and a brief period of chemo-protection that follows. The first model for drug treatment was developed to understand MDA [17]. In developing models for policy, we must be careful about drug taking and its effects because it modifies the relationship between exposure (the EIR) and infection.

13.6.1 The Time Course of an Infection

The time course of infections is complex, and we will need to develop some models that relate parasite densities. In the chapters that follow, we introduce two main kinds of models:

- AoI
- SoI

13.6.2 Intrahost Models

There are two kinds of models we will discuss, but we would like to avoid them in making policy if possible.

- In host models;
- Individual-based models.

13.6.3 Synthesis

In the end, we do not need perfect models of malaria infection and immunity, but we do need a sound understanding of several things to make policy:

- The prevalence of infection by age as a function of exposure and drug-taking;
  - . . . in a cross-section of the population;
  - . . . in the care-seeking population.
- The incidence of malaria by severity and by age;
- The fraction of malaria that is promptly treated by severity and by age;
- The net infectiousness of a population of humans to mosquitoes.

In the chapters that follow, we will develop some models that based on a new concept – the age of the youngest infection – that combine many of the ideas in the chapters above.

Chapter 14

Infection

In the following sections, we walk through several models for the dynamics of malaria infection and immunity in humans. We cover infection and detection, immunity, infectiousness, disease, drug taking, and cohort dynamics.

14.1 Life Stages

14.1.1 Sporozoites

14.1.2 Liver Stages

14.1.3 Asexual Stages

14.1.4 Gametocytes

14.2 Controlled Infections

14.2.1 Malariotherapy

14.2.2 CHMI

14.3 A Simple Infection

14.3.1 Time Course

14.3.2 Duration

14.4 Superinfection

14.5 Process *vs.* Observation

14.5.1 Detection

The state space modeling approach.

Chapter 15

Superinfection

15.1 Parasite Densities and Detection

15.2 Light Microscopy

15.3 Biomarkers and RDTs

15.4 PCR

Chapter 16

Age, Exposure and Malaria Immunity

Immunity to malaria has been described as being *weak* and immunity has a *poor memory*. In quantitative terms, this means that almost every aspect of infection is affected by the history of exposure, and exposure accumulates with age.

In developing our models so far, we have ignored age as a factor in malaria epidemiology. In analyzing the models, we looked for steady states or stable temporal orbits. These analyses give a misleading impression of malaria.

The first model to examine malaria immunity was developed for the Garki Project [16]. We will present the model, for historical reasons, and then we will

it is much easier to formulate models for malaria infection dynamics in cohorts as they age.

16.1 The Garki Model

16.2 Stage-Structured Immunity

16.3 Strain Specific Immunity

16.4 Memory Tracking

16.5 Age *vs.* Prevalence

Chapter 17

Fever and Severe Disease

17.1 Fever

17.2 Anemia

17.3 Severe Disease

Chapter 18

Drug Taking

18.1 Curing Infections

18.2 Chemoprotection

18.3 Adherance

18.4 Treatment Rates

Chapter 19

Gametocytes and Infectiousness

19.1 Gametocytemia

19.2 Anti-Gametocyte Immunity

Chapter 20

Care Seeking

Chapter 21

Modeling Malaria Epidemiology

21.1

Part IV

Transmission

Chapter 22

Blood Feeding

The endpoint

22.0.1 Search and Risk

22.0.2 Search Weights and Availability

To deal with heterogeneous exposure and many other phenomena, we need a sensible way of segmenting humans into population **strata**. Stratification makes it possible to deal with population heterogeneity.

A new model of **blood feeding** is based on a model of blood feeding as the endpoint of a search for a blood host [78].

- Each sub-population has a *search weight* (w), and the total *availability* of humans for blood feeding (W) is the sum of the sizes of the strata weighted by their search weights.
- We also consider the availability of alternative vertebrate species for blood feeding (O).

22.0.3 Functional Response

- Mosquito blood feeding rates are computed using a *functional response* to total availability of vertebrate hosts ($f = F_f(B)$).
- To compute total availability, we add a scaling parameter on alternative hosts, because mosquito preferences can translate into different patterns of search; total availability is $B = W + O^{\varsigma}$.

- The human fraction is proportional to the relative availability of hosts $q = W/B$.

22.0.4 Environmental Heterogeneity

- The *search weights* thus translate into a kind of **Frailty**, which is one component of *heterogeneous exposure.* Important sources of frailty include bednet use, housing type, and age.
- We also want to consider *variability* in exposure within a stratum – what is the distribution of the *expected* number of bites over time? We have already discussed frailties, so this is a different kind of heterogeneous exposure that we call **Environmental Heterogeneity**. This helps us to align models with data: mosquito counts data tend to be described well by *negative binomial* distributions, so it is likely that the distribution of infectious bites also follows a negative binomial distribution. We introduce a function that translate the EIR into the FoI:

$$h = F_h(E)$$

In the Ross-Macdonald model, the underlying assumption is consistent with a Poisson distribution, but we have also derived *negative binomial hazard rates.* Environmental heterogeneity can arise from two sources:

- the aggregated distributions of mosquitoes in micro-habitats, and the redistribution of mosquito populations by wind and weather;
- random movements of humans around mosquito micro-habitats that affect their risk in a way that doesn't tend to change the mean;

22.1 Host Availability

22.2 Blood Feeding Rates

22.3 The Human Fraction

22.4 The Mixing Matrix, β

Chapter 23

Human Mobility

23.1 Modalities

23.2 Time at Risk

23.3 Travel

23.4 Migration

Chapter 24

Spatial Dynamics

Chapter 25

Heterogeneous Transmission

For humans, exposure to malaria means exposure to the bites of infectious mosquitoes. A problem that we'll have to deal with sooner or later is that exposure risk differs among humans over space and time. While this might seem like an odd thing to introduce so early, we will have to tackle the topic sometime. The discussion of [Heterogeneous Biting], in the previous chapter, showed that heterogeneity plays an important in understanding transmission and thresholds. This discussion of heterogeneous exposure (*i.e.*, looking at heterogeneous biting from the human side) is a good way of introducing some of the core concepts that are built into the framework:

- [Heterogeneous Biting] is one way of getting around a conundrum. In models with homogeneous biting, the relationship between *average* mosquito density and the prevalence of infection would lead us to make quantitative predictions about the likely effects of vector control.
- We discuss two different kinds of heterogeneous exposure: frailty, and environmental heterogeneity. In a nutshell, frailty multiplies the mean hazard rate for a sub-population by some amount k. Environmental heterogeneity does not affect the mean, but it changes the distribution of the mean.
- We introduce the idea that we can deal with frailties in human populations by segmenting the population into strata.
- We set the stage for a new model of mosquito **blood feeding** that we introduce in the next chapter.
- In a chapter on [Approximation], we use these models to discuss the problem of model-based inference.

