

Developing a risk map of malaria transmission for East Africa

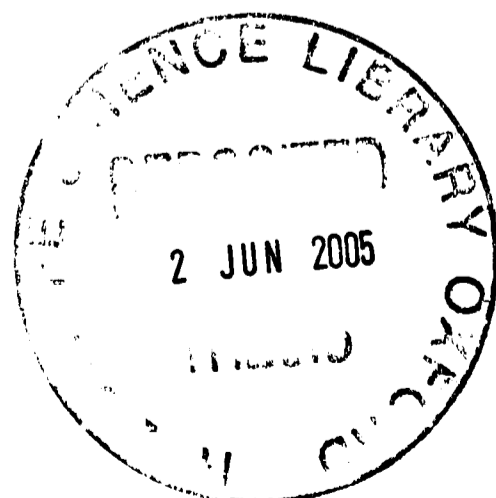
by

Judith A Omumbo

Green College, University of Oxford

**Thesis submitted in partial fulfilment of the requirements
for the degree of Doctor of Philosophy
Department of Zoology
University of Oxford**

**Michaelmas Term
November, 2004**



Acknowledgements

It is a pleasure to thank the many people who have made this work possible. Firstly I thank the Wellcome Trust, UK for financial support for my studies through the award of a Prize studentship (#060063). I also acknowledge the financial support of the International Development Research Centre, Canada and the South Africa Medical Research Council towards data collection. Information used in this study includes data produced through funding from NASA's Earth Observing System Pathfinder Program. The data are provided by the Earth Observing Data and Information System (EODIS), Distributed Active Archive Center (DAAC) at the Goddard Space Flight Centre. I also thank Fred Snijders (FAO ARTEMIS) for permission to use Meteosat. CCD data.

This thesis is the culmination of work carried out over the past 6 years beginning with the MARA team to whom I am indebted. Thanks all of you. I want also to thank the generous colleagues in malaria research and control in Kenya, Uganda, Tanzania, Ethiopia and Sudan for their support, for sparing time to facilitate my data collection, for sharing unpublished data and their hospitality during visits to each country. In particular I am grateful to Dr. Omer, Dr. Dukeen and Dr. Munira of the Sudanese National Malaria Administration, Dr. de Savigny and Rose Lusinde formerly of the Tanzania Essential Health Interventions Project (TEHIP), Albert Kilian (MoH, Uganda), Hugh Reyburn and Chris Drakeley (Gates Malaria Program, London School of Hygiene and Tropical Medicine), Chris Curtis and Ilona Carneiro (LSHTM), Dr. Talisuna and Dr. Langi of the Epidemiological Surveillance Division, MoH, Uganda and Dr. Ouma, Dr. Rapuoda and Mr Kwanya Agak of the Division of Malaria Control, MoH, Kenya, Dr. V. Marsh, Dr. T. Mwangi, Dr. D. Zurovac and Professor Bob Snow of the KEMRI/Wellcome Trust Research Programme, Kenya. The study would not have been possible without your contributions, asanteni sana.

I am deeply indebted to my supervisors Professor Bob Snow, Dr. Simon Hay and Professor David Rogers. I greatly appreciate the considerable amount of

time you spent critically reviewing what seemed like hundreds of drafts. In particular I thank Prof. Snow for keeping things moving, prompt and thorough reading of drafts and for your valuable comments on all aspects of this work. David thanks for all those hours of help with programming. Without you all, this work would not have been possible. Thanks too to Professor Sarah Randolph for your advice and encouragement and Dr. Willy Wint for being able to fix anything.

A special thank you is due to friends in Nairobi and in the UK and to colleagues in the Department of Zoology who made the past four years fun and provided a stimulating environment in which to learn and grow. So Bee, Nyami, Cee, Nyawira, Smoo, Sylvia, Andy, Dayo, Mike, Jana, Allan, Lina, Claudio, Bekky, Liz and Pippi, thanks for the visits, laughter, music, and for helping to keep life balanced. Jane, I thank you for providing a home away from home and for allowing me to enjoy the warm hearth of No. 6 for all those months.

Finally I wish to express my loving appreciation to my family. Mami and Baba, I have benefited greatly over the years from your wisdom and teaching about life's choices. Thank you for your unwavering dedication to all of us. Lucy, Laura, Jim and Joe, you made it all worth it.

Abstract

Background: The distribution of malaria in sub-Saharan Africa is determined largely by climatic influences on the development and survival of *P. falciparum* and its Anopheline vectors. This inter-relationship has been exploited in developing a limited number of predictive maps of malaria's distribution but these climate maps have limitations. Climate alone does not fully describe the complex dynamics of transmission and, in particular, human influences such as urbanization and the use of widespread anti-malarial interventions. The improved accuracy and validation of solely climate-driven maps relies on the availability of robust malariometric training data. To date, such data have been scarce. This study redresses several deficiencies of existing malaria maps for Africa through the collation of an extensive database of empirical *P. falciparum* prevalence data, the investigation of the relationship between prevalence and a widely-used climate-driven map, an assessment of the influence of urbanization on prevalence and finally, through the use of empirical training data to develop an improved malaria map for Kenya, Tanzania and Uganda.

Methods: An extensive published and grey-literature search was conducted between 1996 and 2004 and identified 2003 *P. falciparum* prevalence surveys conducted among childhood populations across East Africa between 1927 and 2003. Stringent criteria were applied to select the best sample data; only randomly sampled community-based surveys, surveys with samples ≥ 50 children, surveys conducted between 1980-2004 and children aged 0-14 years, and surveys which were spatially and temporally unique. The selected data were used to investigate the association between *P. falciparum* prevalence and a fuzzy logic climatic suitability (FCS) map of malaria transmission, the

effect of urbanization on prevalence and to train Fourier-processed multi-temporal climate surrogate data derived from meteorological satellites in order to predict prevalence for un-sampled areas. Using discriminant analysis, the top ten climatic predictor variables that distinguished best between 4 categories of malaria prevalence (0-<5, 5-<25%, 25-<75% & >=75%) were selected and these used to develop a predictive transmission map.

Results: Three hundred and thirty *P. falciparum* prevalence surveys met the criteria for inclusion in subsequent analyses. A weak, but significant, association ($kappa=0.367$; $p<0.0001$) was found between the categorical definitions of malaria prevalence and FCS criteria. Urbanization was shown to have a significant limiting effect on transmission. Studies from urban areas (N=37) had 18% ($p=<0.001$) lower *P. falciparum* prevalence than rural area studies (N=292; 1 study was unclassified). No urban area surveys had a *P. falciparum* prevalence >75%. Environmental predictors that best discriminated between the four categories of malaria prevalence included surrogates of moisture availability (normalised difference vegetation index) and temperature (land surface temperature). The overall accuracy of the predictive model compared with the training data was 62% and a high level of agreement was noted between observed prevalence and predictions ($kappa=0.477$; $tau=0.495$).

Conclusions: Climate continues to be an important driver of transmission intensity in East Africa; however, maps that rely entirely upon climate cannot predict the plethora of human and local ecological factors that coincidentally modify transmission. At least 28% of East Africa's population live in urban areas. The trend towards urbanization is increasing and this will have important impacts on the distribution of malaria. The study

highlights the need for improved data relating to malaria, including more accurate spatial definitions of population settlement and the coverage of anti-malarial interventions. More accurate, higher resolution spatial definitions of malariometric data will be required for future high spatial resolution malaria maps. There is a need for the continued use of predictive mapping as more extensive empirical malariometric databases are unlikely to be available in the near future. This study demonstrates the capacity for improvement of predictive malaria mapping. With better descriptions of the distribution of populations at risk it should be possible to improve the rationalisation of interventions and evaluate their impacts.

Table of Contents

Title page	i
Acknowledgements	ii
Abstract	iv
List of figures	xi
List of tables.....	xiv
List of abbreviations	xvii
Chapter 1: The epidemiology and cartography of malaria in sub-Saharan Africa	1
1.1. Background	2
1.2. The scope of the thesis.....	6
1.3. The biology of <i>P. falciparum</i> malaria.....	9
1.3.1. The life cycle of <i>Plasmodium</i>	9
1.3.2. Anopheline vectors of human malaria in SSA.....	12
1.3.3. The life cycle of the <i>Anopheles</i> mosquito	12
1.3.4. Behavioural characteristics of the <i>Anopheles</i> mosquito.....	13
1.4. Clinical malaria	14
1.5. The malaria burden.....	19
1.6. Challenges for malaria control in SSA countries	21
1.7. Measuring infectious disease transmission.....	22
1.7.1. The Basic Reproduction Number of a disease R_0	23
1.8. Describing the endemicity of malaria.....	25
1.8.1. Definition of endemicity.....	25
1.8.2. Stable and unstable malaria	26
1.8.3. Seasonality of malaria transmission.....	27
1.8.4. Epidemic malaria	28
1.9. Indicators of malaria endemicity: Incidence, prevalence, the force of infection and the Parasite Ratio.....	31
1.9.1. Measuring the Parasite Ratio	32
1.9.2. The classification of endemicity based on the Parasite Ratio.....	34
1.9.3. The relationship between the Entomological Inoculation Rate (EIR) and the Parasite Ratio	35

1.10. Determinants of <i>Plasmodium falciparum</i> endemicity	37
1.10.1. Temperature	38
1.10.2. Altitude	39
1.10.3. Rainfall and humidity	41
1.10.4. Human factors influencing transmission	42
1.10.4.1. Population movement, migration and displacement.....	43
1.10.4.2. Urbanization	44
1.10.4.3. Irrigation, water resources management and rice cultivation, land use and agriculture	46
1.10.4.4. Socio-economic and behavioural factors	47
1.10.4.5. Effects of malaria interventions.....	48
1.11. Malaria models for mapping malaria risk in sub-Saharan Africa	49
1.11.1. Historical global maps of malaria risk	49
1.11.2. Climate data and remote sensing applied to vector-borne disease mapping	51
1.11.3. Mapping the geographical distribution of <i>Anopheles</i> in SSA.....	53
1.11.3.1. Mapping the relative abundance of <i>Anopheles</i> using climate data.....	54
1.11.3.2. Predicting malaria vector species distribution using satellite imagery.....	55
1.11.4. Modelling the seasonality of malaria	58
1.11.4.1. The duration of malaria transmission seasons in SSA.....	58
1.11.4.2. Predicting the onset of malaria seasons in Kenya using clinical data and satellite imagery	59
1.11.5. Modelling malaria risk: how much malaria?	61
1.11.5.1. A fuzzy logic climate suitability model of malaria transmission	61
1.11.5.2. A logistic regression model of malaria for West Africa	64
1.11.5.3. Malaria transmission model for Kenya: estimating ‘at risk’ populations	66
1.11.5.4. Predicting the EIR using satellite imagery	68
1.11.6. Modelling malaria under changing climate scenarios	70
1.12. Discussion of limitations of current malaria maps for SSA	72

Chapter 2: Materials and methods	77
2.1. Background	78
2.2. The data entry proforma	79
2.3. Methods of data collection	80
2.4. Data geo-referencing	83
2.5. Data entry and verification	85
2.6. Data selection and descriptive analysis of data	86
2.6.1. Overall data identification	86
2.6.2. Selection criteria for data in Kenya, Tanzania and Uganda	90
2.7. Construction of model attribute data	98
2.7.1. General principles	98
2.7.2. Satellite sensor derived meteorological surrogate data	99
2.7.2.1. The Normalised Difference Vegetation Index (NDVI)	101
2.7.2.2. Land Surface Temperature (LST)	102
2.7.2.3. Middle Infra-red Reflectance (MIR)	102
2.7.2.4. Cold Cloud Duration (CCD)	103
2.7.2.5. Ordination of time series data	103
2.7.2.6. Temporal Fourier analysis of RS data	104
2.7.3. Altitude	105
2.7.4. Urbanization	105
2.7.5. Ecological features from Africover	106
2.8. Statistical methods	107
2.8.1. Principles of discriminant analysis	109
2.8.2. Predictor variable extraction and manipulation	110
2.8.3. Assessing the accuracy of the predicted model	112
 Chapter 3: Publication 1: <i>Plasmodium falciparum</i> parasite prevalence in East Africa: A review of empirical data 1927-2003	 115
 Chapter 4: Publication 2: The relationship between the <i>Plasmodium falciparum</i> parasite ratio in childhood and climate estimates of malaria transmission in Kenya	 116

Chapter 5: Publication 3: The influence of urbanization on measures of <i>Plasmodium falciparum</i> infection prevalence in East Africa	117
Chapter 6: Publication 4: Updating historical maps of malaria transmission intensity in East Africa using remote sensing	118
Chapter 7: Publication 5: Modelling malaria risk in East Africa at high spatial resolution	119
Chapter 8: Conclusions, limitations and future work	137
8.1. Summary	138
8.2. The limitation of the <i>P. falciparum</i> prevalence training data	140
8.3. A comparison of previous East African malaria maps	145
8.5. Future aims for mapping malaria and suggested steps to make improvements	153
References	155
Appendix: The data collection proforma	187

List of figures

Chapter 1: Figures

Figure 1.1:	The life cycle of malaria parasites	11
Figure 1.2:	Relative risks of infection, morbidity, severe disease and Mortality risks among a population aged 0-15 years located in a stable endemic area of the Kenyan coast	15
Figure 1.3:	The direct, indirect and consequential public health effects of <i>Plasmodium falciparum</i> malaria in Africa	19
Figure 1.4:	A conceptual framework of the dynamics malaria transmission	38
Figure 1.5:	The effect of temperature on parasite development and vector development and survival	39
Figure 1.6:	Altitude ranges of highland malaria epidemics in eastern African countries	41
Figure 1.7:	The geographical limits of malaria.....	50
Figure 1.8:	The geographical distribution of malaria infections.....	50
Figure 1.9:	Sites where members of the <i>Anopheles gambiae</i> complex have been observed mapped against total annual rainfall	54
Figure 1.10:	Distributions of five sibling species of the <i>Anopheles gambiae</i> complex in Africa, predicted from temporal Fourier-processed satellite data	57
Figure 1.11:	Predicting the seasonality of malaria in Kenya.....	60
Figure 1.12:	Fuzzy logic model of climatic suitability for stable <i>P. falciparum</i> transmission in SSA	63
Figure 1.13:	Predicted prevalence of <i>P. falciparum</i> in children aged 2 to 10 years for West Africa	66
Figure 1.14:	Modelled predictions of stable malaria endemicity for Kenya.....	67
Figure 1.15:	Satellite-derived predictions of EIR in Africa.	69

Chapter 2: Figures

Figure 2.1:	MARA organogram and map of regional data collection centres.....	79
Figure 2.2:	MARA data entry screens.....	86
Figure 2.3:	Distribution of parasitological survey data from the East and Horn of Africa between 1907 and 2003 according to dates of surveys.....	88
Figure 2.4:	The spatial extents of surveys included in the analysis of parasite prevalence pre- 1980.....	96
Figure 2.5:	The spatial extents of surveys included in the analysis of parasite prevalence post 1980.....	97

Chapter 3: Figures

Figure 1a:	The spatial distribution of <i>P. falciparum</i> parasite surveys undertaken between 1927 and 1979 against first-level UN approved administrative boundaries for Kenya, Tanzania and Uganda.....	115
Figure 1b:	The spatial distribution of <i>P. falciparum</i> parasite surveys undertaken between 1980 and 2003 against first-level UN approved administrative boundaries for Kenya, Tanzania and Uganda.....	115

Chapter 4: Figures

Figure 1:	MARA Fuzzy Suitability Class (FCS) values categorised into three categories in Kenya showing distribution of selected parasite survey data points.....	116
Figure 2:	<i>Plasmodium falciparum</i> parasite prevalence among children under the age of 15 years surveyed in 217 spatially independent surveys by FCS values extracted for the same spatial areas.....	116
Figure 3:	Box plot of distribution of <i>Plasmodium falciparum</i> parasite prevalence (%) according to three FCS categories.....	116

Chapter 5: Figures

Figure 1:	Distribution of parasite prevalence surveys by malaria ecological zone in East Africa	117
Figure 2:	Distribution of parasite prevalence data from East Africa according to rural vs. urban classification in low and moderate-intense malaria risk areas	117

Chapter 6: Figures

Figure 1:	Relationship between land surface temperature, altitude and historical classifications of malaria transmission intensity	118
Figure 2a:	Historical map of malaria transmission intensity.....	118
Figure 2b:	Map of discriminant analysis predictions of malaria transmission intensity	118
Figure 2c:	Distribution of misclassified pixels in relation to historical malaria transmission map	118

Chapter 7: Figures

Figure 1:	Distribution of PR surveys according to ecological strata	134
Figure 2:	Urban forced prediction with ecological zone stratification.....	135
Figure 3:	Comparison of malaria maps available for East Africa.....	136

Chapter 8: Figures

Figure 8.1a:	Historical map of malaria distribution for Kenya, Uganda and Tanzania	143
Figure 8.1b:	MARA fuzzy climate suitability model	143
Figure 8.1c:	HIMAL epidemic risk model.....	143
Figure 8.1d:	Satellite-derived predictions of EIR for East Africa	143
Figure 8.1e:	Modelled predictions of stable endemicity for Kenya	143
Figure 8.1f:	Satellite-derived predictions of the malaria seasonality for Kenya	143

Figure 8.1g:	Discriminant analysis predictions of malaria transmission intensity in East Africa using satellite-derived surrogates of climate	143
Figure 8.1h:	Empirical model of malaria transmission in East Africa	143
Figure 8.2:	Comparisons between historical classifications of malaria And available maps of malaria distribution for East Africa	147
Figure 8.3:	Comparisons between historical classifications of malaria and empirical maps of malaria distribution	148
Figure 8.3:	Boxplot of PR by malaria prevalence category	152

List of tables

Chapter 1: Tables

Table 1.1:	Summary of behavioural preferences of malaria vector species.....	14
Table 1.2:	Elements of R_0 for malaria	24
Table 1.3:	Classification of malaria.....	35

Chapter 2: Tables

Table 2.1:	Parasite surveys report source descriptive data for Kenya, Tanzania, Uganda, Ethiopia and Sudan.....	89
Table 2.2:	Selection criteria for parasite prevalence survey data in East Africa and description of selected surveys pre-1980	93
Table 2.3:	Selection criteria for parasite prevalence survey data in East Africa and description of selected surveys between 1980 and 2003	94
Table 2.4:	Diagram of a contingency matrix.....	113

Chapter 3: Tables

Table 1:	Decade frequency of community sampled surveys of <i>P. falciparum</i> infection prevalence in East Africa	115
Table 2:	Description of selected parasite prevalence surveys in East Africa	115

Table 3:	First-level administrative boundary median <i>P. falciparum</i> parasite prevalence data for East Africa: 1927-2003	115
----------	---	-----

Chapter 4: Tables

Table 1a:	Agreement between parasite prevalence categories and spatially congruent categorical values of the FCS using a 0.75 FCS threshold for stable endemic malaria	116
-----------	--	-----

Table 1b:	Agreement between parasite prevalence categories and spatially congruent categorical values of the FCS using a 0.6 FCS threshold for stable endemic malaria	116
-----------	---	-----

Chapter 5: Tables

Table 1:	Selection criteria for parasite prevalence survey data in East Africa and description of selected surveys.....	117
----------	--	-----

Table 2:	Spatial resolution of parasitological surveys in Kenya, Uganda and Tanzania	117
----------	---	-----

Table 3:	Malaria ecozone classification for parasite survey data in East Africa according to urban and rural coverage.....	117
----------	---	-----

Table 4:	The mean, median and distribution of parasite prevalence between urban and rural surveys undertaken in two malaria ecological zones in East Africa.....	117
----------	---	-----

Chapter 6: Tables

Table 1:	Correlation coefficients of independent variables.....	118
----------	--	-----

Table 2:	Classification matrix for training data and predicted pixels.....	118
----------	---	-----

Table 3:	Agreement between historical and predicted map of malaria transmission intensity	118
----------	--	-----

Table 4:	Ten most discriminating RS variables by rank	118
----------	--	-----

Chapter 7: Tables

Table 1a:	Classification matrix for urban-forced prediction stratified by ecozone: Ecozone 1	133
-----------	--	-----

Table 1b:	Classification matrix for urban-forced prediction stratified by ecozone: Ecozone 2	133
-----------	---	-----

Chapter 8: Tables

Table 8.1a:	Agreement between Butler, 1959 & Omumbo <i>et al.</i> , 2002	150
Table 8.1b:	Agreement between Butler, 1959 & Craig <i>et al.</i> , 1999	150
Table 8.1c:	Agreement between Butler, 1959 & Snow <i>et al.</i> , 1998	150
Table 8.1d:	Agreement between Craig <i>et al.</i> , 1999 & Snow <i>et al.</i> , 1998	150

List of abbreviations

AVHRR	Advanced Very High Resolution Radiometer
CBS	Central Bureau of Statistics
CCD	Cold cloud duration
CQ	Chloroquine
DA	Discriminant analysis
DALY	Disability-Adjusted Life Year
DDT	Dichlorodiphenyltrichloroethane
DEM	Digital elevation model
EIR	Entomological Inoculation Rate
EMR	Electromagnetic radiation
EMS	Electromagnetic spectrum
FAO	Food and Agriculture Organization
FCS	Fuzzy climate suitability index
GCM	General Circulation Model
GIS	Geographic Information Systems
GRID	Global Resources Information Database
GLM	Generalised linear models
GLR	Generalised linear regression
GLOBE	Global Land One-kilometre Base Elevation
GPS	Global Positioning Systems
GRUMP	Global Rural-Urban Mapping Project
HRPT	High resolution picture transmission
ILRI	International Livestock Research Institute
IPCC	Intergovernmental Panel on Climate Change
IPT	Intermittent presumptive treatment
IRS	Insecticide residual spraying
LAC	Local area coverage
LST	Land surface temperature
MARA/ARMA	Mapping Malaria Risk in Africa/Atlas du Risque de la Malaria en Afrique
MIR	Middle infra red reflectance
MoH	Ministry of Health
MVC	Maximum value composite
NDVI	Normalised Difference Vegetation Index
NGO	Non-governmental Organisation
NOAA	National Oceanic and Atmospheric Administration
PR	Parasite Ratio
RBM	Roll Back Malaria
RMSE	Root mean square error
SALB	Second level administrative boundaries
SI	Spatially interpolated
SSA	sub-Saharan Africa
SSS	Satellite sensor systems
TFA	Temporal Fourier analysis
UNDP	United Nations Development Programme
UNEP	United Nations Environment Project
UNICEF	United Nations Children's Emergency Fund
USGS	United States Geological Survey
WHO	World Health Organization

CHAPTER 1:

The epidemiology and cartography of malaria in sub-Saharan Africa

1.1. Background

The malaria parasite, *Plasmodium* is a unicellular protozoan. Of the four species of the genus *Plasmodium* that cause malaria in humans, *P. falciparum* is responsible for the greatest part of malaria's global morbidity and mortality. *P. falciparum* is also the most widely distributed *Plasmodium* species in tropical Africa, eastern Asia, Oceania and the Amazon compared with temperate zones. *P. vivax*, localised to northern Africa, is the second most widespread in Africa. *P. malariae* and *P. ovale* occur across sub-Saharan Africa (SSA) but their distribution is more focal (Lysenko & Beljaev, 1969). Phylogenetic studies date the origin of *P. falciparum* at least 40,000 years ago (Hume *et al.*, 2003; Joy *et al.*, 2003). A dramatic expansion of African *P. falciparum*, accompanying the development of agriculture 6000 years ago, has been suggested (Coluzzi & Bradley, 1999). The malaria parasite co-evolved with humans, as they migrated out of Africa, becoming adapted to the new environments they occupied (Coluzzi *et al.*, 1979).

Today *Plasmodium falciparum* remains the dominant malaria parasite species of the tropics and sub-tropics affecting 42% of the world's population resident in 83 countries (WHO, International travel and health, situation as on 1 January 2003. Geneva: WHO pg 193). Over two billion people are potentially exposed to infection (Hay *et al.*, 2004). Three hundred to five hundred million cases are reported annually, 90% of them being from SSA (WHO, 1994). Although all-cause mortality may be declining in many areas of SSA, studies suggest that malaria death rates may be rising (Snow *et al.*, 2003). In East and Southern Africa, for example, malaria mortality has increased from ~6.5 per 1000 child-years in 1982-1989 to 11.9 per 1000 child-years

during 1990-1998 (Korenromp *et al.*, 2003). There are multiple contributing factors to this trend that continues despite local efforts to control the disease. What is clear is that the dynamics of transmission are complex and effective control of malaria requires a multi-faceted attack on several points in the transmission cycle at the same time.

Sir Ronald Ross' and Giovanni Batista Grassi's description in 1898 of the complete transmission cycle of *Plasmodium* in avian (Ross, 1928) and human malaria respectively (described in Fantini, 1999) was a significant milestone in the history of malaria control. For the first time, there were hopes of breaking the transmission cycle through interruption of human - vector contact. On the basis of mathematical models of transmission dynamics, the forefathers of malaria control suggested that by widespread elimination of mosquito vectors, transmission could be reduced to below a threshold value such that malaria eradication could be achieved (Ross, 1911; MacDonald, 1956; Macdonald, 1957).

The malaria eradication programme began with extensive geographical reconnaissance to ensure total coverage for insecticide residual house spraying (IRS), and this baseline mapping work was pivotal to the success of eradication where it proved successful (Lepeš, 1974). Equipped with maps, the insecticide dichlorodiphenyltrichloroethane (DDT) and an effective anti-malarial drug, chloroquine (CQ), the dream of eradication began to be realised. The end of the Second World War saw the eradication of malaria from southern Europe, Russia and the more temperate areas of Asia (Guerin *et al.*, 2002). Ross and MacDonald's mathematical models quantified the significant impact of these vector control

programmes. Spurred on by the success of eradication in temperate countries, the World Health Organization's (WHO) 8th World Health Assembly adopted a Global Malaria Eradication Campaign (World Health Organization, 1951). The results of pilot projects were disappointing and the campaign deemed unsustainable in countries of highly intense malaria transmission (Spielman, 1993). By the 1970s, it was generally accepted that eradication would not be achievable in these areas and that instead, control would be more realistic (Litsios *et al.*, 2002).

The Roll Back Malaria (RBM) initiative is a global partnership of donor agencies, the private sector, and Non Governmental Organisations (NGOs) in malaria endemic countries. The initiative, set up in October 1998, was co-founded by the United Nations Development Programme (UNDP), the United Nations Children's Emergency Fund (UNICEF), The World Bank and the WHO. The overall objective of RBM is to halve the global burden of malaria by the year 2010 through multiple sustained interventions (Nabarro & Tayler, 1998; Remme *et al.*, 2001). These strategies include; prompt access to effective treatment, access to insecticide treated nets (ITNs), prevention and control of malaria during pregnancy through intermittent presumptive treatment (IPT) and effective response to epidemics and emergencies (www.rbm.who.int). The emphasis is on the development of interventions that are adapted to local needs and supportive of the health sector development strategy adopted by governments in several SSA countries.

A key failure of the malaria eradication era in the tropics and sub-tropics was the lack of recognition of the geographical differences in the epidemiology of malaria and in vector behaviour (Coluzzi, 1994). The global strategy assumed that a uniform

eradication approach would be effective over a wide range of malaria ecologies and the importance of these differences was only recognised in hindsight. Current control approaches require a renewed examination of the geographical determinants of malaria. New spatial models of malaria risk should form the basis of decision-making for interventions, definitions of target populations and estimations of resource needs. Although malaria distribution maps are recognised as an important epidemiological tool, reliable empirical maps are lacking for most of SSA (Snow *et al.*, 1996). Recent advances in disease mapping using Geographic Information Systems (GIS) coupled with an increased access to improved geo-referenced data on correlates of malaria transmission have rekindled an interest in such epidemiological investigations (Hay, 2000; Snow *et al.*, 1996). In particular, the relationship between climate and vector-borne disease transmission has been exploited in mapping the distribution of these diseases (Hay *et al.*, 2000a; Rogers *et al.*, 2002). Currently available climate-based maps, however, are limited in that they consider only a few of the factors that determine transmission, leaving out key covariates such as urbanization and the effects of malaria control interventions.

The precise global morbid and fatal burden of malaria remains uncertain (Snow *et al.*, 2004). The oft cited statements “three hundred to five hundred million cases of malaria reported globally each year” and “90% of the *Plasmodium falciparum* burden borne by SSA” (World Health Organization, 1994) are at best based upon imperfect data and at worst wild speculation. Recently more informed approaches to defining the Pan African burden of *P. falciparum* have been developed using semi-quantitative models of the spatial extent of infection risk, population distributions and morbid or fatal risks derived from epidemiological studies (Snow *et al.*, 1999a; 2003).

The African continent supports a diverse set of ecological conditions that influence the likelihood and intensity of malaria parasite transmission. Hay *et al.* (2000b) have shown that empirical measures of entomological inoculation rates vary from as little as one new infection per person every few years to several hundred per person per night. In order to characterise the burden of malaria effectively, one must begin with an understanding of the relationship between malaria infection and disease outcome.

This thesis aims to provide some new approaches to defining the cartography of malaria risk in East Africa by employing high spatial resolution remotely sensed data, more comprehensive information on covariates of transmission and better empirical data upon which the models can be developed. The thesis is anchored in the belief that without better descriptions of the spatial extents of quantifiable malaria risk, future iterations of the contribution of malaria to the Disability-Adjusted Life Year (DALY) attributed to malaria will be less informed and less precise and targeted investment in local needs for control less efficient.

1.2. The scope of the thesis

The present chapter, Chapter 1, considers the biology of *P. falciparum*, the principal malaria parasite in SSA, the biology of the parasite's principal vectors and the consequences of *P. falciparum* infection in humans. The chapter then explores how malariologists characterise the parasite's transmission and the existing empirical evidence of biotic and abiotic factors that influence transmission. The chapter ends with a summary of how the determinants of *P. falciparum* transmission have been

recently modelled to map malaria infection risk in Africa, highlighting some of the methodological issues and opportunities for new approaches.

Chapter 2 describes the methods used to capture empirical data, understand their limitations and match them to remotely sensed and other domain environmental data in order to develop predictive models of risk. The rigorous selection criteria adopted to identify model input data and principal statistical models used are also described in detail.

Chapter 3 provides a detailed description of the selected parasite survey data between 1927 and 2003 across East Africa. A simple time-series description of the data is presented by 2nd level administrative boundaries for the East Africa region to highlight the paucity of empirical information from some areas and the contrasting plethora of information from others. The temporal data are also interpreted in the light of changing survey focus and possible changes in malaria endemicity. This work is presented as a publication in the *East African Medical Journal* entitled “*Plasmodium falciparum* parasite prevalence in East Africa: A review of empirical data 1927-2003 (2004).”

Chapter 4 considers a widely cited parameter used during recent presentations of malaria in Africa by RBM, WHO and UNICEF. The Fuzzy Climate Suitability (FCS) index was developed to describe the likelihood of stable *P. falciparum* transmission in 1999 in Africa (Craig *et al.*, 1999). The continental maps have been used by several authors to describe the spatial extent of malaria and the consequent disease burdens. Surprisingly, the FCS has never been compared with empirical, spatially

matched parasitological data. This work is presented as a paper published in the *Malaria Journal*, 3: Article N° 17 entitled “The relationship between the *Plasmodium falciparum* parasite ratio in childhood and climate estimates of malaria transmission in Kenya,” (Omumbo *et al.*, 2004). The results suggest that climate-driven factors are important but other factors might well impinge on the reliability of climate-based data to develop malaria risk maps.

Chapter 5 explores urbanization, one of the major possible influences on recent efforts to map malaria risk. Differences in Entomological Inoculation Rates (EIR) between urban and rural communities have been described previously (Hay *et al.*, 2000b; Robert *et al.*, 2003) but few systematic reviews of the effects of urbanization on the parasite rate have been undertaken. This work is presented as a paper published in *Acta Tropica* entitled “The influence of urbanization on measures of *Plasmodium falciparum* infection prevalence in East Africa (Omumbo *et al.*, 2005)”. The paper also describes some of the limitations of current definitions of urban extents for use in geographic information systems (GIS) analyses but clearly demonstrates the effects of urbanization on the prevalence of *P. falciparum* infection within different ecological strata. The findings emphasize further the need to include covariates other than climate in definitions of malaria transmission.

Chapter 6 provides an example of how new high spatial resolution remotely sensed correlates of climate can be used to map the historical extent of malaria endemicity in East Africa. The chapter introduces the application of remotely sensed imagery in malaria risk mapping and is presented as a paper published in *Photogrammetric Engineering and Remote Sensing*, 68: 161-166, entitled “Updating historical maps of

malaria transmission intensity in East Africa using remote sensing (Omumbo *et al.*, 2002).” The chapter provides an introduction to the next chapter which exploits the use of empirical parasitological data in the model development rather than reliance upon the historical expert opinion maps.

Chapter 7 represents the culmination of the preceding chapters using carefully defined selection criteria for input model data on parasite prevalence, urbanization and other non-climatic factors and newly acquired, high spatial resolution remotely sensed data on climate. The chapter employs discriminant analysis techniques to identify key predictors of endemicity classes useful for public health definitions of disease risk and control options. The work is presented as a paper entitled “Modelling malaria risk in East Africa at high spatial resolution” submitted to *Tropical Medicine and International Health*. A large part of the detailed methodology is presented in Chapter 2.

Finally Chapter 8 provides a summary of the thesis identifying some of the weaknesses and areas that demand further investigation.

1.3. The biology of *P. falciparum* transmission

1.3.1. The life cycle of Plasmodium

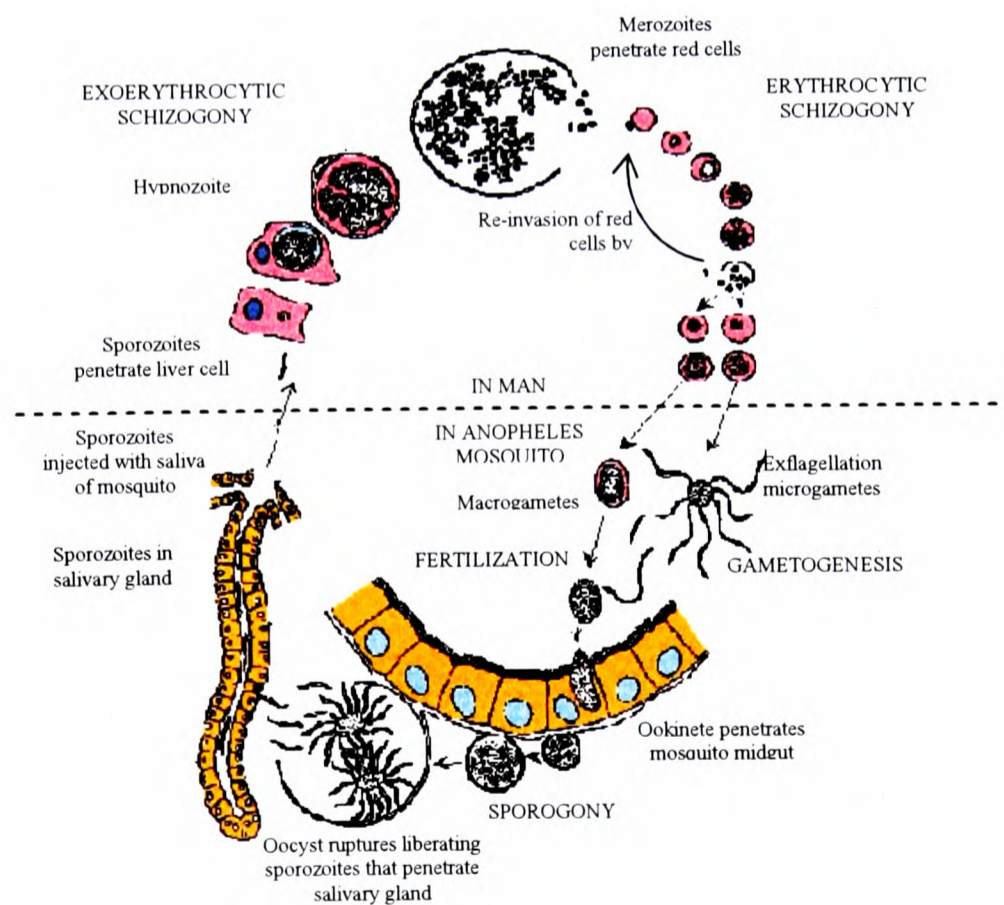
There are four main species of the protozoan *Plasmodium* that cause malaria in humans; *Plasmodium falciparum*, *P. vivax*, *P. ovale* and *P. malariae*. The development of *Plasmodium* takes place in two cycles, a sexual cycle within the

mosquito and an asexual cycle passed in the human host (Figure 1.1). The female *Anopheles* mosquito requires blood for the maturation of her eggs and person-to-person transmission of malaria occurs through the bite of the *Anopheles*. Infection of the human host begins with the inoculation of the asexual stage of the parasite, the sporozoite, from an infected mosquito into the blood stream of a susceptible host. This cycle of blood feeding and egg-laying is known as the *gonotrophic cycle*. After thirty minutes to four hours, sporozoites leave the circulating blood and migrate to the host's liver cells (hepatocytes). An asexual replication process, termed *exoerythrocytic schizogony*, follows. Liver schizonts are produced and these multiply rapidly producing between 10,000 and 30,000 merozoites per hepatocyte. This period, before symptoms are evident, lasts about a week. *P. vivax* and *P. ovale* are able to form a dormant liver stage, the hypnozoite that release merozoites into the peripheral blood system at a later point in time. In this way they can cause relapsing malaria.

The merozoites then invade the red blood cells (RBCs). The early RBC (erythrocytic) stage of the parasite is known as the trophozoite. Trophozoites form schizonts within which between 16-32 merozoites develop and these are released into the blood circulation. Some of the merozoites develop into macrogametes and microgametes (female and male sexual forms respectively). The gametocytes are ingested during a blood meal. The rest of the merozoites invade other erythrocytes, multiplying within them and causing their rupture. The erythrocytes rupture in synchronised 48-hour cycles in *P. falciparum*, *P. vivax* and *P. ovale* infections and in 72 hours cycles for *P. malariae*. Bouts of fever occur when the erythrocytes rupture in response to the release of toxins into the blood stream.

Macro- and microgametes mate in the mosquito midgut and the resulting zygote develops into a motile ookinete. The ookinetes traverse the midgut membrane to lodge in the mid-gut outer wall as oocysts. Over a period of 10-14 days the oocysts undergo asexual multiplication forming thousands of sporozoites. The duration of this process, known as *sporogony*, depends on the *Plasmodium* species and ambient temperature. For example, optimum temperatures for sporogony for *P. falciparum* are between 25-30°C while *P. vivax* prefers lower temperatures of about 25°C (Section 1.10.1). The sporozoites break out of the oocyst and migrate to the mosquito's salivary gland. Here they remain infective for the duration of the mosquito's life or are inoculated into a human host.

Figure 1.1. The life cycle of malaria parasites in the human and mosquito



1.3.2. Anopheline vectors of human malaria in SSA

The subfamily Anophelinae belongs to the family Culicidae of the order Diptera and comprises 3 genera; *Chagisa*, *Bironella* and *Anopheles*. *Anopheles* is the only genus known to be involved in the transmission of human malaria. There are about 430 known species of *Anopheles*. Seventy of these are vectors of malaria and only forty are of medical importance (Service & Townson, 2002). Vectors in areas of intense transmission are able to maintain high levels of abundance, have a high susceptibility to infection and a high longevity (Macdonald, 1957).

Anopheles gambiae sensu lato is the most dominant vector within the Afrotropical and Ethiopian faunal region (Boyd, 1930). It is a seven species complex comprising *Anopheles gambiae* sensu stricto, *An. arabiensis*, *An. melas*, *An. merus*, *An. quadriannulatus*, *An. quadriannulatus* Species B and *An. bwambae* (Coluzzi, 1984). *An. gambiae* s.s. and *An. arabiensis* are found in 70% of the African continent and *An. funestus* is the next most widely distributed vector in tropical Africa (Gillies & Coetzee, 1987).

1.3.3. The life cycle of the *Anopheles* mosquito

The development of the *Anopheles* mosquito (egg, larval and pupal stages) takes place in water and lasts about a week. The adults mate within 2 days of their emergence and females can travel up to 4 kilometres from the breeding site in search of vertebrate hosts (Gillies & De Meillon, 1968). The duration of the gonotrophic cycle (Section 1.3.1) is temperature dependent, lasting longer in cooler climates. In

tropical countries, mosquitoes take a blood meal every 2-3 days and complete 4-5 gonotrophic cycles in their lifetime (Service & Townson, 2002). The longevity of the mosquito is a key determinant of infection risk among exposed and susceptible human populations as described by the basic reproduction number, R_0 (Section 1.7.1).

1.3.4. Behavioural characteristics of the Anopheles mosquito

Key determinants of mosquito distribution are their feeding, resting and habitat preferences. Vectors that prefer to feed on humans are referred to as anthropophilic or anthropophagic. The Human Blood Index (HBI) gives an estimate of the anthropophilia of a vector and is the proportion of mosquitoes sampled that contain human blood (Section 1.7.1). In general, the vectors of African malaria bite at night (Service & Townson, 2002) and after feeding, some vectors rest in shady areas outside houses. Such vectors are termed exophilic. Others are endophilic and remain indoors after feeding. Anthropophagic and endophilic vectors are dominant around human settlements while those with a preference for biting animals (zoophilic) are abundant where livestock is kept. The behaviour traits, however, are not absolute and a zoophilic vector may feed on humans if animal hosts are not readily available. Behavioural preferences differ between the main *Anopheles* species in SSA. These differences are relevant for targeting interventions for their control and determine the relative abundance of sibling species (Table 1.1).

Table 1.1: Summary of behavioural preferences of malaria vector species

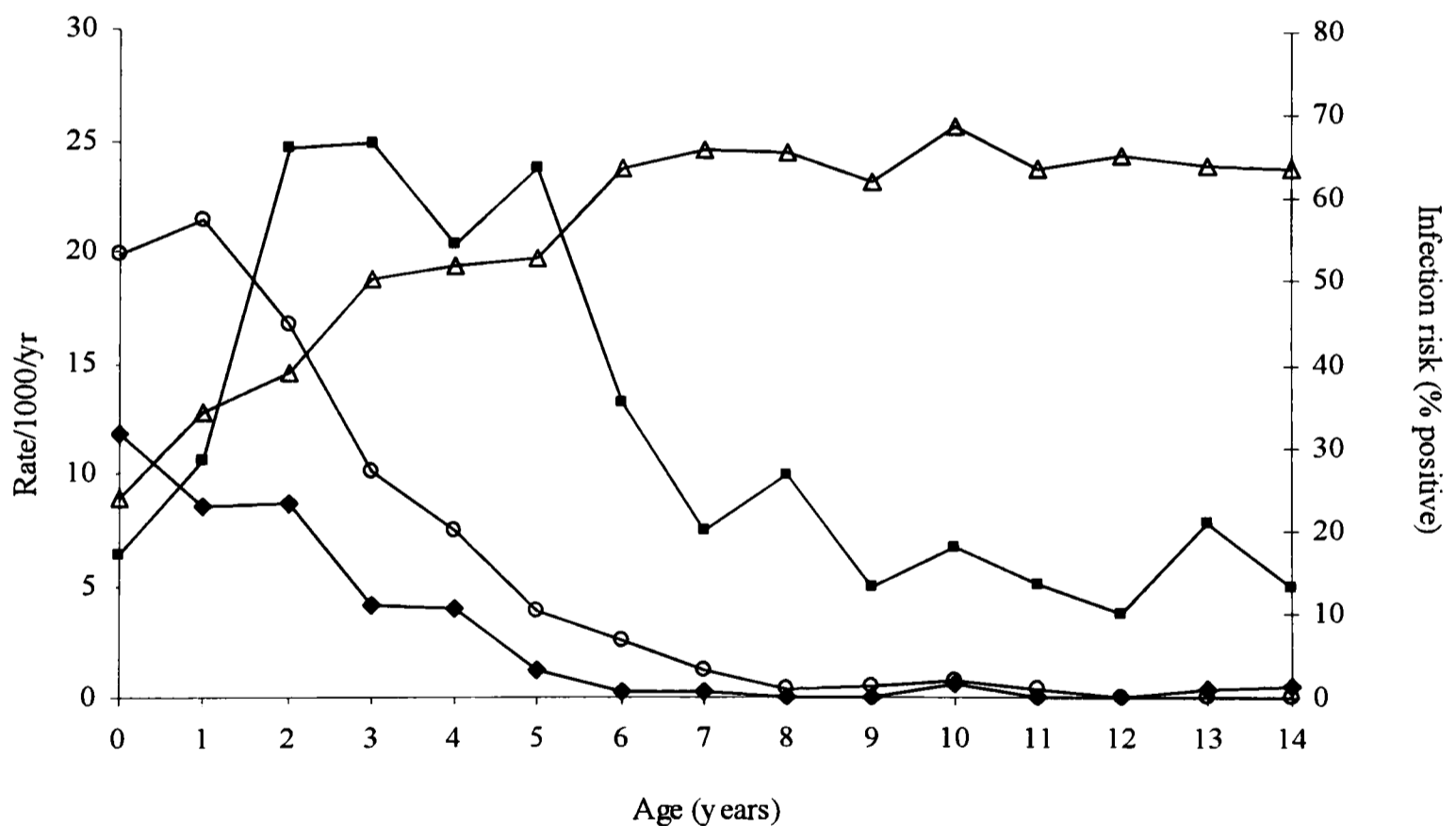
Species	Habitat preferences	Feeding preference	Resting
<i>An. gambiae</i> s.s	Permanent fresh water pools Humid areas	Anthropophagic	Indoors
<i>An. arabiensis</i>	Temporary fresh water pools Arid areas	More readily feeds on cattle than on humans	Outdoors
<i>An. merus</i>	Salt water breeder	Zoophilic Anthropophagic in absence of other hosts	Outdoors
<i>An. melas</i>	Salt water breeder	Zoophilic Anthropophagic in absence of other hosts	Outdoors
<i>An. quadriannulatus</i>	Fresh water pools	Zoophilic	Outdoors
<i>An. bwambae</i>	Mineral waters	Zoophilic	Outdoors
<i>An. funestus</i> Giles	Shaded marshes, streams, rivers & swamps	Anthropophagic	Indoors

1.4. Clinical malaria

It has long been recognised that the relationships between *P. falciparum* infection, morbidity and disease outcome are complex. What we do know is that individuals born into areas of stable *P. falciparum* transmission frequently move between periods of being infected and remaining uninfected. Most individuals will, at some stage in their lives, develop an overt clinical response to an infection, often manifested as a febrile event associated with difficulties in breathing, diarrhoea and vomiting. A comparatively small proportion of these clinical events may progress to severe pathological clinical states, including acidosis, anaemia and altered consciousness. These may naturally resolve or the patient may survive through medical intervention. However, even with optimal clinical management, case fatalities for conditions such as cerebral malaria are often in excess of 15% (Marsh *et al.*, 1995).

The comparative risks of infection, mild morbidity, severe complicated disease and mortality during childhood have been recorded during a series of studies at Kilifi, on the Kenyan Coast. The study population comprised children aged 0–15 years who could expect to receive about 10 new infections each year from the local vectors. The risks of severe malaria warranting admission to hospital and malaria deaths declined during early childhood (<5 years), while the frequency of mild, clinical attacks fell some time later (Figure 1.2).

Figure 1.2: Relative risks of infection (triangles), morbidity (squares), severe disease (circles) and mortality risks (diamonds) among a population aged 0-15 years located in a stable endemic area of the Kenyan Coast (Snow & Gilles, 2002)



Since the risk of infection remained approximately constant throughout childhood, it is reasonable to assume that these age-related findings are associated with the development of functional immunity. It appears that at least three levels of immune acquisition occur: initial protection against severe and fatal outcomes, followed by protection against mild, self-limiting clinical disease; much later, during adulthood, individuals develop an ability to regulate peripheral infection (Snow & Gilles, 2002)

Of relevance to disease control is whether the cumulative risks of disease or death decline with declining parasite exposure. Clearly, at one end of the transmission spectrum where the risk of infection is very low, immunity will be acquired slowly and all age groups are likely to be at risk of both disease and death following infection. Under these conditions the relative risk of a particular disease outcome will simply be a function of the chance encounters with the parasite. Among communities experiencing a wide range of stable, moderate to high intensity transmission conditions, early acquisition of functional immunity becomes important. Recent studies in Africa among children admitted to hospital with severe, life-threatening disease show a marked decline in severe disease with increasing age in areas in which the local prevalence of asymptomatic infection is >70%. In areas with a local prevalence <30%, malaria occurs with similar frequency throughout childhood (Trape *et al.*, 1987; Seboxa & Snow, 1997; Snow *et al.*, 1997; Armengaud *et al.*, 1962; Marsh & Snow, 1999; Snow & Marsh, 2002; Gernaat *et al.*, 1998). These observations support the view that the development of functional immunity may depend upon an ill-defined amount of parasite exposure from birth. Moreover, severe malarial anaemia tends to dominate the clinical picture in areas of high transmission, while

cerebral malaria assumes increasingly greater importance in areas of less intense transmission (Marsh & Snow, 1999).

In areas of high transmission intensity, it is uncertain whether artificially reducing the rate of natural parasite exposure, through sustained vector control or personal protection, will lead to new epidemiological conditions typical of intermediate transmission (Snow *et al.*, 1994; Snow & Marsh, 2002; Smith *et al.*, 2001a; 2001b; Trape & Rogier, 1996; Molineaux, 1988). It seems likely that the result would be a change in age-specific risks of disease outcomes resulting in a higher mean age of severe disease and death. Less certain is whether there would be an overall change in cumulative disease or risk of death throughout childhood.

