Predicting disease vector distributions through space and time using environmental and vector control data

by

Emily Acheson

Thesis submitted to the Faculty of Graduate and Postdoctoral Studies
University of Ottawa
in partial fulfillment of the requirements for the M.Sc. Degree in the
Ottawa-Carleton Institute of Biology

Thèse soumise à la Faculté des études supérieures et postdoctorales,
Université d'Ottawa en vue de l'obtention de la maîtrise en sciences de
L'Institut de Biologie d'Ottawa-Carleton

Acknowledgements

I would first like to thank my supervisor, Dr. Jeremy Kerr, for providing the most patient editing of often-painful manuscripts, opportunities and financial support for conferences, generous recommendation letters, an unforgettable trip to Tanzania, and Basil, the lab therapy dog. I also thank my committee members Dr. Antoine Morin and Dr. Scott Mitchell for their advice and guidance over the past two years.

I am grateful to NSERC and the Queen Elizabeth II Graduate Scholarship for funding my research. I would also like to acknowledge the University of Ottawa for providing me with tuition scholarships and other award opportunities.

To Cassie, Alana, Laura, and Bronwyn, you have each selflessly offered help without hesitation the moment I needed it. A special thanks to Juan Zuloaga, who became a second supervisor when R or ArcGIS just would not work (which was all the time) and Rosana Soares for all the hugs and laughs and pizza binge eating.

I thank my loving boyfriend Andrew. You have never stopped pushing me forward and gently reminding me of what matters most while also serving as an example of someone who works harder than anyone else I know.

My family deserves several pages for all the home-cooked food sent up in suitcases, cars, planes and trains, and for the love and encouragement that have gotten me this far. First and foremost, my parents are the reason I am still pursuing this crazy career. Neither of you ever doubted I could do it and I might not still be trying without your constant support and late-night phone calls. I dedicate this thesis to you both.

Table of Contents

Abstract

Within this thesis, I performed a systematic review of approaches to broad-scale modelling of disease vector distributions and determined the most widely used methods predict current species niches and project the models forward under future climate scenarios without temporal validation. I then provided a forward-looking summary of emerging techniques to improve the reliability and transferability of those models, including historical calibration.

I then predicted *Anopheles* mosquito distributions across Tanzania in 2001 (before large-scale ITN distributions) and compared this model with countrywide ITN use by 2012 to assess where the most suitable mosquito habitats were located and whether ITN rollouts in Tanzania ensured coverage of such areas. I concluded that ITNs in Tanzania did not optimally target areas most at risk of malaria. In doing so, I provided a new approach to monitoring and evaluating vector control interventions across large spatial scales.

Résumé

Dans cette thèse, j'ai effectué une revue systématique des approches de modélisation de distribution de vecteurs de maladie à grande échelle et j'ai déterminé que les méthodes de modélisation les plus couramment utilisées prédisent les niches d'espèces actuelles et produisent des modèles prévisionnels selon des scénarios climatiques futurs sans validation temporelle. J'ai ensuite fourni un résumé des nouvelles techniques pour améliorer la fiabilité et la transférabilité de ces modèles, notamment la calibration au moyen de données rétrospectives.

Ensuite, j'ai prédit les distributions des moustiques *Anophèles* à travers la Tanzanie en 2001 (avant la distribution à grande échelle des moustiquaires imprégnées d'insecticide). Ce modèle a ensuite été comparé avec l'utilisation des moustiquaires en 2012 afin de localiser les habitats les plus favorables aux moustiques et d'évaluer si les déploiements de moustiquaires ont atteint une couverture suffisante dans ces zones. J'ai conclu que les moustiquaires n'ont pas ciblé adéquatement les zones les plus à risque de paludisme. Ainsi, j'ai fourni une nouvelle approche pour le suivi et l'évaluation des interventions antivectorielles à grande échelle.

List of Tables

List of Figures

General Introduction

Vector-borne diseases represent an estimated 17% of the global disease burden (Karunamoorthi 2012), making accurate predictions of vector distributions critical to human health applications. Environmental changes, including climate change and land use changes, have already altered arthropod vector distributions, such as mosquitoes, tsetse flies, and ticks (WHO 2014a). With these environmental changes expected to accelerate over the next century due to increasing greenhouse gas emissions (IPCC 2013) and aggravated by land use intensifications in many parts of the world (Van Asselen and Verburg 2013), spatiotemporal modeling of arthropod vectors is also rapidly growing. Not only are forecasts of vector distributions under these environmental changes important to optimize vector control placements in the future, but spatial analyses comparing current placements of controls with predicted vector distributions is imperative to learn whether past and present-day control systems need to be improved.

The expanding fields of “landscape epidemiology” (Pavlovsky 1966; Kitron 1998; Peterson 2006; Reisen 2010), “medical geography” (Meade and Earickson 2000) and “disease macroecology” (Keith et al. 2012) examine the spatial and temporal variations in the biophysical determinants of disease vectors to better understand past, present, and future disease distributions. With climatic changes accelerating, predicting future disease emergence is a particularly urgent area for advancements in forecast modeling. This includes forecasts of malaria through projections of *Anopheles* mosquito distributions (e.g., Fuller et al. 2012a; Fuller et al. 2012b; Drake and Beier 2014), Lyme disease through projections of tick distributions (e.g., Ogden et al. 2008; Simon et al. 2014), and trypanosomiasis through predictions of tsetse fly distributions (e.g., Messina et al. 2012). Epidemiological models therefore follow many of the same underlying principles and processes as those used in biogeography or macroecology (Kerr et al. 2011).

Spatiotemporal models of disease vectors, be they correlative or mechanistic, must learn the ecology of the disease vector (Fischer et al. 2014). Correlative models, such as species distribution models, estimate a species' realized niche over geographical space by learning relationships between the species' georeferenced presence records and environmental predictors (Elith and Leathwick 2009). Similarly, mechanistic models estimate a species' fundamental niche without the need for species occurrence records but instead using knowledge of physiological tolerances of the species under different environmental conditions (Kearney and Porter 2009). These modelling methods can function strictly through space or, using inputs of predicted future environmental scenarios (e.g., using combinations of General Circulation Models with emissions scenarios created by the Intergovernmental Panel on Climate Change), they can be projected into the future.

Anopheles mosquitoes are an excellent example of insect vectors that are strongly influenced by their environmental conditions and, due to these influences, alter the pathogens and parasites they transmit (e.g. malaria protozoa, lymphatic filariasis worms, and multiple viruses) (Afrane et al. 2012). Only female mosquitoes feed on blood, and in doing so, transmit the pathogens or parasites into the host's blood stream (Clements 2012). Average global temperatures are expected to rise 2.6-4.8°C by 2100 (IPCC 2013), potentially increasing mosquito biting rates, reproduction rates, mosquito larvae development rates, and parasite survival and development within mosquitoes (Blanford et al. 2013) as well as increasing their spread into higher altitudes (Siraj et al. 2014). Areas with increasing rainfall promotes more mosquito breeding grounds (Blanford et al. 2013), and anthropogenic environmental changes such as urbanization, agricultural development, and deforestation also have critical impacts on habitat availability and microclimatic conditions (Afrane et al. 2012). *Anopheles* mosquito species are also generally difficult to control. Their ~460 species worldwide occupy a variety of niches that allow them to inhabit and invade a wealth of environments (Afrane et al. 2012). Many of the 30-40 *Anopheles* species that carry malaria parasites are growing increasingly resilient to insecticides (e.g., Protopopoff et al. 2013; Matowo et al. 2014; Nkya et al. 2014).

Disease vector control is an example of the phrase “prevention is better than cure,” yet the most recent report from the World Health Organization outlines an alarming lack of infrastructure and skills around the world to improve efficiency, monitor impact and maintain coverage and surveillance of vector controls (WHO 2014a). Increased accuracy of statistical models forecasting disease vector distributions can improve the prevention of disease transmission from vectors to humans by predicting where vectors will be in the coming years. Vector control efforts rely on these analyses to target human populations that will be most at risk of disease transmission and must be strengthened to decrease current disease rates and manage any short-term climate risks to later increase resilience to long-term climate change effects (Campbell-Lendrum et al. 2015).

Studies of antimalarial controls, such as insecticide-treated mosquito nets (ITNs), have already begun at local, countrywide, and even near continent-wide scales, but predominantly from socioeconomic perspectives. Distributions of ITNs in sub-Saharan Africa have been found to vary spatially, with almost random distributions and showing no relationship to malaria endemicity (Vanderelst and Speybroeck 2013). These variations are likely a complex combination of socioeconomic, ecological, and cultural factors (Graves et al. 2011). Although malaria cases and deaths have decreased in recent years, *Plasmodium falciparum* malaria – the most dangerous parasitic form of malaria – continues to infect an estimated 200 million people and cause over 550 000 deaths globally every year (WHO 2014b). Approximately 90% of all malaria deaths occur in Africa, with the main victims being children under the age of five and pregnant women (WHO 2014b).

Unfortunately, spatiotemporal models of arthropod disease vector distributions carry many limitations. Most forecast modelling techniques validate their models through space but not through time (Acheson and Kerr 2015). These forecast models therefore work under the assumption that spatial accuracy is equivalent to temporal reliability, often referred to as the space-for-time assumption (Blois et al. 2013). This assumption is often false, however, if a species' observed spatial relationship to environmental characteristics differs from its temporal relationship to these factors (Kerr et al. 2011). To correct for this assumption, forecast analyses should perform historical calibration (Kharouba et al. 2009), which employs a multi-temporal design to validate models through a time period that has already passed before projecting into the future. Yet this approach can also be challenging. Georeferenced disease vector occurrence records are generally rare or insufficient for statistical modelling, and more substantial collections largely contain data lacking systematic sampling across space and through time and suffering from sampling bias (e.g., favouring areas near roadsides, villages, etc.). This lack of systematic, large-scale, continuous and unbiased sampling continues to limit validation of vector forecast analyses (e.g., Martens et al. 1999; Levine et al. 2004; Estrada-Pena and Venzal 2007; Gonzalez et al. 2010; Tonnang et al. 2010; Messina et al. 2012). However, these issues are beginning to be addressed. An increasing number of sources provide large assemblages of freely-accessible occurrence records for vectors, like *Anopheles* mosquitoes, through online sources and primary research publications.

The absence of prolonged, systematically-sampled species observations often limits vector distribution models to purely spatial (i.e. static) ones. These models are no less useful than forecast models, however, and can provide insights into species' responses to spatial environmental gradients while allowing for comparatively more feasible validation. While there is no lack of spatial models of mosquitoes (e.g. Kulkarni et al. 2010), ticks (e.g., Leighton et al. 2012), and tsetse flies (e.g., Matawa et al. 2013), the incorporation of vector control measures into these models has been encouraged (Kelly-Hope et al. 2009; Luz et al. 2010) but barely explored.

This thesis consists of two chapters, each of which addresses a question regarding spatiotemporal modelling of arthropod disease vectors. The first chapter focuses on the temporal aspects of statistical forecast models while the second chapter focuses purely on spatial modelling methods.

Chapter One performs a systematic review of the current methods used in statistical models to forecast arthropod disease vectors under future environmental scenarios, and whether these models are validated temporally. It reviews the forecast and validation methods and explains the benefits or drawbacks of each. Historical calibration is then introduced as a solution to forecast model validation, detailing the multi-temporal design required to validate models through a time period that has already passed before projecting into the future (Kharouba et al. 2009).

Chapter Two questions whether insecticide-treated mosquito net distributions in the United Republic of Tanzania (hereafter referred to as Tanzania) have so far adequately targeted prime habitats for *Anopheles* mosquitoes, the vector of malaria parasites. I hypothesize that if the distribution of ITNs has been optimized, areas of high mosquito habitat suitability before mass distributions of mosquito nets would be followed by increased use of ITNs throughout these areas and thus lower variance (coverage rates should be consistently high in malaria-endemic areas) in ITN use. I combine georeferenced records of *Anopheles* mosquito presences before mosquito net rollouts in 2004 with environmental predictors and use ecological niche modelling to create a layer of relative habitat suitability for these mosquitoes. I then create separate layers depicting mosquito net use distributions in Tanzania as of the most recent surveys from 2011-2012. By comparing the intensity of mosquito net use across Tanzania with the corresponding predicted mosquito habitat suitability in those surveyed areas, I determine how well mosquito net use correlates with mosquito habitats.

The goal of this thesis is to attempt to improve on current forecasting approaches used in disease vector modelling. In addition, it puts forward an interdisciplinary approach of incorporating disease vector control data into ecological niche models. To achieve these goals, a macroecological approach was used. In regards to Chapter Two, this thesis does not attempt to explain why spatial variations may exist in antimalarial interventions; these reasons extend beyond purely ecological ones, such as social inequalities, life style choices, and cultural beliefs (Graves et al. 2011). Regardless of these reasons, the analyses presented in this thesis examine patterns at larger spatial and temporal scales, attempting to target signals from the large variations inherent not only in environmental predictors but also in the socioeconomic factors to assess mosquito net use at a countrywide scale.

Chapter 1:
Looking forward by looking back: Using historical calibration to improve forecasts of human disease vector distributions

Please note: This paper has been published in the peer-reviewed journal Vector-Borne and Zoonotic Diseases, *with the following citation:*

Acheson, E. S. and J. T. Kerr. "Looking Forward by Looking Back: Using Historical Calibration to Improve Forecasts of Human Disease Vector Distributions." *Vector-Borne and Zoonotic Diseases* 15, no. 3 (2015): 173-183.

Abstract

Arthropod disease vectors, most notably mosquitoes, ticks, tsetse flies, and sandflies, are strongly influenced by environmental conditions and responsible for the vast majority of global vector-borne human diseases. The most widely used statistical models to predict future vector distributions model species niches and project the models forward under future climate scenarios. Though these methods address variations in vector distributions through space, their capacity to predict changing distributions through time is far less certain. Here, we review modelling methods used to (1) validate and (2) forecast future distributions of arthropod vectors under the effects of climate change and outline the uses or limitations of these techniques. We then suggest a validation approach specific to temporal extrapolation models that is gaining momentum in macroecological modelling and has great potential for epidemiological modeling of disease vectors. We performed systematic searches in Web of Science, ScienceDirect and Google Scholar to identify peer-reviewed English journal articles that model arthropod disease vector distributions under future environment scenarios. We included studies published up to and including June 2014. We identified 29 relevant articles for our review. The majority of these studies predicted current species niches and projected the models forward under future climate scenarios without temporal validation. Historically-

calibrated forecast models improve predictions of changing vector distributions by tracking known shifts through recently observed time periods. With accelerating climate change, accurate predictions of shifts in disease vectors are crucial to target vector control interventions where needs are greatest.

Introduction

Rising concerns over environmental changes and their potential effects on disease prevalence require more extensive exploration of statistical modelling methods that can accurately predict future disease vector distributions (Mills et al. 2010; Tabachnick 2010). Satellite-based remote sensing and statistical geographic information tools continue to evolve at an extraordinary pace, providing continuous measurements of environmental factors at progressively finer spatial and temporal resolutions (Hay et al. 2000). As a result, increasingly accurate spatial models are being created using present-day data and projected into future time periods to predict disease vector distributions under climate change scenarios (e.g., Tonnang et al. 2010). While many of these models appear accurate across space, their predictions are often highly divergent through time (Araujo et al. 2005b; Pearson et al. 2006; Kharouba et al. 2009; Braunisch et al. 2013). The “temporal transferability” of these models – their ability to predict independent events through time correctly – is highly uncertain (Araujo and Guisan 2006).

The failure of the space-for-time assumption (Blois et al. 2013) imposes strong limits on the reliability of spatial models for temporal prediction. For example, nine modelling techniques used to forecast four South African plant species’ distributions in 2030 predicted between a 92% loss to a 322% gain, but all models successfully described the current spatial distributions of these species (Pearson et al. 2006). Furthermore, environmental factors that correlate positively with species' distributions spatially may show the opposite relationship temporally. For example, models of Canadian butterfly species' richness showed a positive association with human population density spatially but a negative association temporally (White and Kerr 2006). Species distribution modeling methods range from logistic regression to highly specialized machine-learning techniques (Elith and Leathwick 2009), all of which have proven useful under different circumstances. Strong tests of model accuracy are critical to forecasting how changing environmental circumstances will alter the distributions of disease vectors through time.

Accurately projecting distributions of human disease vectors, such as those transmitting malaria, Lyme disease, dengue, and trypanosomiasis, is crucial because of their human health impacts and the related economic consequences of misdirected management interventions. Yet, predictions of future vector distributions often conflict. For example, models of *Anopheles* mosquito distributions, the vectors of the malaria parasite, across the Caribbean did not predict northern range expansions into the Gulf coastal region under warming climates by 2080, contradicting previous studies which predicted future expansions of malaria by as early as 2050 (Fuller et al. 2012a). Conflicting results of future vector distributions mislead control programs and practices that already pose substantial financial challenges for many developing countries (Martens et al. 1999; WHO 2008). Technological advances are unlikely to permit predictions of future species distributions to be tested unambiguously; the definitive, but impractical, method for testing such models is to wait to compare predictions to actual events (Rastetter 1996; Algar et al. 2009). This does not mean forecasting future species distributions, and validating these methods, lacks rigour. It does, however, underscore the need for strong validation methods to increase confidence that such models are reliable (Botkin et al. 2007).

There are extensive parallels between macroecological models, which predict species distributions over broad areas (e.g., regions and continents, see Brown and Maurer 1989; Kerr et al. 2007) and epidemiological models needed to predict changing vector distributions across comparable modelled areas. Both may use correlative and mechanistic methods for projecting changing species distributions (for a comparison of these two methods in vector modelling under future climates, see Fischer et al. 2014). Correlative models can include general linear models or use machine-learning techniques, such as maxent (Phillips et al. 2006), techniques from the Biomod modelling package, such as random forest analyses (Thuiller 2003), and genetic algorithms for rule-set production (Stockwell and Peters 1999) to relate species occurrence records to environmental factors and determine the area the species actually occupies (i.e., realized niche). Mechanistic models, or stochastic models coupled with habitat suitability or population models, use lab- and field-based measures of physiological traits (e.g., behaviour, temperature tolerances, energy turnover) to determine all potential areas where the species may persist (i.e., fundamental niche) (Kearney and Porter 2009). While temporal macroecological models often seek to forecast future species distributions for the purposes of conservation (Trejo et al. 2011), invasive species (Sobek-Swant et al. 2012), or agriculture (Ureta et al. 2012), their forecasting and validation methods show considerable overlap with those predicting changing disease risk. Satellite remote sensing tools provide a common array of baseline measurements for both fields that are revolutionizing statistical modelling in epidemiology (Hay et al. 2000) and macroecology (Kerr and Ostrovsky 2003) by improving detection of spatial and temporal variation in the biophysical determinants of species distributions. Consequent syntheses of techniques from both disciplines has led to the emergence of "eco-epidemiology," "landscape epidemiology" (Pavlovsky 1966; Kitron 1998; Peterson 2006; Reisen 2010), "medical geography" (Meade and Earickson 2000), and "disease macroecology" (Keith et al. 2012).

In attempting to predict changes in species distributions through time, predictive models commonly share a straightforward approach. Generally, these models 1) use current environmental conditions to predict current disease vector distributions, 2) test the accuracy of the model by evaluating its fit to current spatial observations, and 3) project the model to a future time period by substituting anticipated environmental observations into the spatial model. Spatial validation, which improves confidence in model fit in the current time period, does not imply temporal reliability. To ensure the latter, analyses must be multi-temporal (Kharouba et al. 2009). A validation technique that accounts for time (or historical calibration) is needed. Using this approach, models are constructed with observations from different time periods. These models are then projected either backwards or forwards into a time period with independent vector data to test their capacity to predict changing distributions through time. Historical calibration is likely the minimum requirement to establish confidence that these models can be used to forecast changes in vector distributions accurately (e.g., Kharouba et al. 2009; Rapacciuolo et al. 2014). This approach provides spatiotemporal tests of model reliability, rather than spatial tests followed by risky reliance on space-for-time assumptions.

Here, we review the critical field of forecasting future disease vector distributions under changing climate. First, we discuss the different methods used to validate the spatial models of present-day vector distributions. Next, we assess how these spatial models are projected into future time periods under different climatic scenarios. We then introduce the framework we call historical calibration for validation that is specific to temporal extrapolations in macroecology and may increase accuracy of forecasting models in epidemiology. While we focused on disease vectors that directly affect human health for this review, the assessment of the forecast and validation approaches can be applied across spatial scales and broadly among many focal species and hosts, such as livestock (e.g., for advances in forecast modelling of *Culicoides* midges and the bluetongue virus in livestock, see Acevedo et al. 2010 and Guichard et al. 2014).

