University of Southampton Research Repository

Copyright © and Moral Rights for this thesis and, where applicable, any accompanying data are retained by the author and/or other copyright owners. A copy can be downloaded for personal non-commercial research or study, without prior permission or charge. This thesis and the accompanying data cannot be reproduced or quoted extensively from without first obtaining permission in writing from the copyright holder/s. The content of the thesis and accompanying research data (where applicable) must not be changed in any way or sold commercially in any format or medium without the formal permission of the copyright holder/s.

When referring to this thesis and any accompanying data, full bibliographic details must be given, e.g.

Thesis: Author (Year of Submission) "Full thesis title", University of Southampton, name of the University Faculty or School or Department, PhD Thesis, pagination.

Data: Author (Year) Title. URI [dataset]

University of Southampton

Faculty of Environmental and Life Sciences

Geography and Environmental Science

Exploring Spatiotemporal Variation in Host Population Mobility and Vector-Borne Disease Exposure

by

Jessica Rhona Floyd

ORCID ID: 0000-0003-2022-0791

Thesis for the degree of Doctor of Philosophy

October 2020

University of Southampton

Abstract

Faculty of Environmental and Life Sciences

Geography and Environmental Science

Thesis for the degree of Doctor of Philosophy

Exploring Spatiotemporal Variation in Host Population Mobility and Vector-Borne Disease Exposure

by

Jessica Rhona Floyd

Vector-borne diseases are widespread, diverse and disproportionately affect certain populations. It is well-known that the mobility of host populations is critical to vector-borne disease spread and persistence, and understanding spatiotemporal aspects of this mobility can help predict exposure risk at both fine and large scales. This thesis aims to examine variations in host mobility in the context of vector-borne diseases at opposing ends of the spatiotemporal scale in a 'three-paper format'.

The first paper examines the mobility of a small sample population of humans and livestock in a rural area of western Kenya at a very fine spatiotemporal resolution using surveys and GPS loggers. Several important demographic characteristics are linked to movement patterns, and some seasonal differences in time spent on different types of landcover are observed. Individual variations in movement patterns are likely to be causing differential exposure to some types of vector-borne disease.

The second paper further explores the human factors linked to mobility, focusing on the activity-driven movements of the local population in relation to various types of resource access, as well as demographic differences in activity-driven mobility. Both gender and age are found to be linked to activity-driven movements in this small rural population, and women reported spending longer than men accessing health facilities, highlighting how some population subgroups may have differential access to treatments and preventions for vector-borne disease.

The final paper is set at the other end of the spatiotemporal scale and quantifies the movement patterns of the population of Mozambique over several months, combining these with country-wide epidemiological data to examine how large-scale differences in mobility may affect exposure to malaria. Human-mediated parasite movements are shown to be highly heterogeneous across Mozambique, and individual movements between rural and urban areas are likely to be driving malaria transmission in some parts of the country.

This thesis makes important contributions to our understanding of individual differences in mobility patterns and highlights how both small-scale and large-scale perspectives are valuable for understanding the factors that may increase individual risk of exposure to vector-borne diseases. The work concludes that while mobility underpins much of the dynamics of vector-borne diseases, it is also crucial for understanding differences in the mobility of host populations, as these play an important part in perpetuating transmission and therefore contribute to disease persistence.

Table of Contents

Table of Tables

Table of Figures

Research Thesis: Declaration of Authorship

Print name: Jessica Rhona Floyd

Title of thesis: Exploring Spatiotemporal Variations in Host Population Mobility and Vector-Borne Disease Exposure

I declare that this thesis and the work presented in it are my own and has been generated by me as the result of my own original research.

I confirm that:

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

Floyd JR, Ogola J, Fèvre EM, Wardrop N, Tatem AJ, Ruktanonchai NW. (2020) Activity-specific mobility of adults in a rural region of western Kenya. *PeerJ*. 8, e8798

Floyd JR, Ruktanonchai NW, Wardrop N, Tatem AJ, Ogola J, Fèvre EM. (2019) Exploring fine-scale human and livestock movement in western Kenya. *One Health*. 7, 100081

Signature: .... Date: 07/10/2020

Acknowledgements

Throughout this project I have received a great deal of support from family, friends and colleagues, to whom I dedicate this thesis.

To my husband Alex, for your unwavering support throughout these five years of trips to Kenya, part-time jobs, four house moves and a wedding, for the endless cups of tea and proof-reading, and for the adventure you bring to my life.

To my family: parents, sisters, brothers, and in-laws, for believing in me and keeping me sane and for always supporting my wild ideas wherever they may take me.

To Nick Ruktanonchai, for your supervision, support and friendship, for answering my questions at all hours of the day and for never letting me give up.

To my other supervisors, past and present: Andrew Tatem, Nicola Wardrop and Eric Fèvre. Thank you for your mentorship and guidance, which provided the foundations for me to become an independent scientist.

To my friends and colleagues at WorldPop for support and guidance, especially Dana Thompson, Jess Steele, Victor Alegana and Shengjie Lai for valuable discussions, and to Cori Ruktanonchai for your friendship and thesis-writing support.

To Joseph Ganda Ogola, my fieldwork partner and friend who went above and beyond, who looked after me in a Busia hospital and without whom the data collection would not have been possible.

To the rest of my colleagues in Kenya, who provided intellectual and moral support: especially to Allan Ogendo, for fieldwork assistance and country music; to Isaac Ngere, for his rapid diagnosis in my hour of need; to Lilian Achola and Christine Mosoti for keeping me in line administratively; to Catherine Wilson and Laura Falzon for the weekday slaughterhouse visits, weekend trips and candlelit boardgames, and to the many others who made me feel at home in Kenya.

To the county commissioner, the chiefs, assistant chiefs and village elders of Busia County for facilitating the fieldwork, and to the people and ruminants of Busia County for participating in the studies.

And finally, to the University of Southampton and the research councils supporting the Zoonoses and Emerging Livestock Systems (ZELS) programme for providing the opportunity and funding for this project.

Definitions and Abbreviations

CDR	Call Detail Record
DDT	Dichlorodiphenyltrichloroethane
DHS	Demographic and Health Survey
EIR	Entomological Inoculation Rate
IRS	Indoor Residual Spraying
ITN	Insecticide Treated bed Net
GLMM	Generalized Linear Mixed Model
GPS	Global Positioning System
LMM	Linear Mixed Model
MDA	Mass Drug Administration
MDG	Millennium Development Goals
ODK	OpenDataKit
SDG	Sustainable Development Goal
UN	United Nations
WHO	World Health Organisation

Chapter 1 Introduction

1.1 Background

Vector-borne infectious diseases are diverse and found globally, accounting for over 17% of all infectious diseases (World Health Organization, 2014). The majority are transmitted between hosts by blood-feeding arthropods, which act as vectors for parasitic, bacterial and viral pathogens and can infect both human and animal host species depending on the pathogen. Because competent vector populations (including but not limited to mosquitoes, ticks, sand flies and tsetse flies) are widespread in tropical and sub-tropical parts of the world, these diseases disproportionately affect the global poor.

Early efforts to manage vector-borne diseases have been documented as far back as the Greek and Roman times, well before the first transmission route of a vector-borne disease (malaria) was discovered in 1897 (A. L. Wilson et al., 2020). Before the widespread use of residual insecticides, the control of vector-borne diseases relied mainly on vector population and environmental management strategies such as the drainage of aquatic habitats, vegetation clearance, housing improvements and in the case of zoonotic vector-borne diseases, game and habitat destruction (Ault, 1994; A. L. Wilson et al., 2020). Between 1940 and 1960, major discoveries of new insecticides such as DDT led to the introduction of indoor residual spraying (IRS) programmes to control vector populations on a massive scale. Hailed as a breakthrough technology, these were successful at controlling many of the most prevalent vector-borne diseases at the time but came at the cost of a reduction in global funding and resources for their control (WHO, 2017). In the 1970s and 1980s, trials of insecticide-treated bed nets (ITNs) showed how their use resulted in huge reductions in vector-borne disease infection and mortality (World Health Organization, 1989), and mass drug administration (MDA) campaigns also proved effective against mosquito-borne diseases such as malaria and lymphatic filariasis, yet global funding for these diseases reached a low in the 1990s (WHO, 2017; A. L. Wilson et al., 2020).

From the early 2000s, big increases in funding to combat vector-borne disease have seen significant advancements made towards the elimination of diseases like malaria, onchocerciasis and lymphatic filariasis in some areas, but several others have spread or re-emerged in new parts of the world and the WHO now estimates that 80% of the global population is at risk of at least one vector-borne disease, causing 700,000 deaths annually (World Health Organization, 2014). In 2017 the WHO launched the Global Vector Control Response (GVCR), a strategy to tackle vector-borne diseases calling for locally appropriate and integrated control methods as these are more likely to be both effective and sustainable (WHO, 2017). As areas reach low levels of pathogen prevalence, the movement of vector-borne pathogens through human and animal mobility becomes more important to their persistence. The GVCR identifies population movement as one of seven major challenges to the strategy, threatening to impede progress in the fight against vector-borne diseases.

In recent years, the field of population mobility in relation to infectious diseases has expanded significantly, with a number of important reviews focusing on this topic. Vector-borne pathogens rely on competent host and vector populations for their transmission, and the mobility of these two populations is critical to their proliferation. Population mobility causes the movement of pathogens from one area to another through individual host movements which can contribute to the persistence of pathogens in these areas and expose host and vector populations to novel pathogens that they may not have previously encountered. Since many vector populations are not thought to move long distances, much of the current research focuses on the movements of host populations, both human and animal alike, and their role in the spread and persistence of vector-borne diseases. In a recent review of the literature on mobility and malaria, one of the most prevalent and thoroughly-studied vector-borne diseases, Smith and colleagues note that although population mobility is a much-discussed topic in relation to malaria, the factors affecting mobility patterns that change the risk of transmission is understudied (Smith & Whittaker, 2014). This holds true for many other vector-borne diseases, particularly those that fall into the category of neglected tropical diseases (NTDs), which mainly affect populations living in resource-poor areas where data on mobility patterns has historically been sparse.

Host population mobility is itself a broad and complex topic, affected by both anthropological and environmental factors with significant implications for several different aspects of vector-borne disease transmission. Moreover, like all infectious diseases, the occurrence, persistence and spread of vector-borne diseases varies across spatiotemporal scales, from endemic diseases that are widespread, to epidemics that can cause devastation to local human and animal populations over relatively short periods of time. To help quantify mobility in different contexts for vector-borne diseases, several unique methods of data collection have been developed over the years, and improving technology continues to bring new applications. Travel history surveys, GPS loggers and large digital datasets are just a few of the tools that have been employed to quantify host mobility in the context of vector-borne diseases. Understanding mobility can help researchers determine where and how people spend their time and therefore estimate both their risk of exposure to vector-borne diseases and their access to infrastructure and resources that enable prevention and treatment of disease. Moreover, since the majority of mathematical representations of vector-borne disease dynamics assume homogenous mobility and contact rates within populations (if mobility is accounted for at all), understanding factors that affect heterogeneity in individual movement patterns within host populations is key to understanding their impact on vector-borne disease burden.

1.2 Thesis objectives

Vector-borne diseases threaten the livelihoods of people in many parts of the world and are encroaching on new areas as social, demographic and environmental factors continue to change. The mobility of human and animal populations perpetuates vector-borne disease transmission and shows great heterogeneity both spatially and temporally. As technology advances, so too do methods for capturing the mobility data needed to quantify movement patterns across scales, thus improving our knowledge of how vector-borne diseases spread, and which populations are most at risk.

This thesis aims to provide evidence to account for individual differences in movement patterns that may disproportionately increase the burden of vector-borne diseases amongst some subgroups of populations, with the following objectives:

1) To explore demographic and seasonal differences in the movement patterns of a small human and livestock population in rural Kenya at a fine scale;
2) To quantify their activity-driven movement patterns and explore links between these and demographic characteristics; and,
3) To quantify the movement patterns of the population of Mozambique and link these to malaria incidence across the country.

This work follows a “three-paper PhD” format and is organized to address these objectives accordingly. Chapter 1 provides a broad overview of the study of mobility for vector-borne diseases and outlines the structure and objectives of this thesis. Chapter 2 contains a literature review that summarises existing research of vector-borne disease mobility studies, focusing on human and animal studies of mobility, the various methods used to capture mobility data and the factors driving movement patterns at different scales. Chapter 3 gives the study framework, introducing the contexts that the subsequent papers were conducted in and gives an overview of the primary and secondary datasets collected, justification of methods used and describes how this work contributes to the current literature. Chapter 4 is a modified version of the article “Exploring fine-scale human and livestock movement in western Kenya”, published in the journal *One Health* (Floyd et al., 2019). Chapter 5 presents the article “Activity-specific mobility of adults in a rural region of western Kenya”, as published in the journal *PeerJ* (Floyd et al., 2020). Chapter 6 presents a large-scale approach in the final article “Estimating malaria parasite mobility in

Mozambique using mobile phone records”, submitted to the journal *Nature Communications* for publication. Chapter 7 summarises the main findings from the three articles while outlining their limitations and discusses their impact and further research that could provide more evidence to support the broad conclusions of this thesis.

Chapter 2 Literature Review

To address the objectives outlined in the previous section, a detailed understanding of how host mobility contributes to the burden of vector borne disease is required. This chapter will provide a summary of the current literature on the spatiotemporal dynamics of host mobility in relation to vector-borne disease. First, I give a brief overview of vector-borne diseases and discuss some of the environmental and anthropological factors affecting their distribution and burden in different contexts around the world. I then explore the importance of host mobility in vector-borne disease research, focusing on understanding why mobility is important for vector-borne diseases of both humans and animals, how mobility data is collected and analysed, and how environmental and anthropological factors that affect mobility can alter exposure to vector-borne diseases.

2.1 Vector-borne disease in humans and animals

2.1.1 The burden of vector-borne disease in humans and animals

Many common species of arthropods such as mosquitoes, ticks, flies and fleas are capable of carrying and transmitting pathogens that may be bacterial, viral or parasitic between mammalian hosts. These pathogens have the potential to cause disease in both humans and animals. The major vector-borne diseases of humans include malaria, dengue fever, yellow fever, Zika virus disease and Chagas disease, amongst others. Some of these diseases, like malaria, have been circulating in human and vector populations for centuries while others such as the Zika virus have newly emerged. All exert a burden on human populations, and the WHO estimates that vector-borne diseases account for 17% of the global burden of infectious diseases, causing over 700,000 deaths each year (WHO, 2017).

Some vector-borne pathogens found in animals can also cause disease in humans; these are known as zoonotic vector-borne diseases and include Japanese encephalitis, trypanosomiasis and Rift Valley Fever (RVF). As well as causing illness in humans, zoonotic vector-borne diseases exert a burden on animal hosts and can reduce livestock productivity, causing substantial economic losses. In rural areas where many of the world’s poorest communities reside, a significant proportion of people are dependent on livestock for their livelihoods and thus it is widely acknowledged that these diseases perpetuate poverty (WHO & DFID, 2006). RVF, for example, is a mosquito-borne viral disease responsible for sporadic outbreaks in sub-Saharan Africa that primarily cause abortion and death in livestock, resulting in severe economic losses, but RVF can also cause mild to severe disease in humans (Rich & Wanyoike, 2010). Previous outbreaks of RVF have caused the death of thousands of heads of livestock at a time, as well as hundreds of human deaths (Dar et al., 2013; Nanyingi et al., 2015), and the 2006-2007 East African outbreak had an estimated economic cost of USD$541 million (Nanyingi et al., 2015).

Other zoonotic vector-borne pathogens that cause more chronic disease can have a relatively low mortality rate but still exert a high burden on their hosts and cause large-scale reductions in productivity. African Trypanosomiasis is an example of one such disease. Affecting humans, livestock and wildlife, and caused by parasites of the genus *Trypanosoma* spread by infected tsetse flies, trypanosomiasis is endemic in 36 countries in Africa and causes both acute and chronic illness in hosts. In livestock, the disease can present with clinical symptoms ranging from mild disease to death depending on the pathogen species of infection (Morrison et al., 2016). It has been estimated that 50 million cattle live in tsetse-infested areas in Africa, and causes $4.75 billion in economic losses every year (Van den Bossche et al., 2010). Thus, African Trypanosomiasis has been identified by the World Health Organisation (WHO) as “a major obstacle to the economic development of rural areas” (World Health Organization, 2013). Nevertheless, RVF, Trypanosomiasis and many other zoonotic vector-borne diseases fall into the category of ‘neglected tropical diseases’, due to the lack of global funding, research and resources directed towards their control (D. Molyneux et al., 2011) and as such, understanding factors that affect their spread could have important benefits for the populations that suffer most from them.

2.1.2 Environmental and anthropological factors contributing to vector-borne disease

Although these diseases are all vector-borne, they can have vastly different transmission cycles affected by both intrinsic factors such as differences in host and pathogen biology, and extrinsic factors such as environmental conditions and host behaviour. The control of vector-borne diseases relies on disrupting their transmission cycle, so understanding the environmental and anthropogenic factors that influence transmission is a crucial part of implementing effective strategies for their control.

Because vector-borne diseases are reliant on a capable vector population for their spread, they are frequently sensitive to environmental and climatic factors such as rainfall, temperature, land cover and land use. Malaria is one example of such a disease: responsible for over 400,000 deaths and more than 220 million cases annually, malaria is arguably the highest-burden vector-borne disease (Roth et al., 2018; World Health Organization, 2018b). Dependant on transmission by the *Anopheles* genus of mosquitoes, it is highly sensitive to a range of environmental factors which means that its distribution across most of sub-Saharan Africa is far from homogenous. For example, rainfall patterns and altitude have been linked to peaks in the populations of different anopheline species in Kenya (Zhou et al., 2004) and studies have shown that particular types of land cover can lend themselves to increased risk of malaria transmission, especially shallow, shady and stagnant water bodies, as these constitute potential anopheline larval breeding sites (Stefani et al., 2013). Other vector-borne diseases are similarly sensitive to environmental factors: for example, epidemics of yellow fever (a viral disease transmitted by mosquitoes) have been linked to changes in normal rainfall patterns (Hamrick et al., 2017; Ivers & Ryan, 2006; Vasconcelos et al., 2001) and particular land types have been associated with increased risk of Crimean-Congo Haemorrhagic fever (Messina et al., 2015).

Anthropogenic changes in the environment can also affect the spread and distribution patterns of vector-borne diseases. Deforestation (Barros & Honório, 2015), agricultural activities (Mutuku et al., 2009) and gold mining (Moreno et al., 2007) are some examples of changes in land use that have been associated with changes in density of malaria larval breeding sites, often due to inadvertent creation of anopheline breeding sites in suitable aquatic environments such as shallow puddles. Changing land use is an important factor in the spread of other vector-borne diseases too. Urbanization is widely acknowledged to be an important driver in outbreaks of dengue fever (Gubler, 2012), with one study in Taiwan finding a direct link between areas with higher degrees of urbanization and dengue incidence (P. C. Wu et al., 2009). Land use changes can similarly affect zoonotic vector-borne disease prevalence; in the case of tsetse-transmitted livestock trypanosomiasis, the clearing of areas of vegetation and the introduction of domesticated livestock has resulted in drastic reductions in tsetse populations in some areas but increases and the resurgence of trypanosomiasis epidemics in others, due to tsetse populations being forced to switch from wildlife to livestock as their primary source of food (Van den Bossche et al., 2010). The effects of these wide-ranging natural and anthropogenic environmental factors on vector-borne disease distribution vary greatly with setting and are therefore an important topic in studies of vector-borne disease.

Demographic and social factors can also affect the distribution of vector-borne disease. Many studies have shown that some vector-borne diseases are highly linked to high levels of household poverty. There are many reasons for this: relatively poorer households are less able to afford testing and treatment, are more likely to work laborious, time-intensive jobs and live in housing of poorer quality, which means a higher likelihood of vector exposure (Boelaert et al., 2009). Other socio-demographic factors can also affect the distribution of vector-borne diseases, partly due to differing human behaviours associated with these factors. For example, studies have linked higher educational level to earlier care-seeking behaviour and increased uptake of treatment for malaria (Kibusi et al., 2015). Context is extremely important, however. For example, in some settings women have been found to seek care before men (Kaur et al., 2013), while other studies have found that women seek care for infectious diseases later than men (Shayo et al., 2015; Voeten et al., 2004), possibly due to specific barriers to healthcare for women in certain communities. Socio-demographic factors such as these can affect vector-borne disease dynamics and vary widely across contexts, so it is important to have an understanding of the impact of these factors when designing efforts to control vector-borne diseases.

The WHO has stated that anthropogenic and environmental factors such as these have an important impact on the transmission and distribution patterns of vector-borne diseases from local to regional, national and international scales, and calls for this complexity to be taken into account when designing vector control efforts. In their Global Vector Control Response strategy for 2017-2030, population mobility is also highlighted as a major threat to the success of these efforts (WHO, 2017). As vector-borne diseases emerge in new areas of the world while approaching elimination in others, understanding how host mobility and the environmental and anthropological factors driving it affects vector-borne disease distribution and transmission will prove crucial to designing effective control and elimination strategies in different contexts.

2.2 Host mobility and vector-borne disease

A key property that vector-borne and many other infectious pathogens have in common is that their spread is initiated or perpetuated by the movements of their hosts. Individual host mobility dictates where and when populations interact with each other, with vector populations, and with their environment, so understanding the spatiotemporal dynamics of host movement is critical to understanding how outbreaks occur, why pathogens persist in particular places and which populations are most at risk of exposure. In this section, I discuss the importance of host mobility in infectious disease research generally before focusing on aspects of human and animal mobility that are important to vector-borne disease research specifically.

2.2.1 The growing importance of host mobility in infectious disease research

Although there are many hundreds of infectious pathogens that cause disease in humans, a common feature of their transmission is that all rely on some form of direct or indirect contact between hosts in order to be transmitted. For humans, this contact is often driven by mobility on various scales, whether international, regional or local, and the forefront of human mobility research involves the study of the movements of individuals across these scales in space and time. As in animal populations, these movements vary widely on spatial and temporal scales and play an important role in disease transmission (Martens & Hall, 2000; R. Mansell Prothero, 1977).

Moreover, variation in movement within human populations is frequently much greater than in animal populations, and this presents some unique challenges for disease control. Compared to animal populations, humans tend to have larger, more widespread and more connected networks, meaning highly contagious diseases like influenza can rapidly spread globally (Riley, 2007; Stoddard et al., 2009). Furthermore, the number of people moving has increased in recent years, and consequently so has connectivity between individuals and populations. Globally, it has been estimated that over 60 million people fly internationally each week (Dirk Brockmann & Theis, 2008), with hundreds of millions more travelling by car, train and bus. Recent studies of air traffic data suggest that this type of mobility is increasing (Gabrielli et al., 2019), leading inevitably to increasing connectivity. This can be positive and contribute to improvements in quality of life. The obvious benefits include improvement to resource access, transport services and improved social inclusion, to name a few, but new complexities also present new challenges on both local and national/international scales. For example, increasing mobility and connections puts pressure on local traffic networks in urban areas. Understanding this mobility and how it affects populations can help policymakers make decisions about where best to funnel resources to improve networks (Hickman et al., 2013).

Increasing connectivity also has important consequences for infectious disease transmission and has been identified as a crucial factor in the spread of new pandemics such as SARS (Lam et al., 2003; W. Li et al., 2005), H5N1 influenza (Y. Li et al., 2011; Van Borm et al., 2005) and the Ebola virus (Leroy et al., 2005, 2009). For example, in the SARS outbreak of 2002-2003 and the recent COVID-19 outbreak, human movement has been identified as a key factor in the rapid global spread of the virus, specifically through air travel routes (Lai et al., 2020; Poutanen et al., 2003; Wells et al., 2020) but also through movement by other modes of transport on regional and local scales (Kraemer et al., n.d.). Movement restrictions are an obvious strategy to tackle these outbreaks but compared to animal populations, they can be difficult to enforce in human societies to combat the spread of disease (Eric Fèvre et al., 2006; Longini et al., 2005). The restriction of cattle movements between farms in the Netherlands after the foot-and-mouth outbreak in 2001 was shown to reduce the size of the epidemics (Velthuis & Mourits, 2007) and although effective, such restrictions can have severe economic and practical consequences amongst human populations, meaning their implementation is often slow and contentious (Budd et al., 2009). Nevertheless, in the wake of the current COVID-19 pandemic, many countries implemented nation-wide movement restrictions with varying efficacy, demonstrating the importance of host mobility in driving outbreaks of infectious diseases.

For pathogens like influenza and COVID-19 that are primarily spread by direct host-to-host contact, host movements are directly related to risk of exposure to infected individuals and therefore transmission and spread can often be predicted to some extent if host movements are well understood (Riley, 2007). For vector-borne pathogens however, transmission is most often indirect, occurring in a human-vector-human cycle instead of directly between humans. Therefore, understanding host mobility is compounded by the added complexity of exposure to vector populations; for a transmission event to occur, a host must come into contact with a competent and infectious vector population, which is often highly environmentally dependent (Rottier & Ince, 2005). Moreover, vector-borne disease transmission can involve different species of hosts: for zoonotic vector-borne pathogens there is the added dimension of animal movements as well as those of humans. In a disease like trypanosomiasis for example, understanding the movements of the animal host reservoir (frequently cattle but occasionally also small ruminants) can help us understand exposure risk of the human population, because both humans and livestock can move from endemic to pathogen-free areas and introduce trypanosomes to a previously unexposed vector population, thereby causing local outbreaks of human sleeping sickness and animal trypanosomiasis (Morris, 1962; Selby et al., 2013).

Movement ecology has long been studied in the context of infectious disease transmission, and several studies demonstrate how for both human and animal populations, mobility can affect disease transmission in ways beyond exposure driven by host contact. In a seminal paper in 1988, Pulliam outlines how differential birth and death rates of populations can cause habitats to naturally acquire ‘source’ and ‘sink’ properties, whereby sources have a surplus of population and sinks a deficit. Competition for resources drives individual members of these populations to move to areas where resources are sufficient to sustain a species, and this mobility can increase the size of ecological niches in which populations and, by extension, their pathogens, exist (Pulliam, 1988). A study by Cross and colleagues builds on this concept, examining the effect of movements of individuals between subgroups of structured populations on the transmission of acute and chronic diseases and demonstrating that for populations with a grouped structure (as typically found in human and animal populations), movements between and within these groups change the transmission dynamics of diseases with pandemic potential (Cross et al., 2005). A recent review by Boulinier and colleagues, while detailing several different types of wildlife movement, goes further, highlighting that host movement does not necessarily equate to pathogen dispersal (Boulinier et al., 2016). For example, a study by Mazé-Guilmo and colleagues tested the hypothesis that parasite dispersal rates depend strongly on host movements and found that while this is broadly true, other environmental and ecological factors can change the spatial distribution of different genetic compositions found in parasites and therefore affect the coevolution of parasite-host populations, resulting in different spatial distributions (Mazé-Guilmo et al., 2016).

Therefore, mobility of different individuals in a population drives not only host exposure to pathogens, but can change the ecological niches of the pathogens themselves.

To be able to quantify how mobility affects vector-borne pathogen transmission, it is important to understand the types of both human and animal movement patterns involved in vector-borne disease spread. Because of the inherent differences between the two, I will discuss human and animal movement patterns separately.

2.2.2 Spatiotemporal human mobility patterns and vector-borne disease

For vector-borne diseases, increased mobility not only leads to increased connectivity between host populations, but also between host and vector populations. This increased connectivity causes changes in exposure to vector-borne diseases, and can ultimately lead to changes in disease distribution and spread (Stoddard et al., 2009). Because of this added complexity of vector-borne disease, understanding patterns of human mobility is crucial to help us design more effective disease control strategies, yet has historically been understudied for vector-borne pathogens, partly due to the complexity of host-vector-pathogen dynamics but also due to difficulties in collecting detailed datasets on the movement patterns of individual hosts.

A key topic in the field of human mobility for vector-borne disease research is to understand mobility across spatiotemporal scales, particularly how the interactions of populations with their environments can affect their disease risk. In 1977, Prothero identified human mobility as a neglected factor in epidemiology and highlighted both spatial and temporal variation in human movements as fundamental intricacies that affect the spread of infectious diseases (R. Mansell Prothero, 1977). Understanding human mobility across these scales is particularly important to improving our knowledge of where and how people are exposed to vector-borne diseases, because mobility is heterogeneous at both ends of the scale and this varying mobility affects exposure to vector-borne diseases differently. People can be exposed to vector-borne diseases daily through their short-distance movements to particular areas or places on a local scale, and similarly they can be exposed through medium and long-distance movements to wider areas where vector-borne diseases are present. However, questions remain about how fine-scale and large-scale movements may affect individual exposure and vulnerability to vector-borne disease. To help put these into context, Stoddard and colleagues built on work by Prothero and developed a seminal framework to visualise human movements in the context of vector-borne diseases, illustrating the scales at which human movements affect exposure, distribution and spread of vector-borne diseases (Figure 2-1).

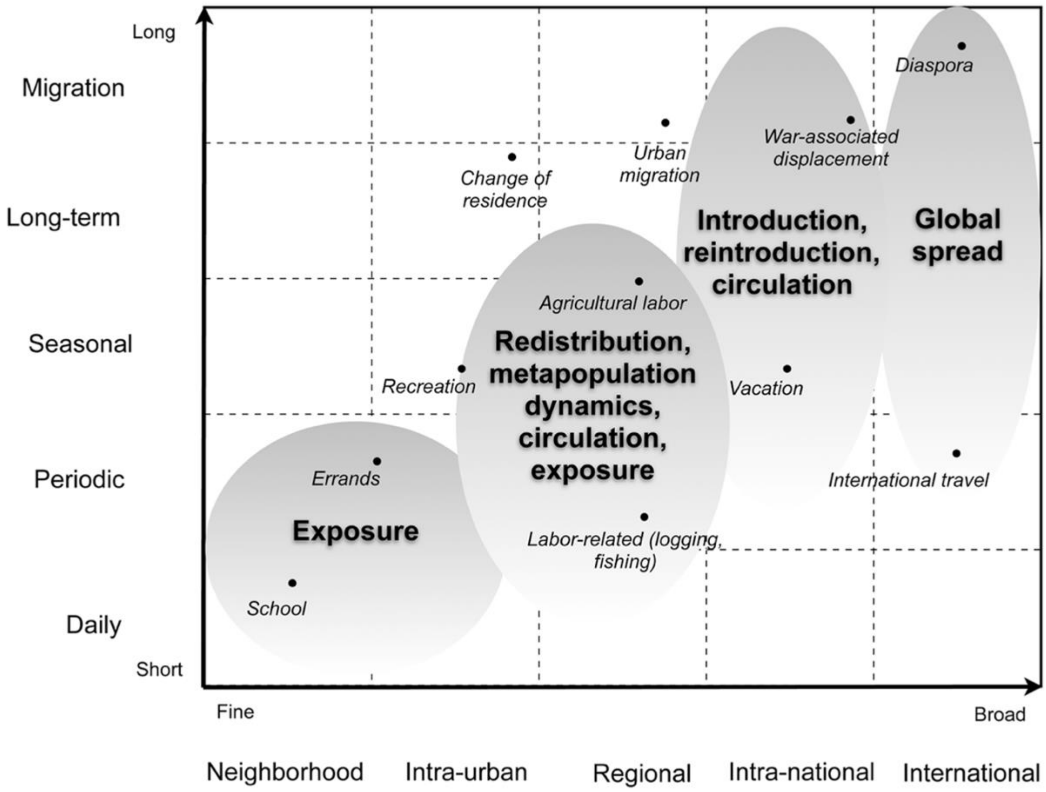


Figure 2-1 A theoretical framework of spatial (x-axis) and temporal (y-axis) scales of human movements that affect vector-borne disease dynamics. As published by Stoddard et al., 2009.

These spatial and temporal differences in human movement patterns affect exposure to vector-borne diseases differently. At the local end of the spatial scale on which humans move, studies have shown that individual movements can determine exposure to vector-borne disease. For example, a related study in Iquitos, Peru looking at dengue fever, a mosquito-borne viral disease, found that movements between households determined an individual's exposure to Dengue fever (Stoddard et al., 2013). Regional movements have also been implicated in the spread of vector-borne disease. Studies in Swaziland and South Africa have identified human movement from Mozambique as risk factors for malaria transmission, and it is likely that these movements are driving malaria transmission in these countries (Raman et al., 2020; Tejedor-Garavito et al., 2017). Finally, at the international and global end of the scale, human movements can lead to the spread of vector-borne diseases between countries in a matter of hours, thereby driving their spread and persistence in new areas (A. J. Tatem et al., 2012).

Temporal variation in human movements is also important to vector-borne disease and can affect individual exposure differently. For example, regular movements such as daily visits to a local area may matter more than infrequent ones in areas where vector populations have an alternative blood meal, as frequency of movement to these areas would increase individual risk of exposure and therefore contribute to transmission. At the other end of the scale, in regions where pathogen and vector populations are limited to specific areas, longer-term human movements may be a more important driver of transmission than shorter ones as these are likely to lead to the carriage of pathogens to new areas and potentially competent new vector populations (N. W. Ruktanonchai et al., 2016; Stoddard et al., 2009).

2.2.3 Spatiotemporal animal mobility patterns and vector-borne disease

Studies of animal mobility in relation to infectious diseases have long highlighted the importance of mobility in disease spread. A recent review of emerging diseases of wildlife presents a list of 34 vertebrate wildlife pathogens thought to have emerged since 2000, and more than half (18) of these have some type of host movement explicitly stated as the cause of disease spread (Tompkins et al., 2015), with movement implicated as a contributor to spread in many of the others. The importance of animal movement has also been repeatedly stressed in studies of pathogens that do not fall under the ‘emerging’ label, such as the risks associated with the movements of dogs and wild animals in the importation of *Echinococcus multilocularis* (Hegglin & Deplazes, 2013) and the transmission cycle of rabies (J. Chen et al., 2015; Cliquet et al., 2014), or the risks associated with the trade and other movements of livestock in the transmission of bovine tuberculosis (Caron et al., 2016; Cowie et al., 2014), brucellosis (Caron et al., 2016; Cowie et al., 2014; Njeru et al., 2016) and cysticercosis (Lian F. Thomas et al., 2013).

For vector-borne diseases, animals can also act as reservoirs of the vector-borne pathogen, presenting an additional layer of complexity and playing an important part in disease distribution and spread. For example, the movement patterns of wild birds have long been implicated in the introduction of vector-borne pathogens to new areas. West Nile virus (WNV) outbreaks in the USA have been linked to the migration corridors of wild birds along the eastern seaboard, and these birds are suspected to be the main amplifying hosts of WNV in nature, acting as a reservoir and spreading the virus to previously unaffected vector populations that can then pass the virus on to humans and domestic livestock (Reed et al., 2003). Similarly, cattle movements are thought to be an important factor in the spread of vector-borne diseases like bluetongue virus and sleeping sickness, though few studies have been able to directly identify these movements as the primary cause of disease spread (E. M. Fèvre et al., 2005; T. Sumner et al., 2017). Quantifying the movements of cattle and other types of domestic livestock and specifically their contact with humans and exposure to different environments could help quantify the risk of vector-borne disease exposure in some populations.

However, these movement patterns can be difficult to quantify as they vary on spatial and temporal scales, each with different implications for disease transmission. Spatially, animal movements vary greatly across the different species of host: wild animals such as buffalo can regularly move long distances in a matter of hours (Funston et al., 1994), while the movements of their livestock cousins, domestic cattle, tend to be restricted to a smaller area controlled by the people that look after them. However, even for the same host species, spatial scales can differ greatly; the relatively restricted movements of cattle, for example, can vary from daily grazing and foraging (Bailey, 1995) to medium- and long-distance human-mediated movements between markets and nomadic movements covering large distances (Natale et al., 2009). These movement patterns on different spatial scales can affect vector-borne disease transmission across a range of contexts. For example, an outbreak of sleeping sickness in a previously unaffected area of south-eastern Uganda was found to have its roots in a cattle market stocked with cattle originating from endemic sleeping sickness areas (EM Fèvre et al., 2001), illustrating the potential of market movements to introduce infection into new areas. The nomadic long-distance movements of cattle have also been linking to vector-borne disease transmission: a study in a rural area of eastern Kenya found that rates of RVF infection were higher in humans living near long-distance routes of nomadic cattle (Tigoi et al., 2020).

Animal movement patterns also vary temporally, which can affect disease transmission in different ways depending on the context. Using ruminant livestock as an example, some movements follow daily patterns like going out to grazing sites, while others occur over a longer term, such as movements between farms or movements to markets. For example, sheep movement networks in Great Britain follow a highly seasonal pattern, with large numbers moving between farms in the summer months of August and September. Analysis of these sheep movement networks has shown that the increased movement during these months is linked to greatly increased risk of disease outbreaks occurring (Kiss et al., 2006). Conversely, studies have shown that other types of movement that follow seasonal patterns, such as the migratory movements of several species of wild animals, can actually help mitigate pathogen spread by reducing host exposure to infected habitats and hampering the migration of infected individuals, thereby reducing pathogen transmission within a population (Altizer et al., 2011). Although dependent on many factors such as the type of pathogen, host population densities and land use patterns, these examples illustrate how context is important even when movements have similar temporality, such as seasonal movements.

Because understanding the spatial and temporal scales of animal movements is critical to disease transmission, several studies have made use of animal mobility data to study movements in the context of infectious diseases. In the past it has been difficult to obtain accurate records of host movement, with many studies relying on survey or direct observation methods but in recent years, advances in technology have enabled more comprehensive measurements of these movements through a variety of different methods, including the use of remote sensing data, mobile phone call records, radio telemetry collars and tags for animals, and Global Positioning Systems (GPS) devices. Data from these methods have proven valuable for providing evidence to aid the control of several animal infectious diseases. For example, a study in Edinburgh collected the movements of dog foxes using radio-based trackers to prevent rabies outbreaks (Kolb, 1984), free-ranging pigs have been tracked to provide evidence for cysticercosis interventions in rural Kenya (L. Thomas et al., 2013), and a study that used GPS tracking devices to monitor fine-scale goat movements was able to identify ‘hotspots’ for *Cryptosporidium* disease transmission in a rural area of Tanzania (Parsons et al., 2014). These are examples of studies undertaken in the context of zoonoses, yet studies that have explored both human and animal movements together are lacking. This is likely because their neglected status means that the resources to control zoonoses are often limited, and collecting mobility data can be time-consuming and expensive. Nevertheless, a review by Fèvre and colleagues highlights how it is clear that the control of many zoonotic vector-borne diseases would benefit from these types of mobility data, particularly when host movement varies across seasons and landscapes (Eric Fèvre et al., 2006).

2.3 How is host mobility measured?

As a complex process resulting from the interplay of many individuals and factors within a system or network, mobility has been the subject of studies for decades (Weiner, 1992) with applications within and across scientific disciplines. From the long-distance migration movements of whole populations to the daily movements of city-dwelling individuals, the types of mobility studied has varied widely over spatial and temporal scales. Archaeologists study the migration patterns of early human populations (Watson & Stoll, 2013), epidemiologists use movement patterns of both humans and animals to predict the spread of disease (Travis et al., 2011), geographers have long studied movement patterns in disaster situations (Lu et al., 2012), anthropologists study mobility in a social context to assess inequalities in resource access (Bardosh et al., 2017), and urban planners use movement data to design transport networks efficiently (Weiner, 1992). All of these situations require detailed knowledge of movement patterns: knowing why, where, when and how people move are crucial to efforts to get resources where they are needed, and this wide range of uses for mobility data has led to the growth of methods for their collection and analysis.

In the past, efforts have been made to estimate human mobility patterns on large scales using proxies such as the movement of banknotes (D. Brockmann et al., 2006), changes in night-time brightness (Bharti et al., 2011) and the movements of domestic livestock (Larson, 2012) to name a few. Other, more direct studies of movement have relied on data collected on mass scales such as air travel (A. J. Tatem et al., 2012) and other transportation data (H. Liu et al., 2015) to estimate movement patterns at a population level. Data on the individual movements of hosts, however, has been more difficult to collect. There are several reasons for this: the sheer time and effort required for study frameworks involving large numbers of individuals recruited to movement studies, the ethical and privacy issues surrounding the tracking of individuals' movement patterns, and the difficulties in compliance and recall bias for retrospective and prospective studies of movement patterns all make individual movement data more difficult to capture than population-level movement data.

The main methods of mobility data collection for humans and animals and the relative spatiotemporal scales that they cover are shown in Figure 2-2 and Figure 2-3. Individual movement data has traditionally been captured through surveys and censuses, both of which can provide high-quality data on movement patterns at an individual level and the same time allow linkage to important demographic factors. The ability to collect high-quality datasets linked to a range of social and demographic datapoints is important because there are many factors, both known and unknown, that affect individual movements and therefore the mobility of a population as a whole. However, survey methods can suffer from a lack of temporal resolution due to their cross-sectional nature. In recent years the use of GPS-based technology has increased, and so too has the ability to collect data from large numbers of individuals on massive scales, through improved technology and the rise of geo-referenced 'big data' and its offshoots. These advantages have not gone unnoticed in the mobility community, with movement patterns being analysed through geo-located tweets (Blanford et al., 2015; Hawelka et al., 2014), data collected from location-sharing services (Z. Chen & Schintler, 2015) and mobile phone records (Jones et al., 2018). For vector-borne diseases, many of these methods have been used to elucidate mobility patterns in various settings. Human mobility, particularly in the context of malaria transmission, has been studied extensively through different methods of mobility data collection and on scales ranging from highly localised transmission between households in villages (Fornace et al., 2019) to transmission through national and cross-border movements (N. W. Ruktanonchai et al., 2016; Tejedor-Garavito et al., 2017).

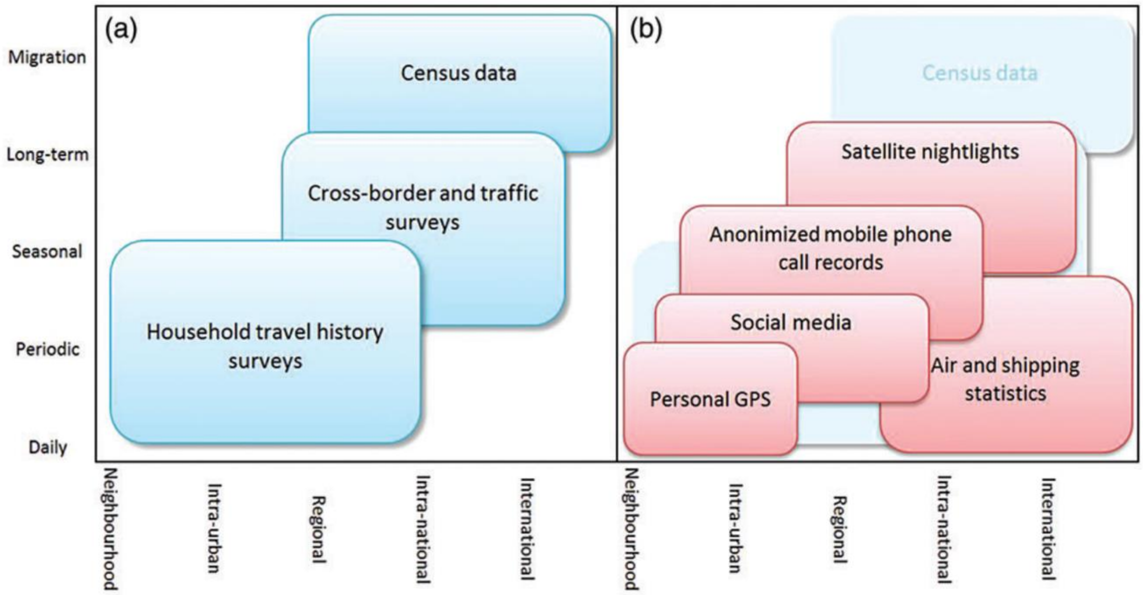


Figure 2-2 Methods of human mobility data capture, as published by Tatem, 2014.

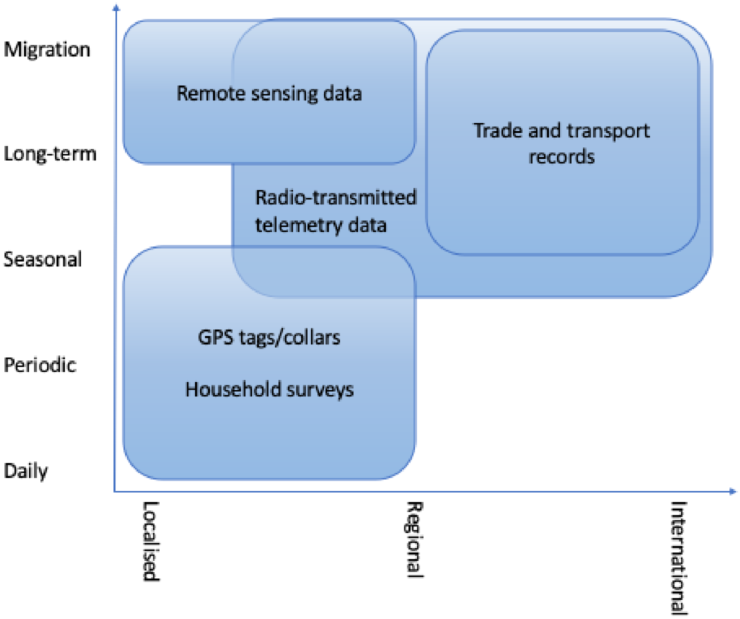


Figure 2-3 Methods of animal mobility data capture with the x-axis indicating approximate spatial scale and the y-axis indicating approximate temporal scale.

Figure 2-2 and Figure 2-3 present a visual summary of some of the most widely used methods for human and animal mobility data capture and the relative spatiotemporal scales covered by these methods. Methods for the direct measurement of individual host movements can be broadly classed into three categories: survey methods, telemetry and GPS devices, and large digital datasets (such as call detail records from mobile phones, remote sensing and social media datasets). In this section, I discuss the unique advantages and disadvantages of movement data arising from each of these. Because of their limited ability to track individual movements through time, I exclude discussion of movement data such as trade and transport statistics from this section.

2.3.1 Movement data collected through surveys

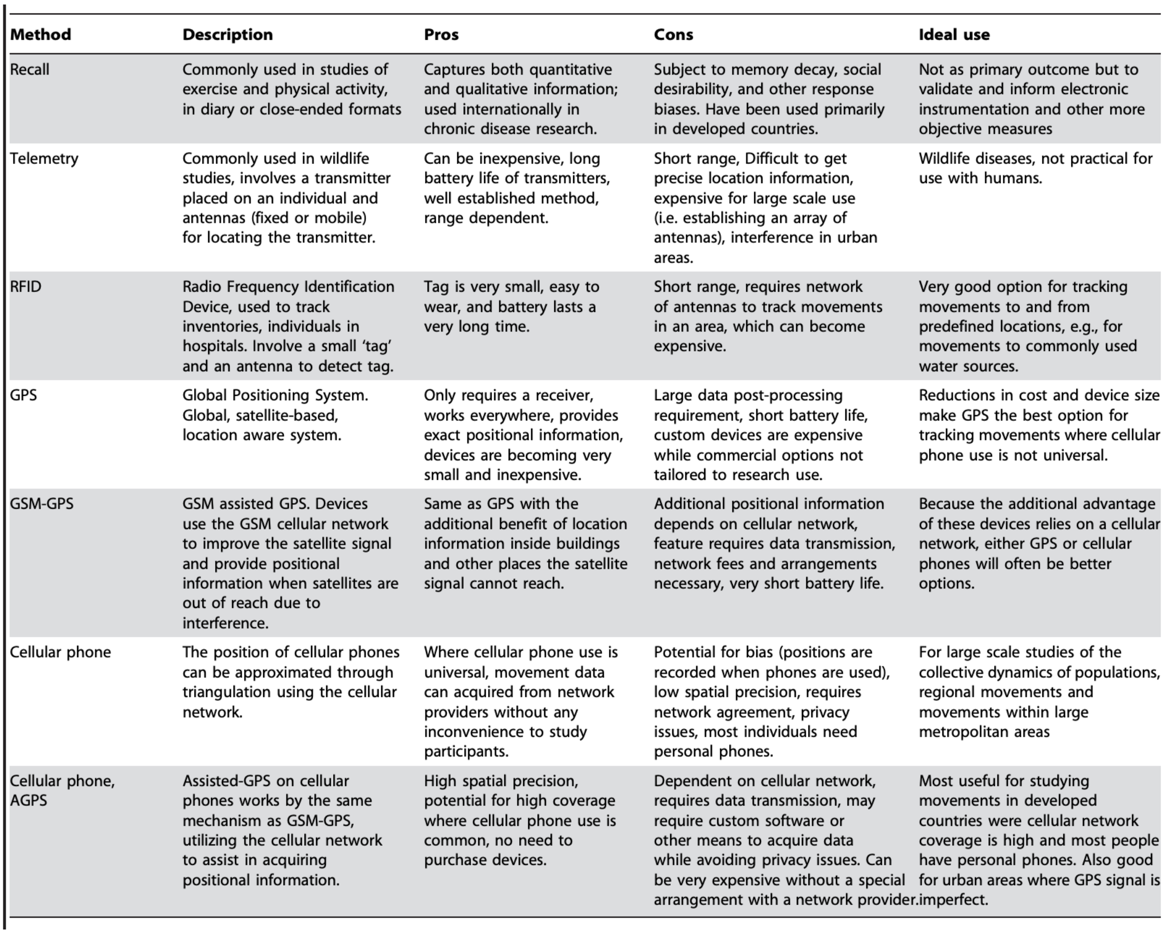
Traditionally, human movements have been measured using various survey methods, including travel history surveys and from data collected as part of larger surveys such as the Demographic and Health Surveys (DHS) and national censuses. Travel history surveys have been widely used in infectious disease studies to determine case history and identify sources of infection (Tejedor-Garavito et al., 2017; Tessema et al., 2019), and in other studies to understand how and where people travel in disaster and conflict situations (L. Wu & Walters, 2016). Data from the DHS program in particular has been used to identify long-term migration patterns of populations. For example, a study in Peru used longitudinal DHS data over six years to investigate the relationship between fertility and migration and were able to link demographic indicators to geographic mobility (White et al., 1995), illustrating an important advantage of surveys over anonymous types of data collection: the ability to also collect demographic information from participants. This key feature of surveys allows researchers to know whether the survey is likely to be representative of the underlying population (thereby reducing bias) and to link these important demographic factors to individual mobility patterns.