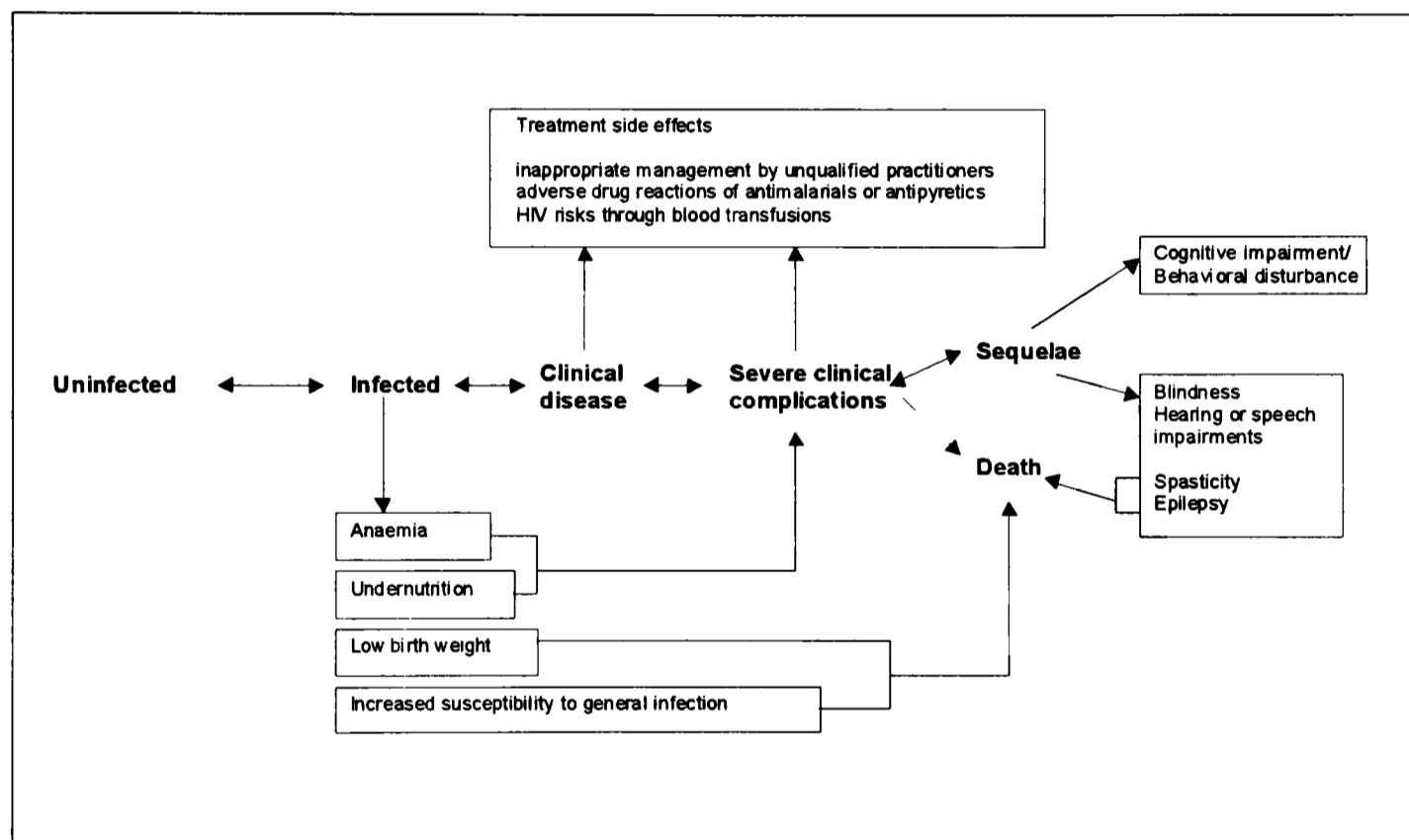
In areas where parasite transmission is high, a reduction in transmission is initially marked by a reduction in disease across all age groups (because a reduced challenge is met with a high level of immunity). This has been shown clearly during short-term studies of insecticide-treated bed nets or indoor residual house-spraying (IRS) in Africa (Lengeler *et al.*, 1998). Children born into the population, thereafter, grow up facing the new reduced level of challenge and presumably develop an immune profile and disease experience similar to that of children born in other areas with similarly lower challenge. The net result may be that, with time, the severe disease and, possibly, malaria-specific death rates re-adjust to those of the new reduced intensity transmission conditions. Conversely, interventions in areas where transmission intensity is already low could provide conditions in which the relationship between parasite exposure, immunity and disease result in a low risk of serious disease. Such

ecological conditions represent the fringes of stable transmission and may be subject to epidemics.

In addition to the morbidity and mortality directly attributed to *P. falciparum* there are other consequential and indirect effects on mortality linked to each step of the infection and disease process (for review see Snow *et al.*, 2003) . Chronic, sub-clinical infections cause anaemia or may lead to under-nutrition which in turn may increase susceptibility to severe clinical outcomes of subsequent infection with malaria or other pathogens. Despite a poor understanding of the precise mechanisms of pathology (Menendez, 1995), the morbid outcomes of malaria infection during pregnancy have also been well described (Brabin, 1983; Guyatt & Snow, 2001a, 2001b; Steketee *et al.*, 1996; 2001). In endemic settings in Africa, pregnant women experience relatively little malaria-specific morbidity (*e.g.* fever illness) but do have increased risk of infection and higher density parasitaemia leading to anaemia and placental sequestration of the parasite. Maternal anaemia has been shown to be an important contributor to maternal mortality with a relative risk for mortality of 1.35 for moderate anaemia and 3.51 for severe anaemia (Brabin *et al.*, 2001). Prematurity and low birth weight (LBW; <2500 grams) is associated with maternal malaria, including the contribution from both malaria-associated maternal anaemia and placental infection. The contribution of malaria during pregnancy to LBW and subsequent mortality in the first year of life has been estimated to range from 3% to 8% of infant mortality (Greenwood *et al.* 1992; Steketee *et al.* 1996; 2001). Patients seek treatment and treatments often carry their own risks of fatal or morbid outcomes (Snow & Gilles, 2002). Patients who survive severe disease may be left with debilitating sequelae, such as spasticity or epilepsy (Mung'ala *et al.*, 2004). Subtler

consequences, including behavioural disturbances or cognitive impairment, have also been described (Holding & Snow, 2001; Holding & Kitsao-Wekulo, 2004; Mung'ala *et al.*, 2004). These combined effects are summarized in Figure 1.3. In the absence of measures aimed to reduce the risk of infection, the risks shown in Figure 1.3 will largely depend upon extrinsic factors that determine the speed with which a population develops acquired immunity, access to effective case-management and intrinsic factors such as host genetics and nutrition.

Figure 1.3. The direct, indirect and consequential public health effects of *Plasmodium falciparum* malaria in Africa (Snow & Gilles, 2002)



1.5. The malaria burden

The most recent estimations of the burden posed by *P. falciparum* in SSA derive from extrapolations of endemicity-specific epidemiological assessments of morbid and fatal risk to populations located in areas of the continent congruent with defined levels of climate suitability for stable transmission (Snow *et al.*, 2003). Further details

of how these risk populations were structured are provided in Chapter 4. The models took no account of urbanization and the implications of this are also discussed.

During 2000 Snow *et al.* (2003) estimated that approximately 1.14 million people in SSA might have died as a direct result of infection with *P. falciparum*. 88% of these deaths would have occurred in areas of stable, endemic malaria and the majority of these would have been among young children. Snow *et al.* (2003; 2004) further explore the empirical evidence for indirect or consequential mortality component by reviewing the evidence of trials aimed at reducing parasite exposure on all-cause mortality and the relationship between parasite prevalence and all-cause mortality. They conclude that at least an additional 10% of mortality not directly attributed to malaria could be a result of infection acting as a risk factor.

One million deaths due to malaria each year in Africa resonates with earlier claims of a similar figure proposed as far back as the 1950's (Bruce-Chwatt, 1952; Greenwood, 1990; Schwartzlander, 1997; Struchler, 1989). These estimations are hard to comprehend without a methodological framework or empirical evidence to support them. The estimations provided by Snow *et al.* (2003) continue to be driven by informed approximations, in part because of the paucity of reliable and accurate data, but also due to the inherent difficulties of uniquely diagnosing malaria. The approach employs empirical epidemiological measures of mortality risks, structured according to age and malaria transmission. It is similar to the approach used previously by Snow *et al.*, (1999a) to define the malaria burden of SSA but is cognisant of new data, refined endemicity classifications, broader health consequences and temporal effects associated with changing anti-malarial drug sensitivity. This data-driven approach

allows for a more transparent review of the evidence. However, it is not without scope for further improvement and one major area is in the definition and modelling of populations exposed to the varied risks of parasite challenge in SSA as is the subject of the present thesis.

1.6. Challenges for malaria control in SSA countries

An estimated US \$12 billion is spent on the direct costs of treatments and prevention for malaria in SSA (Gallup & Sachs, 2001). Despite this investment, malaria causes an annual decline in economic growth of at least 1.3% in these countries (Gallup & Sachs, 2001; World Health Organization, 2002). The Roll Back Malaria (RBM) movement proposes to halve malaria mortality by the year 2010. This goal is operational at a time when existing, affordable therapeutics are rapidly failing, health service provision is breaking down, there are no immediate prospects of widespread vaccination and poverty continues to afflict most endemic countries. There are strong reasons to believe that over the past 15 years malaria-specific mortality has risen and now accounts for an increasing proportion of overall childhood mortality (Snow *et al.*, 2003; Korenromp *et al.*, 2003). It is important to realise that the starting point for new efforts to Roll Back Malaria is not a level playing field, but a mortality burden returning to levels described before Africa gained independence.

Much is understood of the transmission dynamics of the malaria parasite and this has provided a solid epidemiological basis for the development of the appropriate vaccine in the future. Morbidity, mortality, the efficacy of interventions such as ITNs, the spread of drug resistance and the development of anti-malarial immunity are all

driven by the intensity of malaria transmission. The most significant challenge facing the malaria epidemiologist today is to describe malaria health outcomes in Africa and how these risks change from birth through adulthood according to the dependent factors of infection, immunity and control. As part of this, a better cartography of malaria risk is required.

1.7. Measuring infectious disease transmission

The underlying principles of transmission dynamics are common to many infectious diseases and are captured in a variety of mathematical models of the probability of transmission. The infection probability p is the probability that a susceptible host becomes infected on contact with a pathogen. It is basically the product of the contact rate with infective persons \times the probability of transmission per contact \times the duration of infectiousness (Anderson & May, 1992; Halloran, 2001). p can be modelled in many different ways for different diseases but the basic reproduction number R_0 is commonly used in infectious disease epidemiology (Ross, 1911; MacDonald, 1957).

When the transmission cycle of a pathogen involves an insect or other vector, measurements of p consider the vector's ecology. Models of malaria transmission consider the contact rates from mosquito to man and from man to mosquito, both of these transmission probabilities and the duration of infectiousness in both the mosquito and in humans. For malaria transmission, R_0 is also refined to incorporate ecological covariates of vector and parasite dynamics. Many of these variables; including mosquito abundance, feeding frequency, survival and the parasite's extrinsic incubation period, are sensitive to climate. This knowledge of climate

sensitivity is applied to the biological modelling of malaria transmission as defined through R_0 , also commonly known as the Ross-MacDonald Model (Ross, 1911; MacDonald, 1957).

1.7.1. *The basic reproduction number of a disease R_0*

For a pathogen to persist in a population, the pool of susceptible individuals needs to exceed a threshold number such that, on average, each primary case of infection generates at least one secondary case. R_0 is defined as the average number of successful offspring that a parasite can potentially produce (MacDonald, 1957) or the average number of secondary infections produced from one infected individual introduced into a susceptible host population (Anderson & May, 1992). The longevity and abundance of vectors are key factors in the sustenance of malaria transmission. The sporogonic cycle of *Plasmodium* is long relative to the life span of the mosquito and usually only a small proportion of mosquitoes live for long enough to transmit disease. A sizeable population of vectors is therefore required to maintain transmission. R_0 combines measures of mosquito infectivity and survival and, for the parasite to persist, R_0 must be greater than 1.0. A reproduction number > 1.0 indicates an expansion of infections in a population while < 1.0 indicates a decline in infections in the population. R_0 is calculated using the following equation, the elements of which are summarized in Table 1.2.

$$R_0 = \frac{ma^2bce^{-\mu\tau}}{\mu r} \quad \text{(Equation 1.1.)}$$

Once infected, the mosquito remains infective throughout its lifetime. Its duration of infectivity is related to the reciprocal of its mortality rate, $1/\mu$. The human biting rate

is derived from the frequency of bites and the proportion of those blood meals that are taken from humans. Interventions that modify the exponents ‘ μ ’ and ‘ τ ,’ in Equation 1.1. have the most significant impact on R_0 . This model was the basis for the WHO’s decision to initiate widespread anti-vectorial measures during the 1950’s malaria eradication campaign (Garrett-Jones, 1964). Both elements, μ and τ , refer to the malaria vector and are therefore sensitive to climatic influences (Rogers, 1988).

Table 1.2: Elements of R_0 for malaria

Term	Description
N	the size of the human population
M	the size of the female mosquito population
m	= M/N , the number of female mosquitoes per human host or the vector density
a	the rate of biting on humans by a single mosquito (number of bites per unit time)
b	the proportion of bites that result in vector to human infection
c	the proportion of bites that result in human to vector infection
r	the rate of recovery from infection
μ	the vector mortality rate
τ	the extrinsic incubation period (incubation period to infectivity of the parasite within the vector)

An, as yet, un-quantified relationship exists between biotic factors (relating to the vector or the parasite) and abiotic factors (such as ambient temperature and relative humidity). Biotic entomological factors include vector density, survival, mosquito biting habit, duration of the asexual development phase and the proportion of infective mosquitoes. Biotic parasitological factors include the population-based prevalence of the parasite *i.e.* the parasite ratio (PR) and the immunological status of the population. A threshold reservoir of infective carriers in the population is necessary to maintain transmission. The PR is the proportion of population carrying

the infective stage of the parasite and is thus a measure of the infective reservoir in the population. Abiotic factors are mainly meteorological and environmental.

1.8. Describing the endemicity of malaria

The endemicity of malaria has been defined in a variety of ways by various authors using varied measures and categories of these measures. The descriptions are often attempts to describe a continuous variable according to discrete categories. While conditions at the upper and lower extremes of endemicity may be easy to define, conditions in between are not. The result is some degree of arbitrariness and overlap. In some cases, operational definitions are unclear and their interpretation is therefore difficult. Some of the historical definitions have incorporated new terms as additional knowledge has permitted more detailed studies of endemicity. In general, definitions selected for individual studies are chosen to suit the questions being asked by the specific study. The main definitions available from the literature are presented in this section.

1.8.1. Definition of endemicity

A disease is described as endemic when it is normally found in a population. Malaria, for example, is endemic in most of SSA. Areas of non-endemicity are also found in SSA including parts of southern Africa, North Africa and highland areas > 2500 metres above sea level (Service & Townson, 2002). Endemicity can be modified and many non-endemic areas are former areas of endemic malaria where the disease has been eradicated by successful anti-malarial interventions. Following 50 years of

control activities in Egypt and Morocco, for example, malaria was close to eradication by the late 1970's. Control activities were briefly interrupted in the 1980s and the disease remained at low endemicity. Control was reinstated and eradication achieved by 1999. The disease is no longer endemic in the two countries (Beljaev, 2002).

Within endemic areas the intensity of transmission, or endemicity that populations are exposed to, is spatially heterogeneous (Snow & Marsh, 1998; Snow & Gilles, 2002). Transmission can be described as a continuum where at the lowest extreme the risk of infection is minimal, herd immunity is low and all age groups of the population are exposed to an equal risk of disease. At the highest extreme of the continuum, intense year-round infection results in the development of a high degree of immunity achieved by early childhood.

1.8.2. Stable and unstable malaria

The endemicity of malaria has also been defined in terms of its level of stability; *i.e.* stable or unstable, again two extremes on a continuum of epidemiological scenarios (MacDonald, 1957). Where malaria is stable, transmission is often continuous or perennial; the frequency of infection persistently high and immunity, due to frequent infection challenge, develops at a relatively early age among the resident population. Stable malaria is relatively insensitive to environmental changes and control interventions. The mechanisms of this refractoriness are unclear but the high level of community immunity (herd immunity) may play a major rôle. Estimations of herd

immunity are rarely available and this is one of the many terms defining endemicity that have not been operationally defined.

Unstable malaria, on the other hand, is characterised by great spatial and temporal variability. Recessions and recurrences occur *i.e.* periods when disease incidence is low alternate irregularly with high incidence periods with several months separating the periods of recurrence. R_0 also fluctuates either side of 1.0. Because the periods of disease outbreak are brief and infrequent, herd immunity is undeveloped among all age groups. Seasonal fluctuations are a feature of malaria across a wide range of endemicity settings and again, definitions of 'temporal variability' are unclear. In many areas, low or moderate levels of transmission occur year-round with periodic intra-annual upsurges occurring. These 'seasonal' upsurges are not associated with levels of morbidity and mortality as high as those experienced during epidemic periods.

1.8.3. Seasonality of malaria transmission

The inter-annual periodicity of malaria's endemicity occurs in response to climatic seasons. Vector densities and relative abundance, vector biting rates, and the risk of infection respond to these climatic variations. Studies have also suggested that the amplitudes and frequency of these seasonal cycles of infection risk modify the incidence of severe *P. falciparum* malaria in children (Brewster & Greenwood, 1993; Snow *et al.*, 1993). Evidence of relevant seasonal aspects of malaria is found in a range of longitudinal field studies conducted in malaria endemic sites across the world (McKenzie *et al.*, 2001). Pioneering investigations of the seasonality of malaria

transmission were carried out in the 1930's by Gill (1938). He stated that '*the cause of the seasonal periodicity of malaria is still imperfectly understood and the whole question of the influence of climate and season upon the natural history of the disease requires re-examination.*' Although seasonality has been recognised as a key influence of infection incidence since Gill's work, few investigations of seasonality were carried out prior to the 1990's. Today the understanding of seasonal influences on malaria transmission is still limited. Recent studies of malaria seasonality are discussed further in Chapter 2.

1.8.4. Epidemic malaria

Epidemics may be considered a form of highly unstable malaria. By definition, they occur when unusual environmental conditions cause sudden upsurges of transmission among non- or semi-immune populations (Molineaux, 1988; Nájera *et al.*, 1998; Abeku *et al.*, 2003). They are characterized by extremely high mortality among all age groups of a population and are worsened in situations such as in the highlands of East Africa where medical systems are ill equipped to cope with the numbers of the ill and dying (Cox *et al.*, 1999; Hay *et al.*, 2003). Macdonald (1957) defined an epidemic as '*...an acute exacerbation of disease out of proportion to the normal to which a community is subject.*' What constitutes a 'normal' level of disease is often debatable. This, however, is the definition still in use by the WHO with the following additional caveats:

"An epidemic is the result of the disturbance of a previously existing equilibrium of the ecological system comprising human, parasite and

vector populations in a particular environmental niche. Depending on the resilience of the system, and whether or not the disturbance has changed some of its essential components, it will either return to its previous state of equilibrium after the end of the disturbance, or tend to find a new equilibrium, with or without going through a period of oscillation” (Nájera et al., 1998).

In SSA, epidemiological characteristics that predispose to epidemics are found in semi-arid/arid and highland areas where low rainfall or low temperatures respectively tend to limit malaria transmission. Such epidemics have been documented in the areas of the East and Horn of Africa (Heisch & Harper, 1949; Fontaine *et al.*, 1961; Some, 1994; Malakooti *et al.*, 1998; Lindblade *et al.*, 1999; Snow *et al.*, 1999b; Bødker *et al.*, 2000; Abeku *et al.*, 2003) and in 1997 in the arid regions of North Eastern Kenya (Brown *et al.*, 1998; Snow *et al.*, 1999b). The emergence of malaria in previously ‘malaria-free’ zones and highland areas, from which it has been previously eradicated, has sparked a renewed interest in the phenomenon of ‘highland malaria’ (Malakooti *et al.*, 1998; Lindblade *et al.*, 1999; Bødker *et al.*, 2000; Lindblade *et al.*, 2000; Cox *et al.*, 2002). Examples from East Africa reveal how variable definitions of “highland malaria” can be.

Highland areas of northern Tanzania include the Pare and Usambara Mountain ranges (600-1200 metres ASL). Most of Tanzania’s highland malaria data are from the Eastern Usambara area, which historically was considered as having a very low risk of malaria infection (Wilson D.B. & Wilson M.E., 1937; Clyde, 1967; Matola *et al.*, 1987; Cox *et al.*, 1999; Bødker *et al.*, 2000; Lindsay *et al.*, 2000). Estimates of

population based infection prevalence rates have varied markedly between studies from as low as 6% to above 75% (Lelijveld, 1971; Voller *et al.*, 1972; Matola *et al.*, 1987; Ellman *et al.*, 1998; Drakeley & Reyburn, personal communication). Several of these studies have suggested that these marked fluctuations are due to changes in human activities (clearing of forests for tea plantations and movement of migrant workers from surrounding lowland malarious areas) that introduce parasite carriers to non-immune populations. Studies by Ellman *et al.*, (1998) found that, despite low infection challenge, parasite prevalence in this area was consistently high (33-76%) and suggested that malaria was probably not as rare in the Usambara's as previously suggested.

Malaria incidence rates in highland areas of Ethiopia have also increased markedly in the 1980's (Tulu, 1993) but much of this rise may have been due to emerging drug resistance and population resettlement from lowland areas. In Kenya there are very limited malariometric data from highland or epidemic defined areas. Kericho in Western Kenya (1700-2230m ASL) is considered an epidemic-prone area by the Kenya Ministry of Health. Hospital-based studies of malaria by Malakooti *et al.*, (1998) using data from 1991-1997, suggest a highly seasonal malaria profile with epidemics occurring when rainfall exceeds a threshold of 150 mm a month.

A clear definition of epidemic malaria is not available at present and current ones are specific to local regions. No single factor can be considered a precipitator and most variables are interdependent. Modelling the variables related to precipitation of epidemics has an obvious rôle in epidemic detection, early warning and future forecasting (Hay *et al.*, 1998a; Cox *et al.*, 1999; Thomson & Connor, 2001; Hay *et*

al., 2002c; Hay *et al.*, 2003). In recent years, models of highland malaria have been developed for East Africa aimed at addressing information needs for epidemic early warning in the region (Cox *et al.*, 1999).

1.9. Indicators of malaria endemicity: incidence, prevalence, the force of infection and the parasite ratio

This section considers methods used to measure endemicity. Incidence refers to the number of *new* cases of infection occurring in a population within a unit of time and is sometimes referred to as an ‘attack rate.’ Numerical measures of incidence are normally given as the number of cases per population per year. For malaria, incidence of infection is impractical to measure as it requires frequent population-based blood sampling. The true incidence rate of a transmissible disease is a measure of the force of infection, λ , defined as the per capita probability that a susceptible individual acquires infection (Anderson & May, 1992) or the number of new cases of infection/person/year. For a specific community, the value of λ is age-dependent and can be measured epidemiologically by serological conversion rates *i.e.* the rate at which individuals convert from seronegative to seropositive. λ can also be estimated from the number of infected individuals (prevalence (Section 1.9.1.)) if a linear relationship between λ and prevalence can be assumed (MacDonald, 1950a; Molineaux *et al.*, 1978). The value of λ is modified by the balance of the ratios between susceptible persons, those who become infected, recovery rates and those who develop immunity to infection. This relationship can be represented as follows (Anderson & May, 1992):

$$\lambda = \beta \int_0^{\infty} Y(a, t) da \quad \text{(Equation 1.2)}$$

$Y(a,t)$ represents the number of infected individuals 'a' at time 't' ' γ ' represents changes in the number of susceptible hosts due to the development of immunity. When $\gamma=0$ there is life long immunity and if $\gamma=\infty$ there is no immunity. β is a measurement of transmission that incorporates biotic and abiotic factors that drive transmission while ' d_a ' is the mortality rate with respect to the infected individuals. In malaria endemic areas, λ can only truly be measured in infants (0-11 months of age) as they are the only truly susceptibles having not had prior exposure to infection (MacDonald, 1950a).

Estimates of λ have been derived from repeat cross-sectional surveys of infection in infant populations in SSA where values ranged from 0.41 infections/infant/year under the low-moderate seasonal malaria conditions found in Sukuta, The Gambia to >2 infections/infant /year under intense perennial conditions in Kilifi, Kenya (Snow *et al.*, 1994). Such studies are however extremely labour intensive and as such are rare in the literature on malaria endemicity in SSA.

1.9.1 Measuring the parasite ratio

Where infections persist for long periods and super-infection occurs, it is not possible to distinguish between new (incidence) and old (prevalence) infections making measurement of incidence difficult among older children. The term prevalence refers to the proportion of infections in a community at a point in time and is commonly measured according to the PR. The PR is based on random community-based samples and is taken as the proportion of those sampled that have peripheral blood stage parasites. The laboratory confirmation of malaria parasites in blood slides involves

the direct microscopic observation of whole parasite cells, parasite nucleic acids or parasite products. The techniques for this, involving slide preparation, drying, fixing, staining and microscopic examination are outlined in detail in Schute, (1988), World Health Organization (1991) and Gilles (1993b).

Microscopy for malaria parasites demands observer skill and quite often low parasite densities can go undiagnosed. Skill is required to prepare a slide of adequate thickness (thick blood films vs. thin films), to mix the various buffering and staining solutions required and also in distinguishing parasite particles from artefacts. Rules for numbers of high-powered microscopic fields examined are not uniform. The more fields examined, the more likely parasites will be identified. Although 100 to 200 microscopic fields are recommended, the number of fields counted during microscopic examination varies from study to study from as low as 50 fields with some observers only counting until the first parasite is seen. Making comparison between studies can therefore be problematic. Quality control can be applied by re-examining a proportion of slides by a second observer and estimating observer agreement. This, however, is rarely reported in published parasitological surveys.

Although often described as the parasite rate or absolute parasite rate, strictly speaking, it is not a rate, as it is not estimated over a period of time. The term Parasite Ratio (PR) is preferred and is used here. In endemic areas, an endemicity-dependent saturation of the PR occurs in older children due to the development of immunity and the longevity of infection such that only children under the age of 10 years are included in its estimation (Metselaar & van Thiel, 1959). Ideally the PR should be recorded for each *Plasmodium* species and sexual and asexual stages separately but

this is often neglected. The accuracy of the PR as an endemicity measure depends on how well it can be related to β (Equation 1.2) and thus how well it represents the multiple factors that define malaria transmission. Compared with other measures, the PR is limited in many ways but as it has high specificity and is relatively easy to estimate under field conditions, it is widely available. For this reason, for many years, malaria control programmes across SSA have based their epidemiological assessments of malaria on the PR. A more comprehensive description of the PR and the rationale for its use in defining endemicity is given in Chapter 2.

1.9.2. The classification of endemicity based on the parasite ratio

The categorical classification of endemicity has historically been based on cut-off values applied to prevalence indicators described in the previous section. The 1950 WHO malaria conference in Kampala endorsed a classification based on the prevalence of enlarged spleens (splenomegaly), the spleen rate (World Health Organization, 1951)). Splenomegaly occurs commonly with chronic parasitic infections. Such infections are highly prevalent in malaria endemic areas. As the spleen rate has a low specificity, because of competing causes of splenomegaly in Africa, Metselaar & van Thiel (1959) revised the WHO classification and defined four classes of endemicity using the PR (Table 1.3).

Prevalence of infection crudely corresponds to the frequency and duration of parasite exposure but does not provide a precise quantification of the number of new infections received by a child each year. The non-linear relationship between challenge and infection prevalence has long been recognised (MacDondald, 1950b).

Given the long persistence of parasitaemia after a single infection (Eyles & Young, 1951), it might be expected that PR would become an increasingly blunt measurement instrument at higher levels of transmission.

Table 1.3: Classification of malaria (Metselaar & van Thiel, 1959)

Type of endemicity	Parasite Rate	Epidemiological description
Hypoendemic	If the parasite rate in the age group 2-10 years is as a rule under 10%. It may be higher for part of the year	Areas where there is little transmission and the effects, during the average year, upon the general population are unimportant. Can be regarded as unstable malaria
Mesoendemic	If the PR in the age group 2-10 years is as rule between 11-50%. It may be higher for part of the year	Typically found among rural communities in sub-tropical zones when wide geographical variations in transmission risk exist. Can be regarded as unstable malaria in some cases although epidemics are less severe than in hypoendemic areas
Hyperendemic	If the PR in the age group 2-10 years is constantly over 50%	Areas where transmission is intense but seasonal and immunity is insufficient in all age groups
Holoendemic	If the PR is constantly over 75% in the 1-year age group	Perennial, intense transmission resulting in a considerable degree of immunity outside early childhood. Stable malaria

1.9.3. The relationship between the Entomological Inoculation Rate (EIR) and the parasite ratio

The entomological inoculation rate (EIR) is the product of the human biting rate (the number of vectors biting an individual over a fixed period of time, usually expressed per year) and the sporozoite rate (the proportion of vectors with sporozoites in the salivary glands). Estimates of sporozoite challenge across broad geographical areas are fraught with methodological problems. First, the frequency of parasite exposure varies considerably within small geographical areas and even between households (Cattanni *et al.*, 1986; Mbogo *et al.*, 2003). All previous comparisons of the risks of parasite challenge from birth and disease outcomes (Snow *et al.*, 1994; Smith *et al.*, 2001a) and spatial models of EIR (Rogers *et al.*, 2002) have selected one point

estimate of EIR from a single or collection of households to represent the wide areas. Second, whilst the EIR defines the frequency of contact between hosts and infected vectors, methods used to measure it in the field vary (Githeko *et al.*, 1996). Whether sampling of vectors is undertaken indoors or outdoors with light-traps placed close to the bed, with pyrethrum spray catches or through human bait catches affects the quantification of the EIR. The overall consequence is that EIR data are hard to reconcile as standard measures between sites and may convey a degree of precision in the estimation of parasite challenge that simply does not exist. More importantly the overall paucity of reliable data across SSA limits the ability to develop robust models (Hay *et al.*, 2000b).

The PR is a more readily available and widely used index of transmission intensity and it has the advantage over entomological measures of endemicity in that parasitological surveys often cover more comprehensively defined spatial units. Beier and colleagues (1999) have recently reviewed coincidental measures of the PR in childhood and standardised estimates of EIR. Their analysis showed that over 70% of the variance in PR could be explained by the logarithmic classification of the EIR, overall and when analysed by East and West Africa or by “ecological zone”. Broadly there were no studies where a prevalence of infection in childhood was less than 50% that had an EIR greater than 15. Conversely all areas with a prevalence of infection greater than 80% had an EIR in excess of 200. Analysis of the logarithmic classifications of the EIR showed that an EIR of less than 1 infective bite per adult per annum encompassed an interquartile range of infection prevalence between 1 and 25%. EIRs between 1 and 10 had an infection prevalence range of 25-50%, and EIRs

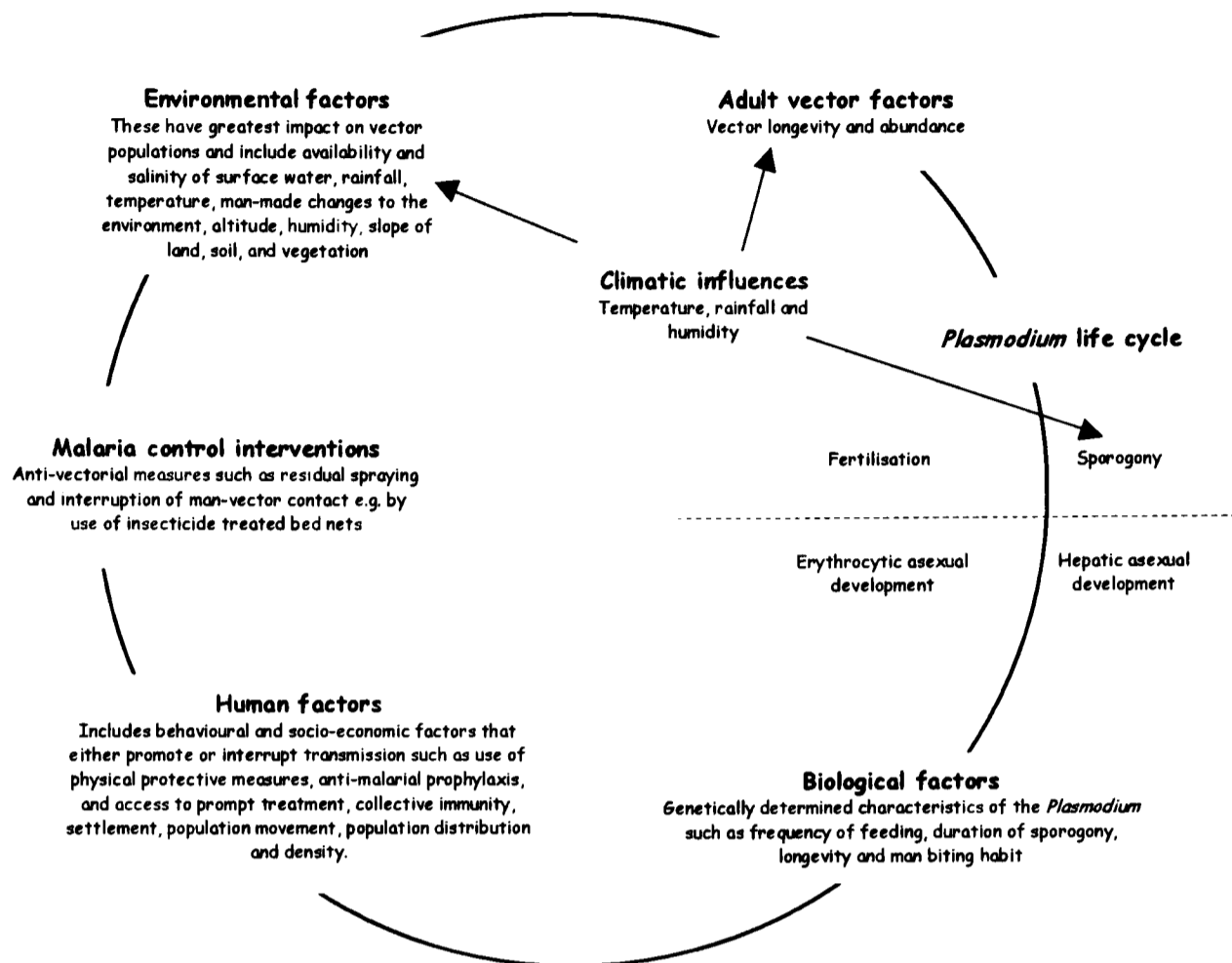
between 11 and 100 covered the prevalence range between 51 and 74%. The highest EIRs (in excess of 100) had infection prevalence in excess of 74%.

The analysis by Beier and colleagues strongly suggests that approaching measures of exposure as categorical rather than continuous variables may be important to overcome the inadequacies in the measures available to estimate parasite challenge. This is important for reasons not only related to the methodological difficulties. The future mapping of the spatial distribution of transmission intensity, populations exposed to risk and the likely impact of intervention strategies will most probably be based upon categorical classes of risk and benefit.

1.10. Determinants of *Plasmodium falciparum* endemicity

A complex interaction exists between man, the parasite, the vector and the environment and this interaction determines malaria's endemicity (Figure 1.4). As shown by R_0 , malaria transmission is highly sensitive to the survival rate of *Anopheles* (Jepson *et al.*, 1947; Macdonald, 1957; Detinova, 1962; Charlwood *et al.*, 1995; Martens, 1997) (Section 1.7.1.)). The influences on transmission dynamics are discussed in this Section.

Figure 1.4. A conceptual framework of the dynamics of malaria transmission (adapted from Wernsdorfer & McGregor (1988))

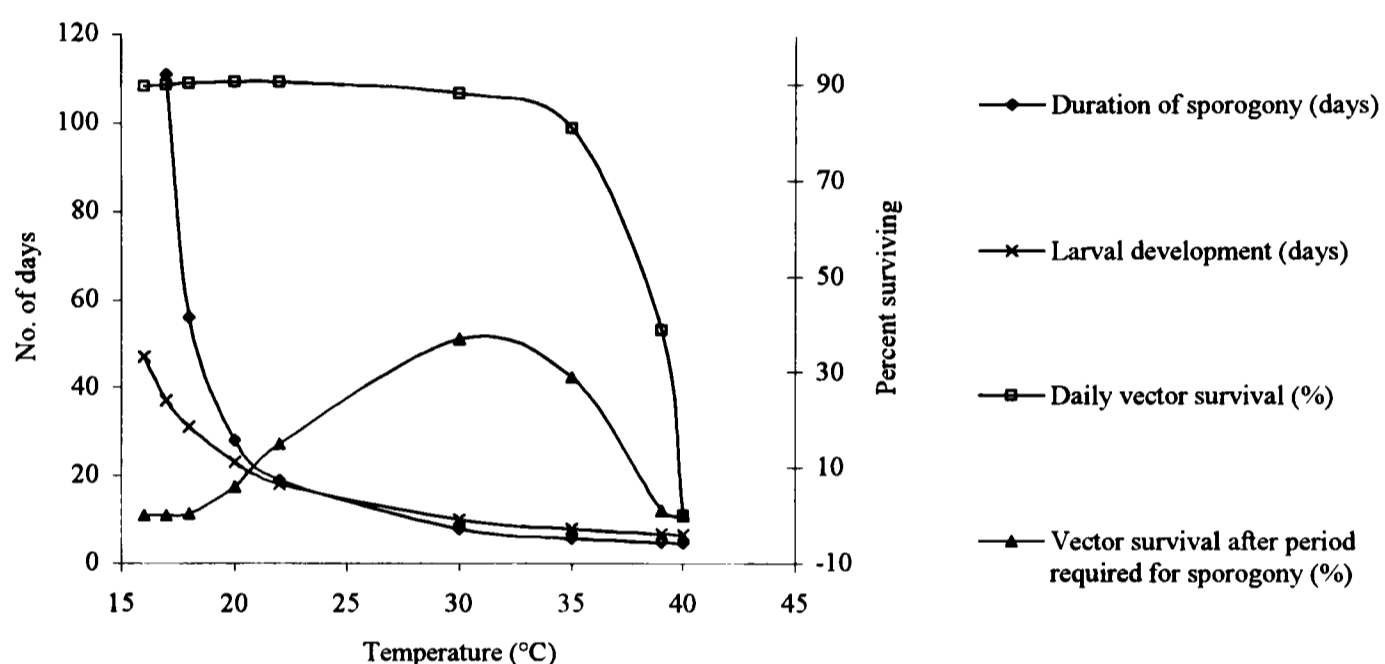


1.10.1. Temperature

The optimal temperature for parasite development (sporogony) is species dependent. Sporogonic development for *P. falciparum*, for example, takes 9 days at 30°C, 10 days at 25°C, 11 days at 24°C and 23 days at 20°C (Boyd, 1949). Below 16°C and above 35°C sporogony ceases for *P. falciparum* (Detinova, 1962). The optimum temperature for sporogony in *P. vivax* is around 25°C. Temperature also has effects on mosquito development (Figure 1.5). Mortality at optimum conditions of temperature (27°C) is 15% compared with 16.1%, 1.9% and 4.9% at 21.2°C, 26.5°C and 29.5°C respectively (Armstrong *et al.*, 1961). Thermal death of mosquitoes occurs at 40-42°C (Dutta & Dutt, 1978). Mosquitoes are less sensitive to low temperatures than malaria parasites and temperature has a greater limiting effect on

the distribution of malaria parasites than on the distribution of mosquitoes. Uninfected mosquitoes can therefore be found in high altitude areas of endemic countries, a phenomenon termed 'anophelism without malaria.' In such areas sudden rises in temperature, permitting the proliferation of malaria parasites may be associated with epidemics of malaria (Loevinsohn, 1994; Matola *et al.*, 1987)

Figure 1.5. The effect of temperature on parasite (Detinova, 1962) and vector (Martens, 1997) development and survival.



1.10.2. Altitude

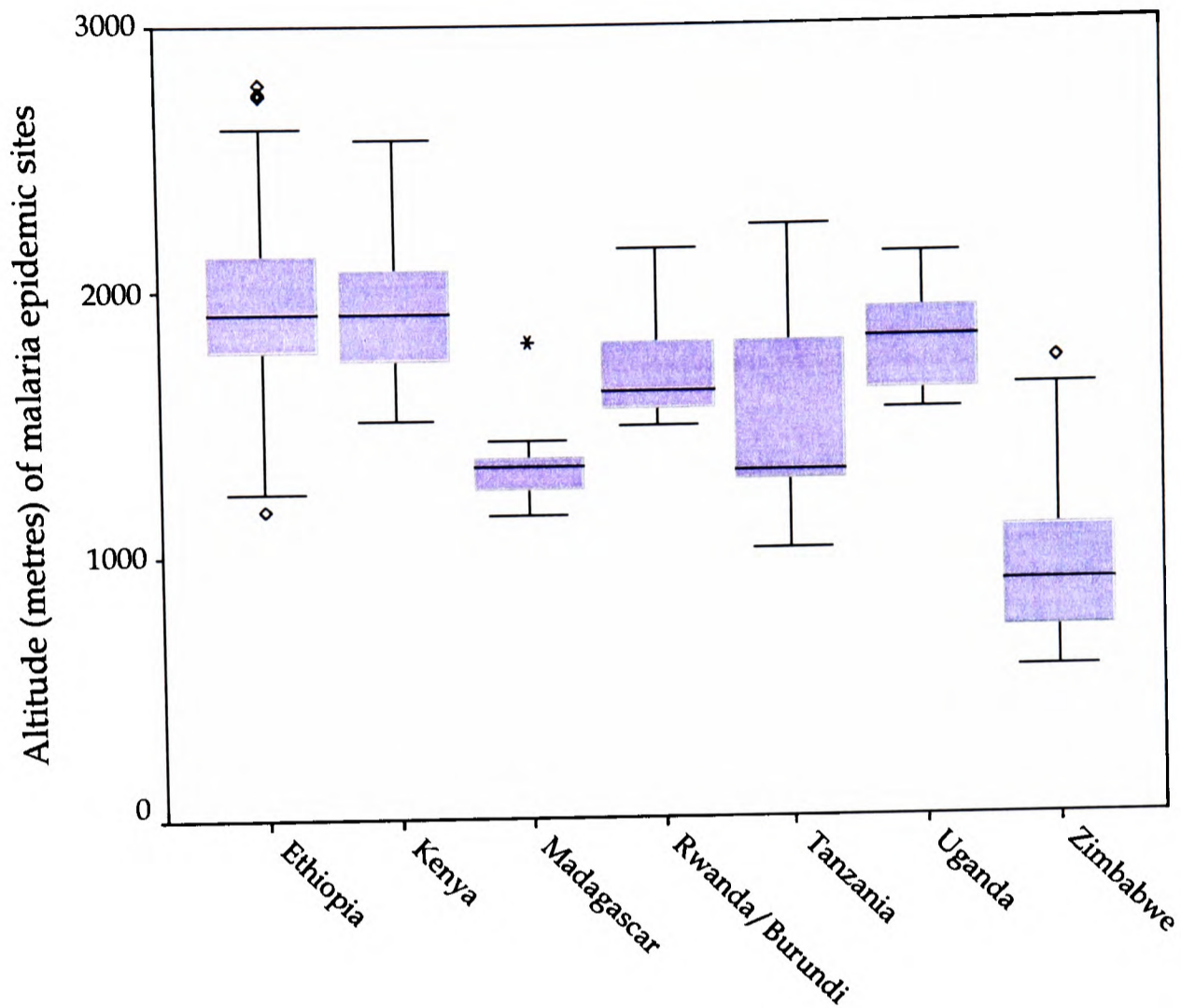
Altitude is inversely related to temperature and in general a 0.5-0.6 °C drop in temperature occurs for every 100-metre increase in altitude (Schwetz, 1942; Cox *et al.*, 1999). This altitude/temperature criterion, however, is very variable and is influenced by latitude, air saturation and distance from large water bodies (Cox *et al.*, 1999). Malaria transmission is dependent on the presence of suitable numbers of parasitised vectors and susceptible hosts. *Anopheles* are generally not found at

altitudes over 2500 metres above sea level (Service & Townson, 2002). As altitude increases, transmission periods become shorter and a more seasonal transmission profile tending towards 'epidemics' may be experienced (Bruce-Chwatt, 1984).

Historical medical reports record an absence of malaria in highland areas of East Africa above 1500m at the turn of the 19th Century (Lindsay & Martens, 1998). Other authors have reported epidemic malaria at 2000m in Kenya (Bruce-Chwatt, 1984). Studies in the Muheza District of Tanzania also showed decreasing malaria transmission risk, as determined by the EIR and PR, with increasing altitude (Ellman *et al.*, 1998; Cox *et al.*, 1999; Lindsay *et al.*, 2000; Drakeley & Reyburn, personal communication). The exact altitude at which transmission ceases is difficult to define. It varies geographically and according to latitude and is modified by a host of local factors such that altitude alone is a poor predictor of epidemic-proneness (Cox *et al.*, 1999; Figure 1.6).

The recent emergence of malaria at higher altitudes causing epidemics of 'highland malaria' has sparked much debate on the possible rôle of climate change and global warming in the aetiology of re-emergent malaria (Reiter, 1998; Dye & Reiter, 2000; Shanks *et al.*, 2000; Hay *et al.*, 2002a; 2002b). As pointed out by Mouchet *et al.*, (1998), Reiter (1998), Lindblade *et al.* (2000), Shanks *et al.* (2000) and Hay *et al.* (2002b), the quality of public health services, irrigation and agricultural activities, land use practices, civil strife, natural disasters, ecological change, population change, use of insecticides, and the movement of people are more likely to be the cause of the increase in malaria at high altitudes than temperature.

Figure 1.6: Altitude ranges of highland malaria epidemics in Eastern African countries (Cox *et al.*, 1999). The box length is the interquartile range; the line within the box is the median. Outliers (open circles) are values between 1.5-3 box lengths from the upper quartile.



1.10.3. Rainfall and humidity

It is difficult to quantify the direct relationship between rainfall and malaria transmission. Several investigations, particularly among non-immune populations, have noted marked increases in infection incidence during periods of unusually heavy rainfall (Bouma & Vanderkaay, 1996; Bouma & Dye, 1997; Brown *et al.*, 1998; Kilian *et al.*, 1999). Breeding site availability is associated with rainfall and increased mosquito survivorship with increasing humidity (Gill, 1920; Dutta & Dutt, 1978). Mosquito abundance is directly related to the prevalence of water bodies available for vector breeding. Dramatic increases in populations of *An. arabiensis* are associated

with short-term increases in rainfall that create temporary breeding pools. Studies in the Kilombero River Valley in Tanzania show that *An. arabiensis* numbers can increase a hundred-fold during the short rainy season compared with dry season numbers (Charlwood *et al.*, 1995). Unusually heavy rainfall, however, can have the effect of flooding out mosquito breeding sites as was seen during the severe flooding in Mozambique in March 2000 (Cox *et al.*, 2002).

Breeding site availability is of particular relevance to malaria vector survival during extended dry periods in arid areas. Vector populations survive the dry seasons in small pools of water, marshes and dry riverbeds (Bruce-Chwatt, 1984; Hamad *et al.*, 2002). The ability of soil to retain moisture may also contribute to vector survival. The prevalence of malaria parasitaemia has been shown to vary locally in relation to soil type with higher parasitaemias associated with more moist soils and it has been shown that vectors may undergo periods of aestivation in moist soils (Chinery, 1997; Patz *et al.*, 1998). The effect of humidity in field conditions is not well understood. Laboratory studies, however, have shown that a humidity of 50-60% is optimum for vector survival (Molineaux, 1988; Beljaev, 2002). The climatic factors described in this section are not independent, their effects being modified by human intervention. Some of these human influences on transmission, notably population settlement, movement and urbanisation, are described in the next section.

1.10.4. Human factors influencing malaria transmission

Although the important rôle of human factors for malaria control has been recognised for many years (Prothero, 1977; MacCormack, 1984), very few studies or

interventions have attempted to address them. This is perhaps because human behaviour is often difficult to predict or measure. Urbanization, population distribution and movement, agricultural practices and land-use, nonetheless are key influences on malaria transmission (Gilles, 1993a). Construction of dams, formation of reservoirs and irrigation systems and agricultural practices provide important influences on human settlement, land-use and disease risk. The relationship between the amount of rainfall and the development of breeding sites is dependent on several factors such as the slope of the land; run-off and soil type and the suitability of these breeding sites will further be affected by the availability of shade, vegetation, predators and level of salinity.

1.10.4.1. Population movement, migration and displacement

Recent interest in human factors has been sparked by the re-emergence of malaria in areas where it either had previously been eradicated or had not been recorded. This phenomenon has been attributed to population mobility (Martens & Hall, 2000). Population movement may be temporary (circulation) or involve a permanent change of residence (migration). In general, migration, as a consequence of war and famine or urbanization, poses the greatest risk for malaria transmission.

The movement of non-immune individuals *en masse* into holoendemic areas could result in malaria epidemics or conversely immune parasite carriers have been seen to introduce the disease to previously malaria-free areas (Prothero, 1961; Mouchet *et al.*, 1998). The greatest risk of malaria increase occurs when mass population re-settlement occurs rapidly and in an unplanned fashion. Such situations are becoming

increasingly prevalent in war and famine stricken countries. According to the United Nations High Commission for Refugees (UNHCR) in 2000, at least 12.1 million people were displaced due to war and famine (United Nations High Commissioner for Refugees, 2001). Of these, 3.6 million refugees were in Africa where the 2nd highest proportion of refugees (30%) live (United Nations High Commissioner for Refugees, 2001). Displaced populations are at increased risk of malaria as they often live in temporary shelters with inadequate sanitation facilities. Epidemics of malaria occur when non-immune populations are moved into highly endemic areas (Crowe, 1997). Although the exact extent of the risk of malaria in these situations is not known (Allan, 2003).

1.10.4.2. Urbanization

According to the United Nations Development Programme's (UNDP) world urbanization prospects report, 37.2% of Africa's 794 million inhabitants lived in urban areas in 2000 (Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat, 2002). This figure is expected to rise by 52.9% to 1.5 billion people by 2030. The urban growth rate during this period is estimated to be 2.1% per annum compared with a 1.14% *p.a.* growth rate in rural areas. SSA, thus, has the highest rates of urbanization in the world (World Bank, 1996). A large part of the urban growth will occur as a consequence of transformation of rural trading centres into towns and cities. Such rapid urban expansion often occurs ahead of infra-structural developments with the result that low standards of housing, health care, water supply and sanitation prevail. In addition, people moving from rural areas to settle around urban centres often retain a rural life style of small-scale

subsistence farming using irrigation and the use of informal building materials. The resulting pattern of malaria transmission may mirror that of rural areas (Warren *et al.*, 1999).