Methods

We performed systematic searches in three databases for peer-reviewed literature: Web of Science (http://thomsonreuters.com/thomson-reuters-web-of-science), ScienceDirect (http://www.sciencedirect.com/) and Google Scholar (http://scholar.google.com). We restricted the search to English journal papers. We included studies published up to and including June 2014. Each search varied depending on the search engine input options, but combined the following three sets of inputs: the first set used “either/or” for the projection methods “forecast*,” “climate scenario*,” “projection”, or “future”; the second set used “either/or” for the model types “correlative”, “mechanistic” or “model”; the third set used “either/or” for the vector species “mosquito*,” “tsetse*,” “tick*,” or “sandfly*.” While the term “forecast” generally refers to predicting forward through time, relative to a baseline period (e.g., using historical data to predict present-day conditions) (Elith and Leathwick 2009), this review only examines papers forecasting future events. The initial searches yielded a total of 1082 papers. Studies were included if they predicted future vector distributions under changing climate or changing non-climatic factors (e.g., land use). We therefore excluded studies that modelled disease cases instead of the arthropod vector itself (e.g., Naish et al. 2013). We included studies using future climate scenarios (e.g., general circulation models, or GCMs) and environmental changes proposed by the study (e.g., incremental increases in temperature). Both correlative and mechanistic models were included. Since the focus of this review is forecasting arthropod vectors that transmit diseases to humans, we excluded models of vectors that solely attack non-humans, such as livestock. We also excluded non-arthropod vectors, such as rodents. Based on these inclusion and exclusion criteria, only 29 studies remained relevant as primary literature in the final review.

Results and Discussion

Validating spatial models

First, it is important to distinguish between model fit – whether a model's predictions reasonably match data available within an observed time period – and true model validation – determining whether the parameters of a model are causally linked with the biological response, which it predicts accurately. The former is the domain of approaches such as the "split data" approach. The latter truly tests model predictions, rather than model fit. However, true model validation is more challenging because it requires independent disease vector data to test models' generality or temporal transferability (i.e. data not used in model training and that are spatially or temporally independent). All too frequently, such data do not exist. While the validation methods reviewed were limited by inevitable autocorrelation issues or data availability, and occasionally were not described at all, these 29 studies generally included strong validation techniques, which are categorized and described below.

The most common test of model fit we observed in these studies is a variation of the "split-data" approach (see Table 1.1). This approach is a subset of cross-validation, which divides datasets into two parts: training data used to calibrate the model and testing data to evaluate its predictive accuracy. The "split-data" approach randomly splits the original data into training and testing data (e.g., 70% of occurrence data are used for training the model, and the remaining 30% are used for testing). While the "split-data" approach is a cross-validation method, it tests model fit, not the model's temporal transferability.

The "split-data" approach is limited by autocorrelation between calibration and validation data sets (Araujo et al. 2005a). Modelling methods assume the modelled occurrences are independent and thus suitable for model validation, but this is not the case when the data used to calibrate and validate the model are collected from the same modelled region. For example, the spatial models of *Anopheles* mosquitoes in the Caribbean were validated by randomly selecting 36 points from the original 350-point data set, which were then treated as "independent test points" (Fuller et al. 2012a). Yet, randomly selected subsets of the original data set do not become independent because the environmental determinants of species distributions are strongly autocorrelated with the data used for model calibration. This issue was shown when model validation using independent vs. non-independent data found the latter consistently resulted in higher estimations of model accuracy (Araujo et al. 2005a). The use of independently collected vector test data from the same modelled area in different time periods (e.g., Ogden et al. 2008) does not correct this issue but only introduces the risk of temporal autocorrelation due to possible lack of independence between data points adjacent through time (Araujo et al. 2005a; Rapacciuolo et al. 2014). Species distribution models may therefore sometimes perform no better than random but predict spatial distributions accurately purely because of underlying spatial or spatiotemporal autocorrelation in predictor variables (Bahn and McGill 2013). It then becomes possible to fit environmental factors to almost any spatially autocorrelated species distribution, making it difficult to determine whether a distribution arises because the vector responds biologically to that factor or because both the species distribution and environmental factors share a similar autocorrelative structure. Species distributions, including those of disease vectors like mosquitoes, depend strongly on environmental factors that are highly autocorrelated both spatially and temporally (e.g., Duncombe et al. 2013). As a consequence, any other variable that shows similar underlying spatial autocorrelative structure may appear, spuriously, as a causal variable in a model. Validation efforts that limit spatial and temporal autocorrelation by drawing model testing data from different areas and/or times are more likely to be effective (Araujo et al. 2005a; Botkin et al. 2007). Developing statistical techniques to measure spatial autocorrelation explicitly can also increase confidence that the model is predicting accurately and capturing critical biological processes (Kerr et al. 2011). Unravelling the relative roles of purely spatial variables, purely environmental causes, and the overlapping contributions of space and time may require additional statistical analyses, such as variance decomposition (Legendre and Legendre 1998). Moran’s I or Geary’s c are also useful to evaluate the strength of spatial autocorrelation in datasets used in these models (see Dormann et al. 2007).

In the absence of independent data, many studies we assembled were forced to rely on AUC (area under the receiver operating characteristic) as a surrogate for model validation, with models considered validated if they received high AUC values for their training and testing data (e.g., Moo-Llanes et al. 2013; Garza et al. 2014; Gonzalez et al. 2014). However, AUC tests model fit and does not perform true model validation. AUC’s measure of a statistical model’s fit is derived from presence-absence species data, and this method can overestimate model accuracy (for details on these problems, see Lobo et al. 2008). The true skill statistic (TSS), however, is threshold-dependent (i.e., uses continuous probabilities of presence) and appears to provide more reliable tests of model fit than AUC (Allouche et al. 2006). Ideally, multiple test statistics should be used to assess model performance (Elith and Leathwick 2009); methods such as AUC and TSS should be used to assess model fit while model tests of spatial and temporal transferability using independent observations should be used to test those models more strongly.

A second method of validation is the use of independent data (i.e., data not used for model training or tests of model fit) and is the ideal validation approach for predictive models to test the model's transferability to independent events (Araujo and Guisan 2006; Botkin et al. 2007). Independent data sets can come from a geographically distinct area (Fielding and Haworth 1995; Randin et al. 2006), or from a different time period (Araujo et al. 2005a). For instance, *Anopheles* mosquito occurrence records collected between 1930 and 1940 from northeastern Brazil were used to validate present-day spatial models of these mosquito species in sub-Saharan Africa (Tonnang et al. 2010). Yet, correlative disease vector forecast models rarely use temporally independent data for model validation (e.g., Brownstein et al. 2005), usually due to data scarcity.

We found that studies using mechanistic models to predict changing disease vector distributions were more likely to employ independent field observations for validation, with six of the nine mechanistic model studies using this approach (Table 1.1). Studies employing correlative models (e.g., Beebe et al. 2009; Drake and Beier 2014) did so more rarely. Another variation of independent validation involved comparison of the study's spatial models with predicted suitability maps from other sources from that same time period. For example, the Tsetse Ecological Distribution model (see Figure 1.1), used to predict the sleeping sickness vector *Glossina morsitans* in Kenya, was validated by comparing it to a combination of tsetse fly habitat suitability maps from other sources (DeVisser et al. 2010; Messina et al. 2012). However, agreement between the model's output and available tsetse habitat suitability maps is low, likely due to variation in how each map was constructed and what each map represents (e.g., maps of historical ranges, species distribution models, etc.).

Many of the studies we reviewed claimed insufficient data for validation (e.g., Martens et al. 1999; Levine et al. 2004; Estrada-Pena and Venzal 2007; Gonzalez et al. 2010; Tonnang et al. 2010; Messina et al. 2012). While mechanistic modelling of disease vectors is one way to counter this problem, independent occurrence data are still needed to validate the models. Paleoecological observations have value for model validation for some species (Martinez-Meyer and Peterson 2006; Willis and Birks 2006; Roberts and Hamann 2012), but their potential for arthropod vectors is limited because some specimens are rarely preserved and because environmental conditions at the time of their preservation can be too uncertain to provide insight into present-day environmental dependence (Borkent and Grimaldi 2004). This also reveals the added challenge of modelling most emerging vector species (e.g., mosquitoes carrying West Nile Virus; e.g., Yoo 2014), especially with the historical calibration approach: historical records will be largely lacking for places of concern. For these cases, historical calibration might not be possible until sufficient vector records have been collected.

A third validation approach, the comparison of occurrence data of arthropod vectors with the locations of infected reservoir species, holds significant promise (Table 1.1). For example, models of phlebotomine sand flies were validated using disease case data on leishmaniasis, for which dogs are a reservoir (Galvez et al. 2011). Leishmaniasis case data were collected from veterinary clinics scattered across 32 villages in Madrid. The number of dogs infected with leishmaniasis, and the locations of the clinic in which they were diagnosed, were recorded. Their locations were then geocoded into a map layer and compared with the map of phlebotomine sand fly vectors and disease presence and predicted phlebotomine distributions were found to be strongly related.

While the presence of a disease suggests the nearby presence of the vector that transmits it, such as leishmaniasis in dogs carried by the phlebotomine fly, areas where the disease is detectable are commonly subsets of a vector's geographical distributions (Kulkarni et al. 2010). Furthermore, disease cases may be aggregated around the health centres where they were diagnosed and recorded, offering biased measurements of the geographical distribution of disease incidence and, consequently, vector distributions. For example, the Human African Trypanosomiasis (HAT) vector *Glossina morsitans* still inhabits many regions of Tanzania, but HAT is now mainly found in the western and northern regions of the country (Matemba et al. 2010; Malele 2011).

Projecting vector distributions into the future

Studies we reviewed projected future changes in disease vector distributions by substituting future environmental changes into their current spatial models in one of three ways: using general circulation models (GCMs), adding a constant increment of environmental change uniformly throughout the region of interest, or extrapolating recent environmental trends linearly into the future (see Table 1.1). First, twenty-four of the 29 papers we reviewed used GCMs to model climate on the basis of its underlying physical processes and used these models to predict geographical variation in future climates under different emissions scenarios. For instance, Lyme disease risk in North America was predicted by modelling *Ixodes scapularis* tick distributions with current climate data and tick occurrence records and then projecting the current spatial model to 2020, 2050, and 2080 using GCMs in combination with emissions scenarios (Brownstein et al. 2005). While GCMs are believed to remain the most advanced tools available for future projections of climate change (Randall et al. 2007), their outputs still fail to address changes in biological processes through time when the GCMs are projected from only one spatial model.

Second, the use of constant environmental increments set by the study, such as the input of 0.1, 2, or 4°C increases in temperature combined with 10% increases or decreases in seasonal rainfall across Africa to predict future *Anopheles* mosquito distributions (Tonnang et al. 2010), assumes the spatial environmental patterns will remain constant through the modelled period. Retaining constant spatial trends in environmental factors using this approach will lead to discontinuities in predicted vector distributions if modelled environmental differences exceed thresholds of tolerance among vector species in some areas. For this reason, detailed knowledge of the vector's ecology is imperative to these modelling approaches. Information on temperature thresholds for survival are gathered through field observations, such as for the Asian tiger mosquito, *Aedes albopictus* (Fischer et al. 2014). Upper and lower critical temperatures can also be tested in laboratory settings and then incorporated in future climate scenarios. However, results for survival thresholds can vary substantially between field and laboratory conditions. Since ecological complexity, such as short-term temperature fluctuations, can affect results of survival thresholds in the field, both field and laboratory results should be taken into account when modelling geographic limitations (e.g., Brady et al. 2013). While such measurements are likely to correlate with species' realized thermal niche limits beyond laboratory confines, species' capacities for behavioural thermoregulation are variables, among species and among environments (Sunday et al. 2014).

Third, extrapolating from recent environmental trends by adding a constant change increment (e.g., +1 °C) is an improvement to setting theoretical increments because its future projections are grounded on observed trends through time. For example, change in irrigated cropland in Sudan and Upper Egypt was first calculated between 2001 and 2009 and then extrapolated to 2050 to predict future *Anopheles arabiensis* mosquito distributions (Fuller et al. 2012b). However, this method assumes these trends will only continue linearly.

Determining which projection method is most realistic will remain uncertain because the future is uncertain, but ensemble forecasting is a promising solution to increasing forecast robustness (Araujo and New 2007). Multiple model outputs are analysed for their range of projections. Forecast models derived from different future scenarios are combined using consensus methods to produce outputs and are predicted to be more accurate than forecasts from single models (Forester et al. 2013). Ensemble forecasts can and are already being used in epidemiological forecast models. For instance, projections of *Ae. albopictus* in Europe used 10 regional climate models and achieved climate projections consistent with the IPCC projections and future hot spots consistent with other published studies (Caminade et al. 2012). But regardless of the method used to project future environmental change, a forecast model will lack credibility unless the magnitude of species' temporal response to environmental change can be approximated.

Each of the projection modelling frameworks we reviewed requires the space-for-time assumption (Pearson et al. 2006; Fisher et al. 2010). That is, observed spatial environmental gradients within the species' current range are assumed to be identical to how it would respond to the temporal environmental changes of similar magnitude. Models are then extrapolated to spatial and temporal domains different from those the models were trained in (Dobrowski et al. 2011; Kerr et al. 2011). However, spatially accurate models may not be temporally transferrable. Slight differences in the current distributions of species predicted by different models become magnified when those distributions are then projected into future climate scenarios, resulting in a variety of predicted range sizes for the same species and region and increasing projection uncertainty (Kharouba et al. 2009).

If species distributions are at or near equilibrium with environmental conditions (i.e., the current areas where the population is found represent the full possible extent of the population's distribution), it is more likely that models will correctly predict their distributions through time (Botkin et al. 2007; Broennimann and Guisan 2008; Elith and Leathwick 2009; Araujo and Peterson 2012; Pagel and Schurr 2012; Sax et al. 2013; Wenger et al. 2013). Yet, disease vectors are often likened to invasive species in their ability to spread rapidly into new environments through human transport and changes in land use and climate and equilibrium assumptions are likely to be unreliable (Crowl et al. 2008; Fuller et al. 2012b). The equilibrium assumption may mislead predictions over broad areas, depending on species' traits (e.g., dispersal), so species' capacities to colonize previously unoccupied areas may affect reliability of model predictions (Araujo and Pearson 2005; Leroux et al. 2013).

Understanding dispersal limitations can be crucial. For example, correlative modelling of the potential spread of *Ae. albopictus* to higher altitudes in the Alps was assessed by relating presences and absences of the mosquito to high resolution (200-m) land surface temperature data (Roiz et al. 2011). A temperature-gradient-based model was created to fill the no-data areas and, combined with threshold conditions governing egg and adult overwintering survival, the models agreed with previous studies' results of annual and January mean temperature as the most crucial to mosquito dispersal. The authors did not discuss their validation method but do admit that model validation with empirical observations is needed and that they will validate with independent field data as they conduct new vector collections in the coming years. Mechanistic modelling of *Glossina morsitans* tsetse flies in Kenya, for example, considers both environmental factors and fly dispersal abilities (Messina et al. 2012). The modelling framework accounts for fly dispersal by setting a “fly front” (i.e., tsetse population movement as a whole, instead of individual tsetse movement rates), where interactions between tsetse distributions and predicted environmentally-suitable areas would be considered tsetse presence. The end result consists of 250-m spatial and 16-day temporal resolution presence-absence tsetse maps (see Figure 1.1), validated with tsetse habitat suitability maps from other sources. A cost analysis estimates that if such a predictive modelling system is implemented it will, over an eight-year tsetse management campaign, save Kenya’s control programs over $19 million or approximately 58% of normal tsetse control expenditures (McCord et al. 2012).

Historical calibration

A validation method in macroecology that relies on historical calibration of model predictions (Kharouba et al. 2009) is achieving general acceptance in conservation and global change research (Rapacciuolo et al. 2014; described in Figure 1.2). We found no studies that have yet used this framework for predicting changing vector distributions through time, so we describe its utility by drawing on examples from macroecological and global change research. The key benefit of this framework is to evaluate the reliability of the space-for-time assumption and provide a direct test of whether spatial models are temporally transferrable in the context of a natural, or pseudo-, experiment (Kerr et al. 2007).

Historically-calibrated distribution models employ the usual tests of model fit, based on cross-validation and statistical accuracy assessment (e.g., AUC and TSS). However, predictions of how a species' distribution will change through time are tested through a period during which such changes can be directly observed. For example, butterfly distributions across Canada were examined in a historical period (1900-1930; Kharouba et al. 2009). Each of these species' range responses to subsequent climate change was tested by substituting observed climate measures from 1960-1990 into this historical model and predicting how their ranges should have changed if historical relationships between climate and species' occurrences remained constant. Predictions for the latter time period (1960-1990) could then be tested in two ways. The first involves recalculating model accuracy statistics based on new occurrence points from that latter time period. The second involves constructing new models from latter observations for comparison with historical model predictions. Predictions of butterfly species distribution changes over the 20th century across Canada were strongly correlated with observed changes in those distributions. These studies assume that species do not adapt to novel environmental conditions through the observed or anticipated time periods. While that assumption appeared reasonable for butterflies in Canada, which often have one generation per year, uncertainties regarding disease vectors' potential to adapt to new environments (particularly since many breed extremely rapidly in far less seasonal environments than Canada) are substantial (Fischer et al. 2014).

The use of multi-temporal species observations and environmental data reduces the uncertainty imposed by the space-for-time assumption. Minimally, using historical data to test (or validate) model predictions through time will reveal clear violations of the assumption, such as instances when models appear spatially accurate but predict nonsensically through time (e.g., Pearson et al. 2006). Yet, predicting future conditions will always require this assumption: changes observed recently will continue to reflect changes anticipated through the future period that is being modelled. Both forecasting (Figure 1.2a) and backcasting (or hindcasting; Figure 1.2b) can provide some measure of historical calibration for model predictions. These historical calibration methods are finding increased application, especially in conjunction with advancing measures of species' dispersal capacities, population growth, and even species' particular behavioural thermoregulatory strategies under continuously varying environmental changes (e.g., Monahan and Tingley 2012; Sunday et al. 2012; Leroux et al. 2013; Sunday et al. 2014). Opportunities to use this framework to gain biological insights into vector responses to global change appear substantial, particularly given the increasing availability of historical species observations.

Advanced modelling techniques, rapidly evolving satellite remote sensing tools, and increased vector surveillance provide the tools necessary for building historically-calibrated vector models. For instance, forecast models of *Anopheles* mosquitoes are advancing with both mechanistic (e.g., Yamana and Eltahir 2013) and correlative (e.g., Tonnang et al. 2010) approaches at macroscales. Satellite-borne remote sensing data sources, such as the Moderate Resolution Imaging Spectroradiometer (MODIS; https://lpdaac.usgs.gov/products/modis_products_table/modis_overview) are increasingly accessible and provide high resolution (up to 250m for MODIS), near-real time measurements of both climate and land use changes over very broad areas (see Figure 1.3). Occurrence records for vectors, like *Anopheles* mosquitoes, are also increasingly accessible through online sources, such as the Global Biodiversity Information Facility (http://www.gbif.org/), Malaria Atlas Project (http://www.map.ox.ac.uk/) and IRMapper (http://www.irmapper.com/), and primary research publications also provide large assemblages of these observations. Nevertheless, occurrences records for many species remain difficult to find (noted in several studies here: Estrada-Pena and Venzal 2007; Gonzalez et al. 2010; Messina et al. 2012) and predicting changes in associated disease risk will remain proportionately more difficult (and interventions harder to implement) while this problem persists.

While historical calibration of species distribution models does not eliminate shortcomings intrinsic to predicting an uncertain future, it reduces or resolves more challenges than it creates. This historical calibration and testing framework reduces the risk that spatial autocorrelation may create misleading fits between environmental predictors and species distributions, but it cannot eliminate potential impacts of autocorrelation, which exists temporally as well as spatially. Yet, a key benefit is a direct test of the temporal transferability of a species distribution model that might otherwise provide only a snapshot of a species' reliance on environmental conditions over broad areas. When accompanied by more reliable projection methods, like ensemble forecasting, historically-calibrated forecast models increasingly address the current dilemma of determining predictive model accuracy over areas where vector range changes have actually been tracked through time.

Conclusions

Validation through historical calibration is increasingly used for macroecological and global change research and could be adapted readily for epidemiological models forecasting disease vector distributions. Tests of model transferability, used with other statistical tests of model fit, will increase confidence in model predictions both spatially and temporally. Availability of reliable species observations over broad areas and at different times is a significant impediment here, as it has been for macroecological and global change research. Increased and consistent disease vector occurrence data collections are urgently needed. Forecasting disease vector distributions is further complicated by inadvertent or careless introductions of vectors into new areas (Sutherst 2004; Fischer et al. 2014). For example, chikungunya was recently found in St. Maarten in December 2013, the first recorded instance of local transmission of the disease in the Western Hemisphere (CDC 2014). Despite continuous improvements in techniques that assess how vector species' distributions and abundances depend on environmental conditions, predicting their future distributions will remain a challenging endeavor. Nevertheless, interdisciplinary opportunities to improve biological understanding of how and when vector species will respond to environmental changes could improve prospects for successful interventions to reduce their impacts on human health. Historical validation is one such opportunity.

Tables

Table 1.1. Different approaches to validation of forecast models for arthropod disease vectors with environmental change.