Survey methods have some inherent disadvantages, however. When conducted on large scales it can be difficult to quantify highly spatially resolved movements using survey data, but this limitation can be acceptable if the study aims are suitable for the data collected. For example, studies of how migrational human mobility is linked to environmental factors commonly use survey methods to identify seasonal patterns (Gray & Mueller, 2012; Murali & Afifi, 2014), because very detailed knowledge of individual movements is not required to elucidate general seasonal movements. Moreover, although survey methods are often limited by their cross-sectional nature, when conducted longitudinally they can be used to understand changes in movement patterns over time. For example, national surveys like the DHS and national censuses are conducted periodically at the household level and generally include questions on current and past places of residence. With several years of these censuses, flows of people can be calculated in a longitudinal fashion and aggregated to quantify population flows (Barbosa et al., 2018). However, censuses only measure change of residence over 1 or 5 year periods and are usually conducted only once a decade, giving them a very coarse temporal resolution compared to most travel history surveys. This limits their use for modelling human mobility for vector-borne diseases, and thus they are only of value for these applications if they correlate with shorter temporal scale flows, as previous studies have been able to do in Haiti (Nick W. Ruktanonchai et al., 2016) and Kenya (A. Wesolowski et al., 2013).

Generally, when only survey methods are used to collect movement data either temporal or spatial resolution is sacrificed, if not both. Nevertheless, this does not mean that these data are necessarily inadequate for mobility studies. A study comparing travel surveys with mobile phone call detail records data from Paris and Chicago found good concordance between the two for daily patterns of movement in these urban areas, although only two days per person were analysed in the study (Schneider et al., 2013). This illustrates how sufficiently detailed survey data can be as accurate as large-scale mobile phone data in some circumstances. A study in Kenya by Wesolowski and colleagues compared census-derived migration data to a corresponding fine-scale movement dataset and showed that in certain contexts, census data can be used to approximate movement patterns on smaller temporal scales than previously thought (A. Wesolowski et al., 2013). Overall, survey methods can be valuable tools to quantify movement patterns, particularly when combined with demographic data and when used with appropriate study aims and design.

2.3.2 Movement data collected by telemetry and GPS devices

Improvements in movement tracking technology as a data collection method have also given rise to a number of mobility studies, particularly in animals but more recently in human populations too. In the past, many of these studies made use of radio telemetry: a technology that involves the remote capturing of data through a device whose function is to measure and record data at specified time intervals and transmit the data via radio waves to a receiver (Tomkiewicz et al., 2010). More recently, this technology has expanded to include Global Positioning System (GPS) data (Perras & Nebel, 2012; Tomkiewicz et al., 2010). These data are most often GPS coordinates with a timestamp, but other metrics can also be included such as elevation, barometric pressure and other environmental measures. The devices can automatically send data back to the observer, thereby allowing real-time capture of movement patterns, or they can store data in onboard memory to be downloaded on retrieval of the device. The types of devices used to track human and animal movements are portable, and those that record GPS data are known as GPS loggers. Stoddard and colleagues compiled some of the most commonly used movement data collection methods in humans (including several different telemetry and GPS-based methods) and outlined their advantages and disadvantages (Stoddard et al., 2009), this is reproduced here in Figure 2-4 for reference.



Method	Description	Pros	Cons	Ideal use
Recall	Commonly used in studies of exercise and physical activity, in diary or close-ended formats	Captures both quantitative and qualitative information; used internationally in chronic disease research.	Subject to memory decay, social desirability, and other response biases. Have been used primarily in developed countries.	Not as primary outcome but to validate and inform electronic instrumentation and other more objective measures
Telemetry	Commonly used in wildlife studies, involves a transmitter placed on an individual and antennas (fixed or mobile) for locating the transmitter.	Can be inexpensive, long battery life of transmitters, well established method, range dependent.	Short range, Difficult to get precise location information, expensive for large scale use (i.e. establishing an array of antennas), interference in urban areas.	Wildlife diseases, not practical for use with humans.
RFID	Radio Frequency Identification Device, used to track inventories, individuals in hospitals. Involve a small 'tag' and an antenna to detect tag.	Tag is very small, easy to wear, and battery lasts a very long time.	Short range, requires network of antennas to track movements in an area, which can become expensive.	Very good option for tracking movements to and from predefined locations, e.g., for movements to commonly used water sources.
GPS	Global Positioning System. Global, satellite-based, location aware system.	Only requires a receiver, works everywhere, provides exact positional information, devices are becoming very small and inexpensive.	Large data post-processing requirement, short battery life, custom devices are expensive while commercial options not tailored to research use.	Reductions in cost and device size make GPS the best option for tracking movements where cellular phone use is not universal.
GSM-GPS	GSM assisted GPS. Devices use the GSM cellular network to improve the satellite signal and provide positional information when satellites are out of reach due to interference.	Same as GPS with the additional benefit of location information inside buildings and other places the satellite signal cannot reach.	Additional positional information depends on cellular network, feature requires data transmission, network fees and arrangements necessary, very short battery life.	Because the additional advantage of these devices relies on a cellular network, either GPS or cellular phones will often be better options.
Cellular phone	The position of cellular phones can be approximated through triangulation using the cellular network.	Where cellular phone use is universal, movement data can acquired from network providers without any inconvenience to study participants.	Potential for bias (positions are recorded when phones are used), low spatial precision, requires network agreement, privacy issues, most individuals need personal phones.	For large scale studies of the collective dynamics of populations, regional movements and movements within large metropolitan areas
Cellular phone, AGPS	Assisted-GPS on cellular phones works by the same mechanism as GSM-GPS, utilizing the cellular network to assist in acquiring positional information.	High spatial precision, potential for high coverage where cellular phone use is common, no need to purchase devices.	Dependent on cellular network, requires data transmission, may require custom software or other means to acquire data while avoiding privacy issues. Can be very expensive without a special arrangement with a network provider.	Most useful for studying movements in developed countries were cellular network coverage is high and most people have personal phones. Also good for urban areas where GPS signal is imperfect.

Figure 2-4 Advantages and disadvantages of commonly used methods of mobility data capture, as published by Stoddard et al., 2009.

Mobility studies that employ telemetry and GPS-based devices as the primary method of data capture usually require the active solicitation of participants, which can be resource-intensive and time consuming and therefore often result in small sample sizes. However, this also means that studies making use of this technology can be designed to take advantage of the strengths of the devices. Studies employing these methods often use devices that are highly specialised and adapted to their purpose. For example, studies of diving seabirds (Browning et al., 2018) have combined GPS loggers with time-depth recorders to gain movement information specific to the study aims, while a study of the long-distance movements of elephant herds combined GPS loggers with temperature sensors to analyse the impact of temperature on movement patterns (Thaker et al., 2019). These demonstrate the versatility of GPS loggers for capturing movement data in combination with other relevant data.

In studies of humans, the requirement for active solicitation means that it is also possible to gather demographic data to link to movement patterns. A study of human movement in relation to malaria elimination in Zambia used GPS loggers to link movement patterns to activity spaces across malaria seasons, with additional collection of demographic data allowing the researchers to ensure their sample was representative of the underlying population (Searle et al., 2017), and to identify differences in movement patterns by age and sex of the participants (Hast et al., 2019). A different study in neighbouring Tanzania successfully used similar devices to track the movements of different species of domesticated animals to assess overlap of landscape use and identify areas where zoonotic disease transmission may be occurring (Parsons et al., 2014). These studies demonstrate the utility of GPS tracking technology for identifying risks of exposure to vector-borne and other infectious diseases, as well as their suitability for use in rural settings where other methods of data collection like mobile phone records may be biased towards wealthier populations. However, a major disadvantage of these devices is that they are typically only used to track relatively small population numbers over short periods of time (see Figure 2-2 and Figure 2-3) and so researchers have looked for ways to track much larger populations over longer timeframes, such as with mobile phone records.

2.3.3 Methods for measuring human movement through mobile phone data

Technological advancements and the rise of ‘big data’ have produced a number of novel datasets from which patterns of human mobility can be inferred. These data can be analysed independently or integrated with other data collection methods and come in many forms, including air traffic data, geo-located social media data, and call detail records from mobile phones. Several of these have been used for mobility studies: for example, big datasets have been used to predict risk of disease importation (Bajardi et al., 2011; Le Menach et al., 2011; A. J. Tatem et al., 2012), to explore seasonal patterns of movements across borders (Blanford et al., 2015; Gabrielli et al., 2019; Recchi et al., 2019), to dynamically map populations (Deville et al., 2014) and to explore urban mobility patterns (Yuan & Raubal, 2012).

Recent increases in mobile phone penetration in particular have allowed the collection of large amounts of high-quality mobility data in near-real time, allowing the ability to respond quickly and appropriately to changes in mobility. Studies of mobility from mobile phones most often analyse call detail records (CDRs) to estimate individual mobility. A CDR consists of an ‘event’ (call or text), the date and time of the event, and the GPS coordinates of the cell tower where the event was recorded. The approximate geolocation of individual subscribers to a mobile phone network can be inferred from these records. CDRs are routinely collected for billing purposes by network operators in real-time and can be used to provide evidence for decision-making. CDR datasets can help countries mitigate the effects of emergencies by enabling policy makers to get resources to populations more efficiently. For example, scientists were able to use a computational model to provide estimates of population displacements just nine days after the 2015 Nepal earthquake using the anonymised CDRs of 12 million people (R. Wilson et al., 2016), which helped policymakers plan effective resource allocation to cope with the impacts of the earthquake. Another study was able to use CDRs to retrospectively measure the impact of travel restrictions in Sierra Leone on human mobility during an Ebola outbreak in 2015 (Peak et al., 2018). Methods employing synthetic datasets also hold promise for future studies: one study was able to model population movements in the New York and Los Angeles metropolitan areas using synthetic CDRs, a method that was shown to have high fidelity with real CDRs (Isaacman et al., 2012) but is currently only useful for areas with high population densities.

CDR datasets can be extensive, often covering whole portions of a population and thus are very valuable for describing complex population movement processes. However, to protect the privacy of individuals, they are commonly anonymised or pseudonymised, and therefore missing demographic data due to ethical and privacy issues. This is an important disadvantage that can contribute to a ‘data gap’, which occurs when datasets are biased toward certain subgroups that are more represented in the dataset than others, and leads to the misrepresentation of populations (Aultman-Hall & Ullman, 2019). Some studies have sought to overcome this obstacle by linking two or more datasets from the same population. For example, one study using mobile phone data from Kenya measured the impact of potential biases (cell phone ownership is known to be biased towards wealthy urban males) and found that they can be identified and corrected for (Amy Wesolowski et al., 2013). Another way to reduce bias is to collect the demographic information at the same time as linking to a mobile phone dataset: a pilot study in the UK conducted a survey to collect demographic information alongside prospective collection of users Google Location History (GLH) data from their mobile phones in conjunction with GPS loggers and found that the GLH data matched well with data from their personally-assigned GPS logger, showing how valuable these data could be on larger scales (N.W. Ruktanonchai et al., 2018).

Several studies have used mobile phone data to gain insight into human movements because these datasets can often be analysed shortly after collection, making them particularly suited to cases where real-time analysis is required. Furthermore, because these data are also frequently large and span several weeks or months, they can be used to gain an understanding of the dynamic nature of mobility in a population. A study using mobile phone data collected over nine days from subscribers to a mobile phone network in a Chinese city used a Dynamic Time Warping algorithm to classify urban areas by the similarity of the mobility patterns that occurred within them, which enabled the identification of ‘outlier’ urban areas that had unusual mobility patterns and gave insight into how mobility in the city changed from day to day (Yuan & Raubal, 2012). In less populated areas, other studies have used similar methods to produce temporally explicit estimates of population densities at different scales (Deville et al., 2014).

These kinds of analyses can give insight into the dynamic nature of mobility in both densely and sparsely populated areas, by providing near real-time information on the processes driving mobility. For infectious disease research, data from these analyses can be used to inform mathematical and computational models which can predict the effects of mobility on disease spread in near real-time. In the absence of highly-resolved datasets, modellers routinely make the assumption that the mobility of the populations in question is homogenous, partly due to a lack of evidence to account for differences in mobility patterns (Smith & Whittaker, 2014). In the next section I discuss some of the factors that affect mobility patterns, particularly as they relate to vector-borne disease exposure.

2.4 Factors driving human mobility patterns and their relevance to vector-borne diseases

Understanding mobility and the movement patterns of a population is valuable in vector-borne disease research because knowing where and when human and animal hosts travel can help identify risks of exposure to disease, and these movement patterns vary both spatially and temporally as previously discussed. There are many different factors driving these variations and it is important to understand these to get an accurate picture of population movements as a whole, yet many studies of human mobility in relation to infectious disease assume that most people in a population travel in the same way (Bansal et al., 2007; Riley, 2007). This assumption is not due to a lack of evidence to suggest otherwise, rather, the studies acknowledge that human mobility is highly complex and is affected by various factors which may be demographic, socioeconomic or environmental in nature.

When such assumptions are made, our understanding of the mobility of a population (and therefore its connectivity, accessibility and use of transport networks) is distorted. This can disproportionately affect the most vulnerable people in society: by ignoring or overlooking their movements and the factors that drive them, our understanding of mobility is biased towards the data we have from particular subgroups and a ‘data gap’ is created (Aultman-Hall & Ullman, 2019). This is often unavoidable and even the biggest datasets like transport data, mobile phone records and geo-located social media data are frequently biased towards the subgroups of people who have access to these services. Studies have repeatedly shown that women (Hanson & Hanson, 1980; McGuckin & Murakami, 1999; Sarmiento, 2000), older people (Collia et al., 2003; Figueroa et al., 2014) and poorer people (Gray & Mueller, 2012) have different movement patterns, use transport differently, and have different drivers of movement compared to others. As a consequence, these and other subgroups of people are likely to access resources, migrate and travel differently. Efforts to disaggregate movements by factors such as gender, race and household wealth can improve our understanding of human mobility in contexts such as disaster management, urban planning and disease control (Smith & Whittaker, 2014).

The factors that underpin these differences in mobility are important for vector-borne disease research because they affect subgroups differently and contribute to the unequal burden of vector-borne disease amongst populations in different settings. In this section, I discuss some of the key factors driving these differential mobility patterns in the context of vector-borne disease, including demographic factors, access to resources, and seasonal drivers. Because these factors largely affect human rather than animal mobility in the context of vector-borne diseases and the focus of this thesis excludes animal vector-borne diseases that are not zoonotic, I primarily discuss how these factors relate to human mobility.

2.4.1 Demographic factors affecting human mobility

For vector-borne diseases, the importance of understanding demographic differences in movement patterns between individuals is becoming increasingly clear. Mathematical models of mobility have enormous value in predicting disease spread and identifying areas of transmission risk, but are often limited by a lack of high-quality mobility data in relevant contexts and therefore make the assumption that populations or individuals in a model move in the same way (Riley, 2007). Identifying factors underpinning individual differences in movement patterns could help improve these models for vector-borne disease because some of these patterns are linked to demographic characteristics. For example, a study of the role of commuting movements of people in a city in Taiwan found that elderly adults and housewives, who tended not to commute, were more likely to cause localised outbreaks of dengue fever than people who were regular commuters, whose movement patterns were linked to dengue epidemics over larger geographic areas (Wen et al., 2012). Although much of the evidence for variations in movement patterns by demographic characteristics comes from people living in urban areas of more economically developed countries, there is clear evidence that travel patterns vary by demographic group in less economically developed countries too (Salon & Aligula, 2012; Salon & Gulyani, 2010). Studies of human movement patterns and malaria infection risk in East Africa have found that movement patterns vary by age group, gender and urban/rural setting, and that these differences in movement patterns are linked to differential rates of malaria parasite prevalence (Pindolia et al., 2013). These examples suggest that disaggregating the mobility of a population by at least some demographic characteristics could help to improve our understanding of population mobility as a whole and also improve the underlying mechanics of vector-borne disease models that incorporate mobility.

2.4.2 Urban-rural differences in human mobility

The difference between rural and urban mobility is one of the many factors that can affect mobility patterns of populations, and there is a considerable body of evidence for differences in movement patterns between these areas. In 1983, Chapman and Prothero highlighted differences between urban and rural human mobility in sub-Saharan Africa, including the different activities people carry out in these contexts and how they might link to differing risks of infectious disease exposure (Chapman & Prothero, 1983). Since then, many studies have looked at differences in urban and rural mobility, and the definitions of these areas have expanded to include a range of different areas in between, such as suburban and peri-urban depending on the context. Unsurprisingly, differences between urban and rural mobility are not always clear-cut. A study comparing urban and rural movements in the USA saw that while mobility was overall higher in rural areas due to a lower density of places of interest, both age and household wealth were also important drivers of mobility inequality (Pucher & Renne, 2005). A study of migrant workers in East Africa found that while urban migrant residents in Kenya and Tanzania had higher connectivity compared to their rural counterparts, urban migrant residents in Uganda had lower connectivity. Moreover, these urban-rural differences, combined with different demographic factors, affected their malaria risk. This study highlighted the importance of stratifying human movements by demographic subgroups and urban-rural context to determine malaria risk (Pindolia et al., 2013).

Although much of the data on the daily patterns of movement in an urban environment comes from developed countries, there is plenty of evidence to indicate that different mobility patterns are observed in similar contexts. Studies in Iquitos, Peru, a resource-poor urban environment with a high population density, found that while people moved mostly near their households, far fewer of these movements were the result of predictable routines compared to those found in urban areas in other studies (Perkins et al., 2014; Vazquez-Prokopec et al., 2013). There are very few studies with this resolution of mobility in other urban areas of developing countries, and these examples illustrate the importance of the context in which movements are made, as well as the danger of making assumptions about mobility in different areas globally.

For vector-borne diseases it is important to understand how mobility varies between urban and rural areas because host and vector population distributions vary greatly between these areas. For example, across much of sub-Saharan Africa, malaria transmission has decreased in urban areas (Robert et al., 2003), while dengue fever is on the rise (Stanaway et al., 2016). Moreover, changing connectivity patterns may have different effects on the risk of introduction of vector-borne diseases depending on the context. Understanding how mobility varies in these different contexts in conjunction with demographic factors affecting mobility could aid disease control efforts by providing evidence on how movement patterns might be leading to a greater risk of vector-borne disease exposure for different demographic groups.

2.4.3 Activity-driven mobility and resource access

Studies in the literature have highlighted the importance of the ‘activity space’, or movements driven by the motivation to carry out different activities as another driver of mobility (Perkins et al., 2014; Stoddard et al., 2009). Activity spaces have been studied for decades in geographic contexts (see Figure 2-5, reproduced from Chapman & Prothero, 1983) and are particularly relevant to issues of resource access. In the context of vector-borne diseases, access to health facilities is of obvious interest, and the link between mobility and health facility access is well researched. For example, studies have used mobility data to explore the methods of transport people use to access health facilities (Tanser et al., 2006) and the demographic differences in healthcare-related mobility (Mumtaz & Salway, 2005). However, other types of resource-driven mobility such as access to treatment, household resources and education are also important for the control of many diseases. Few studies have linked mobility to other types of resource access in the context of vector-borne diseases, but nevertheless the importance of resource access for the control of many vector-borne diseases is clear, such as access to clean water and praziquantel treatment in areas with endemic schistosomiasis cases (Campbell et al., 2014) and access to treatment and education for malaria (Johnson et al., 2012; Mubyazi et al., 2010).

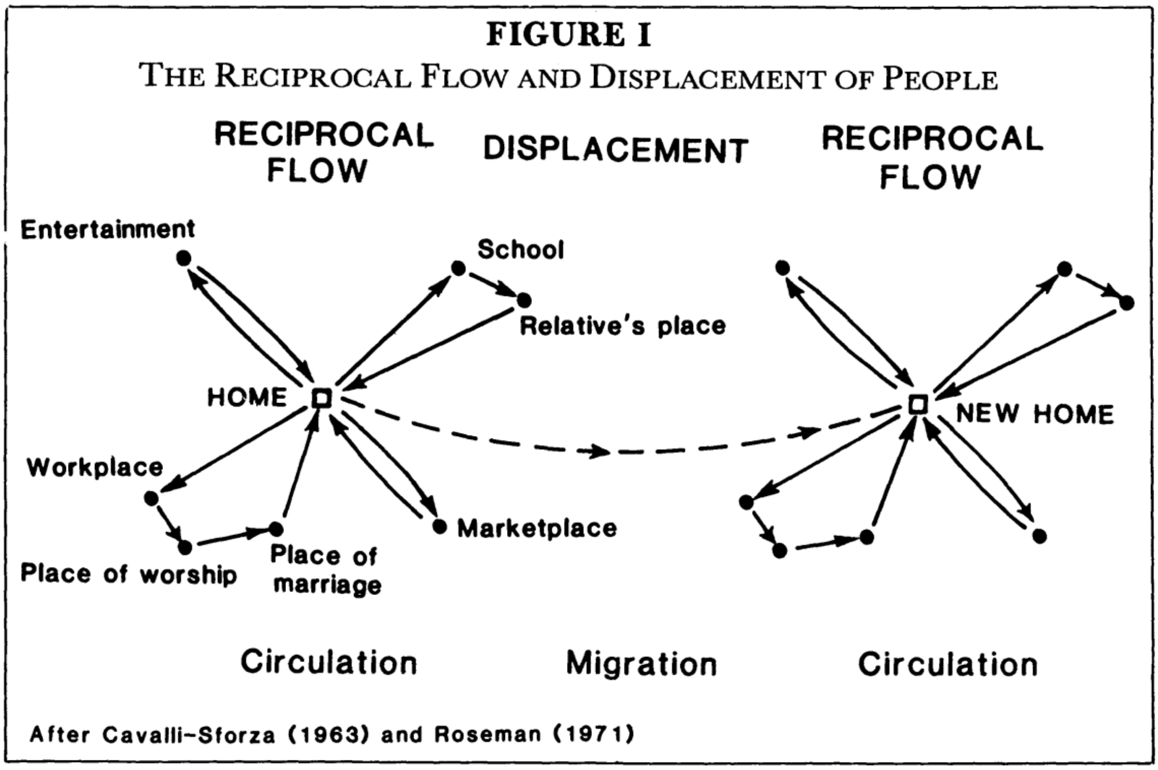


Figure 2-5 Schematic of an activity space ('reciprocal flow'), published by Chapman & Prothero, 1983.

People often travel for specific resource needs or activities such as gathering food and water, livelihood and occupational activities, or accessing healthcare. In low-income settings, travel to these resources can be very time-consuming or expensive meaning that people may forego healthcare, employment, or other resources. As a result, the geographic inaccessibility of vulnerable populations can lead to worse health outcomes, a poorer economic outlook, and can widen spatial inequalities (Alegana et al., 2018; Macintyre et al., 2019; Pearce et al., 2008). In less economically developed areas, resource access is particularly important for poverty reduction in rural populations, as people often have to travel further for resource-related activities than their urban counterparts. It is widely accepted that people in rural areas spend more time accessing resources than people in urban areas, and that this likely contributes to poverty in these areas. For example, market access is important for several household activities such as gathering food, selling crop surpluses and buying medicine, and studies have shown that poor market access contributes to poverty in rural areas (Chamberlin & Jayne, 2013). Similarly, poor water source access means reduced time for making money elsewhere and therefore contributes to household poverty, as well as being linked to poor health (Cook et al., 2016; Whittington et al., 1990). Geographic inaccessibility of healthcare is also a known driver of poor health outcomes, particularly in rural settings (A. M. Noor et al., 2003; C. W. Ruktanonchai et al., 2016). The accessibility of health facilities can be affected by several factors such as poor road quality and lack of public transport options, which can severely impact mobility and therefore healthcare access in rural areas where the most vulnerable populations live (Airey, 1992; Tanser et al., 2006).

Because geographic accessibility is vitally important for ensuring vulnerable populations can utilise healthcare, a significant body of recent research focuses on modelling accessibility across national scales. These models use accessibility surfaces or straight-line distances to predict clinics used and associated travel times, and therefore often assume that people visit their nearest clinic (Alegana et al., 2012). In reality, geographic accessibility is often highly heterogeneous and straight-line distance is not the only factor that impacts whether people can access healthcare in a reasonable time without undue expense. In a 1994 paper, Thaddeus and Maine proposed a framework to describe the factors affecting healthcare accessibility that contribute to maternal mortality in low and middle-income countries. They outlined a three-delay model (Figure 2-6), describing three stages at which barriers and delays to seeking and obtaining healthcare may be faced, and illustrating how several factors can contribute to delays in these contexts (Thaddeus & Maine, 1994). In Kenya, these factors are important to how people access healthcare: studies here have found that people often visit clinics other than their nearest one, for reasons such as the availability of medicines or the perceived effectiveness of the facility (G. M. Mwabu, 1986). One study found that only 54–63% of people surveyed visited their nearest facility, with the rest visiting clinics further away (G. Mwabu et al., 1995). Efforts to improve healthcare accessibility in certain areas could benefit from taking these factors into account but require robust evidence to understand these factors in context.

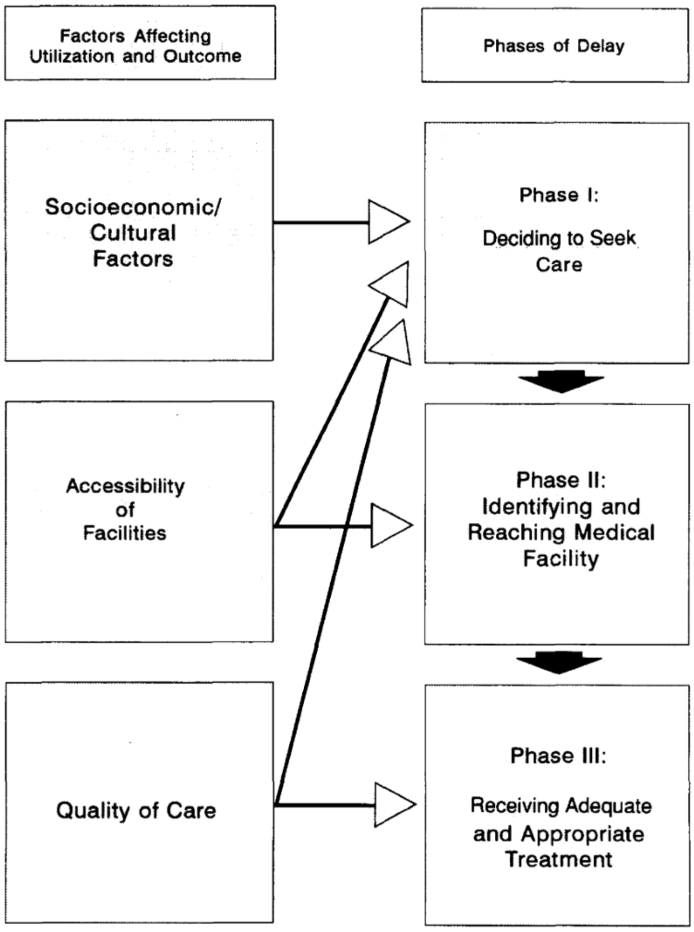


Figure 2-6 The three-delay model, as published by Thaddeus & Maine, 1994.

Moreover, resource-related movement is quantitatively different from other types of mobility, so accessibility models that predict general mobility patterns may not accurately reflect resource seeking behaviour. For example, a recent study in Iquitos, Peru found that residents moved significantly further for commercial and familial reasons than for healthcare (Perkins et al., 2014). Sociodemographic factors could also influence mobility and resource-specific movement: household income, rural/urban context (C. S. Molyneux et al., 1999) and gender are all important determinants of mobility, but spatial models of access often lack this demographic information, instead assuming that all adults access resources such as markets and health facilities identically, regardless of socioeconomic context. More detailed movement datasets could help improve these models by providing evidence to support or reject these key assumptions and better understand how movements driven by different activities vary.

2.4.4 Seasonal factors affecting mobility

Seasonal patterns of movement occur in human and animal populations worldwide due to combinations of economic, social and climatic factors. People make seasonal movements dependent on environmental factors that change with the time of year (for example, to access resources like water, food, livestock markets or grazing land) (Majekodunmi et al., 2013), they may move for work that varies with economic activity (for example, for jobs in industries like agriculture that are more active at certain times of year) (Hampshire, 2015), and social factors such as local holidays can also cause seasonal movements. Similarly, livestock and wildlife movements can vary seasonally, and movement patterns of some species are particularly sensitive to seasonal changes in landscapes. The distribution of vector-borne diseases carried by these hosts, such as malaria and trypanosomiasis, also tend to follow seasonal patterns because of their inherent sensitivity to environmental factors (Craig et al., 1999; Nnko et al., 2017). Thus, it is widely acknowledged that the seasonal movement patterns of hosts can have implications for disease control efforts (Amy Wesolowski, Metcalf, et al., 2015). For instance, disease control programmes can take advantage of seasonal patterns by implementing strategies like mass drug administration when transmission intensity is highest (Gao et al., 2014). Another factor affecting movement patterns with strong environmental links is the local landscape. This can be particularly important in rural areas, which generally have a larger diversity of land types. Indeed, several studies have found that movement patterns can vary between different ecological areas, and that time spent by hosts on particular types of land can affect their risk of vector-borne disease exposure (Chevalier et al., 2004; Fornace et al., 2019).

Seasonal variations in movement patterns can be seen at both ends of the spatiotemporal scale and as previously discussed, both fine-scale movements within a local area and larger-scale, long distance movements can result in differential exposure to vector-borne diseases. The interaction of these variations in movement with seasonal factors can also be important for vector-borne disease dynamics. At the fine spatial scale, short-distance movements of people within their local areas may expose them daily to vector-borne diseases that have a higher transmission intensity in certain seasons. Studies in Borneo have illustrated this, observing that malaria vector densities can vary spatially and seasonally in different ecological areas and that individual human risk of exposure to zoonotic *Plasmodium knowlesi* malaria is highest when human movements span these areas (Fornace et al., 2019; Wong et al., 2015). Although these particular studies found no significant seasonal differences in human movement patterns when only anopheline biting times were considered, they found marked differences in movement patterns between seasons across all hours measured, meaning that seasonal variations in movement could still play a part in differential exposure to other infectious diseases. Researchers have long been aware of the potential for seasonal movements between different ecological areas to result in increased exposure to vector-borne diseases such as trypanosomiasis (Mansell Prothero, 1963; R. Mansell Prothero, 1977) and malaria (Imai et al., 2014; N. W. Ruktanonchai et al., 2016), but studies to quantify these movements are lacking due to difficulties in collecting appropriate data.

At the opposing end of the spatial scale, movements over large distances can also vary seasonally and have implications for infectious disease dynamics. For example, a study in several countries using CDRs to quantify national human movement patterns and model the spread of a novel pathogen found that when seasonal movement patterns in each country were taken into account, the month of pathogen introduction into a country had a significant effect on the speed of its spread (Amy Wesolowski et al., 2017). For malaria, several studies have identified the relatively long-term and long-distance movements of seasonal migrant workers as contributing to malaria transmission in areas where it otherwise might not be sustained (R. M. Prothero, 2001; Schicker et al., 2015). Furthermore, because seasonal movements cause population densities to fluctuate on both small and large scales, measuring these movements can improve our knowledge of population denominators. Spatiotemporal changes in population densities have been shown to explain seasonality in incidence estimates of diseases like measles (Bharti et al., 2011) and malaria (Zu Erbach-Schoenberg et al., 2016). Overall, it is clear that individual movement patterns can affect vector-borne disease dynamics across the spatiotemporal scale, and data to quantify these seasonal patterns could provide valuable insight into the factors driving vector-borne disease transmission.

Chapter 3 Study design

In this Chapter, I present the theoretical framework, aims and research questions for each of the three papers in Chapters 4, 5 and 6; followed by an overview of the study areas chosen for the data collection and the methods used for data analysis. I conclude with a summary of knowledge gaps in the literature and the scientific and intellectual contributions of the thesis.

3.1 Theoretical framework

In Chapter 2, I discussed how the mobility of human and animal populations affects the transmission and distribution patterns of vector-borne diseases, and how these movement patterns can vary across different contexts. Because multiple host types can carry and transmit pathogens of vector-borne diseases, understanding the dynamics of zoonotic vector-borne pathogen transmission requires a research framework that studies both human and animal movements and their interactions with their local environments. As discussed in Chapter 2, these movements vary on spatial and temporal scales and also differ by urban or rural context. For many people living in rural areas affected by some of the highest-burden vector-borne diseases, domestic livestock are a critical resource and both humans and their livestock make regular fine-scale movements and are affected by zoonotic vector-borne diseases such as trypanosomiasis. We know that fine-scale movements differ to large-scale movements, and can affect vector-borne disease exposure differently, yet no studies have measured human and livestock movement simultaneously in the context of zoonoses or vector-borne diseases, so both the feasibility and value of quantifying human and livestock movements together is currently unknown.

In rural areas of sub-Saharan Africa, poor access to local resources such as water, markets and health facilities are linked to a greater burden of infectious diseases. Mobility studies in these areas typically focus on single types of resource access without comparing against travel to other types of resources (Kanuganti et al., 2015; Schröder et al., 2018). As discussed in Chapter 2, data on activity-dependent mobility in rural contexts could provide a richer picture of how people spend their time in these settings and therefore help us understand how movement patterns could exaggerate or mitigate geographic inaccessibility to resources that are critical to vector-borne disease prevention and control. Given the growing populations in rural areas of lower-income countries, detailed knowledge of activity spaces on a fine scale in a rural population could have benefits for better understanding specific types of resource access in similar populations. This is a fundamental issue in infectious disease research, because mobility is intrinsic to interactions with local infrastructure and therefore how people access healthcare and other resources can affect their vulnerability to vector-borne and other diseases (Bardosh et al., 2017).

On a larger scale such as at a country level, understanding the movement patterns of individuals in a large population can be of huge value for the control of widespread vector-borne diseases like malaria. The conditions required for malaria transmission vary spatially and temporally across landscapes and the populations that reside in them (Carter et al., 2000), and factors that determine the capacity of an area to sustain transmission include mosquito dispersal, insecticide resistance and other ecological factors (Ferguson et al., 2010; Killeen et al., 2003), local environment factors such as temperature fluctuations (Paaijmans et al., 2012) and housing quality (J. X. Liu et al., 2014), and human behavioural factors such as care-seeking (Källander et al., 2008) and population mobility (Pindolia et al., 2013). These factors cause local heterogeneity such that across Africa, malaria exists in areas with self-sustaining endemicity characterised by local disease transmission as well as transmission in areas that would be otherwise unable to sustain it (N. W. Ruktanonchai et al., 2016).

Since mosquitoes are unlikely to travel far, the malaria parasite is mainly transported between areas of high and low transmission by human hosts, allowing the parasite to be transmitted in areas where it may not be present otherwise. Malaria can be controlled and subsequently eliminated through combinations of reliable interventions, including insecticide-treated nets (ITNs) (Pryce et al., 2018), indoor residual spraying (IRS) (Kouznetsov, 1977) and mass drug administration (MDA) (Sarr et al., 2011), but knowledge of where and when to implement such interventions is reliant on our understanding of where transmission is occurring. Therefore, while identifying local hotspots of malaria can provide targets for intervention strategies, the movements of humans must also be accounted for in order to efficiently target efforts to control malaria. However, as discussed in Chapter 2, collecting high-quality data to provide evidence about where and when people and vector-borne parasites move on country-wide scales is challenging, and relatively few studies have been able to quantify population movements across a whole country and in the context of a vector-borne disease.

To address these gaps, I conduct three studies to provide insight into the movement patterns, activity spaces and risk of vector-borne disease exposure for humans and their livestock in two settings at opposing ends of the spatiotemporal scale using three different methods of mobility data collection. First, I use GPS loggers in an exploratory study to examine the fine-scale movements of a rural population and their livestock. I explore how mobility varies between people with different demographic characteristics and the potential exposure of both humans and their livestock to zoonotic vector-borne disease through time spent on different types of land cover. In the second study, I use primarily household and individual-level surveys to see how people in the same rural population make activity-driven movements and compare these mobility data to the previously mentioned GPS data in order to identify potential barriers to resource access in this area, which may be increasing individual risk to vector-borne diseases. In the third study I conduct a country-scale analysis of population movements in Mozambique. I use malaria incidence data to map risk of exposure and a CDR dataset to explore how individual movement patterns that affect exposure to malaria translate up to population-level malaria hotspots.

3.2 Study aims and research questions

3.2.1 Chapter 4

The first study in this thesis, entitled *Exploring fine-scale human and livestock movement in western Kenya*, is an exploratory study that aims to quantify the fine-scale movement patterns of a small population of humans and their livestock living in Busia county in western Kenya and explore links in these movement patterns on the local scale to environmental and demographic characteristics.

To do this, I collect GPS data and conduct surveys to understand where people and their livestock move to, how much time they spend in different parts of their landscape, whether these varied by season and whether certain household or environmental characteristics correlate with particular movement patterns. This study has been published in the journal *One Health* (Floyd et al., 2019).

The research questions for this study were:

- How often, for how long and where do humans and livestock move?
- Do human and livestock movements differ significantly between seasons?
- What demographic characteristics, if any, are linked to these movement patterns?

3.2.2 Chapter 5

The second study in this thesis, entitled “*Activity-specific mobility of adults in a rural area of western Kenya*”, aims to explore the activity spaces of the same sample population living in Busia county, western Kenya, identify demographic differences in access to local resources and assess the validity of GPS and survey-reported data on resource access at a local scale.

To do this, I use individual and household-level surveys to gather data on the activity spaces of people living in a rural area of western Kenya and compare these movement data to the GPS data gathered in the previous study. The results of this study help account for demographic differences in access to resources that may be linked to vulnerability to vector-borne diseases and, by comparing the movement patterns reported in the survey to those observed in the GPS data, provide further evidence for the utility of GPS loggers to quantify fine-scale movement patterns. This study has been published in the journal *PeerJ* (Floyd et al., 2020).

To address these aims I will answer the following research questions:

- How often do people visit different types of places in their local area, and how long do they spend there?
- Are demographic characteristics correlated with time spent at particular types of places?
- How well does the time spent at different places calculated from the GPS data reflect the time spent reported in the survey data?

3.2.3 Chapter 6

The aim of the third and final study of this thesis, entitled "*Estimating malaria parasite mobility to inform elimination strategies in southern Mozambique using mobile phone records*", is to quantify the national-scale mobility of the population of Mozambique and estimate malaria parasite movements across the country.

To achieve this, I use a large retrospectively collected dataset of CDRs to quantify human movements in Mozambique and combine these with high-resolution maps of malaria incidence to identify areas of high parasite importation and exportation within Mozambique to answer the following research questions:

- Where are human-mediated malaria parasite movements occurring in Mozambique? Specifically:
  - Which districts are the highest net importers and exporters of malaria parasites?
  - Are there differences in human-mediated malaria parasite movements between rural and urban areas?
- How connected are districts in Mozambique through human-mediated malaria parasite mobility?
- How much of the malaria incidence in districts in southern Mozambique is due to importation from other districts?

3.3 Study areas

3.3.1 Kenya

Vector-borne diseases have the greatest burden in tropical and subtropical parts of the world and are particularly prevalent across much of sub-Saharan Africa. In Kenya, a country lying on the equator in East Africa, there is endemic transmission of malaria, at least four mosquito-borne viruses and numerous other vector-borne pathogens (Eric M. Fèvre et al., 2017; Karungu et al., 2019). Figure 3-1 and Figure 3-2 show the distributions of several pathogens and their vectors across Kenya.

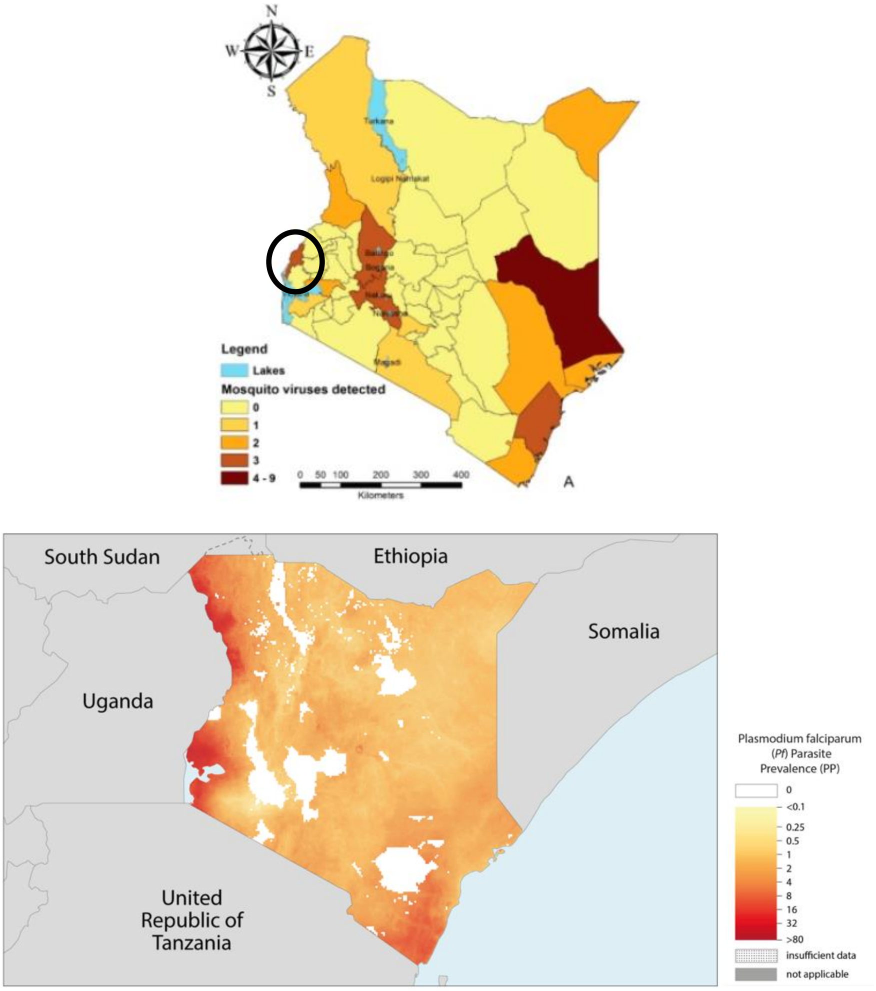


Figure 3-1 Abundance and distribution of selected mosquito-borne pathogens in Kenya.

Top: Abundance of 9 endemic mosquito viruses in Kenya. Dengue, Chikungunya, Ndumu and Sindbis viruses are present in Busia county (circled). As published by Karungu et al., 2009.

Bottom: Predicted *Plasmodium falciparum* prevalence in 2018. As published by the World Health Organisation, 2018a.

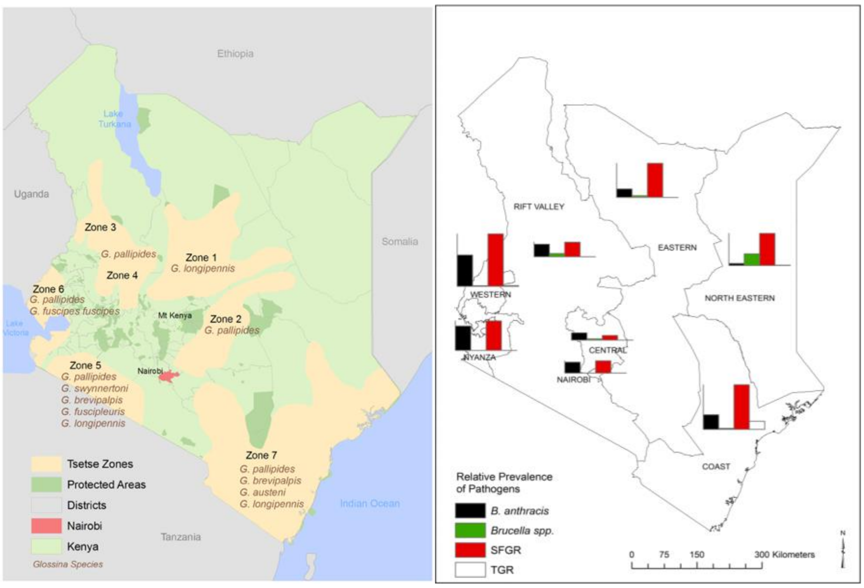


Figure 3-2 Distribution of selected zoonotic pathogens in Kenya

Left: Distribution of Tsetse fly habitats in Kenya. Busia county lies in Zone 6. As published by Grady et al., 2011.

Right: Relative seroprevalences of four zoonotic pathogens in the provinces of Kenya including the zoonotic vector-borne pathogen SFGR (spotted fever group rickettsioses). As published by Omballa et al., 2016.

These and other pathogens cause health-related issues that have been identified as important factors perpetuating poverty in Kenya. Distances to health facilities and other resources have also been identified as spatial determinants of poverty in Kenya, with higher poverty levels found in locations that are further away from good roads, small towns and district hospitals (Okwi, Ndeng'e, et al., 2007). Moreover, 84% of those living in extreme poverty are found in rural areas like Busia county (Kenya National Bureau of Statistics, 2018). These examples illustrate how in Kenya, factors that dictate local-scale movement patterns such as distance from health facilities have been shown to affect health outcomes, although large-scale patterns are also important to vector-borne disease exposure such as the strongly heterogeneous distribution of arboviruses across counties.

Busia county in western Kenya covers an area of 1,697km^2 (circled in top panel of Figure 3-1), which were administratively divided into 10 divisions, 60 locations and 181 sublocations until the administrative system was updated in 2010. The area has bi-annual rainfall, with the long rains occurring between March and May, the short rains between July and October, and a dry season between November and February (Thuranira-McKeever et al., 2010). Busia County has an estimated population of 894,000 people, a population density of 527 people per km^2 (Kenya

National Bureau of Statistics, 2019) and high densities of domestic livestock. Most of the human population in this area of Kenya relies on mixed crop-livestock farming for both food production and income (Kristjanson et al., 2004). Since up to 69.3% of people live below the overall poverty line of 3,252 KSh/$30 USD per month in rural areas or 5,995 KSh/$56 USD per month in urban areas, livestock play an important part in sustaining livelihoods. Busia county has the highest reliance on food of own production of any county in Kenya, with 38.8% of all food being of own production, and over half the population were living below the food poverty line (less than 1,954 KSh/$18 USD per month in rural areas and 2,552 KSh/$24 USD per month in urban areas) in 2015 (Kenya National Bureau of Statistics, 2018).

Data from the 2014 Demographic and Health (DHS) survey in Kenya suggest that 83% of rural households in Busia County own livestock. Of these households, 92% own chickens, 60% own cattle and 29% own goats and/or sheep (small ruminants) (Kenya National Bureau of Statistics et al., 2015). Figure 3-3 and Figure 3-4 show how people and cattle populations are distributed across Kenya, with particularly high densities in western Kenya. Because of the high reliance on livestock production systems, zoonotic diseases that reduce production and affect human and animal health exert a particularly high burden in this region.