Although it is widely held that urban malaria is imported as a result of such population migrations, there is increasing evidence of local transmission in urban centres (Rapuoda & Achola, 1984; Warren *et al.*, 1999). A study in Accra, Ghana attributes increasing vector populations to pools of surface water created by rapid, inadequately planned urbanization (Chinery, 1984). The effects of habitat changes on vectors show some worrying trends. In the study in Accra, *An. gambiae*, normally a fresh water breeder, was noted to adapt to the polluted water habitats found in urban settings (Chinery, 1984). Similarly, increases in vectorial capacity were seen in Cotonou, Benin where urbanization and accompanying changes in housing favoured the proliferation of *An. gambiae* against indigenous *An. melas* (Akogbeto *et al.*, 1992).

Increasing data from hospital-based studies suggest the presence of local urban malaria transmission in major cities; Freetown (Morgan, 1994), urban Mozambique (Thompson *et al.*, 1997; Granja *et al.*, 1998), Brazzaville (Carme *et al.*, 1992; Carme, 1996) Zanzibar (Matteelli *et al.*, 1996) and Kinshasa (Hedberg *et al.*, 1993). In contrast, however, studies of non-mobile or more settled urban populations suggest that urbanization may be associated with a reduced risk of malaria transmission. Decreasing risks of infection associated with urbanization, either entomological or parasitological, have been shown in urban areas in West Africa: Liberia (Björkman *et al.*, 1985), The Gambia (Lindsay, 1990), Senegal (Trape *et al.*, 1992), Ghana

(Gardiner *et al.*, 1984), Burkina Faso (Robert *et al.*, 1985; Rossi *et al.*, 1986; Benasseni *et al.*, 1987; Modiano *et al.* 1999); Central Africa: Cameroon (Fondjo *et al.*, 1992), Republic of Congo (Trape and Zoulani, 1987), Democratic Republic of Congo (Ngimbi *et al.*, 1982; Coene, 1993); the Horn of Africa: Ethiopia (Yohannes & Petros, 1996), Sudan (El Sayed *et al.*, 2000); and Southern Africa: Zambia (Watts *et al.*, 1990). Moreover, in a recent meta-analysis of 159 sites across Africa where annual entomological inoculations rates (EIR) had been recorded, it was shown that people in rural areas received on average 146 *P. falciparum* infected bites per annum compared to only 14 for people resident in urban areas (Hay *et al.*, 2000b; 2005a).

1.10.4.3. Irrigation, water resources management and rice cultivation, land use and agriculture

Labour related to agricultural activities and mining may also pose an increased risk of malaria infection in SSA (Martens & Hall, 2000). Water resource development projects such as irrigation and cultivation of rice create suitable habitats for mosquito breeding (Ijumba *et al.*, 2002a; Henry *et al.*, 2003) although the exact impact of these activities on infection risk is difficult to define. Some studies associate rice cultivation with a decreased risk of malaria infection (Ijumba & Lindsay, 2001). In a study in Bobo-Dioulasso, Burkina Faso, *An. gambiae* densities were found to be ten times higher within rice fields compared with the surrounding savannah area (Carnevale & Robert, 1987). Other studies in Tigray, Ethiopia showed a sevenfold higher incidence of malaria among children living in close proximity to micro-dams compared with those living more than 1500 metres away (Ghebreyesus *et al.*, 1998;

Ghebreyesus *et al.*, 1999). On the other hand, a study in Kenya found no additional risk associated with rice-field irrigation (Githeko *et al.*, 1993).

Anopheles density, survival, human biting rate and malaria infection incidence were monitored at 3 sites in the lower Moshi irrigation area in northern Tanzania. Each site represented different agro-ecosystems namely a sugarcane plantation a smallholder rice irrigation scheme and savannah with subsistence cropping. Although the vectorial capacity was four times greater in the rice field village than in the other 2 sites, the EIR in the rice field village was 61-68% lower. *An. arabiensis* was the principal vector of malaria in all 3 villages and sporozoite rates were respectively 0.01%, 0.1% and 0.12% in the rice field, sugarcane and savannah villages. The lower infection risk in the rice field village was attributed to relatively better socio-economic conditions and therefore greater use of anti-malarials and anti-mosquito measures (Ijumba *et al.*, 2002b). A further study on infection prevalence in this same population supported these findings (Ijumba *et al.*, 2002a).

1.10.4.4. Socio-economic and behavioural factors

The decreased risk of malaria associated with urbanization and rice cultivation is most likely due to the improved economic status and higher education levels of the populations living in these areas (Ijumba *et al.*, 2002a). Other studies have looked at the influences of types of housing on transmission. In rural Gambia, house entry by *Anopheles gambiae*, was reduced by the presence of a ceiling: plywood (59% reduction), synthetic netting (79%), insecticide-treated synthetic netting (78%), plastic insect-screen (80%, $P < 0.001$ in all cases) and closed eaves (37%, ns)

(Lindsay *et al.*, 2003). In a study in Principe, significantly lower malaria infection prevalence was found among people living in houses with eaves compared with those without eaves (RR=1.57; 95% CI=1.22-2.02; p=0.0008).

1.10.4.5. Effects of anti-malarial interventions

There are several examples of successful vector control programmes in SSA that were able to maintain low malaria transmission levels over several years. These include the insecticide residual house spraying interventions in the Pare-Taveta area of Tanzania during the 1950's (Draper & Smith, 1957; 1960; Pringle & Matola, 1967; Draper *et al.*, 1972), in the Gezira area of Sudan within the Blue Nile Project during the 1960s – 1970s (El Gaddal, 1990; 1991) and the highland malaria areas of Ethiopia (Malaria Control Programme Ethiopia, 1991) and Madagascar (Mouchet *et al.*, 1997). The main lesson learned from these projects is that transmission control in SSA needs to be sustained. If the interventions stop, infection risks rapidly revert to pre-intervention levels (Mouchet *et al.*, 1998; Hagmann *et al.*, 2003).

Significant reductions in malaria transmission resulting in up to 60% reductions in all-cause mortality have been associated with the use of ITNs under programme conditions (Alonso *et al.*, 1991; D'Alessandro *et al.*, 1995; Binka *et al.*, 1996; Nevill *et al.*, 1996). ITNs have also been shown to have a transmission reducing effect among communities not using bed nets themselves but living in proximity to bed net using communities (Howard *et al.*, 2000).

1.11. Malaria models for mapping malaria risks in sub-Saharan Africa

1.11.1. Historical global maps of malaria risk

Surprisingly few advances have been made in malaria mapping since Lysenko and Beljaev's 1969 description of the limits of the distribution of the disease in relation to the 60 °F (15.6 °C) and 70 °F (21°C) summer isotherms (Figure 1.7). The two summer isotherms reflect the ideal temperatures required for sporogony in *P. falciparum* (over 18 °C) and *P. vivax* and *P. malariae* (about 15 °C). Available disease atlases are outdated and over the years, relatively little investment has been made in recording and mapping malaria prevalence (Snow *et al.*, 1996). Perhaps the labour-intensive nature of traditional cartography has been a deterrent to mapping efforts particularly in the resource-constrained settings where the greatest burden of malaria is felt. Recently there has been a rekindling of interest in malaria mapping. With today's desktop mapping systems, data storage and updating is facilitated by the use of relational databases. Within these systems, representative malariometric data can be used to train environmental data in making predictions of malaria for wider unsampled sites. Malaria maps are thus evolving from being purely empirical distribution maps based on previously documented relationships between the environment and disease, to more process-based models of transmission dynamics and the influences on transmission described in the previous section.

Several regional malaria control measures have been implemented since the malaria endemicity distribution map of (Lysenko & Beljaev, 1969) was published. Their impact has been varied across ecological settings and today, malaria is localised to the

tropics and sub-tropics. A series of WHO records document changes in the global distribution of stable malaria between 1965 and 2002 (Pampana & Russel, 1955; World Health Organization, 1966; World Health Organization, 1997; Hay *et al.*, 2004) (Figure 1.8).

Figure 1.7. The geographical limits of malaria (Lysenko & Beljaev, 1969)

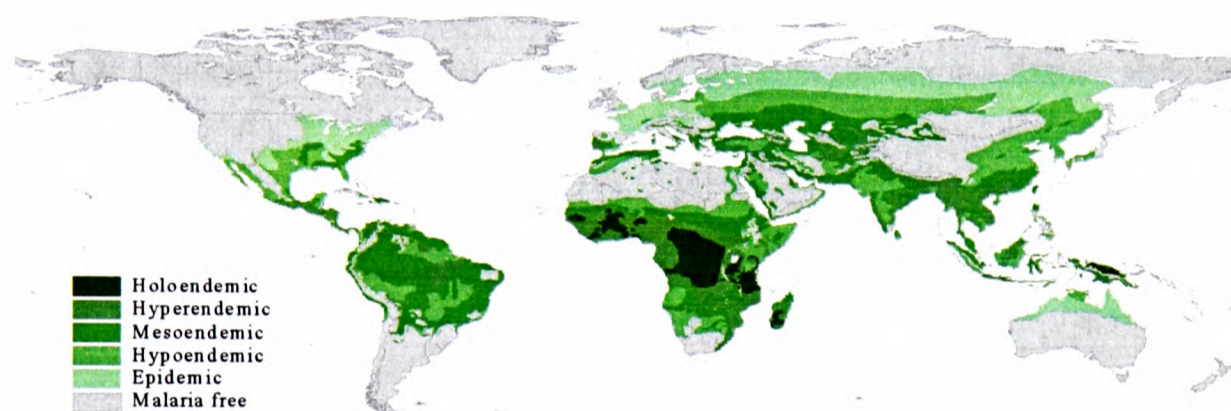
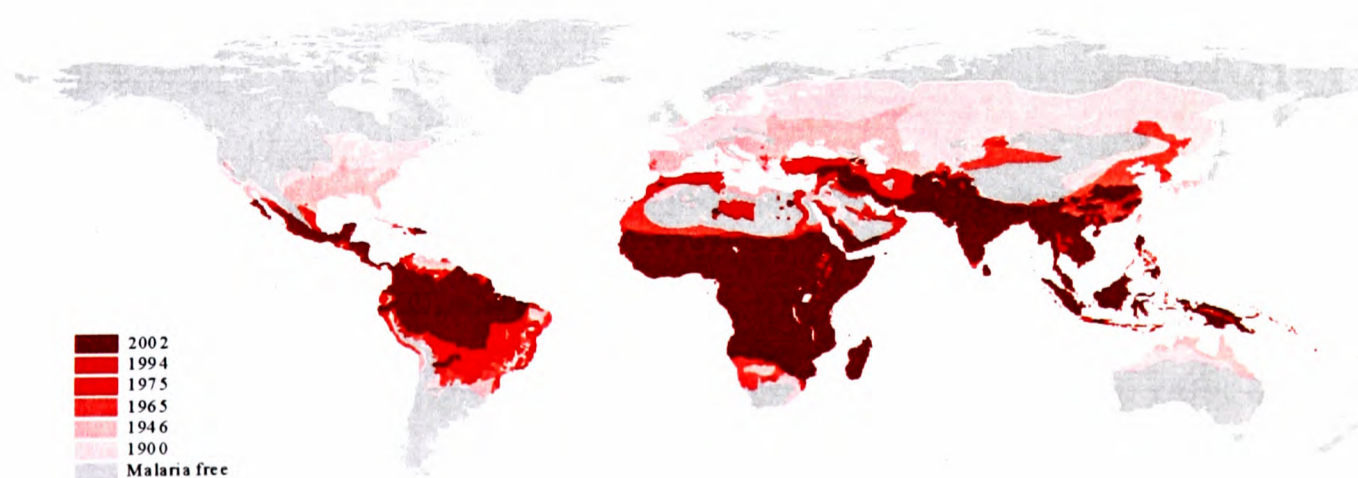


Figure 1.8. The geographical distribution of malaria infections (Hay et al., 2004)



As described in section 1.4, the risks of malaria morbidity and mortality vary within these distributional limits. To date, limited work has been carried out to define the different categories of malaria risk geographically. Major reasons for this are a poor understanding of how infection risk translates to disease outcome and a limited availability of reliable empirical disease data to determine risk for most of SSA. In the absence of data, environmental, mostly climatic, correlates of malaria transmission have been used to model malaria in data deficient areas.

1.11.2. Climate data and remote sensing applied to vector-borne disease mapping

One of the most significant boosts to vector-borne disease mapping has been the increased availability, range and quality of climate-surrogate data measured by satellite sensors positioned in space (Hay *et al.*, 1996; 1997; Rogers, 1998; Hay, 2000; Rogers *et al.*, 2002). Three climatic variables, or composite measures of them, tend to be modelled frequently namely, temperature, vegetation moisture availability/humidity and rainfall. Such models will form the basis of future studies of disease incidence and prevalence. This section is an overview of the evolution of malaria modelling in SSA from historical times to the present day and a critique of its current status.

Vector-borne disease (VBD) maps have proven effective as decision support tools in designing, monitoring and assessment of the impact of disease control interventions for several diseases. Examples include health systems research studies (Tanser, 2002), the monitoring of guinea worm (*Dracunculus medinensis* (Clarke *et al.*, 1991)), lymphatic filariasis (elephantiasis) and schistosomiasis (*Schistosoma*

haematobium) control (Brooker *et al.*, 2001; Malone *et al.*, 2001), and for mapping animal trypanosomiasis distribution (Rogers *et al.*, 1996). The potential of maps is also being recognised increasingly for describing malaria's epidemiology (Hay *et al.*, 1998a; Craig *et al.*, 1999; Kleinschmidt *et al.*, 2000; Kleinschmidt *et al.*, 2001a; Kleinschmidt *et al.*, 2001b; Kleinschmidt *et al.*, 2002) and estimating disease burden (Snow *et al.*, 1998; 1999a) A more complete review of the applications of malaria maps in SSA is provided in (Hay *et al.*, 2000a). Common to all VBDs is the impact of environmental parameters, notably climate, on the distribution, abundance and dynamics of their respective vectors.

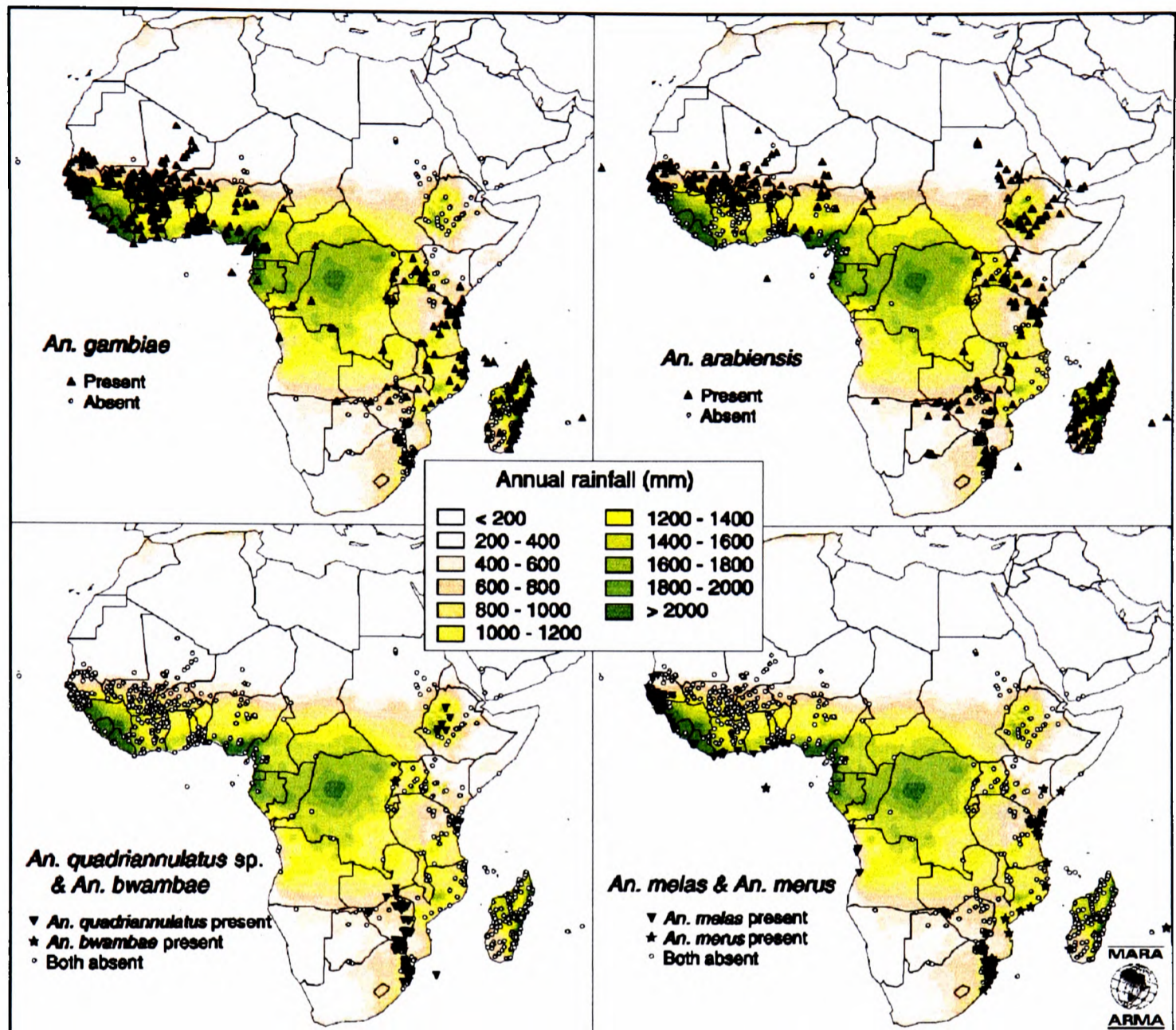
Accurate and verified models that translate climate-related processes into more detailed maps of malaria, however, are scarce. There are some malaria models available and these vary in complexity. The simplest are hypothetical, mapping areas where climatic conditions are suitable for the development of *Plasmodium falciparum* and the *Anopheles* mosquito. A limited number of transmission models have used empirical disease data to define suitable habitats and train climate models (Snow *et al.*, 1998; Kleinschmidt *et al.*, 2000; Kleinschmidt *et al.*, 2001a). The models also vary in the number or types of climatic variables included and the statistical methods applied. Two sources of climate data have been explored for malaria modelling, namely; spatially interpolated meteorological station data and satellite-sensor derived surrogates of climate.

1.11.3. Mapping the geographical distribution of Anopheles in SSA

The earliest map of the distributional limits of the main *Anopheles* species in Africa was based on field observations (Gillies & de Meillon, 1968). This map has since been updated with field studies on the *An. gambiae* species complex (Davidson & Lane, 1981; Gilles & Coetzee, 1987). More recently, the field records (n=2537 studies dating from 1944) have been transferred to a GIS (Coetzee *et al.*, 1993; Coetzee *et al.*, 2000; <http://www.wits.ac.za/fac/med/entomology/medento.htm>; (Figure 1.9)).

Although a very useful and now readily updateable record, this species distribution map is limited in several aspects. Firstly, many areas have not been sampled. Secondly, species were assumed absent at sampling sites if they were not recorded when other species were sampled such that it is possible that species were missed. Finally, species identification methods vary between surveys and identification and locality inconsistencies are inevitable. However, these data have been used to train climate data in predicting species distribution in previously un-surveyed locations (Lindsay *et al.*, 1998; Bayoh *et al.*, 2001; Rogers *et al.*, 2002; Levine *et al.*, 2004). Some of these predictive models are discussed in the next sections.

Figure 1.9. Sites where members of the *Anopheles gambiae* complex have been observed (closed symbols) mapped against total annual rainfall. Open circles show sites where species were not recorded during collections of other species of the *An. gambiae* complex (Coetzee et al., 2000)



1.11.3.1. Mapping the relative abundance of *Anopheles* using climate data

Lindsay *et al.*, (1998) have used annual precipitation and annual and wet season temperature (Hutchinson *et al.*, 1995) to define the ranges of *An. gambiae* s.s. and *An. arabiensis* in Africa. 'Suitable' climatic conditions for the species were defined on the basis vector distribution maps (Davidson & Lane, 1981; White, 1985) that were scanned and digitised. The climatic habitats preferred by *An. gambiae* s.s. and *An. arabiensis* were defined empirically using field data from West Africa. Non-linear

regression techniques were used to predict the relative abundance of the two species for the rest of Africa (Lindsay *et al.*, 1998).

The study found that relative abundance was predicted best by the ratio of five-monthly maximum precipitation to five-monthly potential evapotranspiration (P/PE). Significant differences ($p < 0.001$) were found in habitat preferences. *An. gambiae* s.s. showed a preference for more saturated areas ($P/PE > 1.0$) while *An. arabiensis* predominated in less saturated locations ($P/PE < 1.0$). The model showed high correlation with field observations from Tanzania ($r^2 = 0.745$).

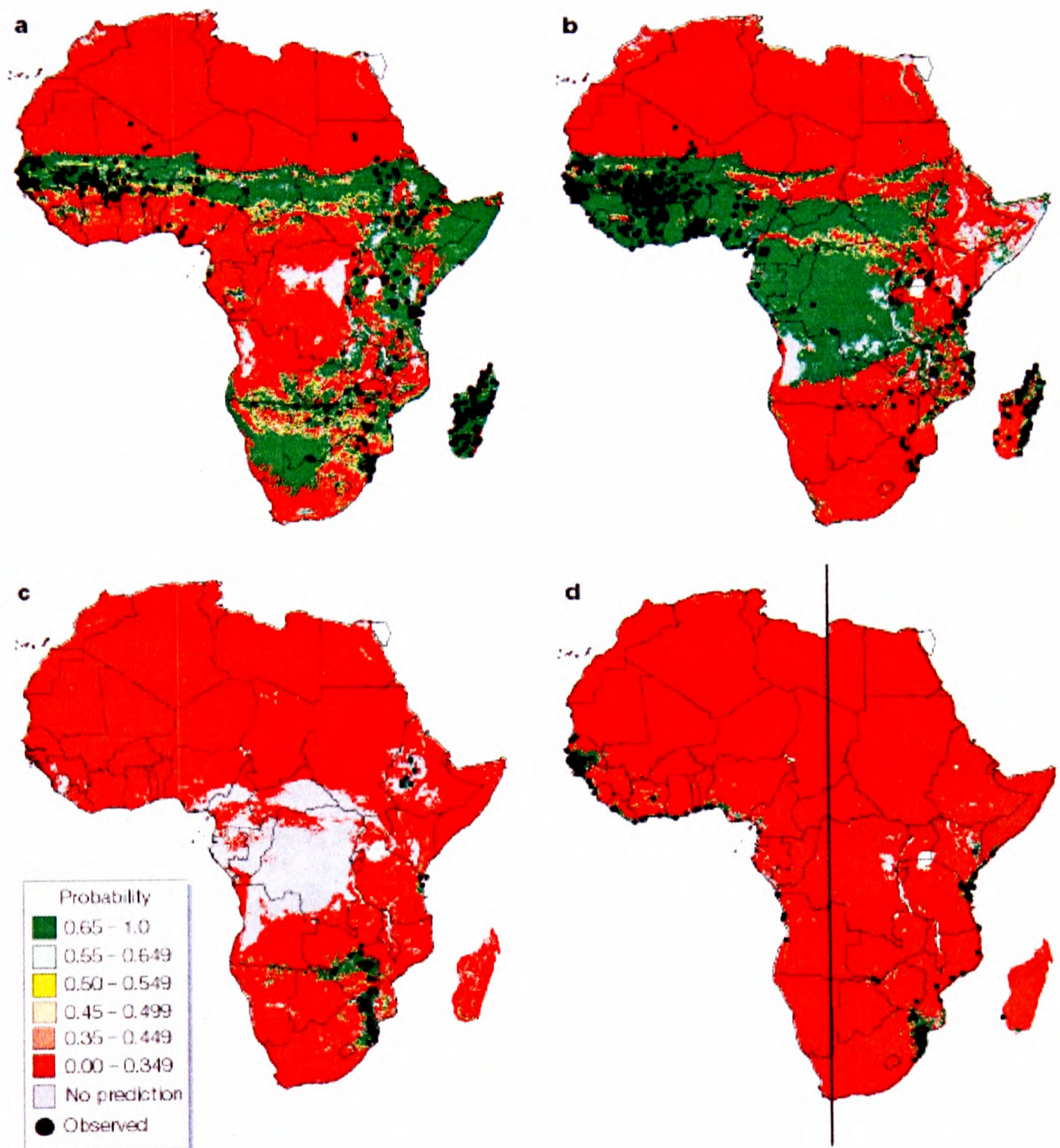
1.11.3.2. Predicting malaria vector species distribution using satellite imagery

Climate models of vector-borne disease are essentially models of environmental suitability for a vector. Areas where environmental conditions are unsuitable are classed as outside the distributional limits of the vector. Within suitable environments, predictions of the distributions of *An. gambiae* s.s., *An. arabiensis*, *An. merus*, *An. melas* and *An. quadrimaculatus* have been made for Africa between latitudes 18°N–30°S using data from sensors of the National Oceanic and Atmospheric Administration's Advanced Very High Resolution Radiometer (NOAA-AVHRR) and Meteosat High Resolution Radiometer (HRR) satellites (Figure 1.10) (Rogers *et al.*, 2002) and a 0.05 degrees spatial resolution digital elevation model (EROS Data Center, 2002). The satellite sensor derived meteorological surrogate data included thermal variables *i.e.* reflectance in the middle infra-red (MIR) channel and land surface temperature (LST) from the AVHRR and a surrogate for rainfall, the cold cloud duration (CCD), from Meteosat's HRR. Species 'presence' sites were

defined as areas within 0.15° of an observation site and 'absence' was determined from sites where only other species recorded. Randomly selected pixels from the presence and absence sites were used to train the environmental data. K-means clustering methods applied to the environmental and altitude data were then used to identify natural presence and absence clusters.

Using maximum likelihood discriminant analysis, 10 environmental variables that best discriminated between absence and presence classes were selected for each species. Selection was carried out in a stepwise manner with variables in each round of selection added on the basis of generation of the maximum index of agreement (*kappa* statistic (Cohen, 1960; Landis & Koch, 1977)). The probabilities of membership in the absence or presence classes were calculated for each pixel and the presence/absence status mapped for the rest of the continent. The predictions were compared with the training data and an assessment of predictive accuracy made. *Kappa* values were above 0.6. Landis & Koch (1977) suggest the following interpretation of observer agreement defined by the *kappa* statistic. Agreement is poor where $k < 0.4$, good where $0.4 < k < 0.75$ and excellent where $kappa \geq 0.75$. Good to excellent agreement was achieved for all species predictions.

Figure 1.10. Distributions of five sibling species of the *Anopheles gambiae* complex in Africa, predicted from temporal Fourier-processed satellite data (Rogers et al., 2002). a) *An. arabiensis*; b) *An. gambiae* s.s.; c) *An. quadriannulatus*; d) *An. melas* (West African coast) and *An. merus* (East African coast) (Rogers et al., 2002)



1.11.4. Modelling the seasonality of malaria

The intra-annual or seasonal cycles of infection risk characteristic of African malaria are also driven by climate (Gill, 1920; 1936; 1938; Onori & Grab, 1980). Seasonality is particularly marked in areas of low infection challenge where effects of immunity are minimal and fluctuations in challenge are manifested in changes in disease levels. In these areas, the morbidity and mortality from malaria infection is concentrated during a defined period of the year, the transmission season.

The length of the transmission season in East African countries was defined in historical expert-opinion malaria maps (Government of Tanganyika, 1956; Butler, 1959; Government of Uganda, 1962). These are discussed in more detail in Chapter 6. Since these early maps, very few seasonality maps have been developed mainly due to a lack of data either on seasonal indicators of challenge or intra-annual changes in disease. Seasonal estimates of vector abundance, for example, are lacking for most of Africa (Rogers *et al.*, 2002) while disease data are either lacking or unreliable (Snow *et al.*, 1999b). Climate-driven predictive models have been used to describe the duration of transmission seasons across Africa (MARA/ARMA collaboration, 1998) and in Kenya (Hay *et al.*, 1998b). These models are described in this section.

1.11.4.1. The duration of malaria transmission seasons in SSA

Interpolated climate surfaces have been used to predict the duration of malaria transmission seasons at a 5x5 kilometre spatial resolution in Africa (MARA/ARMA collaboration, 1998). Temperature and rainfall thresholds necessary to sustain

transmission were defined for map pixels using mean monthly temperature and rainfall data (temperature threshold=19.5 °C; rainfall threshold \geq 80mm monthly rainfall). The number of months in a year where threshold rainfall and temperatures coincided was then recorded for each pixel. A 'seasonality' surface with 3 'length of season' classes (1-3 months, 4-6 months and 7-12 months) was then created. This map provides a crude low resolution classification of the duration of malaria seasons. No attempt was made to verify the map with empirical evidence and so no assessment of its accuracy can be made.

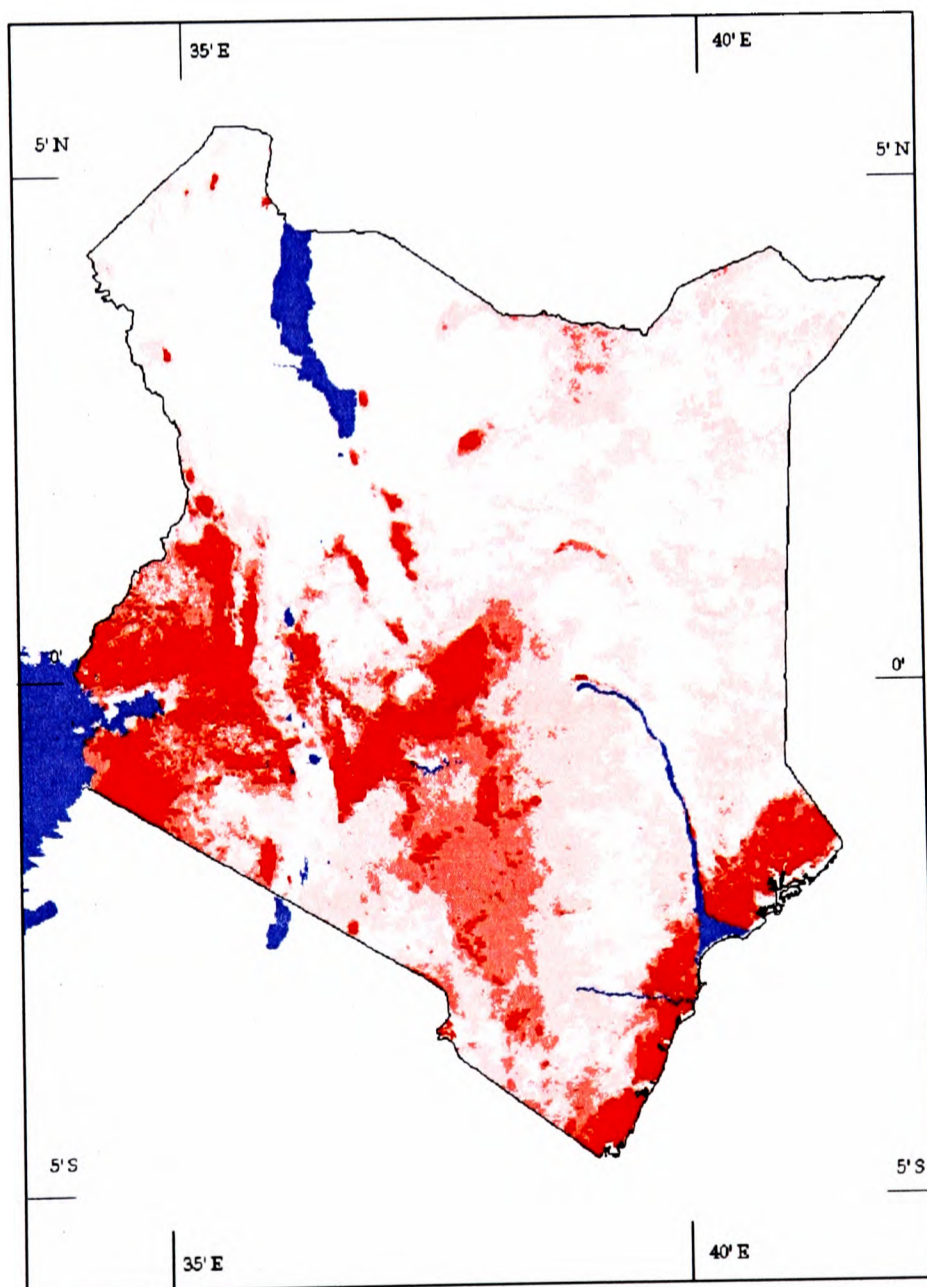
The distribution of seasonal vs. perennial malaria has also been mapped using Boolean thresholds for temperature (16-32 °C) and a moisture index based on the ratio of precipitation to potential evapotranspiration (Martin & Lefebvre, 1995). Areas where threshold conditions were experienced for \geq 7 months in a year were classified as experiencing seasonal transmission.

1.11.4.2. Predicting the onset of malaria seasons in Kenya using clinical data and satellite imagery

The seasonal abundance of *An. gambiae* s.l. in The Gambia has been shown to relate to the NDVI (Thomson *et al.*, 1996). Hay *et al.* (1998a) have also demonstrated a relationship between the NDVI and hospital admission rates for childhood malaria. This has been used to predict the onset and duration of malaria transmission seasons. Monthly malaria admission rates expressed as a percentage of total annual malaria admissions in 3 hospitals in Kenya were compared with contemporaneous monthly NDVIs.

Childhood malaria admissions correlated with the previous month's NDVI ($r^2=0.71$). Malaria seasons (defined as admissions $>5\%$ of annual total malaria admissions) were associated with an NDVI threshold of >0.35 . The duration of a 'malaria season' was computed by counting the number of months during which NDVIs were at or above this threshold. Seasonality was then classified as in historical malaria maps for East Africa (*i.e.* 0-3 months, 3-6 months and >6 months of transmission per year) presented on an 8x8 kilometre resolution map (Figure 1.11).

*Figure 1.11. Predicting the seasonality of malaria in Kenya. The number of months in a year during which *P. falciparum* transmission is possible are predicted based on an NDVI threshold of 0.35 (Hay et al., 1998a). ■ =>6 months, ■ =3-6 months, ■ =<3 months transmission, □ =no transmission.*



1.11.5. Modelling malaria risk: how much malaria?

A further step in malaria mapping has been to model the stability and intensity of malaria transmission within stable endemic areas using interpolated climate surfaces and satellite sensor-derived climate surrogates. Following are examples of these models from SSA countries.

1.11.5.1. A fuzzy logic climate suitability model of malaria transmission

Most ecological measures have some uncertainty and do not necessarily fall into distinct Boolean categories. Climatic conditions, for example, can be suitable for malaria transmission along a gradient of suitability relative to 2 extremes, 'suitable' (S) and 'unsuitable' (U). Such data can be described as *fuzzy sets*. Fuzzy sets define vague concepts and admit the possibility of partial membership. The degree to which an object belongs to a fuzzy set is denoted by a membership value between 0 (U) and 1 (S). The relationship between a variable and its fuzzy membership value can be described according to four types of fuzzy membership curves or functions: sigmoidal, j-shaped, linear or some other user defined curve. The fuzzy functions are controlled by four points ordered from low to high on a measurement scale.

Sigmoidal fuzzy membership curves have been used to rank climate variables derived from spatially interpolated weather station data (Hutchinson *et al.*, 1996) according to a scale of 'suitability' of climate for transmission (Craig *et al.*, 1999; Figure 1.12). Rainfall thresholds were defined from an investigation of long-term monthly rainfall and monthly mean and minimum temperature patterns in 20 sites in Africa where the malaria transmission profiles have been described previously (Smith *et al.*, 1993;

Some, 1994; Dolo *et al.*, 1997; Snow *et al.*, 1997). The sites analysed encompassed perennial, seasonal, epidemic and malaria-free epidemiologies. A threshold number of *Anopheles* must survive winter months to transmit malaria during warmer months. A period of winter temperatures (< 4-6°C) during a year is sufficient to kill vector populations (Detinova, 1962). The model, assigned a fuzzy membership of 0 to these areas. Optimum temperatures for sporogonic development in the mosquito were defined as 22°C - 32°C (suitable; fuzzy suitability =1) while <18°C and >40°C was considered unsuitable (fuzzy suitability=0). A mean monthly rainfall of 80 millimetres for at least 5 months in a year was defined as suitable.

Each map pixel was assigned a fuzzy membership value (y) for each climatic predictor variable according to sigmoidal relationships described by the equation:

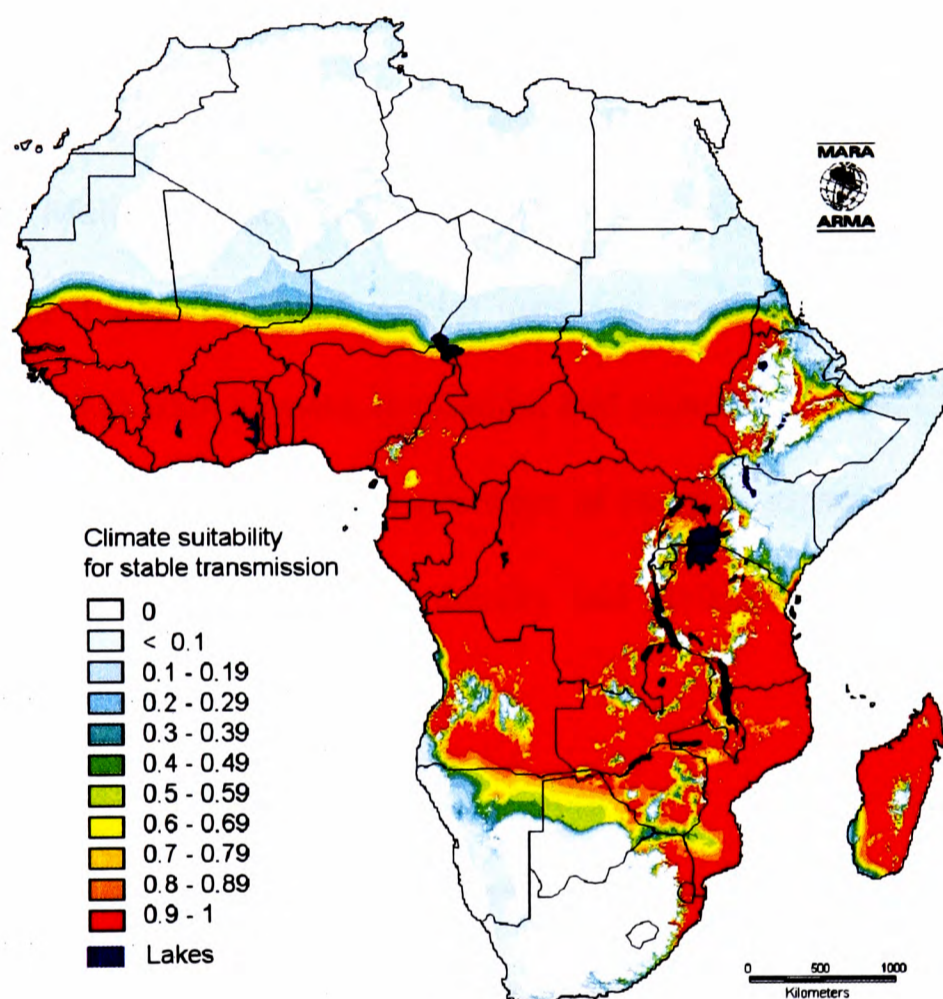
$$y = \cos^2 \left[\frac{x - U}{S - U} \right] \times \frac{\pi}{2} \quad (\text{Equation 1.3})$$

For a positive fuzzy relationship between fuzzy suitability and predictor variable (*e.g.* suitability for sporogony and increasing temp), y is the fuzzy suitability of predictor variable x . (For an inverse relationship the fuzzy suitability is defined by $1-y$). S is the value of x when conditions are suitable for malaria transmission and U the value when conditions are unsuitable.

A suitable rainfall and temperature need to occur concurrently for optimum transmission conditions, the number of months where threshold conditions were met was included as a parameter of the model. Due to ecological differences and variations in vector behaviour, a much shorter period of optimum conditions is required to build up sufficient vector populations in the more arid regions to the north of Africa (above 8°N of latitude) compared with more southerly regions. For the

model, 3 months a year was taken as suitable in northern and 5 months of optimal conditions for the rest of the continent. Using these criteria, the minimum suitability rating was calculated for each 8x8 km map pixel.

Figure 1.12. Fuzzy logic model of climatic suitability for stable P. falciparum transmission in SSA (Craig et al., 1999)



Although such maps delineate the broad geographical limits of areas where malaria may occur, they are purely climate driven and do not measure actual disease risk or any non-climatic factors that modify malaria's distribution. Transmission risk can only be defined by the use of empirical data. Furthermore the model uses long-term average climate data and so does not take into account the effect of seasonal climatic influences on transmission. This model, however, is the only recent map of malaria transmission distribution at a continental level and has been a pioneering tool for

defining the burden of malaria while taking account of differing transmission scenarios on burden estimates (Snow *et al.*, 1999b; 2003). The limitations of the map for Kenya are presented in Chapter 4.

1.11.5.2. A logistic regression model of malaria for West Africa

Empirical data have to a limited extent been used with climate data in making predictions for un-sampled areas. Models applying generalised linear regression are available for Mali (Kleinschmidt *et al.*, 2000) and West Africa (Kleinschmidt *et al.*, 2001a; Figure 1.13). The Mali model uses 101 malaria prevalence surveys dating from 1960-2000 to train climate, ecological and population data. Climatic explanatory variables included were long-term averages of rainfall, mean monthly maximum and minimum temperatures. Population density and distance to nearest water body were also modelled.

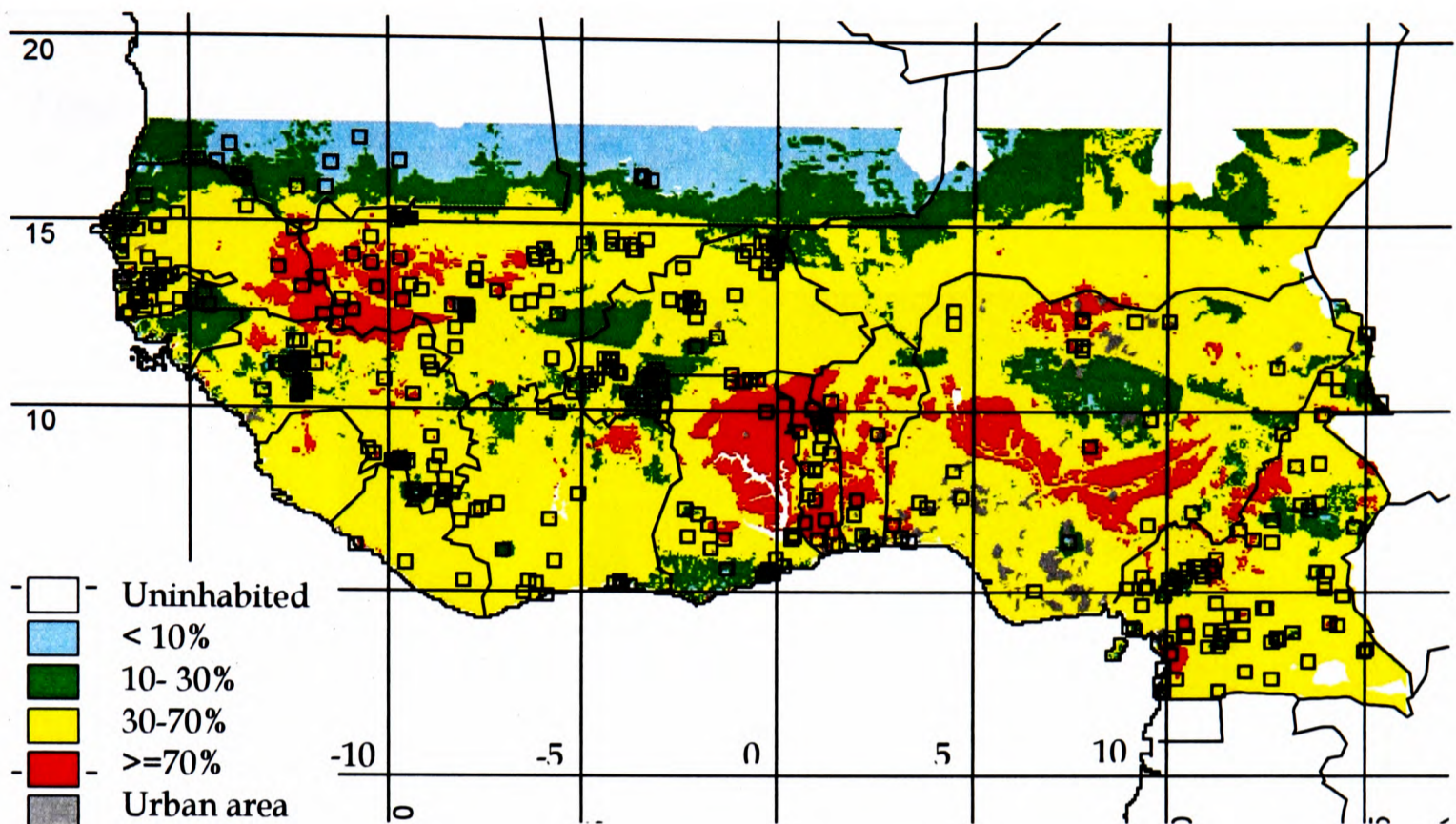
The West African model used step-wise multiple logistic regression to identify climatic and environmental risk factors of malaria prevalence defined using the PR between latitudes 1 °N and 22 °N and longitudes 17 °W and 16 °E. The model applies methodological approaches used by Snow *et al.*, (1998), Craig *et al.*, (1999) and Kleinschmidt *et al.*, (2000). The training data comprised PR surveys of children less than 10 years of age dating from 1970-1999. Climate data (Hutchinson *et al.*, 1996) were long-term average monthly rainfall, monthly averages of daily maximum and minimum temperatures, the NDVI, an index of drainage density and estimated population density. The monthly average climate data were summarized according to quarterly averages to correspond with the wet and dry seasons. This initial primary

model explained only 65% of the variation in malaria prevalence. Kriging was carried out on a local level to improve the predictions made (Kleinschmidt *et al.*, 2001a). Kriging uses a weighted linear combination of weights defined at known points to estimate values at unknown points (Isaaks & Srivastava, 1989).

A further consideration in areas where vector niches differ significantly is that models developed for one ecological zone are likely to perform poorly if applied to an ecologically different area. Studies of the distributions of trypanosomiasis (Rogers *et al.*, 1996) and schistosomiasis (Brooker *et al.*, 2002) highlight the importance of considering ecological zone similarity when extending VBD predictions to previously unpredicted areas. Kleinschmidt *et al.*, (2001a) modelled the interactions between environmental predictors of the PR and agro-ecological zones (AEZ) defined by the Food and Agriculture Organization (FAO) (1978).

The AEZ classification stratifies areas within West Africa according to the length of the period, during a year, that is suitable for growing crops based on the interactions of precipitation, evapotranspiration and soil moisture levels. Four AEZs are recognised, namely, the Sahel, the Sudan savannah, the Guinea savannah and a forest zone. (The Sahel and Sudan Savannah regions were combined for this study as they are similar climatically and also very limited data were available for the Sahel). The prediction method resulted in discontinuities along AEZ boundaries. A statistical model was derived for each AEZ and 4 categories of prevalence were predicted in each AEZ. 77.6% of observed points (N=450) were predicted accurately ($kappa=0.62$; $P<0.0001$).

Figure 1.13. Predicted prevalence of *P. falciparum* in children aged 2 to 10 years for West Africa (Kleinschmidt *et al.*, 2001a). Black squares show distribution of malaria prevalence survey sites

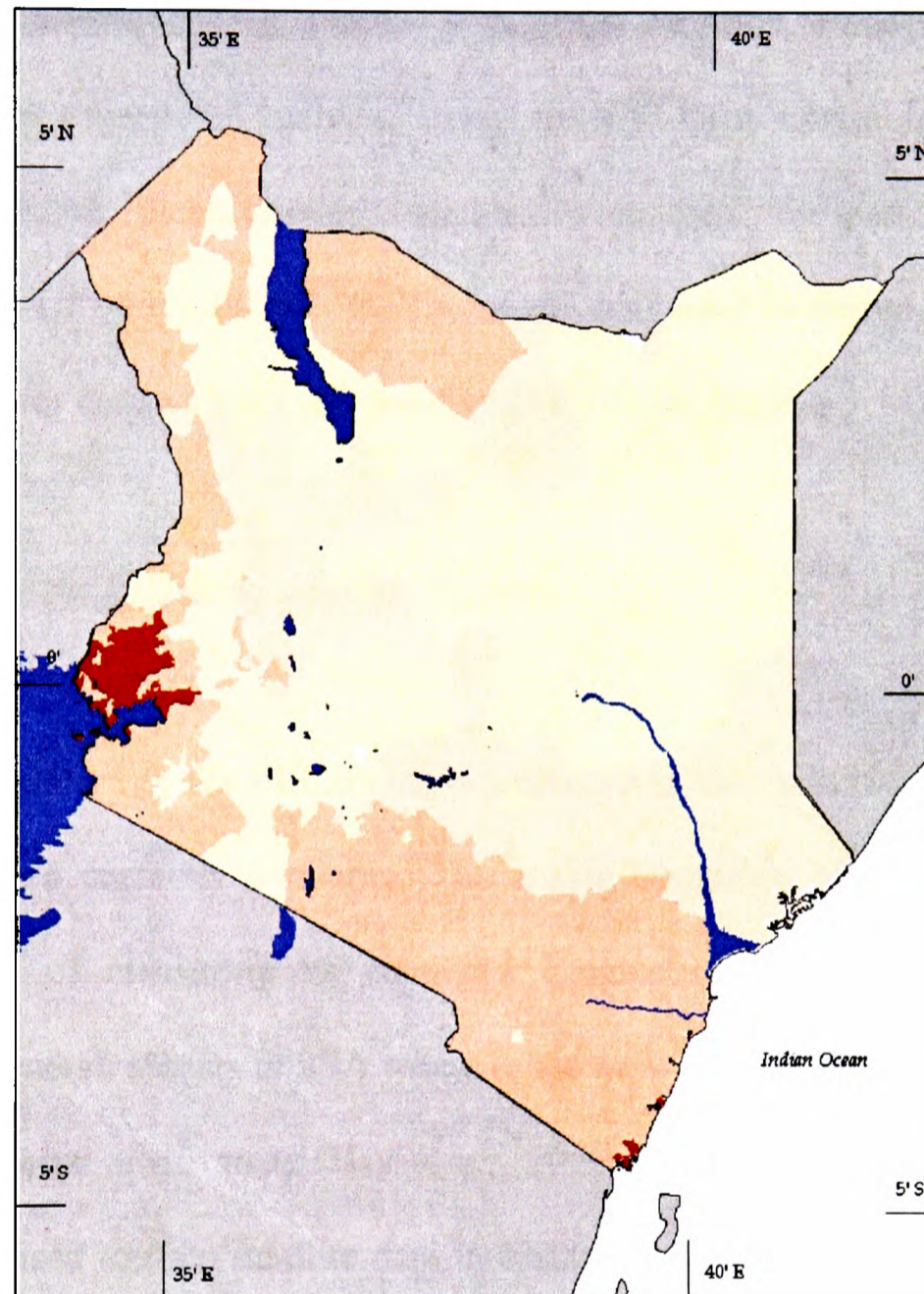


1.11.5.3. Malaria transmission model for Kenya: estimating 'at risk' populations

Snow *et al.*, (1998) have also used empirical data of malaria prevalence (PR), and interpolated climate surfaces to model malaria prevalence in Kenya (Figure 1.13). Community-based surveys (N=124) dating from 1966-1999 of children <10 years of age where at least 100 children were sampled were used in the model. 'Unstable' transmission areas were defined as 4th level administrative units (locations) with a fuzzy probability of malaria transmission <0.05 *i.e.* a less than 1 in 20 chance that climatic conditions within an average year are suitable for malaria transmission (Section 1.11.5.1, Craig *et al.*, 1999). For the rest of the country, linear discriminant analysis techniques were used to assign, each of the country's 1080 locations to one

of 3 endemicity classes based on a categorical classification of the PR; 0-<20%, 20-<70% and >=70% or low, moderate and high respectively.