	Validation approach				Forecast approach		
Reference	Data Splitting	Independent Validation	Validation with disease cases	Validation approach not specified	General circulation model	Environmental changes proposed by study	Project quantity of change in environ-mental variable
Correlative models							
Beebe et al. 2009	X				X		
Brownstein et al. 2005		X			X		
Cross and Hyams 1996				X		X	
Drake and Beier 2014	X				X		
Estrada-Pena and Venzal 2007	X					X	
Feria-Arroyo et al. 2014	X				X		
Fischer et al. 2011	X				X		
Fuller et al. 2012a	X				X		
Fuller et al. 2012b	X						X
Galvez et al. 2011			X		X		
Garza et al. 2014	X				X		
Gonzalez et al. 2010	X				X		
Gonzalez et al. 2014	X				X		
Hongoh et al. 2012	X				X		
Levine et al. 2004	X				X		
Moo-Llanes et al. 2013	X				X		
Peterson and Shaw 2003	X				X		
Roiz et al. 2011				X	X		
Schafer and Lundstrom 2009				X	X		
Tonnang et al. 2010		X				X	
Mechanistic models							

Caminade et al. 2012	X				X	
Erickson et al. 2012		X			X	
Kolivras 2010			X		X	
Medlock et al. 2006				X		X
Messina et al. 2012		X			X	
Morin and Comrie 2010		X			X	
Ogden et al. 2006		X			X	
Ogden et al. 2008		X			X	
Yamana and Eltahir 2013		X			X	

Figures

Figure 1.1. The African Trypanosomiasis Forecasting System (ATcast) modelling outline. Incorporated in this modelling framework is the Tsetse Ecological Distribution (TED) Model, which accounts for spatio-temporal dispersal dynamics of tsetse fly distributions, among various other climatic and non-climatic factors. Projected regional temperature changes are calculated using the regional atmospheric modelling system (RAMS). MODIS = Moderate Resolution Imaging Spectroradiometer; LST = Land Surface Temperature; NDVI = Normalized Difference Vegetation Index. Adapted from Messina et al. 2012.

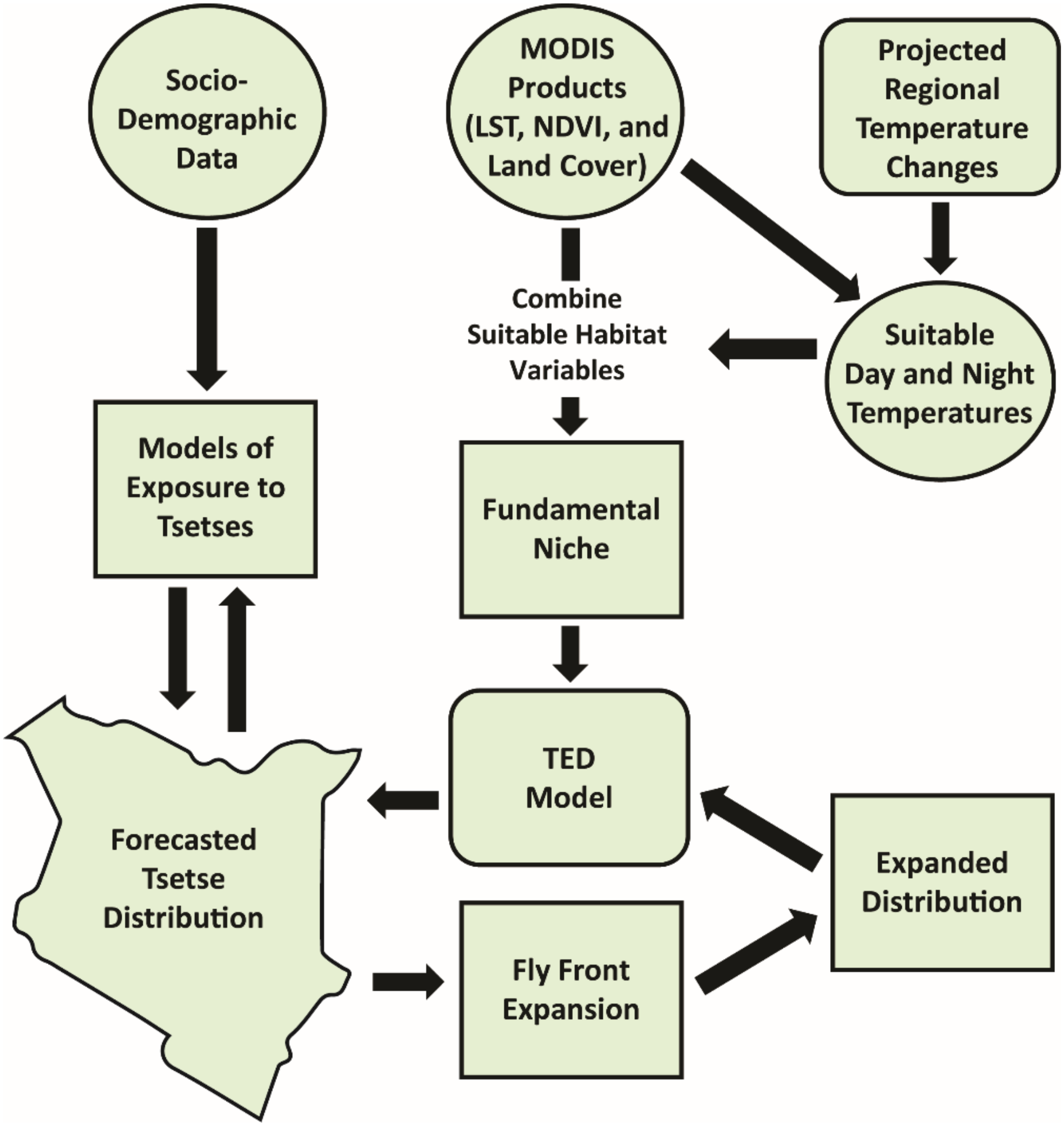

Socio-Demographic Data
MODIS Products (LST, NDVI, and Land Cover)
Projected Regional Temperature Changes
Combine Suitable Habitat Variables
Suitable Day and Night Temperatures
Models of Exposure to Tsetses
Fundamental Niche
TED Model
Forecasted Tsetse Distribution
Expanded Distribution
Fly Front Expansion

Figure 1.2. Methodology applied when using historical species occurrence data to validate species niche models through time (adapted from Kharouba et al. 2009), where models can be (a) forecasted, (b) hindcasted, or both.

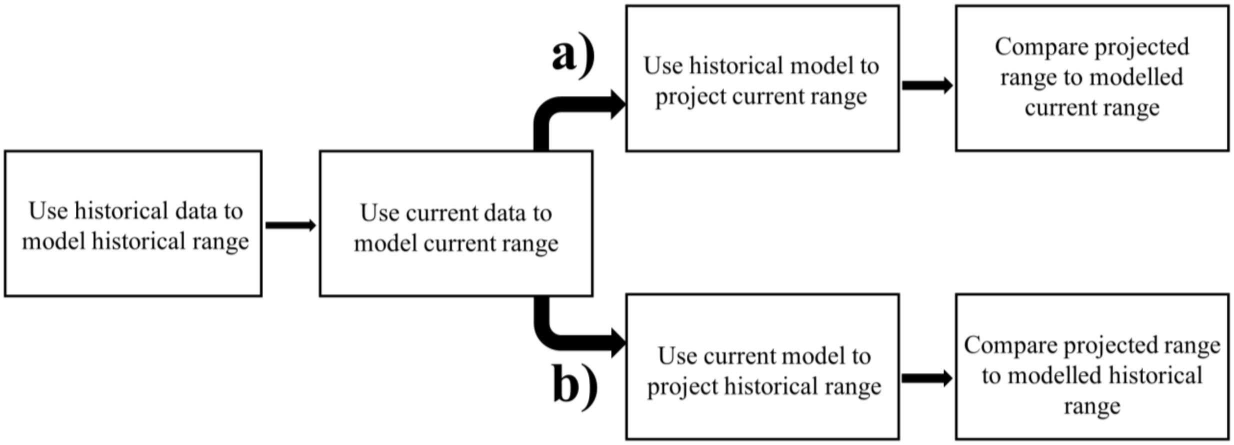
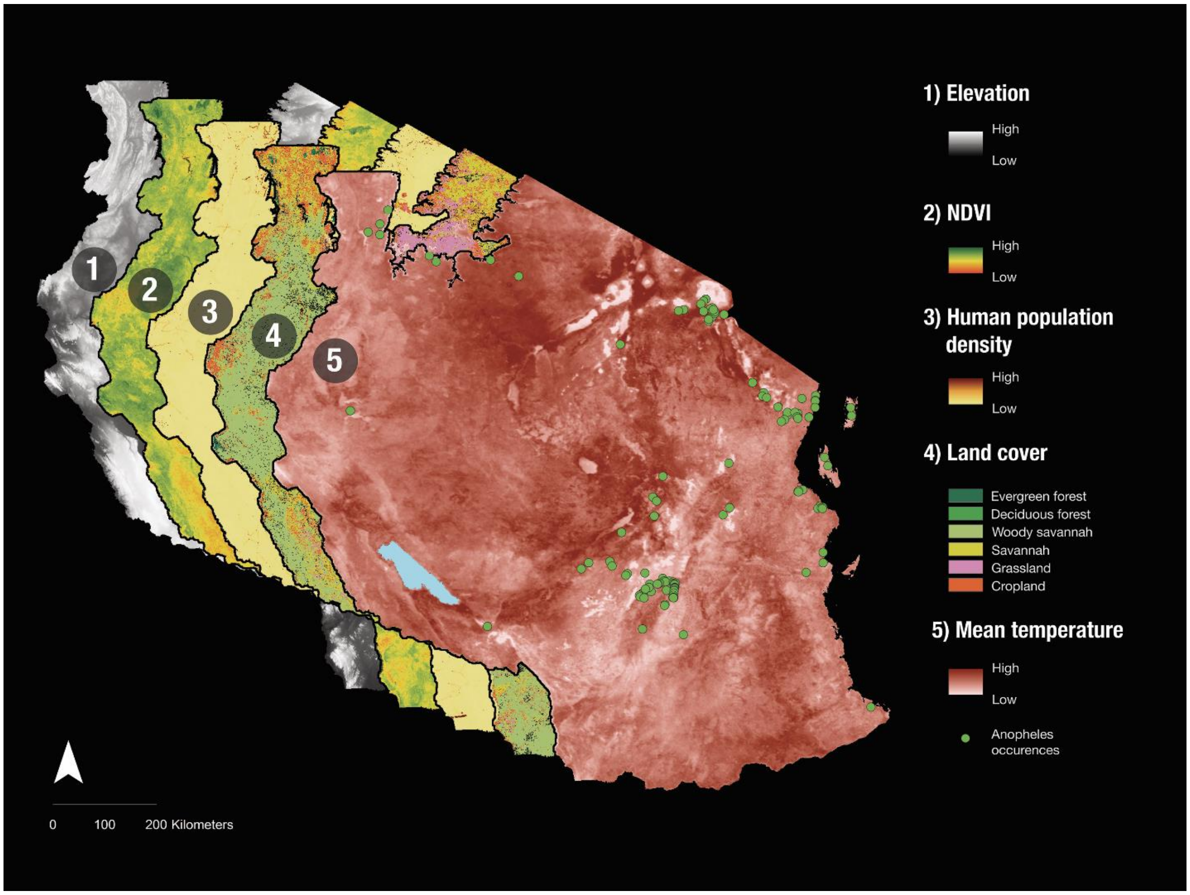

Use historical data to model historical range
Use current data to model current range
a)
Use historical model to project current range
Compare projected range to modelled current range
b)
Use current model to project historical range
Compare projected range to modelled historical range

Figure 1.3. Different satellite imagery layers for mainland Tanzania in 2001 from the Shuttle Radar Topography Mission (SRTM, elevation) and MODIS used in conjunction with *Anopheles* mosquito occurrence records for predictive modelling. These satellite data can often be acquired for multiple years and even in areas where environmental data sampling is sparse.


1
2
3
4
5
1) Elevation
High
Low
2) NDVI
High
Low
3) Human population density
High
Low
4) Land cover
Evergreen forest
Deciduous forest
Woody savannah
Savannah
Grassland
Cropland
5) Mean temperature
High
Low
Anopheles occurences
0
100
200 Kilometers

Chapter 2:
Where have all the mosquito nets gone? Spatial modelling reveals mosquito net distributions across Tanzania do not target optimal *Anopheles* mosquito habitats

Please note: This paper has been submitted to the peer-reviewed journal Malaria Journal *on May 5th, 2015 (and later published on August 17th, 2015) with the following citation:*

Acheson, E. S., A. A. Plowright and J. T. Kerr. Submitted. "Where have all the mosquito nets gone? Spatial modelling reveals mosquito net distributions across Tanzania do not target optimal *Anopheles* mosquito habitats." *Malaria Journal*.

Abstract

Malaria remains the deadliest vector-borne disease despite long-term, costly control efforts. The United Republic of Tanzania has implemented countrywide antimalarial interventions over more than a decade, including national insecticide-treated net (ITN) rollouts and subsequent monitoring. While previous analyses have compared spatial variation in malaria endemicity with ITN distributions, no study has yet compared *Anopheles* habitat suitability to determine proper allocation of ITNs. This study assesses where mosquitoes were most likely to thrive before implementation of large-scale ITN interventions in Tanzania and determine if ITN distributions successfully targeted those areas. The niche model was built with Maxent, incorporating 2001 environmental data with mosquito occurrences from 1999-2003. A 2011-2012 layer of mosquito net use was created using georeferenced data across Tanzania from the Demographic and Health Surveys. We compared baseline mosquito habitat suitability to subsequent ITN use using 1) the average number of ITNs per house and 2) the proportion of households with ≥1 net to compare with national targets. Elevation, land cover, and human population density outperformed variants of temperature and NDVI in the niche model. The spatial distribution of ITNs across

Tanzania was near-random in geographical space (Moran's I=0.070). An average of 2.488 ITNs were used per household and 93.41% of households had ≥1 ITN. Mosquito habitat suitability was not statistically related to numbers of ITNs/household and very weakly to the proportion of households with ≥1 ITN (R^2=0.051). Proportional ITN use/household varied relative to mosquito habitat suitability (Levene's test F=3.0037). We used quantile regression to assess trends in ITN use among households with the highest and lowest 10% of ITN usage. ITN use declined significantly toward areas with the highest vector habitat suitability among households with lowest ITN use (t=-3.38). In areas with lowest habitat suitability, ITN use was consistently higher. ITN use is a key strategy to control malaria in many tropical regions. While Tanzania-wide efforts to distribute ITNs has reduced malaria impacts, gaps and variance in ITN use are unexpectedly large in areas where malaria risk is highest. Supplemental ITN distributions targeting prime *Anopheles* habitats are likely to have disproportionate human health benefits.

Introduction

Malaria remains the most deadly of all vector-borne diseases, infecting an estimated 200 million people and causing over 550 000 deaths globally every year (WHO 2014b). Approximately 90% of all malaria deaths occur in Africa, where an estimated 78% of deaths are in children under the age of five (WHO 2014b). Malaria parasites are transmitted exclusively by *Anopheles* mosquitoes (Kelly-Hope et al. 2009), spurring decades of research not only on the parasite but also on its widespread vector. The spatial variation in vector-borne disease distributions such as malaria can likely be attributed to variations in environmental conditions (e.g., land cover, temperature, precipitation) that disease vectors depend on for survival (Ernst et al. 2006). Ecological niche models (ENMs) correlate these environmental factors with georeferenced occurrences of *Anopheles* mosquitoes to improve detection of spatial and temporal variation in biophysical determinants of these malaria vectors, often at a countrywide or even continent-wide scale (e.g., Kulkarni et al. 2010; Tonnang et al. 2010; Fuller et al. 2012a; Fuller et al. 2012b; Drake and Beier 2014). Although these modelling methods are rapidly advancing using environmental factors, they have yet to incorporate anti-malarial intervention data, though such approaches have been suggested and encouraged (Kelly-Hope et al. 2009). Antimalarial controls are not only time-intensive but also financially burdensome, with $2.7 billion US used for global malaria controls in 2013, although the target was $5.1 billion (WHO 2014b). The synthesis of ecological niche modelling and large-scale mosquito control methods can offer a new perspective on the potential effects of antimalarial controls but has yet to be explored.

Insecticide-treated nets (ITNs) are an excellent example of such vector control methods working on a large scale. Since 2000, global antimalarial initiatives, including

ITNs, have saved an estimated 3.3 million lives and reduced mortality rates by 42% worldwide and 49% in Africa (WHO 2015a). In 2014, a record number of ITNs were distributed to malaria-endemic African countries, with the total number of ITNs delivered to Sub-Saharan Africa reaching 427 million (WHO 2014b). Within the United Republic of Tanzania alone (hereafter referred to as Tanzania), an extensive ITN national plan has been implemented over the past 25 years, supported by both the Global Fund to Fight AIDS, Tuberculosis and Malaria (GFATM) and the USA President's Malaria Initiative (West et al. 2012). Since 2004, the Tanzania National Voucher Scheme has provided subsidized ITNs to pregnant women during antenatal visits (Marchant et al. 2010). Between 2009 and 2010, a mass campaign distributed 8.7 million ITNs across the country, free to families with children under the age of five (Bonner et al. 2011). In 2011, a Universal Coverage Campaign distributed 17.6 million ITNs nationally with the goal to increase global use in the general population to 80% (West et al. 2012). In addition, the GFATM set a national target for Tanzania to increase the proportion of households with at least one ITN to 90% by 2013 (GFATM 2011).

Numerous analyses have been conducted to determine the effects of ITN use on malaria endemicity (e.g., Vanderelst and Speybroeck 2013; Burgert et al. 2014; Larsen et al. 2014; Ngonghala et al. 2014) and on mosquito populations at local scales (e.g., Bayoh et al. 2014; Futami et al. 2014; McCann et al. 2014), but none has yet compared ITN coverage with mosquito habitat suitability to determine if coverage has been optimally allocated. Environments that favour *Anopheles* mosquito survival and reproduction will be environments where malaria risk will be greater. Despite the enormous time- and cost-intensive efforts, antimalarial controls in Tanzania have so far failed to eliminate malaria transmission, likely due to scarcity of information about *Anopheles* ecology and their susceptibility to insecticides (Kabula et al. 2011). With such a mass distribution of ITNs, is the use of ITNs across Tanzania targeting the most at-risk groups in areas with high mosquito habitat suitability? Furthermore, if mosquito nets have been optimized to target the most at-risk groups, areas with greater mosquito habitat suitability would be expected to not only have a greater average number of mosquito nets used per house but also lower variance in mosquito net use (coverage rates should be consistently high in malaria-endemic areas). Yet, ITN use relative to vector suitability remains uncertain even in regions, such as Tanzania, where national household surveys exist to permit such assessments.

Here, we create a habitat suitability model at 1-km resolution of *Anopheles* mosquitoes across Tanzania in 2001 (before large-scale ITN distributions) and compare this model with countrywide ITN use by 2012 (number of ITNs used per house and proportion of houses with at least one ITN) to assess where mosquitoes were most likely to thrive and whether ITN rollouts ensured coverage of such areas. If the distribution of ITNs has been optimized, areas of high mosquito habitat suitability before mass distributions of mosquito nets would be followed by increased use of ITNs throughout these areas and thus lower variance in ITN use. To our knowledge, our study uses the most extensive collection of *Anopheles* mosquito occurrence records in Tanzania to date, collected from over 400 published sources and supplemented by private data collections and the online IR Mapper and the Malaria Atlas Project databases. This study combines ecological niche modelling of *Anopheles* mosquitoes with ITN use and may serve as the starting point for the critical synthesis of *Anopheles* mosquito ecology with broad-scale mosquito control interventions.

Methods

Study area

The study area included all of Tanzania, as well as the group of islands off the mainland's east coast (Figure 2.1). Tanzania occupies 886 100 km^2 (NBS 2013a) and its population exceeded 45 million people by 2012 (WB 2013). Tanzania remains one of the poorest countries in the world, with a per capita average yearly income of $630 US as of 2013 (WB 2013). Of all vector-borne diseases in the country, malaria still causes the highest morbidity and mortality rates (Mboera et al. 2011). Over 85% of cases are caused by *P. falciparum*, the most dangerous malaria parasite (CDC 2013).

The country has a tropical climate with wet and dry seasons that vary due to its topography. The northern and eastern regions experience two wet seasons, with short rains from October to December and long rains from March to May. The western, central, and southern regions mainly have one wet season that lasts from October to April or May. About 80% of the population lives in rural arid regions (Mtega and Ronald 2013) with densities as low as 1 person per km^2, with the remaining population in high-density urban centers, such as Dar es Salaam, of up to 3133 people per km^2 (NBS 2013b). Mosquito net coverage following ITN rollouts has not been found to be associated with poverty (i.e. higher mosquito net coverage was not found in poorer areas) but is intended to prioritize pregnant women and families with children under the age of five (West et al. 2012).

Entomological records

The main malaria vectors in Tanzania are *An. arabiensis*, *An. funestus*, and *An. gambiae* s.s. (Temu et al. 1998). A review of over 400 published literature sources was thus conducted to collect georeferenced occurrence records of Anopheles mosquitoes in Tanzania. Records were collected for Anopheles species for which localities were specified and for any month and year of collection between 1934 and 2014. These data were supplemented by unpublished data collections (personal comm. Dr. Maureen Coetzee) and online databases IR Mapper (http://www.irmapper.com) and the Malaria Atlas Project database (http://www.map.ox.ac.uk). Each coordinate was verified using Google Earth (Version 7.1.2.2019, Google Inc.).