25.1 Overview

Some reasons heterogeneous exposure to malaria have been documented in hundreds of studies. This is an overview.

25.1.1 Age

- Port, Boreham
- Carnevale

25.1.2 Location

25.1.3 House Type

25.1.4 Activities

25.2 Frailty

In general, we define frailty as a multiplicative factor on the FoI. If the average FoI in the population is h, then the FoI in a stratum is hk. The size of the stratum, p_k, is constrained such that:

$$\int_0^\infty k\, p_k\, dk = 1$$

With this constraint, the mean FoI in the population is h.

Continuous distributions are difficult to extend, but we can stratify a population to accomplish some of the same effects.

25.3 Environmental Heterogeneity

Part V

Mosquito Ecology

Chapter 26

Mosquitoes

In a previous chapter, we consider adult mosquito population density as a process closely related to, but also exogenous to the process of parasite transmission. We formulated the models for mosquito density in terms of an emergence rate, $\Lambda(t)$. In many studies, this might be good enough. There are some challenges in vector control, however, that call for a deeper understanding of adult mosquito population dynamics in relation to the population dynamics of immature mosquito populations. It's hard to imagine giving any advice about LSM without a model of mosquito ecology. Adult vector control will reduce mosquito population densities, but it would also reduce egg laying affecting immature mosquito populations [8]. Mosquito populations have their own thresholds for persistence, and their own spatial dynamics.

While it could be argued that these are not the primary concerns for malaria programs – we would not disagree – our mathematical framework must be strong enough to support analytics for integrated vector control, including LSM. Immature mosquito dynamics are one of the core dynamical components in our framework. We introduce the topic here, using a very basic model for mosquito ecology [68].

To extend the previous model, we need to make several changes and additional assumptions:

- We define terms that describe egg-laying by adult mosquitoes;
- We write a basic equation to describe the dynamics of immature mosquitoes in aquatic habitats as a function of egg-laying. These models describe how eggs hatch, how mosquito larvae develop and mature in aquatic habitats, pupate, and then emerge as adults. The models do not need to be complex, but they could be.
- We replace the parameter $\Lambda(t)$ from the Aron-May model with a term that describes emergence of adults from aquatic habitats.

One of the big themes we want to introduce here is that the outcomes of vector control can vary substantially by context because of differences in mosquito ecology.

26.1 Aquatic Dynamics

26.1.1 Equations

Here, we start with the equation for adult ecology from before, leaving out parasite transmission dynamics. We modify the basic slightly by adding one extra term that describes delayed maturation, a different response to crowding [68]. The mosquito maturation rate, $\psi e^{-\sigma L}$ defines the average time from egg to emergence, and we add one parameter that slows down larval development in response to crowding. The per-capita mortality rate, $\psi + \theta L$, has two parts: a density independent term ψ and a response to mean crowding θL. We assume that adults lay χ eggs per feeding cycle.

$$\begin{aligned} \frac{dM}{dt} &= \Lambda - gM \\ \hline \frac{dL}{dt} &= \eta - (\psi e^{-\sigma L} + \phi + \theta L)L \\ \Lambda &= \frac{1}{2}\psi L e^{-\sigma L} \\ \eta &= \chi f M \end{aligned} \tag{26.1}$$

26.1.2 Solutions

```
BasicAquatic_dML = function(t, y, params){with(c(params, as.list(y)),{

  # Terms
  Lambda = psi*exp(-sigma*L)*L/2
  eta = chi*f*M

  # Dynamics
  dM = Lambda - g*M
  dL = eta - (psi*exp(-sigma*L) + phi + theta*L)*L

  return(list(c(dM, dL)))
})}
```

```
makePar_BasicAquatic = function(f=1/2.5, g=1/10, chi = 50, psi = 1/10, phi = 1/10, thet
  list(f=f, g=g, chi=chi, psi=psi, phi=phi, theta=theta, sigma=sigma)
}
```

26.1.3 Regulation

```
tt = seq(0,100, by=1)
y0 = c(M=1, L=1)
params = makePar_BasicAquatic()
params1 = makePar_BasicAquatic(phi=1)
params2 = makePar_BasicAquatic(theta=1/10)
params3 = makePar_BasicAquatic(sigma=1/100, theta=0)
```

This code solves the equations:

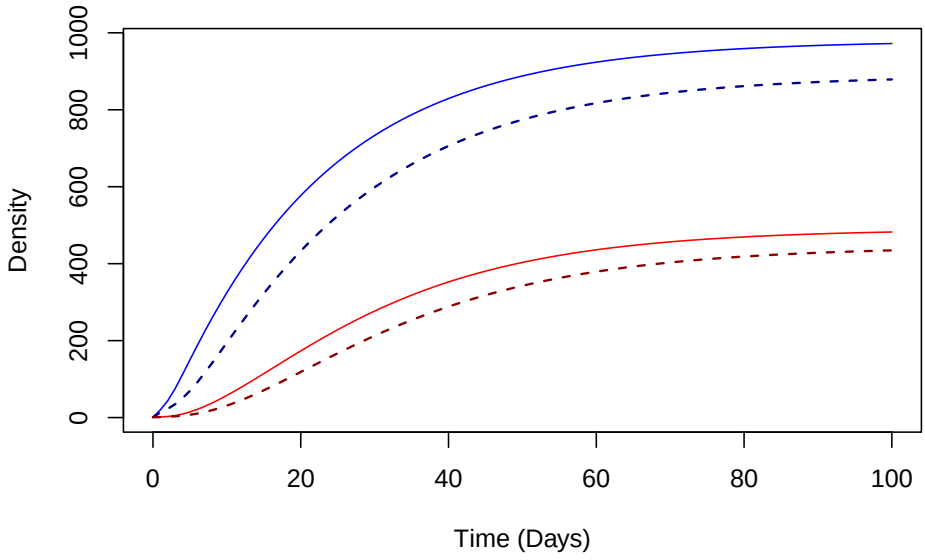
```
require(deSolve)
yout  <- lsode(y=y0, times=tt, func=BasicAquatic_dML, parms=params)
yout1 <- lsode(y=y0, times=tt, func=BasicAquatic_dML, parms=params1)
yout2 <- lsode(y=y0, times=tt, func=BasicAquatic_dML, parms=params2)
yout3 <- lsode(y=y0, times=tt, func=BasicAquatic_dML, parms=params3)
```

```
plotML = function(out, llwd=1, clrL = "blue", clrM="red"){with(data.frame(out),{
  plot(time, L, type = "l", lwd=llwd, col = clrL, xlab = "Time (Days)", ylab = "Density")
  lines(time, M, col = clrM, lwd=llwd)
})}

addML = function(out, llty = 2, llwd=1, clrL = "blue", clrM="red"){with(data.frame(out),{
  lines(time, L, col = clrL, lty=llty, lwd=llwd)
  lines(time, M, col = clrM, lty=llty, lwd=llwd)
})}
```