Figure 1.14. Modelled predictions of stable malaria endemicity for Kenya (Snow et al., 1998). ■ =high, ■ =moderate, ■ =low, ■ =unstable endemicity.



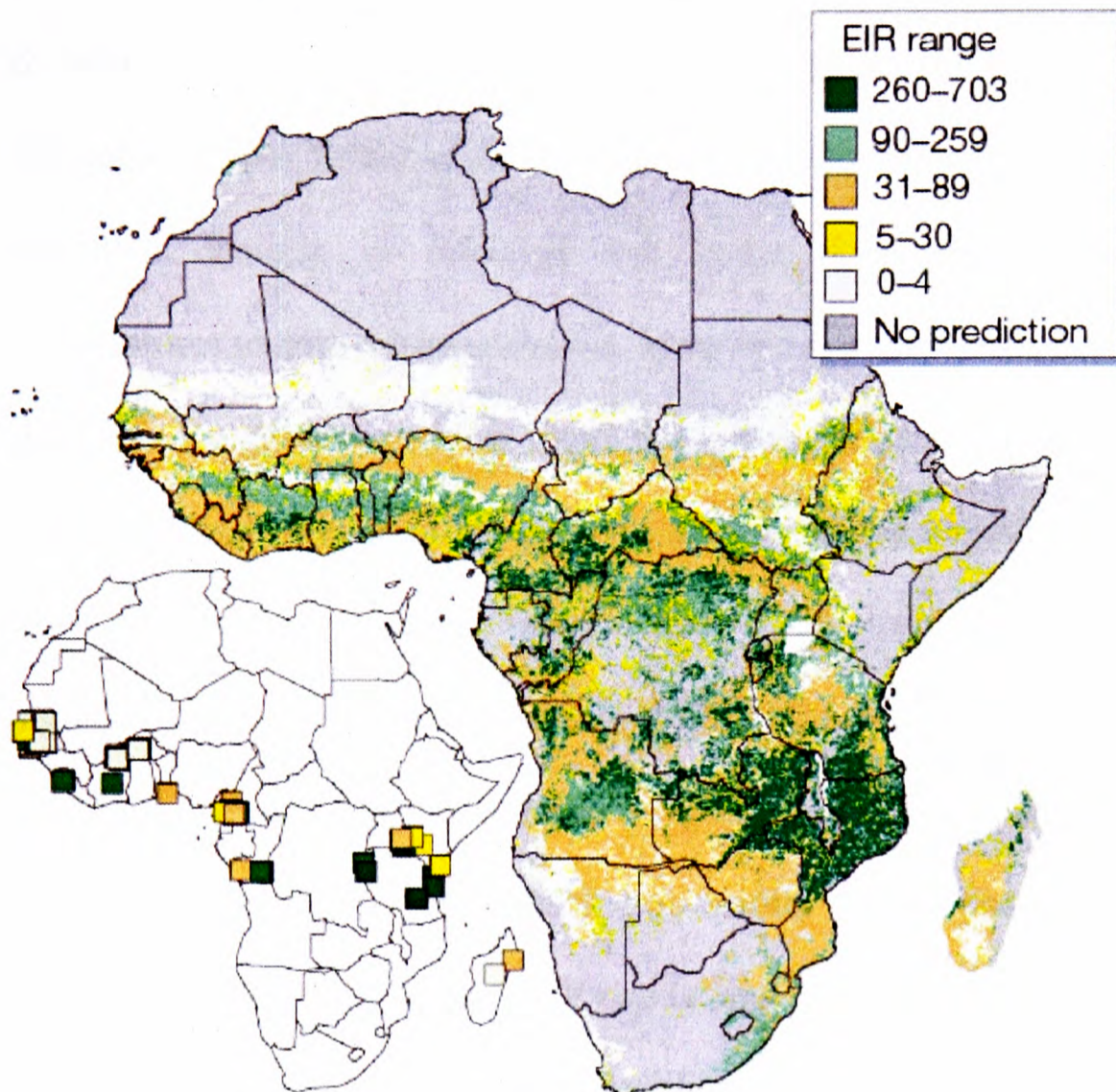
Eight significant ($p \leq 0.05$) discriminating variables that identified the 3 categories of stable malaria prevalence were identified. These were the number of months of transmission risk; rainfall in April and December; NDVI in April; minimum temperatures in April and December; and maximum temperatures in May and November.

A statistically significant correlation was found between the PR and total months of transmission risk (Pearson $r=0.62$, $P< 0.0001$). Linear discriminant criteria correctly identified 77.4% (95% CI: 70-85%) of each endemicity class and overall classification accuracy was 73% (95% CI: 65-80%). Two variables *i.e.*, the months of transmission risk and rainfall in April explained 25% of the variation in malaria prevalence. The combined model explained 46% of the variation in endemicity class. Additional multiple regression analysis, using the PR as a continuous variable, showed that combined, these climate variables, accounted for over half of the variation in the PR ($R^2=0.55$, $P< 0.001$). The model was used to derive estimates of populations at risk for each of the 4 risk categories modelled (Figure 1.14).

1.11.5.4. Predicting the EIR using satellite imagery

As described in Section 1.9, the EIR among populations is the most reliable measure of malaria challenge currently available. The major limitation of the EIR is the practical difficulty of measuring its elements. Estimates of the EIR are rarely available and, in general, studies in SSA where it has been measured have been short-term and have covered small areas (Hay *et al.*, 2000b). The more reliable of these studies have been used to train satellite data in making predictions of EIR for Africa (Rogers *et al.*, 2002; Figure 1.15).

Figure 1.15. Satellite-derived predictions of EIR in Africa. EIR data (map inset) were grouped into five approximately equal-sized classes of mean levels of malaria challenge. The same satellite data layers and analytical methods as used in Figure 1.10 are used to define the probability with which each continental pixel belongs to one of the five challenge categories. Insufficient training data were available to define EIR in those parts of the continent marked grey.



The satellite sensor data and analysis methods used were the same as those applied to the predictions of *Anopheles* distribution described in Section 1.11.3.2. Environmental data were able to discriminate between 5 equal-sized classes of EIR giving a high index of agreement (*kappa* statistic=0.77) between the predicted values and the training data. Predictions were not made for large areas of the continent due to the absence of training data representative of these areas.

1.11.6. Modelling malaria under changing climate scenarios

The potential impact of global warming on the distribution of malaria is a controversial subject (Rogers & Packer, 1993; Martens *et al.*, 1995;1997; McMichael & Martens, 1995; Lindsay & Birley., 1996; Reiter, 2000a; 2000b; Rogers & Randolph, 2000; Reiter *et al.*, 2004). The results from the few studies that have modelled changes in infection risk under different scenarios of increasing temperatures to date are inconclusive. Most of these models are primarily temperature driven and suggest that risk is likely to be greatest for populations living within the latitudinal and altitudinal fringe areas of malaria transmission (Lindsay & Birley, 1996; Lindsay & Martens, 1998). Here small increases in temperature could significantly raise vectorial capacity and the models predict an increase in seasonal and epidemic malaria in these regions (Martin & Lefebvre, 1995; Lindsay & Birley, 1996; Lindsay & Martens, 1998). Loevinsohn (1994) and Matola *et al.* (1987), working in highland areas of Rwanda and Tanzania respectively, have linked increasing temperatures at these sites with increasing malaria incidence rates.

The Intergovernmental Panel on Climate change (IPCC) has postulated six likely future climate scenarios based on different mathematical general circulation models (GCM (Houghton *et al.*, 1996)). Martin & Lefebvre (1995) and Martens *et al.*(1999) have modelled malaria's distribution under some of these scenarios. Although all models predict the spread of seasonal and epidemic malaria into higher altitudes and further latitudes, they do not concur on where the spread is likely to occur. As noted by these studies, many of the parameters of the GCM are poorly elucidated and so introduce biases. In addition, none of these studies were able to define current risk

accurately enough as a basis for prediction of changes in risk. The fit of these models to currently available maps of malaria is poor. Rogers & Randolph (2000) present a multivariate statistical approach to defining current malaria risk according to a relative measure of epidemic potential (EP). Infection risk was modelled by rearranging the R_0 to predict the vector/host ratio. The EP was then taken as the reciprocal of the vector/host ratio and this used to model predictions of malaria under climate change. Very few changes in the distribution of malaria were predicted even under the most extreme scenarios of climate change. The model by Rogers & Randolph (2000) includes covariates of malaria transmission other than temperature (precipitation and saturation vapour pressure) and cite these covariates as limiting factors in the future spread of malaria.

Currently available models of climate models are limited and future predictions of malaria distribution will need more accurate, higher resolution data. Furthermore, the effects of preventive interventions, host related factors and population adaptive processes in fringe areas are not possible to predict into the future. Most importantly an accurate geographical definition of the current distribution of malaria is a necessary pre-requisite for comparing future change scenarios.

1.12. Discussion of limitations of current malaria maps for SSA

An often-cited reason for the failure of malaria eradication was its sole reliance on vector-based measures. This lesson may also be applied to malaria mapping. Whereas vector distribution maps may be useful for identifying potential areas of transmission, the presence of vectors does not necessarily mean the presence of disease. Similarly suitable vector habitats do not necessarily mean malaria transmission occurs. This is perhaps the greatest drawback of solely climate driven models of vector distribution. Given that widely distributed and accurate disease data sets are unlikely to be available in the near future, the development of malaria models will continue to rely on climatic predictors. Climate data, however, have not been exploited to their full potential and there are several limitations of current data sets and prediction methods.

There has been a heavy reliance on spatially interpolated (SI) climate surfaces in malaria modelling (*e.g.* Hutchinson *et al.*, 1996). The effect of SI is to smooth out data and this results in the loss of information. The accuracy of SI depends on the variability of the climate measure *i.e.* less spatially variable measures, such as temperature, interpolate more accurately than very variable variables. Rainfall, in particular, varies markedly over small areas and this variation is lost during smoothing. In order to compensate for this, Hutchinson *et al.*, (1996) interpolated over small areas (tiles); 2 for temperature and 15 for rainfall. Although accuracy is improved by tiling, the rainfall stations from which the data were derived were also irregularly distributed. The interpolated climate surfaces also are limited by data gaps. As noted by the authors, records in many stations were inconsistent, for example rainfall recording was more frequent than temperature (Hutchinson *et al.*, 1996).

Furthermore, means summarised over varied time periods were used in developing single climate surfaces. The periodic aspects of malaria transmission are highly significant in describing the epidemiology of malaria and these aspects are removed when climate averages are used. The rainfall estimate (RFE) integrated approach using both SI and RS meteorological surrogates has markedly improved the accuracy of SI rainfall data (<http://www.cpc.ncep.noaa.gov/products/fews/rfe1.html>). There is much potential for further improvements in rainfall estimation using such combined approaches.

Craig *et al.* (1999) have defined a 'transmission' season based upon the number of months where climatic conditions are suitable for transmission. 'Suitable' was defined as 5 consecutive months of optimal conditions. In SSA, rainfall seasons closer to the equator tend to follow a bimodal profile while farther away, the pattern tends towards unimodal or a single rainfall season each year. The models do not consider which 5 months or take into account the rainfall patterns. The authors, however, did not find that this had a significant effect on the transmission model. GCM-based seasonal models have also been developed. In addition to the well-recognised biases inherent in the climate models themselves, the malaria models are almost exclusively temperature driven and do not consider the effects of covariates of malaria transmission. Furthermore, no attempts have been made to validate the models. Improved seasonal models will require better empirically defined data on the seasonal profile of transmission.

Linking very different multivariate data sets spatially can be problematic. Snow *et al.*, (1998) provide a novel approach to estimating disease risk in Kenya by modelling

disease level based on local data and applying it to up-to-date population maps. The malaria prevalence memberships and population data were summarised according to unequal sized polygons. For each of these, a malaria prevalence membership was assigned based upon the percentage surface area of each polygon located within that prevalence category. The limitations of this method are explained in the following example. A polygon where over 50% of the surface was 'high' was classified as high prevalence. If the population were concentrated in the lower prevalence region of the polygon, their risk category would have been misclassified. It is important, therefore, to consider the spatial distribution of population in such assessments. This model, however, demonstrates a very much-needed application for malaria models. It has been derived for a specific area and purpose and is based upon empirical data. The challenge, once again, is to improve on the data sets used for this purpose.

Kleinschmidt *et al.*, (2001a) have used PR data to train the climate data. As noted by the authors, their limited number poses problems for accurate generalised linear regression (GLM) as GLM is sensitive to sample size. Attempts were made to improve accuracy through kriging (Matheron, 1963; Oliver, 1996). GLM uses only the information of covariates at the specified geographical co-ordinates and in so doing tends to give predicted values that are close to the mean. Kriging, on the other hand, takes into account the values of all the covariates surrounding the specified point and this improves predictions to a degree. In assessing accuracy, another consideration for this model is the PR. The PR is sensitive to age and season and perhaps sample size. A limited attempt was made to control for age by limiting the data to a target age group. No control was applied for seasonality. Sample sizes of the PR studies vary in a non-random way through the data used. Such variations are

likely to affect the confidence of the PR estimate. The authors suggest that it may have been better to weight predicted values according to sample size but this was not done. Both this study and Snow *et al.*, (1998) have combined PR data spanning several decades. It is not clear whether the PR is stable over extended time periods and if this would affect the estimate of prevalence. Climate data for 1920-1980 were matched with malaria data from the late 1960s to 1996. The impacts of age and season on the PR and its effects on the estimation of risk of infection are discussed in greater detail in Chapter 2.

Attempts to validate maps are few and many models have relied on visual comparisons with historical maps. Craig *et al.*, (1999) based their validation on an overall 'visual' comparison with historical distribution maps and incidence data for Southern African countries. The validation of climate models is still in its infancy. There is a need in particular for contemporary malaria data to perform such validations.

As is clear from this review, there are several methods for vector-borne disease modelling already available. What is also clear is the lack of empirical disease data either for training models or for verifying them. Models have used different variables and varied techniques but the message for all is the same, climate can be used to predict malaria. Improvements in modelling will require more accurate climate and disease data sets. Remotely sensed meteorological surrogate data sets have obvious advantages for disease modelling. Although their greater accuracy compared with interpolated climate surfaces has been shown (Hay & Lennon, 1999), their potential for investigating temporal phenomena in malaria transmission is, to-date, under

exploited. RS is also able to capture a much wider range of environmental data than climate alone. Information on water bodies, vegetation, soil types, humidity, and population can be made available.

The important points to note from Section 1.11. are that developments in the techniques of malaria modelling using environmental and empirical data are advanced. Much as these techniques are available, the lack of both climate and empirical datasets hamper the development of accurate high spatial and temporal resolution models that can be of use to malaria control planners at a regional or country level. Current malaria models have many limitations. Empirical database may improve in the future but this is unlikely given the cost and practical difficulties of field sampling. Modelling continues to rely on environmental data and a reliable, untapped source of this is RS. Malaria vector dynamics and *Plasmodium falciparum* transmission are driven by the interaction of several environmental conditions. RS provides the possibility of making predictions of the dynamics of malaria transmission with greater accuracy than has been achieved in the past. The following chapters deal with empirical data generated from East Africa and their use in new models, incorporating new RS data and non-climatic data, to provide more precise estimates of risk for this sub-region.

CHAPTER 2:

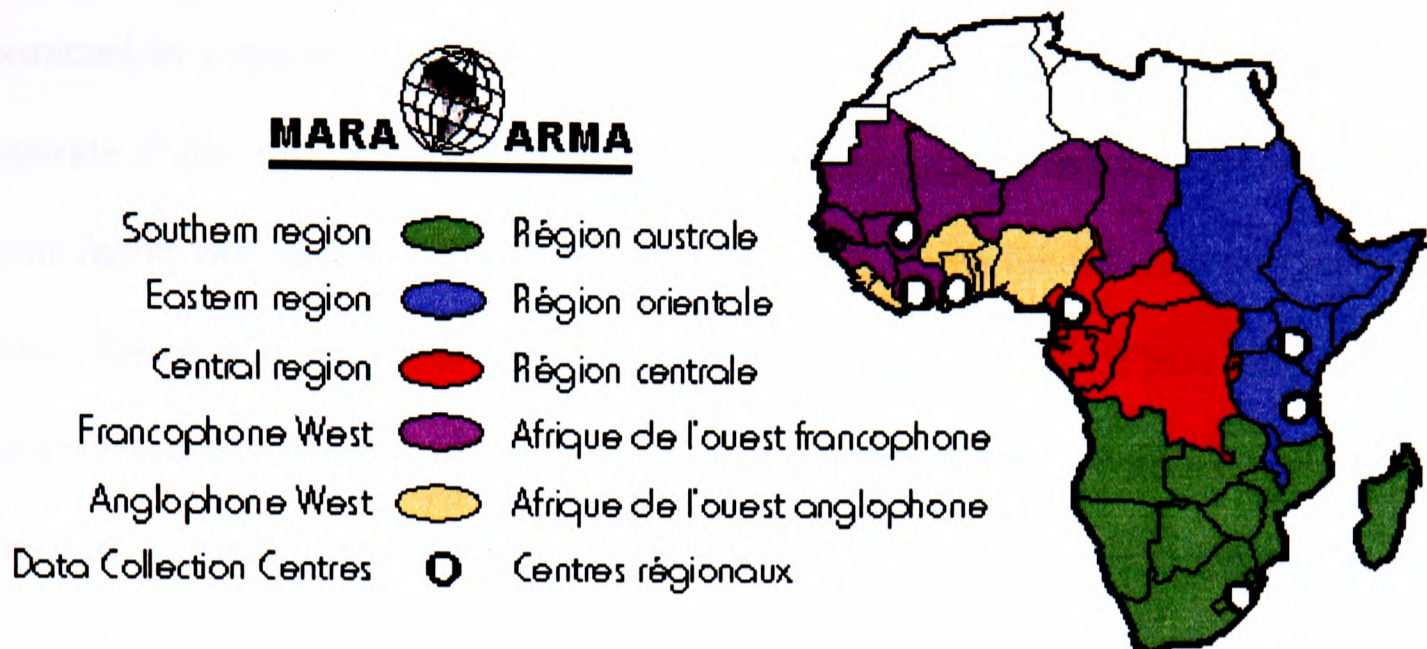
Materials and methods

2.1. Background

This study began in 1998 with an early attempt to collect empirical data on *P. falciparum* prevalence in Kenya (Omumbo *et al.*, 1998). This was prompted by the earlier work of Snow and colleagues (1994a; 1994b; 1997) that characterised the clinical epidemiology of severe disease and death directly attributed to malaria in relation to malaria endemicity showing differences in disease patterns with increasing transmission intensity (Section 1.4). This set the scene for a renewed effort to map malaria endemicity in Africa (Snow *et al.*, 1996). Following discussions with the International Development Research Council of Canada, Kenyan and South African investigators established the beginnings of the Mapping Malaria Risk in Africa (MARA/ARMA) program in 1998 (MARA/ARMA, 1998). The feasibility of building a continental data resource on published and unpublished parasitological data had been established (Omumbo *et al.*, 1998) and a network was formed to replicate this approach across the continent. Research foci were established as per Figure 2.1.

This thesis is based on the data collected as part of MARA for the East and Horn of Africa for which the author was responsible. The original scope of the thesis was to use all the data collected from the region but, as will become clear in Section 2.6, this objective was finally changed to represent modelling work among the three countries of the East African community: Kenya, Tanzania and Uganda. The availability of other data is described at the beginning of the chapter but the focus of further descriptive work and analysis is centred on the three East African countries.

Figure 2.1. MARA/ARMA collaborating centres (<http://www.mara.org.za/>)



2.2. The data entry proforma

The proforma used for data collection was developed in Kenya and modified later by other members of the MARA collaboration during early network meetings in Durban, South Africa. The proforma, shown in Annex A, was designed to capture a broad range of indices related to malaria risk, some of which were unique to special interests in different areas of the continent. For example, a section was included for malaria incidence data for countries in southern Africa who report annual case data at high resolution. Vector abundance and descriptions of entomological inoculation rates were also included for special interests and resulted in an early MARA publication of a vector distribution map as presented in Chapter 1 (Section 1.11.3). The thesis considers only the sections of the proforma related to the PR data.

The proforma included a unique identifier for each source and subsequent suffixes for repeat data from the same source. The first 3 digits referred to the ISO code for the country (for example, 254=Kenya, 255=Tanzania and 256=Uganda), the next 4 digits

referred to the report number and final 5 digits identified the PR data within the report. Each possible source of information was screened to ensure that data could be extracted in a standardised format as shown in Annex A. PR data were regarded as separate if the cross-sectional surveys were conducted at different times within the same report and separate if the same report provided details for different geographical areas. These were collapsed later as described in Section 2.6.2. Whenever possible, data were broken down by as small age groupings as possible. The month and year of the start and end of the survey, season, description of the study site, details on the methods used, *etc.* were all noted from the report. Care was taken to record as many details as possible on the study location (including photocopying maps contained within the report) to facilitate later identification of the study longitude and latitude (Section 2.4).

2.3. Methods of data collection

The data search strategy relied on multiple approaches to ensure a comprehensive coverage of all possible information. Electronic database searches were performed using Medline[®] (SilverPlatter International, Boston, MA, USA 2000), Popline[®] (Johns Hopkins School of Hygiene & Public Health, Baltimore, MD, USA) and EmBase[®] (Elsevier Science, Little Rock, Arkansas, USA 1999-2000). The following keywords were used in the search; *malaria, Africa, East Africa, Sudan, Ethiopia, Djibouti, Somalia, Ethiopia, Eritrea, Kenya, Tanzania, Uganda, parasite and malaria and Africa, Plasmodium falciparum, parasite rate, parasite prevalence and malaria transmission.* For each publication, bibliographies were cross-referenced to identify additional sources of information and other studies. Where additional details could

not be identified through the literature searches, authors were contacted to provide more information on geographical location, timing and age-specific characteristics of their parasitological data.

Pre-electronic journals held at the Wellcome Institute's Library in the grounds of the Kenyatta National Hospital as well as archive collections at the London School of Hygiene and Tropical Medicine and Oxford University were reviewed systematically volume by volume from the earliest available. In addition, national or regional, non-electronically referenced peer-reviewed journals were reviewed manually for additional data. Annual reports and unpublished reports from mission hospitals, drug companies and non-governmental organisations involved in health care delivery or research in each country were opportunistically accessed using local information on who works on what and the where their research work is carried out.

Visits were made to Uganda (December 1997 & January 2001), Sudan (October 1999 & July 2000), Ethiopia (February 2001) and Tanzania (August 2002). Postgraduate theses were searched manually in the libraries of the following departments and universities: University of Nairobi Medical and Chiromo Campuses and the Community Health Department; The Faculty of Medicine, Makerere University, Uganda; Faculty of Science, University of Khartoum, University of Gezira, Lion Hospital and the Wellcome Trust Library in Khartoum, Sudan; and Muhimbili University College of Health Sciences, Dar es Salaam, Tanzania. At these libraries and those of national institutes for medical research (Kenya Medical Research Institute, Nairobi; National Institute of Medical Research, Dar es Salaam) national conference proceedings, annuals reports and institute journals were also identified. Of

note were the annual reports of the pre-1980 WHO-established centres of the East African Institute for Medical Research and East African Institute of Malaria and Vector-Borne Diseases located at Amani, Tanzania.

Results of routine PR surveys undertaken by Ministries of Health in Kenya, Uganda and Sudan were of particular interest. In Kenya, the Division of Vector Borne Diseases (DVBD) was established in 1951 and has been routinely involved in vector-borne disease surveillance, including periodic PR surveys in communities and schools. The reports of these surveys were available at the DVBD headquarters in Nairobi and the archive was thoroughly searched. In addition it was possible to visit 7 provincial headquarters to identify additional material from the 45 field stations operated by DVBD to locate information not forwarded to headquarters. In Uganda, at the completion of the pre-eradication PR surveys, the office of the Malaria Eradication Pilot Project, located in Jinja, was closed and occupied by the MoH's Jinja Laboratory Training centre. While all project documents were removed (and largely destroyed), a small filing cabinet was left and remained unopened for many years. Fortunately, the raw data from Uganda's 1959-1960 pre-eradication pilot project that focused on 3 districts were found in this cabinet. One further extremely useful resource for historical data on PR in Tanzania was the reports provided by David Clyde (1967) in his book entitled *Malaria in Tanzania*.

2.4. Data geo-referencing

All parasite survey data were “geo-referenced” (*i.e.* their latitude and longitude determined). Most survey reports provided only the survey site name. To identify the precise location of each survey site, a variety of techniques and complimentary sources of information were used. First, maps provided in the reports were used to begin the more thorough positioning of each survey, unless the report provided the exact longitude and latitude extents of the survey. The next important sources were national maps of varying resolution and detail; the most useful were 1:50 000 scale topographic maps produced by the Directorate of Overseas Surveys in 1971 for East Africa. For those sites not discernable from topographical maps digital gazetteers were used which include place names and spatial co-ordinates these included: GEOname (GDE Systems Inc., 1995, San Diego California) and Africa Data Sampler (World Resources Institute, 1995). Finally a combination of correspondence with various research groups, particularly the International Livestock Research Institute (ILRI) for Kenya, allowed an approximation of spatial positions for those that could not be spatially defined through the original source report, topographical maps or electronic gazetteers.

Two variables were used to record the spatial position of the parasitological surveys: first the centroid of the survey site (longitude and latitude in decimal degrees) was recorded as a unique identifier to cross-reference spatial overlaps with other surveys; and second, a more detailed definition of the spatial extent of the surveys was coded and stored as polygons within the GIS platforms used for the analysis. For the latter, surveys were classified as representing one of five spatial extents: First, surveys

representing a single village where the central longitude and latitude was used to define point estimates covering the community using ArcView 3.2 (ESRI, Redlands, CA, USA). The second spatial classification reflected surveys that sampled from several villages but presented the data as a single PR for all the villages. In this case, a polygon was created to enclose the villages. The third, fourth and fifth spatial criteria corresponded to surveys undertaken at the 5th (average area covered 9.4 km²), 4th (average area covered 15.4 km²) or 3rd (average area covered 34.1 km²) administrative levels.

The true position of a point in space can be taken as the co-ordinates measured using a Global Positioning System (GPS). The accuracy of other geo-referencing methods can be measured against GPS readings and a root mean square error (RMSE) calculated. The RMSE is determined by calculating the deviations of points from their true position, summing up these residuals and then taking the square root of the

sum *i.e.* $\sum_{i=1}^{i=n} \sqrt{x_i^2 + y_i^2} / n$. Using a recently compiled database of 56 GPS-positioned

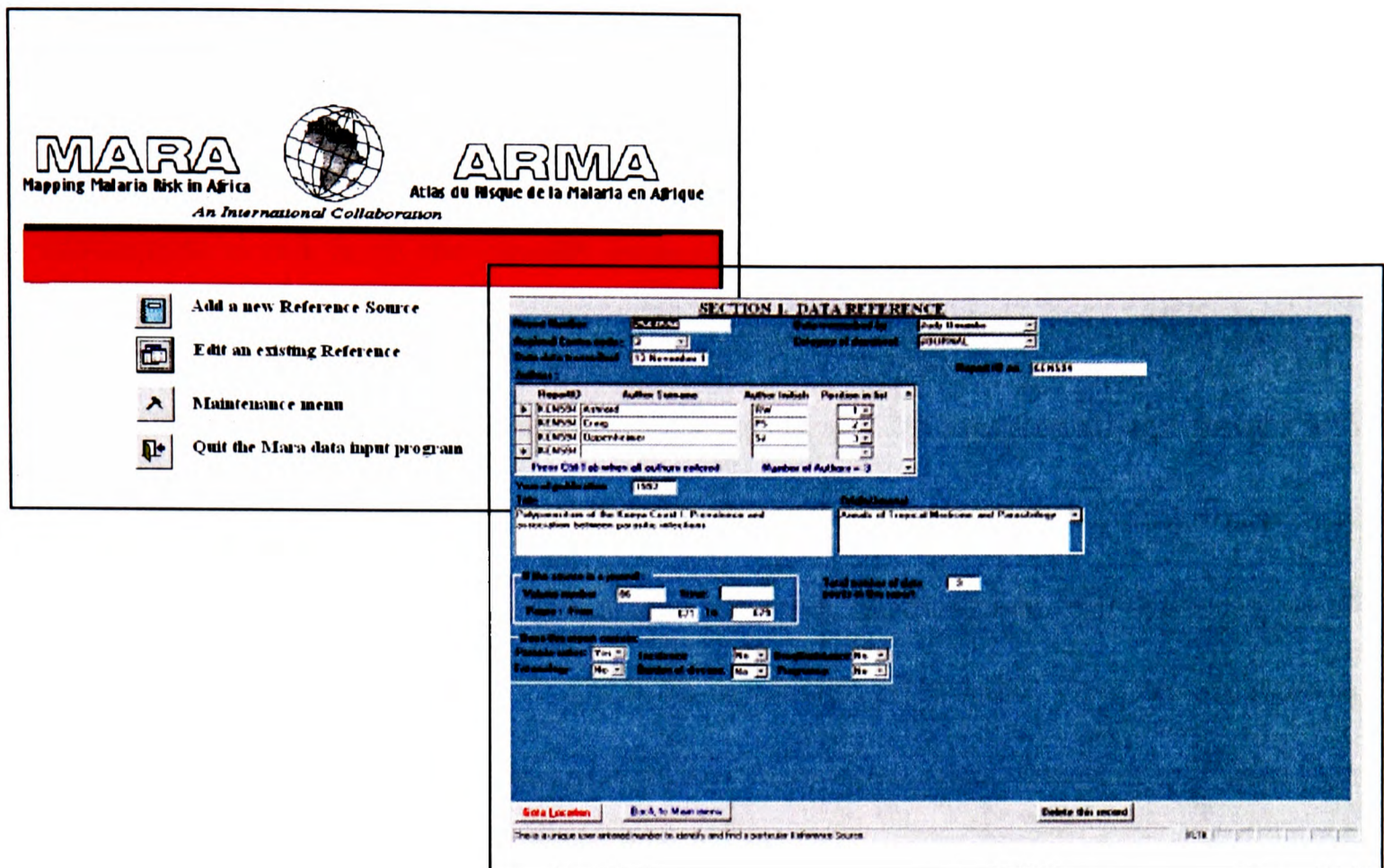
health facilities in Kenya (Noor *et al.*, 2004) RMSE estimates were made to compare longitude and latitude co-ordinates derived from 1:50,000 scale topographical maps and a digital 5th level administrative unit map and digital place name gazetteers. The RMSE were respectively 2.02×10^4 , 5.02×10^4 and 4.22×10^4 decimal degrees.

2.5. Data entry and verification

All the data were entered using a customised data-entry program developed in Microsoft Corporation Access97 Release 2 [Microsoft Corporation, 1989-1996 Seattle, Washington]. The data entry program was first developed by Informatica (Nairobi, Kenya) for the present project and later forwarded to the MARA collaboration where subsequent modifications were made. The data entry screens allowed easy entry of the data and had a number of skip and internal range consistency checks built in to reduce data entry errors (Figure 2.2).

Data were entered twice by different data entry clerks and both files were verified to detect data entry errors. These were corrected using the original proformas in both files by the data entry clerks who re-ran the verification program until both files were identical. Then a program was run to re-check consistency errors, for example inconsistencies in two date fields or district names for a given country. These were corrected on the final temporary file before being saved as an update file. When this file was complete for a number of proformas the data were then posted within the program to the main data file. The main Access file served as the primary source for future extractions. Data were later managed and ordered in Microsoft Excel Version 9 (Microsoft Corporation, 1985-1999).

Figure 2.2. MARA data entry screens



2.6. Data selection and descriptive analysis of data

2.6.1. Overall data identification

By 2004, 2308 spatially and temporally independent PR surveys had been identified for the East and Horn of Africa, with the earliest report being in 1907 (Table 2.1). The majority (43%) of the reports identified were from Kenya (Figure 2.3). This reflects the location of the investigator and also the more intensive investment in malaria research in Kenya compared to its neighbours over the years. Table 2.1 also shows the sources of information for each recorded survey, dates, age ranges and georeferencing methods. Only one report was available from Djibouti, two reports from

Eritrea and four reports from Somalia and these have not been included in the table or further analysis.

In Ethiopia and Sudan a large number of surveys could not be geo-positioned due to inadequate detail from the reports or inadequate reference material to identify the positions of the surveys, 36 (26%) and 42 (25%) respectively. Of those that could be positioned, among studies reported in Ethiopia, over 77% sampled all age groups and the results were not provided for children only, in Sudan the corresponding figure was 64%. Of the remaining 23 surveys in Ethiopia and 45 surveys in Sudan that sampled children only, 5 and 35 respectively were conducted after 1980. For these reasons it was decided to exclude Ethiopia and Sudan from subsequent analyses and focus on the three East African countries for which more appropriate, spatially referenced and contemporary data were available.

Figure 2.3. Distribution of parasitological survey data between 1907 and 2003 according to dates of surveys in the East and Horn of Africa

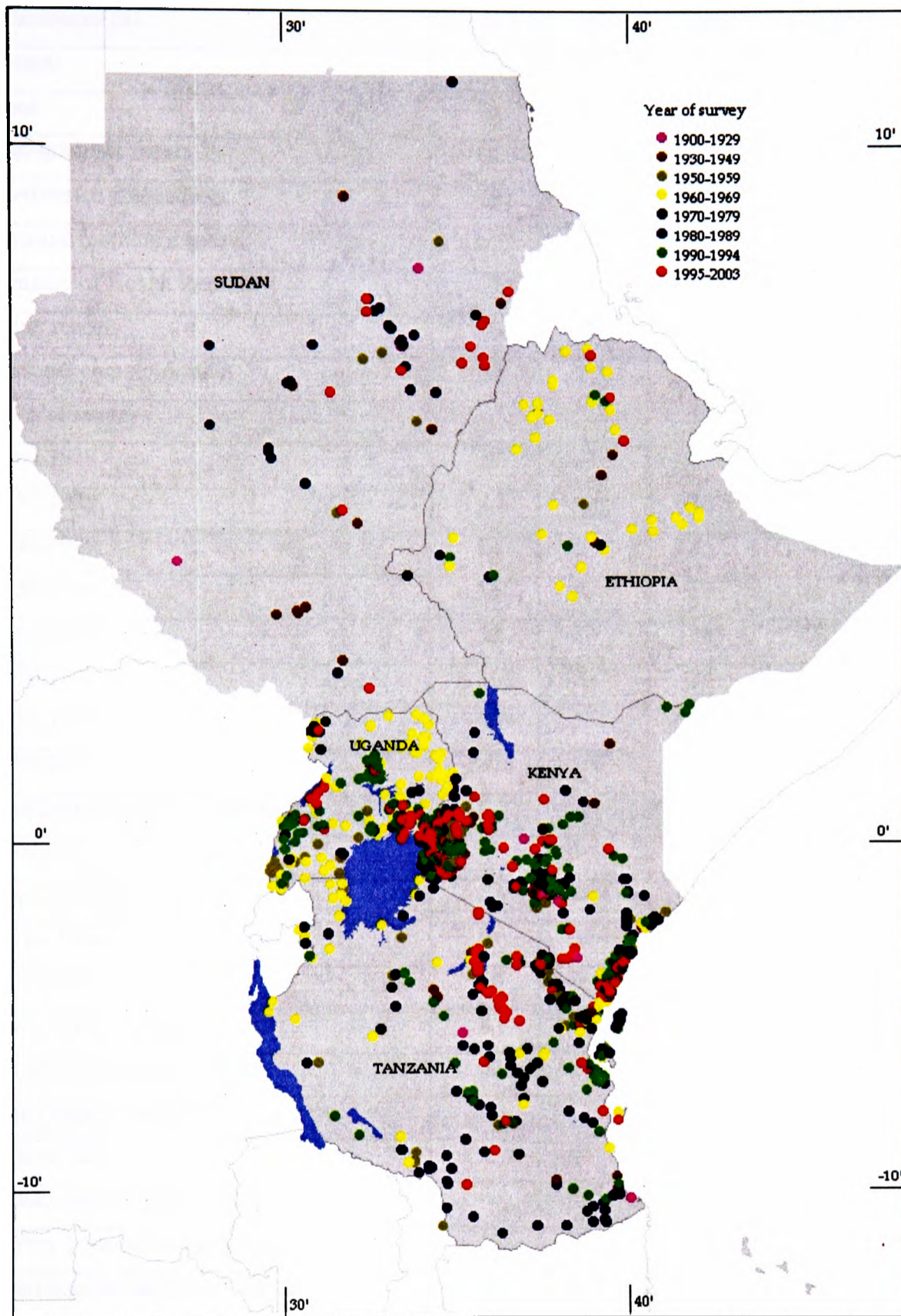


Table 2.1: Parasite survey report source descriptive data for Kenya, Tanzania, Uganda, Ethiopia and Sudan

Country*	K	T	U	E	S	Total
Data source (n)	992	522	489	138	167	2308
Journal	130	121	252	91	12	606
Book	2	66	-	-	-	68
Post-graduate theses	41	10	11	1	68	131
Conference proceedings	29	16	-	-	2	47
National Institute Reports	-	93	36	-	10	139
Ministry of Health Reports	683	35	77	-	59	854
Other reports	76	146	101	46	15	384
Personal communication	31	35	12	-	1	79
Date of surveys						
1900-1929	27	5	-	-	9	41
1930-1939	19		2	1	26	54
1940-1949	21	5	4	4	-	34
1950-1959	76	60	67	3	10	216
1960-1969	25	27	249	76	27	440
1970-1979	102	55	20	2	7	186
1980-1989	249	171	17	7	42	531
1990-2003	428	157	130	45	46	806
Methods of spatial positioning						
Could not be positioned	-	1	-	36	42	79
Topographical maps	585	109	106	16	21	837
Africa Data Sampler	6	394	42	32	3	477
GEOname	60	6	286	35	53	440
Coordinates in report	8	-	27	-	19	54
Other (digital map or GPS)	333	12	28	19	29	421
Age range of surveys						
Infants only	1	5	2	-	2	10
Adults aged 15 years + only	-	3	-	1	-	4
Survey covered all age groups 0-80yrs	52	77	43	35	40	247
Age range including 0-6 year olds	84	132	33	-	3	252
Age range including 0-15 year olds	508	197	179	56	60	1000
Age range including 5-15 years only	302	77	144	-	14	537
Age of sample not given in report	45	31	88	46	48	258

*K=Kenya, T=Tanzania, U=Uganda, E=Ethiopia, S=Sudan

2.6.2. Selection criteria for data in Kenya, Tanzania and Uganda

Rigorous selection criteria were applied to the primary database on the basis that it is important to establish a series of rules for the inclusion or exclusion of primary and secondary data from peer-reviewed and un-reviewed data to ensure that some comparability was achieved in the data to be analysed. Firstly, data were combined from individual survey reports if repeat surveys were undertaken by the same investigators within a 24 month period or for surveys undertaken by several investigators at different times in the exactly same location within a 24 month period. Surveys were spatially mapped in MapInfo Professional Version 7.0 (Troy, New York, USA) to identify a single estimate for one spatial unit. Where two or more surveys were identified in the same community undertaken > 24 months apart or with different age classes, the most congruent age class survey was selected followed by a selection for the more “random” survey (see below), the largest sample and then followed lastly for a criterion based on the most recent survey.

The surveys identified represented a variety of sampling approaches, ranging from household-level random studies to non-random or purposively sampled households that may have included volunteers, surveys of healthy attendees at maternal and child health (MCH) clinics, expanded programme for immunization (EPI) clinics, surveys of school children and anti-malarial drug-resistance screening surveys. In an effort to reduce those surveys which might have been subject to sampling error introduced through selective recruitment, only studies that recruited children at the community level were included in the analysis. These included total population, randomly sampled or longitudinal community surveys. Several community-based surveys,

particularly those undertaken by the Ministry of Health (MoH), provided little detail on the sampling strategy used to select the childhood populations. These surveys were retained but coded separately in the database to distinguish them from truly random selections should there be differences in estimates of PR dependent upon sampling methods.

Age ranges used during the parasitological surveys also varied between surveys (Table 2.1). Infection prevalence changes with age as a result of acquired anti-parasitic immunity resulting in lower prevalence of infection in older age groups in stable endemic areas (Section 1.4). Infants have a reduced prevalence of infection for a number of reasons, including reduced body size (Port & Boreham, 1980) and more importantly passively acquired immunity through maternal IgG (McGregor, 1965). Where possible, survey data have been pooled to provide a single estimate within the age range of 0-15 years but surveys which predominantly included children aged less than two years were excluded. Surveys were excluded if they covered age ranges beyond 15 years, largely those that reported for entire communities.

Sample size is likely to affect the accuracy of the prevalence estimate and to control for this only surveys where at least 50 subjects were sampled were included. Using a cut-off of 50 subjects was based on advised sampling strategies for parasitological surveys (Snow & Gilles, 2002). For all remaining surveys the precision of each estimate was calculated from the standard error of the proportion where the denominator and numerator were provided. For obvious reasons studies for which sample size details were unavailable were excluded.

The spatial extent of the surveys was also an important selection criterion. There were several studies that covered very wide, spatial areas, for example >4th level administrative units, which are areas of governmental administration two levels below the district level. These posed a particular problem for the present analysis as it was unclear the extent of the spatial sampling within these polygons and thus it is not clear how representative the parasitological data were of that polygon. All reports covering greater than 4th level administrative units in Kenya, Uganda and Tanzania were therefore excluded. In addition those that could not be accurately positioned were obviously excluded.

Data were finally reduced to provide spatially unique studies undertaken since 1927. The exclusions and data reduction strategy is shown in Tables 2.2 and 2.3 for the periods 1929-1979 and 1980-2003 respectively. The spatial distribution of the survey data pre-1980 and 1980-2003 are shown in Figures 2.4 and 2.5 respectively. These two time periods have been selected to ensure that modelling work presented in Chapters 4-7 is cognisant of any errors likely to be introduced resulting from long-term temporal changes in predictive variables within the models. Furthermore it is likely that the dependent data, the PR, might also have changed over time within the same spatial areas for areas of East Africa and due largely to growing parasitological failures to commonly used anti-malarial drugs. As such the data for the models considers only data since 1980 (whilst allowing for dates of surveys), the following section (2.6.3) examines the same-site changes in infection prevalence and possible changes over longer time periods are presented in Chapter 3.

Table 2.2: Selection criteria for PR survey data in East Africa and description of selected surveys pre-1980.

	Kenya	Tanzania	Uganda	Total
Total number of surveys identified	236	194	237	667
Exclusion criteria				
Excluded post-1979	756	328	252	1336
Excluded because unable to locate longitude/latitude	0	1	0	1
Excluded because no information on denominator	69	58	3	130
Excluded sample size <50	16	4	25	45
Excluded because non community sampled (or no details given)	38	31	88	157
Excluded sample age groups ≠ 0-15 or infants only	8	8	37	53
Excluded because spatial overlap with more contemporary or larger sample size survey	28	36	44	108
Characteristics of surveys				
Final survey samples included in analysis	77	56	40	173
Surveys reported in peer-reviewed journals or books	25	23	27	75
Surveys reported in Ministry of Health, NGO or donor reports	52	33	11	96
Surveys reporting total population sampling, longitudinal sampling, random sampling	27	42	36	105
Surveys covering period 1920-1929	6	0	0	6
Surveys covering period 1930-1939	2	1	0	3
Surveys covering period 1940-1949	5	0	1	6
Surveys covering period 1950-1959	11	18	2	31
Surveys covering period 1960-1969	17	16	30	63
Surveys covering period 1970-1979	36	21	7	64
Median (interquartile range) of sample sizes	165 [129, 315]	311 [170, 617]	93 [67, 373]	191 [112, 432]
Surveys covering age ranges within 0-5 years	0	9	18	27
Surveys covering age ranges within 0-10 years	60	40	20	120
Surveys covering age ranges within 0-15 years	17	7	2	26

Table 2.3: Selection criteria for PR survey data in East Africa and description of selected surveys between 1980 and 2003

	Kenya	Tanzania	Uganda	Total
Total number of surveys identified	992	522	489	2003
Exclusion criteria				
Excluded pre-1980	236	194	237	667
Excluded because unable to locate longitude/latitude	0	33	2	35
Excluded because no information on denominator	9	1	13	23
Excluded sample size < 50	82	9	69	160
Excluded because non-community sampled ^a	408	128	121	657
Excluded sample age groups ≠ 0-15 or infants only	20	39	19	78
Excluded because spatial overlap with more contemporary or larger sample size survey ^b	20	32	1	53
Characteristics of surveys				
<i>Final survey samples included in analysis</i>	217	86	27	330
Surveys reported in peer-reviewed journals	11	24	7	42
Surveys reported in reports of Ministry of Health, NGO or donor reports	168	34	17	219
Surveys reporting total population sampling, longitudinal sampling, random sampling ^c	143	86	16	245
Surveys covering period 1980-1989	54	32	2	88
Surveys covering period 1990-1994	100	14	14	128
Surveys covering period 1995-2003	63	40	11	114
Median [interquartile range] of sample sizes	221 [119, 431]	194 [146, 443]	88 [76, 265]	204 [120, 427]
Surveys covering age ranges within 0-5 years ^d	32	61	0	93
Surveys covering age ranges within 0-10 years ^d	134	16	17	167
Surveys covering age ranges within 0-15 years ^d	51	9	10	70

- a. Several surveys reported follow-up of patients during drug resistance surveys, or patients sampled during routine EPI/MCH clinic visits. These were excluded on the basis that it was difficult to know whether they represented community-based infection prevalence. In addition surveys conducted during randomized controlled trials of interventions were only included for control populations, intervention arms were excluded.
- b. For surveys which were contiguous with others, the most recent, largest sample size and most detailed in its reporting of sampling strategy was included.
- c. Several surveys, particularly those reported by Vector Control or Malaria Divisions of the Ministry of Health, did not provide adequate sampling strategy details (26.4% (n=88) of the overall survey samples). It is unclear whether these prevalence estimates have the same precision as surveys providing detailed descriptions of sampling and survey method.
- d. The surveys included in the analysis were not all consistent with a single age range. The surveys varied in the age groups sampled, for example 0-4, 1-4, 0-3, 0-9, 1-9, 1-12, 0-15, etc. The table represents the summary of three broad age classes using the upper age range as the defining age for sub-groups. Separate analyses have shown that the categorical description of endemicity based on the PR does not alter significantly within one area by different age groupings within the 0-15 year age range.