When complete, the database comprised 203 unique georeferenced observations of *An. gambiae* s.l. (35.5%), *An. arabiensis* (32.4%), *An. funestus* s.l. (16.2%), and *An. gambiae* s.s. (15.9%) across all years. Identification of mosquitoes within the *An. gambiae* s.l. complex relied on polymerase chain reaction or cytogenetics. The database details species, month and year of collection, citation source, village of collection, and trapping method.

Niche model

Maxent software (Phillips et al. 2006), v.3.3.3k, was used to predict the relative habitat suitability of *Anopheles* mosquitoes across Tanzania. Maxent combines presence-only species occurrence records with environmental data to create a model that predicts areas of relative habitat suitability for a given species. Maxent niche models of acceptable internal accuracy have been constructed before for a variety of arthropod disease vector species, including mosquitoes (e.g., Foley et al. 2010; Kulkarni et al. 2010; Rochlin et al. 2013; Mughini-Gras et al. 2014), ticks (e.g., Porretta et al. 2013; Feria-Arroyo et al. 2014; Giles et al. 2014), tsetse flies (e.g., Matawa et al. 2013; Bouyer et al. 2014; Dicko et al. 2014) and sandflies (e.g., Colacicco-Mayhugh et al. 2010; Fischer et al. 2011; Samy et al. 2014). Despite Maxent’s tendency to be more conservative than other machine-learning techniques

(Moffett et al. 2007), it is one of the most reliable species distribution modelling methods, even with very small numbers of species observations (Pearson et al. 2007; Hernandez et al. 2008).

Since large-scale mosquito net distributions commenced in 2004, we narrowed our niche model building to the *Anopheles* coordinates collected between 1999 and 2003. Many historical records for anopheline mosquitoes in the 1999-2003 period do not distinguish species within the *An. gambiae* s.l. complex (e.g., only three geographically-unique points were recorded as *An. gambiae* s.s. compared to 54 recorded as *An. gambiae* s.l.). We combined all recorded *Anopheles* species to build the Maxent model. Our focus includes any of these malaria vectors, despite niche differences among species, relative to mosquito net use. All eight species within the *An. gambiae* s.l. complex can transmit malaria parasites except *An. amharicus*, which does not occur in Tanzania, and *An. quadriannulatus*, which has not been recorded in the country since 1968 (Coetzee et al. 2013). The absence of spatially detailed data sufficient for these models prevents species-by-species evaluation of habitat suitability for the time period relevant to national ITN distribution. In total, 56 occurrence records were thus used to build the Maxent model (see Appendix A for full list), substantially exceeding minimum requirements (>5 observations) to produce informative predictions (Pearson et al. 2007).

Multiple models were constructed using elevation, human population, land cover, and variations of temperature and NDVI (average, minimum, and maximum). For assessment of model accuracy, occurrence records were randomly partitioned into 75% for model training and 25% for model testing. Ten replicates were run for each model, using 10-fold cross-validation, each with a randomized partitioning of training and testing data, a technique that can reliably test spatial model skill for disease vector distributions (Acheson and Kerr 2015).

The habitat suitability raster maps were then averaged to determine the relative probability of suitability per grid cell.

The model's accuracy was determined using a threshold-dependent binomial omission test and a threshold-independent receiver operating characteristic analysis (Phillips et al. 2006). For the binomial test of omission, we used a fixed threshold for habitat suitability values of 0.1 (Moffett et al. 2007; Kulkarni et al. 2010). For the threshold-independent analysis, we evaluated model skill using area under the receiver operating characteristic (AUC) (e.g., Estrada-Pena and Venzal 2007; Kulkarni et al. 2010; Moo-Llanes et al. 2013). AUC values approaching 1 indicate perfect discrimination between suitable and unsuitable areas for the target species, and values of 0.5 indicate performance no better than random. Each variable's unique and shared contribution to model accuracy was evaluated using jackknife procedures (Phillips et al. 2006). Each variable's relative contribution to overall model predictions was then calculated and presented as a percentage.

Environmental data

The histories of many mosquito-borne diseases, including malaria, indicate that climate is rarely the principal determinant of mosquito distributions in tropical regions and that shorter-term consequences of human activities often exert greater effects (Rogers and Randolph 2000; Reiter 2001). Relevant climatic factors include land surface temperature and precipitation (Moffett et al. 2007; Kulkarni et al. 2010). Models incorporating satellite-based land cover and/or topography as indicators of mosquito habitats can pinpoint local variations in malaria risk (Kerr et al. 2011). Human population density is a critical determinant of relative malaria risk because of density-dependent transmission of parasites among individuals in the presence of anthropophilic *Anopheles* vectors (Moffett et al. 2007).

We used measurements of environmental conditions at 500-1000m resolution across Tanzania derived using the Moderate Resolution Imaging Spectroradiometer (MODIS) sensor on the Terra satellite. We acquired all 8-day Land Surface Temperature composites (LST; MOD11A2, 1000 metre resolution), 12 monthly Normalized Difference Vegetation Index composites (NDVI; MOD11A3, 1000 m), and land cover observations (MCD12Q1, 500 m) for 2001. NDVI measures greenness and photosynthetic activity, which also captures impacts of irrigation (such as rice-growing areas). We also obtained digital elevation data at 90-m resolution from the Shuttle Radar Topography Mission (SRTM, http://www.cgiar-csi.org/data/srtm-90m-digital-elevation-database-v4-1) as well as 2001 human population density data from the OakRidges National Laboratory LandScan database at 1000-m resolution (http://web.ornl.gov/sci/landscan/).

We unprojected the MODIS data from Sinusoidal to geographic coordinates (GCS WGS 1984 datum) using the MODIS Reprojection Tool (MRT, v.4.0 from US Geological Survey). We transformed human population and elevation layers to the same geographical coordinates using ArcGIS v.10.1 (ESRI 2012, Redlands, CA). Some 8-day LST and NDVI mosaics' pixels were flagged as missing or of low quality, usually due to atmospheric haze or cloud cover. We imported data into R v.3.0.2 (http://www.r-project.org/) and the raster brick function (package *raster*) was used to extract mean, minimum and maximum temperature and NDVI, omitting missing and low-quality pixels. The raster-bricked temperature and NDVI layers were then exported back to ArcGIS 10.1 and all environmental layers were clipped to a mask of countrywide Tanzania and converted to ASCII format for use in Maxent.

Mosquito net survey data

We used 2011-2012 data conducted by the AIDS Indicator Survey (AIS) throughout Tanzania, with permission from the Demographic and Health Surveys (DHS) Program (http://www.dhsprogram.com, funded primary by the US Agency for International Development). AIS fieldwork was conducted between December 2011 and May 2012, implemented by Tanzania's National Bureau of Statistics. This survey provided data on number of ITNs used per household, including long-lasting insecticide nets (LLINs). All DHS procedures and questionnaires were reviewed and approved by the ICF International Institutional Review Board. Further details are provided in each survey's final report.

To protect respondent confidentiality, the georeferenced cluster points were each displaced, with shifts in latitude and longitude under set parameters (Perez-Heydrich et al. 2013). Urban clusters were displaced 0-2 km and rural clusters were displaced 0-5 km, with 1% (or every 100th point) displaced from 0-10 km. Shifts in direction and distance were both random. In addition, displacement of the georeferenced clusters was restricted at the district (admin2) level for the 2011-2012 AIS survey (Figure 2.1). During the randomized cluster displacement process, clusters could cross localized administrative boundaries (e.g., ward boundaries) but not a DHS region or district boundary.

Mosquito net use layers

For the 2011-2012 AIS survey for Tanzania, data were provided for 10 040 households and accompanied by the corresponding GPS dataset of 583 georeferenced survey clusters, which each represent 17-18 surrounding households. Ten clusters were removed due to missing data. Of the remaining 573 clusters, 440 were in rural locations and 133 were in urban locations (Figure 2.2). Using this data set, we measured mosquito net use in two ways:

first, the average number of mosquito nets used in each household and, second, the proportion of houses using ≥ 1 mosquito net. While both offer distinct perspectives on mosquito net use, the latter is particularly relevant to the stated national target of 90% of households with at least one mosquito net (GFATM 2011). For each layer, the resulting calculated value from the surrounding households was assigned to their corresponding cluster coordinate.

We expected that mosquito net usage patterns would show very strong spatial autocorrelation and could consequently be used to estimate mosquito net usage rates throughout unsurveyed areas of Tanzania. To test this possibility, we attempted such extrapolations using kriging and inverse distance weighting techniques in ArcGIS. We compared predicted mosquito net usage from extrapolated surfaces to observed mosquito net usage rates using 10-fold crossvalidation, but model fit was very poor. Mosquito net usage shows very little spatial autocorrelation, as measured using Moran's I statistic ($I = 0.07$, $p < 0.05$, where Moran's I scales as the Pearson product-moment correlation between 1 and -1, with 0 indicating no autocorrelation). We restricted further analyses to point-based measurements.

We created buffer zones surrounding each survey locality, with urban clusters having a circular buffer zone with a radius of 2 km and rural clusters having a circular buffer zone with a radius of 5 km for each layer of ITN usage. Since cluster points were randomly relocated but did not cross district boundaries, buffer zones that crossed into an adjacent district could be made more spatially precise by clipping along the district boundary. Buffer zones were consequently retained only within the district particular to each survey locality. This processing resulted in a total of 912 mosquito net polygons. The area of the buffer zone was assigned the value attributed to its cluster point for the average number of mosquito nets and the proportion of houses with ≥1 ITN. For overlapping buffer areas, the averaged value of the individual buffers was calculated.

Statistical analyses

To relate mosquito habitat suitability to mosquito net usage, the habitat suitability map created in Maxent and the mosquito net buffer layers created in ArcGIS 10.1 were imported into R statistical software (v.3.0.2, http://www.r-project.org). Since each buffer zone has either a 2- or 5-km radius, and had been clipped depending on proximity to district boundaries, R was used to calculate the average habitat suitability pixel value for each buffer zone, including overlapping buffer areas. This process attributed one averaged suitability pixel value to each buffer zone. Ordinary least squares (OLS) regression models were then constructed to test whether *Anopheles* habitat suitability predicted either average numbers of mosquito nets used per household or proportions of households with ≥1 mosquito net. For each regression analysis, Levene's test was conducted to determine if variances differed significantly between deciles of suitability. Residual spatial autocorrelation and the associated Moran's I were calculated using R. Results were considered significant if $p<0.05$. In addition, we tested for other possible relationships between the two measures of mosquito net use as a function of mosquito habitat suitability by performing quantile regression analyses in the 0.1 and 0.9 quantiles. Quantile regression is a method for providing a more complete view of possible relationships in ecological processes aside from the mean of the response variable distribution (Cade and Noon 2003).

Results

Habitat suitability model

Elevation, land cover, and human population density provided stronger predictions of mosquito habitat suitability than any temperature- or NDVI-based measurements (Table 2.1). Model fits were comparable whether the latter measurements were omitted (AUC = 0.872) or included (AUC = 0.862-0.875). Elevation made the greatest contribution to the model, followed by human population density and land cover. Jackknife analysis attributed the maximum training gain to elevation (0.828), which indicates this variable contains the most information that is not present in the other variables. Modelled habitat suitability demonstrated concentrated areas of high habitat suitability along low elevation areas, mainly along the eastern coastlines and extending into Kilombero Valley and in the southeast along the Mozambique border (Figure 2.3). Areas surrounding Lake Victoria and north of Lake Malawi also contain extensive areas with high habitat suitability for anopheline mosquitoes. Dense shrub land and cropland/natural vegetation mosaic land covers were associated with higher mosquito habitat suitability. Internal accuracy of the model through the threshold-dependent binomial omission test was also high, with an omission error rate of 7.14% (n = 56). Removing elevation from Maxent models degraded model fit substantially (AUC=0.765), increasing proportional contributions of human population density and land cover (65% and 25.1%, respectively), with minor increases in temperature and NDVI (7.1% and 2.8%, respectively). The model excluding any variant of temperature and NDVI while still maintaining a comparatively high AUC value was retained.

ITN use and mosquito habitat suitability

The relationship between the average number of mosquito nets used per house and mosquito habitat suitability was not significant ($p>0.05$, Figure 2.4a) and model residuals showed practically no spatial autocorrelation in net use, even at short distances (Moran's I = 0.07, $p < 1 \times 10^{-6}$). Average mosquito net use per house did not significantly differ between rural (2.485 mosquito nets) and urban (2.491 mosquito nets) surveyed areas ($t= 0.13$, $p=0.8967$). Although variance appears smaller in only the highest decile of habitat suitability (Figure 2.4b), a Levene's test reveals variance did not significantly differ among deciles of habitat suitability ($F=0.7821$, $p=0.6331$). Both the 0.1 and 0.9 quantiles of the average numbers of mosquito nets per house showed no significant trend with mosquito habitat suitability ($p > 0.05$; Figure 2.4a).

There is a slight tendency for the proportion of households to decrease toward areas with higher mosquito habitat suitability ($R^2 = 0.051$, $p < 1 \times 10^{-6}$; Figure 2.4c). Overall, 93.41% of surveyed households had at least one mosquito net (93.51% in rural households and 93.05% in urban households). A Levene's test revealed that variance did differ significantly between deciles of habitat suitability ($F = 3.0037$, $p = 0.0015$; Figure 2.4d). While the 0.9 quantile of the proportion of households with at least one mosquito net showed no significant trend relative to mosquito habitat suitability, the 0.1 quantile of the proportion of households with at least one mosquito net significantly decreased as mosquito habitat suitability increased ($\beta = -0.0753$, $t = 3.38$, $p < 0.001$; Figure 2.4c).

Discussion

The World Health Organization (WHO) estimates ~50% of African households had access to ≥1 ITN by 2010 but only ~3% in 2004 (WHO 2014b). Our analysis of the DHS survey data across Tanzania from 2011-2012 suggests more optimistic results, with >93% of Tanzanian households having one or more nets, meeting the national target of at least 90% by 2013 (GFATM 2011). There is no statistical difference in ITN use between rural and urban households. However, the effectiveness of ITN distributions varies among communities, many of which have much lower rates of ITN use than national averages. Among such communities, the proportion of households with at least one mosquito net is inversely related to mosquito habitat suitability. There is a clear opportunity to improve ITN effectiveness at a national scale by targeting supplemental net distributions in areas where current use rates are low but mosquito habitat suitability is extremely high. The unexpected decrease in ITN presence in some surveyed areas where malaria risk is high likely contributes to persistent, high malaria morbidity and mortality (20 900 deaths in 2013 (WHO 2015b)). Other factors clearly contribute to malaria persistence, including increasing insecticide resistance among *Anopheles* mosquitoes in Tanzania (Protopopoff et al. 2013; Matowo et al. 2014; Nkya et al. 2014).

Regardless of the causes for whether ITN rollouts achieve national or international targets, it is imperative that they reach households where mosquito habitat suitability, and thus malaria risk, is highest. We expected that areas with higher mosquito habitat suitability would have the most comprehensive ITN coverage and lowest variance in the number and proportion of households with one or more ITNs, but we observed the opposite trends. In areas with the highest anopheline habitat suitability, proportional ITN presence in households declines below more modest (≥80%) international targets (77% of households have ≥1 ITN; Figure 2.4c). These declines were most visible when analyzing the proportion of households within the lowest 10% with ≥1 ITN, indicating that evaluations of performance trends using only measures of central tendencies may fail to identify gaps in ITN use that may have significant health consequences. Over an assemblage of 23 African countries, coarse resolution analyses have revealed similarly uneven mosquito net use relative to malaria endemicity (Vanderelst and Speybroeck 2013). Internal networks used to distribute ITNs and education and promotion programs supporting ITN use in Tanzania are unlikely the causes of local variability in ITN presence or usage rates, as such factors (e.g., road access) should have led to spatially autocorrelated trends, which were marginal in this study (Moran's I = 0.07). In other African countries, variation in net use has multiple causes. These include beliefs in malaria risk, trust in health workers distributing nets, perceived benefits of the nets, education, the number of children under five in the household, and availability and use of other vector control measures (Graves et al. 2011).

Our prediction of mosquito habitat suitability relates strongly to georeferenced malaria cases as well as independent maps of *Anopheles* mosquito distributions in Tanzania. We repeated the binomial test of omission with the fixed threshold of 0.1 to test the model against independently-collected data on unique georeferenced records of *Plasmodium falciparum* across Tanzania from the MAP website. Our prediction of mosquito habitat suitability corresponded strongly to malaria records from 1985-2008 (of 568 geographically-unique *P. falciparum* records, our model had an omission error rate of 12.46%) as well as those specifically between our 1999-2003 study period (of 107 *P. falciparum* records, our model had an omission error rate of 13.21%). In addition, our model's prediction of mosquito habitat suitability (based on the fixed threshold of 0.1) was broadly consistent with the predicted occurrence maps for *An. gambiae*, *An. arabiensis*, *An. funestus*, and *An. gambiae* s.s. as developed by the Malaria Atlas Project across the eastern coastline (including the islands), throughout Kilombero Valley, and extending halfway along the Mozambique border (MAP 2010). However, our model did not predict as extensive a distribution as those provided by the Malaria Atlas Project (MAP), even though our model combined *An. gambiae* s.l. and *An. funestus* s.l. mosquito species. MAP data are coarser and are used to generate continental models that are subsequently clipped by country, whereas our map was built using more detailed satellite information and more extensive vector observation data, which may reveal local variations in suitability of *Anopheles* habitats. Although our *Anopheles* training records did not include the Lake Victoria region, our model correctly predicted high relative habitat suitability there. This region is a malaria hotspot (Gething et al. 2011), with malaria steadily spreading to higher altitudes and aggravated by climate variability and poverty (Wandiga et al. 2010).

The distribution of anopheline mosquitoes and malaria is strongly influenced by land cover (e.g., Foley et al. 2009; Fuller et al. 2012b) and human population density (e.g., Moffett et al. 2007; Ayala et al. 2009) in addition to climatic and topographical factors (Moffett et al. 2007; Mutuku et al. 2009; Ermert et al. 2012; Stefani et al. 2013). Elevation can similarly limit *Anopheles* distributions (e.g., Moffett et al. 2007; Kulkarni et al. 2010; Sinka et al. 2010) and its omission from ecological niche models can severely degrade both model fit and prediction accuracy (Sweeney et al. 2006). In this study, mosquito habitat suitability increased particularly in cropland/natural land cover, reflecting roles for irrigation or water pooling, but also in closed shrublands. Land features, such as soil type, are not directly measured in these satellite land cover data but may help explain such findings through differences in water holding capacity that can alter mosquito breeding success (e.g., Foley et al. 2009). Temperature and NDVI explained little unique variation in habitat suitability, though both can limit mosquito distributions under some circumstances (e.g., Kulkarni et al. 2010; Sinka et al. 2010; Obsomer et al. 2012; Drake and Beier 2014). Temperature shows relatively little spatial variation across Tanzania except along elevation gradients (Goyal and Harmsen 2014), which are more directly measured by the elevation variable. NDVI relates to vegetation greenness, productivity, and moisture availability, which are also strongly related to mosquito reproduction (Yamana and Eltahir 2013). However, satellite-based land cover measurements provide similar measurements that may subsume the unique contributions arising from measures of greenness. Furthermore, highly productive vegetation (i.e. areas with high NDVI) can be found in many regions of Tanzania, eroding the capacity for such measures to discriminate between suitable and unsuitable mosquito habitats at this scale of analysis.

While our dataset includes the most comprehensive collection of spatially-unique georeferenced *Anopheles* records for this region to date, systematic, randomized, broad-scale sampling that identifies species in the *An. gambiae* complex would be extremely valuable. Presence-only species observations assembled from an array of sources that differ in sampling effort and geographical focus (Kearney et al. 2012) could bias models toward areas with easier access (e.g., roadsides). Collections are not fully or equally representative of the cross-section of environments in Tanzania. In addition, lack of species-by-species evaluation for the historical period forced us to combine *Anopheles* species. Increased sampling and differentiation of species within the *An. gambiae* complex is needed to create species-by-species niche models that can be more reliably projected through time and related to measurements of the effectiveness of programs such as ITN rollouts. ITN use can differ depending on which *Anopheles* species is locally common (e.g., *An. arabiensis* is more commonly found outdoors in dry environments and urban areas (Sinka et al. 2010), rendering it less affected by ITNs than *An. gambiae* s.s. and *An. funestus* (Kitau et al. 2012)). Systematic, repeated sampling of *Anopheles* mosquitoes across Tanzania could significantly improve the capacity to target interventions efficiently and relative to health needs and would provide a stronger basis for predicting how distributions of these vectors would change through time (Obsomer et al. 2013).

Conclusions

Insecticide-treated nets have been broadly distributed across Tanzania but are used unevenly and with decreasing frequency in areas where anopheline mosquito distributions are concentrated. The combination of ecological niche models, drawing on high-resolution satellite information, and georeferenced mosquito net data can provide useful perspective on the effectiveness of such interventions in areas where georeferenced vector and vector control data exist. These techniques are transferrable to other vector-borne disease systems. With 3.2 billion people at risk of malaria infection worldwide (WHO 2014b), the need to control malaria continues to be a top global health priority. The incorporation of georeferenced mosquito net data into mosquito habitat suitability modelling is a promising approach to inform control programs how to optimize allocation of limited resources and minimize malaria cases where risk is greatest.