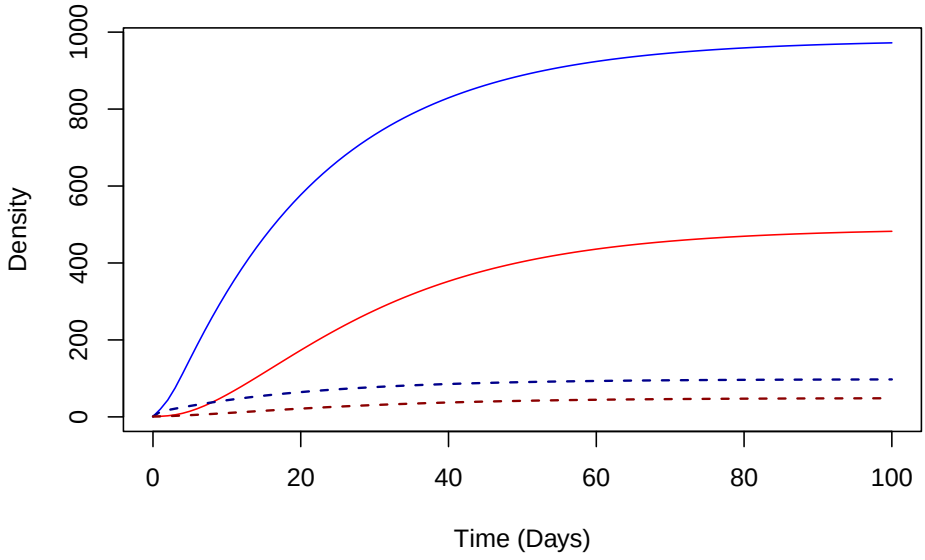
The first thing to point out is that changing the dynamics take some time to approach equilibrium – a couple of months, in this model. The second is that increasing mortality rates for mosquitoes in aquatic habitats by a factor of 10 affects the mosquito densities in minor ways.

```
plotML(yout)
addML(yout1, clrL = "darkblue", clrM="darkred", llwd=1.5)
```



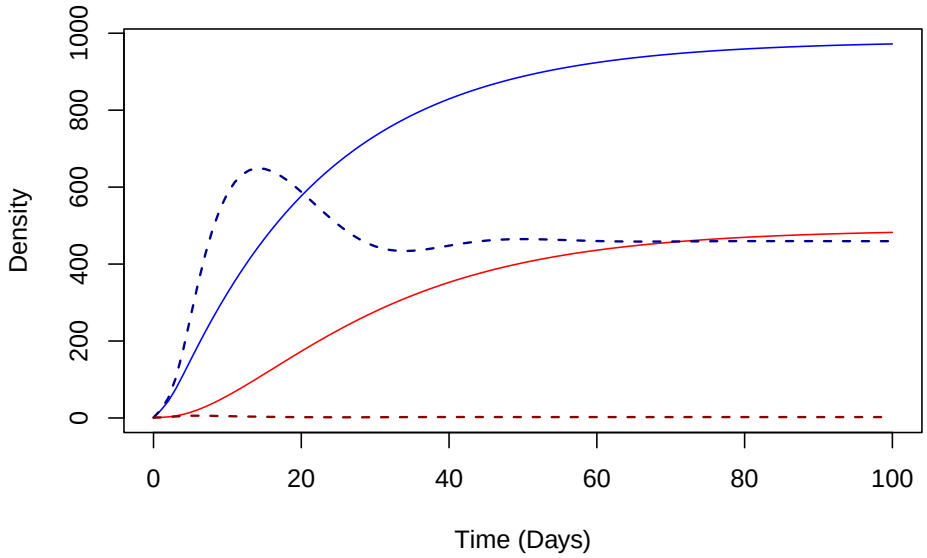
The third is that increasing the response to mean crowding has a very strong effect on the population dynamics.

```
plotML(yout)
addML(yout2, clrL = "darkblue", clrM="darkred", llwd=1.5)
```

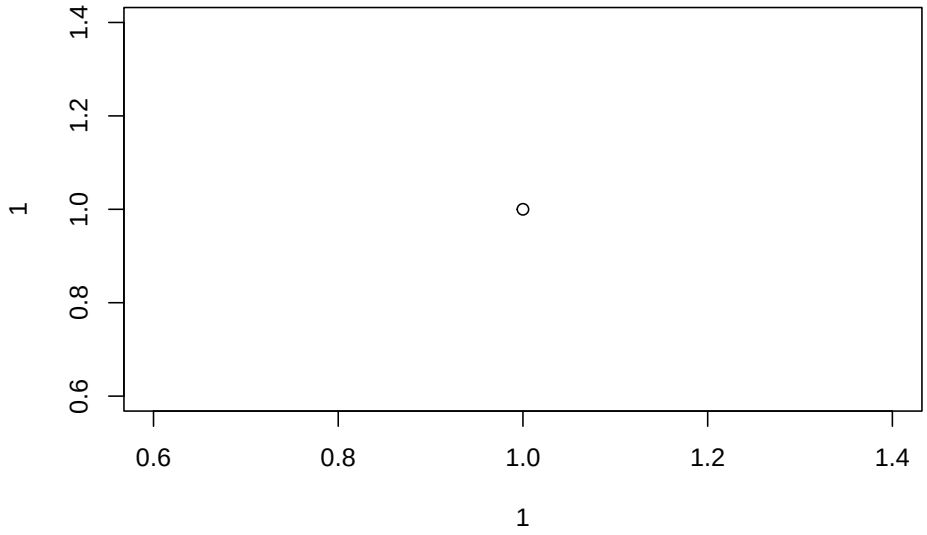


The third point is that density dependent responses to mean crowding cause much larger changes to this system than density-dependent mortality. Aquatic dynamics have a strong tendency to oscillate, but more critically, delayed maturation strongly affects the ratio of adult mosquitoes to adults.

```
plotML(yout)
addML(yout3, clrL = "darkblue", clrM="darkred", llwd=1.5)
```




```
plot(1,1)
```



26.2 Understanding Mosquito Dynamics

This

Chapter 27

Behavioral State Models

Chapter 28

Mosquito Dispersal

Chapter 29

Aquatic Ecology

Chapter 30

Temporal Dynamics

30.1 Exogenous Forcing

30.2 Mosquito Survival through the EIP

Chapter 31

Multi-Species Models

31.1 Vector Competence

Chapter 32

Measuring Mosquitoes

Part VI

Malaria Control

Chapter 33

Malaria Control

33.1 Health Systems

Towards a theory of vector control.