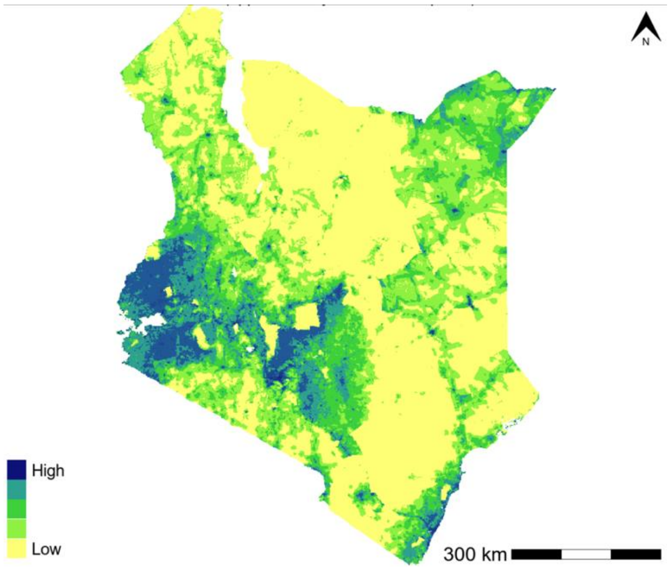
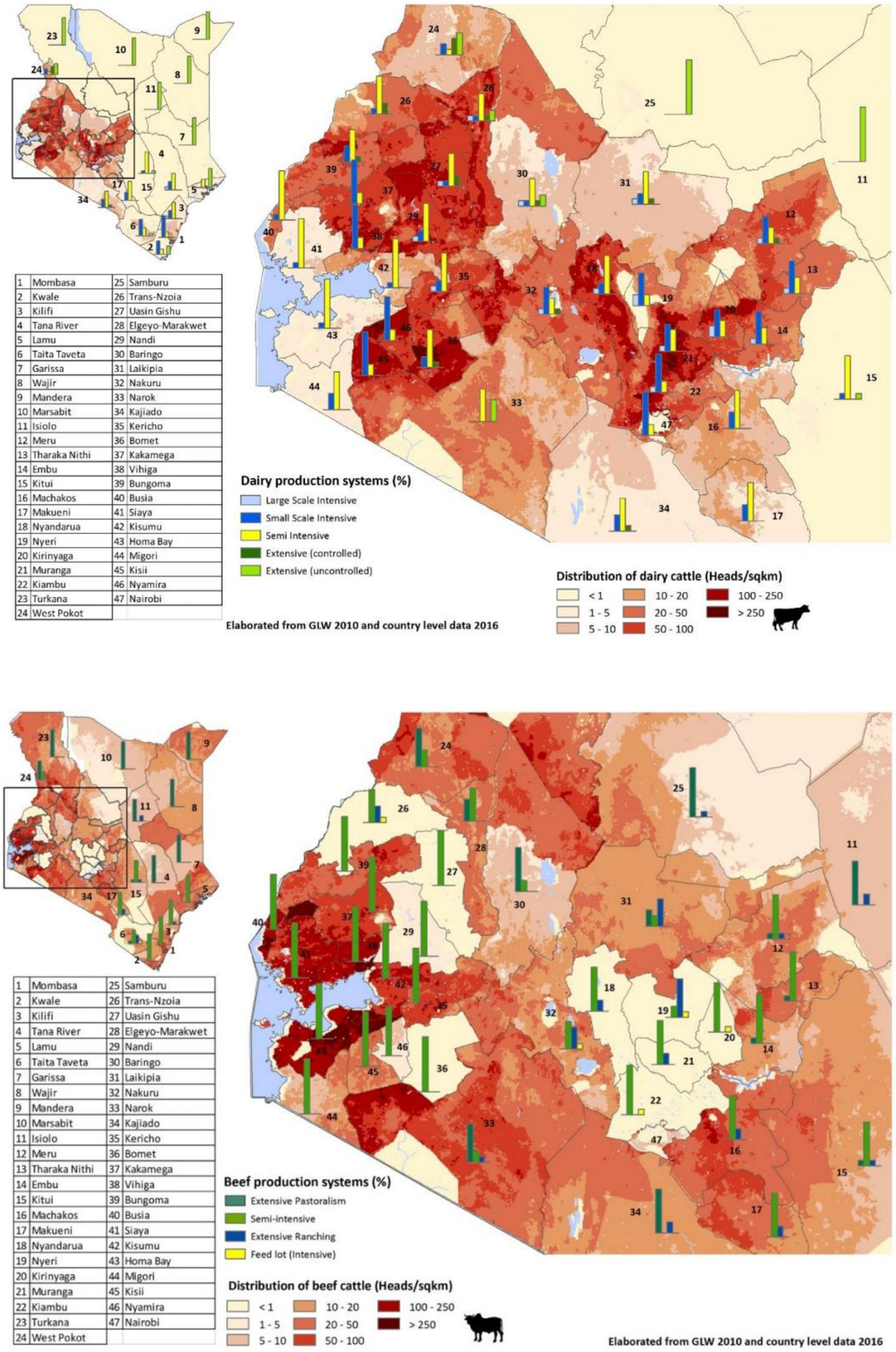


Figure 3-3 Relative human population density in Kenya, 2018. As published by WorldPop, 2018.



Elaborated from GLW 2010 and country level data 2016

Elaborated from GLW 2010 and country level data 2016

Figure 3-4 Cattle production systems and densities of dairy and beef cattle in Kenya.

Top: Dairy production systems and densities of dairy cattle in Kenya (top left), with a focus on western Kenya (right). Approximate location of Busia county marked with number 40.

Bottom: Beef production systems and densities of beef cattle in Kenya (top left), with a focus on western Kenya (right). Approximate location of Busia county marked with number 40.

Adapted from Food and Agriculture Organisation of the United Nations, 2018.

Data collection for Chapters 4 and 5 were conducted in Busia County, western Kenya (Figure 4-1), in collaboration with the Zoonoses in Livestock in Kenya (ZooLink) project run by the zoonotic and emerging disease research group from the University of Liverpool and the International Livestock Research Institute (ILRI).

3.3.2 Mozambique

Mozambique in Southern Africa has an estimated population of 28.9 million people, all of whom live in areas of 'high malaria transmission' as defined by the WHO (World Health Organization, 2018a). Mozambique is administratively divided into 11 provinces including the capital of Maputo city. The provinces of Nampula and Zambezia in the north of the country are the most populous with 6.1 and 5.1 million people respectively, followed by Tete in the north-west (2.8 million) and Maputo in the south with 2.5 million (National Institute of Statistics (INE) (Mozambique), 2017).

Malaria remains one of the biggest causes of death in Mozambique despite recent large reductions in burden. It is estimated that malaria was responsible for 10 million cases (5% of all cases worldwide) and 15,000 deaths in 2017, most of them due to Plasmodium falciparum (the major Plasmodium species in Mozambique, Figure 3-5). More broadly, high prevalence in Mozambique poses a challenge to the relatively low transmission countries nearby. For example, in 2017 the incidence of malaria in Mozambique was estimated to be 219 per 1000 people, while incidence in the bordering countries of Eswatini and South Africa was 2.34 and 2.64 per 1000 people respectively (Moonasar et al., 2016). Malaria prevalence also varies widely within Mozambique, ranging from 1% in the southernmost provinces of Maputo city and Maputo province, to 57% in the most northerly province of Cabo Delgado (Saúde/INS et al., 2019).

The heterogeneity in malaria prevalence in Mozambique has important implications for people living there. While climatic factors such as temperature and precipitation have been shown to affect spatial patterns of transmission (Abellana et al., 2008; Giardina et al., 2015), studies have shown that population mobility is also an important driver of malaria transmission. Across sub-Saharan Africa, malaria exists in heterogeneous areas, including 'hotspots' (areas that have a higher burden than surrounding areas) that are thought to seed transmission in other areas through human or mosquito mobility (N. W. Ruktanonchai et al., 2016; Stresman et al., 2019). In Mozambique, efforts are to control malaria have historically been concentrated in the more populous north of the country, while elimination is targeted for the south (Aide et al., 2019; Moonasar et al., 2016). The impact of population movements between and within these areas on the sustainability of these efforts is unknown.

A meta-analysis of studies from urban areas of sub-Saharan Africa estimated a mean annual entomological inoculation rate (EIR) of 7.1 in city centres, compared to 167.7 in rural areas (Robert et al., 2003) and it is widely accepted that the process of urbanization reduces transmission of malaria. With only 36% of the population of Mozambique living in urban areas (World Bank, 2018), population mobility between rural and urban areas in Mozambique may be an important driver of malaria transmission in urban areas. Identifying hotspots through population mobility and understanding how this mobility varies between rural and urban areas could help researchers and policy-makers identify routes of transmission and quantify the relative importance of rural and urban mobility in Mozambique.

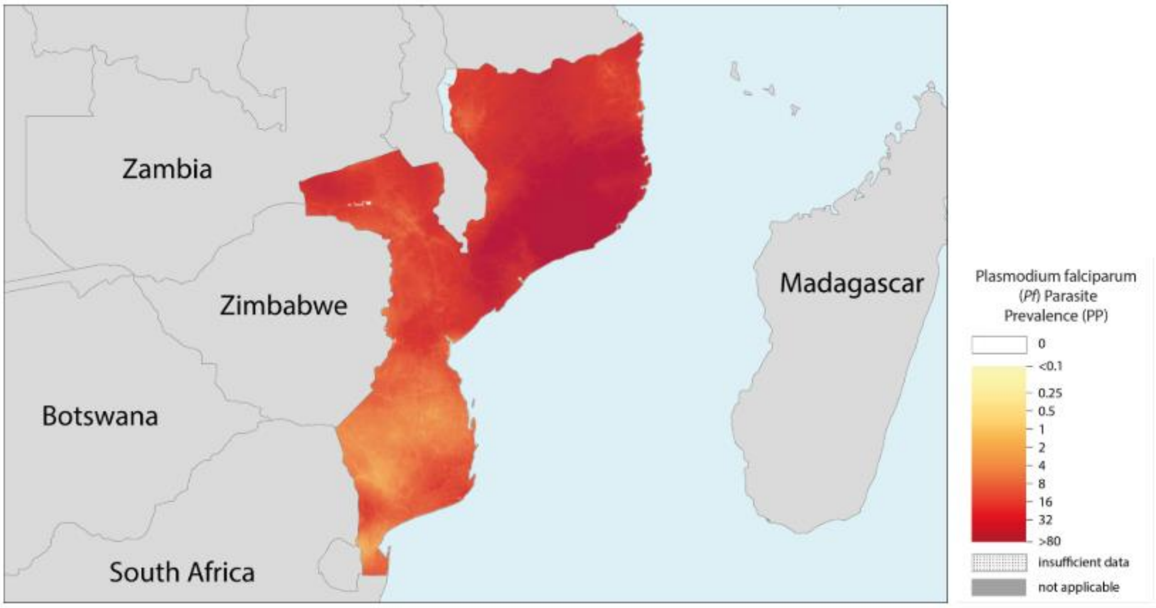


Figure 3-5 *Plasmodium falciparum* prevalence in Mozambique in 2017

As published by the World Health Organisation, 2018b.

3.4 Study methods

The studies presented in Chapters 4 to 6 employed a variety of methods to collect and analyse primary and secondary data. This section outlines the principal methods and datasets used for data collection and analysis in these chapters, including their advantages and limitations.

3.4.1 Household sampling and cross-sectional survey

Over the course of two periods of fieldwork, I aimed to visit approximately 55 households in Busia County. I conducted a survey at all households, and GPS tracking at a subset of these. Because there is little data on the movements of different human populations in this area, I chose the locations of these households using a simple random sampling strategy to select 11 of the 181 sublocations (the smallest administrative unit in Kenya) across Busia County in order to sample from a variety of different populations and landscapes. A previous study that used GPS tracking to explore pig movements in this region used a similar strategy to select one pig per sublocation (L. Thomas et al., 2013); based on this and given the larger individual variation in human movements, I estimated that five households per sublocation could provide useful information about human and livestock movements for the exploratory studies detailed in Chapters 4 and 5. Thus, in each of the 11 sublocations, five households were randomly selected for participation in the study, using random generation of coordinates within each sublocation (selected using the random sampling tool in QGIS), followed by identification of the household closest to the coordinates within 200 metres. Of the 11 sublocations, 6 were chosen at random to conduct the GPS tracking part of the fieldwork (described in 3.4.2). In one sublocation, only two households were sampled due to field team illness. Table 3-1 gives an overview of the survey and GPS data collected in each sublocation. At each sublocation, a village elder accompanied the field team to the households. They had no role or influence on household selection but facilitated the introduction and interview process.

Table 3-1 Overview of data collected in each sublocation

Sublocation ID	Sublocation name	Households visited	People surveyed	People tracked	Livestock tracked
1	Nangoma	5	8	5	5
2	Mundika	5	8	5	1
3	Nakhakina	5	8	0	0
4	Nambale	5	7	0	0
5	Sisenye	5	9	5	4
6	Angorai	2	3	2	2
7	Alupe	5	6	5	3
8	Kakapel	5	5	5	5
9	Kingandole	5	8	0	0
10	Aludeka	5	7	0	0
11	Nanderema	5	6	0	0
Total		**52**	**75**	**27**	**20**

This method of clustered random sampling has several limitations. First, this strategy is likely to bias the sample towards rural households, since the population is not equally distributed across the study area. The small sample size means that statistical power will be limited in data analysis. The presence of village elders may cause participants in the survey to modify or censor their answers to the questionnaires. Given the exploratory nature of the studies, the small sample size and lack of data in the literature on the movements of people from rural households in resource-poor settings, I concluded that these were acceptable limitations.

At each household, all eligible (adults aged 18 or over) and consenting participants were enrolled for the questionnaires. The questionnaires were written in English using OpenDataKit software (Brunette et al., 2013) and translated into local languages before administration. A household questionnaire was administered to the head of the household (identified by oral communication) and individual questionnaires were administered to the rest of the participants. Answers were recorded on a tablet, which also recorded the GPS location of the household. The questionnaire was designed based on previous studies in this area (Eric M. Fèvre et al., 2017; Lian F. Thomas et al., 2013) and included questions to determine demographic characteristics, household characteristics, frequent and infrequent movements, and livestock-related activities. The household and individual questionnaires are given in Appendix A.2.

3.4.2 GPS tracking

To collect high-resolution data on the day-to-day movements of people and their livestock, I used GPS loggers (iGot-U GT600, (Mobile Action Technology Inc., Taipei, Taiwan, 2018) mounted on a lanyard (for the human participants) or a collar (for the livestock), set to record the location of the device as coordinates every 60 seconds. This time interval was chosen to provide a high degree of temporal resolution while maintaining the battery life of the device. The GPS tracking was carried out in two periods spanning July-August and November-December of 2016 in order to compare data across two seasons. The protocol surrounding the use of the GPS devices were guided by previous studies. For example, Parsons and colleagues used a similar device to track the movements of dogs, sheep and goats for 8 days in a village setting in Tanzania across two seasons (Parsons et al., 2014). In the study presented here, participants and their livestock wore the devices for one week and efforts were made to capture data from the same participants in both seasons. The deployment, maintenance and collection of GPS devices was highly labour-intensive, so participants were given the devices to wear for 7 days in order to maximise the number of study participants while prioritising the limited resources of the field team. Previous studies tracking human and animals movements using similar GPS loggers at a fine scale tracked the participants for time spans ranging from 2 days to one month (Parsons et al., 2014; Paz-Soldan et al., 2014; Pray et al., 2016; Lian F. Thomas et al., 2013; Vazquez-Prokopec et al., 2013). Although 7 days was at the short end of this range, I decided that this was acceptable and necessary to maximise the sample size. The relatively short tracking window was limited by the battery life of the device and is likely to make the results susceptible to chance events, particularly in the human populations as more variation in their movements is expected.

3.4.3 Mixed effects modelling

Chapters 4 and 5 used mixed effects models to determine the relationships between demographic characteristics and movement metrics derived from the GPS and survey data. Because of the hierarchical nature of the data collected (individuals nested in clusters, with repeat measurements per individual), I used linear mixed models (LMMs) and generalized LMMs (GLMMs) for the multivariable analyses. These types of models are powerful tools for analysing normal (in a LMM) and non-normal data (in a GLMM) with both fixed and random effects (Bolker et al., 2009). Previous studies using similar GPS datasets have used GLMMs to account for fixed and random effects typically found in study designs using hierarchical sampling (Quaglietta et al., 2012; Schwemmer & Garthe, 2011).

For the mixed models, response variables were movement metrics derived from the GPS and survey data. These included the time spent at different places and on different types of land, frequency of visits to places, distances travelled on trips and home ranges. Explanatory variables were selected based on significance in univariable linear regression against each of the measured movement metrics and *a priori* plausibility. These included age, gender, occupation, season, household wealth, ruminant ownership and number of ruminants owned. All mixed models had the individual household nested within the sublocation as a random effect, to account for variation between individuals from different households and within different sublocations. For response variables that were proportions of time spent in different places, I used a beta regression family in a GLMM to obtain odds ratios for the different effects.

3.4.4 CDR analyses

The analyses comprising Chapter 6 utilise a large CDR dataset to quantify the national-scale movements of individuals across Mozambique. These movements are mapped in an origin-destination matrix, and a high-resolution (5km) gridded malaria incidence dataset is used to estimate human-mediated parasite movements (the estimated movements of the population, weighted by malaria incidence at the origin) across Mozambique. Previous studies have used similar methods to quantify human-mediated malaria parasite movement in other countries in sub-Saharan Africa (Le Menach et al., 2011; N. W. Ruktanonchai et al., 2016; Amy Wesolowski et al., 2012). It is widely acknowledged in the literature that estimating movements from CDRs likely suffer from biases due to heterogeneities on mobile phone ownership and usage, and therefore may overestimate mobility in some areas. However, a notable study by Wesolowski *et al.*, 2013 was able to link a CDR dataset from almost 15 million people in Kenya to a data source on their socioeconomic status and found that although mobile phone ownership was heterogeneous across the country and biased towards wealthier households, the CDRs were able to sufficiently represent distributions of human mobility, including sections of the population without mobile phones. The specific protocol used to estimate human and parasite movements including the advantages and limitations of this particular CDR dataset is described in more detail in Chapter 6 of this thesis.

I also implemented a community structure algorithm to determine inherent community networks through human-mediated parasite movements across Mozambique. The *cluster_walktrap* algorithm in the R package *igraph* (Csardi G, 2006) identifies densely connected communities through a series of random walks (Pons & Lapaty, 2006), such that the number of movements tends to be higher within communities than between them. This approach has an advantage over other algorithms in its speed of computation while accurately describing communities that are highly connected in a weighted network. The use of community structures for identifying communities connected through human-mediated malaria parasite movements has been previously implemented (Strano et al., 2018; Andrew J Tatem et al., 2014) and found to be a useful technique for defining coherent areas of coordination and for targeting surveillance efforts and resources for malaria interventions. In Chapter 6 I present the first use of this approach for a CDR dataset from Mozambique.

Lastly, I estimated the scale of human and human-mediated parasite movements between rural, peri-urban and urban areas using data from the Global Human Settlement Model (GHS-SMOD, (Pesaresi et al., 2019). This dataset is high-resolution (1km gridded), using rigorous methods to classify cells based on settlement type, thereby providing a measure of the degree of urbanization of an area. I quantified movements between areas falling into three different categories of urbanization for human movements alone (the estimated movements of the whole population) and for human-mediated parasite movements in order to compare differences between the two. Chapter 6 details the first use of this approach to quantify human movements between areas of differing urbanicity in Mozambique.

3.5 Scientific contribution

3.5.1 Knowledge gaps

A prominent paper by Woolhouse and colleagues in 1997 described how individual variations in contact patterns can have a profound effect on vector-borne disease transmission. Using a series of mathematical models for different vector-borne diseases, they showed that individual heterogeneities in contact can cause substantial increases in transmission, highlighting the existence of the “80/20 rule”, which purports that 80% of the transmission potential is made by 20% of the population (Woolhouse et al., 1997). In 2009, Stoddard *et al* built on this work, establishing a conceptual framework to investigate how individual movements affect transmission of the mosquito-borne dengue virus and finding that variations in movement patterns can change the relationship between vector abundance and the basic reproduction number (R_0), confirming that variations in movement patterns influence vector-borne disease dynamics. They noted that although human movements are key to vector-borne disease exposure, they remain poorly understood in this context (Stoddard et al., 2009). These studies underline the importance of collecting evidence to account for individual variations in movement patterns, as many disease models still rely on the assumption of spatially homogenous transmission, yet the factors that drive movement pattern variations are still understudied. This calls for a deeper understanding of the factors that drive individual heterogeneities in movement in order to develop targeted interventions.

Lambrechts and colleagues have called for studies to utilise spatially-explicit techniques such as GPS technology to understand how individual variations in movement patterns at a fine scale may affect the transmission of vector-borne diseases (Lambrechts et al., 2009). Stoddard and colleagues noted that although GPS technology has been available for several decades with the capacity to measure movements, technical limitations have constrained their use. However, recent improvements in portability and battery life have made their use more feasible and cost-effective for capturing fine-scale movements over longer periods of time and so several recent studies have used wearable GPS loggers to measure the movements of people in the context of infectious diseases across different settings (Searle et al., 2017; Vazquez-Prokopec et al., 2013). Separately, similar study designs and wearable GPS technology methods have been used to study the movements of livestock in the context of infectious diseases (Barasona et al., 2014; Parsons et al., 2014; Lian F. Thomas et al., 2013) but none have quantified the movements of humans and their livestock together. Studying movement patterns of humans and livestock simultaneously in this way would help to identify important factors that may be driving movement patterns that increase exposure to vector-borne diseases, such as movements to areas of high vector abundance.

Stoddard and colleagues also note the importance of studying individual activity spaces, as both the time spent at and the movements between locations of interest can affect personal exposure to vector-borne diseases (Stoddard et al., 2009). At the same time, studies have highlighted that poor access to resources such as markets and health facilities can perpetuate poverty and affect individual and household vulnerability to infectious diseases, particularly in low-resource regions (Bardosh et al., 2017). Few studies have attempted to quantify this relationship, partially due to difficulties pinpointing what activity might be occurring at a given location, therefore Stoddard et al recommend the use of methods that combine GPS technology with recall techniques to address this issue. However, before the relationship between resource access and individual vulnerability to vector-borne disease can be quantified, high-quality datasets on the activity spaces of individuals within populations are required in order to build a picture of activity spaces and understand potential barriers to resource access.

Finally, Stoddard et al also note that mobile phone data may be useful for measuring movements, but are better for more economically developed countries where coverage is high. In recent years, mobile phone penetration rates in sub-Saharan Africa have increased substantially, making studies that utilise mobile phone data more viable (Global System for Mobile Communications (GSMA), 2019). These datasets offer unique insight into individual movement patterns at a large spatial scale, and therefore when combined with similarly high-resolution data on malaria cases they are particularly suited to determining ‘hotspots’ or areas at increased risk of malaria transmission. Bousema and colleagues argue that targeting malaria hotspots is an efficient way to prioritize resources and is the next logical step for areas where transmission continues despite good coverage of interventions (Bousema et al., 2012). However, in a detailed review Stresman and colleagues consider the impact of hotspots on the epidemiology of malaria and conclude that while they are an intrinsic part of transmission, evidence to support hotspot-targeted interventions is limited and this is likely due to confounding factors like population movements obscuring our ability to develop suitably targeted interventions. They identify population movements as a key question to understanding the dynamics of malaria hotspots (Stresman et al., 2019). Previous studies have gone some way to answering this question (N. W. Ruktanonchai et al., 2016; Amy Wesolowski et al., 2012), but to date none have specifically explored how population movements affect malaria transmission in Mozambique.

In conclusion, although many authors call for multidisciplinary and cross-scale research (Pindolia et al., 2012; Vazquez-Prokopec et al., 2013) into host movements and how they relate to different vector-borne diseases in different contexts, there is a gap in the scientific literature in terms of studying mobility at local, regional and national scales. Although this thesis does not examine both in the same country, it does answer similar research questions at the local and national scales and therefore contributes to furthering our understanding of host mobility and vector-borne disease risk and exposure across these scales.

3.5.2 Contributions to the literature

This thesis aims to address these knowledge gaps in the scientific literature at the interface between human mobility and vector-borne diseases, particularly concerning movements across spatial scales that may be linked to increased exposure or vulnerability to vector-borne disease.

The exploratory study in Chapter 4 demonstrates the feasibility of using GPS loggers in conjunction with survey methods to track the movements of people and their livestock simultaneously in a resource-poor area. It also goes some way towards improving our understanding of the variations in movement patterns of individuals, how these may affect exposure to vector-borne disease and how they link to various demographic characteristics in a resource-poor rural area of Kenya, although further evidence is needed to verify the findings due to the limited representativeness of the sample population.

The analyses presented in Chapter 5 build on this contribution by linking GPS and survey data to demonstrate how individual activity spaces vary with several demographic characteristics. Building on methods outlined by (Perkins et al., 2014; Stoddard et al., 2009) for the quantification of activity spaces in relation to vector-borne diseases, these results constitute the first empirical analysis of the activity spaces of a rural population using both GPS loggers and survey methods and therefore demonstrate the feasibility of these methods for answering specific questions about activity-driven movements at a fine scale, while also highlighting the importance of disaggregating movement patterns by demographic characteristics in order to understand how they affect people's vulnerability to vector-borne diseases differently.

The analyses comprising the final study in Chapter 6 aim to address the question of how population-level movements sustain hotspots of malaria transmission and provide evidence to help account for movements when designing spatially targeted interventions against malaria in Mozambique, particularly in light of ongoing elimination efforts in southern Mozambique. These analyses also contribute further evidence to the literature about the value of large-scale CDR datasets and could potentially encourage future collaborations with telecommunications providers for addressing challenges posed by other infectious diseases such as dengue fever.

This work takes significant steps towards understanding host mobility in the context of vector-borne diseases across scales. For example, in Chapter 4 I find that rural livestock-keeping populations living in Busia county Kenya may be especially vulnerable to vector-borne diseases, and similarly in Chapter 6 I describe how rural populations are likely at higher risk of malaria exposure through their movements than urban populations, demonstrated by the higher numbers of malaria cases moving between rural areas and in the direction of increasing urbanicity. In Chapter 5 I outline the time spent at different places on the local scale that could contribute to differing vulnerability to vector-borne diseases, while in Chapter 6 I observe similar patterns at the national scale, with time spent in different districts linked to differences in malaria importation patterns. Ultimately, I expect that the analyses comprising this thesis will provide much-needed information on the movement patterns of populations in resource-poor areas, highlighting the importance of understanding movement patterns in the context of vector-borne diseases across spatial scales. Additionally, this thesis contributes methodological evidence to confirm the feasibility of using GPS technology to quantify mobility and the value of large-scale CDR datasets for near real-time analysis of movement patterns.

3.5.3 Intellectual contributions

JRF conceptualized the study framework and methods used throughout this thesis, and performed all fieldwork, analysis, and writing of the final thesis and associated manuscripts. JRF also generated all study visualizations and graphs throughout this work. NWR and AJT supervised JRF throughout study conceptualization and application, while other researchers and postgraduate students in the ZooLinK and WorldPop research groups provided feedback on methodological applications and procedures. Other co-authors and collaborators involved in publication of manuscripts comprising this work, along with their specific contributions, are listed individually within the ‘Intellectual contributions’ sections for each corresponding chapter.

Chapter 4 Exploring fine-scale human and livestock movement in western Kenya

4.1 Abstract

Human and livestock mobility are key factors in the transmission of several high-burden vector-borne zoonoses such as rift valley fever and trypanosomiasis, yet our knowledge of this mobility is relatively poor due to difficulty in quantifying population-level movement patterns. Significant variation in the movement patterns of individual hosts means it is necessary to capture their fine-scale mobility in order to gain useful knowledge that can be extrapolated to a population level. Here I explore how the movements of people and their ruminants, and their exposure to various types of land cover, correlate with ruminant ownership and other demographic factors which could affect individual exposure to zoonoses. The study was conducted in Busia County, western Kenya, where the population are mostly subsistence farmers operating a mixed crop/livestock farming system. I used GPS loggers to collect movement data from 27 people and their ruminants for one week per individual in July/August 2016, and the study was repeated at the end of the same year to compare movement patterns between the short rainy and dry seasons respectively. I found that during the dry season, people and their ruminants travelled further on trips outside of the household, and that people spent less time on swampland compared to the short rainy season. These findings also showed that ruminant owners spent longer and travelled further on trips outside the household than non-ruminant owners, and that people and ruminants from poorer households travelled further than people from relatively wealthier households. These results indicate that some individual-level mobility may be predicted by season and by household characteristics such as ruminant ownership and household wealth, which could have practical uses for assessing individual risk of exposure to some vector-borne zoonoses and for future modelling studies of zoonosis transmission in similar rural areas.

4.2 Introduction

Zoonoses cause substantial morbidity and mortality in human and animal populations across the world, threatening less economically developed countries due to the lack of resources needed to efficiently detect and respond to disease events. Because a significant proportion of people are dependent on livestock for their livelihoods in the rural areas where many of the world's poorest communities reside, these diseases are a major factor in perpetuating poverty (Maudlin et al., 2009; D. H. Molyneux, 2014; WHO & DFID, 2006). Zoonotic transmission is ultimately dependent on contact between hosts, whether direct human-to-animal contact or indirect contact through, for example, environmental contamination or vector-borne transmission (Arthur et al., 2017; L. Thomas et al., 2013). Zoonoses that are spread directly from host to host include zoonotic influenza and rabies, others like taeniasis/cysticercosis and brucellosis are primarily spread through contact with infected animal products; and vector-borne zoonoses such as trypanosomiasis and Rift Valley Fever (RVF) can be spread though the bite of infected insects. All these types of contact are driven by the movements of human and animal hosts, since contact (and therefore, pathogen movement) is intrinsically dependent on host movement (Collinge & Ray, 2006).

We know that environmental factors are important in the emergence and maintenance of many zoonotic diseases, some of which are vector-borne (Arthur et al., 2017; Peter et al., 2005; Taylor et al., 2016). However, suitable environmental conditions are important but not sufficient for zoonotic disease spread, as dynamic host populations must also interact through movement and behaviour to facilitate transmission. Host movements across landscapes can lead to the emergence of pathogens in new hosts and environments and contribute to maintenance of endemicity in old ones (Chevalier et al., 2004; Eric Fèvre et al., 2006). Thus, many studies have examined host movement between environments in relation to zoonotic disease spread. For example, vector-borne disease transmission is highly heterogeneous amongst both individual hosts and the landscapes that they move in (Chevalier et al., 2004; de La Rocque & Formenty, 2014; Messina et al., 2015) and outbreaks of vector-borne diseases such as RVF are known to be sensitive to both host movement and landscape characteristics, with factors such as host densities and movement patterns contributing to disease maintenance (Chevalier et al., 2004; Martini et al., 2008).

Understanding the movements of people and their livestock, how they interact with their environments and how these correlate with population characteristics could provide crucial information to aid the control of zoonoses, yet few studies have measured these movements simultaneously. Human mobility is important to zoonotic disease risk because it informs where people may spend time with livestock or wild animal populations, and animal mobility is important because it informs where animals may be a risk of exposure from other animals or the environment, but transmission ultimately occurs when hosts meet, whether directly or indirectly. By measuring mobility simultaneously, it is possible to map when, where, and during what kinds of mobility animal-human interactions occur for both populations. For example, knowledge of how people and their livestock move together could help quantify the amount of contact between them, which is a risk factor for transmission of zoonoses such as *Cryptosporidium spp.* (Stantic-Pavlinic et al., 2003) and zoonotic *Escherichia coli* (Locking et al., 2001). Moreover, the land utilisation of hosts is also related to exposure to some vector-borne zoonoses: for example, the time spent on different types of land has been shown to affect individual risk of infection of Crimean-Congo haemorrhagic fever (CCHF) (Messina et al., 2015). Knowledge of host movement patterns therefore could aid the design of more targeted interventions and behavioural modifications that take into account high risk activities for zoonosis transmission.

Here I present an exploratory study with the aim of measuring the movement patterns of humans and ruminants in a region of western Kenya where several important zoonoses including trypanosomiasis, cysticercosis and Q fever have been previously shown to be endemic (Lian Francesca Thomas et al., 2016; von Wissmann et al., 2011; Wardrop et al., 2016). I collected GPS data (satellite-informed locations taken at regular intervals) from humans and their ruminants. From these, I calculated measures including the home ranges of individuals, the time they spent on different types of land, and the time spent, distance travelled and frequency of trips outside of the household. By exploring how mobility varies between populations with different demographic characteristics, the study shows how GPS data can provide useful information on their movement patterns. Further studies using similar methods could verify the results shown here and thus help inform how future vector-borne zoonotic disease interventions are targeted.

4.3 Methods

4.3.1 Study area and sampling procedure

The study was conducted in Busia County (Figure 4-1) in the Lake Victoria basin region of western Kenya (Eric M. Fèvre et al., 2017). The human population is just over 894,000 people (Kenya National Bureau of Statistics, 2017) mainly composed of subsistence farmers operating mixed crop-livestock farming systems. According to the latest DHS survey, 89% of people in Busia live in a rural area (as defined by the national census) and 83% of rural households own livestock (Kenya National Bureau of Statistics et al., 2015). Of these, 60% own cattle (Kenya National Bureau of

Statistics et al., 2015) usually kept on a mixture of tethered and free grazing systems on common grazing lands (Bronsvoort et al., 2013).

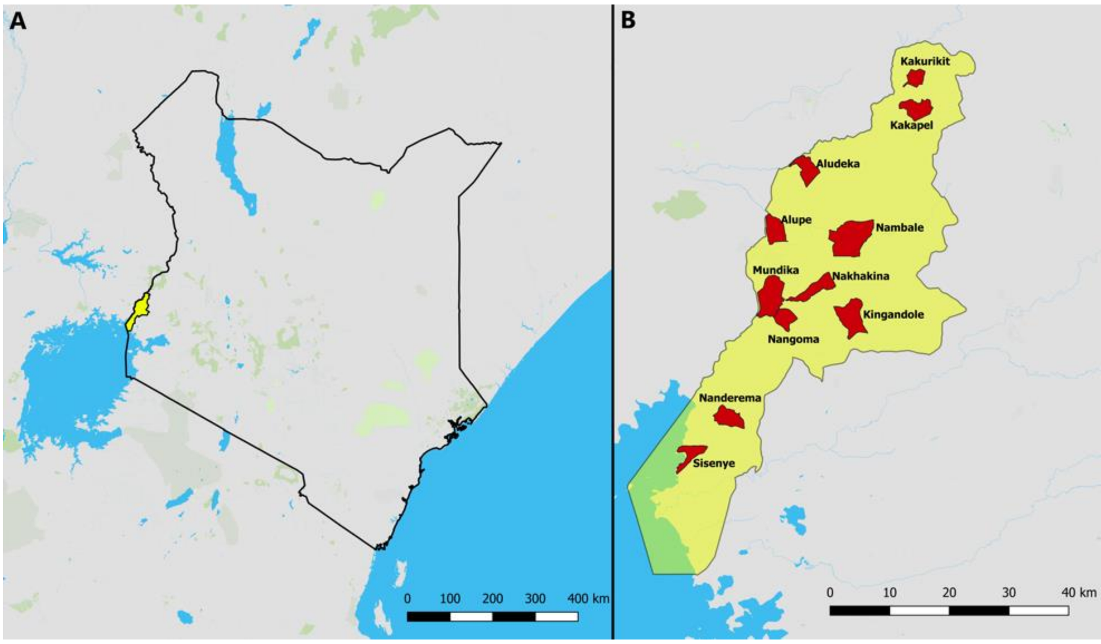


Figure 4-1 Map of the study area in Kenya.

Left: Outline of Kenya with Busia county highlighted. Right: Busia county with the 11 study sublocations highlighted in yellow.

I used clustered random sampling to select 55 households within Busia County. Of the 181 smallest administrative units within the county (known as sublocations), 11 were selected at random to obtain data from a broad geographical range. Within these I visited five households for participation in the study to optimise the use of the limited number of GPS loggers. In the absence of adequate household density data, I used random generation of geographical coordinates within each sublocation followed by identification of the household closest to the coordinates within 200 metres to choose the households, using protocol established by other studies in this area (Eric M. Fèvre et al., 2017). This method will necessarily result in bias toward selection of rural households, which I deemed acceptable given the high proportion of rural households in Busia County. I conducted the survey at all study households and the GPS tracking at 27 randomly chosen of the 55 originally selected households (all of whom agreed to participate), in 6 of the 11 sublocations. The adult who spent the most time tending to the ruminants was asked to wear the logger, or, if there were no ruminants, the household head.

4.3.2 Data collection

The survey (see Appendix A.2) was administered to the consenting adults in each household individually. The questions addressed the following: (a) basic demographic and household information, (b) an individual's movements and (c) their activities involving livestock. The survey allowed estimation of the relative wealth of households by scoring the answers to 10 standardized questions using the 2011 Poverty Probability Index for Kenya (*Kenya | PPI*, 2011).

For the GPS data collection, participants were given a GPS logger (i-gotU GT-600, accuracy of 25m (*Mobile Action Technology Inc., Taipei, Taiwan*, 2018)) fitted to a lanyard for one week, and asked to wear it during the daytime. Simultaneously, if the household kept ruminants then one of these (preferably a cow or bull) was chosen through random number generation to wear an identical waterproofed GPS logger on an adjustable collar. If the household kept no ruminants, only the participant was given a logger to allow for comparison of movement patterns between people with and without ruminants. At the end of the week, the researchers returned to the household to collect the loggers and download the data.

The household survey and GPS tracking were conducted in short rainy season (July/August 2016) and the GPS tracking was repeated during the dry season (November/December 2016) with the same participants where possible, to capture potential differences in movement patterns during different seasons. These seasons are named in accordance with climate classifications for this region of Kenya (Thuranira-McKeever et al., 2010).

4.3.3 Data analysis

Data were mapped in QGIS (v2.18 (QGIS Development Team, 2019)) for inspection, then cleaned and analysed in R software (v3.4.2 (R Core Team, 2017)) using the *trip* (M. D. Sumner, 2016)*, lme4* (Douglas Bates, Martin Maechler, Ben Bolker, 2015) and *glmmTMB* (Brooks et al., 2017) packages. A linear interpolation algorithm was used to clean the data and calculate one point per minute over the collection period. From these data, I calculated five movement measurements for each subject: the time spent, maximum distance travelled and frequency of trips outside of the household, the home range of the subject and the time spent on different types of landcover. A trip was defined as the time when the location of a subject was recorded 100 metres or farther from their household for more than 15 minutes. Home ranges were calculated using the minimum convex polygon method. Land cover data for the region were obtained from Wardrop et al. (Wardrop, 2014) and used to calculate the time spent by people and ruminants on each of 5 types of land cover in the classification: artificial/bare land (defined in the classification as non-vegetated land), crops/grassland, woodland/shrubs, rice paddies and swampland.

I first conducted univariable analyses using linear regression for the four movement measures (time spent, maximum distance travelled and frequency of trips outside of the household), and beta regression for the proportion of time spent in different types of land cover. The covariates used, given in Table 4-1, were selected from covariates used in a previous study of a zoonosis in this region (Wardrop et al., 2016) and had *a priori* plausibility. I log transformed the data for time spent on trips, maximum distance travelled and home range to ensure an approximately normal distribution and used linear models to examine their relationships with the covariates. I then combined the statistically significant factors ($p < 0.05$) from the univariable analyses into a multivariable model to quantify their impact in context of each other. I also tested for interactions between season and statistically significant covariates. Because of the hierarchical nature of the data (individuals nested in clusters, with repeat measurements per individual), I used linear mixed models (LMMs) and generalized LMMs (GLMMs) for the multivariable analyses. All mixed models had the individual household nested within the sublocation as a random effect, to account for variation between individuals from different households and within different sublocations. When analysing time spent on different types of land as the measured outcome, I calculated the number of minutes spent on and off each of the land types, and then used a beta family in a GLMM to obtain odds ratios for the different effects.

4.3.4 Ethics statement

This study was approved by the Institutional Research Ethics Committee and the Institutional Animal Care and Use Committee of the International Livestock Research Institute (IDs: ILRI-IREC2016–11; IACUC-RC2016–14; committees approved by the Kenya National Commission for Science, Technology and Innovation (NACOSTI)) and the Ethics and Research Governance board at the University of Southampton (ID:18984). All participants provided signed, informed consent for their participation in the study and the data were stored securely in accordance with the University of Southampton data storage policy.

4.4 Results

GPS data were successfully collected from 26 humans and 20 ruminants from 26 households in the first season, and from 25 humans and 15 ruminants in the second season (87% of the original tracked individuals), for a total of 86 unique GPS readings. Three of the ruminant GPS units were damaged on retrieval and the data could not be extracted, two suffered battery issues and one was irretrievable. Table 4-1 provides some characteristics of participating households.

Table 4-1 Individual and household characteristics used in analyses.

Demographic covariates	**Number of participants: Survey & GPS**	**Number of participants: GPS only**
Gender		
Male	33 (43.4%)	18 (69.2%)
Female	43 (56.6%)	8 (30.8%)
Age		
18-29	22	7
30-49	26	8
50-69	23	7
70+	5	4
Main occupation		
Farming/agriculture	45	18
Hunting	2	2
Trading	3	1
Other	18	4
Unemployed	8	1
Relative wealth score (PPI Kenya*) of participant's household		
Less than 30	16	9
30 to 50	24	9
51 or more	12	8
Ruminant ownership of participant's household		
No ruminants	12	6
Ruminants	40	20

*PPI Kenya = Poverty Probability Index for Kenya 2011, see supplementary information for details.

4.4.1 Movements beyond the household

Humans spent a mean of 2.4 hours per trip outside of the household while their ruminants spent 4.7 hours. Humans travelled significantly further than their ruminants on these trips, with a mean maximum distance of 1060m travelled compared to 362m for ruminants. They also had larger home ranges, with a mean area of 7.5km^2 compared to 0.1km^2 for ruminants. Finally, humans took more frequent trips away from the household, with a mean of 18 trips per week compared to 9 for ruminants (Figure 4-2).

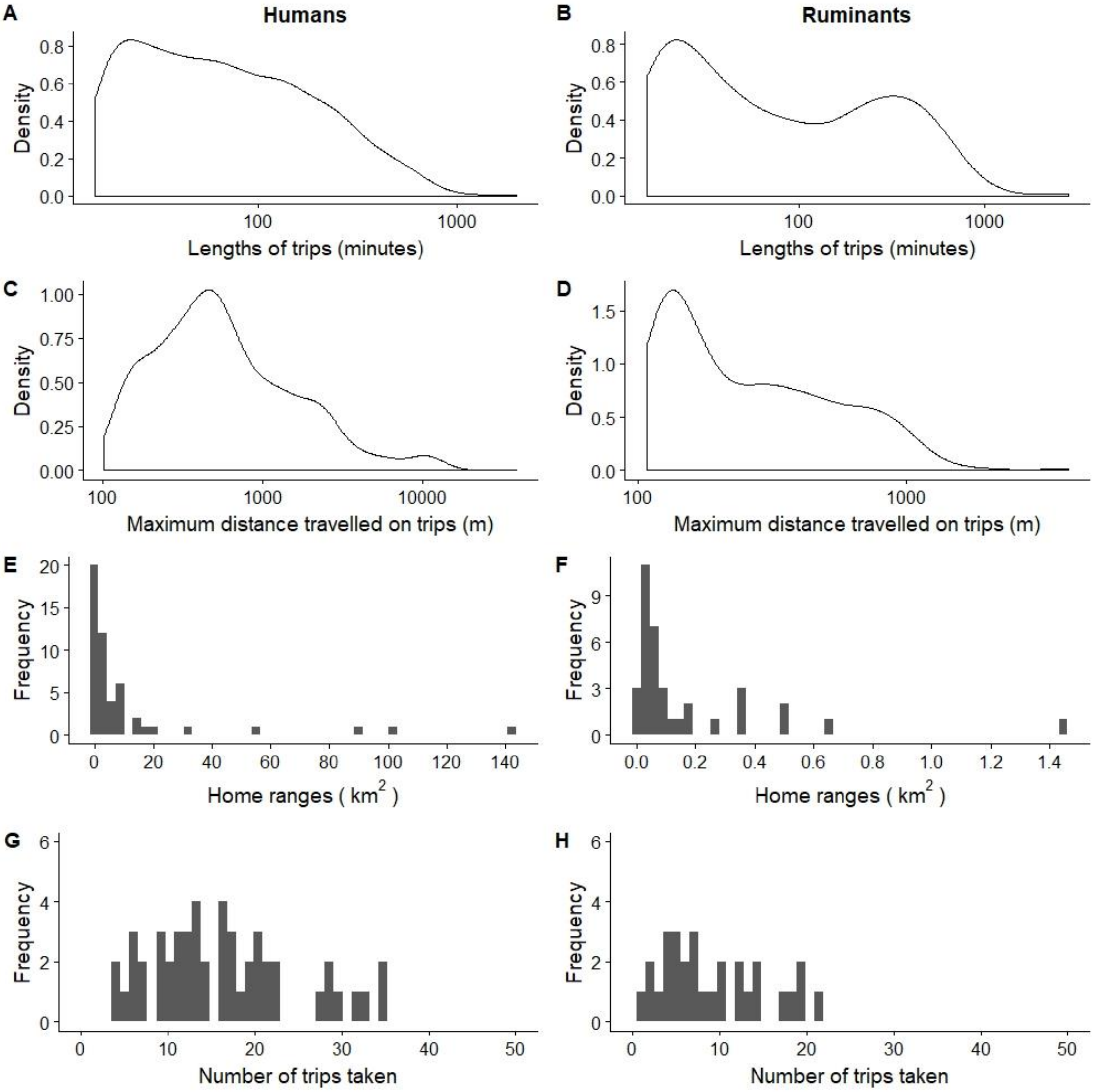


Figure 4-2 Summary of the human (A,C,E,G) and ruminant (B,D,F,H) GPS data.

A,B Lengths of trips outside the household (n=934, n=306). **C,D** Maximum distance travelled on trips outside the household (n=934, n=306). **E,F** Home ranges (n=51, n=36). **G,H** Frequency of trips outside the household (n=51, n=29).

Exploratory analyses using univariable linear regression with random effects showed that some demographic covariates had statistically significant ($p < 0.05$) effects on the movement response variables (Table 4-2). Ruminant owners spent longer and travelled further on trips outside of the household than non-owners. I also observed a link between household wealth and distance travelled on trips, with people travelling 0.9 times farther per 10-point increase in their household wealth. For ruminants, a similar pattern was found: the wealth of the household was associated with both the time spent and frequency of trips taken by ruminants on trips outside of the household, with animals from poorer households spending more time outside of the household and taking more frequent trips. Differences in trip frequency were not associated with any of the covariates tested for humans.

Table 4-2 Univariable linear regression for movement response variables.

The time spent, maximum distance and home ranges were log transformed before modelling, thus these estimates are factor increases and decreases, while estimates for trip frequency are absolute. For ruminants, only the relevant covariates were tested. Figures in square brackets are 95% confidence intervals. *** p < 0.001, ** p < 0.01, * p < 0.05.

Response variable	**Explanatory variable**	**Estimate**	**p-value**
Time spent on trips outside of the household (humans)	Ruminant ownership: yes [Ref = no]	1.47 [1.24, 1.76]	< 0.001 ***
	Number of ruminants	1.12 [1.04, 1.20]	0.002 **
	Gender: male [Ref = female]	1.07 [0.91, 1.26]	0.396
	Occupation: non-farmer [Ref = farmer]	1.02 [0.86, 1.20]	0.836
	Season: dry [Ref = short rainy]	1.21 [1.05, 1.39]	0.008 **
	Household wealth	1.00 [1.00, 1.00]	0.768
	Age (years)	1.01 [1.00, 1.01]	< 0.001 ***
Time spent on trips outside of the household (ruminants)	Number of ruminants	1.32 [1.04, 1.67]	0.023 *
	Season: dry [Ref = short rainy]	0.95 [0.72, 1.27]	0.735
	Household wealth	0.99 [0.98, 1.00]	0.008 **
Maximum distance travelled outside of the household (humans)	Ruminant ownership: yes [Ref = no]	1.50 [1.25, 1.80]	< 0.001 ***
	Number of ruminants	1.10 [1.02, 1.18]	0.015 *
	Gender: male [Ref = female]	1.15 [0.97, 1.37]	0.107
	Occupation: non-farmer [Ref = farmer]	1.21 [1.02, 1.44]	0.034 *
	Season: dry [Ref = short rainy]	1.32 [1.16, 1.51]	< 0.001 ***
	Household wealth	0.99 [0.99, 1.00]	< 0.001 ***
	Age (years)	1.00 [1.00, 1.01]	0.158

Response variable	**Explanatory variable**	**Estimate**	**p-value**
Maximum distance travelled outside of the household (ruminants)	Number of ruminants	1.09 [0.94, 1.26]	0.246
	Season: dry [Ref = short rainy]	0.99 [0.85, 1.15]	0.848
	Household wealth	0.99 [0.99, 1.00]	0.040 *
Home range (humans)	Ruminant ownership: yes [Ref = no]	3.31 [0.81, 14.43]	0.111
	Number of ruminants	1.20 [0.64, 2.24]	0.572
	Gender: male [Ref = female]	1.60 [0.41, 6.28]	0.501
	Occupation: non-farmer [Ref = farmer]	1.33 [0.36, 5.54]	0.677
	Season: dry [Ref = short rainy]	3.28 [1.58, 7.04]	0.004 **
	Household wealth	0.99 [0.95, 1.02]	0.428
	Age (years)	1.03 [1.00, 1.07]	0.049 *
Home range (ruminants)	Number of ruminants	2.09 [1.06, 4.12]	0.049 *
	Season: dry [Ref = short rainy]	0.77 [0.43, 1.35]	0.368
	Household wealth	1.01 [0.98, 1.04]	0.623
Trip frequency (total number of trips taken, humans)	Ruminant ownership: yes [Ref = no]	1.87 [-5.11, 8.71]	0.594
	Number of ruminants	1.05 [-1.96, 3.94]	0.486
	Gender: male [Ref = female]	4.90 [-2.73, 10.96]	0.126
	Occupation: non-farmer [Ref = farmer]	-2.10 [-8.33, 4.14]	0.517
	Season: dry [Ref = short rainy]	-1.11 [-4.79, 2.67]	0.557
	Household wealth	0.00 [-0.16, 0.16]	0.990
	Age (years)	-0.05 [-0.22, 0.11]	0.513
Trip frequency (total number of trips taken, ruminants)	Number of ruminants	1.82 [-1.19, 5.88]	0.285
	Season: dry [Ref = short rainy]	1.82 [-2.22, 5.60]	0.360
	Household wealth	0.14 [0.00, 0.26]	0.038 *

I found significant differences between the two seasons in the distances participants travelled on trips and their overall home ranges (Figure 4-3). Humans spent 1.21 times longer on trips outside of the household ($p = 0.008$), travelled 1.32 times further ($p < 0.001$) and had home ranges that were 3.28 times larger ($p = 0.004$) in the dry season compared to the short rainy season. No link was found between season and trip frequency. The movement measures did not vary significantly by sublocation in either season, and ruminant movement measures showed no significant differences between the seasons.

Using only the significant covariates from the univariable analyses, I then carried out a multivariable linear regression (see Appendix A.1.1). In these analyses, only season remained a significant covariate, confirming the finding that in the dry season people took longer trips, travelled further on them and had larger home ranges compared to the short rainy season.

I tested for interactions between season and the significant covariates from the univariable analyses, and found that the season variable interacted significantly with both ruminant ownership and number of ruminants owned for the time spent and distance travelled outside of the household. The most relevant finding was that in the short rainy season, ruminant owners spent almost twice as long and travelled 1.67 times further on trips outside of the household ($p = 0.002$ and 0.049, respectively) compared to non-owners. See Appendix A.1.2 for full results.