Tables

Table 2.1. Contribution of environmental variables to *Anopheles* Maxent model, including and excluding variants of temperature and NDVI. Variants of temperature and NDVI (mean, maximum, minimum) were attempted in different models, but yielded similar minor contributions to the model. AUC, or area under the curve, values approaching 1 indicate the model's perfect discrimination between suitable and unsuitable areas for the target species, and values near 0.5 indicate performance no better than random. Model 4 was retained for analyses.

	Variable	**% Contribution to model**	**AUC with variable alone**	**AUC without variable**	**Model AUC**
Model 1	Elevation	66.2	0.817	0.772	0.869
	Human population density	28.2	0.727	0.773	
	Land cover	4.2	0.634	0.865	
	Mean temperature	1.2	0.615	0.874	
	Mean NDVI	0	0.593	0.869	
Model 2	Elevation	65.9	0.817	0.769	0.875
	Human population density	28.0	0.727	0.815	
	Land cover	4.3	0.634	0.870	
	Max temperature	1.4	0.546	0.876	
	Max NDVI	0.5	0.542	0.875	
Model 3	Elevation	64.3	0.817	0.766	0.862
	Human population density	27.1	0.727	0.800	
	Land cover	4.0	0.634	0.857	
	Min NDVI	2.9	0.606	0.875	
	Min temperature	1.6	0.575	0.862	
Model 4	Elevation	66.5	0.812	0.764	0.872
	Human population density	28	0.732	0.794	
	Land cover	5.5	0.640	0.876	

Figures

Figure 2.1. United Republic of Tanzania, detailed by districts and elevation. The map shows the 148 Demographic and Health Surveys "admin 2" district boundaries, as well as elevation, several major cities, and bordering countries. The inset map details Tanzania's location within Africa.

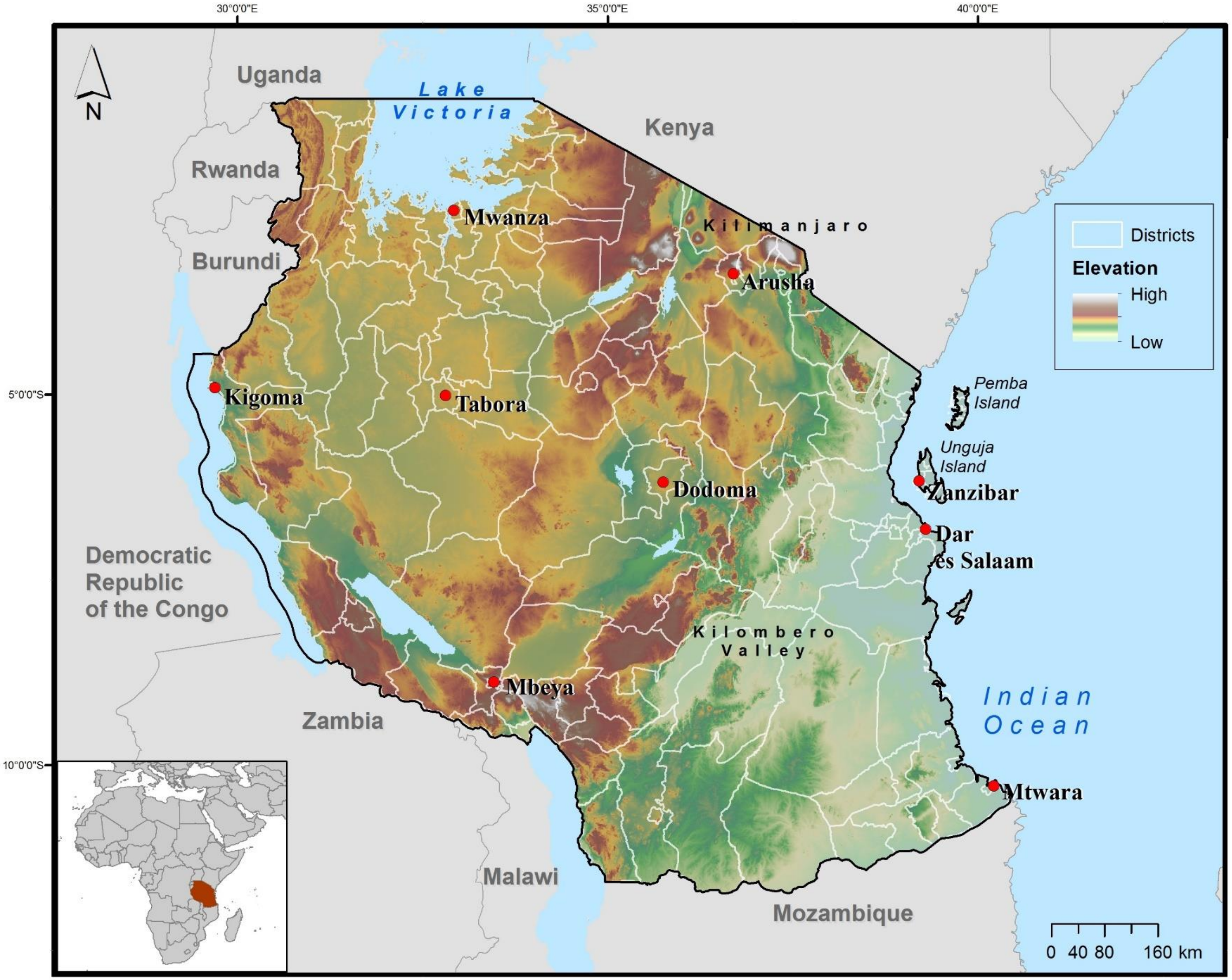

30°0'0"E
35°0'0"E
40°0'0"E
5°0'0"S
10°0'0"S
N
Uganda
Lake Victoria
Kenya
Rwanda
Burundi
Mwanza
Kilimanjaro
Arusha
Kigoma
Tabora
Dodoma
Pemba Island
Unguja Island
Zanzibar
Dar es Salaam
Democratic Republic of the Congo
Kilombero Valley
Mbeya
Zambia
Indian Ocean
Mtwara
Malawi
Mozambique
Districts
Elevation
High
Low
0 40 80 160 km

Figure 2.2. The 573 raw georeferenced cluster points provided by the 2011-2012 DHS survey. Cluster points are colour coded to represent spatial variation in a) average number of mosquito nets used per household or b) the proportion of households with at least one mosquito net across Tanzania.

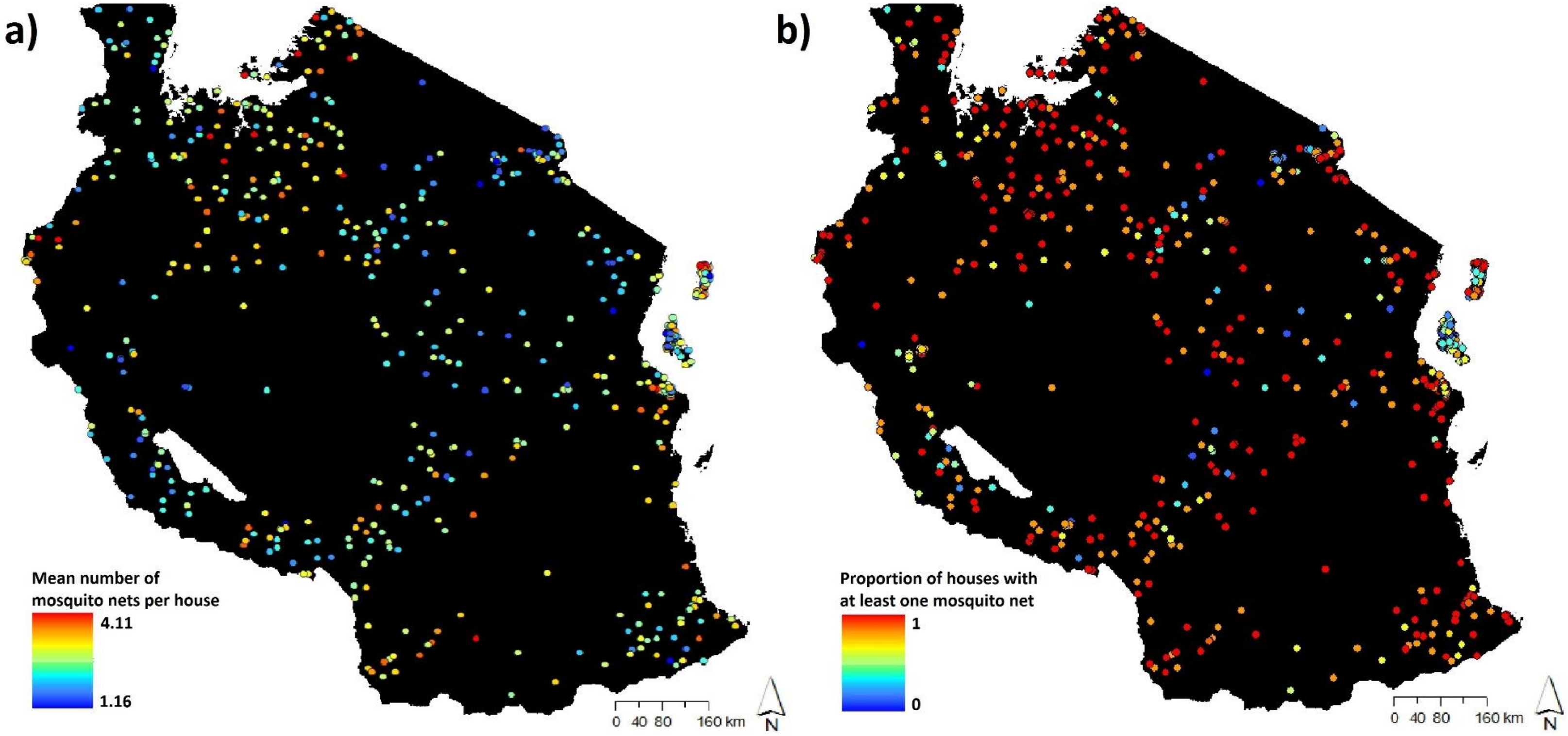

a)
Mean number of
mosquito nets per house
4.11
1.16
0 40 80 160 km
N
b)
Proportion of houses with
at least one mosquito net
1
0
0 40 80 160 km
N

Figure 2.3. Maxent output of relative *Anopheles* mosquito habitat suitability. The habitat suitability map ranges from blue to red, with blue representing lowest predicted relative habitat suitability for *Anopheles* mosquitoes. The suitability map is overlaid with the “admin 2” districts of Tanzania. Predicted relative habitat suitability in many areas are coterminous with the LandScan human population data (e.g., the southeastern polygonal area of low habitat suitability corresponds substantially with the Selous Game Reserve, where human population approaches zero in most areas.)

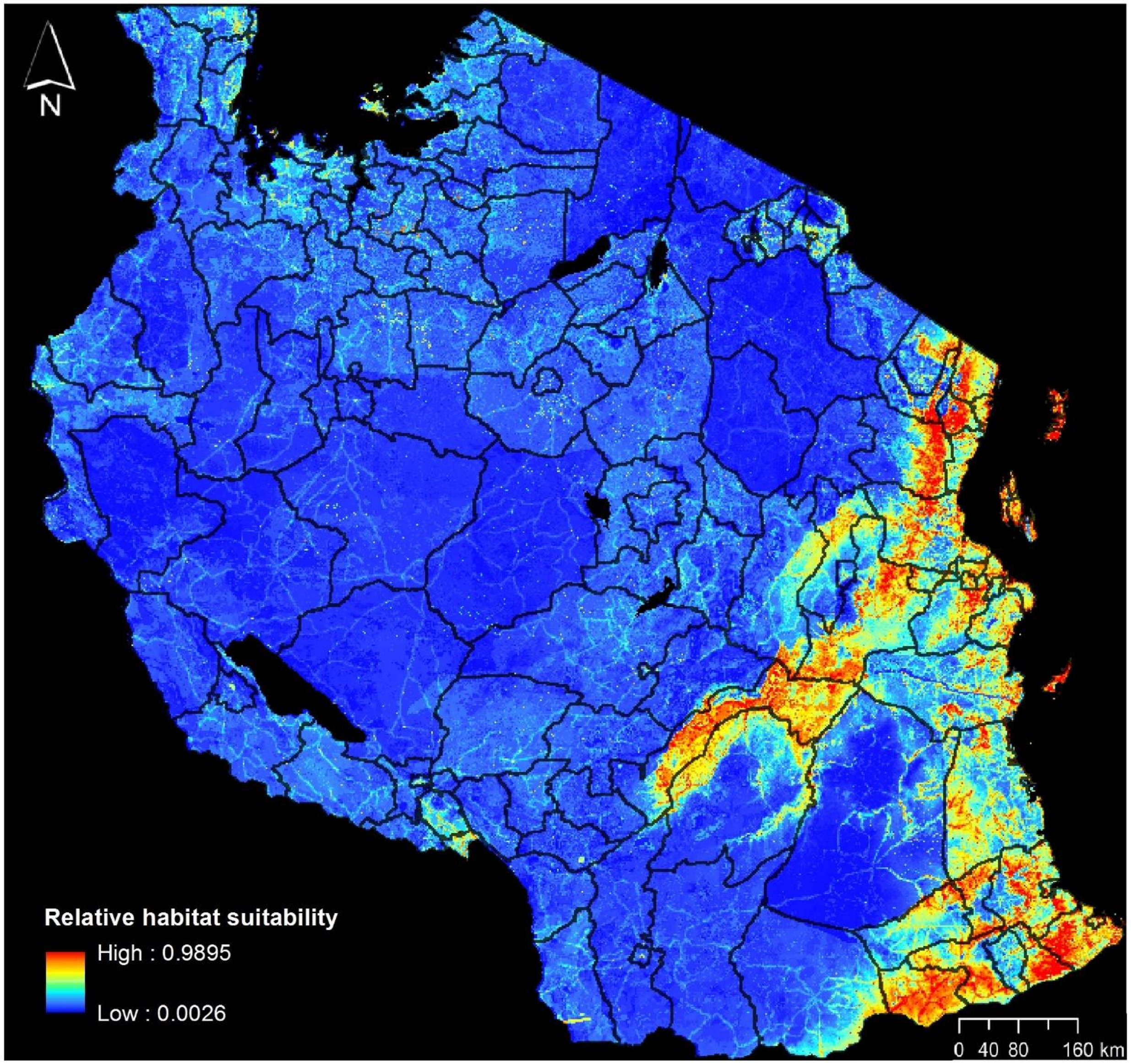
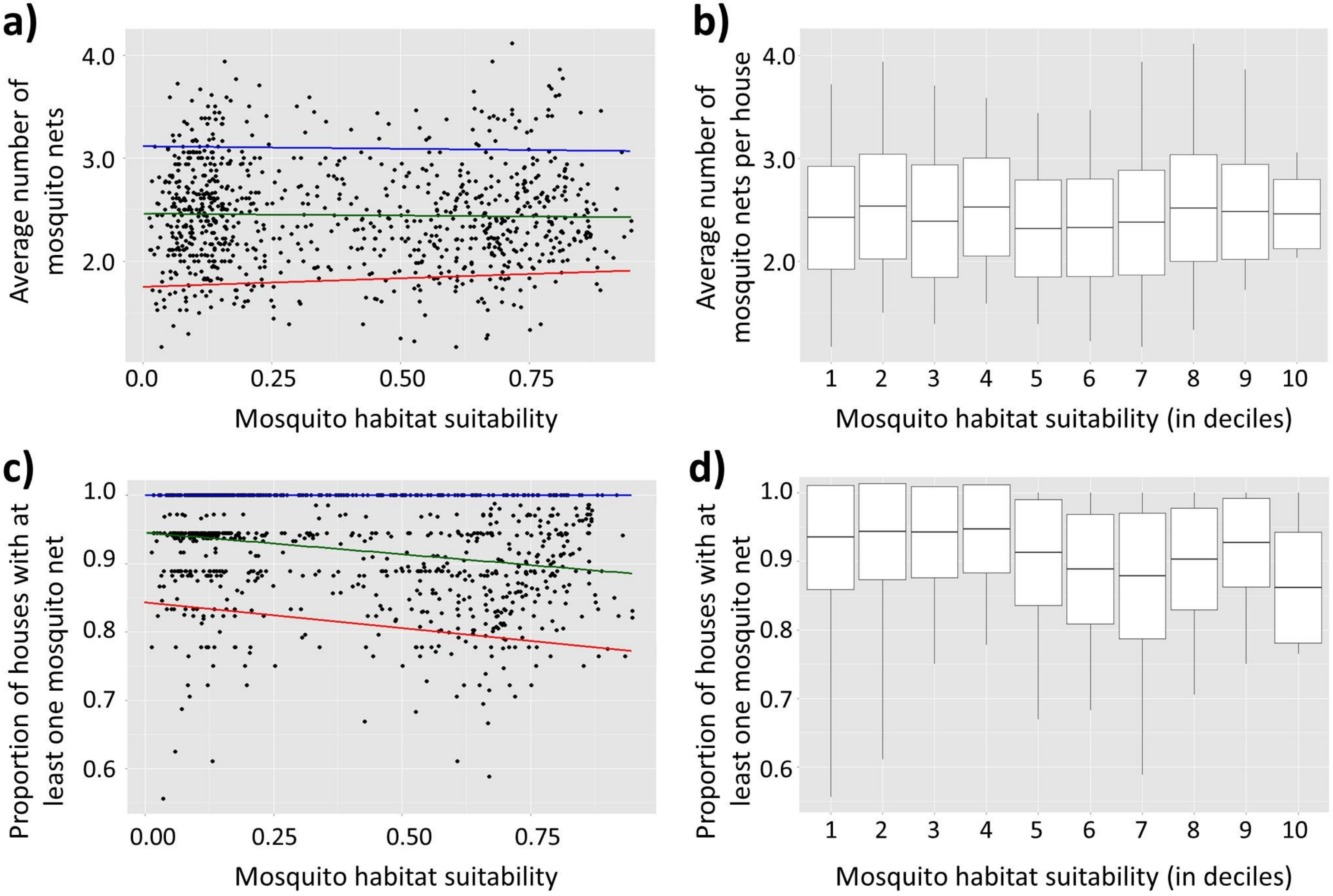

N
Relative habitat suitability
High : 0.9895
Low : 0.0026
0 40 80 160 km

Figure 2.4. Scatterplots and corresponding boxplots of mosquito net use as a function of mosquito habitat suitability. (a) Scatterplot of the average number of mosquito nets per house as a function of relative *Anopheles* habitat suitability with quantile regression and ordinary least squares (OLS) regression lines shown. The 0.9 quantile line is shown in blue, the ordinary least squares regression line is shown in green, and the 0.1 quantile line is shown in red. (b) Boxplot of mean and standard deviation of average number of mosquito nets used per house as a function of *Anopheles* relative habitat suitability, where habitat suitability is divided into deciles (e.g., 1 corresponds to lowest 10%, etc.). Boxplot whiskers extend to the maximum and minimum values. (c) Scatterplot of the proportion of houses with mosquito nets as a function of relative *Anopheles* habitat suitability (i.e. houses with ≥1 mosquito net were assigned a value of 1; otherwise, 0), with quantile regression and ordinary least squares (OLS) regression lines shown. The 0.9 quantile line is shown in blue, the ordinary least squares regression line is shown in green, and the 0.1 quantile line is shown in red. (d) Boxplot of mean and standard deviation of the proportion of houses with mosquito nets as a function of relative *Anopheles* habitat suitability, where habitat suitability is divided into deciles.


a)
Average number of mosquito nets
4.0
3.0
2.0
0.0
0.25
0.50
0.75
Mosquito habitat suitability
b)
Average number of mosquito nets per house
4.0
3.0
2.0
1
2
3
4
5
6
7
8
9
10
Mosquito habitat suitability (in deciles)
c)
Proportion of houses with at least one mosquito net
1.0
0.9
0.8
0.7
0.6
0.00
0.25
0.50
0.75
Mosquito habitat suitability
d)
Proportion of houses with at least one mosquito net
1.0
0.9
0.8
0.7
0.6
1
2
3
4
5
6
7
8
9
10
Mosquito habitat suitability (in deciles)

Conclusions

This thesis has focused on the spatial and temporal modelling of arthropod disease vector distributions. Specifically, it focused on statistical modelling methods that 1) are currently used to forecast arthropod vector distributions under changing environmental scenarios and 2) can potentially be used to evaluate how well georeferenced vector controls, such as insecticide-treated nets in Tanzania, are targeting disease vector distributions.

Chapter One systematically reviewed the current statistical modelling methods used to forecast disease vector distributions under future environmental changes. My results suggested that most models are validated through space but not through time and use the validated spatial model to project into future time periods under different climate change scenarios. These results highlight that spatial validation does not imply temporal reliability. Historical calibration of temporal models using multi-temporal analyses (Kharouba et al. 2009) is a promising solution to ensure the latter. Further investigations into other methods to increase accuracy of spatial and temporal species distribution models would be most useful for future work, including historically-calibrated ensemble forecast models. Historical calibration can strengthen disease vector model forecasts, which in turn could improve prospects for successful interventions to reduce vector-borne impacts on human health.