Figure 2.4. The spatial distribution of parasite survey data 1927-1979

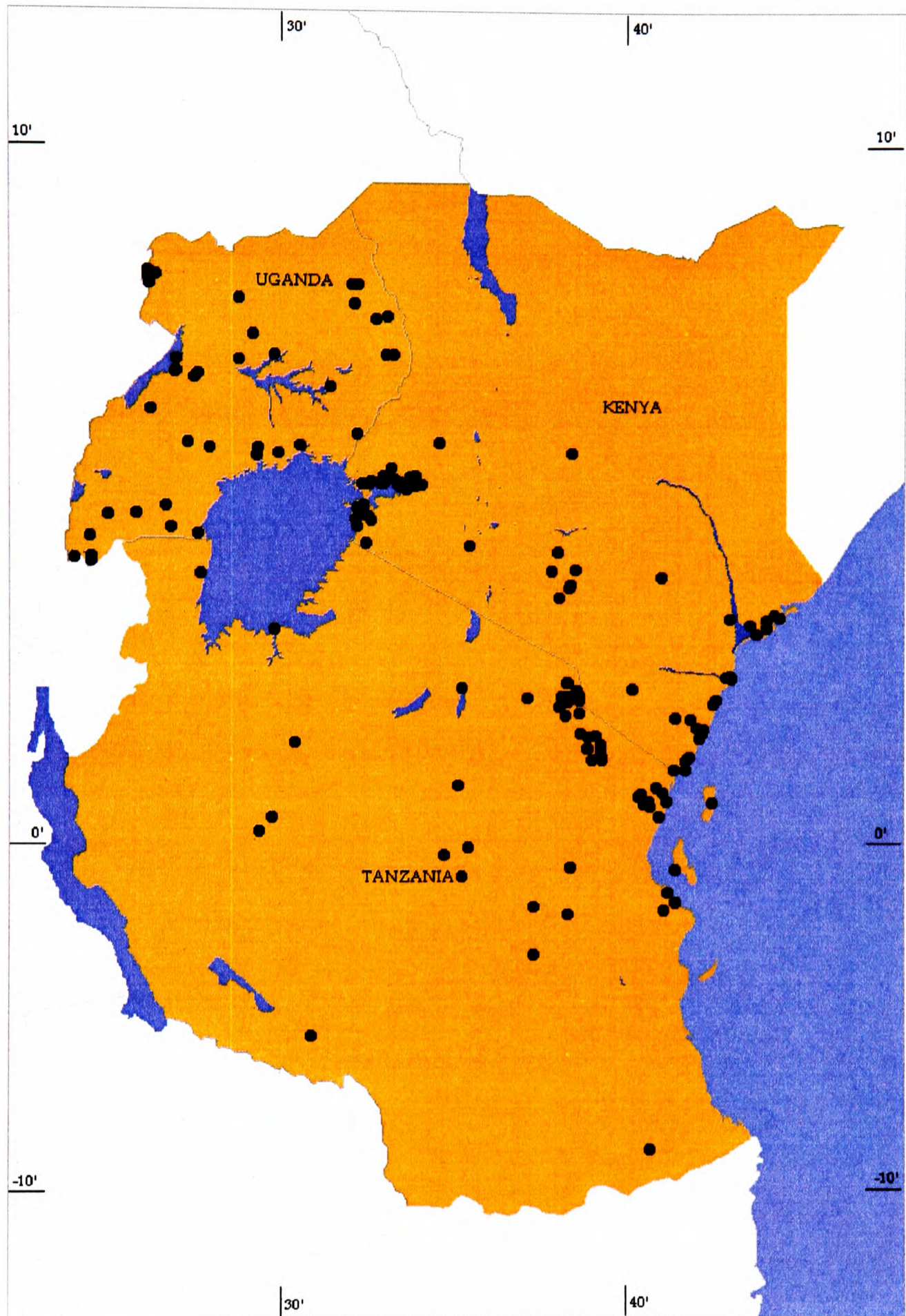
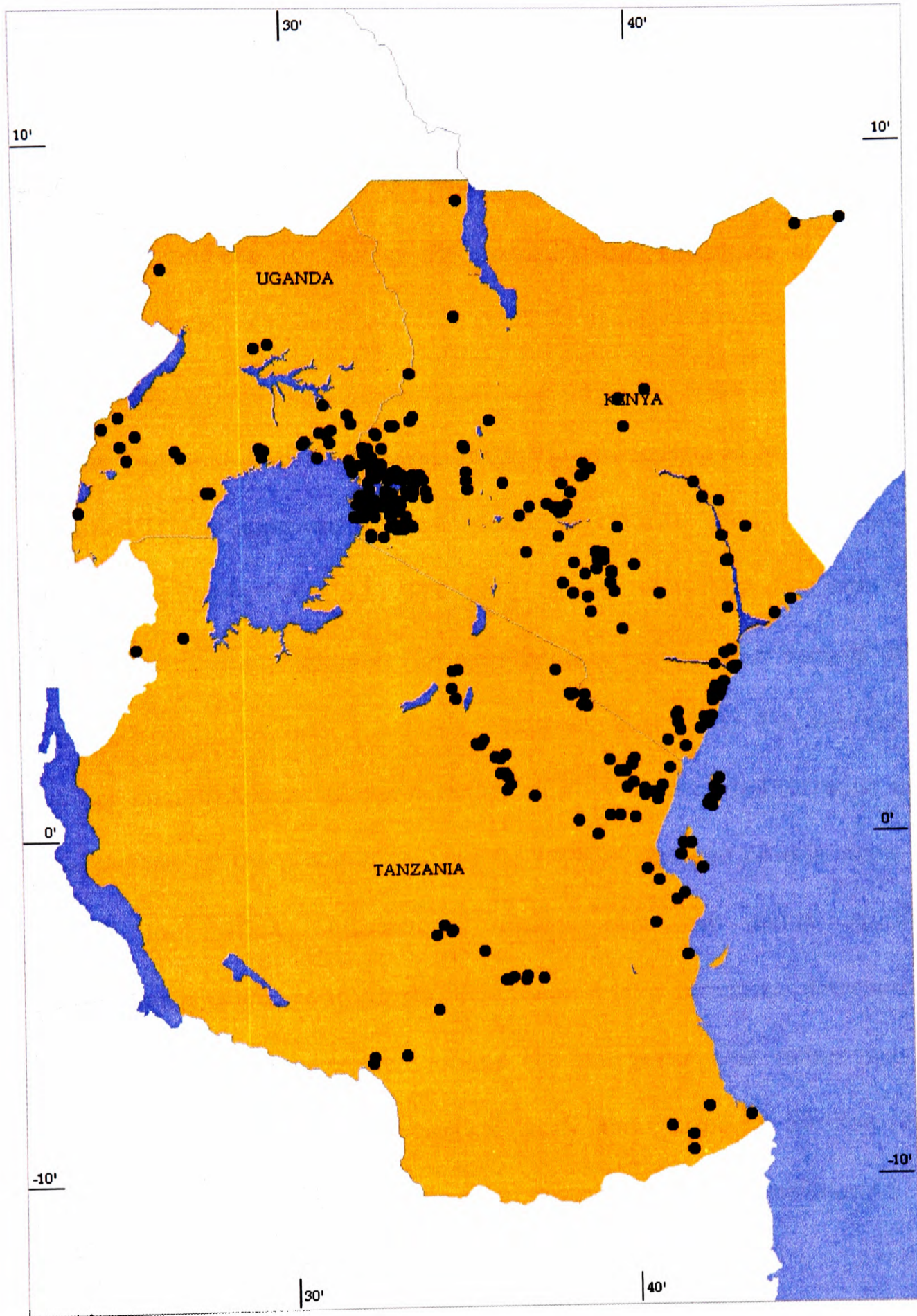


Figure 2.5. The spatial distribution of parasite survey data 1980-2003



2.7. Construction of model attribute data

2.7.1. General principles

The spatial distribution of each parasite survey data point varied and this had to be accommodated for during the matching and extraction of ecological and remotely sensed attribute data. The spatially distinct polygons or points for each PR survey were created using digital administrative maps of East Africa (CBS, 1989; Tanzania Statistics Bureau, 2002; Uganda Bureau of Statistics, 2004), first warped to UN approved national boundaries (SALB, 2004) using ArcView 3.2 (ESRI, Redlands, CA, USA). The spatial extents of attribute data were mapped to single longitudes or latitudes (for point data) or polygons for each of the parasite surveys using ArcView 3.2 (ESRI, Redlands, CA, USA). For surveys covering wider areas (clusters of communities or 3rd-5th level administrative units), the percentage of the polygon covered by various attribute criteria were extracted using the MILA Utilities 3.2 update extension within ArcView 3.2 (<http://www.esri.com/arcscripts>). For vector related proximate determinants a 2.5 km buffer was first created around the polygons. This radius was selected according to previous entomological work summarised by Service (1997) and based on evidence on the average flight range of *Anopheles gambiae* s.l.

2.7.2. Satellite sensor derived meteorological surrogate data

Satellite sensor systems (SSS) can be placed in geostationary or polar-orbits. Geostationary satellites are put into a high orbit (~36,000 km) in the equatorial plane, with a speed equal to that of the Earth's rotation; hence they remain fixed above a particular point on Earth. Polar-orbiting satellites circle the globe repeatedly at much lower altitude orbits (~ 600-900 km) at a high angle to the equator. In the latter situation, successive orbits pass over a different section of the Earth as it rotates (Cracknel & Hayes, 1991). SSS are able to record the intensity of reflected thermal energy. Each sensor measures electromagnetic radiation (EMR) from between specific wavelengths, or channels, of the electromagnetic spectrum (EMS) (Hay, 2000).

As electromagnetic signals pass through the atmosphere and on to the Earth's surface, they may be absorbed, reflected or transmitted. RS techniques use the differences seen in reflectance across the EMS (the *spectral signature*) to make inferences about the physical properties of bodies on the Earth. For example, the spectral signature of a water body is different from that of vegetation. There are a variety of RS systems derived from different satellite sensors and each records a different measure of processes related to the Earth's environment. The sensors differ in their resolution. Resolution refers to the ability of a sensor to distinguish between signals that are close together in space *i.e.* their *spatial resolution*, or between spectrally similar signals *i.e.* their *spectral resolution*. A further difference is the frequency with which sensors record observations, referred to as their *temporal resolution*. Previous reviews of RS applied to the problems of

mosquito and malaria control (Connor, 1999; Dale *et al.*, 1998; Hay *et al.*, 1998a; Roberts & Rodriguez, 1994; Thomson *et al.*, 1997; Washino & Wood, 1994) and to the study and control of a range of arthropod-borne diseases (Hay *et al.*, 1997) have explored the relationships between remotely derived estimates of ecological conditions and parameters of malaria distribution and abundance.

There are several advantages of RS meteorological surrogate data. Routinely collected ground-station meteorological data do not adequately describe variables that are highly spatially variable. RS surrogates of rainfall, however, have shown greater correlation with weather station data than interpolated climate surfaces developed by Hutchinson *et al.* (1995): RS RMSE error=38.0mm, $r^2=0.63$; SI RMSE=93.7mm, $r^2=0.19$ (Hay & Lennon, 1999). Furthermore, improvements in the range and quality of climatic surrogate data are continuing.

A global land coverage of RS data from the NOAA afternoon ascending polar-orbiting satellites were acquired through a world-wide network of 29 high-resolution picture transmission (HRPT) stations and complemented by on-board recording of Local Area Coverage (LAC) where necessary (Eidenshink & Faundeen, 1994). The daily data set is composed of 5-channel, 10-bit, raw AVHRR data, at 1.1-km spatial resolution (at nadir) for each daily orbital pass. These data are available in Goode's interrupted homolosine projection (Steinwand, 1994) for 93 ten-day composites (dekads) from 31 months (April-December 1992, January-September 1993, February-December 1995 and January-May 1996). The data are permanently archived, catalogued, and available from the United States' Geological Survey, Distributed Active Archive Centre

(URL: <http://edcdaac.usgs.gov/1KM/comp10d.html>). All data were re-sampled to a latitude and longitude co-ordinate reference system and coverages for the 3 East African countries subset from the global images using the co-ordinates 19.995 °E to 52.005 °E and latitude 23.755 °N to 13.005 °S. The following sections describe the RS data used in this study.

2.7.2.1. *The Normalised Difference Vegetation Index (NDVI)*

The NDVI is strongly correlated with measures of intensity of vegetation or photosynthetic activity. It is frequently used as a proxy for vegetation density and soil wetness (Justice *et al.*, 1985). The NDVI exploits the differences in light absorption and reflection between the near-infrared (Channel 2) and visible (Channel 1) wavelengths. It is calculated using the formula:

$$\text{NDVI} = \frac{\text{Ch2} - \text{Ch1}}{\text{Ch2} + \text{Ch1}} \quad (\text{Equation 2.1})$$

Values range between -1 and +1. Increasing positive values indicate increasing green vegetation (0-0.2 for desert or bare ground compared with ≥ 0.7 for forest) and negative values indicate non-vegetated features such as water (Hay *et al.*, 1998b). The NDVI has found frequent application in studies of vector-borne disease distribution as an indicator of photosynthetic activity and a surrogate for moisture availability (Hay, 2000; Rogers, 2000). Studies using data for Mali and other West African countries identified the NDVI as a significant explanatory variable of malaria intensity as estimated using the PR (Kleinschmidt *et al.*, 2000).

2.7.2.2. *Land surface temperature (LST)*

LST estimates were calculated from simultaneously collected brightness temperatures (K) from AVHRR channels 4 and 5 and calculated using an equation derived by Price (1984):

$$LST = T_{Ch4} + 3.33(T_{Ch4} - T_{Ch5}) \quad (\text{Equation 2.2})$$

The 3.33 term, is a value determined empirically for the NOAA-7 satellite. This relationship holds because signal attenuation is much greater in channel 5 than in channel 4, so the difference between the channels can be used to estimate, and hence correct for, the amount of atmospheric water vapour attenuation. This equation has been demonstrated to provide LST estimates accurate to $\pm 2-3$ K (Cooper & Asrar, 1989; Sugita & Brutasaert, 1993). The accuracy of this estimation is similar to that of spatially interpolated meteorological station data and has been calculated at $\pm 3-4$ K across Africa (Hay, 1999; 2000).

2.7.2.3. *Middle infrared reflectance (MIR)*

MIR radiance data from AVHRR channel 3 were also processed. Despite being less well documented than other wavebands, MIR wavelengths appear to suffer less atmospheric attenuation than the visible and near infrared wavelengths (Kerber & Schutt, 1986) making these data potentially very suitable for monitoring vegetation in the tropics. Furthermore, MIR has been shown to be more highly correlated with various biophysical properties of vegetation canopies

than the visible or near-infrared wavelengths (Boyd *et al.*, 1996; Boyd & Curran, 1998). Boyd & Ripple (1997) provide an exhaustive review of the proposed mechanisms governing the interaction of MIR and vegetated land surface. MIR reflectance (AVHRR channel 3) has been applied to vegetation mapping and although it is sensitive to both reflected and emitted radiation (Boyd & Curran, 1998) and has advantages in that it suffers less from atmospheric attenuation than channel 4 and channel 5 (Cracknell, 1997).

2.7.2.4. Cold Cloud Duration (CCD)

The CCD index derived from the High Resolution Radiometer (HRR) onboard the European Meteorological Satellite programme's (EUMETSAT) Meteosat satellite series is used as a surrogate measure of rainfall. CCD image pixels represent the number of hours during which cold cloud top temperatures below a geographically variable threshold (Hay & Lennon, 1999) were experienced during a 10-day compositing period. The CCD has been found to have a root mean square error of ± 38 mm when compared with meteorological station recordings across continental Africa (Hay & Lennon, 1999). The 0.01-degree resolution CCD data used for this study were re-sampled to 1x1km resolution.

2.7.2.5. Ordination of time series data

Data from satellite channels are subjected to atmospheric contamination, notably due to clouds. For high temporal resolution satellites (those that take frequent images), the least cloud-contaminated image taken over each 10-day period

(dekad) is selected. The satellite data, summarised according to dekads, is then stored in an image called a Maximum Value Composite (MVC). Meteorological data summarised monthly in this way is highly correlated and data ordination techniques are applied to minimise this effect. In addition to reducing the dimensionality of the data set data ordination is used to identify new meaningful underlying variables. Two main ordination methods are in use, Principal Component Analysis (PCA) and Temporal Fourier Analysis (TFA). Principal component analysis (PCA) involves a mathematical procedure that summarises correlated variables into a number of uncorrelated variables called *principal components*. The first principal component accounts for as much of the variability in the data as possible, and each succeeding component accounts for as much of the remaining variability as possible. TFA is able to retain the information on seasonality within time series data and, for this reason has, been used for ordination of the satellite data used in this study.

2.7.2.6. Temporal Fourier analysis of RS data

The five year series of monthly NDVI, LST, MIR and CCD images were subjected to TFA as described by Rogers & Williams (1994). TFA decomposes seasonal changes in the satellite-sensor variables into the sum of their sinusoidal components with frequencies of one up to six cycles per year, so that the first term in the Fourier expansion gives the annual cycle, the second the biannual cycle, and so on (Rogers *et al.*, 1996). The sum of the different components then describes the observed annual variation at the site in question. The first two Fourier terms capture most of the variation in the signal, and only these were used

in the present study. The analysis also gives information on both the phase (timing in the year of maximum value) and amplitude (maximum range in annual values) of each component and these were used to construct the fitted signal, from which may be extracted the number of months above or below any threshold value.

2.7.3. Altitude

The Global Land 1-km Base Elevation Project (GLOBE) Digital Elevation Model (DEM) Version 1.0. (GLOBE Task Team, 1999) is a 30-arc-second gridded digital elevation model developed by NOAA's National Geophysical Data Center (NGDC (www.ngdc.noaa.gov/seg/topo/globe.shtml)). The DEM is a compilation of several data sets derived from a variety of sources including the GTOPO30 DEM developed by the EROS data center (1993-1997) and the DTED from the US Defence Mapping Agency. GLOBE is the highest resolution global DEM currently available.

2.7.4. Urbanization

Urbanization is a difficult demographic concept to define and there are inconsistencies between international, national and published definitions (Vlahov & Galea, 2002; Tatem & Hay, 2004). A combination of approaches was used to classify urban surveys. First the investigators urban/rural description was accepted as stated in the survey reports. Second, the name of the survey site was used to validate the descriptions, for example, if a study site was reported as

“Malindi town”. Third, national census bureau administrative boundaries of known urban extents against the spatial position of each survey was used to identify those whose survey report or community name did not provide adequate information to classify them as urban or rural. Here urban settlements were classified if they were located in national census bureau definitions of cities, towns or market centres.

Finally a global database of urban extents developed as part of the Global Rural-Urban Mapping Project, GRUMP (Balk *et al.*, 2004; CIESIN/ CIAT 2004; CIESIN/IPFRI/ CIAT, 2004) was used. The urban mask was developed at a 1 x 1 km spatial resolution using data on night-time lights (Sutton *et al.*, 2001) and Landsat satellite imagery in combination with other geographic data (Digital Chart of the World populated places (Mika, 1997), Tactical Pilotage Charts produced from Australian Defence Imagery and Geo-Spatial Organisation, and national census data). The spatially distinct polygons or points for each PR survey were created using digital administrative maps as described earlier using ArcView 3.2 (ESRI, Redlands, CA, USA). The GRUMP urban extents were mapped to the single longitudes or latitudes or polygons of each the parasite surveys using ArcView 3.2 (ESRI, Redlands, CA, USA). If more than 75% of the polygon was classified as either urban or rural the survey was classified accordingly.

2.7.5. Ecological features from Africover

Africover is a project of the United Nation’s Food and Agriculture Organization involved in landcover mapping using satellite data in Eastern Africa (Africover

2003: www.africover.org). The project was operational between 1995 and 2002 and involved ten countries, namely; Burundi, the Democratic Republic of Congo, Egypt, Eritrea, Kenya, Rwanda, Somalia, Sudan, Tanzania and Uganda. The project's aim was to produce user-friendly landcover models for supporting decision makers in planning and management of sustainable development. The Eastern Africa module was the project's first operational component and has so far mapped 8.5 million km² using over 400 Landsat scenes.

The digital database defines 22 landcover/land use classes including cultivated vegetation, natural vegetation, urban areas, settlements (rural & refugee), water bodies (artificial and natural) *etc.* Land use types relevant to this study were water bodies and urban extents. The Africover data were available as ArcView shape files (ESRI, Redlands, CA, USA). Three water body classes (artificial water body, river and lake) and urban areas were extracted from the main database and stored as separate surfaces. The data were converted into grid files within ArcView and the surfaces modified to represent the proportion (%) of each pixel within each of the land use categories. Original vector data were converted into a 100x100 metre resolution grid file and subsequently re-sampled to a 1x1 km resolution grid resulting in surfaces that represent the proportion of each 1x1 km pixel that is urbanised or within a water body.

2.8. Statistical methods

There are two types of models used to link descriptive/field data with meteorological variables. Statistical or empirical models describe wide areas *e.g.*

species distributions and require extensive studies while predictive models model the interactions between biological processes (Rogers, 2000). The latter are also referred to as process-based models. Statistical methods used in predictive modelling aim to reduce a group of often correlated continuous variables to a single linear composite (*i.e.* n -dimensional profiles are reduced to a set of scales along a single axis). The linear association derived comprises values of explanatory variables that maximally distinguish between classes of a variable to be predicted. The techniques then allow the testing of differences between predictor variable group means or centroids. Two techniques are commonly applied in ecological modelling, generalised linear regression (GLR) and discriminant analysis (DA).

In GLR, the relationships between multiple explanatory variables are summarised by a linear relationship that describes how changes in an explanatory variable change according to the variable to be predicted. Linear regression makes two assumptions. Firstly, for any value of an independent variable, it is assumed that the dependent variable is normally distributed. The second assumption is that the dependent variables have equal co-variance matrices (Kirkwood, 1988). Neither of these assumptions can be made of environmental data, which are often spatially correlated. If co-variance matrices are dissimilar the axis separating the explanatory variable groups is not linear. A bias is introduced to predictions resulting in more observations being assigned to the group with the larger co-variance matrix. DA techniques are therefore preferred and are used in the malaria models presented in Chapters 6 and 7.

2.8.1. Principles of discriminant analysis

Discriminant analysis assumes that ‘presence’ of the factor being investigated is described by one multivariate distribution and “absence” of the fact by another multivariate distribution. DA involves defining a best line of discrimination *i.e.* a hypothetical linear axis in multivariate space that maximises the variability or separation between multiple predictor variables (Green, 2002). The difference between the group centroids can then be estimated and finally predictions based on the discriminant function derived can be made. A training data set is used to provide statistics of what predictions are to be expected. The prediction involves assigning a profile to unknown points to the group with the closer centroid based on training data statistics. One approach to discrimination is based on Mahalanobis distances (D^2 (Mahalanobis, 1948)). D^2 is derived from the Euclidean distance (squared) and can be considered a multivariate residual for an observation measuring how far the observation is from the centre of the distribution of all values. D^2 uses information on the means, variances and covariances of the variables being discriminated. Assuming there are 2 mean-corrected centroids, i and j , in a multivariate space of n dimensions, their Mahalanobis distance is defined according to the equation:

$$D_{ij}^2 = (\bar{\mathbf{x}}_i - \bar{\mathbf{x}}_j)' \mathbf{C}_w^{-1} (\bar{\mathbf{x}}_i - \bar{\mathbf{x}}_j) = \mathbf{d}' \mathbf{C}_w^{-1} \mathbf{d} \quad (\text{Equation 2.1})$$

where \mathbf{C}_w^{-1} is the inverse of the pooled within-group covariance matrix (Green, 2002). The Mahalanobis distance of each point to a group centroid is calculated and the point can then be allocated to the group that it is closest to *i.e.* the group

for which D^2 is the minimum. The proportion of correct allocations is used as an indicator of how well the groups can be separated on the basis of the given variables.

2.8.2. Predictor variable data extraction and manipulation

Seven global Fourier analysis processed layers were produced for each of the satellite-derived predictor variables as follows: mean (amp0), minimum (min), maximum (max), amplitude of annual cycle (amp1), amplitude of bi-annual cycle (amp2), phase of annual cycle (ph1), phase of bi-annual cycle (ph2). The resulting 24 satellite variables were stored as IDRISI images (IDRISI Version 2.0, Clarke University Graduate School of Geography, 1997). Data for Eastern Africa (19.995 - 52.005 decimal degrees of longitude and -13.005 - 23.755 decimal degrees of latitude) were selected from the world images using the WINDOW module in IDRISI. Details of sampling criteria used (scale, shift and offset), data type and file type were stored in the documentation file for each image. A sample documentation file for the LST amp1 image for East Africa follows:

```

file title:      East Africa Fourier LST amp 1
DATA type:      integer
file type:      binary
columns:        3201
rows:           3676
ref. system:    latlong
ref. units:     deg
unit dist:      1
min. X:         19.995
max. X:         52.005
min. Y:         -13.005
max. Y:         23.755
pos'n error:    unknown
resolution:     01
min. value:     0
max. value:     1380
value units:    val=(DN/10)

```

value error:	unknown
flag value:	none
flag def'n.:	none
legend cats:	0

Data extraction was carried out in ArcView Version 3.2 (Redlands, CA, USA). The satellite data were exported from IDRISI (Clarke University Graduate School of Geography, 1997) as ASCII files for importation into ArcView as grid files.

Each *P. falciparum* prevalence estimate was represented by a spatial unit representing the area from which samples were drawn. Villages or single communities were represented by 1 km radius area drawn from a central longitude and latitude (centroid). Secondly, in Kenya, where survey reports stated that samples were drawn from known administrative units *i.e.* divisions (3rd level admin. units), locations (4th level units) or sub-locations (5th level units) the administrative boundary was used to represent the sampling area. Respective digital administrative boundaries were obtained from a digital map of 5th level boundaries as used during Kenya's population census in 1989 (Central Bureau of Statistics, 1989). Thirdly, where study samples were drawn from several villages and presented as a single PR estimate for all the villages sampled, a polygon connecting the villages was drawn to represent the study area. A 2.5km buffer was created around these three types of polygon data representing the flight distance of *Anopheles gambiae* (Service, 1997). The buffered polygons were stored as a single shape file in ArcView (Extraction polygons) representing the spatial extents of each survey.

Using the 'zonal statistics' module in ArcView, RS image values were extracted for the extraction polygon areas which were the training data set. An average value of all the pixels within each extraction polygon was extracted and this value stored in a Microsoft Excel file. The relationship between the environmental data and prevalence was defined for the training data using DA. Ideally the training data should represent the entire range of possible environmental conditions across the area for which predictions are to be made. The programs for data analysis and mapping (predictive image outputs) were developed in QuickBasic (Microsoft Corporation, Redmond, Washington) by David Rogers, (personal communication). Each pixel was allocated to a malaria prevalence category according to the similarity of the covariate data related to the respective pixel compared with the training data set. A prediction was not made for pixels that were markedly different from the training data (*i.e.* no predictions were made for pixels that had values greater than 3 standard deviations from the mean).

2.8.3. Assessing the accuracy of the predicted models

The accuracy of a predictive model is assessed in terms of its ability to predict a test data set. Error can be thought of as the difference between a predicted value and the true value at that same location. There are two main approaches to model evaluation. In the first, a single training or calibration data set is used to develop the model and then cross validation techniques are used to assess model accuracy. The alternative method uses two independent data sets, comprising training and evaluation data sets.

A contingency matrix (Table 2.4) is an effective way of representing overall prediction accuracy as well as the accuracy of each individual class as described by the inclusion errors (commission errors) and by the exclusion errors (errors of omission) present in the classification. The main diagonal of the matrix (elements $n_{11}, n_{22}, \dots, n_{kk}$), totals correctly allocated pixels whilst elements outside the diagonal represent incorrect allocations. Overall accuracy is calculated by the sum of main diagonal values divided by the total number of observations, *i.e.*:

$$\text{Overall accuracy} = \frac{\sum_{i=1}^k n_{ii}}{n} \quad (\text{Equation 2.2})$$

Table 2.4: Diagrammatic representation of a contingency matrix

		<i>j</i> =columns (predicted)			Total (row) n_{i+}
		1	2	<i>k</i>	
<i>i</i> =row (training data)	1	n_{11}	n_{12}	n_{1k}	n_{1+}
	2	n_{21}	n_{22}	n_{2k}	n_{2+}
	<i>k</i>	n_{k1}	n_{k2}	n_{kk}	n_{k+}
	Total (column) n_{+j}	n_{+1}	n_{+2}	n_{+k}	n

The contingency matrix can also be used to assess the accuracy of individual classes by calculating the producer's accuracy (omission error *i.e.* the exclusion of a pixel from a class to which it should belong) and the consumer's accuracy

(commission error *i.e.* inclusion of a pixel in a class to which it does not belong (Story & Congalton, 1986)).

$$\text{Producer's accuracy} = \frac{n_{ii}}{n_{i+}} \quad (\text{Equation 2.3})$$

$$\text{Consumer's accuracy} = \frac{n_{jj}}{n_{+j}} \quad (\text{Equation 2.4})$$

Overall accuracy tends to underestimate accuracy as it does not consider the proportion of accuracy that occurs due to chance (Couto, 2003). The *kappa* (k) coefficient (Cohen, 1960) is frequently used as it considers the chance component of agreement. It is based on the difference between the overall agreement and the chance agreement that is indicated by the contingency matrix marginal totals as follows:

$$K = \frac{n \sum_{i=1}^k n_{ii} - \sum_{i=1}^k (n_{i+} \times n_{+i})}{n^2 - n \sum_{i=1}^k (n_{i+} \times n_{+i})} \quad (\text{Equation 2.5})$$

where n is the number of observations, k the number of rows in the matrix, n_{ii} the numbers of observations in row i and column i and n_{i+} and n_{+i} the marginal totals for row i and column i respectively. Values of *kappa* range from 0 (no agreement) to 1 (full agreement). Landis & Koch (1977) suggest the following interpretation of observer agreement defined by the *kappa* statistic: poor where $k < 0.4$, good where $0.4 < k < 0.75$ and excellent where $kappa \geq 0.75$. An advantage of the *kappa* statistic is that confidence intervals can be attached making it a suitable index for hypothesis testing. *Kappa* is a standard measure of accuracy in prediction analyses and is the measure used in this study.

CHAPTER 3:

***Plasmodium falciparum* prevalence in East Africa: A review of empirical data 1927-2003**

JA Omumbo & RW Snow

***East African Medical Journal* (2004), 81: 649-656**

East African Medical Journal Vol 31 No 12 December 2004

PLASMODIUM FALCIPARUM PARASITE PREVALENCE IN EAST AFRICA: A REVIEW

J. A. Omumbo, BDS, MPH, Public Health Group, KEMRI/Wellcome Trust Collaborative Programme P.O. Box 43640, 00100 GPO, Nairobi, Kenya, and R.W. Snow, BSc, MSc, PhD, Public Health Group, KEMRI/Wellcome Trust Collaborative Programme, Kenya and Centre for Tropical Medicine, John Radcliffe Hospital, University of Oxford, UK

Request for reprints to: Dr. J. A. Omumbo, Public Health Group, KEMRI/Wellcome Trust Collaborative Programme, P.O. Box 43640, 00100 GPO, Nairobi, Kenya

PLASMODIUM FALCIPARUM PARASITE PREVALENCE IN EAST AFRICA: A REVIEW

J. A. OMUMBO and R.W. SNOW

ABSTRACT

Objectives: Empirical data on malaria endemicity are rarely available for public domain use to guide effective malaria control. This paper describes the work carried in East Africa since 1997 as part of a pan-African collaboration to map the risk of malaria, Mapping Malaria Risk in Africa (MARA) aimed at redressing deficiency.

Data extraction: Studies of cross-sectional community estimates of *Plasmodium falciparum* prevalence among children aged 0-15 years were identified from a variety of sources including electronic searches of published material, manual review of pre-electronic peer reviewed journals and searches of libraries and archives in Kenya, Tanzania and Uganda. Each survey source, infection prevalence, date, longitude and latitude and survey characteristics were recorded.

Data synthesis: All data were subjected to a number of selection criteria including minimum sample sizes, samples randomly selected, community-based surveys, age ranges of sampled communities within 0-15 years, and surveys that were spatially unique. Of the 2,003 survey data points identified since 1907 in East Africa, only 503 were eligible for inclusion in the analysis dating from 1927 to 2003. The spatial plots of the data demonstrate the paucity of information on malaria prevalence from a number of densely populated areas and highlight the concentration of empirical data in concert with research centres in the sub-region.

Conclusions: Models are required to define malaria risk in areas of East Africa where no empirical data are available so that limited resources can be better targeted to those in greatest need.

INTRODUCTION

Malaria continues to be a major public health problem in the East African sub-region. The disease overwhelms already stretched clinical services and contributes to at least a quarter of all deaths that occur before the fifth birthday(1). Furthermore there is convincing evidence that mortality directly due to *Plasmodium falciparum* is increasing in East Africa since the late 1980's coincidental with the emergence of resistance to widely used anti-malarial drugs(2).

The international community has been effective in raising the importance of malaria control as a priority health investment for Africa through the Roll Back Malaria (RBM) partnership (3,4). Nevertheless five years after the formation of RBM, countries in East Africa still have very low coverage of interventions likely to reduce the burden of malaria in this sub-region. National surveys in East Africa show that less than 5% of children less than five years old are protected by an insecticide-treated net (ITN)(5). In addition there are large disparities in the spatial

components of access to ITN in all these countries. A recent Demographic and Health Survey in Kenya showed that the highest rates of ITN use were in Nairobi, the province least likely to experience malaria risk(6).

Malaria is a vector-borne disease whose geographical extent is driven largely by climatic factors affecting vector and parasite survival. Consequently the large diversity in climate, ecology and urbanisation, characteristic of East Africa, supports a wide range of infection and disease risks. It has often been argued that without at least a basic knowledge of risk, efforts to control malaria will lack the ability to target limited resources to maximise coverage appropriately among those most at risk(7).

There are several approaches to measuring malaria risk, but the most widely used in Africa today was developed following the conference on malaria eradication for Africa held in Kampala in 1950(8). Metselaar and van Theil proposed the use of community-derived estimates of the *P. falciparum* parasite prevalence among children aged 2-10 years(9). They categorised

malaria risk according to the prevalence of parasitaemia [the parasite rate (PR)] among two to ten year olds as follows: Hypoendemic if the PR in the two to ten year age group is <10%; mesoendemic if the PR is 11-50%; hyperendemic if the PR is constantly >50%; and holoendemic if the PR among infants (<1 year old) is constantly <75%. These criteria have been variously applied in different studies and modified according to the data available.

The simplicity of the measure has meant that it rapidly became a common descriptive of malaria risk across Africa. During the 1960's, David Clyde published a series of historical data for Tanzania describing patterns of endemicity by region according to the parasite ratio(10). In Uganda, during the period 1959 to 1960, the WHO/Uganda Malaria Eradication Pilot Project (MEPP) mounted sample surveys of parasite prevalence across the country in preparation for the implementation of the Malaria Eradication Programme(11). Since these early descriptions, there have hardly been any attempts to compile malaria parasite prevalence survey data from East Africa in order to map malaria endemicity. This paper presents the results of a five-year data search and compilation of *P. falciparum* infection prevalence data.

MATERIALS AND METHODS

The feasibility of building a national database of published and unpublished parasitological data was established in Kenya during 1998(12). This led to the formation of a research network called the Mapping Malaria Risk in Africa (MARA) project(13). A data collection proforma was designed to capture a broad range of indices related to malaria risk. The proforma included a unique identifier for each data source and subsequent suffixes for other data from the same source. Parasite prevalence data were regarded as separate if the cross-sectional surveys were conducted at different times within the same report and/or different geographical areas. Care was taken to record as many details as possible on the study location and where maps were provided in the survey report, these were photocopied and attached to the data collection proforma for later determination of the longitude and latitude co-ordinates of the survey site.

Data searches: The data search strategy relied on multiple approaches to ensure a comprehensive coverage of all possible information. Electronic database searches were performed using Medline® (SilverPlatter International, Boston, MA, USA 2000), Popline® (Johns Hopkins School of Hygiene & Public Health, Baltimore, MD, USA) and EmBase® (Elsevier Science, Little Rock, Arkansas, USA 1999-2000). The following keywords were used in the search; malaria, Africa, East Africa, Kenya, Tanzania, Uganda, parasite and malaria and Africa, *Plasmodium falciparum*, parasite rate, parasite prevalence and malaria transmission. For each publication, bibliographies were cross-referenced to identify additional sources of information and other studies. Where additional details could not be identified through the literature searches authors were contacted to provide more information on geographical location, timing and age specific characteristics of their parasitological data.

Pre-electronic journals held at the Wellcome Library (located at the National Public Health Laboratories in the grounds of the Kenyatta National Hospital), archived collections at the London School of Hygiene and Tropical Medicine and at the University of Oxford were reviewed volume by volume systematically from the earliest volume available. In addition, national or regional, non-electronically referenced peer-reviewed journals were manually reviewed for additional data. Annual reports and unpublished reports from mission hospitals, drug companies and non-governmental organisations involved in health care delivery or research in each country were opportunistically accessed using local information on who had undertaken malaria surveys.

Postgraduate theses in the libraries of the University of Nairobi Medical and Chiròmo Campuses and the Community Health Department (Kenya); Makerere University's Department of Child Health, Division of Public Health (Uganda); and Muhimbili University College of Health Sciences (Tanzania) were searched. Visits were also made to national institutes for medical research (Kenya Medical Research Institute, Nairobi; National Institute for Medical Research, Dar es Salaam) to identify published proceedings of national conferences, annual reports and institute journals. These were searched manually to locate additional parasite survey reports. Of particular note were the annual reports of the pre-1980 WHO-established centres of the East African Institute for Medical Research and the East African Community's East African Institute of Malaria and Vector-Borne Diseases whose headquarters were located at Amani, Tanzania.

Results of routine parasite prevalence surveys undertaken by Ministries of Health in Kenya and Uganda were of special interest. In Kenya, the Division of Vector Borne Diseases (DVBD) has been involved in routine vector-borne disease surveillance activities, which include periodic parasite prevalence surveys of communities and schools, since its establishment in 1951. The reports of these surveys were available at the DVBD headquarters in Nairobi. In addition, a visit to six other provincial headquarters was made to identify additional material from 45 field stations operated by DVBD to locate information that may have not been forwarded to the headquarters in Nairobi. At the completion of the pre-eradication parasite prevalence surveys the office of the MEPP located in Jinja, Uganda, was closed and occupied by a paramedical training centre. While all project documents were removed (and largely destroyed), a few items of furniture including a small filing cabinet remained on site and unopened for many years. Fortunately, the raw data from the 1959-1960 eradication pilot project that focused on three districts in Uganda was found in this cabinet.

Spatial positioning of survey data: All parasite survey data were "geo-referenced" (i.e. their latitude and longitude coordinates determined). A limited number of reports provided geographical coordinates but for most survey reports only the survey site name was available. To identify the precise location of these sites, a variety of techniques and complimentary sources of information were applied. First, maps provided in the reports were used to begin the more thorough positioning of each survey. The next important source was topographic maps of varying resolution and detail; in particular, the 1:50000 scale maps produced by the Directorate of Overseas Surveys in 1971 for East Africa. Digital gazetteers(14,15) listing place names and their spatial co-ordinates were used for sites that were not recorded on

topographical maps. Finally, through correspondence with various research groups, particularly the International Livestock Research Institute (ILRI) and the United Nations Environmental Project (UNEP) Global Resources Information Database offices based in Kenya, various digital administrative unit boundary maps were obtained. These allowed an approximation of spatial positions for those that could not be spatially defined though the original source report, topographical maps or electronic gazetteers.

Data entry: All the data were entered using a customised data-entry programme developed in Microsoft Access 97 (Microsoft Corporation, 1989-1996; Seattle, Washington USA). Data were entered twice by different data entry clerks and both files were verified against each other to detect data entry errors which were corrected subsequently using the original proformas. A programme was run to re-check consistency errors, for example inconsistencies in two date fields or district names for a given country. The main Access file served as the primary source for future data extractions. Data were later managed and analysed in Excel (Microsoft Excel version 9.0).

Data selection: Rigorous selection criteria were applied to the primary data. First data were combined from individual survey reports if repeat surveys were undertaken by the same investigators at the same location within a 24 month period or surveys undertaken by different investigators at different times in the exactly same location within a 24 month period. Surveys were spatially mapped using a Geographic Information System (GIS) software, MapInfo Professional Version 7.0 (MapInfo Corporation 1985-2000; Troy, New York) and this was used to select a single estimate for each spatial unit. Where two or more surveys were identified in the same location undertaken > 24 months apart or with different age classes, the most congruent age class survey was selected followed by a selection for surveys reporting the most detailed methodology, the largest sample and then lastly followed a criterion based on the most recent survey within two time-periods (<1980 and >=1980).

The surveys identified represented a variety of sampling approaches, including household-level random studies to non-random or purposively sampled households that may have included volunteers, surveys of healthy attendees at maternal and child health and expanded programme for immunisation clinics (MCH/EPI), surveys of school children and anti-malarial drug-resistance screening surveys. Surveys that might have been subject to sampling error introduced through selective recruitment were excluded and only studies that recruited children at the community level were included in the analysis. Those included comprised total population, randomly sampled or longitudinal community surveys.

Age ranges used during the parasitological surveys also varied between surveys. Where possible, survey data have been pooled to provide a single estimate within the age range of 0-15 years but surveys which included only children aged less than two years were excluded. Surveys were excluded if they covered age ranges above 15 years. Sample size is likely to affect the accuracy of the prevalence estimate and to control for this; only surveys where at least 50 subjects were sampled were included, based on proposed sampling strategies for parasitological surveys (16). Studies for which sample size details were unavailable were excluded.

The spatial extent of the surveys was also an important selection criterion. There were several studies that covered very wide spatial areas, for example >4th level administrative

units. These posed a particular problem for the present analysis as the spatial extent of sampling within these polygons was unclear and thus it was also unclear how representative the parasitological data were of the administrative unit polygon. All reports covering areas larger than 4th level administrative units were therefore excluded for each country. Similarly, studies that could not be accurately positioned were excluded.

RESULTS

The data search identified 2003 *P. falciparum* prevalence surveys undertaken in East Africa since 1907. Eight hundred and fourteen were undertaken among patients attending maternal and child health or expanded programme for immunisation clinics, MCH/EPI, schools or participants of drug trials or intervention studies. These were excluded from the database. Thirty six additional surveys were excluded as they could not be spatially positioned due to a lack of sufficient detail on where they were undertaken or a lack of congruence with digital place names. A further 205 surveys were excluded because they were undertaken among samples of less than 50 subjects. Age ranges varied between surveys and surveys undertaken only among children aged less than two years (n=9) and surveys whose sample age range extended beyond 15 years (n=122) were also excluded. Finally a selection was made for single surveys where more than one survey had been undertaken in the same community >24 months apart, resulting in the exclusion of 161 "spatial duplicates".

The reduced data set contained 503 prevalence surveys undertaken in Kenya (294), Tanzania (142) and Uganda (67) between 1927 and 2003, representing the examination of 202,628 children. Forty three percent of the surveys were conducted before 1980. The frequency of parasite surveys undertaken within ten year time periods is shown in Table 1. We have chosen to divide this time period into two: 1927-1979 and 1980-2003 to provide a historical versus contemporary comparison of survey data in the sub-region being cognisant of major changes that have occurred in the emphasis of malaria control (vertical approaches to vector control pre-1980, to integrated disease control within general health services post-1980) and the emergence of drug resistance. The sources of the survey data, whether undertaken by Ministries of Health, sample sizes and age ranges of the surveys are summarised in Table 2. Ministries of Health were an important source of empirical data representing over 43% of all survey reports. This was particularly true of Kenya. In Uganda there seems to have been a decline in survey effort since 1979 compared to the earlier time periods, possibly reflecting the reduced planning efforts after the national MEPP and a lower malaria research output compared to Kenya and Tanzania. The age ranges and sample sizes included in the surveys did not vary significantly between the two time periods or between countries.

Table 1

Decade frequency of community-sampled surveys of P. falciparum infection prevalence in East Africa

Year of survey	Kenya	Tanzania	Uganda	Total
1920-1929	6	0	0	6
1930-1939	2	1	0	3
1940-1949	5	0	1	6
1950-1959	11	18	1	30
1960-1969	17	16	31	64
1970-1979	36	21	7	64
1980-1989	54	32	2	88
1990-1999	137	24	25	186
2000-2003	26	30	0	56
Total number of surveys	294	142	67	503

Table 2

Description of selected parasite prevalence surveys in East Africa

	Kenya	Tanzania	Uganda	Total
Total number of surveys identified	992	522	489	2003
Excluded	698	380	422	1500
Characteristics of surveys 1927-1979				
Total number of surveys	77	56	40	173
Surveys reported in peer-reviewed journals	25	23	27	75
Surveys undertaken by Ministries of Health	50	13	0	63
Median interquartilerange of sample sizes	165(129;315)	311(170;617)	93(67;373)	191(112;432)
Surveys with age ranges within 0-5 years	0	9	18	27
Surveys with age ranges within 0-10 years	60	40	20	120
Surveys with age ranges within 0-15 years	17	7	2	26
Median IQR of parasite prevalence	40.6(20.6;56.4)	35(16.4;63.0)	49.3(17.9;71.8)	40.1(18.3;62.1)
Lower and upper prevalence recorded	1.5;86.7	0;87.3	0;87.5	0;87.5
Characteristics of surveys 1980-2003				
Total number of surveys	217	86	27	330
Surveys reported in peer-reviewed journals	11	24	7	42
Surveys undertaken by Ministries of Health	116	30	12	158
Median (interquartile range) of sample sizes	221(119;431)	194(146;443)	88(76;265)	204(120;427)
Surveys with age ranges within 0-5 years	32	61	0	93
Surveys with age ranges within 0-10 years	134	16	17	167
Surveys with age ranges within 0-15 years	51	9	10	70
Median (IQR) of parasite prevalence	33.7(14.3;54.0)	34.8(18.5;67.9)	47.5(28.3;79.5)	35.1(17.0;56.3)
Lower and upper prevalence recorded	0;94.5	0;96.1	6.7;90.6	0;96.1

Table 3

First-level administrative boundary median (IQR) (n) *P. falciparum* parasite prevalence survey data for Kenya, Tanzania and Uganda: 1927-1979 and 1981-2003

	1927-1979	1980-2003	Population*
Kenya			
Nairobi		16.7 (11)	2,143,254
Central		2.8(2.4,8.7) (17)	3,724,159
Coast	28.1 (8.3,37.0) (33)	34.0(26.3,54.1) (137)	2,487,264
Western	46.0 (1)	59.4(34.9,77.8) (14)	3,358,776
Nyanza	57.1 (47.9,70.2) (31)	49.2(32.1,59.7) (97)	4,392,196
North Eastern		7.3(5.4,12.1) (16)	962,143
Eastern	20.3 (9.0, 26.5) (7)	20.6(7.1,29.7) (35)	4,631,779
Rift Valley	35.4 (32.0,56.4) (5)	4.8(0.9,17.0) (10)	6,987,036
Tanzania			
Dodoma	10.1 (5.5,15.2)(5)	17.3(11.4,19.2)(7)	1,707,275
Arusha	26.2 (13.5,48.3) (3)	22.4(8.8,39.4)(10)	1,221,890
Kilimanjaro	38.6 (23.3,56.5)(20)	11.3(7.5,14.7)(7)	2,019,963
Tanga	53.8 (37.6,74.3) (11)	54.5(33.7,70.7) (22)	1,742,413
Morogoro	75.4 (73.8,80.8) (5)	85.6(43.6,88.4) (6)	1,783,664
Pwani	87.2 (1)	76.6(60.7,84.1) (4)	848,316
Dar es Salaam	33.9,34.0 (2)	19.7,48.5 (2)	2,547,217
Lindi		73.9 (1)	848,562
Mtwara	8.3 (1)	63.4 (50.9,75.4) (4)	1,079,816
Ruvuma			1,222,242
Iringa		44.1(12.1,75.6) (9)	1,737,791
Mbeya	4.4 (1)		2,235,271
Singida			1,109,005
Tabora	25.0 (24.0,55.0) (3)		1,432,673
Rukwa			1,218,977
Kigoma			1,240,939
Shinyanga			2,615,565
Kagera	20.0 (1)	24.2,36.4 (2)	1,957,921
Mwanza	52.2 (1)		2,665,956
Mara			1,432,476
Manyara			999,729
North Unguja		25.7,45.6 (2)	137,976
South Unguja		18.1 (1)	110,733
West Unguja/Urban	2.0 (1)	46.7 (1)	363,253
North Pemba	0 (1)	34.1(26.3,42.4) (4)	203,137
South Pemba		39.2 (34.7,45.5) (4)	188,695
Uganda			
Central	32.3 (26.9,56.1) (8)	32.0 (23.4,39.8) (8)	6,683,887
Eastern	16.7 (16.5,35.5) (3)	47.6 (39.6,77.2) (10)	6,301,677
Northern	56.6 (44.6,78.2)(17)	80.5 (79.9,82.8) (3)	5,345,964
Western	33.5 (1.3,72.2) (12)	54.9 (31.3,80.4) (6)	6,417,449

*Population data refer to the latest available national census. Kenya (1999: <http://www.cbs.go.ke>); Tanzania (2002: <http://www.tanzania.go.tz/census/census>); Uganda (2002: <http://www.ubos.org>).

National medians and inter quartile ranges of infection prevalence are shown in Table 2 for the two time periods. These ranges of infection prevalence simply serve to demonstrate the diversity of malaria endemicity covered by the surveys included in the database rather than as an indication of national infection risks. The latter is best considered at sub-national resolution. Table 2, however, suggests that surveys undertaken in Tanzania were conducted in areas of higher prevalence compared to surveyed areas

in Kenya and Uganda. To examine the spatial and temporal differences in *P. falciparum* infection prevalence further, we have analysed survey data according to first level administrative units (Kenyan provinces, Tanzanian regions and Ugandan regions) using recent UN approved international country boundaries(17). The longitude and latitude co-ordinates of the centroid of each survey site was used to attribute each survey to one of eight administrative boundaries in Kenya, 26 in Tanzania and four in Uganda (Figures

1a and 1b). Comparison of Figure 1 a (1927-1979) with Figure 1b (1980-2003) further highlights the decline over time in information for Tanzania and Uganda and also the changes in the spatial focus of survey work within countries that have occurred over-time. The studies undertaken pre-1980 are more widespread and so are more representative of the wide range of malaria ecologies

experienced in the countries studied. Later studies have focused on specific ecological sites, notably low malaria risk highland areas in the Usambara Mountains of Tanzania and Kigezi in Uganda and some arid areas of northern Kenya as well as areas of more intense transmission along the Kenyan coast and the Lake Victoria Basin, likely to be a reflection of where research groups operate.