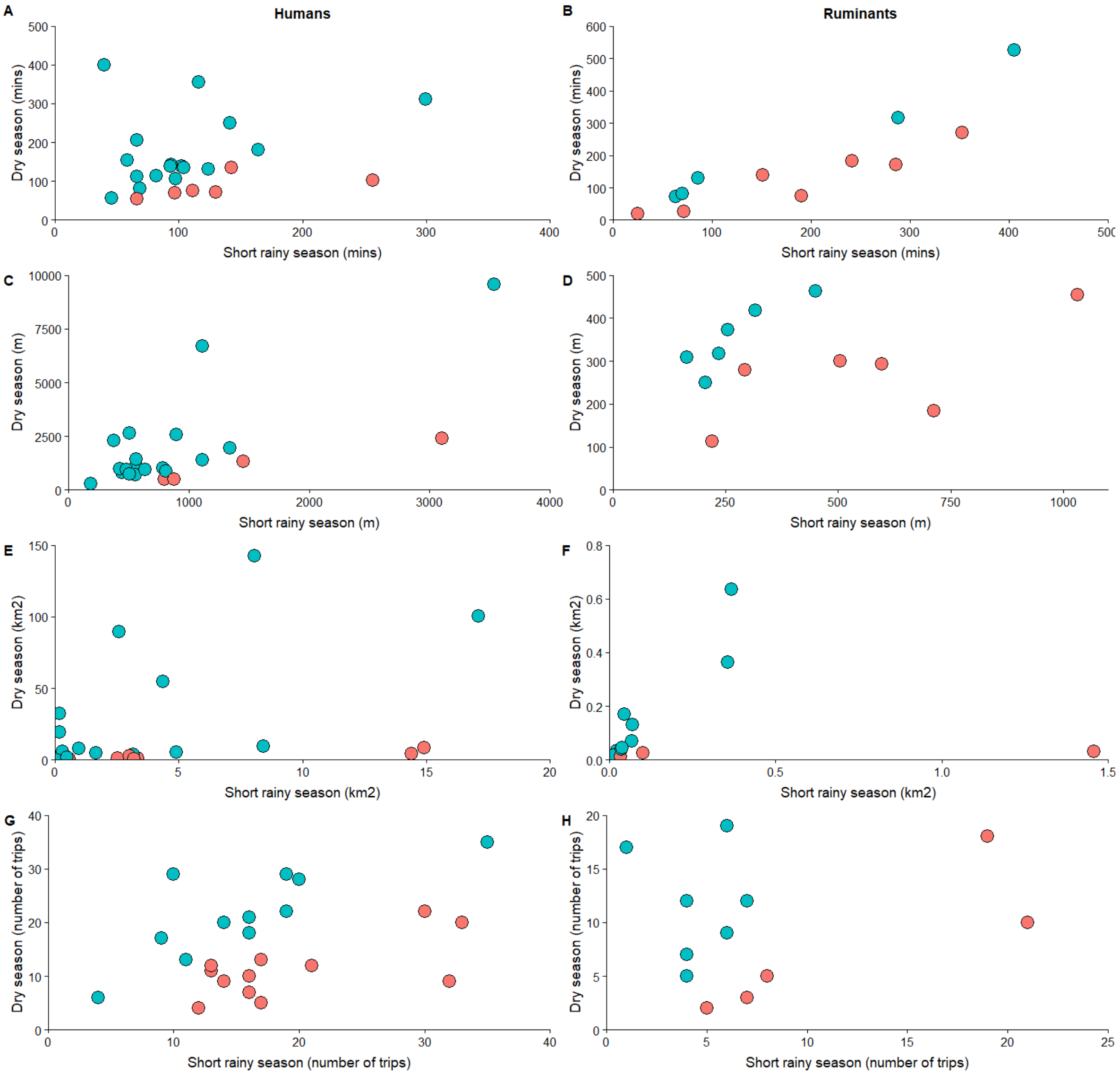


Figure 4-3 Comparison of movement measures between seasons for humans (A,C,E,G) and ruminants (B,D,F,H).

Points are blue where the dry season value is larger than in the short rainy season, and red otherwise. **A,B** Lengths of trips outside of the household. **C,D** Maximum distance travelled on trips outside of the household. **E,F** Home ranges. **G,H** Frequency of trips outside of the household.

4.4.2 Movements and landcover

In univariable beta regression analyses I found that some of the demographic covariates were associated with time spent on different types of land, including ruminant ownership: ruminant owners spent more time on wood/shrubland compared to non-ruminant owners (OR 2.79). Moreover, the odds of a participant spending time on wood/shrubland were 1.38 times higher for each extra ruminant owned. The full univariable results are presented in Appendix A.1.3. The time spent by humans and their ruminants on two types of land cover varied significantly with season. During the dry season, people spent less time on swampland compared to the short rainy season, and their ruminants spent more time on artificial/bare land. The time spent on other types of landcover did not vary significantly by season (Table 4-3). I tested for the interaction of season with ruminant ownership for time spent on different types of land but found few interactions that were significant and relevant to the focus of this paper.

Table 4-3 Comparison of time spent by humans and ruminants on different land types between wet and dry seasons.

Univariable beta regression models with sublocation as a random effect. Figures in square brackets are 95% confidence intervals. *** p < 0.001, ** p < 0.01, * p < 0.05.

Land type	**Host**	**Odds ratio (dry season compared to short rainy season)**	**p-value**
Artificial/bare	Human	1.12 [0.92, 1.36]	0.267
	Ruminant	1.27 [1.03, 1.58]	0.028 *
Crops/grassland	Human	1.01 [0.87, 1.17]	0.929
	Ruminant	1.16 [0.90, 1.51]	0.256
Rice paddies	Human	1.10 [0.95, 1.26]	0.204
	Ruminant	1.00 [0.91, 1.09]	0.981
Swamp	Human	0.83 [0.69, 0.99]	0.034 *
	Ruminant	1.25 [0.96, 1.63]	0.093
Woodland/shrubs	Human	0.96 [0.81, 1.14]	0.642
	Ruminant	1.31 [0.88, 1.95]	0.188

4.5 Discussion

Host movements are critical to the propagation of some of the highest burden vector-borne zoonoses. The movement patterns that contribute to zoonosis risk are relatively unknown, and there is little understanding of how movements vary between different seasons and population groups for both humans and livestock in the context of vector-borne zoonotic disease. Although this was an exploratory study limited by a small sample size and biased towards rural households, the seasonal and demographic differences in the movement patterns of people and their ruminants observed indicate that these patterns are worthy of further study.

In the dry season, I found that people took longer trips, travelled further on them and had larger home ranges compared to the short rainy season, yet these measures for their ruminants remained similar across the two seasons. This suggests the extra distance travelled by people during the dry season was not done with their ruminants, which could be because people had to travel further in the dry season to access resources needed for the household, such as water or forage for their livestock. I also found that ruminant owners in both seasons travelled further and for longer than non-owners, supporting this conclusion. This key difference in movement between seasons has been identified in previous studies: a study done in Zambia using similar methods also found that people living in rural areas travelled further during the dry season (Searle et al., 2017). However, few disease transmission models incorporate this information, likely due to a lack of understanding of how best to account for these differences in movement.

The link between movement and household wealth indicates that people and ruminants from relatively poorer households travelled further than those from wealthier households in Busia County. This could support the relevance of resource access - if people from wealthier households have the option of accessing resources nearer to their household, they may not make as many long-distance trips as people from poorer households. We know that people from poorer households often travel further to access health facilities (Wagstaff, 2002); these findings suggest that this may extend to other types of resources, such as water and forage, although further work is needed to verify the findings. A previous study in Busia County has shown that household wealth is linked to ease of resource access and risk of infection, with poorer households having more difficulty accessing resources and being at higher risk of certain zoonoses (Fèvre, E; unpublished results). Since individual movements are a key factor in zoonotic transmission, it could be that the wider-ranging movement patterns of people and ruminants from poorer households are putting them at additional risk of infection. Further research to identify the types of places people and their herds travel to could pinpoint 'hotspots' of disease transmission.

I show that humans spent more time in swampy areas during the short rainy season than the dry. People may be spending more time on swampland in the short rainy season to take advantage of water sources or grazing areas for their ruminants, or potentially as a water source for the household. Swampland is prone to flooding, and studies of mosquito habitats have shown that it tends to have a higher abundance of larval breeding habitats for zoonotic vectors such as *Aedes aegypti* and *Aedes albopictus* as well as malaria vectors (*Anopheles spp.*) than other types of land

(Omukunda et al., 2012; Sarfraz et al., 2012). Therefore, spending more time in swampland in the short rainy season may be exposing both humans and their ruminants to mosquito-borne diseases, though further analysis with a larger sample size is clearly needed to verify this conclusion across these populations.

The use of GPS devices to track people and animals in urban and rural settings has been previously documented (Parsons et al., 2014; Paz-Soldan et al., 2014) and shown to be a reliable tool for tracking movements, with good portability, weight and battery life. This exploratory study confirms the feasibility of simultaneously tracking human and livestock movements, as this had not previously been tested. Although the sample size was small, the data collected from participants yielded some interesting results over the two seasons studied. In future work, a larger sample size in this area could provide more detailed conclusions and reveal patterns that are generalizable to similar rural populations, particularly those with livestock. To overcome the bias in household selection, high-resolution population density data like those produced by WorldPop (Linard et al., 2012) could be used to weight the selection of various points on the map. Nevertheless, the data obtained by tracking individuals across the two seasons were especially valuable, and thus a longitudinal cohort design where individuals are tracked continuously over the year could be an ideal method to gain data on long-term and long-distance movements, thus facilitating estimation of vector-borne and zoonotic disease risk on wider spatial and temporal scales.

4.6 Intellectual contributions

The author list for this published work is as follows: Jessica R. Floyd, Nick W. Ruktanonchai, Nicola Wardrop, Andrew J. Tatem, Joseph Ogola and Eric M. Fèvre. JRF conceived and designed the fieldwork, performed the fieldwork, analysed the data, prepared figures and tables and wrote and prepared the manuscript for journal submission. JO performed the fieldwork with JRF and provided insight and feedback on the study design. EMF and NW provided insight and feedback on conceptualisation of the fieldwork and the study design. NW, AJT and NWR supervised JRF during the course of this project and oversaw that the analysis was executed in a scientific manner. All authors reviewed the manuscript and provided feedback before final submission to the journal.

Chapter 5 Activity-specific mobility of adults in a rural region of western Kenya

5.1 Abstract

Improving rural household access to resources such as markets, schools and healthcare can help alleviate poverty in low-income settings. Current models of geographic accessibility to various resources rarely take individual variation into account due to a lack of appropriate data yet understanding mobility at an individual level is key to knowing how people access their local resources. This study used both an activity-specific survey and GPS loggers to evaluate how adults in a rural area of western Kenya accessed resources in their local area. I calculated the travel time and time spent at six different types of resource and compared the GPS and survey data to see how well they matched. I found links between several demographic characteristics and the time spent at different resources, and that the GPS data reflected the survey data well for time spent at some types of resource, but poorly for others. I conclude that demography and activity are important drivers of mobility, and a better understanding of individual variation in mobility could be obtained through the use of GPS loggers on a wider scale.

5.2 Introduction

Population mobility is a complex process with great importance in many fields across the social and health sciences (Bajardi et al., 2011; R. Mansell Prothero, 1977; Amy Wesolowski et al., 2012). Often, people travel because of specific resource needs or activities such as gathering food and water, livelihood and occupational activities, or accessing healthcare. In low-income settings, travel to these resources can be very time-consuming or expensive meaning that people may forego healthcare, employment, or other resources. As a result, the geographic inaccessibility of vulnerable populations can lead to worse health outcomes, a poorer economic outlook, and can widen spatial inequalities (Alegana et al., 2018; Macintyre et al., 2019; Pearce et al., 2008). In Kenya, resource access is particularly important for poverty reduction in rural populations, as people often have to travel further for resource-related activities than their urban counterparts. It is widely accepted that people in rural areas spend more time accessing resources than people in urban areas, and that this likely contributes to poverty in these areas. For example, market access is important for several household activities such as gathering food, selling crop surpluses and buying medicine, and studies have shown that poor market access contributes to poverty in rural areas (Chamberlin & Jayne, 2013). Similarly, poor water source access means reduced time for making money elsewhere and therefore contributes to household poverty, as well as being linked to poor health (Cook et al., 2016; Whittington et al., 1990).

Geographic inaccessibility of healthcare is also a known driver of poor health outcomes, particularly in rural settings (A. M. Noor et al., 2003; C. W. Ruktanonchai et al., 2016). Government policies in Kenya have responded accordingly, through measures designed to ensure that everyone lives within 5km of basic healthcare services. In 2003 it was estimated that 82% of the population live within 5km of a primary healthcare and referral service (A.M Noor et al., 2004). Because geographic accessibility is vitally important for ensuring vulnerable populations can utilise healthcare, a significant body of recent research focuses on modelling accessibility across national scales. Often, these accessibility models assume that people visit their nearest clinic (Alegana et al., 2012), using accessibility surfaces or straight-line distances to predict clinics used and associated travel times. In reality, geographic accessibility remains highly heterogeneous across the country despite new clinics in resource-poor areas (Kenya Ministry of Medical Services and Ministry of Public Health & Sanitation, 2013), and straight-line distance is not the only factor that impacts whether people can access healthcare in a reasonable time without undue expense. Other factors such as poor road quality and lack of public transport options can severely impact mobility and therefore healthcare access in rural areas where the most vulnerable populations live (Airey, 1992; Tanser et al., 2006). In Kenya, recent studies have found that people often visit clinics other than their nearest one, for reasons such as the availability of medicines or the perceived effectiveness of the facility (G. M. Mwabu, 1986). One study found that only 54-63% of people surveyed visited their nearest facility, with the rest visiting other clinics (G. Mwabu et al., 1995).

Moreover, resource-related movement is quantitatively different from other types of mobility, so accessibility models that predict general mobility patterns may not accurately reflect resource seeking behaviour. For example, a recent study in Iquitos, Peru found that residents moved significantly further for commercial and familial reasons than for healthcare (Perkins et al., 2014). Sociodemographic factors could also influence mobility and resource-specific movement: household income, rural/urban context (C. S. Molyneux et al., 1999) and gender are all important determinants of mobility, but spatial models of access often lack this demographic information, assuming that all adults access resources such as markets and health facilities identically, regardless of socioeconomic context. More detailed movement datasets could help improve these models by providing evidence to support or reject these key assumptions and better understand how movements driven by different activities vary.

Additionally, mobility studies typically focus on single types of resource access without comparing against travel of other types (Kanuganti et al., 2015; Schröder et al., 2018), particularly in a rural context where activity-dependent mobility models could provide a richer picture of how people spend their time, and how resource-specific movement could exaggerate or mitigate geographic inaccessibility. Traditionally, studies of individual mobility have relied on survey methods, which may be affected by recall bias (Amy Wesolowski, Stresman, et al., 2015). In recent years, specialised tools such as personal Global Positioning System (GPS) tracking devices employed at a household level have facilitated the collection of detailed movement information (Parsons et al., 2014; Searle et al., 2017; Vazquez-Prokopec et al., 2013). Personal GPS loggers have been used in urban and rural settings to document the movements of both humans and animals in a variety of contexts, from investigating the diving behaviour of certain species of birds (Browning et al., 2018; Ryan et al., 2004) to the social structures of cattle in pastoral communities (Moritz et al., 2012) to detailed human movements in urban settings (Vazquez-Prokopec et al., 2013). With good portability, weight and battery life, this type of logger has also been used in research into healthcare access (Siedner et al., 2013) and vector-borne disease (Searle et al., 2017; Vazquez-Prokopec et al., 2013). Given the growing populations in rural areas of lower-income countries, detailed knowledge of activity spaces and health facility access in a rural context could have benefits for better understanding specific types of resource access.

Here, I use surveys of people in a rural area of western Kenya to capture movements to and time spent on different types of activity, and examine links between these and demographic characteristics. I use GPS loggers to explore where the same people spent their time outside of their households. I then compare these two sources of data to see how well they capture proportions of time spent on different types of activity. The results from this study shed light on the activity spaces of people in relation to resource access in a resource-poor setting and provide evidence for how well GPS loggers are able to capture daily activities compared to survey methods. They also help quantify the importance of including sociodemographic and activity-specific movement into geographic accessibility models, particularly for healthcare infrastructure.

5.3 Methods

5.3.1 Study area and population

This study was conducted in the densely populated county of Busia in the Lake Victoria basin region of western Kenya (Figure 4-1). Busia county has a population of just under 900,000 people (Kenya National Bureau of Statistics, 2017), of whom approximately 80% live in rural areas (as defined by the latest DHS survey) and practice smallholder subsistence farming, mainly operating mixed crop-livestock systems with 60% of households in rural areas owning cattle (Kenya National Bureau of Statistics et al., 2015).

I used a clustered random sampling strategy to select 55 households to visit and survey in Busia county. Of the 181 sublocations in the county, 11 were selected at random and within these I selected five households for participation in the study, by randomly generating coordinates within each sublocation, and choosing the household closest to each of the coordinates, up to 200 metres. A surplus of coordinates (up to ten) were generated in each sublocation, so that if a household could not be identified within 200 metres, the next set of coordinates were used. I used QGIS software tools (QGIS Development Team, 2019) for the random selection of sublocations and coordinates. Due to the lack of appropriate data to base a sample size calculation on, I chose to sample 55 households from a broad geographical range based on the manpower and time available for fieldwork while optimising the use of the limited number of GPS loggers. I selected 30 of the 55 households for GPS tracking because this was the maximum number of people that could be tracked for a full week with the resources available.

Households were selected if they were the main residence of at least one consenting adult present at the time of the visit. Inclusion criteria for participation were consenting adults aged 18 years or over who were present at the time of the visit. Some demographic characteristics of the study population are given in Table 4-1. If the household declined to participate in the study, the next household closest to the coordinates was visited. All adults within a household were selected for participation in the survey, but only the adult who spent most time looking after livestock (determined from survey responses and verbal discussion) was selected for participation in the GPS tracking. If the household had no livestock, the head of the household (determined by verbal communication) was selected instead. These choices enabled analysis of the activities of populations that spent time with livestock and therefore may be at higher risk of some zoonotic diseases. This limits the representativeness of the sample population, but I felt this was acceptable given the small sample size.

5.3.2 Data collection (survey)

A structured survey with closed and open questions was administered to all consenting members aged 18 years or over present in the household at the time of visiting (see supplementary information). The survey included questions on demographic characteristics, regular movements outside of the household and activities involving livestock. For movements to places outside the household, I asked about the type of place visited (e.g. school, water source, place of worship, market etc.), how often the respondents visited, the mode of transport used and time spent travelling, and how long they usually spend there. The types of places were pre-defined based on information from previous studies in this area (Eric M. Fèvre et al., 2017; Floyd et al., 2019), and the survey included open-ended questions to identify any other significant types of place. A village elder was present to facilitate introduction to the household and explanation of the study. The survey was written in English and administered through an interpreter. The survey included 10 closed-ended questions from the Poverty Probability Index for Kenya (*Kenya | PPI*, 2011). The answers to these questions were scored to obtain a basic index of household wealth which was used to compare relative wealth between households. The answers to the survey, including collection of GPS coordinates to determine household location, were collected on a tablet using a custom-designed survey built with OpenDataKit (ODK) (Brunette et al., 2013) software and uploaded to a secure server once an internet connection could be established.

5.3.3 Data collection (GPS)

During the same visit, the consenting participant was given a GPS logger (i-gotU GT-600 by Mobile Action (*Mobile Action Technology Inc., Taipei, Taiwan*, 2018)) to wear for one week, fitted to a lanyard, which could be worn around the neck or carried in a pocket. I chose this length of time based on previous studies (Bohte & Maat, 2009; Stopher et al., 2018) and to maximise the use of the limited number of GPS loggers. The time interval used on the loggers was one minute and the devices were programmed to power off if stationary for two minutes, then turn on again when movement was detected. At the end of the week, I returned to the household to collect the loggers and download the data.

The data collection was conducted in two phases: household survey and GPS tracking in July/August 2016, followed by GPS tracking only in November/December 2016. In phase two, the same households were visited as in phase one and the GPS tracking was repeated with the same participants where possible, in order to capture potential differences in movement patterns during different seasons, henceforth called the short rainy season (July/August 2016) and the dry season (November/December 2016) in accordance with climate classifications for this region of

Kenya (Thuranira-McKeever et al., 2010). Of the 26 loggers given out in each season, one device was unrecoverable and two suffered battery issues during the week (these weeks were therefore repeated).

5.3.4 Data analysis

The survey and GPS data were downloaded in .csv format, then cleaned and analysed using R version 3.1.1 software (R Core Team, 2017). Erroneous points in the GPS data were identified by their unlikely speeds and deleted using functions from the *trip* (M. D. Sumner, 2016) package. A linear interpolation algorithm was then applied to obtain locations at regular intervals. Erroneous points can occur due to changing atmospheric conditions and building obstructions and accounted for less than 1% of the dataset. I used survey response, GPS points collected in the field and publicly available datasets of health facilities (Abdisalan M Noor et al., 2009) to group the places where people spent time into six categories: household or residential places, shops and markets, places of worship, health facilities, places where livestock activities occurred, and places where activities related to water (but not livestock) occurred. A central GPS point was identified for each of the places, and a 25-metre radius around that point was used to determine when that place was visited by a person, defined as 5 minutes or more spent within that radius. Due to their larger size, a radius of 50 metres was used for market centres.

I conducted univariable analyses using linear mixed models for the three movement measures calculated from survey responses (frequency of visits, time spent travelling and time spent at places). I log transformed these data to ensure an approximately normal distribution and used linear models to examine their relationships with the covariates. Because of the hierarchical nature of the data (individuals nested in clusters, with repeat measurements per individual), I used linear mixed models (LMMs) which had the individual household nested within the sublocation as a random effect, to account for variation between individuals from different households and within different sublocations. Variables were chosen for analysis in line with those found to be important to movement metrics in previous studies in the area (Eric M. Fèvre et al., 2017; Floyd et al., 2019; Wardrop et al., 2016).

5.3.5 Ethics statement

This study was approved by the Institutional Research Ethics Committee and the Institutional Animal Care and Use Committee of the International Livestock Research Institute (IDs: ILRI-IREC2016–11; IACUC-RC2016–14; committees approved by the Kenya National Commission for Science, Technology and Innovation (NACOSTI)) and the Ethics and Research Governance board at the University of Southampton (ID:18984). All participants provided signed, informed consent for their participation in the study and the data were stored securely in accordance with the University of Southampton data storage policy.

5.4 Results

5.4.1 Frequency of trips outside the household

I used the survey data to explore where people spent most of their time, how long they took to get there and how long they stayed when visiting places outside of the household. When examining the places people travelled to regularly, I found substantial variation in the frequency of visits to different types of places compared to healthcare visits (Figure 5-1). As expected, participants visited health facilities less often than other types of locations, the mode being three to six visits a year with very few visiting more often than monthly. Visits to areas where livestock activities occurred (such as grazing pastures) and to water sources were the most polarised out of all types of travel, with most visits occurring daily or not at all.

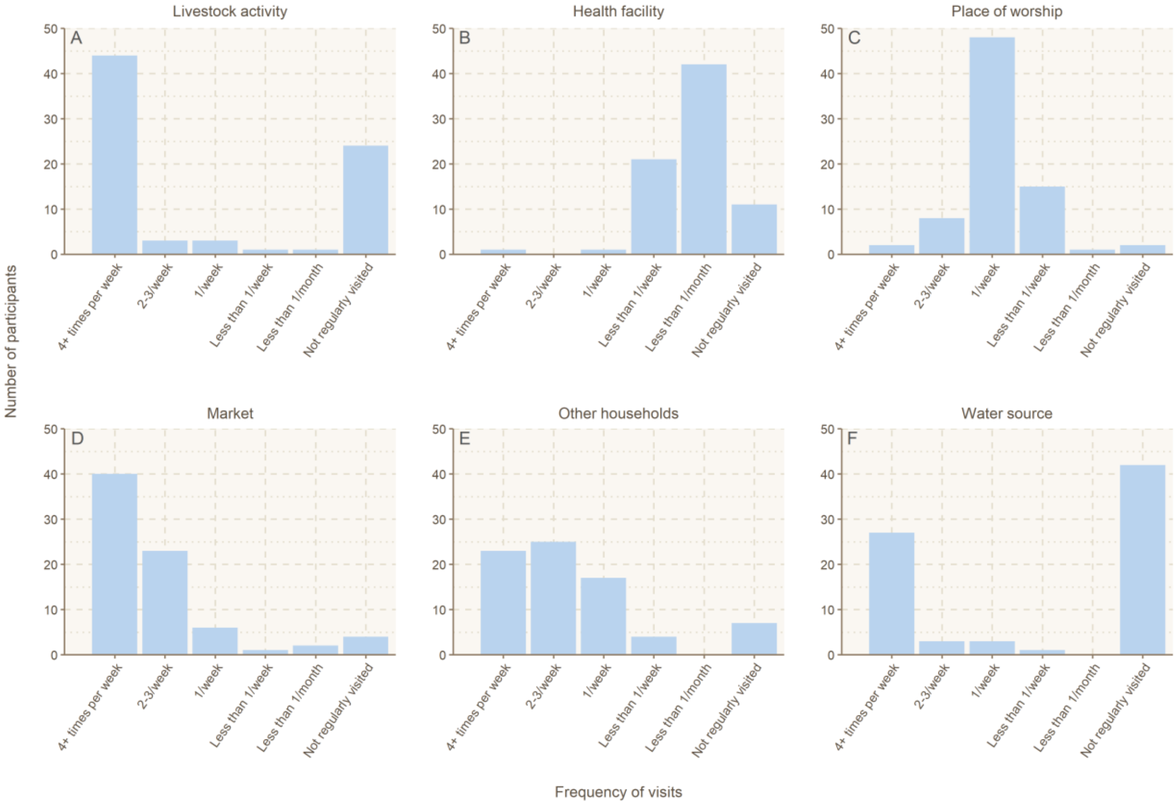


Figure 5-1 Frequencies of visits to different types of places.

5.4.2 Travel times on trips outside the household

I asked participants about how long they spent travelling to different places and found that mean one-way travel times to health facilities ranged from 4 minutes to 2 hours. Participants reported spending the shortest time travelling when tending to livestock (median 9 minutes one way), and the longest times when visiting a market (median 34 minutes one way). I found the greatest range of travel times for visits to health facilities, markets and places of worship, while visits to livestock areas, water sources and other households tended to have shorter travelling times (Figure 5-2).

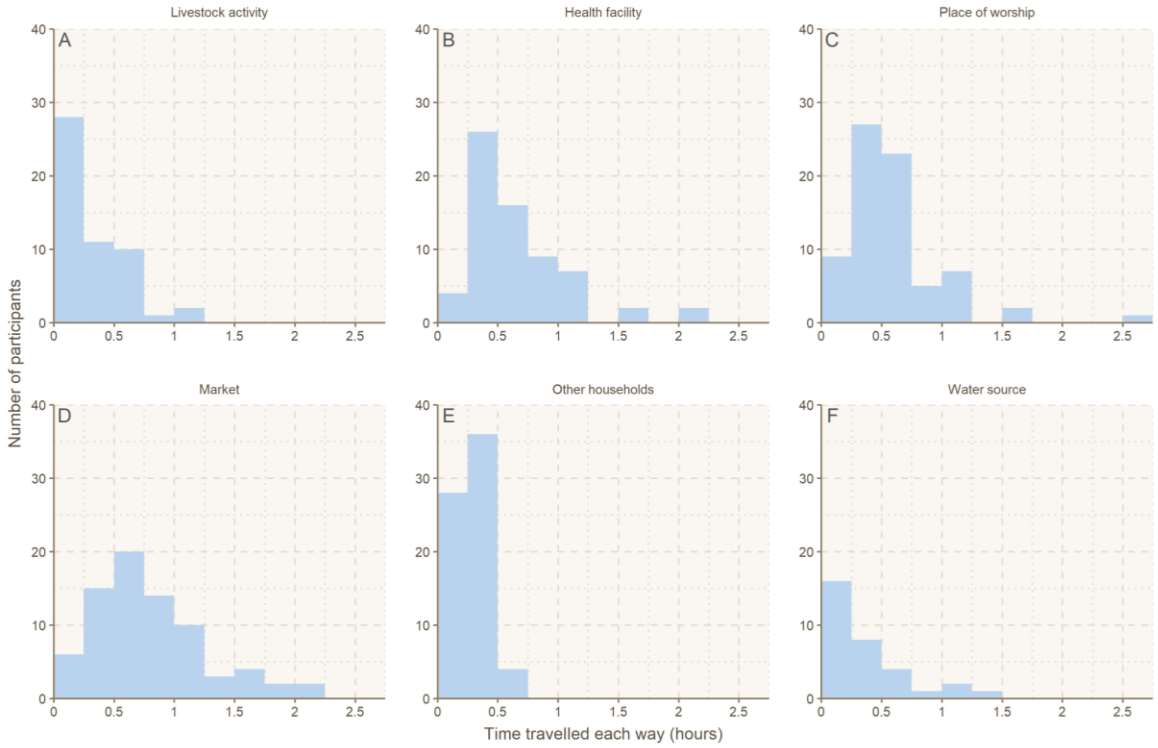


Figure 5-2 Time spent travelling one-way to different types of places.

5.4.3 Time spent at places outside the household

I also asked about the amount of time people spent at different places. Participants reported spending the longest times at health facilities and places of worship (median 3 hours for both), and the shortest times at other households (median 45 minutes) and water sources (median 9 minutes). Figure 5-3 shows how the time spent by participants varied by type of place.

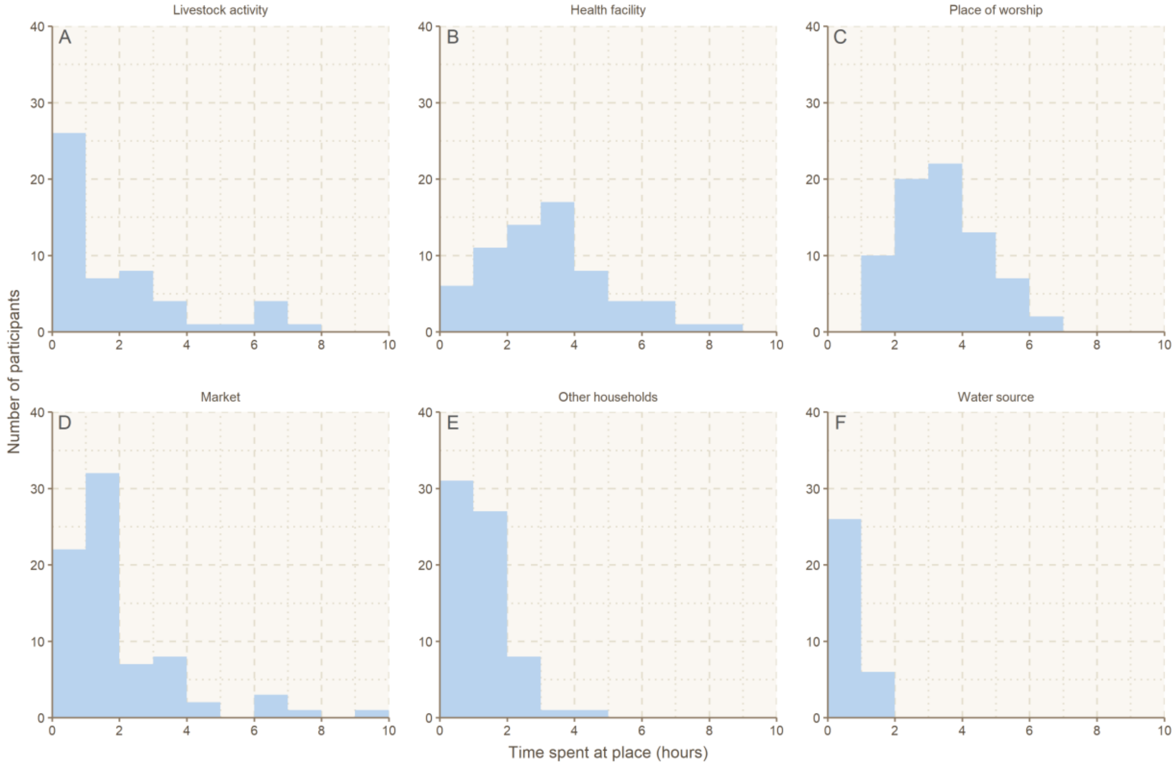


Figure 5-3 Time spent by participants at different types of places.

5.4.4 Demographic links to movement measures

I used the three movement measures (frequency of visits, time spent travelling and time spent at places) and the demographic characteristics collected in the survey to conduct a univariable analysis to explore correlations (Table 5-1). I found that men reported spending more time at and making more frequent visits to places where livestock activities occurred compared to women ($p = 0.017$ and $p = 0.007$, respectively), while women reported spending more time at and making more frequent visits to health facilities than men did ($p = 0.019$ and $p < 0.001$, respectively). Women also made more frequent visits to water sources than men but did not spend significantly more time there ($p = 0.075$). Older people tended to visit health facilities and water sources less often ($p = 0.041$ and $p < 0.001$, respectively) and spend less time at places of worship ($p = 0.003$) than younger people. No significant relationships were observed between time spent travelling to places and any of the demographic characteristics tested; these results are reported in Appendix B.1.1.

Table 5-1 Relationships between movement metrics and demographic characteristics of the surveyed population for different types of activity.

Univariable analysis of movement metrics using a linear mixed model. Estimates are given with 95% confidence intervals in square brackets.
*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

Activity type		**Explanatory variable**	**Estimate**	**p-value**
Livestock activity	Visits per week	Gender: male [Ref = female]	15.78 [2.10, 109.68]	0.007 **
		Occupation: non-farmer [Ref = farmer]	0.08 [0.01, 0.65]	0.022 *
		Household wealth	0.17 [0.01, 3.69]	0.263
		Age	1.04 [0.97, 1.11]	0.245
	Time spent	Gender: male [Ref = female]	5.33 [1.41, 20.22]	0.017 *
		Occupation: non-farmer [Ref = farmer]	0.68 [0.17, 2.79]	0.591
		Household wealth	5.46 [0.95, 31.05]	0.063
		Age	1.03 [0.99, 1.07]	0.206
Health facility	Visits per week	Gender: male [Ref = female]	0.12 [0.04, 0.36]	< 0.001 ***
		Occupation: non-farmer [Ref = farmer]	2.96 [0.93, 9.44]	0.069
		Household wealth	0.25 [0.06, 1.10]	0.071
		Age	0.97 [0.93, 1.00]	0.041 *
	Time spent	Gender: male [Ref = female]	0.65 [0.46, 0.92]	0.019 *
		Occupation: non-farmer [Ref = farmer]	0.88 [0.60, 1.28]	0.504
		Household wealth	0.70 [0.42, 1.18]	0.193
		Age	0.99 [0.98, 1.01]	0.338
Place of worship	Visits per week	Gender: male [Ref = female]	0.50 [0.24, 1.00]	0.054
		Occupation: non-farmer [Ref = farmer]	1.01 [0.48, 2.14]	0.984
		Household wealth	0.65 [0.23, 1.84]	0.425
		Age	1.01 [0.98, 1.03]	0.617
	Time spent	Gender: male [Ref = female]	0.81 [0.65, 1.00]	0.051
		Occupation: non-farmer [Ref = farmer]	1.03 [0.82, 1.28]	0.812
		Household wealth	0.72 [0.54, 0.96]	0.032 *
		Age	0.99 [0.98, 1.00]	0.003 **
Market		Gender: male [Ref = female]	1.09 [0.37, 3.09]	0.875

Activity type		Explanatory variable	Estimate	p-value
	Visits per week	Occupation: non-farmer [Ref = farmer]	0.43 [0.14, 1.36]	0.135
		Household wealth	1.29 [0.27, 6.13]	0.753
		Age	0.97 [0.95, 1.00]	0.121
	Time spent	Gender: male [Ref = female]	1.24 [0.54, 2.87]	0.611
		Occupation: non-farmer [Ref = farmer]	0.64 [0.28, 1.47]	0.299
		Household wealth	0.98 [0.34, 2.84]	0.970
		Age	0.99 [0.96, 1.01]	0.316
Household visits	Visits per week	Gender: male [Ref = female]	1.39 [0.38, 4.85]	0.612
		Occupation: non-farmer [Ref = farmer]	0.92 [0.24, 3.55]	0.909
		Household wealth	1.12 [0.16, 7.49]	0.908
		Age	0.97 [0.93, 1.01]	0.122
	Time spent	Gender: male [Ref = female]	1.22 [0.78, 1.93]	0.387
		Occupation: non-farmer [Ref = farmer]	0.83 [0.51, 1.33]	0.438
		Household wealth	0.87 [0.46, 1.62]	0.655
		Age	1.00 [0.98, 1.01]	0.663
Water activity	Visits per week	Gender: male [Ref = female]	0.02 [0.00, 0.12]	< 0.001 ***
		Occupation: non-farmer [Ref = farmer]	4.80 [0.47, 49.31]	0.191
		Household wealth	0.10 [0.00, 2.66]	0.176
		Age	0.87 [0.81, 0.92]	< 0.001 ***
	Time spent	Gender: male [Ref = female]	2.52 [0.92, 6.65]	0.075
		Occupation: non-farmer [Ref = farmer]	0.86 [0.34, 2.13]	0.755
		Household wealth	0.69 [0.22, 2.17]	0.535
		Age	1.00 [0.97, 1.03]	0.976

5.4.5 GPS validation

Lastly, I used the GPS data to explore the time participants spent outside of their households, and measure how variable overall movements were in the study population. I found that people spent between 5% and 52% of their time outside of their households over the week that they were tracked. For each GPS dataset, I was able to identify the type of place visited for between 2% and 97% of the time recorded outside of the household; some movements could not be characterised because not all locations in the local area were identifiable through the survey or mapping techniques. I selected GPS datasets where the season tracked was the same as when the survey was conducted and where over a third of minutes were identified, resulting in 27 GPS datasets with matching survey data. I compared these to explore how well the GPS data were able to mirror how people spent their time at different places. For this comparison, I only selected types of places that at least 50% of people with GPS data visited at least once a week, apart from the health facility locations, which I included for their public health importance. Unsurprisingly since not all GPS points could be identified, participants reported spending more time at all place types than captured by the GPS (hours reported in survey minus hours recorded by GPS was positive). In most cases, this difference was less than 1 hour of difference per day for each type of place.

Because the GPS data consistently underreported the absolute amount of time spent at different places compared to the survey data, I also compared how well the survey and GPS data were able to capture the proportions of time spent at the different place types. Figure 5-4 shows how analysis of the GPS data found very similar proportions of time spent compared to those reported in the survey for some types of places, such as shops and places where livestock-related activities occurred. Other types of places had large differences in proportion of times spent between the two datasets – specifically, time spent at places of worship and visits to other households and health facilities. Using a paired t-test I found significant differences in the mean proportions of time spent at other households and places of worship ($p = 0.002$ and $p = 0.041$, respectively) between the survey and GPS datasets. The GPS datasets recorded more trips to other households than recounted in the survey, while the survey had more trips to places of worship than recorded in the GPS, potentially due to a recall bias effect: people may be more likely to recall movements that they place higher intrinsic value on. There is also the possibility of a ‘social desirability’ bias: participants may be tempted to overstate the amount of time spent at more desirable locations such as places of worship, particularly as the survey was conducted in the presence of village elders.

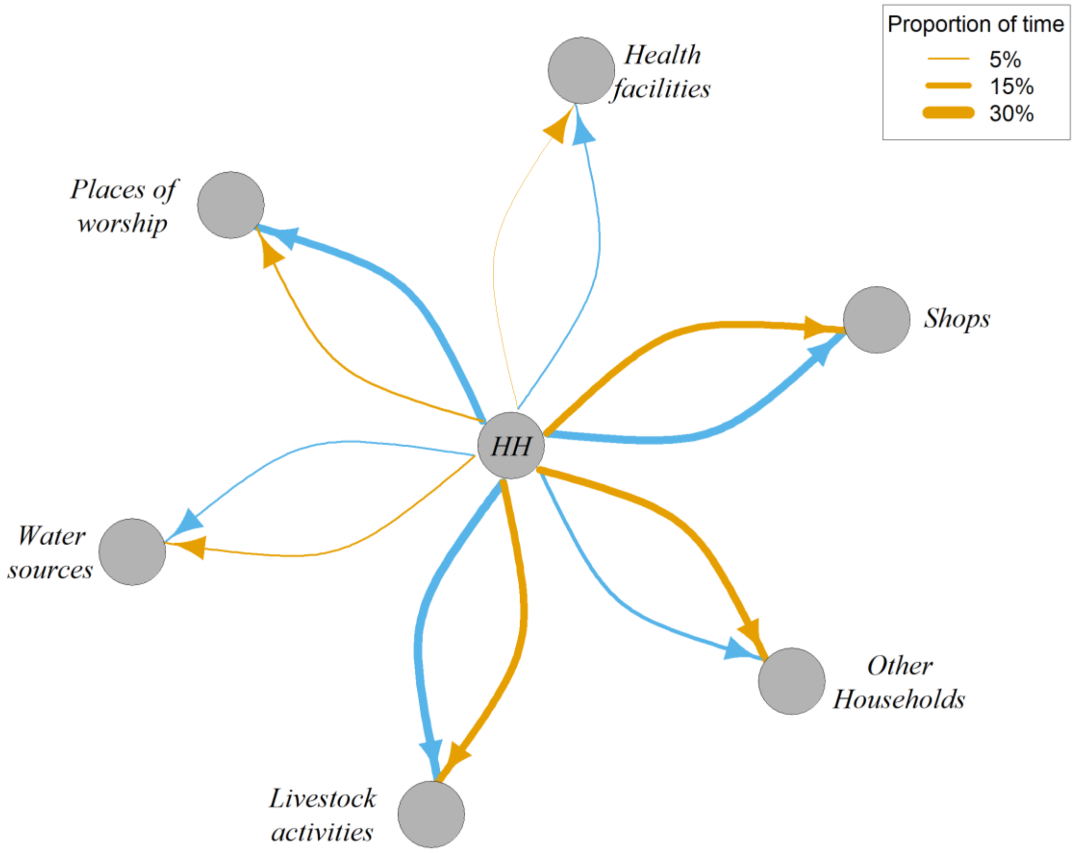


Figure 5-4 Network representation of mean proportions of time spent in different types of places.

Mean proportions of time spent from GPS (orange) and survey (blue) data. HH = Household.

5.5 Discussion

Geographic inaccessibility is a primary factor for many poor communities that are unable to easily use important infrastructure and natural resources, particularly in rural areas (Chamberlin & Jayne, 2013; Clark et al., 2003; Pearce et al., 2008). Even within these communities, accessibility can vary between individuals and can manifest differently when people are traveling for various types of resources (Perkins et al., 2014; Schröder et al., 2018). Studies have highlighted how the day-to-day activities of individuals can change their exposure to vector-borne diseases (Stoddard et al., 2009). Better understanding of how accessibility varies with type of activity and with an individual’s socioeconomic context could help inform future analyses on geographic accessibility. This socioeconomic and activity-based understanding of mobility can also help with identifying populations in communities at especially high risk of being unable to access healthcare and other services, which could help identify subgroups of populations most vulnerable to vector-borne diseases (Bardosh et al., 2017).

When comparing accessibility to a variety of resources I found that time people spent accessing different resources was not homogenous and had some links to demographic characteristics. Specifically, I observed that mobility related to livestock-related activities and health facilities were correlated with factors such as occupation, age and gender, while other types of activities like household visits and market visits were not. The small sample size and bias towards rural households mean that the representativeness of these results is limited, but nevertheless, combined with evidence of social differences in resource access from previous studies in this area (Okwi, Ndeng'e C, et al., 2007), they suggest further research into inequalities in resource access could be beneficial for improving individual and thus population welfare.

Broadly, I also found that people access different resources with variable frequencies and spent varying amounts of time there. For example, people visited places where livestock-related activities occurred and water sources either frequently or not at all, while other types of places like health facilities and places of worship were visited by most participants on a regular basis. The time spent at each type of place also showed large variation, with people spending the longest times at health facilities and places of worship, and the shortest times at other households and water sources. These findings, while unsurprising, demonstrate a heterogeneity in resource access between individuals. Travel times to health facilities in this study were longer than to other types of places – a finding which differs from results in Iquitos, Peru, where people travelled less far for health reasons than for commercial and household reasons (Perkins et al., 2014). This highlights how health facilities in rural communities such as the study area are likely more inaccessible than other resources when compared to access in an urban area such as Iquitos.

Notably, when comparing health facility visits and water source visits (two resources widely regarded as fundamental human rights), I observed that people spent longer both travelling to and at health facilities compared to water sources. However, travelling times to health facilities suggest that most people live within the government target of 5km from a health facility, and previous evidence from rural Kenya found that travel time to health facilities did not affect child mortality in areas with many facilities (Moïsi et al., 2010), leading to calls for more focus on social determinants and quality of care and other factors that influence healthcare access and use. My findings underscore this, as people reported having to spend a mean of 3 hours waiting at health facilities, compared to under 10 minutes to access water. Moreover, such long waiting times at health facilities is likely to be exacerbating vector-borne diseases in the area, as delays such as these are barriers to both seeking and obtaining the healthcare needed for disease prevention and control.

I also found some significant gender disparities in the amount of time people spent at different places, which could influence healthcare utilisation patterns. Men reported spending longer than women on livestock-related activities, while women spent longer than men at health facilities and water sources. Women reported visiting health facilities more frequently than men, possibly because they are usually responsible for their children's healthcare as well as their own (M et al., 1989). The differences in time spent by gender suggests that women are having to wait longer to access healthcare than their male counterparts. Previous studies in rural communities in Kenya and neighbouring Tanzania have found that waiting times, like time spent travelling to the facility, are a major barrier to accessing care and can result in delayed care-seeking behaviour, particularly for women due to the opportunity cost of accessing care over their domestic responsibilities (Mason et al., 2015; Mubyazi et al., 2010; Thaddeus & Maine, 1994). Since women both visit more often on behalf of their families and spend more time at health facilities, measures to reduce waiting times could have a direct benefit to both their and their families' health.

Finally, I compared the travel times and frequencies reported in GPS and survey data, to quantify potential biases in each and address the potential value of both datasets. The survey results are likely affected by recall bias and a social desirability bias. Notably, the GPS data were limited by the short data collection period and the small sample size, failing to capture visits to health facilities in the time available. Nevertheless, for places frequently visited during the data collection period, they more objectively reflected proportions of time spent compared to the survey data. Compared to surveys, GPS tracking technology could give a more complete picture of activity spaces in rural contexts but require long data collection periods and a thorough knowledge of the local area. This knowledge could be obtained through a variety of methods, such as a participatory mapping approach, and could help to reduce recall bias effects. Future studies may be able to overcome the time and cost issues of increasing the data collection period by utilising smartphone technology, which already capture GPS locations of individuals under certain conditions. The collection of smartphone data in combination with GPS loggers has recently been piloted in an urban area of the UK and showed that location histories from smartphones can be valid datasets for exploring individual movements (N.W. Ruktanonchai et al., 2018). In low-income areas of sub-Saharan Africa, this method would have a substantial bias towards wealthy people and is dependent on reliable cell phone network coverage, but smartphone ownership and demand for data plans have both been increasing in recent years (World Bank, 2012), making this a viable option for similar studies in the future.

These results suggest that demography and activity are important drivers of mobility, influencing how scientists should quantify geographic accessibility to resources such as health facilities. Exploring activity spaces of vulnerable populations in this way is the first step to understanding how activity-driven mobility may affect vector-borne disease vulnerability. Because different types of mobility manifest in different ways and occur with various frequencies, data such as GPS and survey data could be used to quantify mobility for specific types of activities, taking into account the advantages of each. Since healthcare-related mobility was particularly time-consuming and appears to be linked to several demographic characteristics in this small study, further research in this area could shed light on how people from different demographic and socioeconomic contexts access healthcare and therefore help to improve access in low-income settings and ultimately aid the control of vector-borne diseases.

5.6 Intellectual contributions

The author list for this published work is as follows: Jessica R. Floyd, Joseph Ogola, Eric M. Fèvre, Nicola Wardrop and Andrew J. Tatem and Nick W. Ruktanonchai. JRF conceived and designed the fieldwork, performed the fieldwork, analysed the data, prepared figures and tables and wrote and prepared the manuscript for journal submission. JO performed the fieldwork with JRF and provided insight and feedback on the study design. EMF and NW provided insight and feedback on conceptualisation of the fieldwork and the study design. AJT and NWR supervised JRF during the course of this project and oversaw that the analysis was executed in a scientific manner. All authors reviewed the manuscript and provided feedback before final submission to the journal.

Chapter 6 Estimating malaria parasite mobility in Mozambique using mobile phone records

6.1 Abstract

Despite large reductions in burden, malaria remains one of the primary causes of death in Mozambique. Prevalence of the major malaria parasite *P. falciparum* varies widely, from 1% in the most southerly provinces to 57% in the most northerly. As Mozambique is a source of malaria cases in Southern Africa, there are elimination efforts in the south of the country aiming to reduce both local burden and importations to relatively low-transmission countries nearby. However, human-mediated parasite movement combined with spatial and seasonal changes in transmission threaten the success of interventions by reintroducing parasites to areas targeted for elimination. Call detail records (CDRs) can provide unique insights into human movements in near real-time and have previously been used to quantify connectivity and identify sinks and sources of malaria in other countries.