Chapter Two tested whether the insecticide-treated mosquito nets distributions as of 2012 have been optimally targetting areas of high mosquito habitat suitability in Tanzania, as determined by a mosquito niche model built from data before large-scale ITN distributions. I not only found that mosquito nets were randomly distributed across the country, but the proportion of households with at least one net was inversely related to mosquito habitat suitability. These results suggest there is an urgent need to increase mosquito net coverage in areas at high risk of malaria in Tanzania. Other broader-scale analyses of mosquito net distributions in sub-Saharan Africa confirmed my findings of a random distribution of ITNs (Vanderelst and Speybroeck 2013), suggesting numerous other socioeconomic and cultural factors may be playing a role in this trend. This finding brings attention to novel interdisciplinary ways of studying the efficacy of vector control allocations by combining georeferenced vector control data with species distribution modelling, compares mosquito net use to national targets for Tanzania (including the proportion of houses with at least one mosquito net), and can be applied to any study region where georeferenced vector and vector control data exist.

The major weakness of my thesis is the requirement to rely on an *Anopheles* niche model. The niche model was constructed on occurrence records collected from a variety of independent sources which demonstrated sampling bias and inconsistent systematic sampling across Tanzania. Yet, independent validation using malaria cases within the same time periods of collection demonstrate the model shows high predictive accuracy and is therefore useful to compare to mosquito net use. The rejection of the hypothesis that insecticide-treated nets have optimally targetted prime *Anopheles* habitats sheds new light on time- and cost-intensive vector controls in Tanzania and potentially in other tropical and sub-tropical countries. Countrywide mosquito net rollouts require increased monitoring and surveillance to manage such large-scale antimalarial interventions and ensure maximum coverage in areas at the greatest risk of malaria. Future work should focus on determining if similar relationships exist between mosquito habitat suitability models and ITN distributions in other tropical and sub-tropical countries.

To my knowledge, this thesis presents the first systematic review of disease vector forecast modelling, as well as the first study to incorporate large-scale antimalarial control survey data into *Anopheles* mosquito distribution modelling. Given the results presented here, vector control interventions can incorporate georeferenced survey data into ecological niche modelling to target vector-borne diseases at large spatial scales. This thesis also stresses that while statistical models of disease vectors may nearly always contain sampling flaws, their thorough validation through space and time still makes them an invaluable contribution to disease vector control.

References

Acevedo, P., F. Ruiz-Fons, R. Estrada, A. L. Marquez, M. A. Miranda, C. Gortazar and J. Lucientes. "A Broad Assessment of Factors Determining Culicoides Imicola Abundance: Modelling the Present and Forecasting Its Future in Climate Change Scenarios." *Plos One* 5, no. 12 (2010): 13.

Acheson, E. S. and J. T. Kerr. "Looking Forward by Looking Back: Using Historical Calibration to Improve Forecasts of Human Disease Vector Distributions." *Vector-Borne and Zoonotic Diseases* 15, no. 3 (2015): 173-183.

Afrane, Y. A., A. K. Githeko and G. Yan. "The Ecology of Anopheles Mosquitoes under Climate Change: Case Studies from the Effects of Deforestation in East African Highlands." *Year in Ecology and Conservation Biology* 1249, (2012): 204-210.

Algar, A. C., H. M. Kharouba, E. R. Young and J. T. Kerr. "Predicting the Future of Species Diversity: Macroecological Theory, Climate Change, and Direct Tests of Alternative Forecasting Methods." *Ecography* 32, no. 1 (2009): 22-33.

Allouche, O., A. Tsoar and R. Kadmon. "Assessing the Accuracy of Species Distribution Models: Prevalence, Kappa and the True Skill Statistic (Tss)." *Journal of Applied Ecology* 43, no. 6 (2006): 1223-1232.

Araujo, M. B. and A. Guisan. "Five (or So) Challenges for Species Distribution Modelling." *Journal of Biogeography* 33, no. 10 (2006): 1677-1688.

Araujo, M. B. and M. New. "Ensemble Forecasting of Species Distributions." *Trends in Ecology & Evolution* 22, no. 1 (2007): 42-47.

Araujo, M. B. and R. G. Pearson. "Equilibrium of Species' Distributions with Climate." *Ecography* 28, no. 5 (2005): 693-695.

Araujo, M. B., R. G. Pearson, W. Thuiller and M. Erhard. "Validation of Species-Climate Impact Models under Climate Change." *Global Change Biology* 11, no. 9 (2005): 1504-1513.

Araujo, M. B. and A. T. Peterson. "Uses and Misuses of Bioclimatic Envelope Modeling." *Ecology* 93, no. 7 (2012): 1527-1539.

Araujo, M. B., R. J. Whittaker, R. J. Ladle and M. Erhard. "Reducing Uncertainty in Projections of Extinction Risk from Climate Change." *Global Ecology and Biogeography* 14, no. 6 (2005): 529-538.

Ayala, D., C. Costantini, K. Ose, G.C. Kamdem, C. Antonio-Nkondjio, J. Agbor, P. Awono-Ambene, D. Fontenille and F. Simard. "Habitat Suitability and Ecological Niche Profile of Major Malaria Vectors in Cameroon." *Malaria Journal* 8, (2009): 307.

Bahn, V. and B.J. McGill. "Testing the Predictive Performance of Distribution Models." *Oikos* 122, no. 3 (2013): 321-331.

Bayoh, M. N., E. D. Walker, J. Kosgei, M. Ombok, G. B. Olang, A. K. Githeko, G. F. Killeen, P. Otieno, M. Desai, N. F. Lobo, J. M. Vulule, M. J. Hamel, S. Kariuki and J. E. Gimnig. "Persistently High Estimates of Late Night, Indoor Exposure to Malaria Vectors Despite High Coverage of Insecticide Treated Nets." *Parasites & Vectors* 7, (2014): 13.

Beebe, N. W., R. D. Cooper, P. Mottram and A. W. Sweeney. "Australia's Dengue Risk Driven by Human Adaptation to Climate Change." *Plos Neglected Tropical Diseases* 3, no. 5 (2009): 9.

Blanford, J. I., S. Blanford, R. G. Crane, M. E. Mann, K. P. Paaijmans, K. V. Schreiber and M. B. Thomas. "Implications of Temperature Variation for Malaria Parasite Development across Africa." *Scientific Reports* 3, (2013): 1300.

Blois, J. L., J. W. Williams, M. C. Fitzpatrick, S. T. Jackson and S. Ferrier. "Space Can Substitute for Time in Predicting Climate-Change Effects on Biodiversity." *Proceedings of the National Academy of Sciences of the United States of America* 110, no. 23 (2013): 9374-9379.

Bonner, K., A. Mwita, P. D. McElroy, S. Omari, A. Mzava, C. Lengeler, N. Kaspar, R. Nathan, J. Ngegba, R. Mtung'e and N. Brown. "Design, Implementation and Evaluation of a National Campaign to Distribute Nine Million Free Llins to Children under Five Years of Age in Tanzania." *Malaria Journal* 10, (2011): 12.

Borkent, A. and D. A. Grimaldi. "The Earliest Fossil Mosquito (Diptera : Culicidae), in Mid-Cretaceous Burmese Amber." *Annals of the Entomological Society of America* 97, no. 5 (2004): 882-888.

Botkin, D. B., H. Saxe, M. B. Araujo, R. Betts, R. H. W. Bradshaw, T. Cedhagen, P. Chesson, T. P. Dawson, J. R. Etterson, D. P. Faith, S. Ferrier, A. Guisan, A. S. Hansen, D. W. Hilbert, C. Loehle, C. Margules, M. New, M. J. Sobel and D. R. B. Stockwell. "Forecasting the Effects of Global Warming on Biodiversity." *Bioscience* 57, no. 3 (2007): 227-236.

Bouyer, F., M. T. Seck, A. H. Dicko, B. Sall, M. Lo, M. J. B. Vreysen, E. Chia, J. Bouyer and A. Wane. "Ex-Ante Benefit-Cost Analysis of the Elimination of a Glossina Palpalis Gambiensis Population in the Niayes of Senegal." *Plos Neglected Tropical Diseases* 8, no. 8 (2014): e3112.

Brady, O. J., M. A. Johansson, C. A. Guerra, S. Bhatt, N. Golding, D. M. Pigott, H. Delatte, M. G. Grech, P. T. Leisnham, R. Maciel-de-Freitas, L. M. Styer, D. L. Smith, T. W. Scott, P. W. Gething and S. I. Hay. "Modelling Adult Aedes Aegypti and Aedes

Albopictus Survival at Different Temperatures in Laboratory and Field Settings." *Parasites & Vectors* 6, (2013): 351.

Braunisch, V., J. Coppes, R. Arlettaz, R. Suchant, H. Schmid and K. Bollmann. "Selecting from Correlated Climate Variables: A Major Source of Uncertainty for Predicting Species Distributions under Climate Change." *Ecography* 36, no. 9 (2013): 971-983.

Broennimann, O. and A. Guisan. "Predicting Current and Future Biological Invasions: Both Native and Invaded Ranges Matter." *Biology Letters* 4, no. 5 (2008): 585-589.

Brown, J. H. and B. A. Maurer. "Macroecology - the Division of Food and Space among Species on Continents." *Science* 243, no. 4895 (1989): 1145-1150.

Brownstein, J. S., T. R. Holford and D. Fish. "Effect of Climate Change on Lyme Disease Risk in North America." *EcoHealth* 2, no. 1 (2005): 38-46.

Burgert, C. R., S. E. K. Bradley, F. Arnold and E. Eckert. "Improving Estimates of Insecticide-Treated Mosquito Net Coverage from Household Surveys: Using Geographic Coordinates to Account for Endemicity." *Malaria Journal* 13, (2014): 13.

Cade, B. S. and B. R. Noon. "A Gentle Introduction to Quantile Regression for Ecologists." *Frontiers in Ecology and the Environment* 1, no. 8 (2003): 412-420.

Caminade, C., J. M. Medlock, E. Ducheyne, K. M. McIntyre, S. Leach, M. Baylis and A. P. Morse. "Suitability of European Climate for the Asian Tiger Mosquito Aedes Albopictus: Recent Trends and Future Scenarios." *Journal of the Royal Society Interface* 9, no. 75 (2012): 2708-2717.

Campbell-Lendrum, D., L. Manga, M. Bagayoko and J. Sommerfeld. "Climate Change and Vector-Borne Diseases: What Are the Implications for Public Health Research and Policy?" *Philosophical Transactions of the Royal Society B-Biological Sciences* 370, no. 1665 (2015): 8.

CDC, "Chikungunya in the Caribbean", CDC: Centers for Disease Control and Prevention http://wwwnc.cdc.gov/travel/notices/watch/chikungunya-saint-martin (accessed Feb. 8 2014).

CDC, "Malaria Information and Prophylaxis, by Country [T]", Centers for Disease Control http://www.cdc.gov/malaria/travelers/country_table/t.html (accessed February 19 2015).

Coetzee, M., R. H. Hunt, R. Wilkerson, A. della Torre, M.B. Coulibaly and N.J. Besansky. "Anopheles Coluzzii and Anopheles Amharicus, New Members of the Anopheles Gambiae Complex." *Zootaxa* 3619, no. 3 (2013): 246-274.

Colacicco-Mayhugh, M. G., P. M. Masuoka and J. P. Grieco. "Ecological Niche Model of Phlebotomus Alexandri and P. Papatasi (Diptera: Psychodidae) in the Middle East." *International Journal of Health Geographics* 9, (2010): 2.

Cross, E. R. and K. C. Hyams. "The Potential Effect of Global Warming on the Geographic and Seasonal Distribution of Phlebotomus Papatasi in Southwest Asia." *Environmental Health Perspectives* 104, no. 7 (1996): 724-727.

Crowl, T. A., T. O. Crist, R. R. Parmenter, G. Belovsky and A. E. Lugo. "The Spread of Invasive Species and Infectious Disease as Drivers of Ecosystem Change." *Frontiers in Ecology and the Environment* 6, no. 5 (2008): 238-246.

DeVisser, M. H., J. P. Messina, N. J. Moore, D. P. Lusch and J. Maitima. "A Dynamic Species Distribution Model of *Glossina* Subgenus *Morsitans*: The Identification of Tsetse Reservoirs and Refugia." *Ecosphere* 1, no. 1 (2010). http://www.esajournals.org/doi/abs/10.1890/ES10-00006.1.

Dicko, A. H., R. Lancelot, M. T. Seck, L. Guerrini, B. Sall, M. Lo, M. J. B. Vreysen, T. Lefrancois, W. M. Fonta, S. L. Peck and J. Bouyer. "Using Species Distribution Models to Optimize Vector Control in the Framework of the Tsetse Eradication Campaign in Senegal." *Proceedings of the National Academy of Sciences of the United States of America* 111, no. 28 (2014): 10149-10154.

Dobrowski, S. Z., J. H. Thorne, J. A. Greenberg, H. D. Safford, A. R. Mynsberge, S. M. Crimmins and A. K. Swanson. "Modeling Plant Ranges over 75 Years of Climate Change in California, USA: Temporal Transferability and Species Traits." *Ecological Monographs* 81, no. 2 (2011): 241-257.

Dormann, C. F. "Promising the Future? Global Change Projections of Species Distributions." *Basic and Applied Ecology* 8, no. 5 (2007): 387-397.

Drake, J. M. and J. C. Beier. "Ecological Niche and Potential Distribution of Anopheles Arabiensis in Africa in 2050." *Malaria journal* 13, (2014): 213.

Duncombe, J., A. Clements, J. Davis, W. B. Hu, P. Weinstein and S. Ritchie. "Spatiotemporal Patterns of Aedes Aegypti Populations in Cairns, Australia: Assessing Drivers of Dengue Transmission." *Tropical Medicine & International Health* 18, no. 7 (2013): 839-849.

Elith, J. and J. R. Leathwick. "Species Distribution Models: Ecological Explanation and Prediction across Space and Time." In *Annual Review of Ecology Evolution and Systematics*, 40, 677-697. Palo Alto: Annual Reviews, 2009.

Erickson, R. A., K. Hayhoe, S. M. Presley, L. J. S. Allen, K. R. Long and S. B. Cox. "Potential Impacts of Climate Change on the Ecology of Dengue and Its Mosquito Vector the Asian Tiger Mosquito (Aedes Albopictus)." *Environmental Research Letters* 7, no. 3 (2012): 6.

Ermert, V., A. H. Fink, A. P. Morse and H. Paeth. "The Impact of Regional Climate Change on Malaria Risk Due to Greenhouse Forcing and Land-Use Changes in Tropical Africa." *Environmental Health Perspectives* 120, no. 1 (2012): 77-84.

Ernst, K. C., S. O. Adoka, D. O. Kowuor, M. L. Wilson and C. C. John. "Malaria Hotspot Areas in a Highland Kenya Site Are Consistent in Epidemic and Non-Epidemic Years and Are Associated with Ecological Factors." *Malaria Journal* 5, (2006): 78.

Estrada-Pena, A. and J. M. Venzal. "Climate Niches of Tick Species in the Mediterranean Region: Modeling of Occurrence Data, Distributional Constraints, and Impact of Climate Change." *Journal of Medical Entomology* 44, no. 6 (2007): 1130-1138.

Feria-Arroyo, Teresa P., Ivan Castro-Arellano, Guadalupe Gordillo-Perez, Ana L. Cavazos, Margarita Vargas-Sandoval, Abha Grover, Javier Torres, Raul F. Medina, Adalberto A. Perez de Leon and Maria D. Esteve-Gassent. "Implications of Climate Change on the Distribution of the Tick Vector Ixodes Scapularis and Risk for Lyme Disease in the Texas-Mexico Transboundary Region." *Parasites & Vectors* 7, (2014): 199.

Fielding, A. H. and P. F. Haworth. "Testing the Generality of Bird-Habitat Models." *Conservation Biology* 9, no. 6 (1995): 1466-1481.

Fischer, D., P. Moeller, S. M. Thomas, T. J. Naucke and C. Beierkuhnlein. "Combining Climatic Projections and Dispersal Ability: A Method for Estimating the Responses of Sandfly Vector Species to Climate Change." *Plos Neglected Tropical Diseases* 5, no. 11 (2011): 13.

Fischer, D., S. M. Thomas, M. Neteler, N. B. Tjaden and C. Beierkuhnlein. "Climatic Suitability of Aedes Albopictus in Europe Referring to Climate Change Projections: Comparison of Mechanistic and Correlative Niche Modelling Approaches." *Eurosurveillance* 19, no. 6 (2014): 34-46.

Fischer, D., S.M. Thomas, F. Niemitz, B. Reineking and C. Beierkuhnlein. "Projection of Climatic Suitability for Aedes Albopictus Skuse (Culicidae) in Europe under Climate Change Conditions." *Global and Planetary Change* 78, no. 1-2 (2011): 54-64.

Fisher, J. A. D., K. T. Frank and W. C. Leggett. "Dynamic Macroecology on Ecological Time-Scales." *Global Ecology and Biogeography* 19, no. 1 (2010): 1-15.

Foley, D. H., T. A. Klein, H. C. Kim, T. Brown, R. C. Wilkerson and L. M. Rueda. "Validation of Ecological Niche Models for Potential Malaria Vectors in the Republic of Korea." *Journal of the American Mosquito Control Association* 26, no. 2 (2010): 210-213.

Foley, D. H., T. A. Klein, H. C. Kim, W. J. Sames, R. C. Wilkerson and L. M. Rueda. "Geographic Distribution and Ecology of Potential Malaria Vectors in the Republic of Korea." *Journal of Medical Entomology* 46, no. 3 (2009): 680-692.

Forester, B. R., E. G. DeChaine and A. G. Bunn. "Integrating Ensemble Species Distribution Modelling and Statistical Phylogeography to Inform Projections of Climate Change Impacts on Species Distributions." *Diversity and Distributions* 19, no. 12 (2013): 1480-1495.

Fuller, D. O., M. L. Ahumada, M. L. Quinones, S. Herrera and J. C. Beier. "Near-Present and Future Distribution of Anopheles Albimanus in Mesoamerica and the Caribbean Basin Modeled with Climate and Topographic Data." *International Journal of Health Geographics* 11, (2012a): 12.

Fuller, D. O., M. S. Parenti, A. N. Hassan and J. C. Beier. "Linking Land Cover and Species Distribution Models to Project Potential Ranges of Malaria Vectors: An Example Using Anopheles Arabiensis in Sudan and Upper Egypt." *Malaria Journal* 11, (2012b): 11.

Futami, K., G. O. Dida, G. O. Sonye, P. A. Lutiali, M. S. Mwania, S. Wagalla, J. Lumumba, J. O. Kongere, S. M. Njenga and N. Minakawa. "Impacts of Insecticide Treated Bed Nets on Anopheles Gambiae S.L. Populations in Mbita District and Suba District, Western Kenya." *Parasites & Vectors* 7, (2014): 63.

Galvez, R., M. A. Descalzo, I. Guerrero, G. Miro and R. Molina. "Mapping the Current Distribution and Predicted Spread of the Leishmaniosis Sand Fly Vector in the Madrid Region (Spain) Based on Environmental Variables and Expected Climate Change." *Vector-Borne and Zoonotic Diseases* 11, no. 7 (2011): 799-806.

Garza, M., T. Patricia Feria Arroyo, E. A. Casillas, V. Sanchez-Cordero, C. Rivaldi and S. Sarkar. "Projected Future Distributions of Vectors of Trypanosoma Cruzi in North America under Climate Change Scenarios." *PLoS neglected tropical diseases* 8, no. 5 (2014): e2818.

Gething, P. W., A. P. Patil, D. L. Smith, C. A. Guerra, I. R. F. Elyazar, G. L. Johnston, A. J. Tatem and S. I. Hay. "A New World Malaria Map: Plasmodium Falciparum Endemicity in 2010." *Malaria Journal* 10, (2011): 16.

GFATM, "The Global Fund to Fight Aids, Tuberculosis and Malaria: Grant Scorecard (Tnz-M-13)" http://portfolio.theglobalfund.org/en/Grant/Index/TNZ-809-G11-M.

Giles, J. R., A. T. Peterson, J. D. Busch, P. U. Olafson, G. A. Scoles, R. B. Davey, J. M. Pound, D. M. Kammlah, K. H. Lohmeyer and D. M. Wagner. "Invasive Potential of Cattle Fever Ticks in the Southern United States." *Parasites & Vectors* 7, (2014): 11.

Gonzalez, C., A. Paz and C. Ferro. "Predicted Altitudinal Shifts and Reduced Spatial Distribution of Leishmania Infantum Vector Species under Climate Change Scenarios in Colombia." *Acta tropica* 129, (2014): 83-90.

Gonzalez, C., O. Wang, S. E. Strutz, C. Gonzalez-Salazar, V. Sanchez-Cordero and S. Sarkar. "Climate Change and Risk of Leishmaniasis in North America: Predictions from Ecological Niche Models of Vector and Reservoir Species." *Plos Neglected Tropical Diseases* 4, no. 1 (2010): 16.

Goyal, M. R. and E. W. Harmsen. "Evapotranspiration: Temperature Versus Elevation Relationships." In *Evapotranspiration: Principles and Applications for Water Management*, 1. Oakville: Apple Academic Press, Inc., 2014.