Figure 1a

The spatial distribution of P. falciparum parasite surveys undertaken between 1927 and 1979(•) against first-level UN approved administrative boundaries (17) for Kenya, Tanzania and Uganda

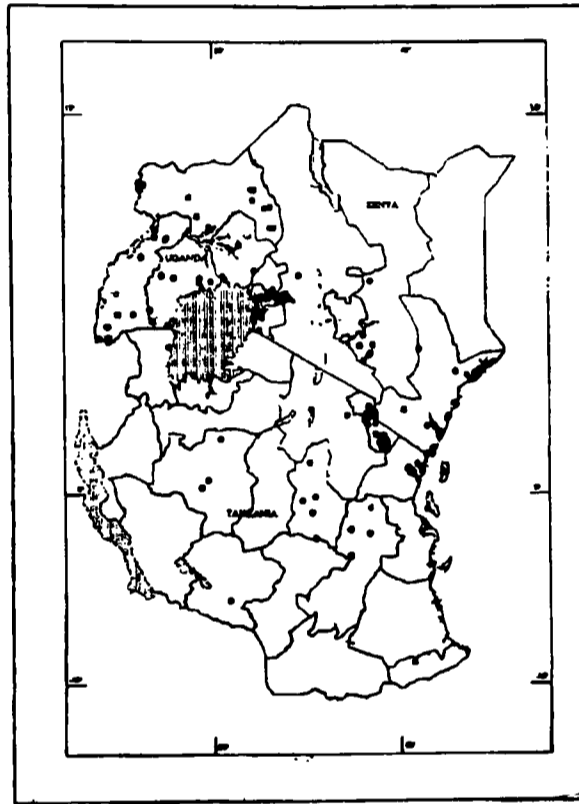
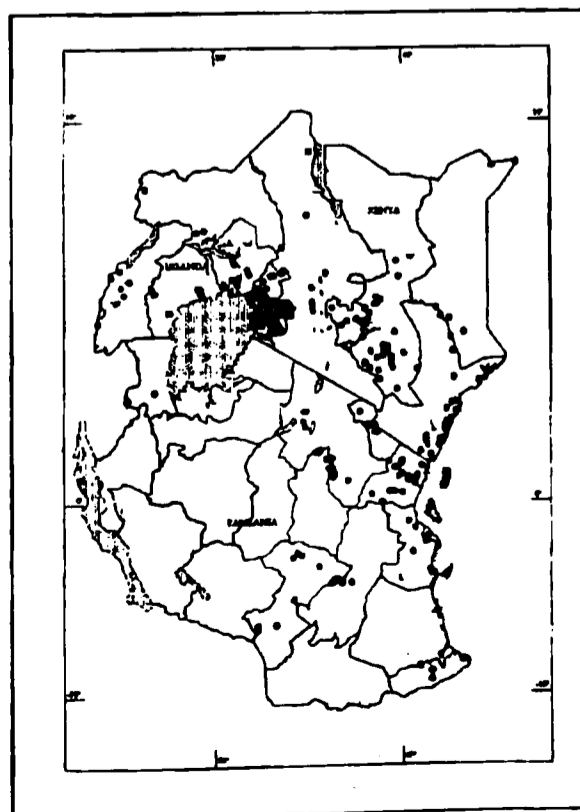


Figure 1b

The spatial distribution of P. falciparum parasite surveys undertaken between 1980 and 2003(•) against first-level UN approved administrative boundaries (17) for Kenya, Tanzania and Uganda



The number of surveys, median of the infection prevalence estimates and ranges of parasite prevalence by administrative unit between 1927-1979 and 1980-2003 are shown in Table 3. The first observation, as shown in Figures 1 a and 1 b, is that there is a paucity of information in several populated areas of East Africa, notably the south and west of Tanzania, the northern regions of Uganda and northern and central parts of Kenya. Secondly, there is a wide range of parasite prevalence estimates within a given administrative boundary, as shown by the interquartile ranges in Table 3 for areas such as Arusha and Tanga in Tanzania. This would be expected given the coarse spatial resolution of first-level administrative boundaries which, in some cases, cover up to 183,000 km² (Rift Valley Province in Kenya) thus encompass a diverse range of altitude and ecology. Nevertheless areas reporting low parasite prevalence tend to be those located either at high altitude (highlands of Tanzania) or arid areas (North Eastern and areas north of Eastern Province in Kenya) and thus conform to our basic understanding of malaria transmission in the sub-region.

DISCUSSION

The Roll Back Malaria (RBM) initiative is a global partnership of donor agencies, Ministries of Health, the private sector and non-Governmental Organisations in malaria endemic countries(3). The overall objective of RBM is to halve the global burden of malaria by the year 2010 through sustained and multi-pronged intervention strategies(3). Equal optimism was expressed during the early part of the last century with the World Health Organisation's 8th World Health Assembly adoption of a Global Malaria Eradication Campaign in 1955 and the expansion of effective drugs for disease management and residual house-spraying(18). Today, strategies for malaria control include; prompt access to effective treatment, access to Insecticide Treated Nets (ITNs), prevention and control of malaria during pregnancy through intermittent presumptive treatment and effective response to epidemics and emergencies(3). Today's emphasis is on the development of interventions that are adopted to local needs and supportive of the health sector development strategy adopted by governments in several sub-Saharan African countries. A key failure of the malaria eradication era in the tropics and sub-tropics was the lack of recognition of the geographical differences in the epidemiology of malaria and in vector behaviour(19). The WHO strategy in the 1950's assumed that a uniform eradication approach would be effective over a wide range of malaria ecologies and the importance of these differences was only recognised in hindsight.

Current control approaches require a renewed examination of the geographical determinants of malaria. Maps of malaria risk are pivotal to the achievement of a spatial dimension to planning malaria control

activities. Several historical maps do exist of malaria risk in East Africa(20-22), however these were developed from "expert opinion" during the late 1950's and early 1960's. These maps were not based upon empirical data and it seems reasonable to assume that risks have changed over the last fifty years. The present review demonstrates two things: firstly, there are data available but often these are not consolidated into centralised databases for public domain access. The collation of the data presented in this study has taken five years and it is the intention of the authors to host the empirical data on an appropriate website. The second important observation is that the contemporary survey data that are available are concentrated in areas where malaria research groups work. This differs from earlier work, before 1980 that aimed to provide national survey data in countries such as Tanzania and Uganda. The limited spatial coverage of survey data prevents easy extrapolation to distant areas that might have very different ecological characteristics lending themselves to very different transmission patterns. There has been a renaissance in malaria mapping and recent advances in disease mapping using Geographic Information Systems (GIS) coupled with an increased access to improved geo-referenced databases on correlates of malaria transmission have rekindled an interest in such epidemiological investigations(23). In particular, the relationship between climate and vector-borne disease transmission has been exploited in mapping the distribution of these diseases(24-25). Currently available climate-based maps are limited in that they consider only a few of the multiple factors that determine malaria's transmission omitting key covariates such as urbanisation, land use and the effects of malaria control interventions (24,26-28).

New spatial models of malaria risk require a source of empirical spatially referenced data for their construction and validation. The data presented in this paper represent one of the largest survey collections for the sub-region and will be used to develop improved, high-resolution models of malaria risk. Such maps should form the basis of decision-making for control interventions; definition of target populations for these interventions and estimation of resource needs within the sub-region. A better geographical awareness of risk will hopefully result in improved use of limited resources available towards achieving RBM goals in East Africa.

ACKNOWLEDGEMENTS

This study received financial support from The Wellcome Trust, UK; International Development Research Centre, Canada; the South African Medical Research Council, the MAR/ARMA collaboration and the Kenya Medical Research Institute. The Wellcome Trust supports JAO and RWS as part of their Prize Studentship (#060063) and Senior Research Fellow (#058992) programmes respectively. The contribution of malaria control

personnel and scientists in East Africa who shared unpublished data. In particular, the Division of Malaria Control (Kenya Ministry of Health), Don de Savigny and Rose Lusinde (Tanzania Essential Health Interventions Project - TEHIP), Hugh Reyburn and Chris Drakeley (Gates Malaria Program, London School of Hygiene & Tropical Medicine), Caroline Maxwell, Chris Curtis and Ilona Carneiro (London School of Hygiene and Tropical Medicine), Ambrose Talisuna, Peter Langi and Albert Kilian (Ministry of Health, Uganda), Vicki Marsh and Tabitha Mwangi (Kenya Medical Research Institute), Dejan Zurovac (MSF-France) and Sporton Kwanya Agak, (KEDHAR). We also acknowledge the help provided by Simon Hay and David Rogers at the Department of Zoology, University of Oxford. This paper is published with the permission of the Director, KEMRI.

REFERENCES

1. Snow, R.W., Korenromp, E.L. and Gouws, E. Paediatric mortality in Africa: *Plasmodium falciparum* malaria as a cause or a risk? *Amer. J. Trop. Med. Hyg.* 2004; 71:16-24
2. Korenromp, E.L., Williams, B.G., Gouws, E., Dye, C. and Snow, R.W. Measurement of trends in childhood malaria mortality in Africa: an assessment of progress toward targets based on verbal autopsy. *Lancet Inf. Dis.* 2003; 3:349-358.
3. World Health Organisation. *Rolling back malaria. The world health report 1999: making a difference.* Geneva: World Health Organization, 1999; 49-63.
4. World Health Organisation/United Nations Children's Fund: *The African malaria report 2003.* Geneva/New York: World Health Organization / United Nations Children's Fund, 2003: 120.
5. Monasch, R., Reinisch, A. and Steketee, R. Child coverage with mosquito nets and malaria treatment from population-based surveys in African countries - a baseline for monitoring progress in Roll Back Malaria. *Amer. Trop. Med. Hyg. J.* 2004; 70:(Suppl) in press.
6. Central Bureau of Statistics. *Kenya Demographic and Health Survey 2003: Preliminary Report.* Central Bureau of Statistics, Nairobi, Kenya, 2003.
7. Snow, R.W., Marsh, K. and Le Sueur, D. The need for maps of transmission intensity to guide malaria control in Africa. *Parasitol. Today.* 1996; 12:455-456.
8. World Health Organisation. *Report on the conference on the eradication of malaria in Equatorial Africa, Kampala.* World Health Organization, Geneva Technical Report Series number. 1951; 38.
9. Metselaar, D. and Van Thiel, P.M. Classification of malaria. *Trop. Geog. Med.* 1959; 11:157-161.
10. Clyde, D.F. *Malaria in Tanzania.* Oxford University Press, London, 1967.
11. Ministry of Health. Uganda/World Health Organization (1959-1962). Annual reports of the Uganda Malaria Eradication Pilot Project.
12. Omumbo, J., Ouma, J., Rapouda, B., Craig, M., le Sueur, D. and Snow, R.W. Mapping malaria transmission intensity using geographic information systems (GIS): an example from Kenya. *Ann. Trop. Med. Parasitol.* 1998; 92:7-21.
13. MARA/ARMA. *Towards an atlas of malaria risk in Africa. First technical report of the MARA/ARMA collaboration.* MARA/ARMA, Durban, 1998.
14. GDE Systems. *Populated places: GeoName Digital Gazetteer v.1 (CD-ROM)* <http://www.GDEsystems.com>, Geographic Information Department (MZ 1211M), GDE Systems Inc., P.O. Box 509009, San Diego, California, 1995.
15. World Resources Institute: *Topographic data: Africa Data Sampler (CD ROM)*, World Resources Institute, Washington, USA, 1995.
16. Snow, R.W. and Gilles, H.M. The epidemiology of malaria. In: *Essential Malariology*, D.A. Warrell and H.M. Gilles. Editors. 2002, Arnold: New York. 85-106.
17. SALB. Second Administrative Level Boundaries (SALB) Project overviews: Concepts, progress and future. URL: http://www3.who.int/whosis/gis/salb/salb_po.htm accessed 29th April 2004.
18. Lepes, T. Present status of the global malaria eradication programme and prospects for the future. *J. Trop. Med. Hyg.* 1974; 77:47-53.
19. Coluzzi, M. Malaria and the afrotropical ecosystems: Impact of man-made environmental changes. *Parasit.* 1994; 36:223-227.
20. Butler R.J. Atlas of Kenya: A comprehensive series of new and authenticated maps prepared from the national survey and other government sources with gazetteer and notes on pronunciation and spelling. Nairobi, Kenya, The Survey of Kenya, 1959.
21. Government of Tanganyika. Atlas of Tanganyika, East Africa. Dar es Salaam, Government Press, Tanganyika, 1956.
22. Government of Uganda. Atlas of Uganda. Uganda, Department of Lands and Surveys, 1962.
23. Hay, S.I. An overview of remote sensing and geodesy for epidemiology and public health application. *Adv. Parasitol.* 2000; 47:1-35.
24. Craig, M.H., Snow, R.W. and Le Sueur, D. A climate-based distribution model of malaria transmission in sub-Saharan Africa. *Parasitol. Today.* 1999; 15:105-111.
25. Hay, S.I., Omumbo, J.A., Craig, M.H. and Snow, R.W. Earth observation, geographic information systems and *Plasmodium falciparum* malaria in sub-Saharan Africa. In: *Adv. Parasitol.* 2000; 47:173-215.
26. Snow, R.W., Gouws, E., Omumbo, J. et al. Models to predict the intensity of *Plasmodium falciparum* transmission: applications to the burden of disease in Kenya. *Trans. Roy. Soc. Trop. Med. Hyg.* 1998; 92:601-606.
27. Kleinschmidt, I., Omumbo, J.A., Briet, O. et al. An empirical malaria distribution map for West Africa. *Trop. Med. Int. Health.* 2001; 6:779-786.
28. Balls, M.J., Bodker, R., Thomas, C.J. et al. Effect of topography on the risk of malaria infection in the Usambara mountains, Tanzania. *Trans. Roy. Soc., Trop. Med. Hyg.* 2004; 98:400-408.

CHAPTER 4:

The relationship between the *Plasmodium falciparum* parasite ratio in childhood and climate estimates of malaria transmission in Kenya

JA Omumbo, SI Hay, CA Guerra, RW Snow

***Malaria Journal* (2004), 3: Article N° 17**

Research

Open Access

The relationship between the *Plasmodium falciparum* parasite ratio in childhood and climate estimates of malaria transmission in Kenya

Judith A Omumbo*^{1,2}, Simon I Hay^{2,1}, Carlos A Guerra² and Robert W Snow^{1,3}

Address: ¹Public Health Group, KEMRI/Wellcome Trust Research Laboratories P.O. Box 43640, 00100 GPO, Nairobi, Kenya, ²ALA Research Group, Linberg Building, Department of Zoology, University of Oxford, South Parks Road, Oxford, UK and ³Centre for Tropical Medicine, John Radcliffe Hospital, University of Oxford, Oxford, UK

Email: Judith A Omumbo* - jomumbo@wtnairobi.mimcom.net; Simon I Hay - simon.hay@zoo.ox.ac.uk; Carlos A Guerra - carlos.guerraloaiza@zoo.ox.ac.uk; Robert W Snow - rsnow@wtnairobi.mimcom.net

* Corresponding author

Published: 17 June 2004

Received: 17 May 2004

Malaria Journal 2004, 3:17 doi:10.1186/1475-2875-3-17

Accepted: 17 June 2004

This article is available from: <http://www.malariajournal.com/content/3/1/17>

© 2004 Omumbo et al; licensee BioMed Central Ltd. This is an Open Access article: verbatim copying and redistribution of this article are permitted in all media for any purpose, provided this notice is preserved along with the article's original URL.

Abstract

Background: *Plasmodium falciparum* morbid and fatal risks are considerably higher in areas supporting parasite prevalence $\geq 25\%$, when compared with low transmission areas supporting parasite prevalence below 25%. Recent descriptions of the health impacts of malaria in Africa are based upon categorical descriptions of a climate-driven fuzzy model of suitability (FCS) for stable transmission developed by the Mapping Malaria Risk in Africa collaboration (MARA).

Methods: An electronic and national search was undertaken to identify community-based parasite prevalence surveys in Kenya. Data from these surveys were matched using ArcView 3.2 to extract spatially congruent estimates of the FCS values generated by the MARA model. Levels of agreement between three classes used during recent continental burden estimations of parasite prevalence (0%, $>0 - <25\%$ and $\geq 25\%$) and three classes of FCS (0, $>0 - <0.75$ and ≥ 0.75) were tested using the kappa (k) statistic and examined as continuous variables to define better levels of agreement.

Results: Two hundred and seventeen independent parasite prevalence surveys undertaken since 1980 were identified during the search. Overall agreement between the three classes of parasite prevalence and FCS was weak although significant ($k = 0.367$, $p < 0.0001$). The overall correlation between the FCS and the parasite ratio when considered as continuous variables was also positive (0.364 , $p < 0.001$). The margins of error were in the stable, endemic (parasite ratio $\geq 25\%$) class with 42% of surveys represented by an FCS < 0.75 . Reducing the FCS value criterion to ≥ 0.6 improved the classification of stable, endemic parasite ratio surveys. Zero values of FCS were not adequate discriminators of zero parasite prevalence.

Conclusion: Using the MARA model to categorically distinguish populations at differing intensities of malaria transmission in Kenya may under-represent those who are exposed to stable, endemic transmission and over-represent those at no risk. The MARA approach to defining FCS values of suitability for stable transmission represents our only contemporary continental level map of malaria in Africa but there is a need to redefine Africa's population at risk in accordance with both climatic and non-climatic determinants of *P. falciparum* transmission intensity to provide a more informed approach to estimating the morbid and fatal consequences of infection across the continent.

Background

In recent years there has been a renaissance in mapping malaria distribution at the national [1-4], continental [5,6] and global scales [7]. These maps have adopted a variety of approaches and data sources in their construction. The most widely cited, contemporary continental resolution map of *Plasmodium falciparum* transmission distribution for Africa was developed by the Mapping Malaria Risk in Africa (MARA) collaboration [5]; <http://www.mara.org.za>. It has formed the basis for several reports by the Roll Back Malaria partnership [8,9] and was used during several recent estimates of the pan-African public health malaria burden to identify population's at-risk [10-12]. The MARA model used a fuzzy membership approach, assigning 5 × 5 km areas to a suitability estimate for stable *P. falciparum* transmission based upon simple rainfall and temperature determinants of the parasite's sporogonic development and mosquito survival. The model did not attempt to define intensities of transmission; rather it determined the likelihood that stable transmission could occur. Using the MARA model, areas of low stable and high stable endemicity across the continent have been distinguished categorically by assuming that the greater the climatic likelihood of stable transmission, the more likely areas would support higher intensities of malaria transmission [11,12]. In this paper, the validity of these assumptions and the likely margins of error are examined by comparing MARA climate suitability values with empirical *P. falciparum* parasite prevalence survey data in Kenya.

Methods

Plasmodium falciparum prevalence surveys data among children aged 0–15 years in Kenya

A search of published and unpublished literature related to malaria infection prevalence surveys in Kenya was conducted as part of the MARA project [13,14]. In brief, electronic database searches were performed using Medline® (SilverPlatter International, Boston, MA, USA 2000), Popline® (Johns Hopkins School of Hygiene & Public Health, Baltimore, MD, USA, 2000) and EmBase® (Elsevier Science Little Rock, Arkansas, USA 1999–2000). The following keywords were used in the search: Kenya, malaria, parasite and malaria and Kenya, *Plasmodium falciparum*, parasite rate, parasite prevalence and malaria transmission. For each publication, bibliographies were cross-referenced to identify additional sources of information from other studies. Where additional details could not be identified through the published sources, authors were contacted to provide more information on geographical location, survey dates and age-specific characteristics of the parasitological data. Postgraduate theses held in the libraries of four departments of the University of Nairobi were also searched (Medical school campuses at the Kenyatta National Hospital and Chiromo, Community

Health Department and the Faculty of Science). Annual reports, journals and conference proceedings of national medical research institutes, and non-governmental organisations were reviewed at respective institute's libraries. Results of routine parasite prevalence surveys undertaken by the Ministry of Health's (MoH) Vector Control Department were manually searched in archives at national headquarters and at seven Provincial offices in Kenya.

Using a variety of sources: 1:50 000 scale topographic maps [15], digital maps of administrative units in Kenya [16] or public domain digital gazetteers [17,18], a longitude and latitude was ascertained for each parasitological survey in decimal degrees. These geo-references were imported into a geographical information system platform ArcView 3.2 (ESRI, Redlands, CA, USA), mapped and overlaid on administrative boundary maps for Kenya [16]. The national administrative boundary maps were first warped within ArcView 3.2 to United Nations approved national boundaries [19] and then used to check for inconsistencies in spatial positioning and to define the spatial coverage of each survey (see below).

All surveys undertaken in Kenya (n = 923) were subjected to a number of selection criteria for inclusion in the analysis. First, to allow for a contemporary assessment of infection risk, surveys were only selected if they were undertaken between 1980 and 2003 (n = 657). The historical data (1927–1979) will be described elsewhere (Omumbo & Snow, in preparation). Second, surveys were excluded if the survey formed part of clinic visits, drug sensitivity testing or included intervention arms of controlled trials. Only total population, randomly sampled or longitudinal community-based surveys were included. Several community-based surveys, particularly those undertaken by the Ministry of Health, provided little detail on the sampling strategy used to select the childhood populations. These surveys were retained but were coded separately in the database to distinguish them from truly random selections should there be differences in estimates of parasite ratio dependent upon sampling methods (n = 74). Third, surveys were excluded if they covered infants only or an age range that extended into adulthood (>=15 years, (n = 12)). Fourthly, a minimum survey sample size of 50 was imposed on the selection to allow for adequate precision in the estimates of infection prevalence [20], or surveys were excluded if there were no details of the denominator or numerator (n = 21). Finally, repeat surveys by the same investigators within a twenty-four month period were combined into a single estimate. Surveys undertaken by several investigators at different times in the same location were reduced to one estimate by selecting the most recent survey, or the one with the largest sample size.

MARA Fuzzy Climate Suitability Index

The MARA model describes climatic conditions that range from unsuitable (0) to completely suitable (1) for stable *P. falciparum* transmission [5]. The Fuzzy Climate Suitability (FCS) index is defined by a series of curves

$$y = \cos^2 \left[\frac{(x - U)}{(S - U)} * \frac{\pi}{2} \right]$$

where x is a climate parameter, U is the value of x when conditions are unsuitable, and S is the value of x when conditions are suitable. When S is greater than U the suitability ($1-y$), increases with x ; when S is less than U the suitability y , decreases as x increases. The model defines a monthly increasing curve ($S = 22$ C, $U = 18$ C) and decreasing curve ($S = 22$ C, $U = 40$ C) for mean diurnal air temperature, a monthly increasing curve ($S = 80$ mm, $U = 0$ mm) for rainfall, and a single increasing curve ($S = 6$ C, $U = 4$ C) for annual minimum temperature.

The FCS values for each parasite ratio data point were extracted from the FCS model developed by Craig *et al.* [5]. To provide a spatially congruent FCS value for each parasite prevalence survey, surveys were classified as representing one of five spatial dimensions: First, for surveys representing a single village, the central longitude and latitude was used to create a 2.5 km buffer surrounding the village using ArcView 3.2 (ESRI, Redlands, CA, USA). The size of the buffer was defined by the average flight range of *Anopheles gambiae* s.l. [21]. Using the MILA Utilities 3.2 update extension within ArcView 3.2 <http://www.esri.com/arcscripts>, the average of all the 5×5 km pixel FCS values within this buffer was extracted to represent the average FCS value for the respective village parasite prevalence survey. The second spatial classification reflected surveys that sampled from several villages but presented the data as a single parasite ratio estimate ($n = 5$). In this case, a polygon was created to connect the villages and a 5 km buffer created around the polygon to represent the wider spatial sampling. Average FCS values within this buffered polygon were summarised. The third, fourth and fifth spatial criteria corresponded to surveys undertaken at the 5th (sub-location), 4th (location) or 3rd

(division) administrative unit levels in Kenya. These spatially distinct polygons were created within ArcView 3.2 using a digital administrative map of Kenya [16] and a 2.5 km buffer was created around each polygon extent before extraction of the average FCS values.

Data entry and statistical methods

All parasite prevalence survey data were entered twice in Microsoft® Access version 7 (Microsoft Corporation 1989–1996; Seattle, Washington, USA). Data entry errors and range inconsistencies were checked and verified against the original material. Data were summarized for matching to other data sources using Microsoft® Excel 2000 version 9.0 and then analysed using SPSS (SPSS, v10.01, 1999, Chicago, Illinois, USA).

Data were then re-classified in accordance with criteria used to define three categorical limits of FCS and parasite ratio risk for burden of disease analysis for the African continent to allow for differences in disease and mortality risks between classes of malaria endemicity [11,12] (Table 1). First, were areas with a FCS value of zero. Second were areas where populations are exposed to marginal risks of malaria transmission and those communities able to support cross-sectional estimates of parasite prevalence below 25% (hypo to meso-endemic malaria transmission). These areas were assumed to be represented by an FCS greater than zero but less than 0.75. Finally, areas of the continent that are described by an FCS of greater than 0.75, which might support parasite prevalence rates of 25% and above (meso- through to holoendemic malaria).

The significance of agreement between the three classes of parasite ratio and FCS was tested using the kappa (k) statistic [22], which is a measure of the agreement between two classifications discounting for the probability that the agreement could be due to chance. Values of kappa range from 0 (no agreement) to 1 (full agreement) and Landis & Koch [23] suggest the following interpretation of agreement defined by the kappa statistic: poor where $k \leq 0.4$, good where $0.4 < k < 0.75$ and excellent where kappa ≥ 0.75 .

Table 1: Agreement between parasite prevalence categories among 217 surveys and spatially congruent categorical values of the Fuzzy Climate Suitability (FCS) using the 0.75 threshold for stable endemic malaria.

	FCS = 0.00	FCS >0 – <0.75	FCS >= 0.75	Totals
Parasite prevalence 0%	3	5	0	8
Parasite prevalence >0 – <25%	6	58	8	72
Parasite prevalence >= 25%	0	58	79	137
Totals	9	121	87	217

The simple correlation between the contiguous measures of the FCS values and their respective estimates of the parasite prevalence was also tested. The continuous association was further tested using weighted least-squares regression with the model weighting the dependent variable (parasite ratio) for its precision using the $\text{Log}_{10}+1$ of the standard error of the parasite ratio. Covariates in this model included the end-year of the study (pre-1990, 1990–94, and 1995–2002), the survey sampling methodology (random, longitudinal or total population surveys versus surveys without precise details) and the maximum age range covered in the sample (up to 5 years, up to 10 years and up to 15 years). The proportion of variation in the parasite ratio explained by each variable in the model was calculated by comparing overall proportions of explained variation (R^2) between subsequent models with additional variables entered in a stepwise fashion allowing for all variables already in the model.

Results

A total of 217 spatially independent parasite prevalence surveys were identified that met the selection criteria. Eleven were reported in peer reviewed journals, 168 were unpublished Ministry of Health reports, NGO/bilateral/multilateral organisation reports, 16 were reported in doctoral or masters theses and 22 were provided as unpublished data by malaria scientists working in Kenya. Twenty-five percent ($n = 54$) of the surveys were undertaken before 1990. The median sample size in the childhood surveys was 220 (inter-quartile range 118, 430). 14.7% ($n = 32$) of the surveys were undertaken among age groups covering the 0–5 year age range; 61.8% ($n = 134$) spanned the age range 0–10 years; and 23.5% ($n = 51$) of surveys included children between the ages of 0 and 15 years. The majority of surveys were regarded as random, longitudinal or total population surveys (65.9%, $n = 143$). The distribution of the survey data against the categorical distinctions for the MARA FCS values is shown in Figure 1.

Table 1 compares the categorical definitions of parasite prevalence with categories of the FCS values used during recent malaria burden estimations [11,12]. Of the parasite surveys reporting a zero prevalence of infection, five out of eight had an FCS value greater than zero. One hundred and thirty seven surveys reported parasite prevalence greater or equal to 25% (*i.e.* stable endemic transmission), however only 79 (58%) had an FCS value \geq of 0.75. For surveys reporting a low parasite prevalence between 1–24% ($n = 72$), 58 (81%) were characterised by an FCS value >0 but less than 0.75. Overall agreement between the categories was poor, $k = 0.367$, although statistically significant ($p < 0.0001$). Table 2 reflects a change in the FCS categories around 0.6 to represent zero, 1–24% and $\geq 25\%$ parasite prevalence. These criteria greatly improved

the levels of agreement with the parasite ratio categories ($k = 0.442$, $p < 0.000$) and 68% of surveys reporting a parasite prevalence $\geq 25\%$ were described by an FCS value ≥ 0.6 . Further changes in the FCS criteria made little difference to the levels of agreement.

Regarding the parasite prevalence and the FCS values as continuous measures produced a weak positive correlation ($R^2 = 0.364$, $p < 0.0001$; Figure 2). Inclusion of covariates, upper age, year of survey or sampling method did not improve the association and nor did weighting the parasite ratio by its standard error during weighted least-squares regression.

Discussion

The analyses presented in this paper compare empirical parasite prevalence data among children sampled in 217 communities across Kenya with a climate-driven model that estimates the likelihood of stable malaria transmission. The results of the study suggest that there is a significant linear, albeit weak, association between these two measures of *P. falciparum* transmission (Figure 2&3). This is perhaps not surprising as the parasite ratio and the FCS value represent two very different transmission criteria. The parasite ratio reflects the intensity of transmission and has been routinely used as a marker of endemicity in Africa since the 1950's [24]. The FCS is a representation of the rainfall and temperature determinants of the parasite and vector's ability to coexist and thus enable stable transmission [5]. Furthermore, the MARA FCS is based on climatological averages for the 1951–1995 period (although recent analyses suggest this to have been surprisingly stable over the last century [25] and the parasite rate sampled in specific years. The potential confounding influence of timing on the parasite prevalence sample is an area of ongoing research.

The positive, albeit weak, correlation with measures of the intensity of transmission lends some support for the MARA model's ability to define populations at-risk of differing intensities of malaria transmission. There were too few surveys reporting zero infection prevalence ($n = 8$) to argue whether the FCS model can correctly distinguish areas of no transmission, however, 5 areas reporting zero prevalence did have a FCS value greater than zero. More striking was the ability of the FCS categories >0 and <0.75 to correctly identify populations at low risk of malaria infection with parasite prevalence's between $>0 - <25\%$ (81%: Table 1). During estimations of malaria burden it has been assumed that populations residing in these areas experience much lower risks of malaria-specific morbidity and mortality compared to populations located in areas described by an FCS ≥ 0.75 [11,12]. Conversely areas described as supporting stable, endemic transmission with parasite prevalence $\geq 25\%$ were less well described

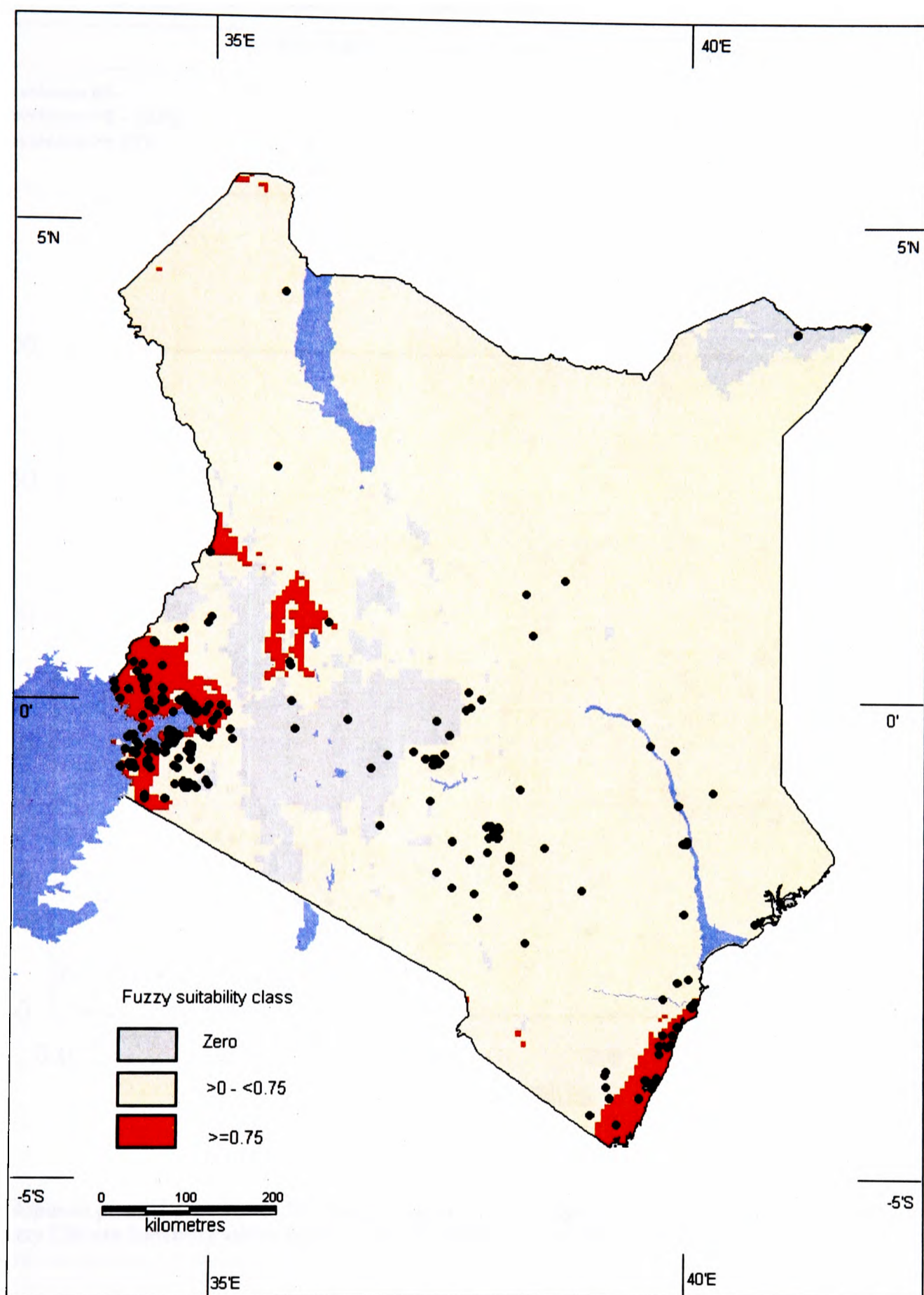


Figure 1
MARA Fuzzy Suitability Class (FCS) values categorised into three categories (light grey, zero FCS; light red FCS > 0 & < 0.75 & red ≥ 0.75) in Kenya showing distribution of selected parasite survey data points (black dots, n = 217)

Table 2: Agreement between parasite prevalence categories among 217 surveys and spatially congruent categorical values of the Fuzzy Climate Suitability (PCS) using the 0.6 threshold for stable endemic malaria.

	FCS = 0.00	FCS >0 - <0.6	FCS ≥ 0.6	Totals
Parasite prevalence 0%	3	4	1	8
Parasite prevalence >0 - <25%	6	56	10	72
Parasite prevalence ≥ 25%	0	44	93	137
Totals	9	104	104	217

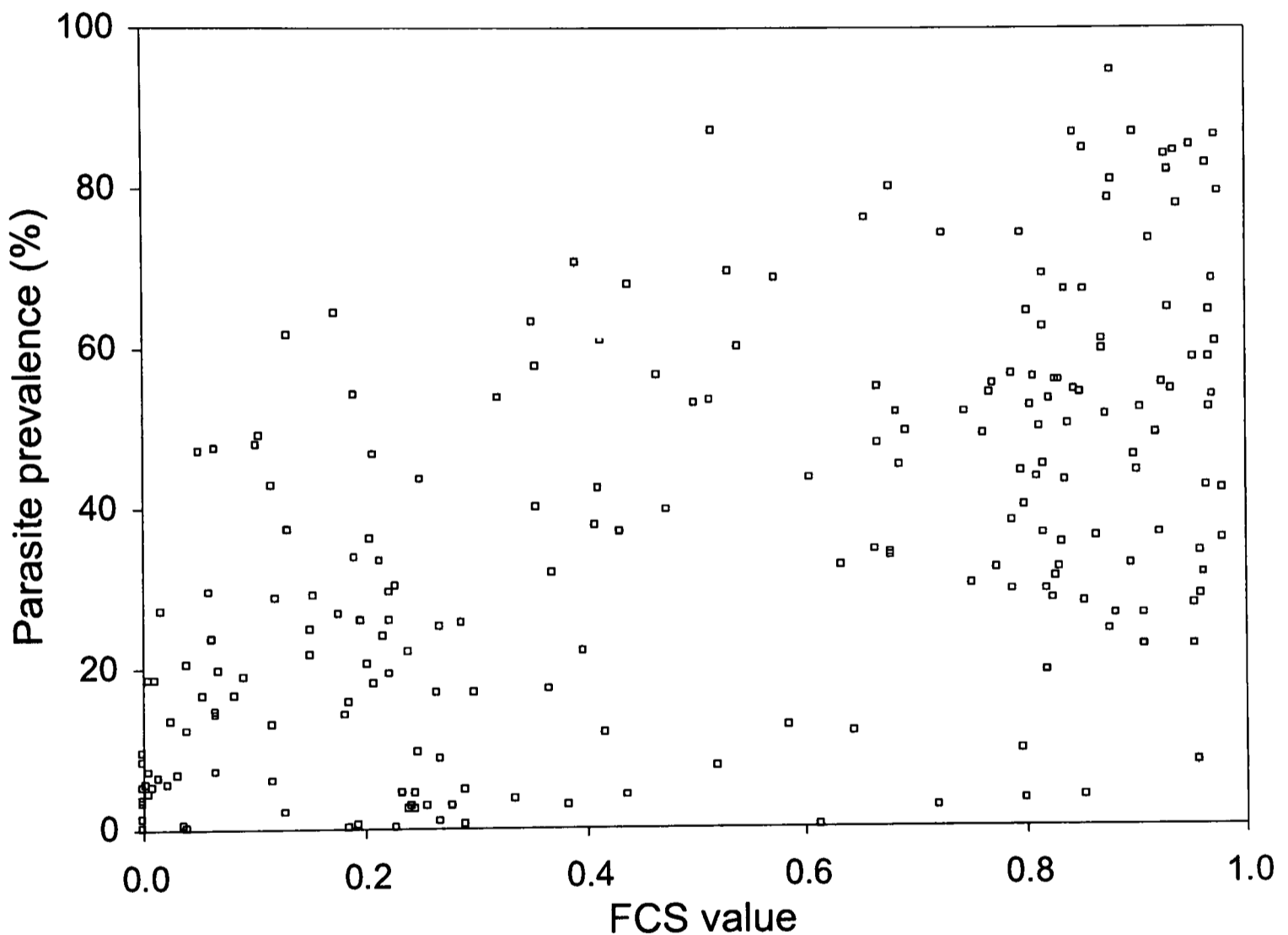
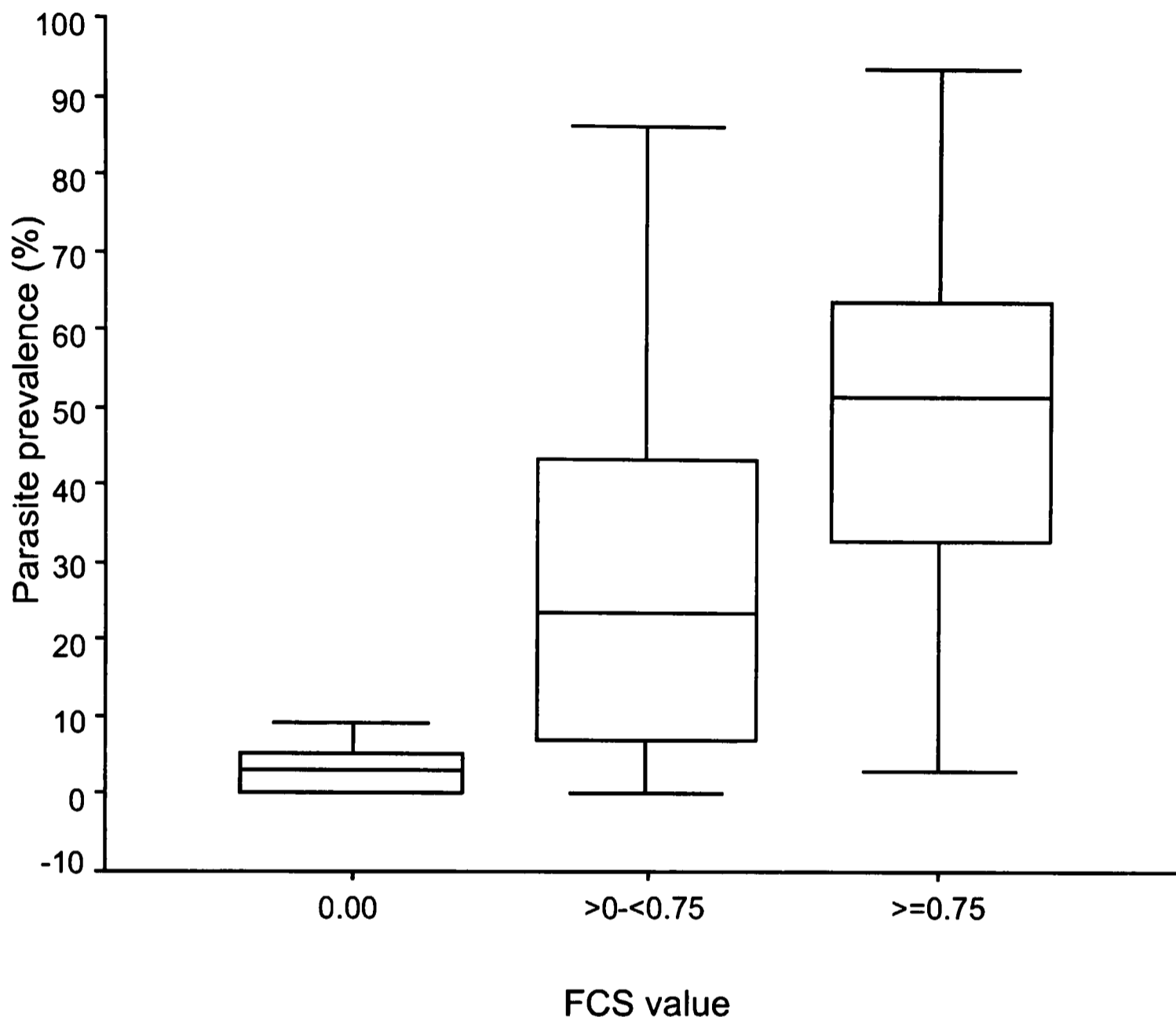


Figure 2
Plasmodium falciparum parasite prevalence (%) among children under the age of 15 years surveyed in 217 spatially independent surveys by Fuzzy Climate Suitability values extracted for the same spatial areas ($R^2 = 0.364$, $p < 0.0001$)

using the criteria $FCS \geq 0.75$. Only 58% of the parasite prevalence surveys reporting a parasite ratio $\geq 25\%$ were classified as having a FCS value ≥ 0.75 , the remaining 42% of surveys were classified as having FCS values >0 but less

than 0.75. Altering the FCS criteria to ≥ 0.6 for stable, endemic transmission improved the classifications of parasite prevalence surveys in Kenya with 68% of surveys correctly identified (Table 2).

**Figure 3**

A box plot of distribution of *Plasmodium falciparum* parasite prevalence (%) according to three FCS categories: >0 , $0 - <0.75$ & ≥ 0.75 . The median (central line), 25 and 75% quartile ranges (box length) and upper and lower limits (bars) are shown.

Conclusions

The results of this study suggest that applying climate suitability modelled estimates of transmission potential to distinguishing populations at differing levels of malaria infection intensity has several limitations. First it may not adequately distinguish populations at no risk of infection. Second, populations exposed to stable, endemic transmission may be poorly defined using criteria of ≥ 0.75 FCS. These results relate only to Kenya and similar validations are required in other settings in Africa. Nevertheless, for Kenya the disease and mortality burden of malaria will

have been under-estimated using the criteria shown in Figure 1. The extent of higher intensity transmission resulting in higher malaria burdens could be wider and the extent of no risk might be more restricted. The model developed by Craig *et al.* [5] represents the only available continental scale map upon which to define populations at risk of *P. falciparum* infection. Our analysis suggests that, at a crude level, it does distinguish communities exposed to differing levels of malaria transmission intensity. Given the large number of assumptions made regarding the paucity of attribute morbidity and mortality data

to define continental scale DALYs for malaria, defining the spatial extents of population denominators is only one part of the problem that is at least soluble. New malaria risk models are being developed which go beyond the climatic determinants of *P. falciparum* transmission and new iterations of malaria risk models will hopefully become more robust with the inclusion of new global scale data on population settlement, land use and ecology.

Authors' contributions

JAO carried out the data collection, entry, checking and analyses and drafted the manuscript. SIH and CAG obtained and extracted the ancillary MARA and urban extents data. RWS, SIH and JAO conceived the study and participated in its design and coordination. All authors read, edited and approved the final manuscript.

Acknowledgements

This study received financial support from The Wellcome Trust, UK; International Development Research Centre, Canada; the South African Medical Research Council, the MARA/ARMA collaboration and the Kenya Medical Research Institute. The Wellcome Trust supports JAO, SIH and RWS as part of their Prize Studentship (#060063), Research Career Development (#069045) and Senior Research Fellow (#058992) programmes respectively. The authors acknowledge the contribution of malaria control personnel and scientists and in Kenya who shared unpublished data. Particular thanks are due to the Division of Malaria Control (Kenya Ministry of Health), Vicki Marsh and Tabitha Mwangi (KEMRI), Dejan Zurovac (MSF-France) and Sporton Kwanya Agak, (KEDHAR). The authors also wish to thank Greg Fegan and Hesbon Ooko for statistical support and Abdisalan M Noor for help with the spatial mapping of the FCS values. This paper is published with the permission of the director of KEMRI.

References

1. Snow RW, Gouws E, Omumbo JA, Rapuoda BA, Craig MH, Tanser FC, le Sueur D, Ouma J: **Models to predict the intensity of *Plasmodium falciparum* transmission: applications to the burden of disease in Kenya.** *Trans Roy Soc Trop Med Hyg* 1998, **92**:601-606.
2. Kleinschmidt I, Omumbo JA, Briët O, van de Giesen N, Sogoba N, Mensah N, Windmeijer P, Moussa M, Teuscher T: **An empirical malaria distribution map for West Africa.** *Trop Med Int Health* 2001, **6**:779-786.
3. Briët OJT, Gunawardena DM, van der Hoek W, Amerasinghe FP: **Sri Lanka malaria maps.** *Malar J* 2003, **2**:22-32.
4. Hay SI: **An overview of remote sensing and geodesy for epidemiology and public health application.** *Adv Parasitol* 2000, **47**:1-35.
5. Craig MH, Snow RW, le Sueur D: **A climate-based distribution model of malaria transmission in sub-Saharan Africa.** *Parasitol Today* 1999, **15**:105-111.
6. PAHO: **Status of malaria programmes in the Americas (based on 2002 data). Report of the 44th directing council, 55th session of the regional committee.** *Pan American Health Organization (PAHO)/World Health Organization (WHO): Washington D.C* 2003.
7. Hay SI, Guerra CA, Tatem AJ, Noor AM, Snow RW: **The global distribution and population at risk of malaria: past, present and future.** *Lancet Infect Dis* 2004, **4**:327-336.
8. World Health Organization: **The Africa Malaria Report 2003.** *World Health Organization/United Nations Children's Fund: Geneva/New York* 2003:120.
9. World Health Organisation: **Rolling back malaria. The world health report 1999: making a difference.** *World Health Organization: Geneva* 1999:49-63.
10. Snow RW, Craig MH, Deichmann U, Marsh K: **Estimating mortality an disability due to malaria among Africa's non-pregnant population.** *Bull World Health Organ* 1999, **77**:624-640.
11. Snow RW, Craig MH, Newton CRJC, Steketee RW: **The public health burden of *Plasmodium falciparum* malaria in Africa: deriving the numbers. Working Paper 11, Disease Control Priorities Project.** *Fogarty International Center, National Institutes of Health: Bethesda, Maryland* 2003 [<http://www.fic.nih.gov/dcpp>].
12. Snow RW, Omumbo JA: **Malaria mortality in sub-Saharan Africa (in press).** In *Disease and mortality in sub-Saharan Africa* Edited by: Jamison D, Feacham R, Makgoba W, Hofman K, Rogo K. Oxford University Press & World Bank; 2004 in press.
13. Omumbo JA, Ouma J, Rapuoda B, Craig MH, le Sueur D, Snow RW: **Mapping malaria transmission intensity using geographical information systems (GIS): an example from Kenya.** *Ann Trop Med Parasitol* 1998, **92**:7-21.
14. MARA/ARMA: **Towards an atlas of malaria risk in Africa. First technical report of the MARA/ARMA collaboration.** *MARA/ARMA. Durban* 1998.
15. Directorate of Overseas Surveys: **1:50,000 topographical maps. Series Y731. Survey of Kenya: Nairobi, Kenya** 1971.
16. Central Bureau of Statistics: **1999 population and housing census: counting our people for development. Volume 1.1 Population distribution by administrative and urban centres.** *Ministry of Finance & Planning, Nairobi, Kenya.* 2001.
17. GDE Systems: **Populated places: GeoName Digital Gazetteer v.1 (CD-ROM).** *Geographic Information Department (MZ 1211-M), GDE Systems Inc., P.O. Box 509009, San Diego, California* 1995.
18. World Resources Institute: **Topographic data: Africa Data Sampler (CD-ROM).** *World Resources Institute, USA* 1995.
19. SALB: **Second Administrative Level Boundaries (SALB) Project overviews: concepts, progress and future.** 2004 [http://www3.who.int/whosis/gis/salb/salb_po.htm].
20. Snow RW, Gilles HM: **The epidemiology of malaria.** In *Bruce Chwatt's Essential Malariology* Edited by: Warrell D, Gilles H. Arnold Publishers: London; 2002:85-106.
21. Service M: **Mosquito (Diptera: Culicidae) dispersal – the long and short of it.** *J Med Entomol* 1997, **34**:579-588.
22. Cohen J: **A co-efficient of agreement for nominal scale.** *Educ Psychol Measurements* 1960, **20**:37-46.
23. Landis J, Koch G: **The measurement of observer agreement for categorical data.** *Biometrics* 1977, **33**:159-174.
24. Metselaar D, Van Theil P: **Classification of malaria.** *Trop Geog Med* 1959, **11**:157-161.
25. Small J, Goetz SJ, Hay SI: **Climatic suitability for malaria transmission in Africa, 1911–1995.** *Proc Natl Acad USA* 2003, **100**:15341-15345.

Publish with **BioMed Central** and every scientist can read your work free of charge

"BioMed Central will be the most significant development for disseminating the results of biomedical research in our lifetime."