In this study I use pseudonymized and aggregated CDRs from a sample of over 5.5 million subscribers of Vodacom Mozambique to quantify human movements over four months (February to May) of 2018. By combining these data with high-resolution malaria incidence maps, I estimate parasite importation and exportation rates and identify potential net sinks and sources of malaria across districts. I use a community structure algorithm to explore the connectedness of districts in Mozambique through human mobility and calculate parasite mobility between urban and rural areas at a high spatial resolution. I find that 61% of human-mediated parasite movements were between rural areas, compared to 36% of all movements. I observe strong community structures and a natural divide between five communities in the north and five in the south. Lastly, I estimate that 92% of cases in Maputo city districts and Maputo province districts are imported from nearby districts. These important findings suggest that parasite importations from the north are unlikely to threaten interventions to eliminate malaria from the south of Mozambique.

6.2 Introduction

Increased human mobility in recent years has led to higher connectivity and therefore higher pathogen transmission between populations, meaning that infectious diseases can spread further and faster than ever before, causing significant threats to population health on global, regional and local scales. For malaria, the conditions required for transmission vary spatially and temporally across landscapes and the populations that reside in them (Carter et al., 2000). This heterogeneity means that malaria exists in areas with self-sustaining endemicity characterised by local disease transmission but also enables transmission elsewhere through host mobility (N. W. Ruktanonchai et al., 2016). Since mosquitoes typically don't travel far (Charlwood et al., 1998; Touré et al., 1998), the malaria parasite is mainly transported from areas of high transmission by human hosts, allowing parasites to move to areas where they may not otherwise be present or abundant. Thus, understanding connectivity between hosts at local, regional and national levels is a key aspect of reducing malaria transmission.

In Mozambique, malaria remains one of the biggest causes of death despite recent large reductions in burden. There were approximately 10 million cases (5% of all cases worldwide) and 15,000 deaths in 2017 (World Health Organization, 2018a) most of them due to *Plasmodium falciparum* (the major *Plasmodium* species in Mozambique). Relatively high malaria transmission in Mozambique also poses a challenge for the lower-transmission countries nearby through population mobility. A recent study in a low prevalence area of South Africa found that the majority of positive malaria cases had travelled from Mozambique (Raman et al., 2020), while another study of the travel patterns of malaria cases in Swaziland found that travel to Mozambique was associated with an increased risk of infection (Tejedor-Garavito et al., 2017). Furthermore, malaria transmission within Mozambique varies widely, with relatively low prevalence in the south and higher prevalence in the north (Bhatt et al., 2015). In southern Mozambique, Maputo province is approaching elimination while Gaza and Inhambane provinces aim to achieve pre-elimination status by 2025. It is generally suspected that transmission in some areas in the south is sustained by host mobility (Moonasar et al., 2016), but evidence to confirm this hypothesis has been lacking. In Magude district in Maputo province a project to demonstrate the feasibility of elimination in southern Mozambique concluded with significant reductions of malaria incidence and prevalence through targeted vector control and mass drug administration (MDA) campaigns, but was unable to interrupt transmission (Aide et al., 2019). Studies have shown that in low transmission settings approaching elimination where most of the pathogens are carried by a relatively small percentage of the population (Bousema et al., 2012), identifying areas where mobility may be causing local transmission can provide valuable targets for intervention strategies (N. W. Ruktanonchai et al., 2016). In southern Mozambique, identifying these areas would help to prioritise limited resources and efficiently target efforts to achieve elimination.

High-resolution maps of malaria incidence can help identify areas where the highest transmission likely occurs, but understanding the factors driving this transmission in high-burden areas requires knowledge of parasite movements. Recent studies have measured parasite mobility directly using genetic data with promising results (Chang et al., 2019; Tessema et al., 2019), but methods are still being developed and gathering sufficient high-quality genetic data remains challenging, so proxies based on human movements have been more widely used (Buckee et al., 2013; Pindolia et al., 2012; Amy Wesolowski et al., 2012). Community structure algorithms can be applied to human mobility flows and malaria incidence data to infer natural communities linked by high numbers of human-mediated malaria parasite movements. This can help identify areas where inherent networks are linked to transmission occurring through host mobility (Nick W. Ruktanonchai et al., 2016; Andrew J Tatem & Smith, 2010) and can also help to minimise importation risk by detecting subcommunities where infected travellers are likely to pass through. However, few studies have been able to examine community structures through human-mediated parasite mobility in low-resource areas, mainly due to a lack of high-quality movement data.

Population mobility has traditionally been estimated using travel history surveys at national and sub-national levels, but these data are often limited by their small sample size and are typically cross-sectional in nature. The recent rise of mobile phone use in sub-Saharan Africa, estimated to have reached 456 million unique users representing 44% of the total population in 2018 (Global System for Mobile Communications (GSMA), 2019), has led to the use of call detail records (CDRs) to estimate mobility on large scales. CDRs are collected by mobile phone operators for billing purposes. They comprise a record of individual events (calls or texts) associated with a timestamp, a pseudonymized user ID (to preserve privacy) and the ID of the cell tower that the event was routed through. The cell towers are geolocated and therefore the approximate locations of individual users at a given time can be deduced through the calls and texts they make. Thus, while imperfect and biased, CDRs provide a unique insight into human movements with high spatial and temporal granularities. They have previously been used to identify sinks and sources of malaria in other countries, ultimately providing evidence to help national control programs more efficiently target efforts to control malaria. Previous studies have used CDRs to measure human movements and identify hotspots of malaria transmission in Namibia (N. W. Ruktanonchai et al., 2016)[2], Kenya (Amy Wesolowski et al., 2012) and Zanzibar (Le Menach et al., 2011).

Here, I use pseudonymized CDRs from over 5.5 million subscribers of a mobile operator in Mozambique from February to May of 2018 to quantify human movements and combine these with high-resolution malaria incidence maps to estimate parasite mobility at a district level. I use a high-resolution map of rural, peri-urban and urban areas in Mozambique to calculate human and parasite mobility between rural and urban areas, and a community structure algorithm to explore connectivity between districts. Lastly, I estimate the percentage of malaria cases that are imported into each district.

6.3 Methods

6.3.1 Population characteristics and mobile phone records in Mozambique

Mozambique is administratively divided into 11 provinces including Maputo city, and 161 districts including city districts, with a total population of 28.9 million in 2017 (National Institute of Statistics (INE) (Mozambique), 2017). Districts in Mozambique have populations ranging from 13,025 to 1,616,267, although most have populations in the range from 50,000 to 150,000. The size of the districts range from 7km^2 to 18,336km^2, with a mean area of 5,302km^2 and a mean distance between districts of 517km (determined through nearest neighbour analysis on district centroids in QGIS version 3.16.3 (QGIS Development Team, 2019)). The provinces of Nampula and Zambezia in the north of the country are the most populous with 6.1 and 5.1 million people respectively, followed by Tete in the north-west (2.8 million) and Maputo in the south (2.5 million) (National Institute of Statistics (INE) (Mozambique), 2017). The majority (63%) of the population reside in rural areas (World Bank, 2019).

6.3.2 Data sources

6.3.2.1 Call detail records

I estimated human mobility in Mozambique using pseudonymized CDRs from mobile phones, provided by Vodacom Mozambique and Vodafone Group. Vodacom has a 44% market share in Mozambique, making it one of the largest telecommunications providers in the country (Shi, 2019). The dataset covered the period between 1st February 2018 and 23rd May 2018 with a total of 10 billion events, which consisted of outgoing and incoming calls or short message services (SMS). I filtered these events to exclude improbable users which gave a total of over 5.5 million unique users. I defined an improbable user as one that had a mean of more than 150 outgoing events per day over the four months of data, as these events were likely to be telesales attempts made by bots. To protect the privacy of users, several safeguards were used: the data were pseudonymized by a third party of data engineers at Vodafone Group, I carried out all data analysis on Vodafone servers, and I excluded data from cell tower pairs where fewer than one person on average travelled daily between them to prevent potential reidentification of users. I carried out the extraction, processing and initial aggregation of the raw CDRs using Python over Apache Spark (Apache Software Foundation, 2018) in a high-performance computing cluster and then carried out further analysis of the aggregated dataset in R version 3.5.1 (R Core Team, 2019).

6.3.2.2 Population data

To scale up the population represented in the CDR dataset to actual population numbers in Mozambique, I used high-resolution human population predictions for 2018 from the WorldPop project. These data are gridded predictions of population at approximately 100m spatial resolution generated using a Random Forest modelling approach to produce estimates based on a combination of census data and a wide range of remotely-sensed and geospatial datasets such as landcover, settlements, satellite nightlights, topography and other sources (WorldPop et al, 2018). I also compared the number of unique users per district in the CDR dataset with estimated district populations calculated from the WorldPop data to see how well the number of users in the CDR data matched estimated populations in districts and found a positive correlation with an R^2 value of 0.713 (see Appendix C.1.1).

6.3.2.3 Malaria incidence data

To estimate the prevalence of malaria in districts and therefore the human-mediated malaria parasite movement between districts, I used maps of estimated monthly gridded malaria incidence for Mozambique modelled at 5km resolution for the months of February to May 2018 inclusive, provided by the Malaria Atlas Project (Colborn & Malaria Atlas Project, 2020). These data are modelled estimates of malaria incidence based on data reported at the health facility level and are subject to limitations. Importantly, the input data used to make these maps comes from a variety of sources and their inherent biases may be carried through to the final output. For example, malaria cases reported at the health facility level may be geographically distant from where cases are acquired, although efforts are made to account for this in the modelling approaches used (Weiss et al., 2019). Figure 6-1 shows the gridded mean malaria incidence estimates between February and May 2018 and the WorldPop gridded population estimates in 2018.

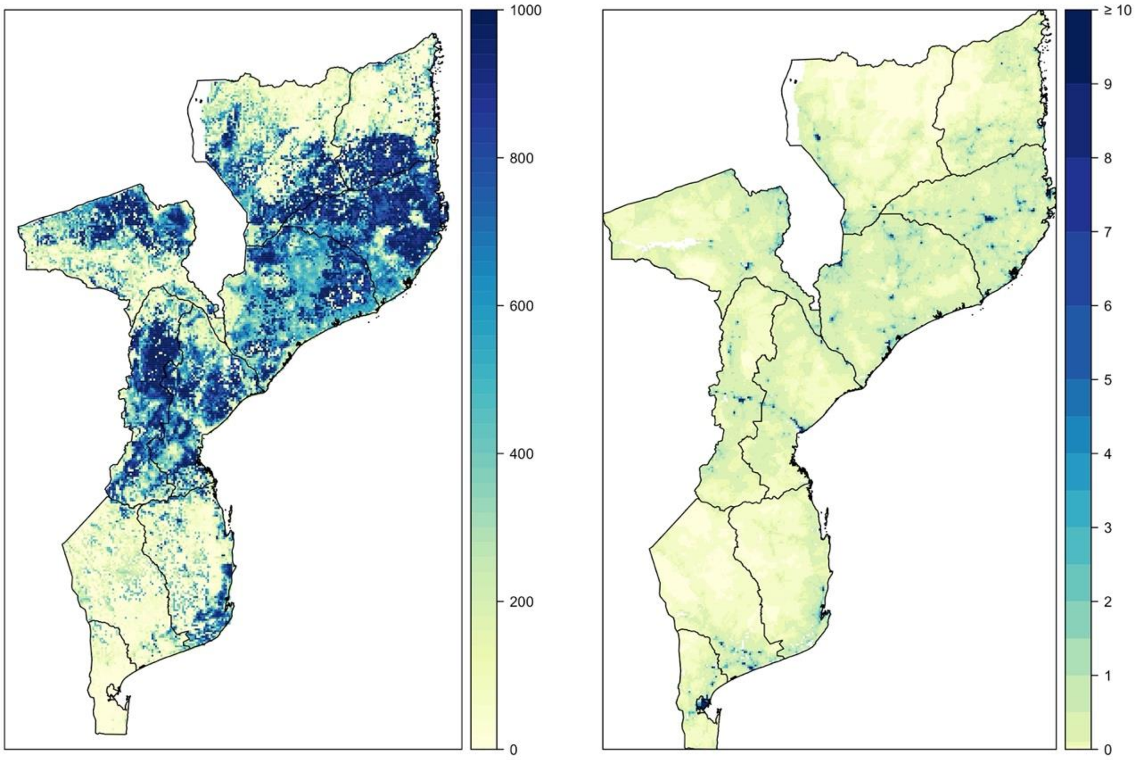


Figure 6-1 Malaria incidence and Population in Mozambique

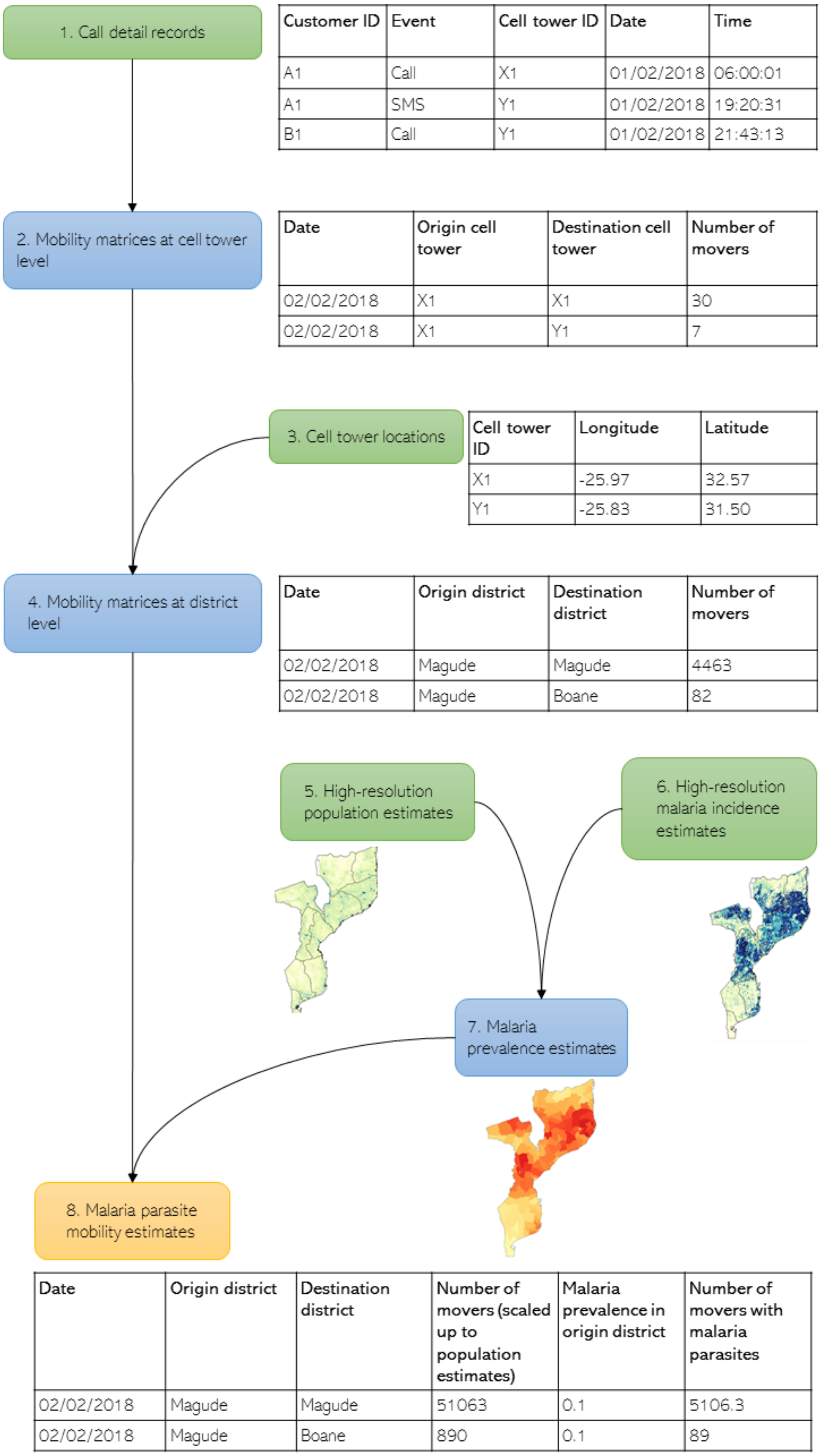
Left: Estimated mean malaria incidence per 1,000 people per year from February to May 2018 at 5km resolution. Right: WorldPop gridded population in 2018 at 100m resolution. Province borders shown in black. White areas indicate a lack of data.

6.3.3 Estimating human and malaria parasite mobility

Each event in the CDR dataset had an ID number that enabled identification of the cell tower that the event was routed through, so I was able to infer individual user movements through these geolocated events (Figure 6-2, Box 1). The exact locations of users were unknown, as antennas have specific beam shapes that define their coverage. For simplicity, I used the locations of the cell towers (Figure 6-2, Box 3) as the locations of users. The coverage of antennas differs with the technology used (2G, 3G, 4G) and the density of antennas varies across the country, with rural areas having lower density of cell towers (and thus more uncertainty as to the exact location of the customer). I used the location of the last event (call or text) each day to estimate where a subscriber spent the night, because the most important vectors of malarial parasites in Mozambique (*An. Gambiae*, *funestus and arabiensis* (World Health Organization, 2018a)) primarily bite at night. If there was no CDR for that day, I assumed the location was the same as the most recently located day. I processed the CDRs and used the last event of each day to calculate the numbers of users moving between cell towers per day (Figure 6-2, Box 2).

To produce mobility matrices, I aggregated the CDR-inferred locations to the district level and calculated the number of people moving between each origin-destination district pair using the *dplyr* package in R (Wickham et al., 2018) (Figure 6-2, Box 4). To account for variation in Vodacom market share in different areas of Mozambique, I scaled up the mobility matrices to WorldPop population estimates by dividing the WorldPop estimate of population in each district by the number of geolocated subscribers in each district to obtain scaling factors. I then multiplied the mobility matrices by these scaling factors in order to estimate actual population movements.

To estimate malaria parasite movements, I overlaid the WorldPop population estimates (Figure 6-2, Box 5) with the malaria incidence maps (Figure 6-2, Box 6) using the *raster* and *sf* packages in R (Hijmans et al., 2014; Pebesma, 2018). I used these overlays to estimate the prevalence of malaria in the population for each district and each month. This was calculated by multiplying the malaria incidence per person by the population estimate in each grid square, then summing these estimates in each district and dividing them by the total population estimate in each district (Figure 6-2, Box 7). I then used these estimates of malaria prevalence to weight the mobility matrices by the amount of malaria in each district and therefore obtain estimates of the numbers of people leaving each district with malaria (Figure 6-2, Box 8).



Customer ID	Event	Cell tower ID	Date	Time
A1	Call	X1	01/02/2018	06:00:01
A1	SMS	Y1	01/02/2018	19:20:31
B1	Call	Y1	01/02/2018	21:43:13

Date	Origin cell tower	Destination cell tower	Number of movers
02/02/2018	X1	X1	30
02/02/2018	X1	Y1	7

Cell tower ID	Longitude	Latitude
X1	-25.97	32.57
Y1	-25.83	31.50

Date	Origin district	Destination district	Number of movers
02/02/2018	Magude	Magude	4463
02/02/2018	Magude	Boane	82

Date	Origin district	Destination district	Number of movers (scaled up to population estimates)	Malaria prevalence in origin district	Number of movers with malaria parasites
02/02/2018	Magude	Magude	51063	0.1	5106.3
02/02/2018	Magude	Boane	890	0.1	89

Figure 6-2 Schematic showing the process of making malaria parasite mobility matrices.

Green boxes indicate raw data from data sources. Blue boxes indicate interim processed data. Yellow box indicates final processed data. Numbering is used to refer to the figure from the text. All data in this figure have been randomly generated and are given for example purposes only.

6.3.4 Urban and rural mobility

Spatial data from the GHS Settlement Model Grid (GHS-SMOD) (Pesaresi et al., 2019) were used to classify Voronoi areas around cell towers as predominantly urban, peri-urban or rural. This was done by overlaying the GHS-SMOD surface (1km resolution) with the WorldPop gridded population surface (100m resolution) and counting the number of people living in the eight different categories per area. I combined some categories to broadly classify areas as urban, peri-urban or rural (see Appendix C.1.5 for details). I then overlaid this map with a Voronoi tessellation of cell towers in Mozambique and used it to calculate the proportion of people in each Voronoi polygon that resided in each area. Voronoi polygons are polygons drawn around cell tower locations such that each polygon contains all possible points on a map closer to that cell tower than any other cell tower. This technique allowed the calculation of movements between rural and urban areas at the finest spatial scale possible (that of the cell towers). The polygons were classified as urban, peri-urban or rural depending on which category had the largest population living in it. I used the origin-destination matrices calculated at a Voronoi spatial resolution to quantify the flows of human and human-mediated parasite movements between the Voronoi polygons. In my analyses, 56% of Voronoi polygons were classed as rural, 42% as urban and 2% as peri-urban.

6.3.5 Community structure algorithm

I used the *cluster walktrap* (Pons & Lapaty, 2006) algorithm from the R package *igraph* (Csardi G, 2006) to determine the community structures in Mozambique through human-mediated parasite mobility. This algorithm identifies densely connected communities through a series of random walks in a network composed of vertices (in this case, the origins and destinations in the mobility matrix) linked by weighted edges (the number of movements between origins and destinations). A walker in the network takes a series of iterative random walks where the probability of moving from vertex A to vertex B is determined by the weight of the edge connecting A and B. The algorithm detects ‘communities’ of vertices linked by their connectivity: the more connected a community is, the more likely it is that the random walk stays in the same community. On specifying the desired number of communities, the algorithm returns the optimal structure for that number of communities and a modularity score, which is a measure of the strength of the division in the network, such that more isolation of communities receives a higher modularity score. I used the origin-destination matrices for human-mediated parasite movements between districts in Mozambique and applied the algorithm to detect the optimum number and structure of communities, determined by the maximum modularity score. Community detection algorithms have been used previously with CDR data to identify communities connected through human-mediated parasite movements as well as potential areas to target resources (Andrew J Tatem et al., 2014).

6.3.6 Local vs. imported cases of malaria

I used the CDR dataset to estimate the amount of time spent by people in each district in every other district, and to infer from this the proportion of malaria cases that are acquired locally. This analysis assumed that people get infected during travel with a probability proportional to the time spent at each location and are reported as cases at health facilities near their homes. Given this assumption, the rate of new cases that manifests near home I_h can be represented as

$$I_a T = I_h \qquad (1)$$

where I_a is the incidence of cases in the places where the transmission occurs, and T is how much time residents of various patches spend in each other patch. Therefore, the observed incidence I_h is the actual incidence I_r after adjusting for time that each community spends in each location with malaria risk.

I used this equation to calculate the *actual* incidence in each location using estimates of observed incidence I_h and time spent T. Further, by comparing the observed incidence in each location with the amount of actual incidence that manifested within the location's resident population ($I_{a(i)} * T_{(i,i)} \div I_{h(i)}$), I calculated the proportion of cases in each location that was acquired locally.

To do this, I first used the CDR dataset to estimate the amount of time spent T by people in each district, where element $T_{i,j}$ is the proportion of time that residents of district i spent in district j. I calculated this matrix by first assigning each person a 'home' cell tower, defined as the cell tower through which the highest proportion of last events of each day occurred. Then I calculated the proportion of last events that occurred at every other cell tower and calculated the mean total amount of time spent at other cell towers j by people from each home cell tower i to give an origin-destination matrix $T_{i,j}$ of time spent at each cell tower. Finally, I aggregated this matrix up to the district level to produce estimates of time spent per district by people in each home district.

With the values of T and I_h, I then calculated the rescaled actual incidence values I_a for each community by solving equation (1) for I_a:

$$I_a = I_h T^{-1} \quad (2)$$

I then calculated the number of local cases that manifested in each patch as the actual incidence in each patch that remained in that patch, or $I_{a,i} * T_{i,i}$.

6.3.7 Ethics statement

Ethical approval for this study was submitted and approved through the University of Southampton Ethics and Research Governance Council (submission ID 48113).

6.4 Results

6.4.1 Identifying sinks and sources of malaria by district

I estimated net daily malaria case movements into and out of each district using the human-mediated parasite mobility matrices scaled up to estimated population counts in each district. Because I observed little monthly variation in the net human-mediated parasite movements between districts (Appendix C.1.2) I used mean daily values over the four months in the dataset and subtracted imported cases from exported cases to make maps of net overall movements from each district (Figure 6-3, panels A and B). I also filtered these data to exclude movements that were less than 25km in distance, on the basis that these short-distance movements are likely to contribute less to sustaining transmission in areas approaching malaria elimination (Figure 6-3, panels C and D). I calculated that 43% of human-mediated parasite movements were over 25km in distance and a study of fine-scale individual movements in neighbouring Zambia defined long-disease movements as movement over 20km from an individual’s home (Searle et al., 2017). Given this and the coarser spatial granularity of the cell towers compared to GPS data, I estimated that 25km was a reasonable distance for this study.

I categorised the districts in Mozambique by whether the district had more human-mediated parasite movement out of it than into it (sources) or vice versa (sinks). While the biggest sources of malaria through human movements are concentrated in the northern provinces, I found that some districts in the south of the country were locally important sources of malaria through human movements, exporting several hundred cases of malaria daily and likely contributing to sustained transmission in the southern provinces. I observed directional movements from source districts into sink districts, which were often smaller districts corresponding to cities. Districts corresponding to the cities of Xai-Xai and Inhambane were the largest net sinks in the southern provinces, with many more cases moving into these areas than out of them.

The relative importance of districts as sources of malaria cases becomes apparent on examining their sink/source status under all movements and movements over 25km. For example, the district of Inharrime in Inhambane province (Figure 6-3, labelled in red) is a net sink when all movements are considered, importing 103 cases net daily, but switches to a net source when only movements over 25km are included, exporting 704 net cases each day and highlighting the importance of such districts in the movements of malaria cases over relatively long distances. For Maputo province (including Maputo city), I found that only 2% of all parasite movements originated from districts outside the province (primarily from Gaza and Inhambane provinces), rising to 29% for movements that were at least 25km in distance.

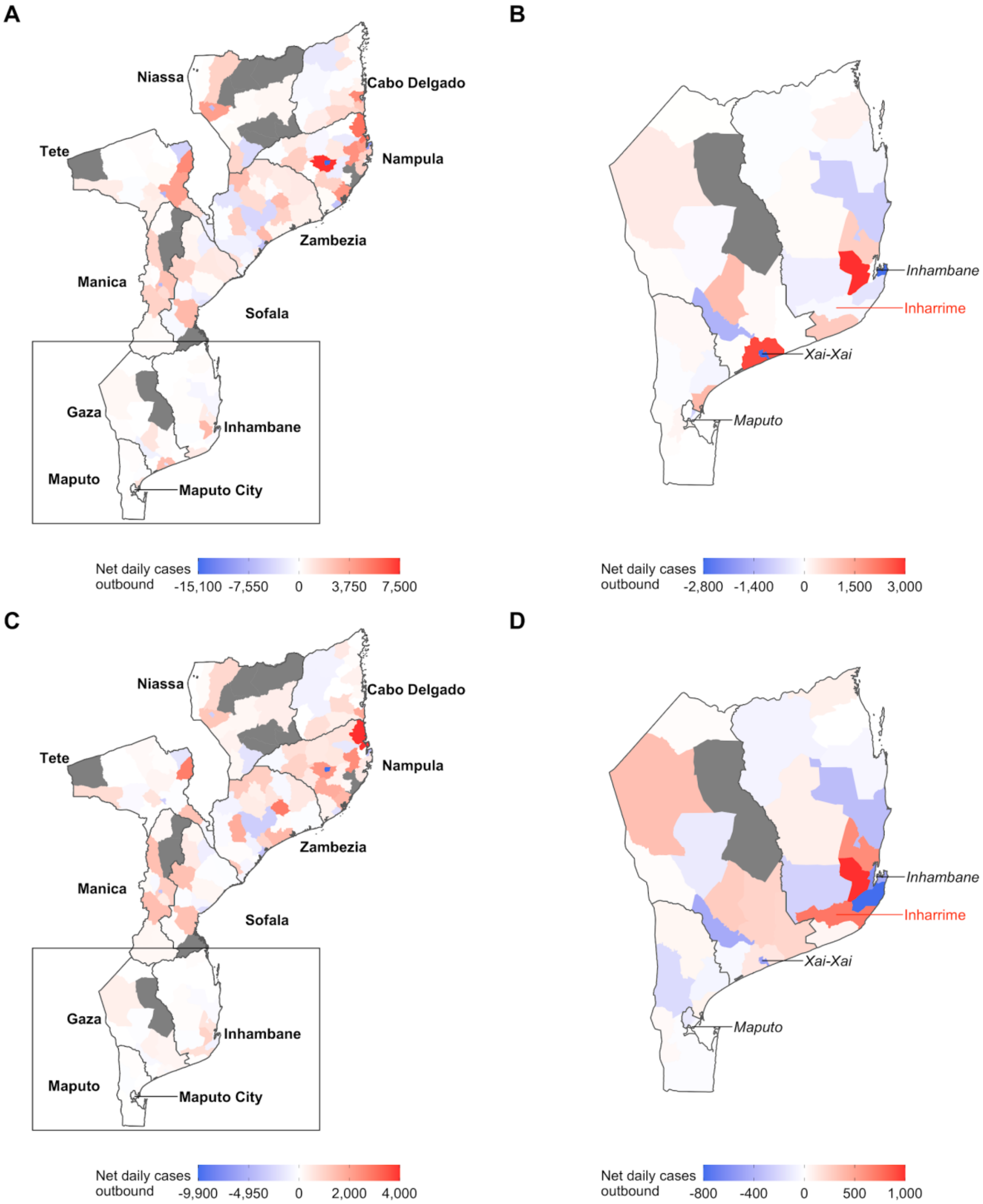


Figure 6-3 Net mean daily movements for human-mediated malaria parasite flows by district for February to May 2018.

A Flows for all districts in Mozambique. **B** Net daily movements for districts in the southern four provinces. **C** Net daily movements that were > 25km in distances for all districts in Mozambique. **D** Net daily movements that were > 25km in distance for districts in the southern four provinces in Mozambique. Grey areas indicate regions where the amount of data available did not meet privacy thresholds as detailed in the methods. Provinces are labelled in bold, major cities in the southern provinces are labelled in italics and the district of Inharrime is labelled in red.

Table 6-1 shows the top ten net sink and net source districts with mean net daily movements over 25km, illustrating how five out of the top ten net sources were in the province of Nampula, which has both a high population and high malaria incidence. Nine out of the top ten net sink districts were cities with relatively high numbers of inbound cases per day. For movements over 25km, the median distances travelled out of the top ten source districts were generally smaller than those into the top ten sink districts, ranging from 30km to 67km out of sources and 34km to 116km into sinks (Table 6-1).

Table 6-1 Top ten net source and net sink districts for daily net mean human-mediated parasite movements.

Top net source districts				**Top net sink districts**			
Province	*District*	*Estimated mean net outbound cases per day*	*Median distance (km)*	*Province*	*District*	*Estimated mean net inbound cases per day*	*Median distance (km)*
Nampula	Memba	3,930	32	Nampula	Cidade De Nampula	9,893	79
Zambezia	Mulevala	2,790	36	Nampula	Nacala-Porto	3,940	52
Tete	Tsangano	2,722	39	Manica	Cidade De Chimoio	3,805	45
Nampula	Monapo	2,433	44	Zambezia	Cidade De Quelimane	3,791	44
Nampula	Rapale	2,284	30	Zambezia	Mocuba	3,175	116
Cabo Delgado	Mecufi	1,864	37	Tete	Cidade De Tete	3,160	94
Nampula	Mogovolas	1,836	61	Niassa	Cidade De Lichinga	3,120	100
Zambezia	Derre	1,800	42	Cabo Delgado	Cidade De Pemba	2,710	92
Nampula	Angoche	1,584	67	Sofala	Cidade Da Beira	2,249	42
Zambezia	Maganja Da Costa	1,554	35	Nampula	Larde	1,967	34

6.4.2 Urban and rural host and parasite mobility

To explore the directional mobility further, I calculated the proportions of both human and human-mediated parasite movements that occurred between urban, peri-urban and rural areas. Table 6-2 shows how although the majority (53%) of all movements were between urban areas, the majority of human-mediated parasite movements were instead between rural areas (61%). Moreover, compared to human movements, parasite movements occurred more often in the direction of increasing urbanicity than vice versa, demonstrating that the sources of malaria tended to be in rural areas.

Table 6-2 Relative human and human-mediated parasite movements between urban, rural and peri-urban areas in Mozambique.

Movements in the direction of increasing urbanicity are highlighted in red.

Type of movement	**Human mobility (% of all movements)**	**Parasite mobility (% of all parasite movements)**
Rural to Rural	36.20	60.72
Urban to Urban	52.99	28.83
Peri-urban to Peri-urban	1.54	0.87
Rural to Urban	3.81	4.77
Urban to Rural	3.84	3.96
Peri-urban to Urban	0.55	0.30
Urban to Peri-urban	0.55	0.24
Rural to Peri-urban	0.26	0.18
Peri-urban to Rural	0.26	0.14

6.4.3 Community structures at district level

To explore connectivity through human-mediated parasite movement in Mozambique and address the question of how separate the northern districts are from the southern districts, I modelled human-mediated malaria parasite connectivity between the 161 districts in Mozambique across the four-month period of study. Using a community structure algorithm to measure connectivity, I found that a 10-community structure best approximated the inherent communities for human-mediated parasite movements between districts in Mozambique, with a modularity score of 0.792 indicating a high degree of isolation. On examining movements between these communities, I observed a natural divide between the most northerly 5 communities, roughly corresponding to the most populous provinces of Nampula, Zambezia, Cabo

Delgado and Niassa, and the five communities south of these (Figure 6-4), with relatively little movement of malaria cases between these two sets of five communities. The most southerly community encompassing Maputo city, Maputo province and Gaza province saw much more movement of malaria cases into the community than out of it, with 87% of human-mediated parasite movements inbound compared to outbound. Of the inbound movements, 47% were from Inhambane province and 41% from Sofala province, suggesting that interventions to reduce transmission in these neighbouring provinces would both reduce local burden and aid elimination efforts in Maputo and Gaza provinces.

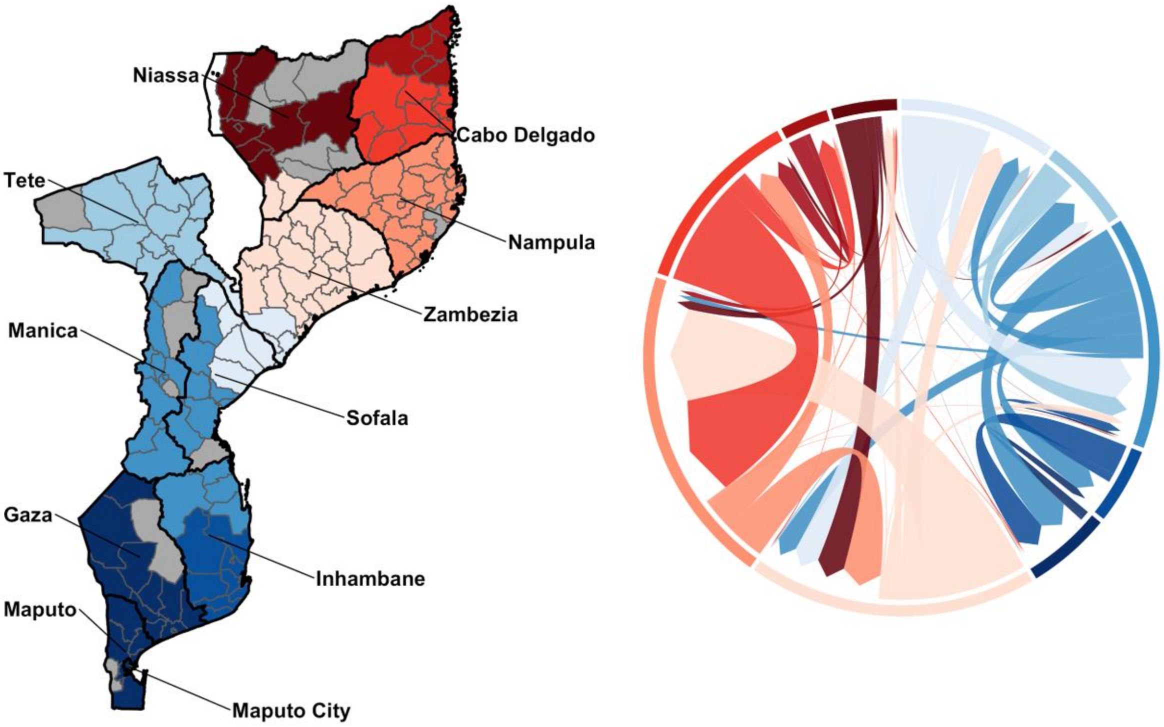


Figure 6-4 Community structure and connectivity though malaria parasite movements in Mozambique.

Left: Map of province borders and community membership through human-mediated parasite movements between districts; modularity score = 0.792.
Right: Directional network of relative flows between communities through human-mediated parasite movements. Colours denote districts belonging to the same subcommunity, grey areas indicate regions where the amount of data available did not meet privacy thresholds as detailed in the methods.

6.4.4 Disaggregating local from imported transmission

An important distinction when considering the impact of mobility on malaria incidence is whether an area has a high number of cases because they are acquired locally or are imported from elsewhere. I aggregated the high-resolution malaria incidence maps from the Malaria Atlas Project to the district level (Figure 6-5, panel A) and used the time spent matrix from the CDRs to produce a map of 'disaggregated' malaria incidence, or incidence weighted by the time spent by people in each district (Figure 6-5, panel B). I used these to estimate the proportion of cases that were imported as opposed to occurring locally (Figure 6-6) and found high proportions of imported cases in the districts in and surrounding the capital city of Maputo. I estimate that all cases in the five *Distritos Urbanos* in Maputo city are imported from other districts, compared to only 33% and 3% of cases in the remaining two districts of KaTembe and KaNyaka, respectively. On average, I found that 92% of cases in Maputo province (including Maputo city) districts are imported from outside the district, almost entirely from districts in the southern provinces, while mean importation proportions for districts in other provinces ranged from 8% to 27%.

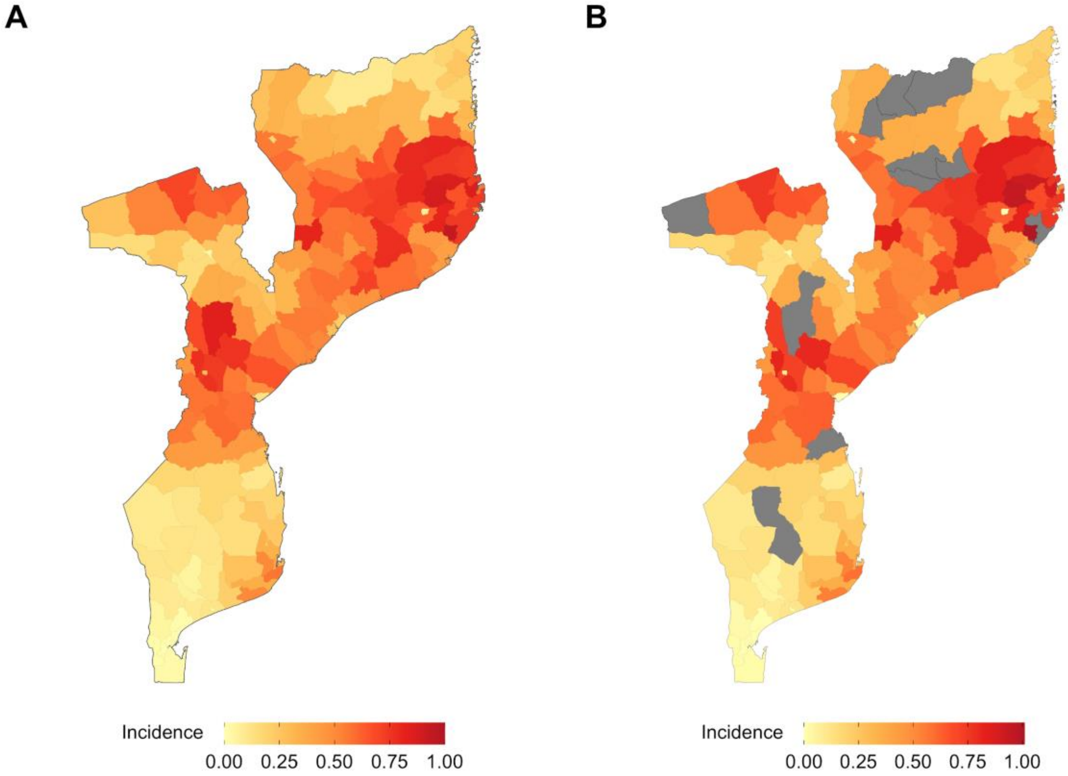


Figure 6-5 Mean estimated malaria incidence by district in Mozambique in 2018

A Estimated mean malaria incidence per person per year, averaged over the months February to May 2018. From (Colborn & Malaria Atlas Project, 2020). **B** Disaggregated incidence from CDRs. Grey areas indicate regions where the amount of data available did not meet privacy thresholds as detailed in the methods.

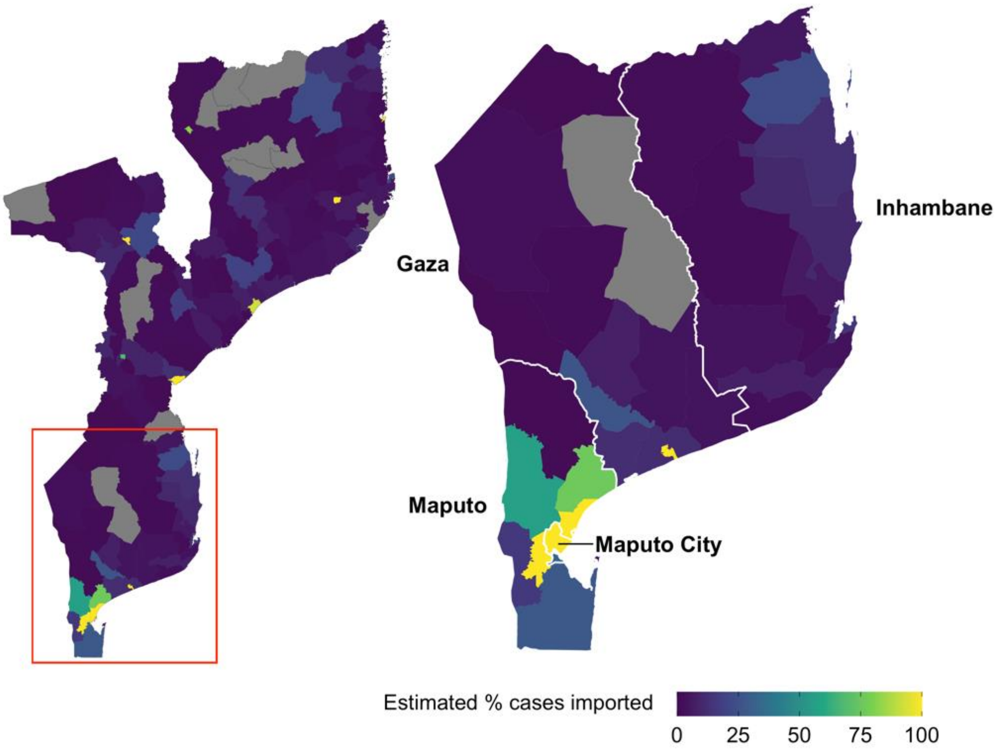


Figure 6-6 Estimated percentage of imported malaria cases for districts in Mozambique.

Grey areas indicate regions where the amount of data available did not meet privacy thresholds as detailed in the methods.

6.5 Discussion

In this study, I quantified the connectivity of mobility networks in Mozambique through *P. falciparum* malaria parasite movements, the proportion of human and parasite movements that were between urban and rural areas and the rates of parasite importation into districts. To the best of my knowledge, this is the first study to analyse pseudonymized and aggregated CDRs from Mozambique to address these issues and helps improve understanding of how mobility contributes to the spread and persistence of malaria across the country.

This analysis of the mean daily numbers of malaria cases moving into and out of districts showed high heterogeneity between districts, with some exporting many more cases than they were importing and vice versa. I also observed directionality between source and sink districts, with eight out of the top ten sink districts corresponding to cities. When focusing on movements of malaria parasites over relatively long distances, I saw that three districts in Inhambane province had up to 1,000 potential malaria cases moving over 25km daily out of the districts. Current elimination efforts in the south of Mozambique focus on Maputo province, but as this area approaches malaria elimination, targeting interventions to nearby districts such as these that export the highest numbers of malaria cases in the south could help to efficiently reduce malaria transmission here.

I quantified the movement of malaria cases between urban and rural areas and found that while the majority of human movements were between urban areas, most human-mediated parasite movements were between rural areas. I also observed a pattern in movements along the rural to peri-urban to urban spectrum, with human-mediated parasite movements higher in the direction of increasing urbanicity than vice versa. The difference in malaria transmission between rural and urban areas in sub-Saharan Africa has been previously documented, with rural areas tending to sustain much higher entomological inoculation rates (EIR) compared to urban areas. As such, much of the burden of malaria falls on rural populations (Hay et al., 2000; Robert et al., 2003). The findings suggest that while the bulk of malaria movements are between rural areas where most people live, human mobility from rural to urban areas is likely an important driver of malaria transmission in urban areas in Mozambique.

I observed strong community structures through human-mediated parasite movements between districts, suggesting that these movements occur mainly in distinct communities that could serve as targets for control. The ten-community structure had a remarkably high modularity (0.792) over the four months examined in this study indicating a high level of isolation and suggesting that malaria may be more easily controlled if interventions target these areas as foci rather than targeting individual districts, due to both the connectivity within communities and to the relatively low risk of importation from other communities. On examining parasite movements between communities, I found an inherent north-south divide occurring between the five communities in the north-east of the country and the five communities to the south of these, with relatively few malaria movements between the two. Furthermore, the ten communities identified did not always correspond to province borders, suggesting that province-based efforts to control malaria may be undermined by movement across province borders. This is the first study in the literature to quantify community structures through malaria parasite mobility in Mozambique, so combining these data with parasite genetics or other epidemiological data would help to validate the communities identified here. A previous study in southern Africa using similar community analysis found good concordance between community structures from CDRs with parasite genetic data and demonstrated that identification of these communities can provide valuable targets for resources to control and eliminate malaria (Tessema et al., 2019).

Finally, I disaggregated imported cases from locally acquired ones and found high proportions of imported cases in cities, particularly in the south. I estimate that 92% of the malaria incidence in districts within Maputo province (including Maputo city) is due to importation from other districts. At the same time, my estimates of parasite mobility suggested that 98% of all parasite movements into districts in Maputo province originated from districts within the province. Therefore, the majority of case importations into districts in Maputo province are likely to be from districts within the province. The high level of isolation through parasite movements of the community of Maputo and Gaza provinces supports this hypothesis, and suggests that elimination here would not be impeded by importation of cases from other provinces. Current efforts to eliminate malaria from the south of Mozambique rely on the assumption that the importation rate of malaria from the north of the country through human movements is low enough to achieve sustained elimination, with a particular focus on the provinces of Maputo, Gaza and Inhambane (Moonasar et al., 2016). My findings support this hypothesis, although sources of malaria in Inhambane province may undermine efforts to achieve pre-elimination status in this community. Targeting resources towards districts in Inhambane province could reduce transmission both locally and in nearby districts.

CDRs are a valuable source of mobility data due to their large size, wide coverage area and fairly high penetration rate in Mozambique. However, I acknowledge that there are limitations to their use. The CDR dataset only spanned 16 weeks, limiting my ability to observe potential effects of seasonality of malaria transmission on the net movements of malaria cases. Specifically, these weeks represent the high transmission season, and may not be representative of movements in the low transmission or other seasons. Although high, the spatial resolution of the data is limited to the density and coverage of the cell towers of the mobile operator. The temporal resolution is dependent on how often events (call or text) occur for each user. I found that the density of cell towers is higher in urban areas, with 42% of all Voronoi polygons being classed as urban, which introduces bias towards people living in urban areas and may overestimate the importance of movements into and out of these areas compared with rural areas. Moreover, the high-resolution population and malaria incidence datasets are estimates which are subject to biases inherent to the input data used to produce the estimates. For example, malaria cases may be reported to health facilities far from where the infection was actually acquired. Finally, the lack of demographic data from the subscriber population makes it impossible to know how representative the sample population is of the whole population of Mozambique – it is highly likely, for example that the movements of children are underrepresented in the dataset, which is detrimental to the results because children are at higher risk of severe malaria than the general population due to their lack of immunity. Moreover, it is likely that women are underrepresented in the sample population, given recent statistics on mobile phone ownership by gender (60% of men and 45% of women own a mobile phone in Mozambique, a gender gap of 24% (GSMA, 2019)). Nevertheless, the large sample size (over 5.5 million subscribers from the total population of 28.9 million people) helps to mitigate these issues.

These results are particularly important in the broader context of malaria elimination in Southern Africa. Since much of the malaria in Eswatini and South Africa is likely transported across the border from Mozambique, work like this that can provide evidence to help the National Malaria Control Programme better allocate resources to reduce malaria transmission are valuable not only to Mozambique, but to the wider region of Southern Africa. Globally, this study demonstrates the value of CDR datasets such as these and provides a strong case for their more general use for various applications in public health and beyond.

6.6 Intellectual contributions

The author list for this work to be submitted for publication is as follows: Jessica R Floyd, Nick W Ruktanonchai, Pedro Rente Lourenço, Nuria Oliver, James Colborn, Baltazar Candrinho and Andrew J Tatem. JRF conceived and designed the study, performed the data extraction and analysis, prepared figures and tables and wrote and prepared the manuscript for journal submission. PRL performed data extraction and provided insight and feedback on the study design. NO provided insight and feedback on the study design and analysis. JC and BC provided feedback on the study design and on the implications of study findings relevant to the study region. AJT and NWR supervised JRF during the course of the project and oversaw that the analysis was executed in a scientific manner. All authors reviewed the manuscript and provided feedback before final submission to the journal.