Graves, P. M., J. M. Ngondi, J. Hwang, A. Getachew, T. Gebre, A. W. Mosher, A. E. Patterson, E. B. Shargie, Z. Tadesse, A. Wolkon, R. Reithinger, P. M. Emerson and F. O. Richards. "Factors Associated with Mosquito Net Use by Individuals in Households Owning Nets in Ethiopia." *Malaria Journal* 10, (2011): 354.

Guichard, S., H. Guis, A. Tran, C. Garros, T. Balenghien and D.J. Kriticos. "Worldwide Niche and Future Potential Distribution of Culicoides Imicola, a Major Vector of Bluetongue and African Horse Sickness Viruses." *Plos One* 9, no. 11 (2014): e112491.

Hay, S., S. E. Randolph and D. J. Rogers. "Remote Sensing and Geographical Information Systems in Epidemiology." 47, (2000): 1-367.

Hernandez, P. A., I. Franke, S. K. Herzog, V. Pacheco, L. Paniagua, H. L. Quintana, A. Soto, J. J. Swenson, C. Tovar, T. H. Valqui, J. Vargas and B. E. Young. "Predicting Species Distributions in Poorly-Studied Landscapes." *Biodiversity and Conservation* 17, no. 6 (2008): 1353-1366.

Hongoh, V., L. Berrang-Ford, M. E. Scott and L. R. Lindsay. "Expanding Geographical Distribution of the Mosquito, Culex Pipiens, in Canada under Climate Change." *Applied Geography* 33, no. 1 (2012): 53-62.

IPCC. "Climate Change 2013: The Physical Science. Summary for Policymakers." 2013.

Kabula, B., Y. A. Derua, P. K. Tungu, D. J. Massue, E. Sambu, G. Stanley, F. W. Mosha and W. N. Kisinza. "Malaria Entomological Profile in Tanzania from 1950 to 2010: A Review of Mosquito Distribution, Vectorial Capacity and Insecticide Resistance." *Tanzania Journal of Health Research* 13, no. 5 (2011).

Karunamoorthi, K. "Impact of Global Warming on Vector-Borne Diseases: Implications for Integrated Vector Management." *Journal of Socialomics* 2, (2012): e113.

Kearney, M. R., A. Matzelle and B. Helmuth. "Biomechanics Meets the Ecological Niche: The Importance of Temporal Data Resolution." *J Exp Biol* 215, no. 6 (2012): 922-933.

Kearney, M. and W. Porter. "Mechanistic Niche Modelling: Combining Physiological and Spatial Data to Predict Species' Ranges." *Ecology Letters* 12, no. 4 (2009): 334-350.

Keith, S. A., T. J. Webb, K. Bohning-Gaese, S. R. Connolly, N. K. Dulvy, F. Eigenbrod, K. E. Jones, T. Price, D. W. Redding, I. P. F. Owens and N. J. B. Isaac. "What Is Macroecology?" *Biology Letters* 8, no. 6 (2012): 904-906.

Kelly-Hope, L. A., J. Hemingway and F. E. McKenzie. "Environmental Factors Associated with the Malaria Vectors Anopheles Gambiae and Anopheles Funestus in Kenya." *Malaria Journal* 8, (2009): 8.

Kerr, J., M. Kulkarni and A. Algar. "Integrating Theory and Predictive Modeling for Conservation Research." In *Predictive Species and Habitat Modeling in Landscape Ecology*, edited by C. Drew, Y.F. Wiersma and F. Huettmann, 9-28. New York: Springer, 2011.

Kerr, J. T., H. M. Kharouba and D. J. Currie. "The Macroecological Contribution to Global Change Solutions." *Science* 316, no. 5831 (2007): 1581-1584.

Kerr, J. T. and M. Ostrovsky. "From Space to Species: Ecological Applications for Remote Sensing." *Trends in Ecology & Evolution* 18, no. 6 (2003): 299-305.

Kharouba, H. M., A. C. Algar and J. T. Kerr. "Historically Calibrated Predictions of Butterfly Species' Range Shift Using Global Change as a Pseudo-Experiment." *Ecology* 90, no. 8 (2009): 2213-2222.

Kitau, J., R. M. Oxborough, P. K. Tungu, J. Matowo, R. C. Malima, S. M. Magesa, J. Bruce, F. W. Mosha and M. W. Rowland. "Species Shifts in the Anopheles Gambiae Complex: Do Llins Successfully Control Anopheles Arabiensis?" *Plos One* 7, no. 3 (2012): e31481.

Kitron, U. "Landscape Ecology and Epidemiology of Vector-Borne Diseases: Tools for Spatial Analysis." *Journal of Medical Entomology* 35, no. 4 (1998): 435-445.

Kolivras, K. N. "Changes in Dengue Risk Potential in Hawaii, USA, Due to Climate Variability and Change." *Climate Research* 42, no. 1 (2010): 1-11.

Kulkarni, M. A., R. E. Desrochers and J. T. Kerr. "High Resolution Niche Models of Malaria Vectors in Northern Tanzania: A New Capacity to Predict Malaria Risk?" *Plos One* 5, no. 2 (2010): e9396.

Lafferty, K. D. "The Ecology of Climate Change and Infectious Diseases." *Ecology* 90, no. 4 (2009): 888-900.

Larsen, D. A., P. Hutchinson, A. Bennett, J. Yukich, P. Anglewicz, J. Keating and T. P. Eisele. "Community Coverage with Insecticide-Treated Mosquito Nets and Observed Associations with All-Cause Child Mortality and Malaria Parasite Infections." *American Journal of Tropical Medicine and Hygiene* 91, no. 5 (2014): 950-958.

Legendre, P. and L. F. J. Legendre. *Numerical Ecology*. Amsterdam: Elsevier Science, 1998.

Leroux, S. J., M. Larrivee, V. Boucher-Lalonde, A. Hurford, J. Zuloaga, J. T. Kerr and F. Lutscher. "Mechanistic Models for the Spatial Spread of Species under Climate Change." *Ecological Applications* 23, no. 4 (2013): 815-828.

Levine, R. S., A. T. Peterson and M. Q. Benedict. "Geographic and Ecologic Distributions of the Anopheles Gambiae Complex Predicted Using a Genetic Algorithm." *American Journal of Tropical Medicine and Hygiene* 70, no. 2 (2004): 105-109.

Lobo, J. M., A. Jimenez-Valverde and R. Real. "Auc: A Misleading Measure of the Performance of Predictive Distribution Models." *Global Ecology and Biogeography* 17, no. 2 (2008): 145-151.

Luz, Paula M., Claudio J. Struchiner and Alison P. Galvani. "Modeling Transmission Dynamics and Control of Vector-Borne Neglected Tropical Diseases." *Plos Neglected Tropical Diseases* 4, no. 10 (2010): e761.

Malele, I. I. "Fifty Years of Tsetse Control in Tanzania: Challenges and Prospects for the Future." *Tanzania Journal of Health Research* 13, (2011).

MAP, "Tanzania (United Republic of)" http://www.map.ox.ac.uk/explore/countries/TZA/#!mosquito-vectors.

Marchant, T., D. Schellenberg, R. Nathan, J. Armstrong-Schellenberg, H. Mponda, C. Jones, Y. Sedekia, J. Bruce and K. Hanson. "Assessment of a National Voucher Scheme to Deliver Insecticide-Treated Mosquito Nets to Pregnant Women." *Canadian Medical Association Journal* 182, no. 2 (2010): 152-156.

Martens, P., R. S. Kovats, S. Nijhof, P. de Vries, M. T. J. Livermore, D. J. Bradley, J. Cox and A. J. McMichael. "Climate Change and Future Populations at Risk of Malaria." *Global Environmental Change-Human and Policy Dimensions* 9, (1999): S89-S107.

Martinez-Meyer, E. and A. T. Peterson. "Conservatism of Ecological Niche Characteristics in North American Plant Species over the Pleistocene-to-Recent Transition." *Journal of Biogeography* 33, no. 10 (2006): 1779-1789.

Matawa, F., K. S. Murwira and W. Shereni. "Modelling the Distribution of Suitable Glossina Spp. Habitat in the North Western Parts of Zimbabwe Using Remote Sensing and Climate Data." *Geoinformatics and Geostastics: An Overview*, no. S1 (2013).

Matemba, L. E., E. M. Fevre, S. N. Kibona, K. Picozzi, S. Cleaveland, A. P. Shaw and S. C. Welburn. "Quantifying the Burden of Rhodesiense Sleeping Sickness in Urambo District, Tanzania." *Plos Neglected Tropical Diseases* 4, no. 11 (2010): 10.

Matowo, J., C. M. Jones, B. Kabula, H. Ranson, K. Steen, F. Mosha, M. Rowland and D. Weetman. "Genetic Basis of Pyrethroid Resistance in a Population of Anopheles

Arabiensis, the Primary Malaria Vector in Lower Moshi, North-Eastern Tanzania." *Parasites & Vectors* 7, (2014): 274.

Mboera, L. E. G., B.K. Mayala, E. J. Kweka and H.D. Mazigo. "Impact of Climate Change on Human Health and Health Systems in Tanzania: A Review." *Tanzania Journal of Health Research* 13, (2011).

McCann, R. S., E. Ochomo, M. N. Bayoh, J. M. Vulule, M. J. Hamel, J. E. Gimnig, W. A. Hawley and E. D. Walker. "Reemergence of Anopheles Funestus as a Vector of Plasmodium Falciparum in Western Kenya after Long-Term Implementation of Insecticide-Treated Bed Nets." *American Journal of Tropical Medicine and Hygiene* 90, no. 4 (2014): 597-604.

McCord, P. F., J. P. Messina, D. J. Campbell and S. C. Grady. "Tsetse Fly Control in Kenya's Spatially and Temporally Dynamic Control Reservoirs: A Cost Analysis." *Applied Geography* 34, (2012): 189-204.

Meade, M. S. and R. Earickson. *Medical Geography*: Guilford Press, 2000.

Medlock, J. M., D. Avenell, I. Barrass and S. Leach. "Analysis of the Potential for Survival and Seasonal Activity of Aedes Albopictus (Diptera : Culicidae) in the United Kingdom." *Journal of Vector Ecology* 31, no. 2 (2006): 292-304.

Messina, J. P., N. J. Moore, M. H. DeVisser, P. F. McCord and E. D. Walker. "Climate Change and Risk Projection: Dynamic Spatial Models of Tsetse and African Trypanosomiasis in Kenya." *Annals of the Association of American Geographers* 102, no. 5 (2012): 1038-1048.

Mills, J. N., K. L. Gage and A. S. Khan. "Potential Influence of Climate Change on Vector-Borne and Zoonotic Diseases: A Review and Proposed Research Plan." *Environmental Health Perspectives* 118, no. 11 (2010): 1507-1514.

Moffett, A., N. Shackelford and S. Sarkar. "Malaria in Africa: Vector Species' Niche Models and Relative Risk Maps." *Plos One* 2, no. 9 (2007): e824.

Monahan, W. B. and M. W. Tingley. "Niche Tracking and Rapid Establishment of Distributional Equilibrium in the House Sparrow Show Potential Responsiveness of Species to Climate Change." *Plos One* 7, no. 7 (2012): e42097.

Moo-Llanes, D., C. N. Ibarra-Cerdena, E. A. Rebollar-Tellez, S. Ibanez-Bernal, C. Gonzalez and J. M. Ramsey. "Current and Future Niche of North and Central American Sand Flies (Diptera: Psychodidae) in Climate Change Scenarios." *Plos Neglected Tropical Diseases* 7, no. 9 (2013): e2421.

Morin, C. W. and A. C. Comrie. "Modeled Response of the West Nile Virus Vector Culex Quinquefasciatus to Changing Climate Using the Dynamic Mosquito Simulation Model." *International Journal of Biometeorology* 54, no. 5 (2010): 517-529.

Mtega, W. P. and B. Ronald. "The State of Rural Information and Communication Services in Tanzania: A Meta-Analysis" *International Journal of Information and Communication Technology Research* 3, no. 2 (2013).

Mughini-Gras, L., P. Mulatti, F. Severini, D. Boccolini, R. Romi, G. Bongiorno, C. Khoury, R. Bianchi, F. Montarsi, T. Patregnani, L. Bonfanti, G. Rezza, G. Capelli and L. Busani. "Ecological Niche Modelling of Potential West Nile Virus Vector Mosquito Species and Their Geographical Association with Equine Epizootics in Italy." *Ecohealth* 11, no. 1 (2014): 120-132.

Mutuku, F. M., M. N. Bayoh, A. W. Hightower, J. M. Vulule, J. E. Gimnig, J. M. Mueke, F. A. Amimo and E. D. Walker. "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, (2009): 19.

Naish, S., K. Mengersen, W. B. Hu and S. L. Tong. "Forecasting the Future Risk of Barmah Forest Virus Disease under Climate Change Scenarios in Queensland, Australia." *Plos One* 8, no. 5 (2013): 7.

NBS. *Tanzania in Figures 2012*. Tanzania: National Bureau of Statistics, 2013a.

NBS. *2012 Population and Housing Census*. Dar Es Salaam: National Bureau of Statistics, 2013b.

Ngonghala, C. N., S. Y. Del Valle, R. J. Zhao and J. Mohammed-Awel. "Quantifying the Impact of Decay in Bed-Net Efficacy on Malaria Transmission." *Journal of Theoretical Biology* 363, (2014): 247-261.

Nkya, T. E., I. Akhouayri, R. Poupardin, B. Batengana, F. Mosha, S. Magesa, W. Kisinza and J. David. "Insecticide Resistance Mechanisms Associated with Different Environments in the Malaria Vector Anopheles Gambiae: A Case Study in Tanzania." *Malaria Journal* 13, (2014): 28.

Obsomer, V., P. Defourny and M. Coosemans. "Predicted Distribution of Major Malaria Vectors Belonging to the Anopheles Dirus Complex in Asia: Ecological Niche and Environmental Influences." *Plos One* 7, no. 11 (2012): e50475.

Obsomer, V., N. Titeux, C. Vancustem, G. Duveiller, J. Pekel, S. Connor, P. Ceccato and M. Coosemans. "From Anopheles to Spatial Surveillance: A Roadmap through a Multidisciplinary Challenge." In *Anopheles Mosquitoes: New Insights into Malaria Vectors*, edited by Sylvie Manguin: CC BY, 2013.

Ogden, N. H., A. Maarouf, I. K. Barker, M. Bigras-Poulin, L. R. Lindsay, M. G. Morshed, C. J. O'Callaghan, F. Ramay, D. Waltner-Toews and D. F. Charron. "Climate Change and the Potential for Range Expansion of the Lyme Disease Vector Ixodes Scapularis in Canada." *International Journal for Parasitology* 36, no. 1 (2006): 63-70.

Ogden, N. H., L. St-Onge, I. K. Barker, S. Brazeau, M. Bigras-Poulin, D. F. Charron, C. M. Francis, A. Heagy, L. R. Lindsay, A. Maarouf, P. Michel, F. Milord, C. J. O'Callaghan, L. Trudel and R. A. Thompson. "Risk Maps for Range Expansion of the Lyme Disease Vector, Ixodes Scapularis, in Canada Now and with Climate Change." *International Journal of Health Geographics* 7, (2008): 15.

Pagel, J. and F. M. Schurr. "Forecasting Species Ranges by Statistical Estimation of Ecological Niches and Spatial Population Dynamics." *Global Ecology and Biogeography* 21, no. 2 (2012): 293-304.

Pavlovsky, E. N. *Natural Nidality of Transmissible Diseases: With Species Reference to the Landscape Epidemiology of Zooanthroponoses*: University of Illinois Press, 1966.

Pearson, R. G., C. J. Raxworthy, M. Nakamura and A. T. Peterson. "Predicting Species Distributions from Small Numbers of Occurrence Records: A Test Case Using Cryptic Geckos in Madagascar." *Journal of Biogeography* 34, no. 1 (2007): 102-117.

Pearson, R. G., W. Thuiller, M. B. Araujo, E. Martinez-Meyer, L. Brotons, C. McClean, L. Miles, P. Segurado, T. P. Dawson and D. C. Lees. "Model-Based Uncertainty in Species Range Prediction." *Journal of Biogeography* 33, no. 10 (2006): 1704-1711.

Perez-Heydrich, C., J.L. Warren, C.R. Burgert and M.E. Emch. *Guidelines on the Use of Dhs Gps Data: Dhs Spatial Analyst Reports 8*. Calverton, Maryland: USAID, 2013.

Peterson, A. T. "Ecologic Niche Modeling and Spatial Patterns of Disease Transmission." *Emerging Infectious Diseases* 12, no. 12 (2006): 1822-1826.

Peterson, A. T. and J. Shaw. "Lutzomyia Vectors for Cutaneous Leishmaniasis in Southern Brazil: Ecological Niche Models, Predicted Geographic Distributions, and Climate Change Effects." *International Journal for Parasitology* 33, no. 9 (2003): 919-931.

Phillips, S. J., R. P. Anderson and R. E. Schapire. "Maximum Entropy Modeling of Species Geographic Distributions." *Ecological Modelling* 190, no. 3-4 (2006): 231-259.

Porretta, D., V. Mastrantonio, S. Amendolia, S. Gaiarsa, S. Epis, C. Genchi, C. Bandi, D. Otranto and S. Urbanelli. "Effects of Global Changes on the Climatic Niche of the Tick Ixodes Ricinus Inferred by Species Distribution Modelling." *Parasites & Vectors* 6, (2013): 1-8.

Protopopoff, N., J. Matowo, R. Malima, R. Kavishe, R. Kaaya, A. Wright, P. A. West, I. Kleinschmidt, W. Kisinza, F. W. Mosha and M. Rowland. "High Level of Resistance in the Mosquito Anopheles Gambiae to Pyrethroid Insecticides and Reduced Susceptibility to Bendiocarb in North-Western Tanzania." *Malaria Journal* 12, (2013): 8.

Randall, D. A., R. A. Wood, S. Bony, R. Colman, T. Fichefet, J. Fyfe, V. Kattsov, A. Pitman, J. Shukla, J. Srinivasan, Stouffer, R. J., A. Sumi and K. E. Taylor. "Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change." In *Climate Models and Their Evaluation*, 589-662. United Kingdom: Cambridge University Press, 2007.

Randin, C. F., T. Dirnboeck, S. Dullinger, N. E. Zimmermann, M. Zappa and A. Guisan. "Are Niche-Based Species Distribution Models Transferable in Space?" *Journal of Biogeography* 33, no. 10 (2006): 1689-1703.

Rapacciuolo, G., D. B. Roy, S. Gillings and A. Purvis. "Temporal Validation Plots: Quantifying How Well Correlative Species Distribution Models Predict Species' Range Changes over Time." *Methods in Ecology and Evolution* 5, no. 5 (2014): 407-420.

Rastetter, E. B. "Validating Models of Ecosystem Response to Global Change." *Bioscience* 46, no. 3 (1996): 190-198.

Reisen, W. K. "Landscape Epidemiology of Vector-Borne Diseases." In *Annual Review of Entomology*, 55, 461-483. Palo Alto: Annual Reviews, 2010.

Reiter, P. "Climate Change and Mosquito-Borne Disease." *Environmental Health Perspectives* 109, (2001): 141-161.

Roberts, D. R. and A. Hamann. "Method Selection for Species Distribution Modelling: Are Temporally or Spatially Independent Evaluations Necessary?" *Ecography* 35, no. 9 (2012): 792-802.

Rochlin, I., D. V. Ninivaggi, M. L. Hutchinson and A. Farajollahi. "Climate Change and Range Expansion of the Asian Tiger Mosquito (Aedes Albopictus) in Northeastern USA: Implications for Public Health Practitioners." *Plos One* 8, no. 4 (2013): e60874.

Rogers, D. J. and S. E. Randolph. "The Global Spread of Malaria in a Future, Warmer World." *Science* 289, no. 5485 (2000): 1763-1766.

Roiz, D., M. Neteler, C. Castellani, D. Arnoldi and A. Rizzoli. "Climatic Factors Driving Invasion of the Tiger Mosquito (Aedes Albopictus) into New Areas of Trentino, Northern Italy." *Plos One* 6, no. 4 (2011): e14800.

Samy, A. M., L. P. Campbell and A. T. Peterson. "Leishmaniasis Transmission: Distribution and Coarse-Resolution Ecology of Two Vectors and Two Parasites in Egypt." *Revista Da Sociedade Brasileira De Medicina Tropical* 47, no. 1 (2014): 57-62.

Sax, D. F., R. Early and J. Bellemare. "Niche Syndromes, Species Extinction Risks, and Management under Climate Change." *Trends in Ecology & Evolution* 28, no. 9 (2013): 517-523.

Schafer, M. L. and J. O. Lundstrom. "The Present Distribution and Predicted Geographic Expansion of the Floodwater Mosquito Aedes Sticticus in Sweden." *Journal of Vector Ecology* 34, no. 1 (2009): 141-147.