Sir Paul Nurse, Cancer Research UK

Your research papers will be:

- available free of charge to the entire biomedical community
- peer reviewed and published immediately upon acceptance
- cited in PubMed and archived on PubMed Central
- yours — you keep the copyright

Submit your manuscript here:

http://www.biomedcentral.com/info/publishing_adv.asp



CHAPTER 5:

The influence of urbanization on measures of *Plasmodium falciparum* infection prevalence in East Africa

JA Omumbo, CA Guerra, SI Hay, RW Snow

***Acta Tropica* (2005), 93: 11-21**



The influence of urbanisation on measures of *Plasmodium falciparum* infection prevalence in East Africa

J.A. Omumbo^{a,*}, C.A. Guerra^b, S.I. Hay^{a,b}, R.W. Snow^{a,c}

^a Public Health Group, KEMRI/Wellcome Trust Research Programme, P.O. Box 43640, 00100 GPO, Nairobi, Kenya

^b Department of Zoology, TALA Research Group, Tinbergen Building, University of Oxford, South Parks Road, Oxford, UK

^c Centre for Tropical Medicine, John Radcliffe Hospital, University of Oxford, Oxford, UK

Received 1 June 2004; received in revised form 5 August 2004; accepted 17 August 2004

Available online 27 October 2004

Abstract

There is a growing interest in the effects of urbanisation in Africa on *Plasmodium falciparum* risks and disease outcomes. We undertook a review of published and unpublished literature to identify parasite survey data from communities in East Africa. Data were selected to represent the most reliable and contemporary estimates of infection prevalence and were categorised by urban or rural status using a number of approaches. We identified 329 spatially distinct surveys undertaken since 1980 in the sub-region of which 37 were undertaken in urban settlements and 292 in rural settlements. Overall rural settlements reported significantly higher parasite prevalence among children aged 0–14 than urban settlements (on average 10% higher infection rates; $p < 0.05$). No urban settlements recorded parasite prevalence in excess of 75%. In areas of East Africa where climatic conditions are likely to support higher parasite transmission, the rural–urban difference was most marked. There was a significant trend towards documenting higher classes of parasite prevalence in rural compared to urban settlements ($p < 0.05$) and the mean difference between rural and urban samples was 18% ($p < 0.001$). These results further highlight the need to better define urban extents in Africa in order to capture the non-climatic determinants of infection and disease risk and provide a more informed approach to describing the burden of disease across the continent.

© 2004 Elsevier B.V. All rights reserved.

Keywords: Malaria; Urbanisation; Disease burden; *Plasmodium falciparum* prevalence; East Africa; GIS; Maps

1. Introduction

The majority of the Disability-Adjusted Life Years (DALY) for sub-Saharan Africa (SSA) is driven by high mortality in early childhood from infectious diseases (diarrhoeal and respiratory diseases, measles and malaria) and high mortality in young adults from HIV

* Corresponding author. Tel.: +254 20 2720163;
fax: +254 20 2711673.

E-mail address: jomumbo@wtairobi.mimcom.net
(J.A. Omumbo).

infection and tuberculosis (Murray and Lopez, 1997; World Health Organization, 2002). The DALY approach to prioritising disease intervention has led to improved efforts to define disease-specific mortality rates. Despite the paucity of data, pan-African extrapolations of specialized survey data on acute respiratory tract infections (Williams et al., 2002), HIV (Walker et al., 2002) and malaria (Snow et al., 1999; Snow and Omumbo, 2004) have been used to re-estimate the contribution made by these infectious diseases to the mortality burden in SSA. Only the estimations for malaria have allowed for the spatial determinants of populations-at-risk of infection and their effect upon subsequent disease risks.

At the population level, the chances and frequency of *Plasmodium falciparum* infection risk influence age-specific disease outcomes: low infection risks result in lower malaria mortality rates and are not influenced by acquired immunity; increasing intensities of transmission support increasing mortality risks and are markedly influenced by age-specific acquired immunity (Snow and Marsh, 2002). The transmission patterns of *P. falciparum* in Africa vary between countries, and more importantly for disease burden estimation, within national borders. The estimates of the malaria burden in Africa have used populations-at-risk of different climate-determined patterns of transmission to reflect the spatial heterogeneity of infection risk on disease outcome (Craig et al., 1999; Snow et al., 1998; Snow and Omumbo, 2004).

One important caveat to this approach has been the inability of the current continental-scale malaria risk maps to reflect other influences upon transmission, notably urbanisation. Two recent reviews suggest that urban populations in Africa receive, on average, 10 times fewer new infections from vectors compared with their rural counterparts (Hay et al., 2000; Robert et al., 2003). In 2000, 38% of Africa's 784 million inhabitants were urban dwellers and this is expected to increase to 55% by 2030 (United Nations, 2002). This shift in human populations from environments supporting predominantly communicable disease burdens to those lending themselves to non-communicable diseases will have profound epidemiological and public health impacts (Harpham, 1997). In this paper we examine differences in *P. falciparum* prevalence between urban and rural settlements of East Africa to explore the likely impact of urbanisation on malaria burdens in this sub-region.

2. Material and methods

2.1. *P. falciparum* prevalence survey data among children aged 0–15 years in East Africa

A search was conducted of published and unpublished literature related to malaria infection prevalence surveys in the East Africa region. Electronic database searches were performed using Medline[®] (SilverPlatter International, Boston, MA, USA, 2000), Popline[®] (Johns Hopkins School of Hygiene and Public Health, Baltimore, MD, USA, 2000) and EmBase[®] (Elsevier Science, Little Rock, AR, USA, 1999–2000). The following keywords were used in the search: malaria, parasite and malaria and Kenya, Tanzania, Uganda, *P. falciparum*, parasite rate, parasite prevalence and malaria transmission. For each publication, bibliographies were cross-referenced to identify additional sources of information and other studies. Where additional details could not be identified through the published sources, authors were contacted to provide more information on geographical location, survey dates and age-specific characteristics of the parasitological data. Postgraduate theses held in the libraries of Nairobi, Makerere and Muhimbili Universities were also searched. Annual reports, journals and conference proceedings of national medical research institutes, and non-governmental organisations were reviewed in each country. Results of routine parasite prevalence surveys undertaken by the Ministry of Health's (MoH) Vector or Malaria Control Departments in Kenya and Uganda were manually searched during field visits to archives at headquarters in each country, and in Kenya at provincial offices.

Using a variety of sources: 1:50,000 scale topographic maps (Directorate of Overseas Surveys, 1971), digital maps of administrative units in Kenya (Central Bureau of Statistics, 1989) or public domain digital gazetteers (GDE Systems, GeoName, 1995; World Resources Institute, Africa Data Sampler, 1995), a longitude and latitude was ascertained for each parasitological survey in decimal degrees. This centroid served as a unique identifier for each survey and, combined with a description of the survey, was used to describe the spatial extent of each sample. Surveys were classified as representing one of five spatial dimensions. First, surveys representing a single village where the central longitude and latitude was used to define point

estimates covering the community using ArcView 3.2 (ESRI, Redlands, CA, USA). The second spatial classification reflected surveys that sampled from several villages but presented the data as a single parasite ratio (five and seven surveys in Kenya and Tanzania, respectively). In this case, a polygon was created to connect the villages. The third, fourth and fifth spatial criteria corresponded to surveys undertaken at the fifth (sub-location: average area covered 9.4 km²), fourth (location: average area covered 15.4 km²) or third (division: average area covered 34.1 km² administrative levels in Kenya (Hay et al., in press).

All surveys undertaken in Kenya, Tanzania and Uganda ($n=2003$) were subjected to a number of selection criteria for inclusion in the present analysis (Table 1). The selection allowed for contemporary assessments of infection risk (>1979); inclusion of community-based surveys; age ranges within 0–15 years (excluding surveys where only infants sampled); minimum sample sizes of 50 children; and adequate details on location, and denominators. Finally, repeat surveys by the same investigators within a 24-month period were combined to a single estimate and surveys undertaken by several investigators at different times in the same location were reduced to one estimate by excluding either the earlier survey or smaller sample size.

2.2. Defining urban and rural extents

Urbanisation is a difficult demographic concept to define and there are inconsistencies between international, national and published definitions (Vlahov and Galea, 2002; Tatem and Hay, in press). We have used a combination of approaches to classify urban surveys. First we accepted the investigators urban/rural description as stated in the survey reports. Second, we used the name of the survey site to validate the descriptions, e.g., if a study site was reported as “Malindi town”. Third, we compared national census bureau administrative boundaries (Central Bureau of Statistics, Kenya, 1989; Statistics Department Uganda, 1995; Bureau of Statistics, Tanzania, 2002) of known urban extents against the spatial position of each survey to identify those whose survey report or community name did not provide adequate information to classify them as urban or rural. Here we classified urban settlements if they were located in national census bureau definitions of cities, towns or market centres.

Finally we used a global database of urban extents developed as part of the Global Rural–Urban Mapping Project (GRUMP) (Balk et al., 2004; CIESIN/CIAT, 2004; CIESIN/IPFRI/CIAT, 2004). The urban mask was developed at a 1 km \times 1 km spatial resolution using data on night time lights (Sutton et al., 2001) and Landsat satellite imagery in combination with other geographic data (Digital Chart of the World populated places (Mika, 1997), Tactical Pilotage Charts produced from Australian Defence Imagery and Geo-Spatial Organisation, and national census data). The spatially distinct polygons or points for each parasite prevalence survey were created using a digital administrative map of Kenya (CBS, 1989), first warped to UN approved national boundaries (SALB, 2004) using ArcView 3.2 (ESRI, Redlands, CA, USA). The GRUMP urban extents were mapped to the single longitudes or latitudes or polygons of each the parasite surveys using ArcView 3.2 (ESRI, Redlands, CA, USA). For surveys covering wider areas (clusters of communities or third–fifth level administrative units), the percentage of the polygon covered by urban extents was extracted using the MILA Utilities 3.2 update extension within ArcView 3.2 (<http://www.esri.com/arcscrips>). If more than 75% of the polygon was classified as either urban or rural the survey was classified accordingly.

2.3. Malaria ecological zonation of East Africa

To compare urban–rural differences in parasite prevalence within broad, climate-derived malaria risk classes we have used an 8 km \times 8 km resolution malaria risk map developed to approximate to the historical descriptions of malaria transmission in East Africa (Omumbo et al., 2002). In brief, Fourier processed data from remotely sensed images related to land surface temperature (LST), rainfall and moisture availability as inferred by cold cloud duration (CCD), and normalised difference vegetation index (NDVI) data were used to construct a model of malaria risk. Discriminant analysis of the models ability to predict the historical categorization of four malaria risk zones in Kenya, Tanzania and Uganda showed that malaria free areas could be predicted with a 96% accuracy, while areas where transmission occurs only near water, moderate malaria areas and intense malaria transmission areas were predicted with accuracies of 90, 72 and 87%, respectively. For the purposes of the

Table 1
Selection criteria for parasite prevalence survey data in East Africa and description of selected surveys

	Kenya	Tanzania	Uganda	Total
Total number of surveys identified	992	522	489	2003
Exclusion criteria				
Excluded pre-1980	236	194	237	667
Excluded because unable to locate longitude/latitude	0	33	2	35
Excluded because no information on denominator	9	1	13	23
Excluded sample size <50	82	9	69	160
Excluded because non-community sampled ^a	408	128	121	657
Excluded sample age groups \neq 0–15 or infants only	20	39	19	78
Excluded because spatial overlap with more contemporary or larger sample size survey ^b	20	32	1	53
Characteristics of surveys				
Final survey samples included in analysis	217	86	27	330
Surveys reported in peer-reviewed journals	11	24	7	42
Surveys reported in reports of Ministry of Health, NGO or donor reports	168	34	17	219
Surveys reporting total population sampling, longitudinal sampling, random sampling ^c	143	86	16	245
Surveys covering period 1980–1989	54	32	2	88
Surveys covering period 1990–1994	100	14	14	128
Surveys covering period 1995–2003	63	40	11	114
Median (interquartile range) of sample sizes	221 (119, 431)	194 (146, 443)	88 (76, 265)	204 (120, 427)
Surveys covering age ranges within 0–5 years ^d	32	61	0	93
Surveys covering age ranges within 0–10 years ^d	134	16	17	167
Surveys covering age ranges within 0–15 years ^d	51	9	10	70

^a Several surveys reported follow-up of patients during drug resistance surveys, or patients sampled during routine EPI/MCH clinic visits. These were excluded on the basis that it was difficult to know whether they represented community-based infection prevalence. In addition surveys conducted during randomized controlled trials of interventions were only included for control populations, intervention arms were excluded.

^b For surveys which were contiguous with others, the most recent, largest sample size and most detailed in its reporting of sampling strategy was included.

^c Several surveys, particularly those reported by Vector Control or Malaria Divisions of the Ministry of Health, did not provide adequate sampling strategy details (26.4% ($n = 88$) of the overall survey samples). It is unclear whether these prevalence estimates have the same precision as surveys providing detailed descriptions of sampling and survey method.

^d The surveys included in the analysis were not all consistent with a single age range. The surveys varied in the age groups sampled, e.g. 0–4, 1–4, 0–3, 0–9, 1–9, 1–12, 0–15, etc. The table represents the summary of three broad age classes using the upper age range as the defining age for sub-groups. Separate analyses have shown that the categorical description of endemicity based on the parasite ratio does not alter significantly within one area by different age groupings within the 0–15 years age range (Omumbo et al., in preparation).

present analysis we have grouped the malaria risk areas: malaria free with malaria only near water and moderate with intense transmission areas to distinguish between low and moderate-to-high infection risks (Fig. 1).

Parasite survey data were classified in accordance with the two malaria ecozones based on the polygon spatial extents of each survey (Table 3) with a 2.5 km buffer (approximating to the average flight distances of *Anopheles gambiae s.l.* (Service, 1997)) using ArcView 3.2 (ESRI, Redlands, CA, USA). Parasite survey data were classified as only one ecozone category

where more than 75% of the surface area was dominated by that zone class.

2.4. Data entry and statistical methods

All parasite survey data were recorded on standard proformas and entered twice in Microsoft[®] Access version 7 (Microsoft Corporation 1989–1996; Seattle, WA, USA). Data entry errors and range inconsistencies were checked and verified against original material. Data were extracted for matching to other data sources using Microsoft[®] Excel 2000 version 9.0. Data

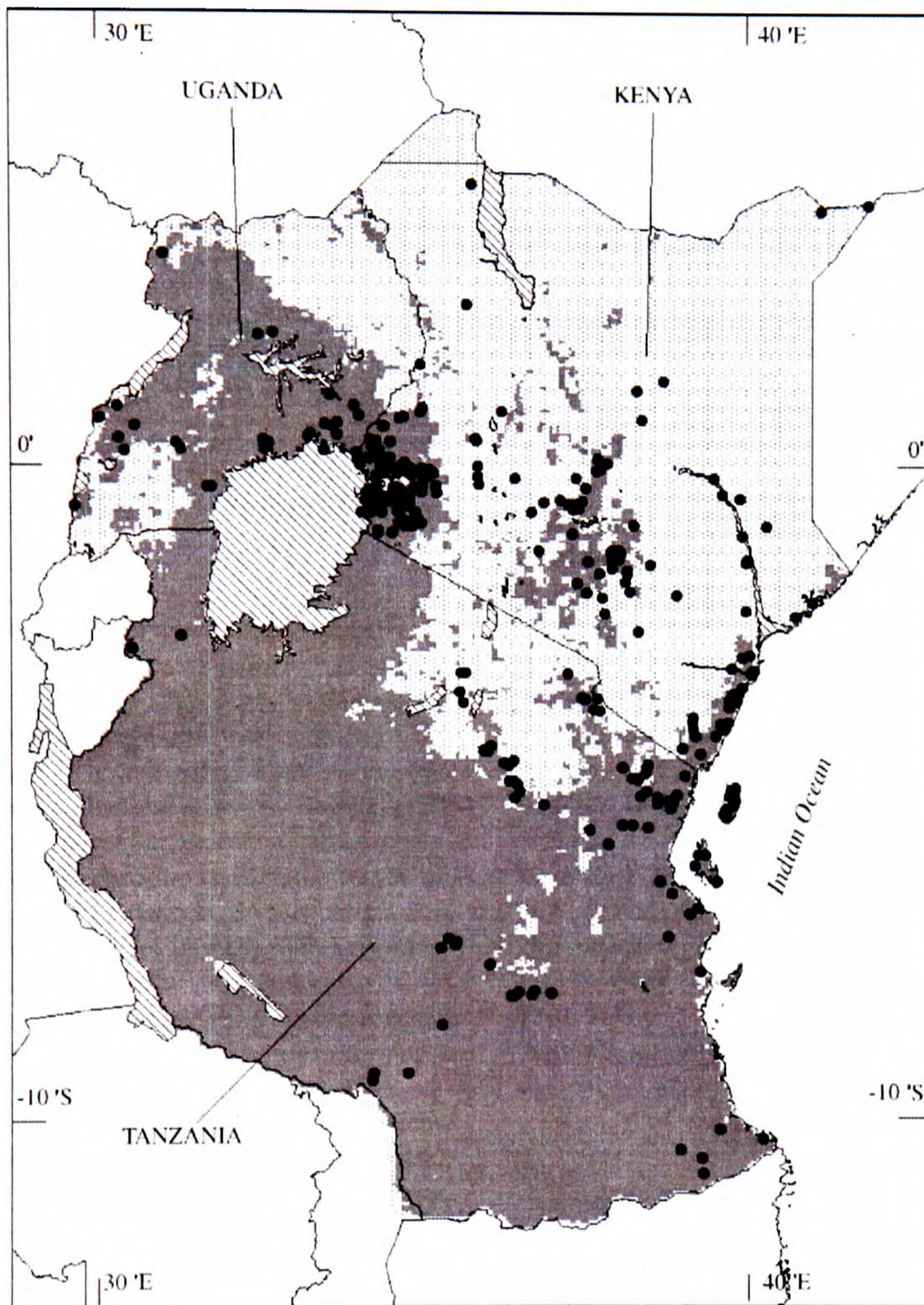


Fig. 1. Distribution of parasite prevalence surveys by malaria ecological zone in East Africa. Malaria-free or transmission occurring only near water (light grey), and moderate to high transmission areas (dark grey). Parasite prevalence survey sites included in the analysis are shown as black circles.

were analysed using SPSS for Windows Release 10.01 (SPSS, 1999, Chicago, IL, USA). Mean differences between urban and rural survey reports of parasite prevalence were tested using analysis of variance (ANOVA) and categorical descriptions of the ranges of parasite prevalence between urban and rural settlements were tested using a chi-square for trend.

3. Results

A total of 330 spatially independent parasite surveys were identified that met the selection criteria across East Africa. The exclusions and characteristics of these surveys are shown in Tables 1 and 2. The surveys in the three countries were distributed across the two

Table 2
Spatial resolution of 329 *P. falciparum* parasitological surveys among children in Kenya, Uganda and Tanzania

Description	Kenya	Uganda	Tanzania
Village/community ^a	85	27	79
Cluster of villages/communities	5	0	7
Sub-location	112	0	0
Location	7	0	0
Division	8	0	0
Surveys in sub-locations, locations, divisions excluded as urban/rural extent could not be classified	1	0	0
Total sample included in analysis	216	27	86

^a Eight studies in Uganda and five studies in Zanzibar/Pemba were recorded as villages or single communities where data were presented at sub-fifth level administrative boundary extents but covered dispersed areas.

principal malaria ecological zones as shown in Fig. 1 and Table 3. The step-wise classification of urban/rural settlements left one survey sufficiently ambiguous at the final classification stage and was excluded from subsequent analysis. Three hundred twenty-nine surveys were included in the final analysis comprising 37 urban and 292 rural surveys. Almost all urban settlements were adequately described within the survey reports, survey site name or national census bureau definitions of urban settlements ($n=28$). GRUMP identified a further nine areas that were ambiguous from the survey reports and were not defined in the national census bureau urban centres. However, we were struck by omissions, within GRUMP, of obvious urban settlements ($n=15$) (e.g., within the city limits of Mombasa, large towns such as Zanzibar, Kilifi and Malindi and market centres in Busia and Ifakara). Thus, using a combined series of approaches to defining urban settlements was deemed an appropriate strategy and highlights the continued lack of comprehensive coverages of urban settlements in the sub-region.

Table 3
Malaria ecozone (Omumbo et al., 2002) classification for 329 independent parasite survey data in East Africa according to urban and rural coverage

Ecozone	Kenya (216)		Uganda (27)		Tanzania (86)		Total	
	Rural	Urban	Rural	Urban	Rural	Urban	Rural	Urban
Malaria free or only near water	58	4	7	0	19	1	84	5
Moderate-to-intense transmission	137	17	11	9	60	6	208	32
Total	195	21	18	9	79	7	292	37

A comparison of the mean parasite prevalence in urban and rural settlements demonstrated that rural surveys reported on average a 10% higher prevalence of infection in childhood compared to urban surveys (Table 4: ANOVA, F statistic = 5.37, $p=0.021$). This difference persisted when rural and urban surveys were compared between areas of low malaria risk (8% lower in urban areas) and areas of moderate-to-high malaria risk (18% lower in urban areas) (Table 4). Only the rural–urban differences in the ecologically classified areas of moderate-to-high transmission, however, were statistically different (ANOVA, F statistic = 11.06, $p=0.001$). The median and interquartile ranges of the overall and malaria ecozone urban and rural samples are shown in Fig. 2.

Table 4 also shows the comparison of the categorical ranges of parasite prevalence reported by surveys undertaken in rural and urban areas. No surveys undertaken in urban settlements reported parasite prevalence in excess of 75%. Overall the trend toward describing higher parasite prevalence (≥ 25 –74% or $\geq 75\%$) in rural (192/292, 65.8%) compared to urban (23/37, 62.2%) settlements was not significant ($\chi^2=5.48$, d.f. = 2, $p=0.065$). Comparison of urban–rural parasite distributions in areas of low transmission was not possible due to the small numbers of samples in this category. However, in moderate-to-high transmission malaria risk areas the difference between rural (161/208, 77.4%) and urban (22/32, 68.8%) reporting higher parasite prevalence was significantly different ($\chi^2=6.37$, d.f. = 2, $p=0.041$).

4. Discussion

We have set out to examine the differences in *P. falciparum* infection prevalence among rural and urban children sampled as part of parasitological descriptions of

Table 4

The mean, (95% CI) percentage parasite prevalence and distribution of infection prevalence^d between rural and urban surveys undertaken in two malaria ecological zones in East Africa

	Rural	Urban
Overall mean (95% CI) parasite prevalence	39.5 (36.4–42.6)	29.0 (23.9–34.1) ^b
Low malaria risk areas, mean (95% CI) parasite prevalence	22.2 (17.3–27.1)	14.5 (6.7–22.3)
Moderate-to-high transmission areas, mean (95% CI) parasite prevalence	46.5 (43.1–49.9)	31.3 (25.6–37.0) ^c
Overall number (%) surveys reporting prevalence <25%	100 (34.2%)	14 (37.8%)
Overall number (%) surveys reporting prevalence ≥25% to <75%	154 (52.7%)	23 (62.2%)
Overall number (%) surveys reporting prevalence ≥75%	38 (13.0%)	0 (0.0%)
Moderate-to-high transmission areas, number (%) surveys reporting prevalence <25%	47 (22.6%)	10 (31.3%) ^b
Moderate-to-high transmission areas, number (%) surveys reporting prevalence ≥25% to <75%	127 (61.1%)	22 (68.8%)
Moderate-to-high transmission areas, number (%) surveys reporting prevalence ≥75%	34 (16.4%)	0 (0.0%)

^a The distribution of infection prevalence within low malaria risk areas is not shown, as there were only five urban surveys in this malaria risk class.

^b ANOVA comparison of means between urban and rural samples *p* value <0.05.

^c Chi-square test comparison between urban and rural samples differences in prevalence criteria *p* value <0.05.

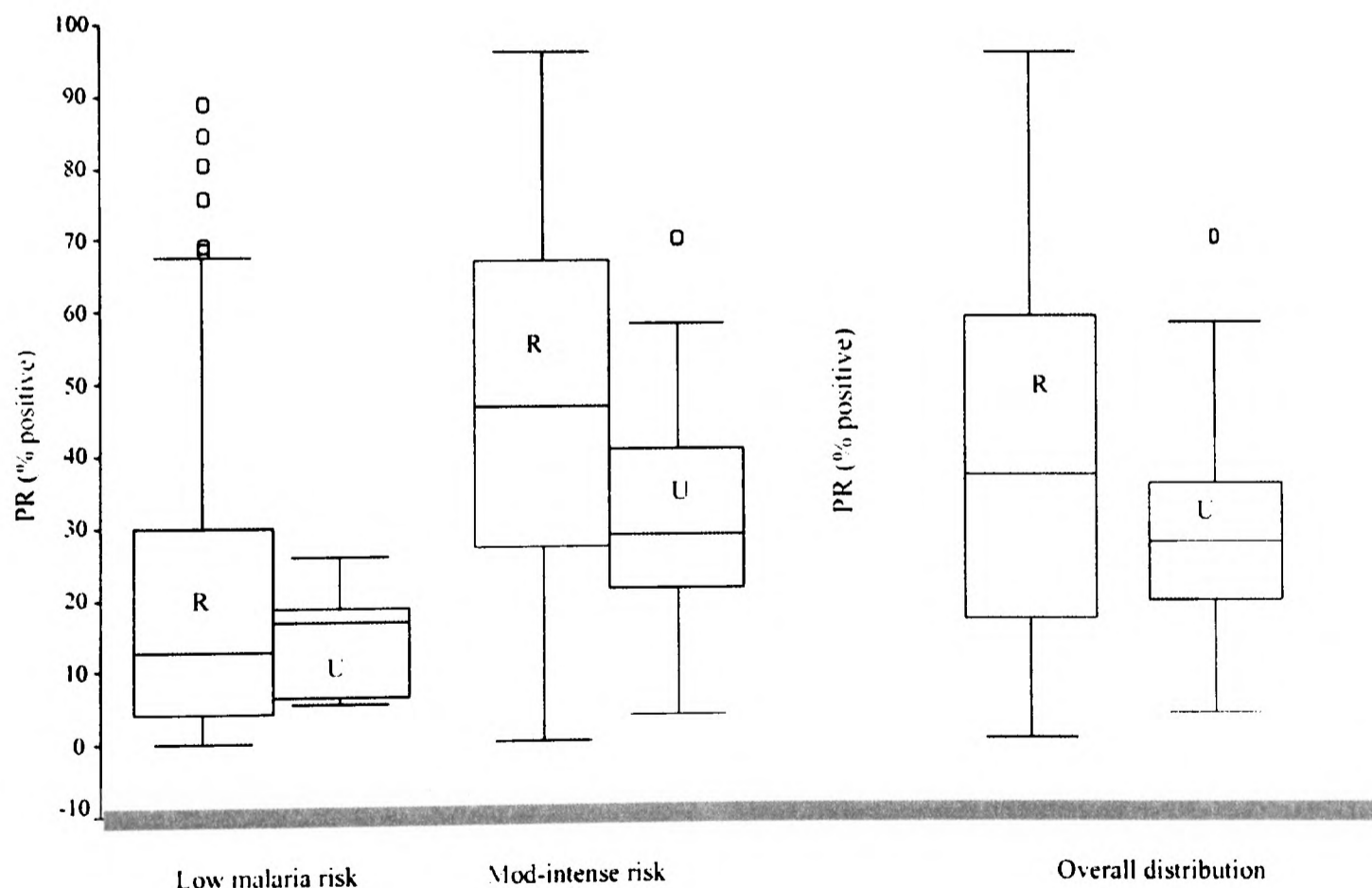


Fig. 2. Distribution of parasite prevalence data from 329 surveys in East Africa according to rural (R) vs. urban (U) classification in low and moderate-intense malaria risk areas and overall. Outliers (open circles) are cases with values between 1.5 and 3 box lengths from the upper edge of the box. The box length is the interquartile range and the line within the box represents the median.

malaria endemicity in East Africa. We have applied rigorous selection criteria to the data identified through an exhaustive search of the literature (Table 1). We believe that reviews of previously published or unpublished literature should apply strict selection criteria to ensure comparability in the data being analysed. It is important to note that our final sample of surveys was, by definition, not a spatially random series of surveys and subsequent analysis will always be limited by where the majority of data derive. The filtering and attribution of surveys to urban or rural classifications resulted in only 37 surveys undertaken in urban areas compared to 292 in rural areas. The limited information from urban settlements probably reflects the general consensus that malaria is a rural problem and epidemiologists often focus their investigations on these communities. Nevertheless, for those surveys identified, among childhood populations in East Africa there were significant differences in the parasite prevalence estimates reported for rural and urban settlements. Overall rural settlements reported 10% higher parasite prevalence compared to urban communities and this difference was larger (18%) between rural and urban communities located in moderate-to-high malaria ecological risk areas of East Africa (Table 4; Fig. 2).

These findings are consistent with micro-epidemiological studies of urban malaria across Africa: decreasing risks of infection associated with urbanisation, either entomological or parasitological, have been shown in urban areas in West Africa: Liberia (Björkman et al., 1985), The Gambia (Lindsay et al., 1990), Senegal (Trape et al., 1992), Ghana (Gardiner et al., 1984), Burkina Faso (Robert et al., 1986; Rossi et al., 1986; Benasseni et al., 1987; Modiano et al., 1999); Central Africa: Cameroon (Fondjo et al., 1992), Republic of Congo (Trape and Zoulani, 1987), Democratic Republic of Congo (Ngimbi et al., 1982; Coene, 1993); the Horn of Africa: Ethiopia (Yohannes and Petros, 1996), Sudan (El Sayed et al., 2000); and Southern Africa: Zambia (Watts et al., 1990). Moreover, in a recent meta-analysis of 159 sites across Africa where annual entomological inoculations rates (EIR) had been recorded, it was shown that people in rural areas received on average 146 *P. falciparum* infected bites per annum compared to only 14 for people resident in urban areas (Hay et al., 2000).

Recent years have seen a satellite imagery-fuelled renaissance in attempts to produce risk-maps of vector-

borne diseases including malaria (Hay, 2000). To a large extent, both historical and contemporary malaria mapping have focussed on extrinsic climatic determinants of disease distribution (Craig et al., 1999; Rogers et al., 2002), ignoring other spatial determinants of malaria risk, most notably the effects of urbanisation. This is primarily because these parameters have been harder to define at the coarse spatial resolution of the climate data currently used to model malaria risk. This is likely to influence the estimation of disease burden. The risks of all-cause childhood mortality are substantially and consistently lower in urban compared to rural areas of Africa (MEASURE, 2004). Our results and those of others have consistently shown malaria infection risks are lower in urban settings compared to rural settings. Severe malaria morbidity (Snow et al., 1997), malaria mortality (Snow and Marsh, 1995) and all-cause mortality (Snow and Marsh, 2002; Snow and Omumbo, 2004) are all substantially lower in areas of low parasite prevalence (childhood infection prevalence less than 25%) compared to areas of higher stable endemicity in Africa. In addition to the biological constraints on infection risk, urban communities are more likely to have either the financial means or better access to measures which further reduce the risks of infection, such as insecticide treated nets (Holtz et al., 2002; Mugisha and Arinawe, 2003), intermittent presumptive treatment during pregnancy (Guyatt et al., 2004), or better malaria case-management and anti-malarial treatment (Rasheed et al., 2000; Molyneux et al., 1999).

The categorical descriptions of endemicity using the parasite prevalence in our analysis of previously published surveys demonstrated that no urban settlements were able to support an infection prevalence in excess of 75% and that there was a significant difference in areas with a climate suitable for stable, endemic, transmission between urban and rural settlements supporting parasite prevalence ratios in excess of 25%. In Kenya, Tanzania and Uganda, approximately 28% of the population reside in urban settlements (UN, 2002) and it seems reasonable to assume therefore that in this sub-region a significant proportion of the population will be misclassified in terms of infection and subsequent disease risk on the basis of climate-driven cartography of risk alone.

For the first time spatially explicit estimates of population living in urban and rural environments are

becoming available in the public domain. These surfaces have been implemented at global scales (Schneider et al., 2003; Balk et al., 2004; Tatem et al., in press). While these resources open significant new possibilities for exploring the impact of urbanization on malaria transmission over large areas, an important step in testing and validating their precision locally is needed as evidenced by discrepancies noted in our urban classification of parasitological survey data. This is warranted due to the novelty of these digital themes and the various techniques used to designate urban and rural land-use classes both as input training data (Tatem and Hay, in press) and output map products (Tatem et al., in press). These precision estimates for Kenya, Tanzania and Uganda are the subject of on-going research, including the use of higher spatial resolution optical and synthetic aperture radar (SAR) satellite imagery. New, higher resolution, better-defined spatial descriptions of urban settlements will ultimately improve our continental scale estimation of malaria risk and disease burden.

Acknowledgements

This study received financial support from The Wellcome Trust, UK; International Development Research Centre, Canada; the South African Medical Research Council, the MARA/ARMA collaboration and the Kenya Medical Research Institute. The Wellcome Trust supports JAO, SIH and RWS as part of their Prize Studentship (#060063), Research Career Development (#056642) and Senior Research Fellow (#058992) programmes, respectively. The authors acknowledge the contribution of malaria control personnel and scientists and in East Africa who shared unpublished data. In particular we thank the Division of Malaria Control (Kenya Ministry of Health), Don de Savigny and Rose Lusinde (Tanzania Essential Health Interventions Project-TEHIP), Hugh Reyburn and Chris Drakeley (Gates Malaria Program, London School of Hygiene and Tropical Medicine), Caroline Maxwell, Chris Curtis and Ilona Carneiro (London School of Hygiene and Tropical Medicine), Ambrose Talisuna, Peter Langi and Albert Kilian (Ministry of Health, Uganda), Vicki Marsh and Tabitha Mwangi (Kenya Medical Research Institute), Dejan Zurovac (MSF-France) and Sporton Kwanya Agak, (KEDHAR). Deborah Balk and Gre-

gory Yetman are also thanked for access and discussion regarding urban rural classification globally. We also acknowledge the help provided by Greg Fegan for statistical advice, Abdisalan M Noor for help with the spatial mapping and Kevin Marsh for comments on an earlier draft. This paper is published with the permission of the director of KEMRI.

References

- Balk, D., Pozzani, F., Yetman, G., Nelson, A., Diechmann, U., 2004. Methodologies to Improve Global Population Estimates in Urban and Rural Areas. CIESIN, Columbia University, New York, USA.
- Benasseni, R., Gazin, P., Carnevale, P., Baudon, D., 1987. Le paludisme urbain à Bobo-Dioulasso. 1. Étude de la morbidité palustre. *Cah. O.R.S.T.O.M. Ser. Entomol. Med. Parasitol.* 25, 165–170.
- Björkman, A., Hedman, P., Brohult, J., Willcox, M., Diamant, I., Pehrsson, P.O., Rombo, L., Bengtsson, E., 1985. Different malaria control activities in an area of Liberia—effects on malarious parameters. *Ann. Trop. Med. Parasitol.* 79, 239–246.
- Bureau of Statistics, 2002. Serikali wa Jamhuri wa Tanzania. Sensa 2002 Takwimu. Dar es Salaam, United Republic of Tanzania.
- Central Bureau of Statistics, 1989. Kenya population census—1989, vol. 1. Central Bureau of Statistics, Ministry of Economic Planning and Development, Government of Kenya, Nairobi, 1989.
- CIESIN/CIAT, 2004. Gridded Population of the World (GPW), version 3, Palisades, New York CIESIN, Columbia University, Centre for International Earth Science Information Network (CIESIN), Columbia University; Centro Internacional de Agricultura Tropical (CIAT) (<http://beta.sedac.ciesin.columbia.edu/gpw>).
- CIESIN/IPFRI/CIAT, 2004. Global Rural-Urban Mapping Project (GRUMP): Gridded Population of the World, version 3, with urban reallocation (GPW-UR). Palisades, New York CIESIN, Columbia University, Centre for International Earth Science Information Network (CIESIN), Columbia University; International Food Policy Research Institute (IPFRI), the World Bank; and Centro Internacional de Agricultura Tropical (CIAT).
- Coene, J., 1993. Malaria in urban and rural Kinshasa: the entomological input. *Med. Vet. Entomol.* 7, 127–137.
- Craig, M.H., Snow, R.W., le Sueur, D., 1999. A climate-based distribution model of malaria transmission in sub-Saharan Africa. *Parasitol. Today* 15, 105–111.
- Directorate of Overseas Surveys, 1971. 1:50,000 topographical maps. Series Y731. Surveys of Kenya, Nairobi, Kenya.
- El Sayed, B.B., Arnot, D.E., Mukhtar, M.M., Baraka, O.Z., Dafalla, A.A., Elnaiem, D.E.A., Nugud, A.H.D., 2000. A study of the urban malaria transmission problem in Khartoum. *Acta Trop.* 75, 163–171.
- Fondjo, E., Robert, V., Le Goff, G., Toto, J.C., Carnevale, P., 1992. Le paludisme urbain à Yaoundé (Cameroun). 2. Étude entomologique dans deux quartiers peu urbanisés. *Bull. Soc. Pathol. Exot.* 85, 57–63.

- Gardiner, C., Biggar, R.J., Collins, W.E., Nkrumah, F.N., 1984. Malaria in urban and rural areas of southern Ghana: a survey of parasitaemia, antibodies, and antimalarial practices. *Bull. World Health Organ.* 62, 607–613.
- GDE Systems, 1995. Populated places: GeoName Digital Gazetteer v.1 (CD-ROM), Geographic Information Department (MZ 1211-M), GDE Systems Inc., P.O. Box 509009, San Diego, California (<http://www.gdesystems.com/>).
- Guyatt, H.L., Noor, A.M., Ochola, S.A., Snow, R.W., 2004. Use of intermittent presumptive treatment and insecticide treated bed nets by pregnant women in four Kenyan districts. *Trop. Med. Int. Health* 9, 255–261.
- Harpham, T., 1997. Urbanisation and health in transition. *Lancet* 349 (Suppl. 3), 11–13.
- Hay, S.I., 2000. An overview of remote sensing and geodesy for epidemiology and public health application. *Adv. Parasitol.* 47, 1–35.
- Hay, S.I., Rogers, D.J., Toomer, J.F., Snow, R.W., 2000. Annual *Plasmodium falciparum* entomological inoculation rates (EIR) across Africa: literature survey, Internet access and review. *Trans. R. Soc. Trop. Med. Hyg.* 94, 113–127.
- Hay, S.I., Noor, A.M., Nelson, A., Tatem, A.J., in press. The influence of large-scale population maps on the accuracy of burden of disease estimates. *Am. J. Epidemiol.*
- Holtz, T.H., Marum, L.H., Mkandala, C., Chizani, N., Roberts, J.M., Macheso, A., Parise, M.E., Kachur, S.P., 2002. Insecticide-treated bed net use, anaemia and malaria parasitaemia in Blantyre, Malawi. *Trop. Med. Int. Health* 7, 220–230.
- Lindsay, S.W., Campbell, H., Adiamah, J.H., Greenwood, A.M., Bangali, J.E., Greenwood, B.M., 1990. Malaria in a peri-urban area of The Gambia. *Ann. Trop. Med. Parasitol.* 84, 553–562.
- MEASURE Program, 2004. Demographic and Health Surveys for sub-Saharan Africa databases derived using STATCompiler, <http://www.measuredhs.org/>, accessed 30.4.2004.
- Mika, A.M., 1997. Three decades of Landsat instruments. *Photogramm. Eng. Remote Sens.* 63, 839–852.
- Modiano, D., Sirima, B.S., Sawadogo, A., Sanou, I., Pare, J., Konate, A., Pagnoni, F., 1999. Severe malaria in Burkina Faso: urban and rural environment. *Parasitology* 41, 251–254.
- Molyneux, C.S., Mung'ala-Odera, V., Harpham, T., Snow, R.W., 1999. Maternal responses to childhood fevers: a comparison of rural and urban residents in coastal Kenya. *Trop. Med. Int. Health* 4, 836–845.
- Mugisha, F., Arinatwe, J., 2003. Sleeping arrangements and mosquito net use among under fives: results from the Uganda Demographic and Health survey. *Malaria J.* 2, 40–50.
- Murray, C.J.L., Lopez, A.D., 1997. Mortality by cause for eight regions of the world: global burden of disease study. *Lancet* 349, 1269–1276.
- Ngimbi, N.P., Beckers, A., Wery, M., 1982. Aperçu de la situation épidémiologique du paludisme à Kinshasa (République du Zaïre). *Ann. Soc. Belg. Med. Trop.* 62, 121–137.
- Omumbo, J.A., Hay, S.I., Goetz, S.J., Snow, R.W., Rogers, D.J., 2002. Updating historical maps of malaria transmission intensity in East Africa using remote sensing. *Photogramm. Eng. Remote Sens.* 68, 161–166.
- Rasheed, S., Johnston, H., Dongier, P., Moreau, R., Lee, C., Lambert, J., Schaefer, C., 2000. Economic impact of febrile morbidity and use of permethrin-impregnated bed nets in a malarious area. I. Study of demographics, morbidity and household expenditures associated with febrile morbidity in the Republic of Benin. *Am. J. Trop. Med. Hyg.* 62, 173–180.
- Robert, V., Gazin, P., Ouédraogo, V., Carnevale, P., 1986. Le paludisme urbain à Bobo-Dioulasso (Burkina Faso). 1. Étude entomologique de la transmission. *Cah. O.R.S.T.O.M. Ser. Entomol. Med. Parasit.* 24, 121–128.
- Robert, V., McKintyre, K., Keating, S., Trape, J.F., Duchemin, J.B., Wilson, M., Beier, J.C., 2003. Malaria transmission in urban sub-Saharan Africa. *Am. J. Trop. Med. Hyg.* 68, 169–176.
- Rogers, D.J., Randolph, S.E., Snow, R.W., Hay, S.I., 2002. Satellite imagery in the study and forecast of malaria. *Nature* 415, 710–715.
- Rossi, P., Belli, A., Mancini, L., Sabatinelli, G., 1986. Enquête entomologique longitudinale sur la transmission du paludisme à Ouagadougou (Burkina Faso). *Parasitology* 28, 1–15.
- SALB, 2004. Second Administrative Level Boundaries (SALB) Project overviews: concepts, progress and future. URL: http://www3.who.int/whosis/gis/salb/salb_po.htm accessed 29th April 2004.
- Schneider, A., Friedl, M.A., McIver, D.K., Woodcock, C.E., 2003. Mapping urban areas by fusing multiple sources of coarse resolution remotely sensed data. *Photogramm. Eng. Remote Sens.* 69, 1377–1386.
- Service, M.W., 1997. Mosquito (Diptera: Culicidae) dispersal—the long and short of it. *J. Med. Entomol.* 34, 579–588.
- Snow, R.W., Marsh, K., 1995. Will reducing *Plasmodium falciparum* transmission alter mortality among African children? *Parasitol. Today* 11, 188–190.
- Snow, R.W., Omumbo, J.A., Lowe, B., Molyneux, S.M., Obiero, J.O., Palmer, A., Weber, M.W., Pinder, M., Nahlen, B., Obonyo, C., Newbold, C., Gupta, S., Marsh, K., 1997. Relation between severe malaria morbidity in children and level of *Plasmodium falciparum* transmission in Africa. *Lancet* 349, 1650–1654.
- Snow, R.W., Gouws, E., Omumbo, J.A., Rapuoda, B.A., Craig, M.H., Tanser, F.C., le Sueur, D., Ouma, J., 1998. Models to predict the intensity of *Plasmodium falciparum* transmission: applications to the burden of disease in Kenya. *Trans. R. Soc. Trop. Med. Hyg.* 92, 601–606.
- Snow, R.W., Craig, M.H., Deichmann, U., Marsh, K., 1999. Estimating mortality, morbidity and disability due to malaria among Africa's non-pregnant population. *Bull. World Health Organ.* 77, 624–640.
- Snow, R.W., Marsh, K., 2002. The consequences of reducing *Plasmodium falciparum* transmission in Africa. *Adv. Parasitol.* 52, 235–264.
- Snow, R.W., Omumbo, J.A., 2004. Malaria mortality in sub-Saharan Africa. In: Jamison, D., Feacham, R., Makgoba, W., Hofman, K., Rogo, K. (Eds.), *Disease and Mortality in sub-Saharan Africa*. Oxford University Press and World Bank.
- Statistics Department Uganda, 1995. The 1991 Population and Housing Census: Final Results (Main Release). Statistics Department, Ministry of Finance and Economic Planning, The Republic of Uganda, Entebbe.

- Sutton, P., Roberts, D., Elvidge, C., Baugh, K., 2001. Census from Heaven: an estimate of the global human population using nighttime satellite imagery. *Int. J. Remote Sens.* 22, 3061–3076.
- Tatem, A.J., Hay, S.I., in press. Measuring urbanization pattern and extent for malaria research: a review of remote sensing approaches. *J. Urban Health*.
- Tatem, A.J., Noor, A.M., Hay, S.I., in press. Defining approaches to settlement mapping for public health management in Kenya using medium spatial resolution satellite imagery. *Remote Sens. Environ.*
- Trape, J.F., Zoulani, A., 1987. Malaria and urbanization in Central Africa: the example of Brazzaville Part II Results of entomological surveys and epidemiological analysis. *Trans. R. Soc. Trop. Med. Hyg.* 81 (Suppl 2), 34–42.
- Trape, J.F., Lefebvre-Zante, E., Legros, F., Ndiaye, G., Bouganali, H., Druille, P., Salem, G., 1992. Vector density gradients and the epidemiology of urban malaria in Dakar, Senegal. *Am. J. Trop. Med. Hyg.* 47, 181–189.
- United Nations, 2002. *World Urbanisation Prospects, 2002 Revision*. New York, United Nations.
- Vlahov, D., Galea, S., 2002. Urbanization, urbanicity and health. *J. Urban Health* 79, 1–12.
- Walker, N., Schwartlander, B., Bryce, J., 2002. Meeting international goals in child survival and HIV/AIDS. *Lancet* 360, 284–289.
- Watts, T.E., Wray, J.R., Ng'andu, H., Draper, C.C., 1990. Malaria in and urban and rural area of Zambia. *Trans. R. Soc. Trop. Med. Hyg.* 84, 196–200.
- Williams, B.G., Gouws, E., Boschi-Pinto, C., Bryce, J., Dye, C., 2002. Estimates of world-wide distribution of child deaths from acute respiratory infections. *Lancet Infect. Dis.* 2, 25–32.
- World Health Organization, 2002. *World Health Report 2002. Reducing Risks, Promoting Healthy Life*. Geneva: World Health Organization.
- World Resources Institute, 1995. *Topographic data: Africa Data Sampler (CD-ROM)*. World Resources Institute, Washington, DC, USA.
- Yohannes, M., Petros, B., 1996. Urban malaria in Nazareth Ethiopia: parasitological studies. *Ethiopian Med. J.* 34, 83–90.

CHAPTER 6:

Updating historical maps of malaria transmission intensity in East Africa using remote sensing

JA Omumbo, SI Hay, SJ Goetz, RW Snow, DJ Rogers

Photogrammetric Engineering and Remote Sensing
(2002), 68: 161-166

Updating Historical Maps of Malaria Transmission Intensity in East Africa Using Remote Sensing

J.A. Omumbo, S.I. Hay, S.J. Goetz, R.W. Snow, and D.J. Rogers

Abstract

Remotely sensed imagery has been used to update and improve the spatial resolution of malaria transmission intensity maps in Tanzania, Uganda, and Kenya. Discriminant analysis achieved statistically robust agreements between historical maps of the intensity of malaria transmission and predictions based on multitemporal meteorological satellite sensor data processed using temporal Fourier analysis. The study identified land surface temperature as the best predictor of transmission intensity. Rainfall and moisture availability as inferred by cold cloud duration (CCD) and the normalized difference vegetation index (NDVI), respectively, were identified as secondary predictors of transmission intensity. Information on altitude derived from a digital elevation model significantly improved the predictions. "Malaria-free" areas were predicted with an accuracy of 96 percent while areas where transmission occurs only near water, moderate malaria areas, and intense malaria transmission areas were predicted with accuracies of 90 percent, 72 percent, and 87 percent, respectively. The importance of such maps for rationalizing malaria control is discussed, as is the potential contribution of the next generation of satellite sensors to these mapping efforts.

Introduction

Human malaria is caused by the parasites *Plasmodium falciparum*, *P. vivax*, *P. ovale*, and *P. malariae*. The Plasmodium life cycle involves an asexual stage in human hosts and a sexual stage in mosquitoes of the genus *Anopheles*. *Anopheles gambiae sensu stricto* and *An. funestus* are the most widely distributed and important malaria vectors in sub-Saharan Africa (SSA). Ninety percent of the global burden of malaria, predominantly due to *P. falciparum*, is borne by the population of SSA

J.A. Omumbo and S.I. Hay are with the Trypanosomiasis and Land-Use in Africa (TALA) Research Group, Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, United Kingdom and with the Kenya Medical Research Institute/Wellcome Trust collaborative programme, P.O. Box 43640, Nairobi, Kenya (judith.omumbo@zoology.oxford.ac.uk).

S.J. Goetz is with the Department of Geography, University of Maryland, College Park, Maryland, MD 20742-8225.

R.W. Snow is with the Kenya Medical Research Institute/Wellcome Trust collaborative programme, P.O. Box 43640, Nairobi, Kenya and with the Centre for Tropical Medicine, University of Oxford, John Radcliffe Hospital, Headington, Oxford, OX3 9DU, United Kingdom.

D.J. Rogers is with the Trypanosomiasis and Land-Use in Africa (TALA) Research Group, Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, United Kingdom.

(WHO, 1999), resulting in approximately 1 million deaths due to *P. falciparum* each year (Snow *et al.*, 1999). The clinical spectrum of *P. falciparum* infection ranges in severity from mild, often self-limiting, fever, chills, and joint pains to a life-threatening illness.

Malaria transmission is often quantified through mathematical models. The basic reproductive number of a disease (R_0) is derived from a generic infectious disease model and is defined as the average number of successful offspring that a parasite is intrinsically capable of producing in a completely susceptible population (Macdonald, 1957). The vectorial capacity (VC), derived from R_0 , reflects the mean number of probable inoculations transmitted from one case of malaria per unit time (Garrett-Jones, 1964) and it is expressed as

$$VC = \frac{ma^2p^n}{-\log_e p} \quad (1)$$

where m is the relative density of female anophelines, a is the probability that the mosquito will take a human blood meal during a particular day, and p^n is the proportion of vectors surviving the parasite's incubation period (i.e., p is the daily probability of vector survival and n is the duration of parasite sexual development within the mosquito or sporogony). All of these factors, with the possible exception of a , are affected by climate and hence environmental proxy information derived from satellite sensors (Hay *et al.*, 1996; Hay *et al.*, 1999) can be of use in predicting malaria transmission intensities (Hay *et al.*, 1997; Rogers *et al.*, 2002).

Optimum temperatures for *P. falciparum* (sporogony) are between 25 and 30 °C. Below 16 °C and above 35 °C sporogony ceases (Detinova, 1962), and thermal death of mosquitoes occurs at 40 to 42 °C (Dutta *et al.*, 1978). Breeding site availability is associated with rainfall and increased mosquito survivorship with increasing humidity (Gill, 1920; Dutta *et al.*, 1978). Non-climatic factors, including proximity to permanent water bodies, urbanization, population distribution, agricultural practices, and other human factors, further modify transmission (Gilles, 1993). The spatial heterogeneity of these factors contributes substantially to variations in transmission intensity. For example, one marker of malaria transmission intensity, the annual entomological inoculation rate (EIR) defined as the average number of infective mosquito bites received per

Photogrammetric Engineering & Remote Sensing
Vol. 68, No. 2, February 2002, pp. 161–166.