Chapter 7 Conclusion

The analysis of individual-level mobility data has proven a valuable tool for assessing vulnerability and exposure to vector-borne disease across spatiotemporal scales, but few studies have been able to incorporate this knowledge, particularly in low-resource settings. In this thesis, I aimed to address this knowledge gap by exploring how individual-level mobility patterns vary at different ends of the spatiotemporal scale, and how these patterns are often linked to demographic and landscape characteristics with important implications for resource access and exposure to vector-borne diseases. Specifically, I aimed to 1) explore demographic and seasonal differences in the movement patterns of a human and livestock population in rural Kenya at a fine scale; 2) quantify the activity-driven movement patterns of a population in rural Kenya and explore links between these and demographic characteristics; and 3) quantify the movement patterns of a large population in Mozambique and link these to incidence of malaria. In this section, I summarise the key findings of the thesis, discuss the contributions made to the literature and outline the limitations of the work. I then discuss further research that could build on the work presented here and the potential impact of the findings before reviewing the final conclusions.

7.1 Key findings

Table 7-1 presents the key findings of this thesis, broken down by study objective for the three papers presented in Chapters 4,5 and 6, and includes a brief description of the research questions, gaps in the literature addressed by the study and the main datasets and methods used. In Chapter 4, I used survey methods and GPS loggers to collect primary data on the movements of people and their ruminants and found significant differences in movement patterns of humans between seasons, with humans spending more time in swampland during the short rainy season compared to the dry season. Moreover, I saw that humans took longer trips, travelled further and had larger home ranges in the dry season compared to the rainy season, a pattern that was not observed in their ruminants and could have important implications for seasonal transmission of vector-borne diseases. Finally, I found the people and ruminants from relatively poorer households travelled further that wealthier households in the study area.

To address the second objective of the thesis, in Chapter 5 I analysed the survey and GPS data collected from a sample of the human population in Busia county and found that the travel times and time spent on trips outside of the household varied depending on the category of place that they were visiting, and that people spent longest at health facilities. These travel times and time spent were correlated with some demographic characteristics, and women in particular spent longer than men at health facilities and at water sources, which are likely to reflect their different exposure and vulnerability to vector-borne diseases. On comparing the survey and GPS data collected from the sample population, I found significantly different reported times spent for particular categories of places, reflecting how the GPS loggers were likely able to more objectively capture the activity spaces of people than the self-reported survey data.

In the third and final paper in Chapter 6, I analysed a CDR dataset from over 5.5 million people in Mozambique to quantify movement patterns in the context of malaria incidence across the country, and found directional flow of malaria parasites through human-mediated malaria parasite movements in a urban to rural direction. I estimated that 92% of malaria cases in districts of Maputo province and Maputo city are imported from other districts. Finally, I discovered a natural north-south divide between communities linked by human-mediated parasite movements, suggesting that malaria may be eliminated from the southern five communities identified with only limited surveillance required to prevent resurgence.

Table 7-1 Key research findings broken down by study objective

Study objective	Research questions	Gaps in knowledge	Data and methods	Key findings
Chapter 4 To explore demographic and seasonal differences in the movement patterns of a human and livestock population in rural Kenya at a fine scale	• How often, for how long and where do humans and livestock move? • Do human and livestock movements differ significantly between seasons? • What demographic characteristics, if any, are linked to these movement patterns?	• Vector-borne disease transmission depends on host movement, but host mobility in this context is poorly understood • No previous studies have measured human and livestock movement simultaneously	• Generalized linear mixed model (GLMM) with fixed and random effects to measure relationship between movement metrics (Time spent on trips, maximum distance travelled, home range and trip frequency) and covariates • Univariable beta regression to estimate odds of spending time on different types of land cover • Survey data at individual and household levels from 75 participants • GPS dataset collected from 26 humans and 20 ruminants in the short rainy season, and 25 humans and 15 ruminants in the dry season	• In the dry season, people took longer trips, travelled further on them and had larger home ranges compared to the short rainy season • People and ruminants from relatively poorer households travelled further than those from wealthier households • Humans spent more time in swampy areas during the short rainy season than the dry season • No specific seasonal differences in movement patterns for the ruminant population studied

Study objective	Research questions	Gaps in knowledge	Data and methods	Key findings
Chapter 5 To quantify the activity-driven movement patterns of a population in rural Kenya and explore links between these and demographic characteristics	• How often do people visit different categories of places in their local area, and how long do they spend there? • Are demographic characteristics correlated with time spent at particular categories of places? • How well does the time spent at different places calculated from the GPS data reflect the time spent reported in the survey data?	• Access to resources is linked to population vulnerability to vector-borne disease (Bardosh et al., 2017) • Studies of mobility and resource access tend to focus on one type of resource • The activity spaces of rural populations in resource-poor settings is particularly understudied	• Linear mixed models (LMMs) with fixed and random effects were used to measure relationship between the movement metrics and categories of place visited • Paired t-test to test for significant differences between survey-reported and GPS-captured data • Survey data at individual and household levels from 75 participants • GPS dataset collected from 26 participants	• People spent longest at health facilities, market and places of worship compared to other types of places • Mobility related to livestock activities, health facilities, water sources and places of worship were correlated with demographic factors while mobility to other types of place were not • Women spent significantly longer than men at health facilities • The GPS data more objectively reflected proportions of time spent at different categories of places compared to the survey data

Study objective	Research questions	Gaps in knowledge	Data and methods	Key findings
Chapter 6 To quantify the movement patterns of a large population in Mozambique and link these to incidence of malaria.	• Where are human-mediated malaria parasite movements occurring in Mozambique? Specifically: • Which districts are the highest net importers and exporters of malaria parasites? • Are there differences in human-mediated malaria parasite movements between rural and urban areas? • How connected are districts in Mozambique through human-mediated malaria parasite mobility? • How much of the malaria incidence in districts in southern Mozambique is due to importation from other districts?	• *P. falciparum* prevalence in Mozambique varies widely across the country, from 57% in the north to under 1% in the south • CDR datasets have been previously used to quantify population movements in the context of malaria • No studies have explored how population mobility is linked to malaria incidence in Mozambique	• Call detail record (CDR) dataset with over 10 billion records from over 5.5 million unique users • 5km malaria incidence maps from MAP (Colborn & Malaria Atlas Project, 2020) • High-resolution gridded population maps from the WorldPop project (WorldPop, 2018) • Spatial data on urban, peri-urban and rural areas from GHS Settlement Model Grid (GHS-SMOD) (Pesaresi et al., 2019) • Cluster walktrap algorithm from the R package *igraph* (Csardi G, 2006)	• Identified net sources and sinks of malaria parasites in Mozambique, with top net sink districts corresponding to cities • 92% of malaria cases in Maputo province districts and Maputo city districts are imported from other (primarily local) districts • Most human movements are between urban areas, but most human-mediated malaria parasite movements are between rural areas • Observed a natural north-south divide amongst district communities in Mozambique

7.2 Thesis contributions

This thesis explored the fine-scale movement patterns of a small rural population, including their activity spaces and demographic characteristics, and quantified the large-scale movement patterns of the population of Mozambique in the context of malaria transmission. Consequently, this thesis makes some important contributions to the current scientific literature.

In Chapter 4, I used GPS loggers and surveys to track the movements of humans and their livestock in two separate seasons. While previous studies have used GPS loggers to quantify the movements of humans and animals separately in the context of infectious diseases (Parsons et al., 2014; Searle et al., 2017), this study was novel in its use of GPS loggers to track both simultaneously. Moreover, the study contributes valuable data by tracking the same individuals across two seasons, allowing for the comparison of fine-scale movement patterns across seasons. Host movements (both human and animal) are key to vector-borne disease transmission yet are poorly understood, particularly in resource-poor areas where the burden of disease is highest. This study sheds light on the movement patterns of a small rural population in western Kenya, where the burden of zoonotic and vector-borne diseases is high and provides evidence to support further investigation into the link between movement patterns and vector-borne diseases on a fine scale.

Chapter 5 presents a study of the activity spaces of the same rural population described in Chapter 4 and is the first study to quantify the time spent at different types of place using GPS loggers in a rural population. Bardosh *et al.*, 2017 described how the accessibility of various resources is linked to household vulnerability to vector-borne diseases, and previous studies have explored the links between health facility access and various vector-borne diseases, particularly malaria (Alegana et al., 2012). However, the accessibility of local infrastructure for people in rural communities in relation to vector-borne disease is particularly understudied, and as such these analyses present novel findings and underscore the importance of understanding how movement patterns are linked to demographic characteristics and affect not only health facility access, but also access to several different types of resource and therefore are likely to be linked to vulnerability to vector-borne disease. This highlights the importance of disaggregating movement patterns by demographic characteristics in order to understand how they affect people's vulnerability to vector-borne diseases differently.

The analyses comprising the final study presented in Chapter 6 are the first to quantify population movement in the context of malaria transmission in Mozambique. Previous studies have used similar datasets and methods to analyse population movements and relate them to malaria transmission in other countries in sub-Saharan Africa, but this study was the first to do so for Mozambique, a country where all 28.9 million residents are at high risk of malaria and 15,000 die of it each year. In addition, the analysis of this large-scale CDR dataset contributes further evidence to the literature about the value of such datasets and could potentially encourage future collaborations with telecommunications providers for addressing challenges posed by other infectious diseases such as dengue fever. Finally, the use of community structure algorithm in this study identified distinct communities in terms of malaria parasite movements, which provides vital information to decision-makers because the relative isolation of districts in the south means that elimination efforts here are unlikely to be undermined by movements from the north, and suggests that only limited surveillance is necessary to sustain elimination.

7.3 Limitations

The analyses in the preceding chapters were subject to several limitations, as outlined in the discussion sections of each chapter. Broadly, all three studies likely suffered from spatial bias inherent in the sampling methods. The studies in Chapters 4 and 5 that employed a simple random sampling strategy for household selection were biased towards people living in rural areas, because the selection of households was not weighted by population density. To overcome the bias in household selection in future studies, high-resolution population density data like those produced by the WorldPop project (WorldPop (www.worldpop.org - School of Geography and Environmental Science, University of Southampton; Department of Geography and Geosciences, University of Louisville; Departement de Geographie, Université de Namur) and Center for International Earth Sc, 2018) could be used to weight the selection of various points on the map. The study in Chapter 6 likely suffered from the opposite bias, in that people living in urban areas in Mozambique are more likely to own mobile phones than people living in rural areas (Saúde/INS et al., 2019), and therefore are likely to be overrepresented in the CDR dataset although I took steps to account for this limitation by multiplying the observed population in each district by a scaling factor weighted by the number of people residing in the district.

Chapters 4 and 5 relied on the use of GPS loggers to track the fine-scale movements of people and their animals. The use of GPS devices to track people and animals in urban and rural settings has been previously documented (Parsons et al., 2014; Searle et al., 2017) and shown to be a reliable tool for tracking movements, with good portability, weight and battery life. However, the use of these devices for humans is subject to several limitations, one of which is the users' agreement to wearing the device. For humans, full informed consent was obtained before issuing the GPS logger and participants were politely requested to wear the device as much as possible, but it was ultimately down to the participant's choice where and when to wear the device, and therefore there were likely periods during the day where the device was not worn. The motion-detection function of the GPS device helped to account for this limitation, by allowing the identification of periods during daytime hours when the device was likely not being worn.

Survey methods such as the ones used in Chapters 4 and 5 that rely on the ability of participants to recall places they have visited are inherently limited by recall bias: participants are more likely to remember places that hold higher importance to them, and less likely to recall places that they perceive to be less important. I attempted to minimise this limitation by using structured surveys with few open-ended questions and by using memory recall aids such as providing categories of places participants may have visited using publicly available data on local infrastructure and results of previous surveys in the area. I also compared the results of the survey that relied on participants recollection of places to the GPS results to see how well they matched and found similar patterns of visitation for some categories of place, as detailed in Chapter 5.

The studies in Chapters 4 and 5 had small sample sizes due to resource constraints, which limited the ability to draw generalisable conclusions from the results. Furthermore, since people and their animals were only tracked for one week at a time, it is inadvisable to extrapolate these results temporally. Nevertheless, the tracking of the same individuals in different seasons and comparison to answers given in the travel history section of the individual surveys helped to mitigate these limitations, by extending the utility of the data available.

The analyses in Chapter 6 relied on a CDR dataset to quantify people’s movements. CDRs are inherently limited in terms of spatial resolution by the density and coverage of the cell towers of the mobile operator. Secondly, CDR datasets can lack temporal resolution when long periods of time pass between call and text events. Therefore, in contrast to the GPS methods used in studies in Chapters 4 and 5 the CDR datasets only allowed for the detection of movements between the general regions surrounding cell towers, and only recorded a location when a call or text was made. Finally, the lack of demographic data from the subscriber population makes it impossible to know how representative the sample population is of the whole population of Mozambique – it is highly likely, for example, that the movements of children are underrepresented in the dataset, which is detrimental to the results because children are at higher risk of severe malaria than the general population due to their lack of immunity. Moreover, it is likely that women are underrepresented in the sample population, given recent statistics on mobile phone ownership by gender (60% of men, 45% of women own a mobile phone in Mozambique, a gender gap of 24% (GSMA, 2019)). Nevertheless, the large sample size (over 5.5 million subscribers out of the total population of 28.9 million people), penetration rate and area covered by the dataset allowed for a unique insight into the large-scale movements of people over a relatively long time period compared to the studies in the preceding chapters and provided a level of detail sufficient to answer the research questions.

7.4 Future research

Future research in this area should aim to build on the findings discussed above, particularly with regard to expanding and strengthening the datasets collected in resource-poor settings. There are few studies of human movement patterns in these areas, particularly rural areas where many people are at highest risk of some vector-borne diseases.

The exploratory nature of the study in Chapter 4 confirms the feasibility of simultaneously tracking human and livestock movements, as this had not previously been tested. Although the sample size was small, the data collected from participants yielded some interesting results over the two seasons studied. To address this limitation in future work, a larger sample size in this area could provide more detailed conclusions and reveal patterns that are generalizable to similar rural populations, particularly those with livestock. Nevertheless, the data obtained by tracking individuals across the two seasons were especially valuable, and thus a longitudinal cohort design where individuals are tracked continuously over the year could be an ideal method to gain data on long-term and long-distance movements, thus facilitating estimation of zoonotic and vector-borne disease risk on wider spatial and temporal scales.

Future studies may be able to overcome the time and cost issues of increasing the data collection period by utilising smartphone technology, which already captures the GPS locations of individuals under certain conditions. The collection of smartphone data in combination with GPS loggers has recently been piloted in an urban area of the UK and demonstrated that location histories from smartphones can be valid datasets for exploring individual movements (Ruktanonchai et al., 2018). In low-income areas of sub-Saharan Africa, this method would have a substantial bias towards wealthy people and is dependent on reliable cell phone network coverage, but smartphone ownership and demand for data plans have both been increasing in recent years (World Bank, 2012), making this a viable option for similar studies in the future.

Finally, it would be beneficial to combine methods used in these studies, for example by exploring the impact of fine-scale human movements on malaria incidence within a much smaller area, which would require equally fine-scale malaria incidence data but may yield insight into the impact of local movements and particularly which demographic characteristics are linked to these movements. When analysed in conjunction with the results of large-scale studies such as those that use CDRs to examine mobility patterns, findings from complementary smaller-scale studies could help researchers to fill in the gaps and identify the mobility patterns that are most important to the spread and persistence of vector-borne diseases at different scales.

7.5 Potential impact and policy implications

The findings presented in Chapters 4 to 6 have some potential implications for policymakers, particularly those working to control vector-borne diseases on fine scales, and separately for those working to control malaria in Mozambique such as the National Malaria Control Program.

As previously mentioned, the analyses in Chapters 4 and 5 describing the movements of a small population in rural Kenya demonstrate the feasibility and value of using GPS loggers to monitor movements relatively soon after the movements occur. Datasets such as the ones I collected in rural Kenya could be used to ground-truth and design further studies of fine-scale movements and would be particularly valuable in conjunction with epidemiological data. Doing so could allow disease control programs to know how local human and domestic livestock populations of interest move and utilise their local infrastructure, which could in turn be used to design targeted interventions aimed at the subgroups most vulnerable to vector-borne diseases. This could have particularly high impact if used for vector-borne diseases that are highly localised, as fine-scale movements are most relevant where disease transmission occurs in small areas.

The analyses in Chapter 6, at the opposing end of the spatiotemporal scale covering a much larger area and population could inherently have a larger impact as they provide empirical evidence of malaria parasite movements in Mozambique. The identification of districts in Mozambique that are big net sources of malaria parasites should be communicated to researchers and policy-makers to ensure that interventions to control and reduce the burden of malaria here target areas that likely export the highest number of parasites, which may not be those with the highest burden. In light of the ongoing elimination efforts in the south of Mozambique, the analyses describing the likely malaria movements into southern provinces in Mozambique are particularly relevant. The high proportion of imported cases compared to locally acquired cases in the southern districts in Maputo province suggest that elimination is feasible here, if importations from nearby districts can be effectively reduced. As elimination is already targeted for Maputo province this year, and pre-elimination status for Gaza and Inhambane provinces by 2025, the next step for elimination programs will be to target these interventions to the communities that are most connected to the southern provinces through human-mediated parasite mobility.

In terms of broader eradication goals for vector-borne diseases, high-quality data on the mobility of both human and animal populations will prove key to preventing and controlling outbreaks of disease worldwide. As the world becomes more connected and populations more mobile, vector-borne disease control programs will look to mobility studies that can describe individual movements on finer and finer scales. Mobile phone technologies that allow the geolocation of large numbers of subscribers will become increasingly valuable as penetration rates in some of the most disadvantaged and isolated areas of the world continue to grow. The results presented in this thesis provide empirical evidence of using mobility patterns to aid control of vector-borne diseases and should encourage future collaborations with telecommunications providers to address global challenges posed by other infectious diseases such as dengue fever.

7.6 Final remarks

The work presented in this thesis broadly addressed the question of how individual-level movements can affect exposure and vulnerability to vector-borne diseases on fine and large scales. Starting with the fine-scale movements of a rural population, I explored how individual movement patterns oat the fine-scale were linked to some demographic characteristics and saw how land use correlated with these movement patterns, which may be affecting exposure to vector-borne disease. Secondly, an examination of the activity spaces of the same rural population identified differences in resource access between individuals, which may be affecting vulnerability to vector-borne disease. Finally, an analysis of the large-scale movements of individuals in a population saw how malaria parasites are moved around the country, with increased movements of parasites from rural to urban areas and the identification of communities of human-mediated parasite movements that are likely to be good targets for elimination measures.

The analyses in Chapters 4 and 5 have been published in peer-reviewed journals and the analysis in Chapter 6 will be submitted for publication this year, demonstrating their contributions to the scientific literature. The findings of these studies highlight the value of understanding human movements at opposite ends of the spatiotemporal scale in the context of vector-borne disease, particularly in low-resource settings. They also emphasize the utility of the methods employed, particularly those involving the use of new technology for tracking individual movements, paving the way for future studies to utilise high-resolution individual movement datasets to tackle the challenges posed by vector-borne diseases in low-resource settings.

Appendix A Chapter 4 Supplementary Information

A.1 Chapter 4 supplementary tables

A.1.1 Multivariable linear regression for movement response variables.

The time spent, maximum distance and home ranges were log transformed before modelling, thus these estimates are factor increases and decreases. Figures in square brackets are 95% confidence intervals. *** p < 0.001, ** p < 0.01, * p < 0.05.

Response variable	**Explanatory variable**	**Estimate**	**p-value**
Time spent on trips outside of the household (humans)	Ruminant ownership: yes [Ref = no]	0.92 [0.52, 1.87]	0.788
	Number of ruminants	1.17 [0.91, 1.46]	0.198
	Season: dry [Ref = short rainy]	1.23 [1.07, 1.41]	0.003 **
	Age	1.01 [1.00, 1.01]	0.171
Time spent on trips outside of the household (ruminants)	Number of ruminants Household wealth	0.95 [0.62, 1.46] 1.00 [0.98, 1.01]	0.804 0.648
Maximum distance travelled outside of the household (humans)	Ruminant ownership: yes [Ref = no]	1.35 [0.56, 2.98]	0.494
	Number of ruminants	1.09 [0.78, 1.53]	0.644
	Occupation: non-farmer [Ref = farmer]	1.37 [0.88, 2.19]	0.201
	Season: dry [Ref = short rainy]	1.38 [1.21, 1.57]	< 0.001 ***
	Household wealth	1.00 [0.99, 1.01]	0.628
Home range (humans)	Season: dry [Ref = short rainy]	3.20 [1.55, 6.83]	0.004 **
	Age	1.03 [1.00, 1.07]	0.056

A.1.2 Univariable linear regression for movement response variables including interaction terms for significant covariates.

The time spent, maximum distance and home ranges were log transformed before modelling, thus these estimates are factor increases and decreases. Figures in square brackets are 95% confidence intervals. *** p < 0.001, ** p < 0.01, * p < 0.05.

Appendix A

Response variable	**Interactions tested**	**Estimate**	**p-value**
Time spent on trips outside of the household (humans)	Ruminant ownership: yes [Ref = no]	1.88 [1.32, 2.69]	0.002 **
	Season: dry [Ref = short rainy]	2.12 [1.53, 2.94]	< 0.001 ***
	Ruminant ownership: yes *Season: dry	0.52 [0.36, 0.74]	< 0.001 ***
	Number of ruminants	1.29 [1.12, 1.51]	0.002 **
	Season: dry [Ref = short rainy]	1.77 [1.37, 2.29]	< 0.001 ***
	Number of ruminants*Season: dry	0.79 [0.69, 0.91]	0.001 **
	Age	1.01 [1.00, 1.01]	0.235
	Season: dry [Ref = short rainy]	0.96 [0.63, 1.44]	0.836
	Age*Season: dry [Ref = short rainy]	1.01 [1.00, 1.01]	0.197
Time spent on trips outside of the household (ruminants)	Number of ruminants	0.95 [0.66, 1.38]	0.789
	Season: dry [Ref = short rainy]	0.78 [0.49, 1.24]	0.292
	Number of ruminants*Season: dry	1.09 [0.72, 1.65]	0.689
	Household wealth	1.00 [0.98, 1.01]	0.756
	Season: dry [Ref = short rainy]	0.82 [0.41, 1.67]	0.578
	Household wealth*Season: dry	1.00 [0.98, 1.02]	0.940
Maximum distance travelled outside of the household (humans)	Ruminant ownership: yes [Ref = no]	1.67 [1.02, 2.73]	0.049 *
	Season: dry [Ref = short rainy]	1.95 [1.42, 2.66]	< 0.001 ***
	Ruminant ownership: yes *Season: dry	0.66 [0.47, 0.93]	0.018 *
	Number of ruminants	1.22 [0.99, 1.50]	0.078
	Season: dry [Ref = short rainy]	1.70 [1.33, 2.17]	< 0.001 ***
	Number of ruminants*Season: dry	0.88 [0.77, 1.00]	0.054
	Occupation: non-farmer [Ref = farmer]	1.31 [0.82, 2.13]	0.271
	Season: dry [Ref = short rainy]	1.45 [1.25, 1.69]	< 0.001 ***
	Occupation: non-farmer *Season: dry	0.83 [0.62, 1.13]	0.243
	Household wealth	0.99 [0.98, 1.01]	0.366
	Season: dry [Ref = short rainy]	1.22 [0.87, 1.72]	0.251
	Household wealth*Season: dry	1.00 [1.00, 1.01]	0.430
Maximum distance travelled outside of	Household wealth	1.01 [1.00, 1.01]	0.183
	Season: dry [Ref = short rainy]	1.30 [0.86, 1.99]	0.216
	Household wealth*Season: dry	0.99 [0.98, 1.00]	0.098

Response variable	Interactions tested	Estimate	p-value
the household (ruminants)			
Home range (humans)	Age	1.03 [0.99, 1.07]	0.133
	Season: dry [Ref = short rainy]	2.46 [0.32, 20.63]	0.406
	Age*Season: dry	1.01 [0.96, 1.05]	0.794
Home range (ruminants)	Number of ruminants	2.29 [1.10, 4.75]	0.040
	Season: dry [Ref = short rainy]	1.47 [0.22, 9.50]	0.697
	Number of ruminants*Season: dry	0.72 [0.28, 1.84]	0.504

A.1.3 Univariable beta regression for time spent by humans and ruminants on different types of land.

Figures in square brackets are 95% confidence intervals. *** p < 0.001, ** p < 0.01, *p < 0.05.

Land type	Explanatory variable	Odds ratio	p-value
Artificial or bare land (humans)	Ruminant ownership: yes [Ref = no]	1.63 [0.36, 7.37]	0.527
	Number of ruminants	0.93 [0.63, 1.38]	0.727
	Gender: male [Ref = female]	1.08 [0.32, 3.68]	0.900
	Occupation: non-farmer [Ref = farmer]	1.21 [0.36, 4.05]	0.758
	Season: dry [Ref = short rainy]	1.12 [0.92, 1.36]	0.267
	Household wealth	0.97 [0.95, 1.00]	0.093
	Age	1.02 [0.99, 1.05]	0.193
Artificial or bare land (ruminants)	Number of ruminants	Invalid model	NA
	Season: dry [Ref = short rainy]	1.27 [1.03, 1.58]	0.028 *
	Household wealth	0.96 [0.92, 1.00]	0.041 *
Crops or grassland (humans)	Ruminant ownership: yes [Ref = no]	0.74 [0.46, 1.17]	0.196
	Number of ruminants	0.98 [0.85, 1.13]	0.773
	Gender: male [Ref = female]	0.87 [0.57, 1.35]	0.545
	Occupation: non-farmer [Ref = farmer]	1.45 [0.94, 2.22]	0.091

Land type	Explanatory variable	Odds ratio	p-value
	Season: dry [Ref = short rainy]	1.01 [0.87, 1.17]	0.929
	Household wealth	1.01 [1.00, 1.02]	0.152
	Age	1.00 [0.98, 1.01]	0.459
Crops or grassland (ruminants)	Number of ruminants	1.00 [0.75, 1.34]	0.991
	Season: dry [Ref = short rainy]	1.16 [0.90, 1.51]	0.256
	Household wealth	1.01 [0.99, 1.02]	0.215
Rice paddies (humans)	Ruminant ownership: yes [Ref = no]	1.07 [0.87, 1.32]	0.493
	Number of ruminants	1.05 [1.00, 1.11]	0.073
	Gender: male [Ref = female]	0.96 [0.81, 1.13]	0.633
	Occupation: non-farmer [Ref = farmer]	0.94 [0.80, 1.12]	0.492
	Season: dry [Ref = short rainy]	1.10 [0.95, 1.26]	0.204
	Household wealth	Invalid model	NA
	Age	1.00 [0.99, 1.00]	0.541
Rice paddies (ruminants)	Number of ruminants	1.08 [1.03, 1.13]	0.001 **
	Season: dry [Ref = short rainy]	1.00 [0.91, 1.09]	0.981
	Household wealth	1.00 [1.00, 1.00]	0.518
Swampland (humans)	Ruminant ownership: yes [Ref = no]	1.19 [0.62, 2.27]	0.609
	Number of ruminants	1.19 [0.89, 1.57]	0.240
	Gender: male [Ref = female]	1.57 [0.86, 2.85]	0.141
	Occupation: non-farmer [Ref = farmer]	0.45 [0.18, 1.10]	0.079
	Season: dry [Ref = short rainy]	0.83 [0.69, 0.99]	0.034 *
	Household wealth	1.01 [0.99, 1.04]	0.385
	Age	0.99 [0.96, 1.01]	0.273
Swampland (ruminants)	Number of ruminants	1.68 [0.83, 3.42]	0.149
	Season: dry [Ref = short rainy]	1.25 [0.96, 1.63]	0.093
	Household wealth	1.01 [0.98, 1.04]	0.688
	Ruminant ownership: yes [Ref = no]	2.79 [1.11, 6.98]	0.029 *

Land type	**Explanatory variable**	**Odds ratio**	**p-value**
Woodland or shrubs (humans)	Number of ruminants	1.38 [1.06, 1.80]	0.017 *
	Gender: male [Ref = female]	1.17 [0.49, 2.82]	0.718
	Occupation: non-farmer [Ref = farmer]	0.94 [0.38, 2.34]	0.901
	Season: dry [Ref = short rainy]	0.96 [0.81, 1.14]	0.642
	Household wealth	1.01 [0.99, 1.04]	0.315
	Age	1.00 [0.98, 1.03]	0.746
Woodland or shrubs (ruminants)	Number of ruminants	1.32 [0.89, 1.98]	0.173
	Season: dry [Ref = short rainy]	1.31 [0.88, 1.95]	0.188
	Household wealth	1.02 [1.00, 1.04]	0.067

A.1.4 Univariable beta regression for time spent by humans on different types of land including interaction between ruminant ownership and season.

Figures in square brackets are 95% confidence intervals. ***p < 0.001, **p < 0.01, *p < 0.05.

Land type	**Explanatory variable**	**Odds ratio**	**p-value**
Artificial or bare land	Ruminant ownership: yes [Ref = no]	2.89 [0.63, 13.33]	0.173
	Season: dry [Ref = short rainy]	3.97 [1.50, 10.53]	0.006 **
	Ruminant ownership: yes *Season: dry	0.30 [0.11, 0.84]	0.023 *
Crops or grassland	Ruminant ownership: yes [Ref = no]	0.88 [0.60, 1.30]	0.529
	Season: dry [Ref = short rainy]	1.08 [0.78, 1.50]	0.630
	Ruminant ownership: yes *Season: dry	0.88 [0.61, 1.27]	0.501
Rice paddies	Ruminant ownership: yes [Ref = no]	1.29 [0.47, 3.55]	0.619
	Season: dry [Ref = short rainy]	1.72 [0.51, 5.81]	0.381
	Ruminant ownership: yes *Season: dry	0.76 [0.19, 2.94]	0.686
Swampland	Ruminant ownership: yes [Ref = no]	2.94 [0.77, 11.30]	0.116
	Season: dry [Ref = short rainy]	2.19 [0.98, 4.87]	0.056
	Ruminant ownership: yes *Season: dry	0.36 [0.15, 0.84]	0.018 *

Land type	**Explanatory variable**	**Odds ratio**	**p-value**
Woodland or shrubs	Ruminant ownership: yes [Ref = no] Season: dry [Ref = short rainy] Ruminant ownership: yes *Season: dry	Invalid model	NA

A.2 Household and individual survey

Individual questions

- Sublocation
- Household ID
- Collect the GPS coordinates of this household.
- Language of administration
- Name of respondent
- Age of respondent
- Gender of respondent
- Tribal origin
- Principal religion
- Marital status
- Level of education reached
- Years lived in current village
- Major occupation

Individual movement questions

Do you regularly visit any of these places?

- School
- Place of work (livestock-related)
- Place of work (other)
- Health facility
- Place of worship
- Market - animal
- Market - other
- Shop
- Another household
- Less regular, e.g. annual visits

For each place:

- Does the place have a name?
- How often do you typically go there?
- How do you usually travel there?
- How much does it cost?
- How long does it take you to get there (minutes)?
- How long do you typically spend there (hours)?
- Do any children go with you?
- How many?

Household questions

- How many members does this household have?
- Gender of household member
- Age of household member
- What is the highest school grade that the female head/spouse has completed?
- What is the main occupation of the male head/spouse?
- How many habitable rooms does this household occupy in its main dwelling?
- What is the floor of the main dwelling predominantly made of?
- What is your source of water for cooking?
- What is your source of water for drinking?
- What is the main source of lighting fuel for the household?
- Is there a latrine in the household?
- What type(s)?
- Is there evidence of scrounging by animals around the latrine?
- Transport: does your household own any of the following?
- Does your household own any irons (charcoal or electric?)
- How many mosquito nets does your household own?
- How many towels does your household own?
- How many frying pans does your household own?
- Do you grow crops?
- Why do you grow crops?
- How do you get to your crops?
- How much does it cost?
- Approximately how long does it take you to get to your crops in minutes?
- Where does the majority of your household access medical facilities?
- How do you normally travel to the medical facility?
- How long does it normally take to get there (minutes)?

Household livestock questions

- Does this homestead keep any animals?
- Where do you access veterinary services?
- Have you used the veterinary services in the last 12 months?
- Does this homestead keep cattle?
- Cattle questions
- How many males?
- How many females?
- Why do you keep cattle?
- Do you ever buy cattle or have you received cattle as a gift from outside the household?
- How long ago did you last buy/receive new cattle?
- Where do you usually buy/receive new cattle from?
- Do your cows/bulls engage in communal breeding?
- Are cattle herded with goats or sheep?
- How do you graze/feed your cattle?
- What is the water source for your cattle?
- Do you use medicine to prevent or treat disease in your cattle?
- Where do you get medicine for your cattle from?
- Have any cattle in the home been given any vaccinations?

- Do you know which vaccines have been given?
- Do you always pasteurize your milk before consumption?

Cattle questions

- Are you involved in feeding the cattle?
- Are you involved in milking the cattle?
- Are you involved in taking cattle to water?
- Are you involved in birthing the cattle?
- Have you ever experienced abortion in your herd?
- When was the last abortion?
- Do you ever handle cattle abortion material?
- What do you do with the aborted material?
- Are you involved in handling cattle manure?
- Do you ever buy cattle from a market?
- Do you ever take the cattle to a market?
- Are you involved in cattle skinning?
- Are you involved in cattle burial?
    - For each of the above:
    - How often do you do this activity?
    - Do you have to travel outside the household to do this activity?
    - What type of place do you go to for this activity?
    - If yes, how do you travel there?
    - How much does it cost?
    - How long does it take you to travel there (minutes)?
    - How long do you spend there (hours)?
    - How many children go with you?
    - What is the name of the place?

Pig questions

- Does this homestead keep pigs?
- How many piglets?
- How many adult males?
- How many adult females?
- Why do you keep pigs?
- Do you buy pigs from a market?
- What is the name of the market?
- How do you normally travel to the market?
- How much does it cost?
- How long does it normally take to get there (minutes)?
- How do you house the pigs?
- How do you feed the pigs?
- Are pigs fed waste?
- If pigs are fed waste, is it cooked prior to feeding it to the pig?
- Are pigs housed during any season?
- What is the flooring in the pig housing?
- Do you use medicine to prevent or treat disease in your pigs?
- Where do you purchase medicine for your pigs?
- Have the pigs on the homestead been vaccinated against anything?
- Does the participant know which vaccine was given to the pigs?

- Name/purpose of vaccine
- Are there any significant problems with your pigs?
- Are you involved in feeding the pigs?
- Are you involved in taking pigs to water?
- Are you involved in birthing the pigs?
- Are you involved in handling pig manure?
- Are you involved in taking pigs to market?
- Are you involved in pig slaughter?
- Are your pigs inspected at slaughter?
- Who inspects the meat for cysts?
- If cysts are found, what do you do with the meat?
- Are you involved in pig burial?
- For each of the activities above:
    - How often do you do this activity?
    - Do you have to travel outside the household to do this activity?
    - What type of place do you go to for this activity?
    - If yes, how do you travel there?
    - How much does it cost?
    - How long does it take you to travel there (minutes)?
    - How long do you spend there (hours)?
    - How many children go with you?
    - What is the name of the place?

Sheep/goat questions

- Does this homestead keep sheep/goats?
- How many sheep/goats?
- Do you buy sheep/goats from a market?
- What is the name of the market?
- How do you normally travel to the market?
- How much does it cost?
- How long does it normally take to get there (minutes)?
- How do you house the sheep/goats?
- How do you graze/feed your sheep/goats?
- Are you involved in feeding the sheep/goats?
- Are you involved in milking the sheep/goats?
- Are you involved in taking sheep/goats to water?
- Are you involved in birthing the sheep/goats?
- Are you involved in handling sheep/goat abortion material?
- Are you involved in handling sheep/goats manure?
- Are you involved in taking sheep/goats to market?
- Are you involved in sheep/goat slaughter?
- Are you involved in sheep/goat skinning?
- Are you involved in sheep/goat burial?
- For each of the activities above:
    - How often do you do this activity?
    - Do you have to travel outside the household to do this activity?
    - What type of place do you go to for this activity?
    - If yes, how do you travel there?
    - How much does it cost?

    - How long does it take you to travel there (minutes)?
    - How long do you spend there (hours)?
    - How many children go with you?
    - What is the name of the place?

Final questions

- Does this homestead keep chickens?
- Do you feed your chickens?
- Do livestock have access to the buildings you sleep in?
- Which livestock have access to the buildings you sleep in?
- In the last 12 months, have you seen wildlife around the home?
- What wildlife have you seen?
- Tracker given?
- Tracker ID

Appendix B Chapter 5 Supplementary Information

B.1 Chapter 5 supplementary table

B.1.1 Relationships between travelling time and demographic characteristics of the surveyed population for different types of activity.

Figures in square brackets are 95% confidence intervals. *** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$.

Activity type		Explanatory variable	Estimate	p-value
Livestock activity	Travelling time	Gender: male [Ref = female]	1.43 [0.75, 2.86]	0.287
		Occupation: non-farmer [Ref = farmer]	1.50 [0.79, 2.81]	0.215
		Household wealth	2.25 [0.74, 6.84]	0.162
		Age	1.01 [0.99, 1.03]	0.555
Health facility	Travelling time	Gender: male [Ref = female]	0.80 [0.60, 1.09]	0.153
		Occupation: non-farmer [Ref = farmer]	0.90 [0.64, 1.29]	0.559
		Household wealth	0.94 [0.55, 1.60]	0.824
		Age	1.00 [0.99, 1.01]	0.655
Place of worship	Travelling time	Gender: male [Ref = female]	0.74 [0.51, 1.10]	0.130
		Occupation: non-farmer [Ref = farmer]	0.80 [0.52, 1.20]	0.276
		Household wealth	0.83 [0.47, 1.48]	0.538
		Age	1.00 [0.99, 1.01]	0.940
Market	Travelling time	Gender: male [Ref = female]	0.94 [0.45, 1.96]	0.877
		Occupation: non-farmer [Ref = farmer]	0.58 [0.28, 1.20]	0.148
		Household wealth	1.01 [0.40, 2.55]	0.989
		Age	0.99 [0.97, 1.02]	0.574

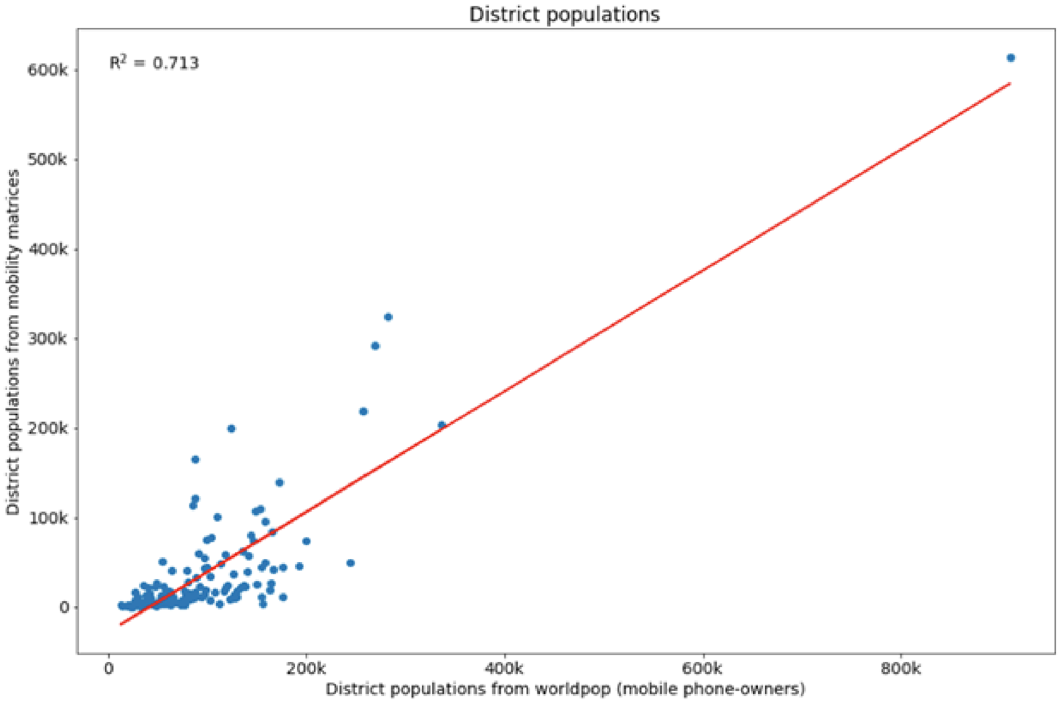
Appendix B

Activity type		**Explanatory variable**	**Estimate**	**p-value**
Household visits	Travelling time	Gender: male [Ref = female]	1.14 [0.83, 1.58]	0.423
		Occupation: non-farmer [Ref = farmer]	0.81 [0.59, 1.11]	0.198
		Household wealth	0.96 [0.64, 1.43]	0.826
		Age	1.01 [1.00, 1.02]	0.074
Water activity	Travelling time	Gender: male [Ref = female]	2.52 [0.92, 6.65]	0.075
		Occupation: non-farmer [Ref = farmer]	0.86 [0.34, 2.13]	0.755
		Household wealth	0.69 [0.22, 2.17]	0.535
		Age	1.00 [0.97, 1.03]	0.976

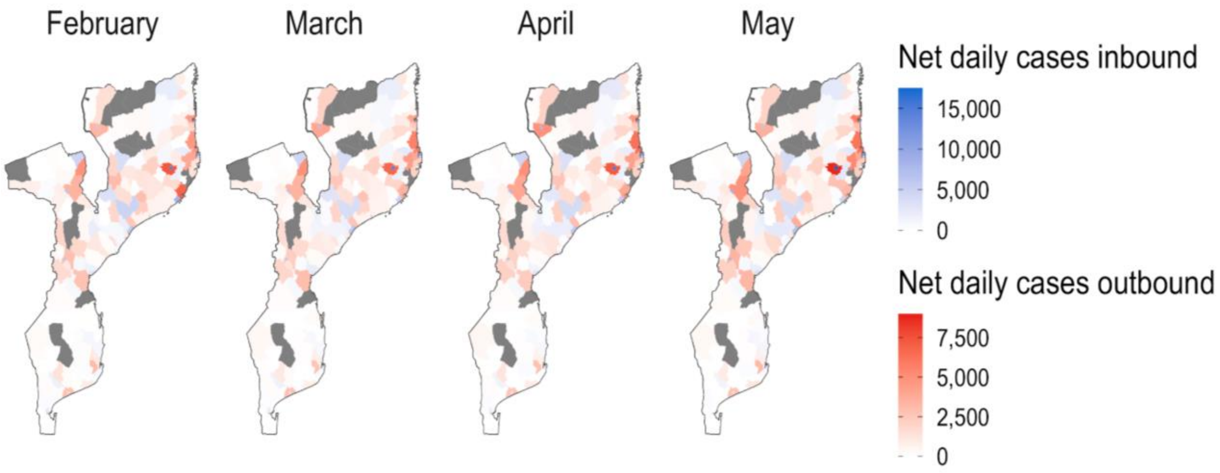
Appendix C Chapter 6 Supplementary Information

C.1 Estimating subscriber bias

I used gridded human population data at 100m resolution from the WorldPop project (WorldPop (www.worldpop.org - School of Geography and Environmental Science, University of Southampton; Department of Geography and Geosciences, University of Louisville; Departement de Geographie, Universite de Namur) and Center for International Earth Sc, 2018) to obtain estimates of the number of people residing in each district of Mozambique. I compared these to the number of geolocated subscribers in each district, which were extracted from the mobility matrices by calculating the mean number of people in each that spent the night in each district over the time span of the dataset (Appendix C.1.1). I divided the WorldPop estimate of population in each district by the number of subscribers in each district to obtain scaling factors, and later used these to scale estimates of people moving out of the districts to estimate actual population movements.



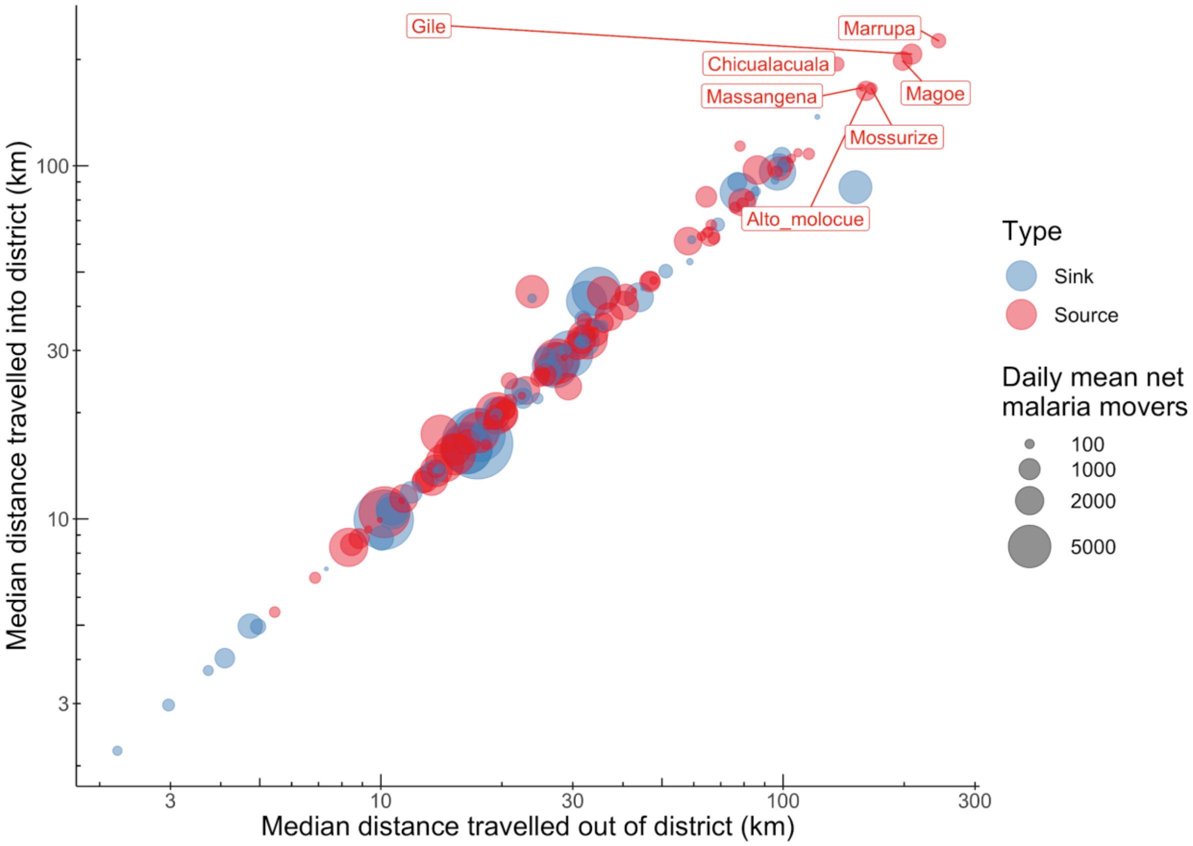
C.1.1 Populations estimated by the WorldPop project vs. populations estimated from the CDR dataset for each district.



C.1.2 Net mean daily movements for human-mediated malaria parasite flows by district for February to May 2018.

Grey areas indicate regions where the amount of data available did not meet privacy thresholds as detailed in the methods.

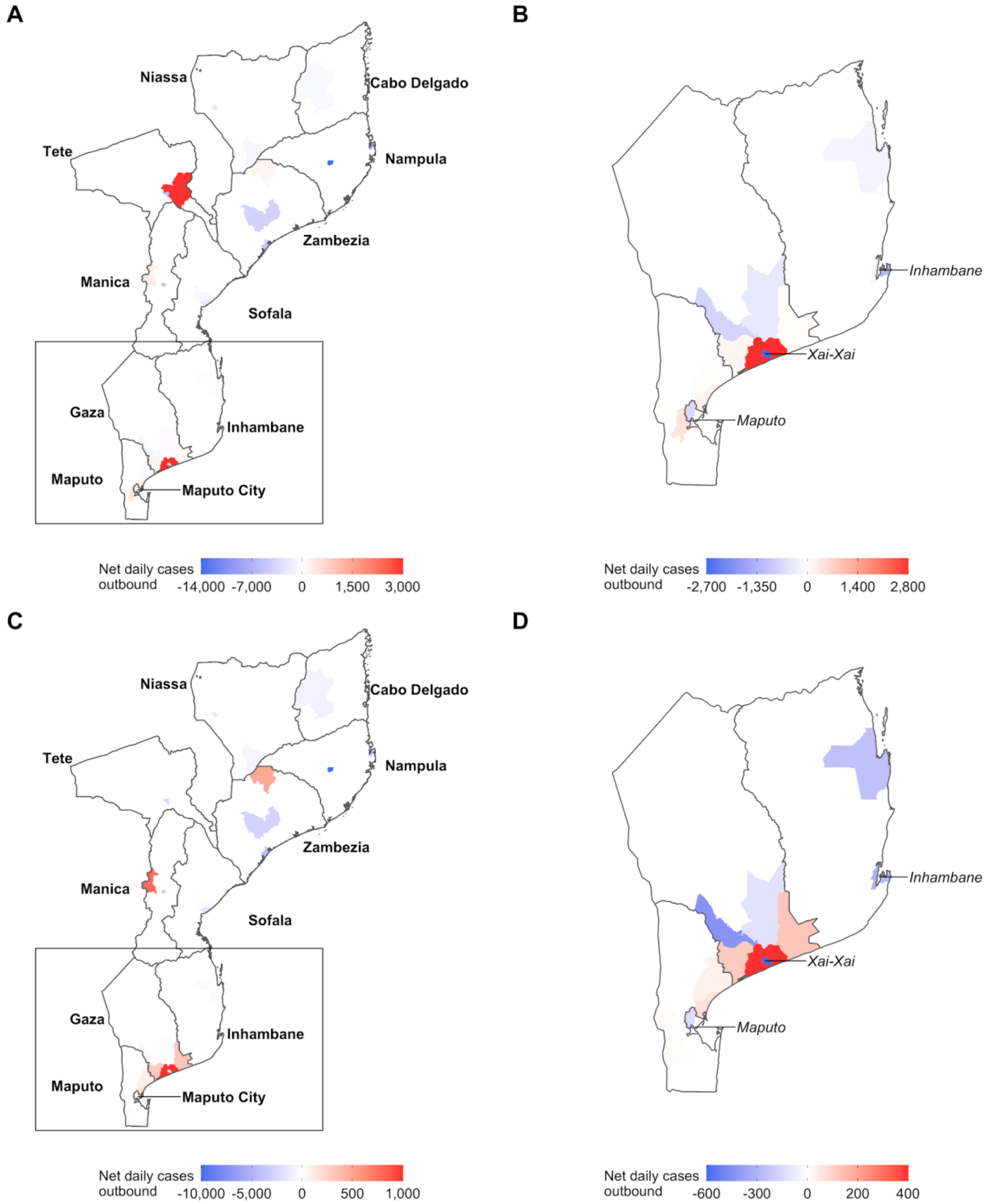
Figure C.1.3 compares the distances travelled into and out of districts and shows how generally the median distance travelled into and out of each district were similar, although these were influenced by both tower distribution and the size and shape of districts. Districts with the shortest median distances were net sinks while districts with the largest distances were net sources. I identified seven districts that could be important sources of malaria, which were net sources with people travelling median distances of more than 130 km out of the district (double the median distance travelled out of all districts). Four of these seven districts lie on the borders with Zimbabwe and South Africa and are likely exporting malaria cases into these countries, although I was unable to measure cross-border movement.