Simon, Julie A., Robby R. Marrotte, Nathalie Desrosiers, Jessica Fiset, Jorge Gaitan, Andrew Gonzalez, Jules K. Koffi, Francois-Joseph Lapointe, Patrick A. Leighton, Lindsay R. Lindsay, Travis Logan, Francois Milord, Nicholas H. Ogden, Anita Rogic, Emilie Roy-Dufresne, Daniel Suter, Nathalie Tessier and Virginie Millien. "Climate Change and Habitat Fragmentation Drive the Occurrence of Borrelia Burgdorferi, the Agent of Lyme Disease, at the Northeastern Limit of Its Distribution." *Evolutionary Applications* 7, no. 7 (2014): 750-764.

Sinka, M. E., M. J. Bangs, S. Manguin, M. Coetzee, C. M. Mbogo, J. Hemingway, A. P. Patil, W. H. Temperley, P. W. Gething, C. W. Kabaria, R. M. Okara, T. Van Boeckel, H. C. J. Godfray, R. E. Harbach and S. I. Hay. "The Dominant Anopheles Vectors of Human Malaria in Africa, Europe and the Middle East: Occurrence Data, Distribution Maps and Bionomic Precis." *Parasites & Vectors* 3, (2010): 117.

Sobek-Swant, S., D. A. Kluza, K. Cuddington and D. B. Lyons. "Potential Distribution of Emerald Ash Borer: What Can We Learn from Ecological Niche Models Using Maxent and Garp?" *Forest Ecology and Management* 281, (2012): 23-31.

Stefani, A., I. Dusfour, M. C. B. Cruz, N. Dessay, A. K. R. Galardo, C. D. Galardo, R. Girod, M. S. M. Gomes, H. Gurgel, A. C. F. Lima, E. S. Moreno, L. Musset, M. Nacher, A. C. S. Soares, B. Carme and E. Roux. "Land Cover, Land Use and Malaria in the Amazon: A Systematic Literature Review of Studies Using Remotely Sensed Data." *Malaria Journal* 12, (2013): 192.

Stockwell, D. and D. Peters. "The Garp Modelling System: Problems and Solutions to Automated Spatial Prediction." *International Journal of Geographical Information Science* 13, no. 2 (1999): 143-158.

Sunday, J. M., A. E. Bates and N. K. Dulvy. "Thermal Tolerance and the Global Redistribution of Animals." *Nature Climate Change* 2, no. 9 (2012): 686-690.

Sunday, J. M., A. E. Bates, M. R. Kearney, R. K. Colwell, N. K. Dulvy, J. T. Longino and R. B. Huey. "Thermal-Safety Margins and the Necessity of Thermoregulatory Behavior across Latitude and Elevation." *Proceedings of the National Academy of Sciences of the United States of America* 111, no. 15 (2014): 5610-5615.

Sutherst, R. W. "Global Change and Human Vulnerability to Vector-Borne Diseases." *Clinical Microbiology Reviews* 17, no. 1 (2004): 136.

Sweeney, A. W., N. W. Beebe, R. D. Cooper, J. T. Bauer and A. T. Peterson. "Environmental Factors Associated with Distribution and Range Limits of Malaria Vector Anopheles Farauti in Australia." *Journal of Medical Entomology* 43, no. 5 (2006): 1068-1075.

Tabachnick, W. J. "Challenges in Predicting Climate and Environmental Effects on Vector-Borne Disease Episystems in a Changing World." *Journal of Experimental Biology* 213, no. 6 (2010): 946-954.

Temu, E. A., J. N. Minjas, M. Coetzee, R. H. Hunt and C. J. Shiff. "The Role of Four Anopheline Species (Diptera : Culicidae) in Malaria Transmission in Coastal Tanzania." *Transactions of the Royal Society of Tropical Medicine and Hygiene* 92, no. 2 (1998): 152-158.

Thuiller, W. "Biomod - Optimizing Predictions of Species Distributions and Projecting Potential Future Shifts under Global Change." *Global Change Biology* 9, no. 10 (2003): 1353-1362.

Tonnang, H. E. Z., R. Y. M. Kangalawe and P. Z. Yanda. "Predicting and Mapping Malaria under Climate Change Scenarios: The Potential Redistribution of Malaria Vectors in Africa." *Malaria Journal* 9, (2010): 111.

Trejo, I., E. Martinez-Meyer, E. Calixto-Perez, S. Sanchez-Colon and R. V. de la Torre. "Analysis of the Effects of Climate Change on Plant Communities and Mammals in Mexico." *Atmosfera* 24, no. 1 (2011): 1-14.

Ureta, C., E. Martinez-Meyer, H. R. Perales and E. R. Alvarez-Buylla. "Projecting the Effects of Climate Change on the Distribution of Maize Races and Their Wild Relatives in Mexico." *Global Change Biology* 18, no. 3 (2012): 1073-1082.

Van Asselen, S. and P. H. Verburg. "Land Cover Change or Land-Use Intensification: Simulating Land System Change with a Global-Scale Land Change Model." *Global Change Biology* 19, no. 12 (2013): 3648-3667.

Vanderelst, D. and N. Speybroeck. "An Adjusted Bed Net Coverage Indicator with Estimations for 23 African Countries." *Malaria journal* 12, (2013): 457-457.

Wandiga, S. O., M. Opondo, D. Olago, A. Githeko, F. Githui, M. Marshall, T. Downs, A. Opere, C. Oludhe, G. O. Ouma, P. Z. Yanda, R. Kangalawe, R. Kabumbuli, J. Kathuri, E. Apindi, L. Olaka, L. Ogallo, P. Mugambi, R. Sigalla, R. Nanyunja, T. Baguma and P. Achola. "Vulnerability to Epidemic Malaria in the Highlands of Lake Victoria Basin: The Role of Climate Change/Variability, Hydrology and Socio-Economic Factors." *Climatic Change* 99, no. 3-4 (2010): 473-497.

WB, "Data: Tanzania", The World Bank: IBRD and IDA http://data.worldbank.org/country/tanzania (accessed February 19 2015).

Wenger, S. J., N. A. Som, D. C. Dauwalter, D. J. Isaak, H. M. Neville, C. H. Luce, J. B. Dunham, M. K. Young, K. D. Fausch and B. E. Rieman. "Probabilistic Accounting of Uncertainty in Forecasts of Species Distributions under Climate Change." *Global Change Biology* 19, no. 11 (2013): 3343-3354.

West, P. A., N. Protopopoff, M. W. Rowland, M. J. Kirby, R. M. Oxborough, F. W. Mosha, R. Malima and I. Kleinschmidt. "Evaluation of a National Universal Coverage Campaign of Long-Lasting Insecticidal Nets in a Rural District in North-West Tanzania." *Malaria Journal* 11, (2012): 273.

White, P. and J. T. Kerr. "Contrasting Spatial and Temporal Global Change Impacts on Butterfly Species Richness During the 20th Century." *Ecography* 29, no. 6 (2006): 908-918.

WHO. *A Global Brief on Vector-Borne Diseases*: World Health Organization, 2014a. http://apps.who.int/iris/handle/10665/111008 (accessed May 13, 2015)

WHO, "United Republic of Tanzania: Who Statistical Profile", WHO and UN, 2015b http://www.who.int/gho/countries/tza.pdf?ua=1 (accessed March 22 2015).

WHO. "World Health Report 2008." (2008): 1-190.

WHO, "World Malaria Day, 25 April 2014", World Health Organization, 2015a http://www.who.int/campaigns/malaria-day/2014/event/en/ (accessed March 14 2015).

WHO. *World Malaria Report 2014*. Geneva: World Health Organization, 2014b.

Willis, K. J. and H. J. B. Birks. "What Is Natural? The Need for a Long-Term Perspective in Biodiversity Conservation." *Science* 314, no. 5803 (2006): 1261-1265.

Yamana, T. K. and E. A. B. Eltahir. "Incorporating the Effects of Humidity in a Mechanistic Model of Anopheles Gambiae Mosquito Population Dynamics in the Sahel Region of Africa." *Parasites & Vectors* 6, (2013): 235.

Yamana, T. K. and E. A. B. Eltahir. "Projected Impacts of Climate Change on Environmental Suitability for Malaria Transmission in West Africa." *Environmental Health Perspectives* 121, no. 10 (2013): 1179-1186.

Yoo, E. "Site-Specific Prediction of West Nile Virus Mosquito Abundance in Greater Toronto Area Using Generalized Linear Mixed Models." *International Journal of Geographical Information Science* 28, no. 2 (2014): 296-313.

Appendix A

Anopheles *occurrence records used in Chapter 2.*

Table A.1: List of the 56 geographically-unique *Anopheles* mosquito occurrence records used in Chapter 2. Corresponding references are provided in Appendix B.

Mosquito species	Year start	Year end	References	District	Specific area	Trap method(s)
An. arabiensis *An. funestus* s.l.	2001	2002	Kulkarni et al. 2006	Hai	Foo	Light traps
An. arabiensis *An. funestus* s.l.	2001	2002	Kulkarni et al. 2006	Hai	Sonu	Light traps
An. arabiensis *An. funestus* s.l.	2001	2002	Kulkarni et al. 2006	Hai	Kware	Light traps Pit Pyrethrum spray
An. arabiensis	2000	2000	McCall et al. 2001	Lower Moshi	Masingi	Resting catches
An. arabiensis *An. funestus* s.l.	2001	2002	Kulkarni et al. 2006	Hai	Rundugai	Light traps Pyrethrum spray Pit
An. gambiae s.l.	2000	2002	Graham et al. 2005	Muheza	Station	Experimental hut
An. gambiae s.s. *An. funestus* s.s.	2003	2003	Kweka et al. 2008	Muheza	Mkuzi	Light traps Pyrethrum spray
An. gambiae s.l. *An. funestus* s.l.	2003	2003	Meyrowitsch et al. 2011	Tanga	Kirare	Light traps
An. gambiae s.l. *An. funestus* s.l.	2001	2001	Simonsen et al. 2004	Tanga	Masaika	Light traps
An. gambiae s.l.	1999	1999	Kulkarni et al. 2007	Dodoma	Dodoma	Resting catches
An. gambiae s.l.	1999	1999	Kulkarni et al. 2007	Morogoro	Mvomero	Resting catches

An. funestus s.l.	2002	2002	Temu et al. 2007	Bagamoyo	Matimbwa	Light traps
An. funestus s.l.	2002	2002	Temu et al. 2007	Bagamoyo	Kongo	Light traps
An. gambiae s.l.	2002	2002	Mboera et al. 2006	Mpwapwa	Kibakwe	Light traps
An. gambiae s.l.	2002	2002	Mboera et al. 2006	Mpwapwa	Chogola	Light traps
An. gambiae s.l. *An. funestus* s.l.	2002	2002	Mboera et al. 2006	Iringa	Mangalali	Light traps Pyrethrum spray
An. funestus s.l. *An. gambiae* s.l.	2002	2002	Mboera et al. 2006	Iringa	Idodi	Light traps Pyrethrum spray
An. arabiensis *An. gambiae* s.s.	2003	2004	Kigadye et al. 2010	Rufiji	NA	Light traps
An. gambiae s.l.	2002	2002	Mboera et al. 2008	Iringa	Tosamaganga	Pyrethrum spray
An. gambiae s.l.	2002	2002	Mboera et al. 2008	Iringa	Makifu	Pyrethrum spray
An. funestus s.l. *An. gambiae* s.l.	2001	2003	Huho et al. 2012	Rufiji	Rufiji	Light traps
An. gambiae s.l.	2002	2002	Mboera et al. 2008	Iringa	Lulanzi	Pyrethrum spray
An. gambiae s.l.	2002	2002	Mboera et al. 2006	Iringa	Kilolo	Pyrethrum spray
An. gambiae s.l.	2001	2003	Killeen et al. 2007	Idete	Godawn	Light traps
An. gambiae s.l.	2000	2000	Schwartz and Koella 2002	Ifakara	Limaomao	Resting catches
An. gambiae s.l.	2001	2003	Killeen et al. 2007	Idete	Msumbiji	Light traps
An. gambiae s.l.	1999	1999	Pinto et al. 2007	Ifakara	Ifakara	Not specified
An. gambiae s.l.	2001	2003	Killeen et al. 2007	Namwawala	Videnge	Light traps
An. arabiensis *An. funestus* s.l. *An. gambiae* s.l.	1999	2000	Drakeley et al. 2003	Ifakara	Town centre	Light traps
An. arabiensis, *An. funestus* s.l. *An. gambiae* s.l.	1999	2000	Drakeley et al. 2003	Ifakara	East	Light traps
An. gambiae s.l.	2001	2003	Killeen et al. 2007	Ifakara	Namwawala Site B	Light traps

An. arabiensis *An. funestus* s.l. *An. gambiae* s.l.	1999	2000	Drakeley et al. 2003	Ifakara	South	Light traps
An. gambiae s.l.	2001	2003	Killeen et al. 2007	Ifakara	Namwawala Site A	Light traps
An. gambiae s.l.	2001	2003	Killeen et al. 2007	Kivukoni	Chikago	Light traps
An. gambiae s.l.	2001	2003	Killeen et al. 2007	Kivukoni	Ramba	Light traps
An. gambiae s.l.	2001	2003	Killeen et al. 2007	Mbingu	Mbingu	Light traps
An. gambiae s.l.	2001	2003	Killeen et al. 2007	Kivukoni	Butiama	Light traps
An. gambiae s.l.	2001	2003	Killeen et al. 2007	Kivukoni	Gezaulole	Light traps
An. gambiae s.l.	2000	2000	Schwartz and Koella 2002	Ifakara	Njagi	Resting catches
An. gambiae s.l.	2001	2003	Killeen et al. 2007	Mavimba	Manjole	Light traps
An. gambiae s.l.	2001	2003	Killeen et al. 2007	Kichangani	Mahongole	Light traps
An. gambiae s.l.	2001	2003	Killeen et al. 2007	Minepa	Kisakimbali	Light traps
An. gambiae s.l.	2001	2003	Killeen et al. 2007	Mavimba	Manjole	Light traps
An. gambiae s.l.	2001	2003	Killeen et al. 2007	Mchombe	Mchombe	Light traps
An. gambiae s.l.	2001	2003	Killeen et al. 2007	Lukolongo	Lukolongo	Light traps
An. gambiae s.l.	2001	2004	Huho et al. 2012	Ulanga	Ulanga	Light traps
An. gambiae s.l.	2001	2003	Killeen et al. 2007	Mkangawalo	Itongoa A	Light traps
An. gambiae s.l.	2001	2003	Killeen et al. 2007	Mkangawalo	Mkangawalo	Light traps
An. gambiae s.l.	2001	2003	Killeen et al. 2007	Lupiro	Ndoro	Light traps
An. gambiae s.l.	2001	2003	Killeen et al. 2007	Lupiro	Lupiro Kati	Light traps
An. gambiae s.l.	2001	2003	Killeen et al. 2007	Lupiro	Libaratula	Light traps
An. gambiae s.l.	2001	2003	Killeen et al. 2007	Lupiro	Madibila	Light traps
An. gambiae s.l.	2001	2003	Killeen et al. 2007	Mkangawalo	Mgudeni	Light traps
An. gambiae s.l.	2001	2003	Killeen et al. 2007	Kidugalo	Kidugalo	Light traps
An. gambiae s.l.	2001	2003	Killeen et al. 2007	Iragua	Igunda	Light traps
An. gambiae s.l.	1999	1999	Kulkarni et al. 2007	Mtwara	Mtwara	Resting catches

Appendix B

References for Appendix A.

Drakeley, C., D. Schellenberg, J. Kihonda, C. A. Sousa, A. P. Arez, D. Lopes, J. Lines, H. Mshinda, C. Lengeler, J. A. Schellenberg, M. Tanner and P. Alonso. "An Estimation of the Entomological Inoculation Rate for Ifakara: A Semi-Urban Area in a Region of Intense Malaria Transmission in Tanzania." *Tropical Medicine & International Health* 8, no. 9 (2003): 767-774.

Graham, K., M. H. Kayedi, C. Maxwell, H. Kaur, H. Rehman, R. Malima, C. F. Curtis, J. D. Lines and M. W. Rowland. "Multi-Country Field Trials Comparing Wash-Resistance of Permanet (Tm) and Conventional Insecticide-Treated Nets against Anopheline and Culicine Mosquitoes." *Medical and Veterinary Entomology* 19, no. 1 (2005): 72-83.

Huho, B. J., G. F. Killeen, H. M. Ferguson, A. Tami, C. Lengeler, J. D. Charlwood, A. Kihonda, J. Kihonda, S. P. Kachur, T. A. Smith and S. M. K. Abdulla. "Artemisinin-Based Combination Therapy Does Not Measurably Reduce Human Infectiousness to Vectors in a Setting of Intense Malaria Transmission." *Malaria Journal* 11, (2012): 12.

Kigadye, E. S. P., G. Nkwengulila, S. M. Magesa and S. Abdulla. "Diversity, Spatial and Temporal Abundance of Anopheles Gambiae Complex in the Rufiji River Basin, South-Eastern Tanzania." *Tanzania Journal of Health Research* 12, no. 1 (2010): 68-72.

Killeen, G. F., A. Tami, J. Kihonda, F. O. Okumu, M. E. Kotas, H. Grundmann, N. Kasigudi, H. Ngonyani, V. Mayagaya, R. Nathan, S. Abdulla, J. D. Charlwood, T. A. Smith and C. Lengeler. "Cost-Sharing Strategies Combining Targeted Public Subsidies with Private-Sector Delivery Achieve High Bednet Coverage and Reduced Malaria Transmission in Kilombero Valley, Southern Tanzania." *Bmc Infectious Diseases* 7, (2007): 121.

Kulkarni, M. A., E. Kweka, E. Nyale, E. Lyatuu, F. W. Mosha, D. Chandramohan, M. E. Rau and C. Drakeley. "Entomological Evalution of Malaria Vectors at Different Altitudes in Hal District, Northeastern Tanzania." *Journal of Medical Entomology* 43, no. 3 (2006): 580-588.

Kulkarni, M. A., R. Malima, F. W. Mosha, S. Msangi, E. Mrema, B. Kabula, B. Lawrence, S. Kinung'hi, J. Swilla, W. Kisinza, M. E. Rau, J. E. Miller, J. A. Schellenberg, C. Maxwell, M. Rowland, S. Magesa and C. Drakeley. "Efficacy of Pyrethroid-Treated Nets against Malaria Vectors and Nuisance-Biting Mosquitoes in Tanzania in Areas with Long-Term Insecticide-Treated Net Use." *Tropical Medicine & International Health* 12, no. 9 (2007): 1061-1073.

Kweka, E. J., A. M. Mahande, W. M. M. Nkya, C. Assenga, E. E. Lyatuu, E. Nyale, F. W. Mosha, S. B. Mwakalinga and E. A. Temu. "Vector Species Composition and Malaria Infectivity Rates in Mkuzi, Muheza District, North-Eastern Tanzania." *Tanzania Journal of Health Research* 10, no. 1 (2008): 46-9.

Mboera, L. E. G., S. M. Magesa and F. Molteni. "Indoors Man-Biting Mosquitoes and Their Implication on Malaria Transmission in Mpwapwa and Iringa Districts, Tanzania." *Tanzania Health Research Bulletin* 8, no. 3 (2006): 141-4.

Mboera, L. E. G., M. L. Kamugisha, S. F. Rumisha, W. N. Kisinza, Kesheni P. Senkoro and Andrew Y. Kitua. "Malaria and Mosquito Net Utilisation among Schoolchildren in Villages with or without Healthcare Facilities at Different Altitudes in Iringa District, Tanzania." *African Health Sciences* 8, no. 2 (2008): 114-119.

McCall, P. J., F. W. Mosha, K. J. Njunwa and K. Sherlock. "Evidence for Memorized Site-Fidelity in Anopheles Arabiensis." *Transactions of the Royal Society of Tropical Medicine and Hygiene* 95, no. 6 (2001): 587-590.

Meyrowitsch, D. W., E. M. Pedersen, M. Alifrangis, T. H. Scheike, M. N. Malecela, S. M. Magesa, Y. A. Derua, R. T. Rwegoshora, E. Michael and P. E. Simonsen. "Is the Current Decline in Malaria Burden in Sub-Saharan Africa Due to a Decrease in Vector Population?" *Malaria Journal* 10, (2011): 9.

Schwartz, A. and J. C. Koella. "Melanization of Plasmodium Falciparum and C-25 Sephadex Beads by Field-Caught Anopheles Gambiae (Diptera : Culicidae) from Southern Tanzania." *Journal of Medical Entomology* 39, no. 1 (2002): 84-88.

Simonsen, P. E., D. W. Meyrowitsch, D. A. Mukoko, E. M. Pedersen, M. N. Malecela-Lazaro, R. T. Rwegoshora, J. H. Ouma, N. Masese, W. G. Jaoko and E. Michael. "The Effect of Repeated Half-Yearly Diethylcarbamazine Mass Treatment on Wuchereria Bancrofti Infection and Transmission in Two East African Communities with Different Levels of Endemicity." *American Journal of Tropical Medicine and Hygiene* 70, no. 1 (2004): 63-71.

Temu, E. A., J. N. Minjas, N. Tuno, H. Kawada and M. Takagi. "Identification of Four Members of the Anopheles Funestus (Diptera : Culicidae) Group and Their Role in Plasmodium Falciparum Transmission in Bagamoyo Coastal Tanzania." *Acta Tropica* 102, no. 2 (2007): 119-125.