33.2 Vector Control

Chapter 34

Insecticide-Treated Bednets

Chapter 35

Indoor Residual Spraying

Chapter 36

Larval Source Management

Chapter 37

Attractive Toxic Sugar Baits

Chapter 38

Integrated Vector Control

Part VII

Health Systems

Chapter 39

Cohort Dynamics

We need a way of incorporating age into our models.

39.1 Boxcar Models

39.2 Delay

Chapter 40

Demography

40.1 Migration

Chapter 41

Human Travel and Malaria Importation

Chapter 42

Stratification

Chapter 43

Pharmaceutical Interventions

43.1 SMC

43.2 MDA

43.3 Drugs

43.4 Vaccines

Part VIII

Analytics

Chapter 44

Malaria Analytics

Chapter 45

Malaria Metrics

If we want to use evidence to make policy, we need to understand transmission and disease in its local context, which we learn about from assembling and analyzing the evidence. In malaria, the evidence is usually one or more of a core set of metrics. One useful metric is the *Pf*PR from cross sectional surveys; these are available from research studies, but they are also available from demographic health surveys (DHS) or malaria indicator surveys (MIS). If we are lucky, we will also have mosquito catch counts data (by some method), which we scale by effort and report as something like the *Pf*HBR. We estimate the *Pf*SR from some of these caught mosquitoes, and we take their product, the *Pf*EIR. From other designs, we might have estimtes of the *Pf*FoI, or if serology was done, we might know the *Pf*SCR.

From routine clinical surveillance, we could get case counts or the *Pf*SPR. We can learn more if we stratify by age, or if we follow populations over time.

If we are to make sense of malaria as a changing baseline that has been modified by control, then we must know something about how transmission has been modified by control. First and foremost, we must understand anti-malarial drug taking, including the frequency of drug taking (for any purpose), adherence to prescribed regimens, and drug taking by severity. To understand transmission, we must have data describing intervention coverage and its timing. This would include patterns of bed net ownership and use over time, IRS and its coverage. We can't make sense of vector control coverage data without knowing something about the patterns of insecticide resistance. We discuss [Measuring Vector Control] in the next chapter.

For planning, we would like to know the burden of malaria, and we would like to have an estimate of transmission – something like a malaria reproductive number.

In most places, we will have routine surveillance data. We would like to piece

together a coherent and accurate picture of transmission from this and from any other data we may have. To do so, we would like to know how these statistics compare in settings where two or more of them have been measured at the same place and at the same time. From studies like this, we can start to understand the **scaling relationships** among metrics, and the factors that influence them. What we have learned from studying these relationships is that there the relationship are highly non-linear [27], and that some of the factors affecting this relationship are context dependent. We have already mentioned drug taking, and we must acknowledge the difficulty of estimating drug taking by age, in part, because it is often highly heterogeneous in populations. Other factors that have been implicated are heterogeneous biting, environmental heterogeneity, seasonality, travel and mobility.

Sitting between us and the truth is an observational process, and we must attempt to quantify bias and error in this sampling process if we are to have any hope of getting an accurate picture of transmission in context.

To get there, we need to organize our algorithms:

- The *Pf*PR varies in a population by age, sex, and the diagnostic method; so we simulate malaria in cohorts as they age, and we output the predicted *Pf*PR by age, sex, and diagnostic method.
- The *Pf*PR is modified by drug-taking; the effect modification varies in a population by age, access to care, and adherence to drug regimens. In our algorithms, we must consider various patterns of drug taking and chemoprotection.
- Exposure is *seasonal*
  - seasonality *would* affect the true *Pf*PR and the true *Pf*EIR over time;
  - seasonality *could* be a source of error in the estimated *Pf*PR (especially with respect to age) and estimated *Pf*EIR, particularly if the *Pf*PR is reported in young children, and mismatched even slightly with *Pf*EIR.
- Heterogeneous exposure:
  - exposure varies by age;
  - there are other identifiable sources of frailty that we can deal with through stratification (*e.g.* house type, location, bed net use), and there are many frailties we will not be able to identify;
  - the relationship between exposure and infection is affected by *environmental heterogeneity;* because of variability in mosquito populations at a fine spatial grain over time and space, and the trajectories of individual humans who are put at risk,
- The *Pf*EIR near home is obscured by human mobility through time at risk and exposure away from home;

- Travel and imported malaria can increase the *Pf*PR, even if there is no local exposure.
- Measurement errors in the *Pf*EIR
  - There are *house effects* that bias the estimation of the *Pf*EIR
  - Different methods of catching mosquitoes give different estimates for the HBR.

These effects are incorporated into the framework. Frailties are dealt with by stratifying by age ($\omega(a)$) and other traits and assigning *biting weights*; when combined with mobility, these become part of the mixing matrix, β. Environmental heterogeneity modifies the relationship between the *Pf*EIR by stratum and the *Pf*FoI (through F_h). Seasonality in mosquito density ($Z(t)$) translates into seasonality in exposure ($S(t)$). Drug taking and immunity modify the *Pf*PR after exposure.

We would like to understand mosquito populations, but we will rely on whatever data we have to *estimate* the *Pf*EIR (see Figure 4.1). This is a core algorithm for simulation-based analytics, and the framework was designed to serve these ends.

In the following sections, we use the framework to illustrate the effects and effect sizes of these factors. The task of estimating is for another book, **Robust Analytics for Malaria Policy.** In this study, we make full use of modularity. We start by *stratifying* our human population by age, and possibly by other factors. We write algorithms that pass the density of blood feeding mosquitoes in a place, and our algorithms predict mosquito catch counts, depending on the method. From this, we compute the total FoI, which includes the *local* FoI (at and around home), and the *travel* FoI. These exposure rates, and algorithms describing drug taking, are translated into estimates of the true *Pf*PR. The models output the predicted *Pf*PR by age, sex, and method; and we also output the predicted patterns of care seeking and associated metrics at health facilities in an HMIS, by age and sex.

45.1 Realistic Bounds

An important reality check is that we can set sensible expectations about upper bounds on the *Pf*PR in relation to exposure. The maximum value of the *Pf*PR will tend to be observed in untreated populations with homogeneous exposure. To identify these bounds, we assume that we are passing the *true Pf*EIR, but not necessarily the *local Pf*EIR.