C.1.3 Districts of Mozambique by distance travelled into and out of the district, size of net daily movements and type (net sink vs. net source).

Labelled: 7 districts that were net sources exporting cases over relatively high distances (>130km, the median distance travelled by people leaving all districts).

Appendix C

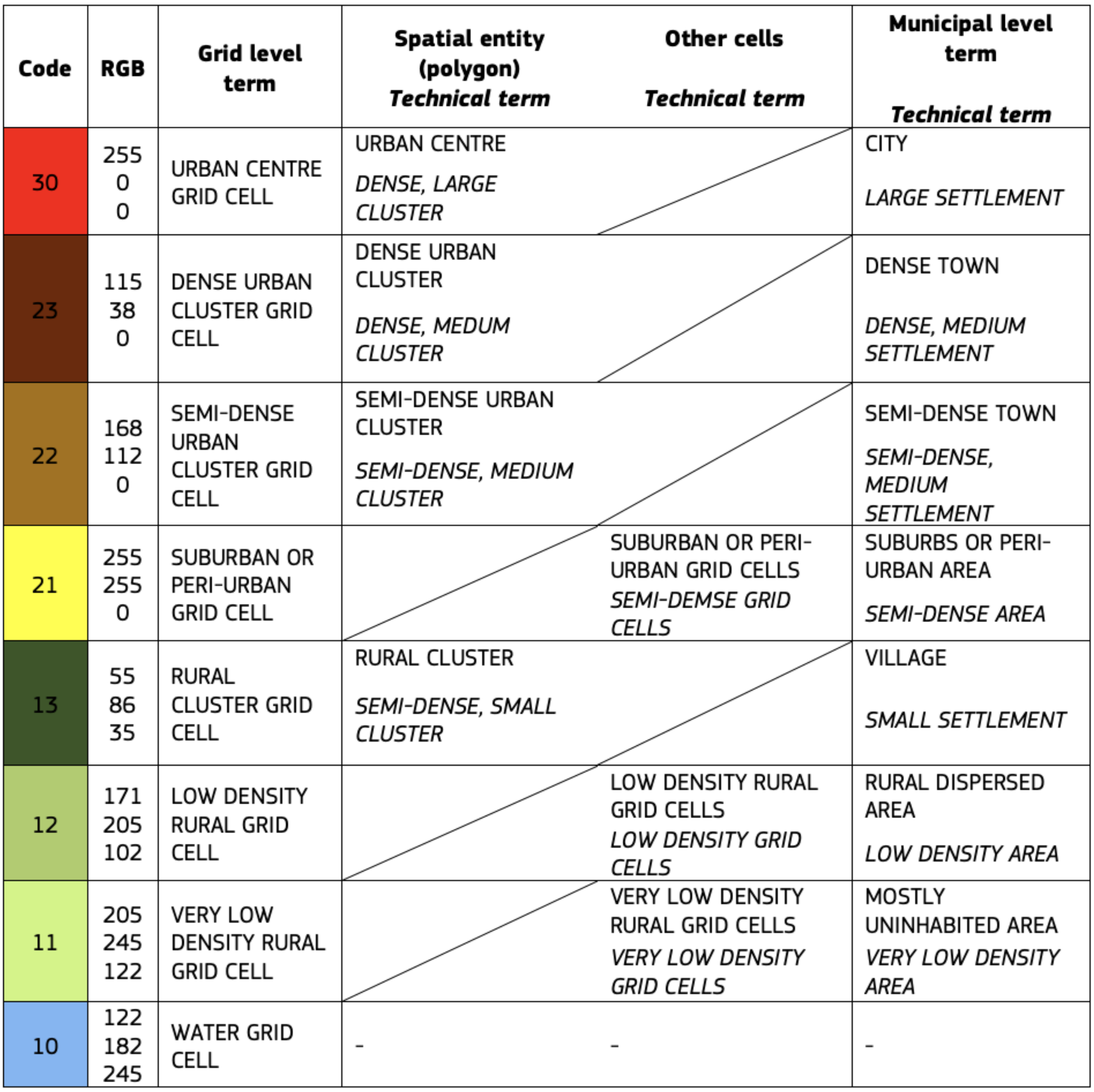


C.1.4 Net mean daily movements in the direction of increasing urbanicity (rural to peri-urban to urban) for human-mediated malaria parasite flows by district for February to May 2018.

A Net daily movements for all districts in Mozambique. **B** Net daily movements for districts in southern four provinces. **C** Net daily movements that were > 25km in distances for all districts in Mozambique. **D** Net daily movements that were > 25km in distance for districts in the southern four provinces in Mozambique. Grey areas indicate regions where the amount of data available did not meet privacy thresholds as detailed in the methods.

C.1.5 Classification of grid cells used by GHS-SMOD

Grid cells used for the urban-rural analysis in Chapter 6 were obtained from GHS-SMOD (Pesaresi et al., 2019). Cells with the codes 30, 23 and 22 were used to classify an area as urban. Cells with the code 21 were used to classify an area as peri-urban. Cells with the codes 13,12 and 11 were used to classify an area as rural. Cells with the code 10 were excluded.



Code	RGB	Grid level term	Spatial entity (polygon) *Technical term*	Other cells *Technical term*	Municipal level term *Technical term*
30	255 0 0	URBAN CENTRE GRID CELL	URBAN CENTRE *DENSE, LARGE CLUSTER*		CITY *LARGE SETTLEMENT*
23	115 38 0	DENSE URBAN CLUSTER GRID CELL	DENSE URBAN CLUSTER *DENSE, MEDUM CLUSTER*		DENSE TOWN *DENSE, MEDIUM SETTLEMENT*
22	168 112 0	SEMI-DENSE URBAN CLUSTER GRID CELL	SEMI-DENSE URBAN CLUSTER *SEMI-DENSE, MEDIUM CLUSTER*		SEMI-DENSE TOWN *SEMI-DENSE, MEDIUM SETTLEMENT*
21	255 255 0	SUBURBAN OR PERI-URBAN GRID CELL		SUBURBAN OR PERI-URBAN GRID CELLS *SEMI-DEMSE GRID CELLS*	SUBURBS OR PERI-URBAN AREA *SEMI-DENSE AREA*
13	55 86 35	RURAL CLUSTER GRID CELL	RURAL CLUSTER *SEMI-DENSE, SMALL CLUSTER*		VILLAGE *SMALL SETTLEMENT*
12	171 205 102	LOW DENSITY RURAL GRID CELL		LOW DENSITY RURAL GRID CELLS *LOW DENSITY GRID CELLS*	RURAL DISPERSED AREA *LOW DENSITY AREA*
11	205 245 122	VERY LOW DENSITY RURAL GRID CELL		VERY LOW DENSITY RURAL GRID CELLS *VERY LOW DENSITY GRID CELLS*	MOSTLY UNINHABITED AREA *VERY LOW DENSITY AREA*
10	122 182 245	WATER GRID CELL	-	-	-

Appendix C

Bibliography

Abellana, R., Ascaso, C., Aponte, J., Saute, F., Nhalungo, D., Nhacolo, A., & Alonso, P. (2008). Spatio-seasonal modeling of the incidence rate of malaria in Mozambique. *Malaria Journal*, *7*(1), 1–11. https://doi.org/10.1186/1475-2875-7-228

Aide, P., Candrinho, B., Galatas, B., Munguambe, K., Guinovart, C., Luis, F., Mayor, A., Paaijmans, K., Fernández-Montoya, L., Cirera, L., Bassat, Q., Mocumbi, S., Menéndez, C., Nhalungo, D., Nhacolo, A., Rabinovich, R., MacEte, E., Alonso, P., & Saúte, F. (2019). Setting the scene and generating evidence for malaria elimination in Southern Mozambique. In *Malaria Journal* (Vol. 18, Issue 1, pp. 1–11). BioMed Central Ltd. https://doi.org/10.1186/s12936-019-2832-9

Airey, T. (1992). The impact of road construction on the spatial characteristics of hospital utilization in the Meru district of Kenya. *Social Science & Medicine*, *34*(10), 1135–1146. https://doi.org/10.1016/0277-9536(92)90287-Z

Alegana, V. A., Maina, J., Ouma, P. O., Macharia, P. M., Wright, J., Atkinson, P. M., Okiro, E. A., Snow, R. W., & Tatem, A. J. (2018). National and sub-national variation in patterns of febrile case management in sub-Saharan Africa. *Nature Communications*, *9*(1), 4994. https://doi.org/10.1038/s41467-018-07536-9

Alegana, V. A., Wright, J. A., Pentrina, U., Noor, A. M., Snow, R. W., & Atkinson, P. M. (2012). Spatial modelling of healthcare utilisation for treatment of fever in Namibia. *International Journal of Health Geographics*, *11*(1), 6. https://doi.org/10.1186/1476-072X-11-6

Altizer, S., Bartel, R., & Han, B. A. (2011). Animal migration and infectious disease risk. *Science (New York, N.Y.)*, *331*(6015), 296–302. https://doi.org/10.1126/science.1194694

Apache Software Foundation. (2018). Apache Spark^{TM} - Unified Analytics Engine for Big Data. In *Apache Spark*.

Arthur, R. F., Gurley, E. S., Salje, H., Bloomfield, L. S. P., & Jones, J. H. (2017). Contact structure, mobility, environmental impact and behaviour: the importance of social forces to infectious disease dynamics and disease ecology. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *372*(1719), 20160454. https://doi.org/10.1098/rstb.2016.0454

Ault, S. K. (1994). Environmental management: A re-emerging vector control strategy. *American Journal of Tropical Medicine and Hygiene*, *50*(6 SUPPL.), 35–49. https://doi.org/10.4269/ajtmh.1994.50.35

Bibliography

Aultman-Hall, L., & Ullman, H. (2019). Long-distance and intercity travel: Who participates in global mobility? In *Mapping the Travel Behavior Genome* (pp. 187–207). Elsevier. https://doi.org/10.1016/B978-0-12-817340-4.00011-5

Bailey, D. W. (1995). Daily selection of feeding areas by cattle in homogeneous and heterogeneous environments. *Applied Animal Behaviour Science*, *45*(3), 183–200. https://doi.org/10.1016/0168-1591(95)00586-H

Bajardi, P., Poletto, C., Ramasco, J. J., Tizzoni, M., Colizza, V., & Vespignani, A. (2011). Human Mobility Networks, Travel Restrictions, and the Global Spread of 2009 H1N1 Pandemic. *PLoS ONE*, *6*(1), e16591. https://doi.org/10.1371/journal.pone.0016591

Bansal, S., Grenfell, B. T., & Meyers, L. A. (2007). When individual behaviour matters: Homogeneous and network models in epidemiology. *Journal of the Royal Society Interface*, *4*(16), 879–891. https://doi.org/10.1098/rsif.2007.1100

Barasona, J. A., Latham, M. C., Acevedo, P., Armenteros, J. A., Latham, A. D. M., Gortazar, C., Carro, F., Soriguer, R. C., & Vicente, J. (2014). Spatiotemporal interactions between wild boar and cattle: implications for cross-species disease transmission. *Veterinary Research*, *45*(1), 122. https://doi.org/10.1186/s13567-014-0122-7

Barbosa, H., Barthelemy, M., Ghoshal, G., James, C. R., Lenormand, M., Louail, T., Menezes, R., Ramasco, J. J., Simini, F., & Tomasini, M. (2018). Human mobility: Models and applications. *Physics Reports*, *734*, 1–74. https://doi.org/10.1016/J.PHYSREP.2018.01.001

Bardosh, K. L., Ryan, S., Ebi, K., Welburn, S., & Singer, B. (2017). Addressing vulnerability, building resilience: Community-based adaptation to vector-borne diseases in the context of global change. In *Infectious Diseases of Poverty* (Vol. 6, Issue 1, pp. 1–21). BioMed Central Ltd. https://doi.org/10.1186/s40249-017-0375-2

Barros, F. S. M., & Honório, N. A. (2015). Deforestation and Malaria on the Amazon Frontier: Larval Clustering of Anopheles darlingi (Diptera: Culicidae) Determines Focal Distribution of Malaria. *The American Journal of Tropical Medicine and Hygiene*, *93*(5), 939–953. https://doi.org/10.4269/ajtmh.15-0042

Bharti, N., Tatem, A. J., Ferrari, M. J., Grais, R. F., Djibo, A., & Grenfell, B. T. (2011). Explaining Seasonal Fluctuations of Measles in Niger Using Nighttime Lights Imagery. *Science*, *334*(6061), 1424 LP – 1427. https://doi.org/10.1126/science.1210554

Bhatt, S., Weiss, D. J., Cameron, E., Bisanzio, D., Mappin, B., Dalrymple, U., Battle, K., Moyes, C. L., Henry, A., Eckhoff, P. A., Wenger, E. A., Briët, O., Penny, M. A., Smith, T. A., Bennett, A., Yukich, J., Eisele, T. P., Griffin, J. T., Fergus, C. A., … Gething, P. W. (2015). The effect of malaria control on Plasmodium falciparum in Africa between 2000 and 2015. *Nature*, *526*(7572), 207–211. https://doi.org/10.1038/nature15535

Blanford, J. I., Huang, Z., Savelyev, A., & MacEachren, A. M. (2015). Geo-Located Tweets. Enhancing Mobility Maps and Capturing Cross-Border Movement. *PloS One*, *10*(6), e0129202. https://doi.org/10.1371/journal.pone.0129202

Boelaert, M., Meheus, F., Sanchez, A., Singh, S. P., Vanlerberghe, V., Picado, A., Meessen, B., & Sundar, S. (2009). The poorest of the poor: a poverty appraisal of households affected by visceral leishmaniasis in Bihar, India. *Tropical Medicine & International Health*, *14*(6), 639–644. https://doi.org/10.1111/j.1365-3156.2009.02279.x

Bohte, W., & Maat, K. (2009). Deriving and validating trip purposes and travel modes for multi-day GPS-based travel surveys: A large-scale application in the Netherlands. *Transportation Research Part C: Emerging Technologies*, *17*(3), 285–297. https://doi.org/10.1016/J.TRC.2008.11.004

Bolker, B. M., Brooks, M. E., Clark, C. J., Geange, S. W., Poulsen, J. R., Stevens, M. H. H., & White, J. S. S. (2009). Generalized linear mixed models: a practical guide for ecology and evolution. In *Trends in Ecology and Evolution* (Vol. 24, Issue 3, pp. 127–135). Elsevier Current Trends. https://doi.org/10.1016/j.tree.2008.10.008

Boulinier, T., Kada, S., Ponchon, A., Dupraz, M., Dietrich, M., Gamble, A., Bourret, V., Duriez, O., Bazire, R., Tornos, J., Tveraa, T., Chambert, T., Garnier, R., & McCoy, K. D. (2016). Migration, Prospecting, Dispersal? What Host Movement Matters for Infectious Agent Circulation? *Integrative and Comparative Biology*, *56*(2), 330–342. https://doi.org/10.1093/icb/icw015

Bousema, T., Griffin, J. T., Sauerwein, R. W., Smith, D. L., Churcher, T. S., Takken, W., Ghani, A., Drakeley, C., & Gosling, R. (2012). Hitting Hotspots: Spatial Targeting of Malaria for Control and Elimination. *PLoS Medicine*, *9*(1), e1001165. https://doi.org/10.1371/journal.pmed.1001165

Brockmann, D., Hufnagel, L., & Geisel, T. (2006). The scaling laws of human travel. *Nature*, *439*(7075), 462–465. https://doi.org/10.1038/nature04292

Brockmann, Dirk, & Theis, F. (2008). Money Circulation, Trackable Items, and the Emergence of Universal Human Mobility Patterns. *IEEE Pervasive Computing*, *7*(4), 28–35. https://doi.org/10.1109/MPRV.2008.77

Bibliography

Bronsvoort, B. M. de C., Thumbi, S. M., Poole, E. J., Kiara, H., Tosas Auguet, O., Handel, I. G., Jennings, A., Conradie, I., Mbole-Kariuki, M. N., Toye, P. G., Hanotte, O., Coetzer, J., & Woolhouse, M. E. (2013). Design and descriptive epidemiology of the Infectious Diseases of East African Livestock (IDEAL) project, a longitudinal calf cohort study in western Kenya. *BMC Veterinary Research*, *9*(1), 171. https://doi.org/10.1186/1746-6148-9-171

Brooks, M. E., Kristensen, K., van Benthem, K. J., Magnusson, A., Berg, C. W., Nielsen, A., Skaug, H. J., Maechler, M., & Bolker, B. M. (2017). glmmTMB Balances Speed and Flexibility Among Packages for Zero-inflated Generalized Linear Mixed Modeling. *The R Journal*, *9*(2), 378–400.

Browning, E., Bolton, M., Owen, E., Shoji, A., Guilford, T., & Freeman, R. (2018). Predicting animal behaviour using deep learning: GPS data alone accurately predict diving in seabirds. *Methods in Ecology and Evolution*, *9*(3), 681–692. https://doi.org/10.1111/2041-210X.12926

Brunette, W., Sundt, M., Dell, N., Chaudhri, R., Breit, N., & Borriello, G. (2013). Open data kit 2.0. *Proceedings of the 14th Workshop on Mobile Computing Systems and Applications - HotMobile '13*, 1. https://doi.org/10.1145/2444776.2444790

Buckee, C. O., Wesolowski, A., Eagle, N. N., Hansen, E., & Snow, R. W. (2013). Mobile phones and malaria: Modeling human and parasite travel. *Travel Medicine and Infectious Disease*, *11*(1), 15–22. https://doi.org/10.1016/J.TMAID.2012.12.003

Budd, L., Morag Bell, & Brown, T. (2009). Of plagues, planes and politics: Controlling the global spread of infectious diseases by air. *Political Geography*, *28*(7), 426–435. https://doi.org/10.1016/j.polgeo.2009.10.006

Campbell, S. J., Savage, G. B., Gray, D. J., Atkinson, J.-A. M., Soares Magalhães, R. J., Nery, S. V., McCarthy, J. S., Velleman, Y., Wicken, J. H., Traub, R. J., Williams, G. M., Andrews, R. M., & Clements, A. C. A. (2014). Water, Sanitation, and Hygiene (WASH): A Critical Component for Sustainable Soil-Transmitted Helminth and Schistosomiasis Control. *PLoS Neglected Tropical Diseases*, *8*(4), e2651. https://doi.org/10.1371/journal.pntd.0002651

Caron, A., Cornelis, D., Foggin, C., Hofmeyr, M., & de Garine-Wichatitsky, M. (2016). African Buffalo Movement and Zoonotic Disease Risk across Transfrontier Conservation Areas, Southern Africa. *Emerging Infectious Diseases*, *22*(2), 277–280. https://doi.org/10.3201/eid2202.140864

Carter, R., Mendis, K. N., & Roberts, D. (2000). Spatial targeting of interventions against malaria. *Bulletin of the World Health Organization*, *78*, 1401–1411. https://doi.org/10.1590/S0042-96862000001200007

Chamberlin, J., & Jayne, T. S. (2013). Unpacking the Meaning of 'Market Access': Evidence from Rural Kenya. *World Development*, *41*, 245–264. https://doi.org/10.1016/J.WORLDDEV.2012.06.004

Chang, H. H., Wesolowski, A., Sinha, I., Jacob, C. G., Mahmud, A., Uddin, D., Zaman, S. I., Hossain, M. A., Faiz, M. A., Ghose, A., Sayeed, A. A., Rahman, M. R., Islam, A., Karim, M. J., Rezwan, M. K., Shamsuzzaman, A. K. M., Jhora, S. T., Aktaruzzaman, M. M., Drury, E., … Buckee, C. (2019). Mapping imported malaria in Bangladesh using parasite genetic and human mobility data. *ELife*, *8*. https://doi.org/10.7554/eLife.43481

Chapman, M., & Prothero, R. M. (1983). Themes on circulation in the Third World. *Circulation in Third World Countries*, 1–26. https://doi.org/10.1177/019791838301700402

Charlwood, J. D., Mendis, C., Thompson, R., Begtrup, K., Cuamba, N., Dgedge, M., Gamage-Mendis, A., Hunt, R. H., Sinden, R. E., & Høgh, B. (1998). Cordon sanitaire or laissez faire: Differential dispersal of young and old females of the malaria vector Anopheles funestus Giles (Diptera: Culicidae) in southern Mozambique. *African Entomology*, *6*(1), 1–6. https://journals.co.za/content/ento/6/1/AJA10213589_251

Chen, J., Zou, L., Jin, Z., Ruan, S., Lozano, R., Naghavi, M., Foreman, K., Wang, X., Huang, J., Yu, Y., Zhang, Y. Z., Xiong, C. L., Xiao, D. L., Song, M., Tang, Q., Wang, D. M., Zhang, J., Jin, Z., Sun, G., … Penny, M. A. (2015). Modeling the Geographic Spread of Rabies in China. *PLOS Neglected Tropical Diseases*, *9*(5), e0003772. https://doi.org/10.1371/journal.pntd.0003772

Chen, Z., & Schintler, L. A. (2015). Sensitivity of location-sharing services data: evidence from American travel pattern. *Transportation*, *42*(4), 669–682. https://doi.org/10.1007/s11116-015-9596-z

Chevalier, V., de la Rocque, S., Baldet, T., Vial, L., & Roger, F. (2004). Epidemiological processes involved in the emergence of vector-borne diseases: West Nile fever, Rift Valley fever, Japanese encephalitis and Crimean-Congo haemorrhagic fever. *REVUE SCIENTIFIQUE ET TECHNIQUE DE L OFFICE INTERNATIONAL DES EPIZOOTIES*, *23*(2), 535–555.

Clark, G. L., Gertler, M. S., & Feldman, M. P. (2003). *The Oxford handbook of economic geography*. Oxford University Press. https://books.google.co.uk/books?id=TzZ_oByXYhkC&pg=PA173&lpg=PA173&dq=Climate,+coastal+proximity,+and+development&source=bl&ots=x4v8qWXIjm&sig=ACfU3U3Fg7GZGRBlSnP8ovSrCKVE3Ju1IQ&hl=en&sa=X&ved=2ahUKEwiiyKKH-7PiAhUBXRUIHdj1DBMQ6AEwBHoECAkQAQ#v=onepage&q=Climate%2C coastal proximity%2C and development&f=false

Cliquet, F., Picard-Meyer, E., & Robardet, E. (2014). Rabies in Europe: what are the risks? *Expert Review of Anti-Infective Therapy*, *12*(8), 905–908. https://doi.org/10.1586/14787210.2014.921570

Colborn, J., & Malaria Atlas Project. (2020). *Personal communication pending publication: 5km monthly malaria incidence rasters for Mozambique.*

Collia, D. V., Sharp, J., & Giesbrecht, L. (2003). The 2001 National Household Travel Survey: A look into the travel patterns of older Americans. *Journal of Safety Research*, *34*(4), 461–470. https://doi.org/10.1016/J.JSR.2003.10.001

Collinge, S. K., & Ray, C. (2006). *Disease ecology : community structure and pathogen dynamics*. Oxford University Press. https://books.google.co.uk/books?hl=en&lr=&id=P6bFjZhD-JIC&oi=fnd&pg=PA6&dq=host+contact+is+intrinsically+driven+by+host+movement+&ots=sFOXmKNCba&sig=gsiMpiY1ATdzwkOousiyU3BItJ4#v=onepage&q&f=false

Cook, J., Kimuyu, P., & Whittington, D. (2016). The costs of coping with poor water supply in rural Kenya. *Water Resources Research*, *52*(2), 619–1539. https://doi.org/10.1002/2015WR017468

Cowie, C. E., Marreos, N., Gortázar, C., Jaroso, R., White, P. C. L., & Balseiro, A. (2014). Shared risk factors for multiple livestock diseases: A case study of bovine tuberculosis and brucellosis. *Research in Veterinary Science*, *97*(3), 491–497. https://doi.org/10.1016/j.rvsc.2014.09.002

Craig, M., Le Sueur, D., & Snow, B. (1999). A climate-based distribution model of malaria transmission in sub-Saharan Africa. *Parasitology Today*, *15*(3), 105–111. https://doi.org/10.1016/S0169-4758(99)01396-4

Cross, P. C., Lloyd-Smith, J. O., Johnson, P. L. F., & Getz, W. M. (2005). Duelling timescales of host movement and disease recovery determine invasion of disease in structured populations. In *Ecology Letters* (Vol. 8, Issue 6, pp. 587–595). https://doi.org/10.1111/j.1461-0248.2005.00760.x

Csardi G, N. T. (2006). The igraph software package for complex network research. *InterJournal*, *Complex Sy*, 1695. https://www.researchgate.net/publication/221995787_The_Igraph_Software_Package_for_Complex_Network_Research

Dar, O., McIntyre, S., Hogarth, S., & Heymann, D. (2013). Rift Valley fever and a new paradigm of research and development for zoonotic disease control. *Emerging Infectious Diseases*, *19*(2), 189–193. https://doi.org/10.3201/eid1902.120941

de La Rocque, S., & Formenty, P. (2014). Applying the One Health principles: a trans-sectoral coordination framework for preventing and responding to Rift Valley fever outbreaks. *Revue Scientifique et Technique (International Office of Epizootics)*, *33*(2), 555–567. http://www.ncbi.nlm.nih.gov/pubmed/25707183

Deville, P., Linard, C., Martin, S., Gilbert, M., Stevens, F. R., Gaughan, A. E., Blondel, V. D., & Tatem, A. J. (2014). Dynamic population mapping using mobile phone data. *Proceedings of the National Academy of Sciences of the United States of America*, *111*(45), 15888–15893. https://doi.org/10.1073/pnas.1408439111

Douglas Bates, Martin Maechler, Ben Bolker, S. W. (2015). Fitting Linear Mixed-Effects Models Using lme4. *Journal of Statistical Software*, *67*(1), 1–48. https://doi.org/10.18637/jss.v067.i01

Ferguson, H. M., Dornhaus, A., Beeche, A., Borgemeister, C., Gottlieb, M., Mulla, M. S., Gimnig, J. E., Fish, D., & Killeen, G. F. (2010). Ecology: A Prerequisite for Malaria Elimination and Eradication. *PLoS Medicine*, *7*(8), e1000303. https://doi.org/10.1371/journal.pmed.1000303

Fèvre, E. M., Picozzi, K., Fyfe, J., Waiswa, C., Odiit, M., Coleman, P. G., & Welburn, S. C. (2005). A burgeoning epidemic of sleeping sickness in Uganda. *Lancet*, *366*(9487), 745–747. https://doi.org/10.1016/S0140-6736(05)67179-6

Fèvre, EM, Coleman, P., Odiit, M., Magona, J., Welburn, S., & Woolhouse, M. (2001). The origins of a new Trypanosoma brucei rhodesiense sleeping sickness outbreak in eastern Uganda. *The Lancet*, *358*(9282), 625–628. https://doi.org/10.1016/S0140-6736(01)05778-6

Fèvre, Eric, Bronsvoort, B. M. de C., Hamilton, K. A., & Cleaveland, S. (2006). Animal movements and the spread of infectious diseases. *Trends in Microbiology*, *14*(3), 125–131. https://doi.org/10.1016/j.tim.2006.01.004

Fèvre, Eric M., de Glanville, W. A., Thomas, L. F., Cook, E. A. J., Kariuki, S., & Wamae, C. N. (2017). An integrated study of human and animal infectious disease in the Lake Victoria crescent small-holder crop-livestock production system, Kenya. *BMC Infectious Diseases*, *17*(1), 457. https://doi.org/10.1186/s12879-017-2559-6

Figueroa, M. J., Nielsen, T. A. S., & Siren, A. (2014). Comparing urban form correlations of the travel patterns of older and younger adults. *Transport Policy*, *35*, 10–20.

https://doi.org/10.1016/J.TRANPOL.2014.05.007

Floyd, J. R., Ogola, J., Fèvre, E. M., Wardrop, N., Tatem, A. J., & Ruktanonchai, N. W. (2020). Activity-specific mobility of adults in a rural region of western Kenya. *PeerJ*, *8*, e8798. https://doi.org/10.7717/peerj.8798

Floyd, J. R., Ruktanonchai, N. W., Wardrop, N., Tatem, A. J., Ogola, J., & Fèvre, E. M. (2019). Exploring fine-scale human and livestock movement in western Kenya. *One Health*, *7*. https://doi.org/10.1016/j.onehlt.2019.100081

Fornace, K. M., Alexander, N., Abidin, T. R., Brock, P. M., Chua, T. H., Vythilingam, I., Ferguson, H. M., Manin, B. O., Wong, M. L., Ng, S. H., Cox, J., & Drakeley, C. (2019). Local human movement patterns and land use impact exposure to zoonotic malaria in Malaysian borneo. *ELife*, *8*. https://doi.org/10.7554/eLife.47602

Funston, P. J., Skinner, J. D., & Dott, H. M. (1994). Seasonal variation in movement patterns, home range and habitat selection of buffaloes in a semi-arid habitat. *African Journal of Ecology*, *32*(2), 100–114. https://doi.org/10.1111/j.1365-2028.1994.tb00562.x

Gabrielli, L., Deutschmann, E., Natale, F., Recchi, E., & Vespe, M. (2019). Dissecting global air traffic data to discern different types and trends of transnational human mobility. *EPJ Data Science*, *8*(1), 26. https://doi.org/10.1140/epjds/s13688-019-0204-x

Gao, D., Amza, A., Nassirou, B., Kadri, B., Sippl-Swezey, N., Liu, F., Ackley, S. F., Lietman, T. M., & Porco, T. C. (2014). Optimal seasonal timing of oral azithromycin for malaria. *American Journal of Tropical Medicine and Hygiene*, *91*(5), 936–942. https://doi.org/10.4269/ajtmh.13-0474

Giardina, F., Franke, J., & Vounatsou, P. (2015). Geostatistical modelling of the malaria risk in Mozambique: Effect of the spatial resolution when using remotely-sensed imagery. *Geospatial Health*, *10*(2), 232–238. https://doi.org/10.4081/gh.2015.333

Global System for Mobile Communications (GSMA). (2019). GSMA The Mobile Economy Sub-Saharan Africa 2019 - The Mobile Economy. In *GSMA*. https://www.gsma.com/mobileeconomy/sub-saharan-africa/

Gray, C., & Mueller, V. (2012). Drought and Population Mobility in Rural Ethiopia. *World Development*, *40*(1), 134–145. https://doi.org/10.1016/J.WORLDDEV.2011.05.023

GSMA. (2019). *The Mobile Gender Gap Report 2019*. 52. https://www.gsma.com/mobilefordevelopment/wp-content/uploads/2019/02/GSMA-The-

Mobile-Gender-Gap-Report-2019.pdf

Gubler, D. (2012). Dengue, urbanization and globalization: The unholy trinity of the 21st century. *International Journal of Infectious Diseases*, *16*, e2. https://doi.org/10.1016/j.ijid.2012.05.009

Hampshire, K. (2015). Fulani on the move: Seasonal economic migration in the Sahel as a social process. *Labour Mobility and Rural Society*, 15–36.

Hamrick, P. N., Aldighieri, S., Machado, G., Leonel, D. G., Vilca, L. M., Uriona, S., & Schneider, M. C. (2017). Geographic patterns and environmental factors associated with human yellow fever presence in the Americas. *PLOS Neglected Tropical Diseases*, *11*(9), e0005897. https://doi.org/10.1371/journal.pntd.0005897

Hanson, S., & Hanson, P. (1980). Gender and Urban Activity Patterns in Uppsala, Sweden. *Geographical Review*, *70*(3), 291. https://doi.org/10.2307/214257

Hast, M., Searle, K. M., Chaponda, M., Lupiya, J., Lubinda, J., Sikalima, J., Kobayashi, T., Shields, T., Mulenga, M., Lessler, J., Moss, W. J., & Research, for the S. and C. A. I. C. of E. for M. (2019). The use of GPS data loggers to describe the impact of spatio-temporal movement patterns on malaria control in a high-transmission area of northern Zambia. *International Journal of Health Geographics*, *18*(1), 19. https://doi.org/10.1186/s12942-019-0183-y

Hawelka, B., Sitko, I., Beinat, E., Sobolevsky, S., Kazakopoulos, P., & Ratti, C. (2014). Geo-located Twitter as proxy for global mobility patterns. *Cartography and Geographic Information Science*, *41*(3), 260–271. https://doi.org/10.1080/15230406.2014.890072

Hay, S. I., Rogers, D. J., Toomer, J. F., & Snow, R. W. (2000). Annual Plasmodium falciparum entomological inoculation rates (EIR) across Africa: literature survey, internet access and review. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, *94*(2), 113–127. https://doi.org/10.1016/S0035-9203(00)90246-3

Hegglin, D., & Deplazes, P. (2013). Control of Echinococcus multilocularis: Strategies, feasibility and cost–benefit analyses. *International Journal for Parasitology*, *43*(5), 327–337. https://doi.org/10.1016/j.ijpara.2012.11.013

Hickman, R., Hall, P., & Banister, D. (2013). Planning more for sustainable mobility. *Journal of Transport Geography*, *33*, 210–219. https://doi.org/10.1016/J.JTRANGEO.2013.07.004

Hijmans, R. J., Etten, J. van, Mattiuzzi, M., Sumner, M., Greenberg, J. A., Lamigueiro, O. P., Bevan, A., Racine, E. B., & Shortridge, A. (2014). Package “raster.” *R*.

Bibliography

Imai, N., White, M. T., Ghani, A. C., & Drakeley, C. J. (2014). Transmission and Control of Plasmodium knowlesi: A Mathematical Modelling Study. *PLoS Neglected Tropical Diseases*, *8*(7). https://doi.org/10.1371/journal.pntd.0002978

Isaacman, S., Becker, R., Cáceres, R., Martonosi, M., Rowland, J., Varshavsky, A., & Willinger, W. (2012). Human mobility modeling at metropolitan scales. *Proceedings of the 10th International Conference on Mobile Systems, Applications, and Services - MobiSys '12*, 239. https://doi.org/10.1145/2307636.2307659

Ivers, L. C., & Ryan, E. T. (2006). Infectious diseases of severe weather-related and flood-related natural disasters. *Current Opinion in Infectious Diseases*, *19*(5), 408–414. https://doi.org/10.1097/01.qco.0000244044.85393.9e

Johnson, A., Goss, A., Beckerman, J., & Castro, A. (2012). Hidden costs: The direct and indirect impact of user fees on access to malaria treatment and primary care in Mali. *Social Science & Medicine*, *75*(10), 1786–1792. https://doi.org/10.1016/J.SOCSCIMED.2012.07.015

Jones, K. H., Daniels, H., Heys, S., & Ford, D. V. (2018). Challenges and Potential Opportunities of Mobile Phone Call Detail Records in Health Research: Review. *JMIR MHealth and UHealth*, *6*(7), e161. https://doi.org/10.2196/mhealth.9974

Källander, K., Hildenwall, H., Waiswa, P., Galiwango, E., Peterson, S., & Pariyo, G. (2008). Delayed care seeking for fatal pneumonia in children aged under five years in Uganda: a case-series study. *Bulletin of the World Health Organization*, *86*(5), 332–338. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=2647445&tool=pmcentrez&rendertype=abstract

Kanuganti, S., Sarkar, A. K., Singh, A. P., & Arkatkar, S. S. (2015). Quantification of accessibility to health facilities in rural areas. *Case Studies on Transport Policy*, *3*(3), 311–320. https://doi.org/10.1016/J.CSTP.2014.08.004

Karungu, S., Atoni, E., Ogalo, J., Mwaliko, C., Agwanda, B., Yuan, Z., & Hu, X. (2019). Mosquitoes of Etiological Concern in Kenya and Possible Control Strategies. *Insects*, *10*(6). https://doi.org/10.3390/insects10060173

Kaur, M., Sodhi, S. K., Kaur, P., Singh, J., & Kumar, R. (2013). Gender differences in health care seeking behaviour of tuberculosis patients in Chandigarh. *Indian Journal of Tuberculosis*, *60*(4), 217–222. http://medind.nic.in/ibr/t13/i4/ibrt13i4p217.pdf

Kenya | PPI. (2011). Poverty Probability Index. https://www.povertyindex.org/country/kenya

Kenya Ministry of Medical Services and Ministry of Public Health & Sanitation. (2013). *HEALTH SECTOR STRATEGIC AND INVESTMENT PLAN (KHSSP)*. http://www.health.go.ke

Kenya National Bureau of Statistics. (2017). *Busia County Statistical Abstract*. https://www.knbs.or.ke/download/busia/

Kenya National Bureau of Statistics. (2018). *Kenya Integrated Household Budget Survey (2015/2016): Basic Report On Well Being In Kenya*. Kenya National Bureau of Statistics. https://www.knbs.or.ke/

Kenya National Bureau of Statistics. (2019). 2019 Kenya Population and Housing Census Volume 1: Population by County and Sub-County. In *2019 Kenya Population and Housing Census*.

Kenya National Bureau of Statistics, Health/Kenya, M. of, Council/Kenya, N. A. C., Institute, K. M. R., & Development/Kenya, N. C. for P. and. (2015). *Kenya Demographic and Health Survey 2014*. http://dhsprogram.com/pubs/pdf/FR308/FR308.pdf

Kibusi, S. M., Kimunai, E., & Hines, C. S. (2015). Predictors for uptake of intermittent preventive treatment of malaria in pregnancy (IPTp) in Tanzania. *BMC Public Health*, *15*(1), 540. https://doi.org/10.1186/s12889-015-1905-0

Killeen, G. F., Knols, B. G., & Gu, W. (2003). Taking malaria transmission out of the bottle: implications of mosquito dispersal for vector-control interventions. *The Lancet Infectious Diseases*, *3*(5), 297–303. https://doi.org/10.1016/S1473-3099(03)00611-X

Kiss, I. Z., Green, D. M., & Kao, R. R. (2006). The network of sheep movements within Great Britain: network properties and their implications for infectious disease spread. *Journal of The Royal Society Interface*, *3*(10), 669–677. https://doi.org/10.1098/rsif.2006.0129

Kolb, H. H. (1984). Factors Affecting the Movements of Dog Foxes in Edinburgh. *The Journal of Applied Ecology*, *21*(1), 161. https://doi.org/10.2307/2403044

Kouznetsov, R. L. (1977). Malaria Control by Application of Indoor Spraying of Residual Insecticides in Tropical Africa and Its Impact on Community Health. *Tropical Doctor*, *7*(2), 81–91. https://doi.org/10.1177/004947557700700216

Kraemer, M. U. G., Yang, C.-H., Gutierrez, B., Wu, C.-H., Klein, B., Pigott, D. M., Covid-19 Data, O., Group, W., Du Plessis, L., Faria, N. R., Li, R., Hanage, W. P., Brownstein, J. S., Layan, M., Vespignani, A., Tian, H., Dye, C., Pybus, O. G., & Scarpino, S. V. (n.d.). *The effect of human mobility and control measures on the COVID-19 epidemic in China*. http://science.sciencemag.org/

Bibliography

Kristjanson, P., Krishna, A., Radeny, M., & Nindo, W. (2004). *Pathways Out of Poverty in Western Kenya and the Role of Livestock (Pro-Poor Livestock Policy Initiative Working Paper No. 14)* (Issue 14). http://www.fao.org/ag/pplpi.html

Lai, S., Ruktanonchai, N. W., Zhou, L., Prosper, O., Luo, W., Floyd, J. R., Wesolowski, A., Santillana, M., Zhang, C., Du, X., Yu, H., & Tatem, A. J. (2020). Effect of non-pharmaceutical interventions to contain COVID-19 in China. *Nature*, 1–7. https://doi.org/10.1038/s41586-020-2293-x

Lam, W. K., Zhong, N. S., & Tan, W. C. (2003). Overview on SARS in Asia and the world. *Respirology (Carlton, Vic.)*, *8 Suppl*, S2-5. http://www.ncbi.nlm.nih.gov/pubmed/15018125

Lambrechts, L., Knox, T. B., Wong, J., Liebman, K. A., Albright, R. G., & Stoddard, S. T. (2009). Shifting priorities in vector biology to improve control of vector-borne disease. *Tropical Medicine & International Health*, *14*(12), 1505–1514. https://doi.org/10.1111/j.1365-3156.2009.02401.x

Larson, G. (2012). Using pigs as a proxy to reconstruct patterns of human migration. In *Population Dynamics in Prehistory and Early History: New Approaches Using Stable Isotopes and Genetics* (pp. 31–40). DE GRUYTER. https://doi.org/10.1515/9783110266306.31

Le Menach, A., Tatem, A. J., Cohen, J. M., Hay, S. I., Randell, H., Patil, A. P., & Smith, D. L. (2011). Travel risk, malaria importation and malaria transmission in Zanzibar. *Scientific Reports*, *1*(1), 93. https://doi.org/10.1038/srep00093

Leroy, E. M., Epelboin, A., Mondonge, V., Pourrut, X., Gonzalez, J.-P., Muyembe-Tamfum, J.-J., & Formenty, P. (2009). Human Ebola outbreak resulting from direct exposure to fruit bats in Luebo, Democratic Republic of Congo, 2007. *Vector Borne and Zoonotic Diseases (Larchmont, N.Y.)*, *9*(6), 723–728. https://doi.org/10.1089/vbz.2008.0167

Leroy, E. M., Kumulungui, B., Pourrut, X., Rouquet, P., Hassanin, A., Yaba, P., Délicat, A., Paweska, J. T., Gonzalez, J.-P., & Swanepoel, R. (2005). Fruit bats as reservoirs of Ebola virus. *Nature*, *438*(7068), 575–576. https://doi.org/10.1038/438575a

Li, W., Shi, Z., Yu, M., Ren, W., Smith, C., Epstein, J. H., Wang, H., Crameri, G., Hu, Z., Zhang, H., Zhang, J., McEachern, J., Field, H., Daszak, P., Eaton, B. T., Zhang, S., & Wang, L.-F. (2005). Bats are natural reservoirs of SARS-like coronaviruses. *Science (New York, N.Y.)*, *310*(5748), 676–679. https://doi.org/10.1126/science.1118391

Li, Y., Liu, L., Zhang, Y., Duan, Z., Tian, G., Zeng, X., Shi, J., Zhang, L., & Chen, H. (2011). New avian influenza virus (H5N1) in wild birds, Qinghai, China. *Emerging Infectious Diseases*, *17*(2), 265–267. https://doi.org/10.3201/eid1702.100732

Linard, C., Gilbert, M., Snow, R. W., Noor, A. M., & Tatem, A. J. (2012). Population Distribution, Settlement Patterns and Accessibility across Africa in 2010. *PLoS ONE*, *7*(2), e31743. https://doi.org/10.1371/journal.pone.0031743

Liu, H., Chen, Y. H., & Lih, J. S. (2015). Crossover from exponential to power-law scaling for human mobility pattern in urban, suburban and rural areas. *European Physical Journal B*, *88*(5), 117. https://doi.org/10.1140/epjb/e2015-60232-1

Liu, J. X., Bousema, T., Zelman, B., Gesase, S., Hashim, R., Maxwell, C., Chandramohan, D., & Gosling, R. (2014). Is Housing Quality Associated with Malaria Incidence among Young Children and Mosquito Vector Numbers? Evidence from Korogwe, Tanzania. *PLoS ONE*, *9*(2), e87358. https://doi.org/10.1371/journal.pone.0087358

Locking, M. E., O'Brien, S. J., Reilly, W. J., Wright, E. M., Campbell, D. M., Coia, J. E., Browning, L. M., & Ramsay, C. N. (2001). Risk factors for sporadic cases of Escherichia coli O157 infection: the importance of contact with animal excreta. *Epidemiol Infect*, *127*(2), 215–220. https://doi.org/10.1017/S0950268801006045

Longini, I. M., Nizam, A., Xu, S., Ungchusak, K., Hanshaoworakul, W., Cummings, D. A. T., & Halloran, M. E. (2005). Containing Pandemic Influenza at the Source. *Science*, *309*(5737). http://science.sciencemag.org/content/309/5737/1083

Lu, X., Bengtsson, L., & Holme, P. (2012). Predictability of population displacement after the 2010 Haiti earthquake. *Proceedings of the National Academy of Sciences of the United States of America*, *109*(29), 11576–11581. https://doi.org/10.1073/pnas.1203882109

M, P., M, B., & JC, T. (1989). *Womens agricultural work child care and infant diarrhea in Rural Kenya.* (pp. 217–236). Boulder Colorado Westview Press 1989. https://www.popline.org/node/364943

Macintyre, S., Maciver, S., Sooman, A., Anderson, A., Blane, D., Ecob, R., Ford, G., Forsyth, A., Green, G., Kelly, M., Mullen, K., Phillimore, P., & Wight, D. (2019). Area, Class and Health: Should we be Focusing on Places or People?*. *NETSCC, On*, *22*, 213–234. https://doi.org/10.1017/S0047279400019310

Majekodunmi, A. O., Fajinmi, A., Dongkum, C., Picozzi, K., Macleod, E., Thrusfield, M. V., M Shaw, A. P., & Welburn, S. C. (2013). Social factors affecting seasonal variation in bovine trypanosomiasis on the Jos Plateau, Nigeria. *Parasites and Vectors*, *6*(1). https://doi.org/10.1186/1756-3305-6-293

Mansell Prothero, R. (1963). Population Mobility and Trypanosomiasis in Africa*. In *Bull. Org. mond. Sante* (Vol. 28).