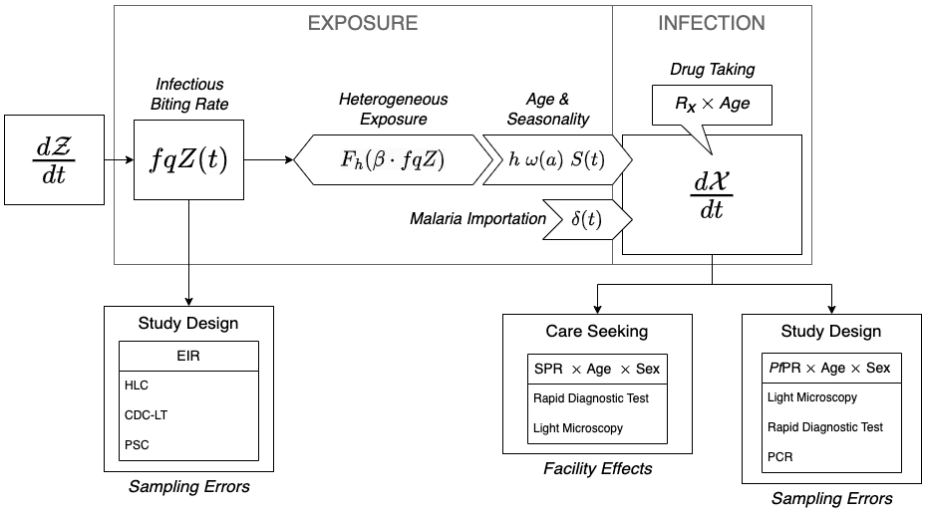


Figure 45.1: Figure 4.1: A generalized framwork for understanding the relationship between the *Pf*EIR and the *Pf*PR:

45.2 The Local Fraction

45.2.1 Travel

45.2.2 Mobility

45.3 Drug Taking

45.4 Seasonality

45.5 Frailty

At low intensity, frailty can affect the prevalence

45.6 Environmental Heterogeneity

45.7 *Pf*EIR *vs.* *Pf*PR in Data

45.8 *Pf*EIR *vs.* *Pf*FoI in Data

45.9 Synthesis

Chapter 46

Discrete Time

Chapter 47

Spatial Concepts and Connectivity

Chapter 48

Mosquito Microecology

Modeling mosquito population dynamics on point sets.

Chapter 49

Microsimulation

Chapter 50

Spatial Control

Chapter 51

Model Libraries

Chapter 52

Base Models

Chapter 53

Built-in Analytics

Chapter 54

Stochasticity

Part IX

Supplements

Chapter 55

References

If you want a PDF and can't find it at the link provided, let us know and we can help you find a copy.

Bibliography

[1] P Armitage. "A note on the epidemiology of malaria." In: *Trop. Dis. Bull.* 50.10 (1953), pp. 890–892. ISSN: 0041-3240.

[2] Joan L Aron and Robert M May. "The population dynamics of malaria". In: *The Population Dynamics of Infectious Diseases: Theory and Applications*. Ed. by Roy M Anderson. Boston, MA: Springer US, 1982, pp. 139–179. ISBN: 978-1-4899-2901-3. DOI: 10.1007/978-1-4899-2901-3_5. URL: https://doi.org/10.1007/978-1-4899-2901-3_5.

[3] Nicolas Bacaër. "Daniel Bernoulli, d'Alembert and the inoculation of smallpox (1760)". In: *A short history of mathematical population dynamics*. Springer, 2011, pp. 21–30. URL: https://link.springer.com/chapter/10.1007/978-0-85729-115-8_4.

[4] Norman T J Bailey. *The Biomathematics of Malaria*. en. Oxford: Charles Griffin & Company Ltd., 1982. ISBN: 978-0-85264-266-5. URL: https://play.google.com/store/books/details?id=4MCAQgAACAAJ.

[5] Daniel Bernoulli and Sally Blower. "An attempt at a new analysis of the mortality caused by smallpox and of the advantages of inoculation to prevent it". In: *Rev. Med. Virol.* 14.5 (2004). Publisher: John Wiley & Sons, Ltd., pp. 275–288. ISSN: 1052-9276. URL: https://www.ingentaconnect.com/content/jws/rmv/2004/00000014/00000005/art00002?crawler=true.

[6] M J Bockarie, A A Gbakima, and G Barnish. "It all began with Ronald Ross: 100 years of malaria research and control in Sierra Leone (1899–1999)". In: *Ann. Trop. Med. Parasitol.* 93.3 (Apr. 1999). Publisher: Taylor & Francis, pp. 213–224. ISSN: 0003-4983. DOI: 10.1080/00034983.1999.11813416. URL: https://doi.org/10.1080/00034983.1999.11813416.

[7] Mark F Boyd, ed. *Malariology. A Comprehensive Survey of all Aspects of this Group of Diseases from a Global Standpoint*. Philadelphia, Pa. and London, W. B. Saunders Co., 1949. URL: https://www.cabdirect.org/cabdirect/abstract/19501000283.

[8] Oliver J. Brady et al. "Adult vector control, mosquito ecology and malaria transmission". eng. In: *International Health* 7.2 (Mar. 2015), pp. 121–129. ISSN: 1876-3405. DOI: 10.1093/inthealth/ihv010.

[9] Leonard Jan Bruce-Chwatt. "History of malaria from prehistory to eradication". In: *Malaria: Principles and Practice of Malariology*. Ed. by Walther

H Wernsdorfer and Ian A McGregor. Ch 1. New York: Churchill Livingstone, Mar. 1988, pp. 1–59. ISBN: 978-0-443-02417-7. URL: http://www.cabdirect.org/abstracts/19900861468.html.

[10] Richard Carter. "Spatial simulation of malaria transmission and its control by malaria transmission blocking vaccination". en. In: *International Journal for Parasitology* 32.13 (Dec. 2002), pp. 1617–1624. ISSN: 00207519. DOI: 10.1016/S0020-7519(02)00190-X. URL: https://linkinghub.elsevier.com/retrieve/pii/S002075190200190X (visited on 01/06/2023).

[11] Francis E G Cox. "History of the discovery of the malaria parasites and their vectors". In: *Parasit. Vectors* 3.1 (Feb. 2010). Publisher: parasitesandvectors.biomedcentral ..., p. 5. ISSN: 1756-3305. DOI: 10.1186/1756-3305-3-5. URL: http://www.parasitesandvectors.com/content/3/1/5.

[12] T E Dempster. *Notes on the Application of the Test of Organic Disease of the Spleen, as an Easy and Certain Method of Detecting Malarious Localities in Hot Climates*. en. Wm. H. Haycock, Secundra Orphan Press, 1848. URL: https://play.google.com/store/books/details?id=3j9pAAAAcAAJ.

[13] K Dietz. "The first epidemic model: a historical note on PD En'ko". English. In: *Aust. J. Stat.* 30A.1 (May 1988). Publisher: Blackwell Publishing Ltd, pp. 56–65. ISSN: 0004-9581. DOI: 10.1111/j.1467-842X.1988.tb00464.x. URL: http://doi.wiley.com/10.1111/j.1467-842X.1988.tb00464.x.

[14] K. Dietz. "The incidence of infectious diseases under the influence of seasonal fluctuations". In: *Mathematical models in medicine*. Ed. by Jürgen Berger et al. Berlin, Heidelberg: Springer Berlin Heidelberg, 1976, pp. 1–15. ISBN: 978-3-642-93048-5.

[15] K Dietz and K P Hadeler. "Epidemiological models for sexually transmitted diseases". en. In: *J. Math. Biol.* 26.1 (1988), pp. 1–25. ISSN: 0303-6812. URL: https://www.ncbi.nlm.nih.gov/pubmed/3351391.

[16] K Dietz, Louis Molineaux, and A Thomas. "A malaria model tested in the African savannah". In: *Bull World Health Organ* 50.3-4 (1974), pp. 347–357. URL: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=4613512.

[17] Klaus Dietz. "Models for parasitic disease control". In: *Bull. Int. Stat. Inst* 46.1 (1975), pp. 531–544.

[18] Klaus Dietz. "Models for vector-borne parasitic diseases". In: *Vito Volterra Symposium on Mathematical Models in biology*. Ed. by Claudio Barigozzi and Simon A Levin. Berlin: Springer-Verlag, 1980, pp. 264–277.

[19] Christopher Dye and G Hasibeder. "Population dynamics of mosquito-borne disease: effects of flies which bite some people more frequently than others". en. In: *Trans. R. Soc. Trop. Med. Hyg.* 80.1 (1986). Publisher: No longer published by Elsevier, pp. 69–77. ISSN: 0035-9203. DOI: 10.1016/0035-9203(86)90199-9. URL: http://dx.doi.org/10.1016/0035-9203(86)90199-9.

[20] Paul E M Fine. "Ross's a priori pathometry - a perspective". en. In: *Proc. R. Soc. Med.* 68.9 (Sept. 1975), pp. 547–551. ISSN: 0035-9157. URL: http://www.ncbi.nlm.nih.gov/entrez/query.fcgi?cmd=Retrieve&db=PubMed&dopt=Citation&list_uids=1105597.

[21] Paul E M Fine. "Superinfection - a problem in formulating a problem". In: *Tropical Diseases Bulletin* 75.6 (1975), pp. 475–488.

[22] P.C.C. Garnham. "History of Discoveries of Malaria Parasites and of Their Life Cycles". In: *History and Philosophy of the Life Sciences* 10.1 (1988). Publisher: Stazione Zoologica Anton Dohrn - Napoli, pp. 93–108. ISSN: 0391-9714. URL: https://www.jstor.org/stable/23329001 (visited on 06/28/2023).

[23] B Grassi. *Studi di uno zoologo sulla malaria*. Roma : R. Accademia dei lincei, 1901. DOI: 10.5962/bhl.title.37999. URL: http://www.biodiversitylibrary.org/bibliography/37999 (visited on 06/28/2023).

[24] B Grassi, G Bastianelli, and A Bignami. "Ulteriori ricerche sul ciclo dei parassiti malarici umani sul corpo del zanzarone". In: *Rend. Acad. Lincei* (1898).

[25] Weidong Gu et al. "An individual-based model of *Plasmodium falciparum* malaria transmission on the coast of Kenya". en. In: *Trans. R. Soc. Trop. Med. Hyg.* 97.1 (Jan. 2003), pp. 43–50. ISSN: 0035-9203. DOI: 10.1016/s0035-9203(03)90018-6. URL: http://dx.doi.org/10.1016/s0035-9203(03)90018-6.

[26] G Hasibeder and Christopher Dye. "Population dynamics of mosquito-borne disease: persistence in a completely heterogeneous environment". en. In: *Theor. Popul. Biol.* 33.1 (Feb. 1988). Publisher: Academic Press, pp. 31–53. ISSN: 0040-5809. DOI: 10.1016/0040-5809(88)90003-2. URL: http://dx.doi.org/10.1016/0040-5809(88)90003-2.

[27] Simon I Hay, David L Smith, and Robert W Snow. "Measuring malaria endemicity from intense to interrupted transmission". en. In: *Lancet Infect. Dis.* 8.6 (June 2008), pp. 369–378. ISSN: 1473-3099. DOI: 10.1016/S1473-3099(08)70069-0. URL: http://dx.doi.org/10.1016/S1473-3099(08)70069-0.

[28] John M. Henry. "A hybrid model for the effects of treatment and demography on malaria superinfection". en. In: *Journal of Theoretical Biology* 491 (Apr. 2020), p. 110194. ISSN: 0022-5193. DOI: 10.1016/j.jtbi.2020.110194. URL: https://www.sciencedirect.com/science/article/pii/S0022519320300497 (visited on 06/27/2023).

[29] A Laveran. "Note sur un nouveau parasite dans le sang de plusieurs malades atteints de fièvre palustre". In: *Note communiquée a l'Académie de Medicine, seance du* 23 (1880).

[30] Alphonse Laveran. *Traité des fièvres palustres*. Octave Doin, 1884.

[31] Alfred J Lotka. "Quantitative Studies in Epidemiology". In: *Nature* 88.2206 (Feb. 1912), pp. 497–498. ISSN: 0028-0836, 1476-4687. DOI: 10.1038/088497b0. URL: https://doi.org/10.1038/088497b0.

[32] Alfred J. Lotka. "Contribution to the Analysis of Malaria Epidemiology. I. General Part". en. In: *American Journal of Epidemiology* 3.supp1 (Jan. 1923), pp. 1–36. ISSN: 1476-6256, 0002-9262. DOI: 10.1093/oxfordjournals.aje.a118963. URL: https://academic.oup.com/aje/article-lookup/doi/10.1093/oxfordjournals.aje.a118963 (visited on 01/02/2023).

[33] Alfred J. Lotka. "Contribution to the Analysis of Malaria Epidemiology. II. General Part (continued). Comparison of Two Formulae given by Sir Ronald Ross." en. In: *American Journal of Epidemiology* 3.supp1 (Jan.

1923), pp. 38–54. ISSN: 1476-6256, 0002-9262. DOI: 10.1093/oxfordjournals.aje.a118965. URL: https://academic.oup.com/aje/article-lookup/doi/10.1093/oxfordjournals.aje.a118965 (visited on 01/02/2023).

[34] Alfred J. Lotka. "Contribution to the Analysis of Malaria Epidemiology. III. Numerical Part". en. In: *American Journal of Epidemiology* 3.supp1 (Jan. 1923), pp. 55–95. ISSN: 1476-6256, 0002-9262. DOI: 10.1093/oxfordjournals.aje.a118966. URL: https://academic.oup.com/aje/article-lookup/doi/10.1093/oxfordjournals.aje.a118966 (visited on 01/02/2023).