Martens, P., & Hall, L. (2000). Malaria on the move: human population movement and malaria transmission. *Emerging Infectious Diseases*, *6*(2), 103–109. https://doi.org/10.3201/eid0602.000202

Martini, V., Chevalier, V., Ceccato, P., Anyamba, A., De Simone, L., Lubroth, J., De La Rocque, S., & Domenech, J. (2008). The impact of climate change on the epidemiology and control of Rift Valley fever. *OIE Revue Scientifique et Technique*, *27*(2), 413–426. https://doi.org/10.20506/rst.27.2.1802

Mason, L., Dellicour, S., Ter Kuile, F., Ouma, P., Phillips-Howard, P., Were, F., Laserson, K., & Desai, M. (2015). Barriers and facilitators to antenatal and delivery care in western Kenya: a qualitative study. *BMC Pregnancy and Childbirth*, *15*(1), 26. https://doi.org/10.1186/s12884-015-0453-z

Maudlin, I., Eisler, M. C., & Welburn, S. C. (2009). Neglected and endemic zoonoses. *Philosophical Transactions of the Royal Society of London. Series B, Biological Sciences*, *364*(1530), 2777–2787. https://doi.org/10.1098/rstb.2009.0067

Mazé-Guilmo, E., Blanchet, S., McCoy, K. D., & Loot, G. (2016). Host dispersal as the driver of parasite genetic structure: a paradigm lost? *Ecology Letters*, *19*(3), 336–347. https://doi.org/10.1111/ele.12564

McGuckin, N., & Murakami, E. (1999). Examining Trip-Chaining Behavior: Comparison of Travel by Men and Women. *Transportation Research Record: Journal of the Transportation Research Board*, *1693*(1), 79–85. https://doi.org/10.3141/1693-12

Messina, J. P., Pigott, D. M., Golding, N., Duda, K. A., Brownstein, J. S., Weiss, D. J., Gibson, H., Robinson, T. P., Gilbert, M., William Wint, G. R., Nuttall, P. A., Gething, P. W., Myers, M. F., George, D. B., & Hay, S. I. (2015). The global distribution of Crimean-Congo hemorrhagic fever. *Transactions of The Royal Society of Tropical Medicine and Hygiene*, *109*(8), 503–513. https://doi.org/10.1093/trstmh/trv050

Mobile Action Technology Inc., Taipei, Taiwan. (2018). Mobile Action Technology Inc. http://global.mobileaction.com/support/support_igotU_Faq.jsp

Moïsi, J. C., Gatakaa, H., Noor, A. M., Williams, T. N., Bauni, E., Tsofa, B., Levine, O. S., & Scott, J. A. G. (2010). Geographic access to care is not a determinant of child mortality in a rural Kenyan setting with high health facility density. *BMC Public Health*, *10*(1), 142. https://doi.org/10.1186/1471-2458-10-142

Molyneux, C. S., Mung'ala-Odera, V., Harpham, T., & Snow, R. W. (1999). Maternal responses to childhood fevers: a comparison of rural and urban residents in coastal Kenya. *Tropical Medicine and International Health*, *4*(12), 836–845. https://doi.org/10.1046/j.1365-3156.1999.00489.x

Molyneux, D. H. (2014). Neglected tropical diseases: now more than just 'other diseases'--the post-2015 agenda. *International Health*, *6*(3), 172–180. https://doi.org/10.1093/inthealth/ihu037

Molyneux, D., Hallaj, Z., Keusch, G. T., McManus, D. P., Ngowi, H., Cleaveland, S., Ramos-Jimenez, P., Gotuzzo, E., Kar, K., Sanchez, A., Garba, A., Carabin, H., Bassili, A., Chaignat, C. L., Meslin, F.-X., Abushama, H. M., Willingham, A. L., & Kioy, D. (2011). Zoonoses and marginalised infectious diseases of poverty: where do we stand? *Parasites & Vectors*, *4*(1), 106. https://doi.org/10.1186/1756-3305-4-106

Moonasar, D., Maharaj, R., Kunene, S., Candrinho, B., Saute, F., Ntshalintshali, N., & Morris, N. (2016). Towards malaria elimination in the MOSASWA (Mozambique, South Africa and Swaziland) region. In *Malaria Journal* (Vol. 15, Issue 1, pp. 1–5). BioMed Central Ltd. https://doi.org/10.1186/s12936-016-1470-8

Moreno, J. E., Rubio-Palis, Y., Páez, E., Pérez, E., & Sánchez, V. (2007). Abundance, biting behaviour and parous rate of anopheline mosquito species in relation to malaria incidence in gold-mining areas of southern Venezuela. *Medical and Veterinary Entomology*, *21*(4), 339–349. https://doi.org/10.1111/j.1365-2915.2007.00704.x

Moritz, M., Galehouse, Z., Hao, Q., & Garabed, R. B. (2012). Can One Animal Represent an Entire Herd? Modeling Pastoral Mobility Using GPS/GIS Technology. *Human Ecology*, *40*(4), 623–630. https://doi.org/10.1007/s10745-012-9483-6

Morris, K. R. S. (1962). The epidemiology of sleeping sickness in East Africa. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, *56*(4), 316–338. https://doi.org/10.1016/0035-9203(62)90055-X

Morrison, L. J., Vezza, L., Rowan, T., & Hope, J. C. (2016). Animal African Trypanosomiasis: Time to Increase Focus on Clinically Relevant Parasite and Host Species. *Trends in Parasitology*, *32*(8), 599–607. https://doi.org/10.1016/J.PT.2016.04.012

Mubyazi, G. M., Bloch, P., Magnussen, P., Olsen, Ø. E., Byskov, J., Hansen, K. S., & Bygbjerg, I. C. (2010). Women's experiences and views about costs of seeking malaria chemoprevention and other antenatal services: a qualitative study from two districts in rural Tanzania. *Malaria Journal*, *9*(1), 54. https://doi.org/10.1186/1475-2875-9-54

Mumtaz, Z., & Salway, S. (2005). 'I never go anywhere': extricating the links between women's mobility and uptake of reproductive health services in Pakistan. *Social Science & Medicine*, *60*(8), 1751–1765. https://doi.org/10.1016/J.SOCSCIMED.2004.08.019

Murali, J., & Afifi, T. (2014). Rainfall variability, food security and human mobility in the Janjgir-Champa district of Chhattisgarh state, India. *Climate and Development*, *6*(1), 28–37. https://doi.org/10.1080/17565529.2013.867248

Mutuku, F., Bayoh, M., Hightower, A., Vulule, J., Gimnig, J., Mueke, J., Amimo, F., & Walker, E. (2009). A supervised land cover classification of a western Kenya lowland endemic for human malaria: associations of land cover with larval Anopheles habitats. *International Journal of Health Geographics*, *8*(1), 19. https://doi.org/10.1186/1476-072X-8-19

Mwabu, G. M. (1986). Health care decisions at the household level: Results of a rural health survey in Kenya. *Social Science and Medicine*. https://doi.org/10.1016/0277-9536(86)90129-2

Mwabu, G., Mwanzia, J., & Liambila, W. (1995). User charges in government health facilities in Kenya: Effect on attendance and revenue. *Health Policy and Planning*, *10*(2), 164–170. https://doi.org/10.1093/heapol/10.2.164

Nanyingi, M. O., Munyua, P., Kiama, S. G., Muchemi, G. M., Thumbi, S. M., Bitek, A. O., Bett, B., Muriithi, R. M., & Njenga, M. K. (2015). A systematic review of Rift Valley Fever epidemiology 1931-2014. *Infection Ecology & Epidemiology*, *5*, 28024. http://www.pubmedcentral.nih.gov/articlerender.fcgi?artid=4522434&tool=pmcentrez&rendertype=abstract

Natale, F., Giovannini, A., Savini, L., Palma, D., Possenti, L., Fiore, G., & Calistri, P. (2009). Network analysis of Italian cattle trade patterns and evaluation of risks for potential disease spread. *Preventive Veterinary Medicine*, *92*(4), 341–350. https://doi.org/10.1016/j.prevetmed.2009.08.026

National Institute of Statistics (INE) (Mozambique). (2017). *Mozambique Population and Housing*

Census 2017.

Njeru, J., Wareth, G., Melzer, F., Henning, K., Pletz, M. W., Heller, R., & Neubauer, H. (2016). Systematic review of brucellosis in Kenya: disease frequency in humans and animals and risk factors for human infection. *BMC Public Health*, *16*(1), 853. https://doi.org/10.1186/s12889-016-3532-9

Nnko, H. J., Ngonyoka, A., Salekwa, L., Estes, A. B., Hudson, P. J., Gwakisa, P. S., & Cattadori, I. M. (2017). Seasonal variation of tsetse fly species abundance and prevalence of trypanosomes in the Maasai Steppe, Tanzania. *Journal of Vector Ecology*, *42*(1), 24–33. https://doi.org/10.1111/jvec.12236

Noor, A. M., Zurovac, D., Hay, S. I., Ochola, S. A., & Snow, R. W. (2003). Defining equity in physical access to clinical services using geographical information systems as part of malaria planning and monitoring in Kenya. *Tropical Medicine and International Health*, *8*(10), 917–926. https://doi.org/10.1046/j.1365-3156.2003.01112.x

Noor, A.M, Gikandi, P. ., Hay, S. ., Muga, R. ., & Snow, R. . (2004). Creating spatially defined databases for equitable health service planning in low-income countries: the example of Kenya. *Acta Tropica*, *91*(3), 239–251. https://doi.org/10.1016/J.ACTATROPICA.2004.05.003

Noor, Abdisalan M, Alegana, V. A., Gething, P. W., & Snow, R. W. (2009). A spatial national health facility database for public health sector planning in Kenya in 2008. *International Journal of Health Geographics*, *8*(1), 13. https://doi.org/10.1186/1476-072X-8-13

Okwi, P. O., Ndeng’e C, G., Kristjanson, P., Arunga, M., Notenbaert, A., Omolo, A., Henninger, N., Benson, T., Kariuki, P., & Owuor, J. (2007). *Spatial determinants of poverty in rural Kenya*. www.pnas.org/cgi/content/full/

Okwi, P. O., Ndeng’e, G., Kristjanson, P., Arunga, M., Notenbaert, A., Omolo, A., Henninger, N., Benson, T., Kariuki, P., & Owuor, J. (2007). Spatial determinants of poverty in rural Kenya. *Proceedings of the National Academy of Sciences*, *104*(43), 16769–16774. https://doi.org/10.1073/PNAS.0611107104

Omukunda, E., Githeko, A., Ndong A, M. F., Mushinzimana, E., & Yan, G. (2012). Effect of swamp cultivation on distribution of anopheline larval habitats in Western Kenya. *Journal of Vector Borne Diseases*, *49*(2), 61–71. http://www.ncbi.nlm.nih.gov/pubmed/22898476

Paaijmans, K. P., Blanford, S., Chan, B. H. K., & Thomas, M. B. (2012). Warmer temperatures reduce the vectorial capacity of malaria mosquitoes. *Biology Letters*, *8*(3), 465–468.

https://doi.org/10.1098/rsbl.2011.1075

Parsons, M. B., Gillespie, T. R., Lonsdorf, E. V, Travis, D., Lipende, I., Gilagiza, B., Kamenya, S., Pintea, L., & Vazquez-Prokopec, G. M. (2014). Global positioning system data-loggers: a tool to quantify fine-scale movement of domestic animals to evaluate potential for zoonotic transmission to an endangered wildlife population. *PloS One*, *9*(11), e110984. https://doi.org/10.1371/journal.pone.0110984

Paz-Soldan, V. A., Reiner, R. C., Morrison, A. C., Stoddard, S. T., Kitron, U., Scott, T. W., Elder, J. P., Halsey, E. S., Kochel, T. J., Astete, H., & Vazquez-Prokopec, G. M. (2014). Strengths and Weaknesses of Global Positioning System (GPS) Data-Loggers and Semi-structured Interviews for Capturing Fine-scale Human Mobility: Findings from Iquitos, Peru. *PLoS Neglected Tropical Diseases*, *8*(6), e2888. https://doi.org/10.1371/journal.pntd.0002888

Peak, C. M., Wesolowski, A., Zu Erbach-Schoenberg, E., Tatem, A. J., Wetter, E., Lu, X., Power, D., Weidman-Grunewald, E., Ramos, S., Moritz, S., Buckee, C. O., & Bengtsson, L. (2018). Population mobility reductions associated with travel restrictions during the Ebola epidemic in Sierra Leone: use of mobile phone data. *International Journal of Epidemiology*, *47*(5), 1562–1570. https://doi.org/10.1093/ije/dyy095

Pearce, J., Witten, K., Hiscock, R., & Blakely, T. (2008). Regional and Urban–Rural Variations in the Association of Neighbourhood Deprivation with Community Resource Access: A National Study. *Environment and Planning A: Economy and Space*, *40*(10), 2469–2489. https://doi.org/10.1068/a409

Pebesma, E. (2018). sf: Simple Features for R. *R Package Version 0.6-1*.

Perkins, T. A., Garcia, A. J., Paz-Soldan, V. A., Stoddard, S. T., Reiner, R. C., Vazquez-Prokopec, G., Bisanzio, D., Morrison, A. C., Halsey, E. S., Kochel, T. J., Smith, D. L., Kitron, U., Scott, T. W., & Tatem, A. J. (2014). Theory and data for simulating fine-scale human movement in an urban environment. *Journal of The Royal Society Interface*, *11*(99), 20140642–20140642. https://doi.org/10.1098/rsif.2014.0642

Perras, M., & Nebel, S. (2012). Satellite Telemetry and its Impact on the Study of Animal Migration. *Nature Education Knowledge*, *3*(12), 4.

Pesaresi, M., Florczyk, Anet Schiavina Marcello Melchiorri, M., & Maffenini, L. (2019). *GHS settlement grid, updated and refined REGIO model 2014 in application to GHS-BUILT R2018A and GHS-POP R2019A, multitemporal (1975-1990-2000-2015), R2019A.* European Commission, Joint Research Centre (JRC). https://doi.org/10.2760/062975

Peter, R. J., Van Den Bossche, P., Penzhorn, B. L., & Sharp, B. (2005). Tick, fly, and mosquito control - Lessons from the past, solutions for the future. *Veterinary Parasitology*, *132*(3-4 SPEC. ISS.), 205–215. https://doi.org/10.1016/j.vetpar.2005.07.004

Pindolia, D. K., Garcia, A. J., Huang, Z., Smith, D. L., Alegana, V. A., Noor, A. M., Snow, R. W., & Tatem, A. J. (2013). The demographics of human and malaria movement and migration patterns in East Africa. *Malaria Journal*, *12*(1), 397. https://doi.org/10.1186/1475-2875-12-397

Pindolia, D. K., Garcia, A. J., Wesolowski, A., Smith, D. L., Buckee, C. O., Noor, A. M., Snow, R. W., & Tatem, A. J. (2012). Human movement data for malaria control and elimination strategic planning. In *Malaria Journal* (Vol. 11, Issue 1, pp. 1–16). BioMed Central. https://doi.org/10.1186/1475-2875-11-205

Pons, P., & Lapaty, M. (2006). Computing communities in large networks using random walks. *Journal of Graph Algorithms and Applications*, *10*(2), 191–218. http://jgaa.info/accepted/2006/PonsLatapy2006.10.2.pdf

Poutanen, S. M., Low, D. E., Henry, B., Finkelstein, S., Rose, D., Green, K., Tellier, R., Draker, R., Adachi, D., Ayers, M., Chan, A. K., Skowronski, D. M., Salit, I., Simor, A. E., Slutsky, A. S., Doyle, P. W., Krajden, M., Petric, M., Brunham, R. C., … Detsky, A. (2003). Identification of Severe Acute Respiratory Syndrome in Canada. *New England Journal of Medicine*, *348*(20), 1995–2005. https://doi.org/10.1056/NEJMoa030634

Pray, I. W., Swanson, D. J., Ayvar, V., Muro, C., Moyano, L. M., Gonzalez, A. E., Garcia, H. H., O'Neal, S. E., & Cysticercosis Working Group in Peru. (2016). GPS Tracking of Free-Ranging Pigs to Evaluate Ring Strategies for the Control of Cysticercosis/Taeniasis in Peru. *PLoS Neglected Tropical Diseases*, *10*(4), e0004591. https://doi.org/10.1371/journal.pntd.0004591

Prothero, R. M. (2001). Migration and malaria risk. *Health, Risk and Society*, *3*(1), 19–38. https://doi.org/10.1080/713670171

Prothero, R. Mansell. (1977). Disease and Mobility: A Neglected Factor in Epidemiology. *International Journal of Epidemiology*, *6*(3), 259–267. https://doi.org/10.1093/ije/6.3.259

Pryce, J., Richardson, M., & Lengeler, C. (2018). Insecticide-treated nets for preventing malaria. *Cochrane Database of Systematic Reviews*. https://doi.org/10.1002/14651858.CD000363.pub3

Bibliography

Pucher, J., & Renne, J. L. (2005). Rural mobility and mode choice: Evidence from the 2001 National Household Travel Survey. *Transportation*, *32*(2), 165–186. https://doi.org/10.1007/s11116-004-5508-3

Pulliam, H. R. (1988). Sources, sinks and population regulation. *American Naturalist*, *132*(5), 652–661. https://doi.org/10.1086/284880

QGIS Development Team. (2019). *QGIS Geographic Information System.* Open Source Geospatial Foundation Project.

Quaglietta, L., Martins, B. H., de Jongh, A., Mira, A., & Boitani, L. (2012). A Low-Cost GPS GSM/GPRS Telemetry System: Performance in Stationary Field Tests and Preliminary Data on Wild Otters (Lutra lutra). *PLoS ONE*, *7*(1), e29235. https://doi.org/10.1371/journal.pone.0029235

R Core Team. (2017). *R: A language and environment for statistical computing.* (3.4.2). R Foundation for Statistical Computing. https://www.r-project.org/

R Core Team. (2019). R: A language and environment for statistical computing. In *R Foundation for Statistical Computing*.

Raman, J., Gast, L., Balawanth, R., Tessema, S., Brooke, B., Maharaj, R., Munhenga, G., Tshikae, P., Lakan, V., Mwamba, T., Makowa, H., Sangweni, L., Mkhabela, M., Zondo, N., Mohulatsi, E., Nyawo, Z., Ngxongo, S., Msimang, S., Dagata, N., … Moonasar, D. (2020). High levels of imported asymptomatic malaria but limited local transmission in KwaZulu-Natal, a South African malaria-endemic province nearing malaria elimination. *Malaria Journal*, *19*(1), 152. https://doi.org/10.1186/s12936-020-03227-3

Recchi, E., Deutschmann, E., & Vespe, M. (2019). Estimating Transnational Human Mobility on a Global Scale. *SSRN Electronic Journal*. https://doi.org/10.2139/ssrn.3384000

Reed, K. D., Meece, J. K., Henkel, J. S., & Shukla, S. K. (2003). Birds, migration and emerging zoonoses: west nile virus, lyme disease, influenza A and enteropathogens. In *Clinical medicine & research* (Vol. 1, Issue 1, pp. 5–12). Marshfield Clinic. https://doi.org/10.3121/cmr.1.1.5

Rich, K. M., & Wanyoike, F. (2010). An assessment of the regional and national socio-economic impacts of the 2007 Rift Valley fever outbreak in Kenya. *The American Journal of Tropical Medicine and Hygiene*, *83*(2 Suppl), 52–57. https://doi.org/10.4269/ajtmh.2010.09-0291

Riley, S. (2007). Large-scale spatial-transmission models of infectious disease. *Science*, *316*(5829), 1298–1301. https://doi.org/10.1126/science.1134695

Robert, V., Macintyre, K., Keating, J., Trape, J.-F., Duchemin, J.-B., Warren, M., & Beier, J. C. (2003). Malaria transmission in urban sub-Saharan Africa. *The American Journal of Tropical Medicine and Hygiene*, *68*(2), 169–176. http://www.ncbi.nlm.nih.gov/pubmed/12641407

Roth, G. A., Abate, D., Abate, K. H., Abay, S. M., Abbafati, C., Abbasi, N., Abbastabar, H., Abd-Allah, F., Abdela, J., Abdelalim, A., Abdollahpour, I., Abdulkader, R. S., Abebe, H. T., Abebe, M., Abebe, Z., Abejie, A. N., Abera, S. F., Abil, O. Z., Abraha, H. N., … Murray, C. J. L. (2018). Global, regional, and national age-sex-specific mortality for 282 causes of death in 195 countries and territories, 1980–2017: a systematic analysis for the Global Burden of Disease Study 2017. *The Lancet*, *392*(10159), 1736–1788. https://doi.org/10.1016/S0140-6736(18)32203-7

Rottier, E., & Ince, M. (2005). Chapter 2 Disease and disease transmission. In *Controlling and Preventing Disease* (pp. 7–27).

Ruktanonchai, C. W., Ruktanonchai, N. W., Nove, A., Lopes, S., Pezzulo, C., Bosco, C., Alegana, V. A., Burgert, C. R., Ayiko, R., Charles, A. S., Lambert, N., Msechu, E., Kathini, E., Matthews, Z., & Tatem, A. J. (2016). Equality in Maternal and Newborn Health: Modelling Geographic Disparities in Utilisation of Care in Five East African Countries. *PLOS ONE*, *11*(8), e0162006. https://doi.org/10.1371/journal.pone.0162006

Ruktanonchai, N. W., DeLeenheer, P., Tatem, A. J., Alegana, V. A., Caughlin, T. T., zu Erbach-Schoenberg, E., Lourenço, C., Ruktanonchai, C. W., & Smith, D. L. (2016). Identifying Malaria Transmission Foci for Elimination Using Human Mobility Data. *PLoS Computational Biology*, *12*(4), e1004846. https://doi.org/10.1371/journal.pcbi.1004846

Ruktanonchai, N.W., Ruktanonchai, C. W., Floyd, J. R., & Tatem, A. J. (2018). Using Google Location History data to quantify fine-scale human mobility. *International Journal of Health Geographics*, *17*(1). https://doi.org/10.1186/s12942-018-0150-z

Ruktanonchai, Nick W., Bhavnani, D., Sorichetta, A., Bengtsson, L., Carter, K. H., Córdoba, R. C., Le Menach, A., Lu, X., Wetter, E., Zu Erbach-Schoenberg, E., & Tatem, A. J. (2016). Census-derived migration data as a tool for informing malaria elimination policy. *Malaria Journal*, *15*(1), 273. https://doi.org/10.1186/s12936-016-1315-5

Ryan, P. G., Petersen, S. L., Peters, G., & Gremillet, D. (2004). GPS tracking a marine predator: the effects of precision, resolution and sampling rate on foraging tracks of African Penguins. *Marine Biology*, *145*(2), 215–223. https://doi.org/10.1007/s00227-004-1328-4

Bibliography

Salon, D., & Aligula, E. M. (2012). Urban travel in Nairobi, Kenya: analysis, insights, and opportunities. *Journal of Transport Geography*, *22*, 65–76. https://doi.org/10.1016/J.JTRANGEO.2011.11.019

Salon, D., & Gulyani, S. (2010). Mobility, Poverty, and Gender: Travel 'Choices' of Slum Residents in Nairobi, Kenya. *Transport Reviews*, *30*(5), 641–657. https://doi.org/10.1080/01441640903298998

Sarfraz, M. S., Tripathi, N. K., Tipdecho, T., Thongbu, T., Kerdthong, P., & Souris, M. (2012). Analyzing the spatio-temporal relationship between dengue vector larval density and land-use using factor analysis and spatial ring mapping. *BMC Public Health*, *12*(1), 853. https://doi.org/10.1186/1471-2458-12-853

Sarmiento, S. (2000). Household, Gender, and Travel. *Women's Travel Issues, Proceedings from the Second National Conference.* https://www.fhwa.dot.gov/ohim/womens/chap3.pdf

Sarr, M. D., Sylla, M., Foy, B. D., Chapman, P. L., & Kobylinski, K. C. (2011). Ivermectin Mass Drug Administration to Humans Disrupts Malaria Parasite Transmission in Senegalese Villages. *The American Journal of Tropical Medicine and Hygiene*, *85*(1), 3–5. https://doi.org/10.4269/ajtmh.2011.11-0160

Saúde/INS, I. N. de, Estatistica/INE, I. N. de, Malária/PNCM, P. N. de C. da, & ICF. (2019). *Moçambique Inquérito Nacional sobre Indicadores de Malária (IIM) 2018*. https://www.dhsprogram.com/publications/publication-mis33-mis-final-reports.cfm

Schicker, R. S., Hiruy, N., Melak, B., Gelaye, W., Bezabih, B., Stephenson, R., Patterson, A. E., Tadesse, Z., Emerson, P. M., Richards, F. O., & Noland, G. S. (2015). A Venue-Based Survey of Malaria, Anemia and Mobility Patterns among Migrant Farm Workers in Amhara Region, Ethiopia. *PLOS ONE*, *10*(11), e0143829. https://doi.org/10.1371/journal.pone.0143829

Schneider, C. M., Belik, V., Couronné, T., Smoreda, Z., & González, M. C. (2013). Unravelling daily human mobility motifs. *Journal of the Royal Society Interface*, *10*(84), 20130246. https://doi.org/10.1098/rsif.2013.0246

Schröder, L., Flägel, K., Goetz, K., & Steinhäuser, J. (2018). Mobility concepts and access to health care in a rural district in Germany: a mixed methods approach. *BMC Family Practice*, *19*(1), 47. https://doi.org/10.1186/s12875-018-0733-6

Schwemmer, P., & Garthe, S. (2011). Spatial and temporal patterns of habitat use by Eurasian oystercatchers (Haematopus ostralegus) in the eastern Wadden Sea revealed using GPS data loggers. *Marine Biology*, *158*(3), 541–550. https://doi.org/10.1007/s00227-010-1579-1

Searle, K. M., Lubinda, J., Hamapumbu, H., Shields, T. M., Curriero, F. C., Smith, D. L., Thuma, P. E., & Moss, W. J. (2017). Characterizing and quantifying human movement patterns using GPS data loggers in an area approaching malaria elimination in rural southern Zambia. *Royal Society Open Science*, *4*(5), 170046. https://doi.org/10.1098/rsos.170046

Selby, R., Bardosh, K., Picozzi, K., Waiswa, C., & Welburn, S. C. (2013). Cattle movements and trypanosomes: restocking efforts and the spread of Trypanosoma brucei rhodesiense sleeping sickness in post-conflict Uganda. *Parasites & Vectors*, *6*(1), 281. https://doi.org/10.1186/1756-3305-6-281

Shayo, E. H., Rumisha, S. F., Mlozi, M. R. S., Bwana, V. M., Mayala, B. K., Malima, R. C., Mlacha, T., & Mboera, L. E. G. (2015). Social determinants of malaria and health care seeking patterns among rice farming and pastoral communities in Kilosa District in central Tanzania. *Acta Tropica*, *144*, 41–49. https://doi.org/10.1016/J.ACTATROPICA.2015.01.003

Shi, W. (2019). *Connecting Africa - Tmcel: The New Name in Mozambique's Comms Market*. http://www.connectingafrica.com/author.asp?section_id=725&doc_id=748991&

Siedner, M. J., Lankowski, A., Tsai, A. C., Muzoora, C., Martin, J. N., Hunt, P. W., Haberer, J. E., & Bangsberg, D. R. (2013). GPS-measured distance to clinic, but not self-reported transportation factors, are associated with missed HIV clinic visits in rural Uganda. *AIDS (London, England)*, *27*(9), 1503–1508. https://doi.org/10.1097/QAD.0b013e32835fd873

Smith, C., & Whittaker, M. (2014). Beyond mobile populations: A critical review of the literature on malaria and population mobility and suggestions for future directions. *Malaria Journal*, *13*(1), 1–10. https://doi.org/10.1186/1475-2875-13-307

Stanaway, J. D., Shepard, D. S., Undurraga, E. A., Halasa, Y. A., Coffeng, L. E., Brady, O. J., Hay, S. I., Bedi, N., Bensenor, I. M., Castañeda-Orjuela, C. A., Chuang, T. W., Gibney, K. B., Memish, Z. A., Rafay, A., Ukwaja, K. N., Yonemoto, N., & Murray, C. J. L. (2016). The global burden of dengue: an analysis from the Global Burden of Disease Study 2013. *The Lancet Infectious Diseases*, *16*(6), 712–723. https://doi.org/10.1016/S1473-3099(16)00026-8

Stantic-Pavlinic, M., Xiao, L., Glaberman, S., Lal, A. A., Oražen, T., Rataj-Verglez, A., Logar, J., & Berce, I. (2003). Cryptosporidiosis associated with animal contacts. *Wiener Klinische Wochenschrift*, *115*(3–4), 125–127. https://doi.org/10.1007/BF03040292

Stefani, A., Dusfour, I., Corrêa, A. P. S., Cruz, M. C., Dessay, N., Galardo, A. K., Galardo, C. D., Girod, R., Gomes, M. S., Gurgel, H., Lima, A. C. F., Moreno, E. S., Musset, L., Nacher, M., Soares, A. C., Carme, B., & Roux, E. (2013). Land cover, land use and malaria in the Amazon: a systematic literature review of studies using remotely sensed data. *Malaria Journal*, *12*(1), 192. https://doi.org/10.1186/1475-2875-12-192

Stoddard, S. T., Forshey, B. M., Morrison, A. C., Paz-Soldan, V. A., Vazquez-Prokopec, G. M., Astete, H., Reiner, R. C., Vilcarromero, S., Elder, J. P., Halsey, E. S., Kochel, T. J., Kitron, U., & Scott, T. W. (2013). House-to-house human movement drives dengue virus transmission. *Proceedings of the National Academy of Sciences of the United States of America*, *110*(3), 994–999. https://doi.org/10.1073/pnas.1213349110

Stoddard, S. T., Morrison, A. C., Vazquez-Prokopec, G. M., Soldan, V. P., Kochel, T. J., Kitron, U., Elder, J. P., & Scott, T. W. (2009). The role of human movement in the transmission of vector-borne pathogens. *PLoS Neglected Tropical Diseases*, *3*(7), e481. https://doi.org/10.1371/journal.pntd.0000481

Stopher, P. R., Daigler, V., & Griffith, S. (2018). Smartphone app versus GPS Logger: A comparative study. *Transportation Research Procedia*, *32*, 135–145. https://doi.org/10.1016/J.TRPRO.2018.10.026

Strano, E., Viana, M. P., Sorichetta, A., & Tatem, A. J. (2018). Mapping road network communities for guiding disease surveillance and control strategies. *Scientific Reports*, *8*(1), 4744. https://doi.org/10.1038/s41598-018-22969-4

Stresman, G., Bousema, T., & Cook, J. (2019). Malaria Hotspots: Is There Epidemiological Evidence for Fine-Scale Spatial Targeting of Interventions? In *Trends in Parasitology* (Vol. 35, Issue 10, pp. 822–834). Elsevier Ltd. https://doi.org/10.1016/j.pt.2019.07.013

Sumner, M. D. (2016). *trip: Tools for the Analysis of Animal Track Data. R package version 1.5.0.* https://cran.r-project.org/package=trip

Sumner, T., Orton, R. J., Green, D. M., Kao, R. R., & Gubbins, S. (2017). Quantifying the roles of host movement and vector dispersal in the transmission of vector-borne diseases of livestock. *PLOS Computational Biology*, *13*(4), e1005470. https://doi.org/10.1371/journal.pcbi.1005470

Tanser, F., Gijsbertsen, B., & Herbst, K. (2006). Modelling and understanding primary health care accessibility and utilization in rural South Africa: An exploration using a geographical information system. *Social Science & Medicine*, *63*(3), 691–705. https://doi.org/10.1016/J.SOCSCIMED.2006.01.015

Tatem, A. J., Huang, Z., Das, A., Qi, Q., Roth, J., & Qiu, Y. (2012). Air travel and vector-borne disease movement. *Parasitology*, *139*(14), 1816–1830. https://doi.org/10.1017/S0031182012000352

Tatem, Andrew J, Huang, Z., Narib, C., Kumar, U., Kandula, D., Pindolia, D. K., Smith, D. L., Cohen, J. M., Graupe, B., Uusiku, P., & Lourenço, C. (2014). Integrating rapid risk mapping and mobile phone call record data for strategic malaria elimination planning. *Malaria Journal*, *13*, 52. https://doi.org/10.1186/1475-2875-13-52

Tatem, Andrew J, & Smith, D. L. (2010). International population movements and regional Plasmodium falciparum malaria elimination strategies. *Proceedings of the National Academy of Sciences of the United States of America*, *107*(27), 12222–12227. https://doi.org/10.1073/pnas.1002971107

Taylor, D., Hagenlocher, M., Jones, A. E., Kienberger, S., Leedale, J., & Morse, A. P. (2016). Environmental change and Rift Valley fever in eastern Africa: projecting beyond HEALTHY FUTURES. *Geospatial Health*, *11*(1s).

Tejedor-Garavito, N., Dlamini, N., Pindolia, D., Soble, A., Ruktanonchai, N. W., Alegana, V., Le Menach, A., Ntshalintshali, N., Dlamini, B., Smith, D. L., Tatem, A. J., & Kunene, S. (2017). Travel patterns and demographic characteristics of malaria cases in Swaziland, 2010–2014. *Malaria Journal*, *16*(1), 359. https://doi.org/10.1186/s12936-017-2004-8

Tessema, S., Wesolowski, A., Chen, A., Murphy, M., Wilheim, J., Mupiri, A.-R., Ruktanonchai, N. W., Alegana, V. A., Tatem, A. J., Tambo, M., Didier, B., Cohen, J. M., Bennett, A., Sturrock, H. J., Gosling, R., Hsiang, M. S., Smith, D. L., Mumbengegwi, D. R., Smith, J. L., & Greenhouse, B. (2019). Using parasite genetic and human mobility data to infer local and cross-border malaria connectivity in Southern Africa. *ELife*, *8*. https://doi.org/10.7554/eLife.43510

Thaddeus, S., & Maine, D. (1994). Too Far To Walk : Maternal Mortality in Context. *Social Science and Medicine*, *38*(8), 1091–1110. https://ac.els-cdn.com/0277953694902267/1-s2.0-0277953694902267-main.pdf?_tid=bb92a818-75e2-47e4-be70-ee939118e6a9&acdnat=1548672198_0a6d8722bf63445c15de9f69724f169e

Thaker, M., Gupte, P. R., Prins, H. H. T., Slotow, R., & Vanak, A. T. (2019). Fine-Scale Tracking of Ambient Temperature and Movement Reveals Shuttling Behavior of Elephants to Water. *Frontiers in Ecology and Evolution*, *7*, 4. https://doi.org/10.3389/fevo.2019.00004

Thomas, L., de Glanville, W. A., Cook, E. A., & Fèvre, E. M. (2013). The spatial ecology of free-ranging domestic pigs (Sus scrofa) in western Kenya. *BMC Veterinary Research*, *9*(1), 46.

https://doi.org/10.1186/1746-6148-9-46

Thomas, Lian F., de Glanville, W. A., Cook, E. A., & Fèvre, E. M. (2013). The spatial ecology of free-ranging domestic pigs (Sus scrofa) in western Kenya. *BMC Veterinary Research*, *9*, 46. https://doi.org/10.1186/1746-6148-9-46

Thomas, Lian Francesca, Harrison, L. J. S., Toye, P., de Glanville, W. A., Cook, E. A. J., Wamae, C. N., & Fèvre, E. M. (2016). Prevalence of Taenia solium cysticercosis in pigs entering the food chain in western Kenya. *Tropical Animal Health and Production*, *48*(1), 233–238. https://doi.org/10.1007/s11250-015-0949-6

Thuranira-McKeever, C., Shaw, A., Machila, N., Eisler, M., Welburn, S., & Maudlin, I. (2010). Seasonal influences on livestock keeping in a sedentary crop–livestock system. *Tropical Animal Health and Production*, *42*(4), 705–717. https://doi.org/10.1007/s11250-009-9478-5

Tigoi, C., Sang, R., Chepkorir, E., Orindi, B., Arum, S. O., Mulwa, F., Mosomtai, G., Limbaso, S., Hassan, O. A., Irura, Z., Ahlm, C., & Evander, M. (2020). High risk for human exposure to Rift Valley fever virus in communities living along livestock movement routes: A cross-sectional survey in Kenya. *PLOS Neglected Tropical Diseases*, *14*(2), e0007979. https://doi.org/10.1371/journal.pntd.0007979

Tomkiewicz, S. M., Fuller, M. R., Kie, J. G., & Bates, K. K. (2010). Global positioning system and associated technologies in animal behaviour and ecological research. *Philosophical Transactions of the Royal Society B: Biological Sciences*, *365*(1550), 2163–2176. https://doi.org/10.1098/rstb.2010.0090

Tompkins, D. M., Carver, S., Jones, M. E., Krkošek, M., & Skerratt, L. F. (2015). Emerging infectious diseases of wildlife: a critical perspective. *Trends in Parasitology*, *31*(4), 149–159. https://doi.org/10.1016/j.pt.2015.01.007

Touré, Y. T., Dolo, G., Petrarca, V., Traoré, S. F., Bouaré, M., Dao, A., Carnahan, J., & Taylor, C. E. (1998). Mark-release-recapture experiments with Anopheles gambiae s.l. in Banambani village, Mali, to determine population size and structure. *Medical and Veterinary Entomology*, *12*(1), 74–83. https://doi.org/10.1046/j.1365-2915.1998.00071.x

Travis, D. A., Watson, R. P., & Tauer, A. (2011). The spread of pathogens through trade in wildlife. *Revue Scientifique et Technique (International Office of Epizootics)*, *30*(1), 219–239. http://www.ncbi.nlm.nih.gov/pubmed/21809766

Van Borm, S., Thomas, I., Hanquet, G., Lambrecht, B., Boschmans, M., Dupont, G., Decaestecker, M., Snacken, R., & van den Berg, T. (2005). Highly pathogenic H5N1 influenza virus in smuggled Thai eagles, Belgium. *Emerging Infectious Diseases*, *11*(5), 702–705. https://doi.org/10.3201/eid1105.050211

Van den Bossche, P., Rocque, S. de La, Hendrickx, G., & Bouyer, J. (2010). A changing environment and the epidemiology of tsetse-transmitted livestock trypanosomiasis. *Trends in Parasitology*, *26*(5), 236–243. https://doi.org/10.1016/J.PT.2010.02.010

Vasconcelos, P. F., Costa, Z. G., Travassos Da Rosa, E. S., Luna, E., Rodrigues, S. G., Barros, V. L., Dias, J. P., Monteiro, H. A., Oliva, O. F., Vasconcelos, H. B., Oliveira, R. C., Sousa, M. R., Barbosa Da Silva, J., Cruz, A. C., Martins, E. C., & Travassos Da Rosa, J. F. (2001). Epidemic of jungle yellow fever in Brazil, 2000: implications of climatic alterations in disease spread. *Journal of Medical Virology*, *65*(3), 598–604. http://www.ncbi.nlm.nih.gov/pubmed/11596099

Vazquez-Prokopec, G. M., Bisanzio, D., Stoddard, S. T., Paz-Soldan, V., Morrison, A. C., Elder, J. P., Ramirez-Paredes, J., Halsey, E. S., Kochel, T. J., Scott, T. W., & Kitron, U. (2013). Using GPS technology to quantify human mobility, dynamic contacts and infectious disease dynamics in a resource-poor urban environment. *PloS One*, *8*(4), e58802. https://doi.org/10.1371/journal.pone.0058802

Velthuis, A. G. J., & Mourits, M. C. M. (2007). Effectiveness of movement-prevention regulations to reduce the spread of foot-and-mouth disease in The Netherlands. *Preventive Veterinary Medicine*, *82*(3–4), 262–281. https://doi.org/10.1016/j.prevetmed.2007.05.023

Voeten, H. A. C. M., O'Hara, H. B., Kusimba, J., Otido, J. M., Ndinya-Achola, J. O., Bwayo, J. J., Varkevisser, C. M., & Habbema, J. D. F. (2004). Gender Differences in Health Care-Seeking Behavior for Sexually Transmitted Diseases: A Population-Based Study in Nairobi, Kenya. *Sexually Transmitted Diseases*, *31*(5), 265–272. https://doi.org/10.1097/01.OLQ.0000124610.65396.52

von Wissmann, B., Machila, N., Picozzi, K., Fèvre, E. M., deC. Bronsvoort, B. M., Handel, I. G., & Welburn, S. C. (2011). Factors Associated with Acquisition of Human Infective and Animal Infective Trypanosome Infections in Domestic Livestock in Western Kenya. *PLoS Neglected Tropical Diseases*, *5*(1), e941. https://doi.org/10.1371/journal.pntd.0000941

Wagstaff, A. (2002). Poverty and health sector inequalities. *Bulletin of the World Health Organization*, *80*, 97–105. https://doi.org/10.1590/S0042-96862002000200004

Wardrop, N. A. (2014). *Landcover classification: western Kenya, 2010.*

https://doi.org/10.5258/SOTON/383135

Wardrop, N. A., Thomas, L. F., Cook, E. A. J., de Glanville, W. A., Atkinson, P. M., Wamae, C. N., & Fèvre, E. M. (2016). The Sero-epidemiology of Coxiella burnetii in Humans and Cattle, Western Kenya: Evidence from a Cross-Sectional Study. *PLOS Neglected Tropical Diseases*, *10*(10), e0005032. https://doi.org/10.1371/journal.pntd.0005032

Watson, J. T., & Stoll, M. (2013). Gendered Logistic Mobility Among the Earliest Farmers in the Sonoran Desert. *Latin American Antiquity*, *24*(4), 433–450. https://doi.org/10.7183/1045-6635.24.4.433

Weiner, E. (1992). *Urban transportation planning in the United States: an historical overview*. https://books.google.com/books?hl=en&lr=&id=HB5PAAAAMAAJ&oi=fnd&pg=PR1&ots=pQXTWogyIw&sig=M7rf9dDBNgpF30Nl86-R8C95JQc

Weiss, D. J., Lucas, T. C. D., Nguyen, M., Nandi, A. K., Bisanzio, D., Battle, K. E., Cameron, E., Twohig, K. A., Pfeffer, D. A., Rozier, J. A., Gibson, H. S., Rao, P. C., Casey, D., Bertozzi-Villa, A., Collins, E. L., Dalrymple, U., Gray, N., Harris, J. R., Howes, R. E., … Gething, P. W. (2019). Mapping the global prevalence, incidence, and mortality of Plasmodium falciparum, 2000–17: a spatial and temporal modelling study. *The Lancet*, *394*(10195), 322–331. https://doi.org/10.1016/S0140-6736(19)31097-9

Wells, C. R., Sah, P., Moghadas, S. M., Pandey, A., Shoukat, A., Wang, Y., Wang, Z., Meyers, L. A., Singer, B. H., & Galvani, A. P. (2020). Impact of international travel and border control measures on the global spread of the novel 2019 coronavirus outbreak. *Proceedings of the National Academy of Sciences of the United States of America*, *117*(13), 7504–7509. https://doi.org/10.1073/pnas.2002616117

Wen, T.-H., Lin, M.-H., & Fang, C.-T. (2012). Population Movement and Vector-Borne Disease Transmission: Differentiating Spatial–Temporal Diffusion Patterns of Commuting and Noncommuting Dengue Cases. *Annals of the Association of American Geographers*, *102*(5), 1026–1037. https://doi.org/10.1080/00045608.2012.671130

Wesolowski, A., Buckee, C. O., Pindolia, D. K., Eagle, N., Smith, D. L., & Garcia, A. J. (2013). The Use of Census Migration Data to Approximate Human Movement Patterns across Temporal Scales. *PLoS ONE*, *8*(1), no pagination. http://www.plosone.org/article/fetchObjectAttachment.action?uri=info%3Adoi%2F10.1371%2Fjournal.pone.0052971&representation=PDF%5Cnhttp://ovidsp.ovid.com/ovidweb.cgi?T=JS&PAGE=reference&D=emed15&NEWS=N&AN=368101025

Wesolowski, Amy, Eagle, N., Noor, A. M., Snow, R. W., & Buckee, C. O. (2013). The impact of biases in mobile phone ownership on estimates of human mobility. *Journal of The Royal Society Interface*, *10*(81), 20120986. https://doi.org/10.1098/rsif.2012.0986

Wesolowski, Amy, Eagle, N., Tatem, A. J., Smith, D. L., Noor, A. M., Snow, R. W., & Buckee, C. O. (2012). Quantifying the Impact of Human Mobility on Malaria. *Science*, *338*(6104), 267–270. https://doi.org/10.1126/science.1223467

Wesolowski, Amy, Metcalf, C. J. E., Eagle, N., Kombich, J., Grenfell, B. T., Bjørnstad, O. N., Lessler, J., Tatem, A. J., & Buckee, C. O. (2015). Quantifying seasonal population fluxes driving rubella transmission dynamics using mobile phone data. *Proceedings of the National Academy of Sciences of the United States of America*, *112*(35), 11114–11119. https://doi.org/10.1073/pnas.1423542112

Wesolowski, Amy, Stresman, G., Eagle, N., Stevenson, J., Owaga, C., Marube, E., Bousema, T., Drakeley, C., Cox, J., & Buckee, C. O. (2015). Quantifying travel behavior for infectious disease research: a comparison of data from surveys and mobile phones. *Scientific Reports*, *4*(1), 5678. https://doi.org/10.1038/srep05678

Wesolowski, Amy, Zu Erbach-Schoenberg, E., Tatem, A. J., Lourenço, C., Viboud, C., Charu, V., Eagle, N., Engø-Monsen, K., Qureshi, T., Buckee, C. O., & Metcalf, C. J. E. (2017). Multinational patterns of seasonal asymmetry in human movement influence infectious disease dynamics. *Nature Communications*, *8*(1), 1–9. https://doi.org/10.1038/s41467-017-02064-4

White, M. J., Moreno, L., & Guo, S. (1995). The Interrelation of Fertility and Geographic Mobility in Peru: A Hazards Model Analysis. *International Migration Review*, *29*(2), 492–514. https://doi.org/10.1177/019791839502900207

Whittington, D., Mu, X., & Roche, R. (1990). Calculating the value of time spent collecting water: Some estimates for Ukunda, Kenya. *World Development*, *18*(2), 269–280. https://doi.org/10.1016/0305-750X(90)90052-Y

WHO. (2017). Global vector control response 2017–2030. In *World Health Organization*.

WHO, & DFID. (2006). *The control of neglected zoonotic diseases—a route to poverty alleviation.* World Health Organization.

Wickham, H., Romain, F., Henry, L., Müller, K., & RStudio. (2018). Package “dplyr.” In *R package version 0.8.0.1*.

Bibliography

Wilson, A. L., Courtenay, O., Kelly-Hope, L. A., Scott, T. W., Takken, W., Torr, S. J., & Lindsay, S. W. (2020). The importance of vector control for the control and elimination of vector-borne diseases. In *PLoS Neglected Tropical Diseases* (Vol. 14, Issue 1, pp. 1–31). Public Library of Science. https://doi.org/10.1371/journal.pntd.0007831

Wilson, R., Zu Erbach-Schoenberg, E., Albert, M., Power, D., Tudge, S., Gonzalez, M., Guthrie, S., Chamberlain, H., Brooks, C., Hughes, C., Pitonakova, L., Buckee, C., Lu, X., Wetter, E., Tatem, A., & Bengtsson, L. (2016). Rapid and Near Real-Time Assessments of Population Displacement Using Mobile Phone Data Following Disasters: The 2015 Nepal Earthquake. *PLoS Currents*, *8*. https://doi.org/10.1371/currents.dis.d073fbece328e4c39087bc086d694b5c

Wong, M. L., Chua, T. H., Leong, C. S., Khaw, L. T., Fornace, K., Wan-Sulaiman, W. Y., William, T., Drakeley, C., Ferguson, H. M., & Vythilingam, I. (2015). Seasonal and Spatial Dynamics of the Primary Vector of Plasmodium knowlesi within a Major Transmission Focus in Sabah, Malaysia. *PLoS Neglected Tropical Diseases*, *9*(10). https://doi.org/10.1371/journal.pntd.0004135

Woolhouse, M. E. J., Dye, C., Etard, J. F., Smith, T., Charlwood, J. D., Garnett, G. P., Hagan, P., Hii, J. L. K., Ndhlovu, P. D., Quinnell, R. J., Watts, C. H., Chandiwana, S. K., & Anderson, R. M. (1997). Heterogeneities in the transmission of infectious agents: Implications for the design of control programs. *Proceedings of the National Academy of Sciences of the United States of America*, *94*(1), 338–342. https://doi.org/10.1073/pnas.94.1.338

World Bank. (2012). *The transformational use of information and communication technologies in Africa*. http://siteresources.worldbank.org/EXTINFORMATIONANDCOMMUNICATIONANDTECHNOLOGIES/Resources/282822-1346223280837/MainReport.pdf

World Bank. (2018). CountryProfile | World Development Indicators. In *World Development Indicators database* (p. 1). https://doi.org/10.1371/journal.pone.0122123

World Bank. (2019). *Urban Population statistics - Urban population (% of Total)*. All Countries and Economies Data. https://data.worldbank.org/indicator/SP.URB.TOTL.IN.ZS?locations=ZG%0Ahttps://data.worldbank.org/indicator/SP.URB.TOTL.IN.ZS?view=map%0Ahttps://data.worldbank.org/indicator/SP.URB.TOTL.IN.ZS%0Ahttp://data.worldbank.org/indicator/SP.URB.TOTL.IN.ZS?order=wb

World Health Organization. (1989). *The use of impregnated bednets and other materials for* *vector-borne disease control*.

World Health Organization. (2013). Trypanosomiasis, Human African (sleeping sickness). In *World Health Organization* (p. Fact Sheet 259). https://www.who.int/news-room/fact-sheets/detail/trypanosomiasis-human-african-(sleeping-sickness)

World Health Organization. (2014). A global brief on vector-borne diseases. In *World Health Organization* (Issue September). https://doi.org/WHO/DCO/WHD/2014.1

World Health Organization. (2018a). *Mozambique Malaria Country Profile*. https://www.who.int/malaria/publications/country-profiles/profile_moz_en.pdf?ua=1

World Health Organization. (2018b). *WORLD MALARIA REPORT 2018*. www.who.int/malaria

WorldPop (www.worldpop.org - School of Geography and Environmental Science, University of Southampton; Department of Geography and Geosciences, University of Louisville; Departement de Geographie, Universite de Namur) and Center for International Earth Sc, C. U. (2018). *WorldPop :: Population Density*. https://doi.org/https://dx.doi.org/10.5258/SOTON/WP00645

Wu, L., & Walters, G. (2016). Chinese Travel Behavior in Response to Disastrous Events: The Case of the Japan Earthquake. *Journal of China Tourism Research*, *12*(2), 216–231. https://doi.org/10.1080/19388160.2016.1197870

Wu, P. C., Lay, J. G., Guo, H. R., Lin, C. Y., Lung, S. C., & Su, H. J. (2009). Higher temperature and urbanization affect the spatial patterns of dengue fever transmission in subtropical Taiwan. *Science of the Total Environment*, *407*(7), 2224–2233. https://doi.org/10.1016/j.scitotenv.2008.11.034

Yuan, Y., & Raubal, M. (2012). Extracting dynamic urban mobility patterns from mobile phone data. *Lecture Notes in Computer Science (Including Subseries Lecture Notes in Artificial Intelligence and Lecture Notes in Bioinformatics)*, *7478 LNCS*, 354–367. https://doi.org/10.1007/978-3-642-33024-7_26

Zhou, G., Minakawa, N., Githeko, A., & Yan, G. (2004). Spatial Distribution Patterns of Malaria Vectors and Sample Size Determination in Spatially Heterogeneous Environments: A Case Study in the West Kenyan Highland. *Journal of Medical Entomology*, *41*(6), 1001–1009. https://doi.org/10.1603/0022-2585-41.6.1001

Zu Erbach-Schoenberg, E., Alegana, V. A., Sorichetta, A., Linard, C., Lourenço, C., Ruktanonchai, N. W., Graupe, B., Bird, T. J., Pezzulo, C., Wesolowski, A., & Tatem, A. J. (2016). Dynamic denominators: The impact of seasonally varying population numbers on disease incidence estimates. *Population Health Metrics*, *14*(1), 1–10. https://doi.org/10.1186/s12963-016-0106-0