[35] Alfred J. Lotka. "Contribution to the Analysis of Malaria Epidemiology. V. Summary". en. In: *American Journal of Epidemiology* 3.supp1 (Jan. 1923), pp. 113–121. ISSN: 1476-6256, 0002-9262. DOI: 10.1093/oxfordjournals.aje.a118964. URL: https://academic.oup.com/aje/article-lookup/doi/10.1093/oxfordjournals.aje.a118964 (visited on 01/02/2023).

[36] G Macdonald. "The analysis of equilibrium in malaria". en. In: *Trop. Dis. Bull.* 49.9 (Sept. 1952), pp. 813–829. ISSN: 0041-3240. URL: https://www.ncbi.nlm.nih.gov/pubmed/12995455.

[37] G Macdonald. "The analysis of infection rates in diseases in which superinfection occurs". en. In: *Trop. Dis. Bull.* 47.10 (Oct. 1950). Publisher: cabdirect.org, pp. 907–915. ISSN: 0041-3240. URL: https://www.ncbi.nlm.nih.gov/pubmed/14798656.

[38] G Macdonald. "The analysis of the sporozoite rate". en. In: *Trop. Dis. Bull.* 49.6 (June 1952), pp. 569–586. ISSN: 0041-3240. URL: https://www.ncbi.nlm.nih.gov/pubmed/14958825.

[39] G Macdonald and G W Goeckel. "The malaria parasite rate and interruption of transmission". en. In: *Bull. World Health Organ.* 31 (1964), pp. 365–377. ISSN: 0042-9686. URL: https://www.ncbi.nlm.nih.gov/pubmed/14267746.

[40] George Macdonald. "The analysis of malaria parasite rates in infants". In: *Tropical diseases bulletin* 47.10 (1950), pp. 915–938.

[41] George Macdonald. *The epidemiology and control of malaria*. Oxford university press, 1957. URL: https://www.cabdirect.org/cabdirect/abstract/19582900392.

[42] Louis Molineaux and Klaus Dietz. "Review of intra-host models of malaria". en. In: *Parassitologia* 41.1-3 (Sept. 1999). Publisher: researchgate.net, pp. 221–231. ISSN: 0048-2951. URL: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=10697860&retmode=ref&cmd=prlinks.

[43] Louis Molineaux et al. "*Plasmodium falciparum* parasitaemia described by a new mathematical model". en. In: *Parasitology* 122.Pt 4 (Apr. 2001). Publisher: researchgate.net, pp. 379–391. ISSN: 0031-1820. DOI: 10.1017/s0031182001007533. URL: http://dx.doi.org/10.1017/s0031182001007533.

[44] T Alex Perkins et al. "Heterogeneity, mixing, and the spatial scales of mosquito-borne pathogen transmission". en. In: *PLoS Comput. Biol.* 9.12 (Dec. 2013). Publisher: Public Library of Science, e1003327.

[45] Robert C Reiner Jr et al. "A systematic review of mathematical models of mosquito-borne pathogen transmission: 1970-2010". en. In: *J. R. Soc. Interface* 10.81 (Apr. 2013), p. 20120921.

[46] Andrew Robinson. "Did Einstein really say that?" en. In: *Nature* 557.7703 (Apr. 2018). Bandiera_abtest: a Cg_type: Books And Arts Number: 7703 Publisher: Nature Publishing Group Subject_term: Physics, Politics, Society, Publishing, pp. 30–30. DOI: 10.1038/d41586-018-05004-4. URL: https://www.nature.com/articles/d41586-018-05004-4 (visited on 06/29/2023).

[47] R Ross. "The logical basis of the sanitary policy of mosquito reduction". en. In: *Science* 22.570 (Dec. 1905), pp. 689–699. ISSN: 0036-8075. DOI: 10.1126/science.22.570.689. URL: http://dx.doi.org/10.1126/science.22.570.689.

[48] Ronald Ross. "An application of the theory of probabilities to the study of *a priori* pathometry. Part I". In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science* 92.638 (1916), pp. 204–230.

[49] Ronald Ross. "An improved method for the microscopical diagnosis of intermittent fever". In: *The Lancet* 161.4141 (1903), p. 86. URL: http://scholar.google.com/scholar?q=related:G2VdXAzaHkoJ:scholar.google.com/&hl=en&num=30&as_sdt=0,5.

[50] Ronald Ross. "Extermination of Malaria". en. In: *Ind. Med. Gaz.* 34.7 (July 1899). Publisher: Pandeya Publications, pp. 231–232. ISSN: 0019-5863. URL: https://www.ncbi.nlm.nih.gov/pmc/articles/PMC5145409/.

[51] Ronald Ross. "Inaugural Lecture on the Possibility of Extirpating Malaria from Certain Localities by a New Method." In: *British medical journal* 2.2009 (1899), pp. 1–4. URL: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=20758555&retmode=ref&cmd=prlinks.

[52] Ronald Ross. *Malaria in Greece*. Tech. rep. Washington, DC: Government Printing Office, 1909, pp. 697–710.

[53] Ronald Ross. *Mosquito brigades and how to organise them*. London: George Philip & Son, 1902.

[54] Ronald Ross. "On some peculiar pigmented cells found in two mosquitos fed on malarial blood". In: *British medical journal* 2.1929 (1897), pp. 1786–1788. URL: http://www.bmj.com/content/2/1929/1786.full.pdf.

[55] Ronald Ross. *Report on the Prevention of Malaria in Mauritius*. en. London: Waterlow, 1908.

[56] Ronald Ross. "Some *a priori* pathometric equations". en. In: *Br. Med. J.* i.2830 (Mar. 1915), pp. 546–547. ISSN: 0007-1447. DOI: 10.1136/bmj.1.2830.546. URL: http://dx.doi.org/10.1136/bmj.1.2830.546.

[57] Ronald Ross. "Some quantitative studies in epidemiology". In: *Nature* 87 (1911), pp. 466–467.

[58] Ronald Ross. *The Prevention of Malaria*. en. 2nd. London: John Murray, 1911.

[59] Ronald Ross. "The prevention of malaria in British possessions, Egypt, and parts of America". In: *The Lancet* (1907), pp. 879–887.

[60] Ronald Ross. "The role of the mosquito in the evolution of the malarial parasite". In: *Lancet* 152.3912 (1898). Publisher: Elsevier, pp. 488–490. ISSN: 0140-6736. URL: https://scholar.archive.org/work/vejsrzltgva5nmudubd5oebje4/access/ia_file/crossref-pre-1909-scholarly-works/10.1016%252Fs0140-6736%252801%252981394-5.zip/10.1016%252Fs0140-6736%252801%252981400-8.pdf.

[61] Ronald Ross. *The thick film process for the detection of organisms in the blood*. Tech. rep. Backup Publisher: University Press of Liverpool. Liverpool, United Kingdom: University Press of Liverpool, June 1903.

[62] Ronald Ross and H P Hudson. "An application of the theory of probabilities to the study of *a priori* pathometry. Part II". In: *Proceedings of the Royal Society of London Series a-Mathematical Physical and Engineering Sciences* 93.650 (1917), pp. 212–225.

[63] Ronald Ross and Hilda Hudson. "An application of the theory of probabilities to the study of *a priori* pathometry. Part III". In: *Proceedings of the Royal Society A: Mathematical, Physical and Engineering Science* 93.650 (Feb. 1917), pp. 225–240. URL: http://adsabs.harvard.edu/abs/1917RSPSA..93..225R.

[64] M W Service. "A short history of early medical entomology". In: *Journal of Medical Entomology* 14.6 (1978), pp. 603–626. URL: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=357723&retmode=ref&cmd=prlinks.

[65] F. R. Sharpe and Alfred J. Lotka. "Contribution to the Analysis of Malaria Epidemiology. IV. Incubation lag". en. In: *American Journal of Epidemiology* 3.supp1 (Jan. 1923), pp. 96–112. ISSN: 1476-6256, 0002-9262. DOI: 10.1093/oxfordjournals.aje.a118967. URL: https://academic.oup.com/aje/article-lookup/doi/10.1093/oxfordjournals.aje.a118967 (visited on 01/02/2023).

[66] H E Shortt and P C C Garnham. "The pre-erythrocytic stage of human malaria, Plasmodium vivax". en. In: *Br. Med. J.* 1.4550 (Mar. 1948). Publisher: ncbi.nlm.nih.gov, p. 547. ISSN: 0007-1447. DOI: 10.1136/bmj.1.4550.547. URL: http://dx.doi.org/10.1136/bmj.1.4550.547.

[67] David L. Smith and Simon I. Hay. "Endemicity response timelines for *Plasmodium falciparum* elimination". en. In: *Malaria Journal* 8.1 (Apr. 2009), p. 87. ISSN: 1475-2875. DOI: 10.1186/1475-2875-8-87. URL: http://www.malariajournal.com/content/8/1/87/abstract (visited on 04/09/2013).

[68] David L. Smith et al. "Mosquito Population Regulation and Larval Source Management in Heterogeneous Environments". en. In: *PLOS ONE* 8.8 (Aug. 2013). Publisher: Public Library of Science, e71247. ISSN: 1932-6203. DOI: 10.1371/journal.pone.0071247. URL: https://journals.plos.org/plosone/article?id=10.1371/journal.pone.0071247 (visited on 05/05/2020).

[69] David L Smith and F Ellis McKenzie. "Statics and dynamics of malaria infection in Anopheles mosquitoes." In: *Malaria Journal* 3 (2004), p. 13. DOI: 10.1186/1475-2875-3-13. URL: http://eutils.ncbi.nlm.nih.gov/entrez/eutils/elink.fcgi?dbfrom=pubmed&id=15180900&retmode=ref&cmd=prlinks.

[70] David L Smith et al. "A New Test of a Theory about Old Mosquitoes". en. In: *Trends Parasitol.* 37.3 (Nov. 2021). Publisher: The Authors, pp. 185–194. ISSN: 1471-4922, 1471-5007. DOI: 10.1016/j.pt.2020.10.011. URL: https://pubmed.ncbi.nlm.nih.gov/33250441/.

[71] David L Smith et al. "Recasting the theory of mosquito-borne pathogen transmission dynamics and control". en. In: *Trans. R. Soc. Trop. Med. Hyg.* 108.4 (Apr. 2014). Publisher: Oxford University Press, pp. 185–197. ISSN: 0035-9203, 1878-3503. DOI: 10.1093/trstmh/tru026. URL: http://dx.doi.org/10.1093/trstmh/tru026.

[72] David L Smith et al. "Ross, Macdonald, and a theory for the dynamics and control of mosquito-transmitted pathogens". en. In: *PLoS Pathog.* 8.4 (Apr. 2012), e1002588. ISSN: 1553-7366, 1553-7374. DOI: 10.1371/journal.ppat.1002588. URL: http://dx.doi.org/10.1371/journal.ppat.1002588.

[73] Neal R Smith et al. "Agent-based models of malaria transmission: A systematic review". en. In: *Malar. J.* 17.1 (Aug. 2018), p. 299. ISSN: 1475-2875. DOI: 10.1186/s12936-018-2442-y. URL: http://dx.doi.org/10.1186/s12936-018-2442-y.

[74] Andrew J Tatem and David L Smith. "International population movements and regional *Plasmodium falciparum* malaria elimination strategies". en. In: *Proc. Natl. Acad. Sci. U. S. A.* 107.27 (July 2010), pp. 12222–12227. ISSN: 0027-8424, 1091-6490. DOI: 10.1073/pnas.1002971107. URL: http://dx.doi.org/10.1073/pnas.1002971107.

[75] H Waite. "Mosquitoes and malaria. A study of the relation between the number of mosquitoes in a locality and the malaria rate". In: *Biometrika* 7.4 (1910). Publisher: [Oxford University Press, Biometrika Trust], pp. 421–436. ISSN: 0006-3444. DOI: 10.2307/2345376. URL: http://www.jstor.org/stable/2345376.

[76] G A Walton. "On the control of malaria in Freetown, Sierra Leone; I. *Plasmodium falciparum* and *Anopheles gambiae* in relation to malaria occurring in infants". en. In: *Ann. Trop. Med. Parasitol.* 41.3-4 (Dec. 1947), pp. 380–407. ISSN: 0003-4983, 1364-8594. DOI: 10.1080/00034983.1947.11685341. URL: http://www.tandfonline.com/doi/full/10.1080/00034983.1947.11685341.

[77] Walther H Wernsdorfer and Ian A McGregor, eds. *Malaria: Principles and Practice of Malariology.* Vol. 1-2. New York: Churchill Livingstone, Mar. 1988. ISBN: 978-0-443-02417-7. URL: http://www.worldcat.org/title/malaria-principles-and-practice-of-malariology/oclc/15793746.

[78] Sean L. Wu et al. "Spatial dynamics of malaria transmission". en. In: *PLOS Computational Biology* 19.6 (June 2023). Publisher: Public Library of Science, e1010684. ISSN: 1553-7358. DOI: 10.1371/journal.pcbi.1010684. URL: https://journals.plos.org/ploscompbiol/article?id=10.1371/journal.pcbi.1010684 (visited on 06/27/2023).

[79] Sean L Wu et al. "Vector bionomics and vectorial capacity as emergent properties of mosquito behaviors and ecology". en. In: *PLoS Comput. Biol.* 16.4 (Apr. 2020), e1007446. ISSN: 1553-734X, 1553-7358. DOI: 10.1371/

journal.pcbi.1007446. URL: http://dx.doi.org/10.1371/journal.pcbi.1007446.