0099-1112/02/6800-161\$3.00/0

© 2002 American Society for Photogrammetry
and Remote Sensing

person in a year) ranges from 0 to ~1000 across sub-Saharan Africa (Hay *et al.*, 2000b). This heterogeneity is best captured by maps which themselves should form an important tool for rationalizing malaria control (Snow *et al.*, 1996).

Several authors have attempted to describe the global distribution of malaria using expert opinion and climate data. Early maps were based on the position of summer isotherms (Boyd, 1930; Lysenko *et al.*, 1969); the known latitudinal extent of the disease (Gill, 1920); combinations of temperature, elevation, and rainfall (Dutta *et al.*, 1978); and interpolation of limited malariometric data (Clyde, 1967; Onori, 1967). In recent years there has been a renaissance in mapping malaria in Africa (Snow *et al.*, 1996; MARA/ARMA collaboration, 1998; Hay *et al.*, 2000a; Rogers *et al.*, 2002). This has been facilitated by advancements in geographic information systems (GIS) and increased availability of public domain, remotely sensed (RS) satellite data. RS surrogates of climate have been used to develop high spatial resolution maps of transmission intensity and epidemic potential (Lindsay *et al.*, 1996; Snow *et al.*, 1998; Craig *et al.*, 1999; Thomson *et al.*, 1999; Kleinschmidt *et al.*, 2000). In this paper we discuss the utility of remote sensing for re-visiting historical efforts to describe malaria distribution in East Africa at high spatial resolution.

Materials and Methods

Historical Malaria Distribution Maps

At the time of independence, concerted efforts were made to provide atlases for use in administration and development planning in Tanzania, Uganda, and Kenya (Government of Tanganyika, 1956; Government of Uganda, 1962; Government of Kenya, 1970). The atlases included maps of synoptic rainfall and temperature, health facilities, agriculture, natural resources, and malaria. The rules used to categorize malaria transmission differed slightly between the three countries, but all were derived from expert opinion on climate and local malariology.

Uganda's map (Government of Uganda, 1962) describes four malaria epidemiological zones as follows: (1) *normally malaria free*: highland areas above 1500 meters where transmission is limited by low temperatures, (2) *malarious near water*: arid sparsely populated areas where transmission occurs only around permanent water bodies, (3) *moderately malarious*: areas experiencing warm moist climate with evenly distributed moderate rainfall where malaria transmission is seasonal, and (4) *intensely malarious*: areas below 1500 meters where transmission occurs year round, increasing after periods of heavy rainfall. In both Kenya (Government of Kenya, 1970) and Tanzania (Government of Tanganyika, 1956), five zones are defined according to the duration of the transmission season as follows: (i) *Transmission for less than three months of a year*; (ii) *transmission for three to six months*; (iii) *transmission for more than six months of a year*; (iv) *malarious near water*; and (v) *malaria free*: highland areas, although a precise altitude limit is not indicated in the definition.

To provide uniform categories of malaria transmission for analysis, slight reclassifications were necessary so that maps of malaria transmission intensity in Kenya and Tanzania could be directly compared with those for Uganda. This was justified by recent work suggesting that the "intensity" of transmission is strongly correlated with the length of the malaria transmission season (Tanser, 2000). Areas experiencing transmission for more than 6 months were reclassified as "intensely malarious," areas with transmission for less than 3 months and for 3 to 6 months were combined and reclassified as "moderately malarious," and areas defined as malaria-free and malarious near water remained unchanged. The historical maps were digitized using the GIS MapInfo (MapInfo Corporation, 1985–2000)

and the vector map was rasterized using IDRISI version 2.1 (Clarke University, 1997).

The Pathfinder Advanced Very High Resolution Radiometer Land (PAL) Data

This study uses data derived from the National Oceanic and Atmospheric Administration (NOAA) Advanced Very High Resolution Radiometer (AVHRR) sensor on board the NOAA series of polar-orbiting satellites. Public domain 8- by 8-km spatial resolution satellite data were obtained through the Pathfinder AVHRR Land (PAL) program (James *et al.*, 1994; URL: <http://daac.gsfc.nasa.gov/data/dataset/AVHRR/index.html>). These data were processed using standard quality control and cloud reducing procedures outlined in Hay *et al.* (1999).

Normalized Difference Vegetation Index

Of the many spectral vegetation indices available, the Normalized Difference Vegetation Index (NDVI) (Tucker, 1979) has found most application in epidemiological studies (Hay *et al.*, 1998; Patz *et al.*, 1998; Snow *et al.*, 1998; Kleinschmidt *et al.*, 2000). As a robust indicator of photosynthetic activity (Tucker *et al.*, 1986; Myneni *et al.*, 1995), it is thought to provide information on the response of a landscape to precipitation (Nicholson *et al.*, 1990) and is hence of use in malaria studies. The NDVI is derived from AVHRR channels 1 (visible) and 2 (near-infrared) as follows:

$$NDVI = \frac{Ch2 - Ch1}{Ch2 + Ch1} \quad (2)$$

Mid-Infrared (MIR) Reflectance

AVHRR channel 3 (mid-infrared) has been shown to be sensitive to both reflected and emitted radiation, although the interpretation of this signal is less well understood than the other channels (Boyd *et al.*, 1998). MIR suffers less from atmospheric attenuation than do channels 4 and 5 of the AVHRR (Cracknell, 1997) and has found limited use for vegetation mapping.

Land Surface and Air Temperature Indices

The land surface brightness temperature (LST) was calculated using the "split-window" equation of Price (1984) as follows:

$$LST = T_{Ch4} + 3.33(T_{Ch4} - T_{Ch5}) \quad (3)$$

This has been found to be a relatively accurate proxy of ground-based meteorological measurements of air temperature with a root-mean-square error of ± 4 °C over continental Africa (Hay *et al.*, 2000a). A further air temperature variable (T_{air}) was estimated using a contextual combination of vegetation indices and LST estimates reviewed by Goetz *et al.* (2000).

Altitude

A 1- by 1-km spatial resolution digital elevation model (DEM) for Africa was obtained from the United States Geological Survey (USGS) (Gesch *et al.*, 2000; URL: <http://edcdaac.usgs.gov/gtopo30/README.html>). Elevation is recorded in meters above sea level with a root-mean-square error reported at ± 100 meters.

Rainfall

A proxy for rainfall can be derived from the European Organisation for the Exploitation of Meteorological Satellites (EUMESAT) Meteosat series. The exact threshold temperature associated with rain-bearing clouds and the quantity of rain they deposit varies temporally and spatially so that it must be established empirically. This has been done for Africa by the Tropical Applications of Meteorology using Satellite and other data (TAMSAT) programme of the Department of Meteorology, University of Reading (URL: <http://www.met.reading.ac.uk/famsat>). The northern and southern extent of the calibration

exercise represents the most northern and southern limits of the Inter Tropical Convergence Zone (ITCZ) where convective atmospheric processes dominate (Dugdale *et al.*, 1995). There are set thresholds of -50°C in the summer and -60°C in the winter, used for areas north of the ITCZ, and -40°C throughout the year for regions south. These results were used by the FAO-ARTEMIS project to generate monthly cold cloud duration (CCD) images, where each pixel represents the number of hours during which cold cloud-top temperatures were below these thresholds during a 10-day compositing period. The CCD has been found to have a root-mean-square error of ± 38 mm when compared with meteorological station recordings of rainfall across continental Africa (Hay *et al.*, 1999).

Data Preparation and Analysis

The AVHRR data are supplied in the equal-area Goode's interrupted homolosine projection. All images were reprojected to a Latlong coordinate system. The area covering East Africa (29°E to 42°E and 11.8°S to 5.5°N) was subset for further analysis. Temporal Fourier analysis (TFA) (Rogers *et al.*, 1996; Rogers, 2000) was applied to the five-predictor variables, NDVI, MIR, LST, T_{air} , and CCD. TFA summarizes the time series data using a series of sine and cosine functions. The periodic part of the time series is isolated and described by its frequency (annual, biannual, or triannual), amplitude (amp), and phase (phs). Ten Fourier images (mean, maximum, minimum, and variance; amplitudes of annual, biannual, and triannual cycles; and phases of annual, biannual, and triannual cycles) were produced for each predictor variable.

A randomly sampled "training data" set was selected from the historical distribution maps. Image pixels along the boundaries of malaria categories were excluded from the sampling frame because such areas have a high probability of misclassification of malaria intensity. Areas covered by large permanent water bodies were also excluded. Nine-hundred seventy-five training data pixels were selected; three-hundred for each of the categories "malaria near water," "moderate malaria," and "intense malaria." The "malaria-free" category covered a much smaller area, so only 75 training data pixels were selected. Corresponding values of the predictor variables were extracted for each of the 975 training data pixels.

The *kappa* statistic (Cohen, 1960) was selected as a measure of agreement. Values of *kappa* less than 0.4 indicate poor agreement, values between 0.4 and 0.75 suggest good agreement, and values above 0.75, excellent agreement (Landis *et al.*, 1977). Using discriminant analysis (DA), predictor variables were selected in a forward step-wise fashion by iteratively including the variable that causes the maximum increase in the *kappa* statistic (Cohen, 1960) compared with the other variables in each round of analysis until ten variables were selected. As the environmental characteristics within the different malaria regions are dissimilar, DA was carried out with the assumption of different covariance matrices between the malaria groups. It was also assumed that the *a priori* probabilities of group membership were equal. Based on the discriminant function, each pixel was assigned a malaria category. A map of the *a posteriori* probabilities was then plotted for East Africa and compared with the historical map.

Results

Pixels contaminated with water and those with predictor variable values greater than 6 standard deviations from means were excluded ($n = 4$). Nine-hundred seventy-one pixels were used in the analysis. Temperature variables were strongly correlated (Pearson correlation coefficients of 0.99 for mean Price LST and mean MIR reflectance; 0.81 for mean LST and mean T_{air} ; and 0.77 for mean MIR reflectance and mean T_{air} (Table 1)). A strong negative correlation (-0.67) was found between altitude and T_{air} . Moisture availability surrogates were not as

TABLE 1. CORRELATION COEFFICIENTS OF INDEPENDENT VARIABLES

	Altitude	Mean MIR reflectance	Mean Price LST	Mean NDVI	NDVI amp1
Mean MIR reflectance	-0.50				
Mean Price LST	-0.59	0.99			
Mean T_{air}	-0.67	0.77	0.81		
Mean CCD				0.64	
CCD amp1					0.71

strongly correlated, the highest correlations being CCD amp1 and NDVI amp1 (0.71) and NDVI mean and CCD mean (0.64). Non-significant correlations were noted between malaria intensity class and the NDVI and CDD variables (coefficients less than 0.5).

Plots of the predictor variables by malaria intensity class showed considerable overlap in environmental conditions between the four malaria categories except with regard to temperature (mean Price LST) and altitude. Average and ranges of mean LST distinguish well between the "malaria-free" and "malaria near water" categories compared with other malaria classes (Figure 1). The altitude differences between these categories are also shown in Figure 1.

In the discriminant analysis overall agreement between training data (observed) and predicted pixels was 89.9 percent (Cohen's *kappa* = 0.775; Klecka's *tau* = 0.773), with greatest agreement in the categories "malaria-free" and "malaria near water" (Table 2). The model over predicted "malaria-free" areas (false positive rate = 26.3 percent) and under predicted moderate malaria (false negative rate = 27.7 percent (Table 3)).

Table 4 lists the top ten RS variables selected by the DA, and minimum, maximum, and mean values for these. Six of these are temperature variables (LST, DEM, T_{air} , and MIR reflectance) while CCD variables appear three times. The DA was repeated omitting some of the strongly correlated variables (T_{air} and/or mean MIR reflectance). In each case the mean Price LST was most discriminating. Considerably higher misclassification rates were obtained when T_{air} or MIR reflectance variables were omitted from the discriminant analysis. The DA was also done using only the first five discriminating variables; however, the resulting *a posteriori* predictions were poorer (*kappa* = 0.706). Including all ten variables produced the best fit, with an overall *kappa* of 0.775.

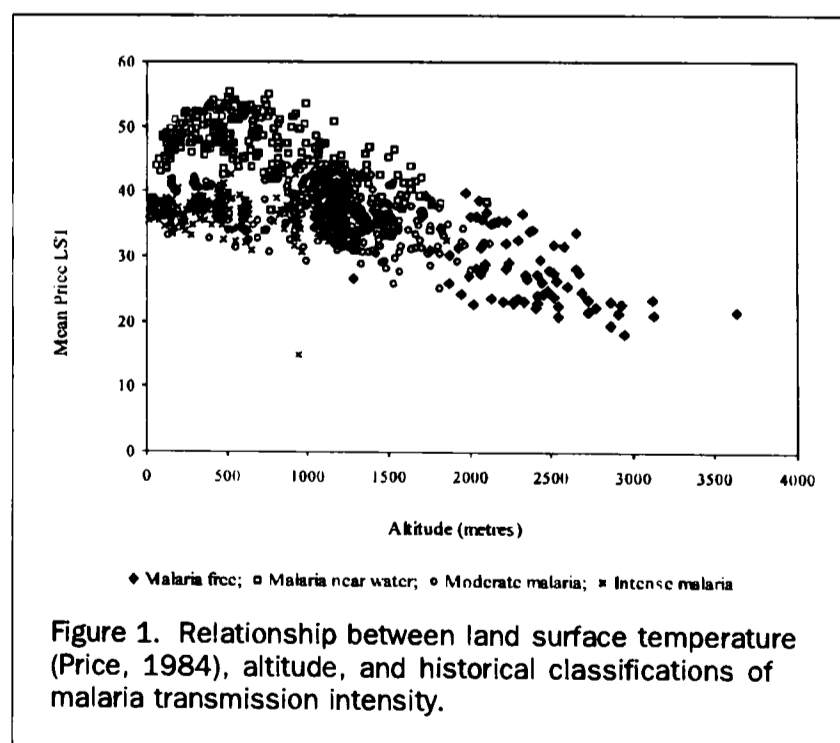


Figure 1. Relationship between land surface temperature (Price, 1984), altitude, and historical classifications of malaria transmission intensity.

TABLE 2. CLASSIFICATION MATRIX FOR TRAINING DATA AND PREDICTED PIXELS

	Malaria intensity	Predicted				Total
		Malaria free	Malaria near water	Moderate	Intense	
Observed	Malaria free	70	0	2	1	73
	Malaria near water	9	270	12	9	300
	Moderate malaria	15	15	217	53	300
	Intense malaria	1	2	37	258	298
	Total	95	287	268	321	971

Cohen's kappa: 0.775 (Cohen, 1960), Klecka's tau: 0.773. Also see Table 3.

Figure 2a shows the rasterized historical transmission map with major water bodies overlaid. (The national boundaries used are as they were at the time the historical maps were printed, and changes have since occurred along Kenya's north-western border with Sudan.) The Tana River in eastern Kenya drains into the Indian Ocean through an arid, low altitude (less than 400 m above sea level) region. According to the historical map, intense malaria transmission occurs within a narrow band (less than 5 km) along the riverbank. Areas of differing malaria transmission intensity are color coded as follows: malaria free (white), malaria near water (light grey), moderate malaria (dark grey), and intense malaria (black). The "malaria-free" areas on the map have an average altitude of 2335 meters (range 1282 to 3634 meters) while the "malaria near water" areas average 680 meters (range 14 to 1800 meters).

Discussion

The DA was carried out using different numbers and combinations of predictor variables. The prediction resulting in the highest kappa statistic (Figure 2b) is compared with the historical map (Figure 2a). Predictions were best at the extremes of transmission and in areas that experience very characteristic environmental or climatic conditions. For example, arid areas (transmission occurring only in proximity to permanent water bodies) experience year-round high temperatures, and the thermal indices therefore identify these areas well. Similarly, "malaria-free" areas are typically at high altitude and are characterized by much lower year-round temperatures (Figure 1). In moderate and intense transmission areas, such distinction of environmental characteristics is not as pronounced. Although the main areas of malaria transmission intensity are well predicted, the boundaries within moderate transmission areas are much less distinct, such as the Indian Ocean coastline on Kenya's southeastern border, parts of western Uganda, and large areas of central Tanzania. Incorporation of an environmental zone classification variable into the DA as an additional predictor variable may allow more of the environmental differences between moderate and intense transmission classes to be isolated (Brooker *et al.*, 2002).

It is important to note that the historical transmission maps were based largely on expert opinion and were intended to provide only a broad representation of available information on the

distribution and intensity of malaria. Criteria used for classification varied between the three countries. Tanzania's map was based predominantly on the expert opinion of a malaria epidemiologist and, on this basis, areas along rivers are classed as intense transmission areas (Figure 2c) as was done with the Tana River in Kenya. Most misclassified pixels are found in Tanzania mainly along rivers and on the margins of transmission categories (Figure 2c), suggesting that additional factors, not identified by climatic predictors, determine transmission intensity. The masking of a row of 8- by 8-km pixels along the boundaries of the transmission categories excluded the Tana River banks from the sampling frame, resulting in the omission of the river on the predictive map. It is possible that several other areas on the historical map where transmission occurs alongside rivers (particularly in Tanzania) should have been left out for similar reasons. Uganda's map is influenced largely by expert knowledge of altitude and rainfall distribution and their relationship to malaria transmission. Regions of similar climate are grouped together; as for example, areas experiencing "Lake Victoria weather" are described. Malaria intensity regions were classified on the basis of such similarity of climatic conditions. Kenya's map was designed to identify persons "at risk" according to ethnic groupings, and the malaria map has been influenced by the pattern of these settlements.

There are several advantages of the approach used here over previous attempts at modeling. First, RS data provide a continuous surface, thus avoiding the use of interpolated climate data with its inherent interpretation biases; also, high temporal resolution climate data have been used. Fourier processed imagery, by summarizing climatic data according to its natural biological cycles, better relates to the biological processes involved in malaria transmission. Rainfall patterns in East Africa show a bimodal distribution (two rainy seasons in a year) in most areas, and this may be related to the selection of the timing of CCD biannual cycles and NDVI by the discriminant analysis. Although temperatures in arid areas tend to be high

TABLE 4. TEN MOST DISCRIMINATING RS VARIABLES BY RANK

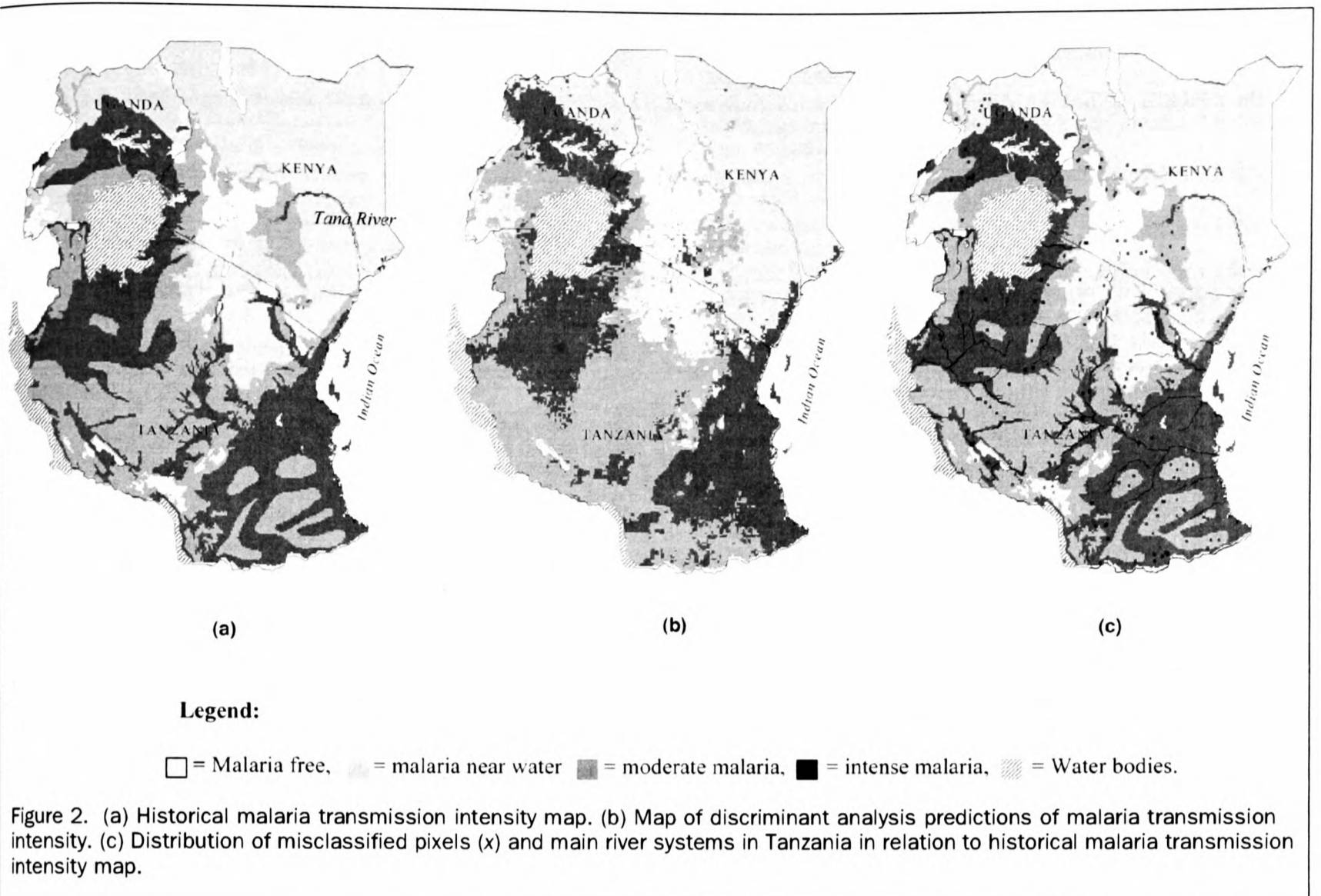
Rank	Predictor variable	kappa statistic
1	Mean Price LST	0.464
2	CCD variance	0.562
3	CCD phase3	0.639
4	DEM	0.691
5	CCD phase2	0.712
6	LST phase1	0.735
7	T _{air} phase1	0.752
8	MIR amp3	0.763
9	NDVI amp2	0.779
10	T _{air} amp3	0.782

LST = Price land surface temperature, CCD = cold cloud duration, T_{air} = air temperature, MIR = mid-infrared, amp2 = amplitude of biannual cycle, amp3 = amplitude of triannual cycle, phase1 = timing of annual cycle, phase2 = timing of biannual cycle, phase3 = timing of triannual cycle.

TABLE 3. AGREEMENT BETWEEN HISTORICAL AND PREDICTED MAP OF MALARIA TRANSMISSION INTENSITY

Malaria intensity	Agreement (%)	False positive rates (%) ^a	False negative rates (%) ^b
Malaria free	95.9	26.3	4.1
Malaria near water	90.0	5.9	10.0
Moderate malaria	72.3	19.0	27.7
Intense malaria	86.6	19.6	13.4
All categories	83.9	16.1	16.1

a. Omission error; b. commission error.



all year round, they usually peak annually, and the DA picks up the timing of the annual cycles of LST and T_{air} in these areas. Furthermore, aside from providing good definitions of areas that do not support stable transmission, stable transmission areas are defined more clearly than models that have used other statistical procedures (Craig *et al.*, 1999).

There is additional scope for the refinement of the climate-based model in several areas; the rationale for this would depend on the level of detail of the information required for use in decision support for malaria control purposes. The inclusion of additional databases such as rivers and other smaller permanent water bodies and the use of higher spatial resolution RS data may improve prediction in the moderate and intense transmission areas in Tanzania and Uganda. Historical maps, however, are likely to be inaccurate at higher resolution and, perhaps, appropriate empirical data could be used to improve the accuracy of predictions. The aim of this study was to demonstrate the utility of RS data in predicting malaria transmission intensity, and the challenge of future research will be to use empirical data to assess the accuracy of these predictions.

Acknowledgments

Information used in this study includes data produced through funding from the Earth Observing System Pathfinder Program of NASA's Mission to Planet Earth in cooperation with National Oceanic and Atmospheric Administration. The data were provided by the Earth Observing Data and Information System (EODIS), Distributed Active Archive Center at Goddard Space Flight Center which archives, manages, and distributes this dataset. Thanks are also due to Fred Snidjers of the Food and

Agriculture Organisation's (FAO) Africa Real Time Environmental Monitoring using Imaging Satellites (ARTEMIS) project for permission to use Meteosat CCD data. We also thank Jennifer Small for help in implementing methods developed at the University of Maryland Department of Geography for the derivation of air temperature estimates from AVHRR data. J.A. Omumbo is supported by a Wellcome Trust prize studentship (#060063), S.I. Hay is supported on a Wellcome Trust Advanced Training Fellowship (#056642), and R.W. Snow is on a Wellcome Trust Senior Fellowship (#033340). This paper is published with the permission of the Director, KEMRI.

References

- Boyd, D., and P. Curran, 1998. Using remote sensing to reduce uncertainties in the global carbon budget: The potential of radiation acquired in the middle infrared wavelengths, *Remote Sensing Reviews*, 16:293-327.
- Boyd, M.F., 1930. *An Introduction to Malariaology*, Harvard University Press, Cambridge, Massachusetts, 1438 p.
- Brooker, S., S.I. Hay, and D.A.P. Bundy, 2002. Tools from ecology to evaluate infection risk models, *Trends in Parasitology*, in press.
- Clarke University, 1997. *IDRISI Version 2.0 for Windows*, Clarke University, Graduate School of Geography, Worcester, Massachusetts, <http://www.idrisi.clarku.edu>.
- Clyde, D.F., 1967. *Malaria in Tanzania*, Oxford University Press, London, New York.
- Cohen, J., 1960. A co-efficient of agreement for nominal scale, *Educational and Psychological Measurements*, 20:37-46.
- Cracknell, A., 1997. Channel 3: The neglected channel, *The Advanced Very High Resolution Radiometer*, Taylor & Francis, London, pp. 321-342.

- Craig, M.H., R.W. Snow, and D. Le Sueur, 1999. A climate-based distribution model of malaria transmission in sub-Saharan Africa, *Parasitology Today*, 15(3):105–111.
- Detinova, T.S., 1962. *Age Grouping Methods in Diptera of Medical Importance, with Special Reference to Some Vectors of Malaria*, World Health Organization, Geneva, Switzerland, 210 p.
- Dugdale, G., V. McDougall, and V. Thorne, 1995. The TAMSAT method for estimating rainfall over Africa and its implications, RSS95: *Proceedings of the 21st Annual Conference of the Remote Sensing Society*, 11–14 September, Southampton, England, pp. 773–780.
- Dutta, H.M., and A.K. Dutt, 1978. Malaria ecology: a global perspective, *Social Science in Medicine Part D (Medical Geography)*, 12:69–84.
- Garrett-Jones, C., 1964. Prognosis for interruption of malaria transmission through assessment of the mosquito's vectorial capacity, *Nature*, 204:1173–1175.
- Gesch, D.B., and K.S. Larson, 2000. *Techniques for Development of Global 1-Kilometer Digital Elevation Models*, United States Geological Survey—EROS data centre, Sioux Falls, South Dakota, (<http://ededaac.usgs.gov/gtopo30/papers/gesch.txt.html>).
- Gill, C.A., 1920. The role of meteorology in malaria, *Indian Journal of Medical Research*, 8:633–693.
- Gilles, H., 1993. Epidemiology of malaria, *Bruce-Chwatt's essential malariology* (E. Arnold, editor), Oxford University Press., Inc., New York, N.Y., pp. 124–163.
- Goetz, S., S. Prince, and J. Small, 2000. Advances in satellite remote sensing of environmental variables for epidemiological applications, *Advances in Parasitology*, 47:289–307.
- Government of Kenya, 1970. *National Atlas of Kenya*, Survey of Kenya, Nairobi, Kenya, 44 p.
- Government of Tanganyika, 1956. *Atlas of Tanganyika, East Africa*, Government Press, Dar es Salaam, Tanganyika, 29 p.
- Government of Uganda, 1962. *Atlas of Uganda*, Department of Lands and Surveys, Kampala, Uganda, 83 p.
- Hay, S.I., and J.J. Lennon, 1999. Deriving meteorological variables across Africa for the study and control of vector-borne disease: A comparison of remote sensing and spatial interpolation of climate, *Tropical Medicine & International Health*, 4(1):58–71.
- Hay, S.I., J.A. Omumbo, M.H. Craig, and R.W. Snow, 2000a. Earth observation, geographic information systems and *Plasmodium falciparum* malaria in sub-Saharan Africa, *Advances in Parasitology*, 47:173–215.
- Hay, S.I., M.J. Packer, and D.J. Rogers, 1997. The impact of remote sensing on the study and control of invertebrate intermediate hosts and vectors for disease, *International Journal of Remote Sensing*, 18(14):2899–2930.
- Hay, S.I., D.J. Rogers, J.F. Toomer, and R.W. Snow, 2000b. Annual *Plasmodium falciparum* entomological inoculation rates (EIR) across Africa: Literature survey, internet access and review, *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 94:113–127.
- Hay, S.I., R.W. Snow, and D.J. Rogers, 1998. Predicting malaria seasons in Kenya using multi-temporal meteorological satellite sensor data, *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 92:12–20.
- Hay, S.I., C.J. Tucker, D.J. Rogers, and M.J. Packer, 1996. Remotely sensed surrogates of meteorological data for the study of the distribution and abundance of arthropod vectors of disease, *Annals of Tropical Medicine and Parasitology*, 90:1–19.
- James, M.E., and S.N.V. Kalluri, 1994. The Pathfinder AVHRR land data set—an improved coarse resolution data set for terrestrial monitoring, *International Journal of Remote Sensing*, 15: 3347–3363.
- Kleinschmidt, I., M. Bagayoko, G.P.Y. Clarke, M.H. Craig, and D. Le Sueur, 2000. A spatial statistical approach to malaria mapping, *International Journal of Epidemiology*, 29(2):355–361.
- Landis, J., and G. Koch, 1977. The measurement of observer agreement for categorical data, *Biometrics*, 33:159–174.
- Lindsay, S.W., and M.H. Birley, 1996. Climate change and malaria transmission, *Annals of Tropical Medicine and Parasitology*, 90(6):573–588.
- Lysenko, A.Y., and A.E. Beljaev, 1969. An analysis of the distribution of *Plasmodium ovale*, *Bulletin of the World Health Organization*, 40:383–394.
- Macdonald, G., 1957. *The Epidemiology and Control of Malaria*, Oxford University Press, London, 310 p.
- MapInfo Corporation, 1985–2000. *MapInfo Professional*, MapInfo Corporation, Troy, New York (software).
- MARA/ARMA collaboration, 1998. *Towards an Atlas of Malaria Transmission in Africa: First Technical Report of the MARA/ARMA*, South African Medical Research Council, IDRC, Canada and Wellcome Trust, UK, Durban, South Africa, 30 p.
- Myneni, R.B., S. Maggion, J. Jaquinta, J.L. Privette, N. Gobron, B. Pinty, D.S. Kimes, M.M. Verstraete, and D.L. Williams, 1995. Optical remote sensing of vegetation: Modeling, caveats and algorithms, *Remote Sensing of Environment*, 51:169–188.
- Nicholson, S.E., M.L. Davenport, and A.R. Malo, 1990. A comparison of vegetation response to rainfall in the Sahel and East Africa using Normalized Difference Vegetation Index from NOAA-AVHRR, *Climate Change*, 17:209–241.
- Onori, E., 1967. Distribution of *Plasmodium ovale* in the Eastern, Western and Northern regions of Uganda, *Bulletin of the World Health Organization*, 37(4):665–668.
- Patz, J.A., K. Strzepek, S. Lele, M. Hedden, S. Greene, B. Noden, S.I. Hay, L. Kalkstein, and J.C. Beier, 1998. Predicting key malaria transmission factors, biting and entomological inoculation rates, using modelled soil moisture in Kenya, *Tropical Medicine & International Health*, 3(10):818–827.
- Price, J., 1984. Land surface temperature measurements from the split window channels of the NOAA Advanced Very High Resolution Radiometer, *Journal of Geophysical Research*, 89:7231–7237.
- Rogers, D.J., 2000. Satellites, space, time and the African trypanosomiasis, *Advances in Parasitology*, 47:129–171.
- Rogers, D.J., S.I. Hay, and M.J. Packer, 1996. Predicting the distribution of tsetse-flies in West Africa using temporal Fourier processed meteorological satellite data, *Annals of Tropical Medicine and Parasitology*, 90:225–241.
- Rogers, D.J., S.E. Randolph, R.W. Snow, and S.I. Hay, 2002. Satellite imagery in the study and forecast of malaria, *Nature*, in press.
- Snow, R.W., M.H. Craig, U. Deichmann, and K. Marsh, 1999. Estimating mortality and disability due to malaria among Africa's non-pregnant population, *Bulletin of the World Health Organization*, 77(8):624–640.
- Snow, R.W., E. Gouws, J. Omumbo, B. Rapuoda, M.H. Craig, F.C. Tanser, D. le Sueur, and J. Ouma, 1998. Models to predict the intensity of *Plasmodium falciparum* transmission: applications to the burden of disease in Kenya, *Transactions of the Royal Society of Tropical Medicine and Hygiene*, 92(6):601–606.
- Snow, R.W., K. Marsh, and D. Le Sueur, 1996. The need for maps of transmission intensity to guide malaria control in Africa, *Parasitology Today*, 12:455–456.
- Tanser, F.C., 2000. *The Application of Geographical Information Systems to Infectious Diseases and Health Systems in Africa*, Faculty of Medicine, University of Natal, Durban, South Africa, 168 p.
- Thomson, M.C., S.J. Connor, U. D'Alessandro, B. Rowlinson, P. Diggle, M. Cresswell, and B.M. Greenwood, 1999. Predicting malaria infection in Gambian children from satellite data and bed net use surveys: The importance of spatial correlation in the interpretation of results, *American Journal of Tropical Medicine and Hygiene*, 61(1):2–8.
- Tucker, C.J., 1979. Red and photographic infrared linear combinations for monitoring vegetation, *Remote Sensing of Environment*, 8:127–150.
- Tucker, C.J., and P. Sellers, 1986. Satellite remote sensing of primary production, *International Journal of Remote Sensing*, 7: 1395–1416.
- WHO, 1999. *The World Health Report 1999: Making a Difference*, World Health Organization, Geneva, Switzerland, 122 p.

CHAPTER 7:

Modelling malaria risk in East Africa at high spatial resolution

JA Omumbo, SI Hay, AJ Tatem, RW Snow, DJ Rogers

Tropical Medicine and International Health, in press

SUMMARY

Objectives: *Malaria risk maps have re-emerged as an important tool for appropriately targeting the limited resources available for malaria control and evaluating progress towards targets set for coverage of anti-malarial interventions. In sub-Saharan Africa such maps are few and in most cases are not based upon empirical data or standardized criteria.*

Methods: *Statistical techniques applied to high spatial resolution remotely sensed climate surrogate data; digital human settlement surfaces and land use data have been used to predict the intensity of malaria transmission as defined categorically according to the childhood parasite ratio (PR) in East Africa. Discriminant analysis was used to train the environmental and settlement predictor variables to distinguish between four classes of PR risk that have been previously shown to relate to disease outcomes in the region.*

Results: *330 independent empirical estimates of the PR were identified from Kenya, Tanzania and Uganda. Surrogate markers of climate recorded on-board earth orbiting satellites, population settlement, elevation and water bodies all contributed significantly to the predictive models of malaria transmission intensity in the sub-region. The accuracy of the model was increased by stratifying East Africa into two ecological zones. In addition, forcing the discriminant analysis to include urbanization as a predictor of malaria prevalence improved the consistency of the predictive map with expert opinion malaria maps. The overall accuracy achieved with ecological zone and urban stratification was 62.1% with surrogates of precipitation (normalised difference vegetation index) and temperature (land surface temperature) being among the most discriminating predictors of the PR.*

Conclusion: *Although a high degree of statistical accuracy was achieved in the predictive map presented, the study revealed several difficulties in mapping malaria at high spatial resolution with currently available malariometric data. Current high spatial resolution empirically-driven models of malaria are limited by the poor resolution of malaria data. This can only be addressed in small areas studies and is perhaps not suitable or necessary for maps that are aimed for use in intervention targeting. It is suggested that lower resolution models could be more appropriate for current intervention targeting needs in East Africa.*

INTRODUCTION

The risks of morbidity and mortality due to *Plasmodium falciparum* vary spatially and temporally across the African continent (Trape & Rogier, 1996; Snow *et al.*, 1997; Snow & Marsh, 2002). There are a number of factors that determine the age-specific patterns of disease risk following infection but the most significant is the role played by acquired functional immunity which is dependent in turn on the frequency and duration of infection from birth (Snow & Marsh, 1998). A number of indices are used to describe the frequency of encounters with malaria infection; these include the entomological inoculation rate (EIR), the vectorial capacity (VC), infant parasite conversion rates and the parasite ratio (PR). The relationship between these markers of infection risk is complex (Macdonald, 1953; Smith & McKenzie, 2004) and most of them, with the exception of the PR, are rarely recorded in Africa (Hay *et al.*, 2000;

2004a). The PR has been used since the 1950's to characterise the intensity of malaria transmission. It is determined through cross-sectional surveys of children and expressed as the percentage found to be infected with *P. falciparum* (Metselaar & Van Theil, 1959). This measure has been shown to correspond categorically to the log of the annual EIR (Beier *et al.*, 1999; Hay *et al.*, 2004a) and, more recently, has been used to discriminate between malaria morbidity and mortality rates across Africa and globally (Snow *et al.*, 1999; Snow *et al.*, 2004; Snow & Marsh, 2002).

The need for maps of malaria's distribution has been recognised by malaria epidemiologists for as long as the disease has been studied (Gill, 1914; 1920; Boyd, 1930; Lysenko & Semashko, 1968; Lysenko & Beljaev, 1969; Dutta & Dutt, 1978). More recently, advances in Geographic Information Systems (GIS) and Remote Sensing (RS) have fuelled a renaissance in malaria risk mapping in Africa (Snow *et al.*, 1996; Hay *et al.*, 1998; Hay *et al.*, 2000; Lindsay *et al.*, 1998; Craig *et al.*, 1999; Thomson *et al.*, 1999; Rogers *et al.*, 2002; Tanser *et al.*, 2003). Relatively few of the available models of the spatial determinants of malaria risk have used empirical data (Snow *et al.*, 1998; Kleinschmidt *et al.*, 2000; 2001). Those available for Kenya and Mali, for example, have not fully exploited the wealth of newer high spatial resolution RS imagery available and have been based almost solely on climatic determinants of risk. Furthermore, influences of human settlement and water bodies are considered inadequately. In this paper we re-examine the distribution and intensity of malaria transmission across East Africa through the use of high spatial resolution satellite sensor imagery, rigorously selected PR data (Omumbo & Snow, 2004), ecozonation (Omumbo *et al.*, 2002), as well as, urbanization, water body and land use parameters that are known to affect transmission (Omumbo *et al.*, 2005).

MATERIALS AND METHODS

Parasite prevalence

A continuous comprehensive search of published and unpublished literature for malariometric data in the East Africa region has been maintained since 1996 as part of the Mapping Malaria Risk in Africa (MARA) collaboration (Snow *et al.*, 1996; <http://www.mara.org.za>). The methods used to identify the PR data in the East Africa region and the criteria used for their selection are presented elsewhere (Omumbo *et al.*, 1998; 2004; Omumbo & Snow, 2004). In brief, electronic database searches, contacts with local malariologists, extraction of data from Ministry of Health archives, manual searches of local journals, conference proceedings and post-graduate theses all provided the basis of the search strategy. Kenyan, Tanzanian and Ugandan survey data were selected for the present analysis according to the following inclusion criteria: more up to date assessments of infection risk (>1979); community-based surveys; age ranges within 0-15 years (surveys where only infants were sampled were excluded); surveys that sampled a minimum of 50 children and those that provided adequate details on survey location and denominators. Finally, repeat cross-sectional surveys on the same populations by the same investigators within a 24 month period were pooled to a single estimate while surveys undertaken by varied investigators at different times in the same location were reduced to one estimate by excluding either the earlier survey or smaller sample size.

Longitude and latitude co-ordinates (in decimal degrees) were determined for each parasitological survey using a variety of sources: 1:50,000 scale topographic maps (Directorate of Overseas Surveys, 1971), digital administrative unit maps in Kenya (UNEP- Global Resources Information Database, 1992) or public domain digital gazetteers (GDE Systems Inc., 1995; World Resources Institute, 1995). The centroid co-ordinates served as a unique identifier for each survey and, combined with a description of the survey, were used to describe the spatial extent of each sample. Surveys were classified as representing one of five spatial dimensions: First, using ArcView 3.2 (ESRI, Redlands, CA, USA), for surveys representing a single village, the central longitude and latitude was used to define an area of 1 km radius encompassing the community. The second spatial classification reflected surveys that sampled from several villages but presented the data as a single PR (5 and 7 surveys in Kenya and Tanzania respectively). In this case, a polygon was created to enclose the villages. The third, fourth and fifth spatial criteria corresponded to surveys undertaken at the 5th (sub-location: average area covered 9.4 km²), 4th (location: average area covered 15.4 km²) or 3rd (division: average area covered 34.1 km²) administrative levels in Kenya (Hay *et al.*, 2004b).

Predictor variables

Satellite sensor derived data at 1 x 1 km spatial resolution were obtained for East Africa (19.995 °E -52.005 °E and 23.755 °N - 13.005 °S) from the United States' Geological Survey, Distributed Active Archive Centre (URL: <http://edcdaac.usgs.gov/1KM/comp10d.asp>). These data were archived from the Advanced Very High Resolution Radiometer (AVHRR) on-board the National Oceanic and Atmospheric Administration's (NOAA) series of afternoon ascending polar-orbiting meteorological satellites (Eidenshink & Faundeen, 1994). All ten bands of raw channel and quality control data were downloaded for the 93 ten-day composites (dekads) from 31 months (Apr-Dec 1992, Jan-Sep 1993, Feb-Dec 1995 and Jan-May 1996) and re-sampled to a latitude and longitude co-ordinate reference system. Data quality control flags, solar and zenith scan angle correction and maximum value compositing procedures were implemented with ERDAS Imagine 8.5 (Leica Geosystems GIS & Mapping, Atlanta, USA). For more comprehensive details on these procedures see Green & Hay, (2002) and Hay *et al.*, (2000).

Monthly time-series of three primary predictor variables were derived from these images for analysis: (i) the normalized difference vegetation index (NDVI); an indicator of photosynthetic activity and surrogate for moisture availability (Hay *et al.*, 1998; 2000); (ii) the land surface temperature (LST); whose accuracy is similar to that of spatially interpolated temperature data obtained from ground meteorological stations in Africa (Hay & Lennon, 1999; Hay *et al.*, 2000); and (iii) the middle infrared reflectance (MIR); a satellite "temperature" band that is sensitive to both reflected and emitted radiation (Boyd & Curran, 1998) and included since it suffers less from atmospheric attenuation than LST (Cracknell, 1997).

Cold Cloud Duration (CCD) data were also derived from the High Resolution Radiometer (HRR) onboard the European Meteorological Satellite programme's (EUMETSAT) Meteosat satellite series and used as a surrogate measure of rainfall. CCD image pixels represent the number of hours during which cold cloud top temperatures below a geographically variable threshold were experienced during a 10-

day compositing period. CCD threshold temperatures have been derived empirically for areas of Africa between 0–27°N of latitude (Snijders, 1991). The CCD has been found to have a root mean square error of ± 38 mm when compared with meteorological station recordings across continental Africa (Hay, 2000; Hay & Lennon, 1999). The 0.01-degree spatial resolution CCD data used for this study were re-sampled to 1 x 1 km spatial resolution.

Landcover data at full spatial resolution (1:100,000) were requested and downloaded from the Africover project's website (<http://www.africover.org>). Africover urban and water body themes were produced from visual interpretation of digitally enhanced Landsat Thematic Mapper (TM) images (bands 4, 3, 2) acquired, mainly in 1999, across Kenya. Previous work has shown these variables to be important local modifiers of the PR (Omumbo *et al.*, 2004). The urban area and water body polygons were overlaid on a 1x1 km grid of the same dimensions as the satellite imagery and the percentage area of each pixel occupied by the land-cover class calculated. Altitude data available and obtained from a global digital elevation model (Hastings & Dunbar, 1998) were resampled to 1x1 km spatial resolution using ERDAS Imagine 8.5 (Leica Geosystems GIS & Mapping, Atlanta, USA).

Previous statistical modelling efforts for vector-borne diseases in eastern Africa have shown that predictions are improved markedly by clustering data according to areas of ecological similarity (Brooker *et al.*, 2001; 2002a; 2002b). The factors that influence the distribution and/or abundance of vectors vary between ecological zones (ecozones). Studies of tsetse fly distributions, for example, show that at the limits of the continental range of flies, a single climatic variable can be used to explain tsetse distributions while in areas of greater fly abundance, several climatic variables are needed for an adequate description (Rogers & Robinson, 2004).

In the present analysis we have defined two broad ecological zones across East Africa (Omumbo *et al.*, 2002; 2004). Ecozone 1 (Figure 1) represents areas at the edge of the distribution of malaria vectors where climatic conditions do not favour the propagation of vectors for most parts of the year. Within these areas it is possible to define vector distributions or abundance using a single climatic variable. Thus Ecozone 1 is composed of arid areas where transmission is limited by low rainfall (vector breeding is restricted to rainy seasons or to areas near water bodies (Figure 1, Area A)), and highland areas where low temperatures limit vector and parasite development (Figure 1, Area B)). The ecology of Ecozone 2 is diverse and in general climatic conditions favour the proliferation of malaria vectors and parasites allowing for longer transmission seasons. This ecozone can be described as being well within the climatic range of malaria vector distribution and more than a range of environmental variables are required to describe malaria transmission. Ecozone 2 tends also to be more densely populated (particularly in Uganda) and better served by perennial rivers.

Statistical analysis

Since monthly climate data are serially correlated and thus have information redundancy, they were pre-processed using temporal Fourier analysis (TFA). TFA summarises the correlated data and in so doing, captures epidemiologically important seasonal variations and these are then described in statistically uncorrelated outputs.

The origins, mathematical basis and arguments for the biological appropriateness of TFA are developed in detail elsewhere (Rogers *et al.*, 1996; 2002; Rogers & Randolph, 2000). 14 outputs were recorded from the TFA for each satellite time-series variable. These included the amplitudes of the annual (a1), bi-annual (a2) and tri-annual (a3) cycles; the phases (in months) of annual (p1), bi-annual (ph2) and tri-annual (p3) cycles; the proportion of the variance in the original time-series described by the annual (d1), bi-annual (d2), tri-annual (d3) and all cycles (da), as well as the maximum (mx), minimum (mn), mean (a0) and variance (vr) of the cycle recomposed from the first three cycles. All these variables and in addition, the elevation derived from the DEM and percentage of urban and water bodies derived from Africover, were available to the discriminant analysis.

Discriminant analysis (DA) was used to identify environmental and population variables that discriminated best between four categories of prevalence namely; 0-<5, 5-<25%, 25-<75%, and \geq 75%). These classes correspond closely to categorical descriptions of malaria morbidity and mortality burden that have been described previously for Africa (Snow *et al.*, 2003). Ideally the training data should include surveys from areas where malaria transmission does not occur. Such data, however, are rare for malaria prevalence studies as surveys are generally not conducted in places where and at times when malaria is absent. Furthermore, in areas of seasonal transmission, a prevalence of zero is not indicative of the absence of transmission as it may have been recorded outside the transmission season. It is difficult to define any sites in East Africa, other than unpopulated areas or mountains, as truly malaria-free. Very few of the PR surveys identified in this study recorded a true zero prevalence (N=10). All ten studies were from areas described according to expert opinion as high altitude or of seasonal transmission. A zero category was not used in this analysis due to the small number of studies and this group has been included in the lowest infection risk class (PR=0-<5%).

The DA procedures used have been applied previously to trypanosomiasis distribution mapping and are outlined in detail by Rogers *et al.*, (1996), Rogers & Randolph, (2000) and Rogers & Robinson, (2004). Variables were selected iteratively on the basis of the generation of the maximum *kappa* statistic (Cohen, 1960; Jensen, 1996) compared with the other variables in each round of selection. The procedure was repeated until either the maximum *kappa* value was generated or 10 most discriminating variables were selected. It was assumed that the probabilities of membership in any of the four categories of prevalence were equal (*i.e.* the prior probabilities were 1/number of prevalence categories). Predictions were completed for all data and the effect of forcing urbanization was tested separately within each ecological zone.

For each training data point, a predictor variable value was extracted using IDRISI Version 2 (IDRISI project, Clarke Labs, Worcester, MA) and output in a database for use in the DA. A customised mapping program (written in Microsoft QuickBASIC Version 4.0) was used to derive a discriminant function based on comparing the training data with the predictor variables. Each image pixel was assigned to a one of the 4 PR prevalence categories on the basis of the discriminant function and a predictive image derived. Predictions were not made for pixels where environmental conditions were significantly different from the training data pixels.

The accuracy of a predictive map is influenced by several factors, notably the sample sizes of data available in each class (in some cases these may not be large enough to provide a high confidence level in the prediction). Accuracy and agreement was assessed by comparison of observed values with predicted values by means of a classification or confusion matrix (Jensen, 1996). Agreement was estimated using the *kappa* (K) and *tau* (T) statistics; related indices of agreement that compensate for the agreement that would be expected due to chance. Values of *kappa* <40% indicate poor agreement; values between 40% and 75% suggest good agreement and values above 75%, excellent agreement (Landis & Koch, 1977). Three measures of accuracy were determined; the overall accuracy ((OA) determined by the proportion of pixels in the main diagonal divided by the total number of training data pixels), the producer's accuracy (PA) and the consumer's accuracy (CA). PA and CA measure accuracy for individual prediction classes. PA (or omission error) refers to the probability of a training data pixel being correctly classified and is the proportion of training data (observed) pixels in a category of prevalence that are classified correctly. CA (or commission error) is a measure of the probability that predicted pixels represent the true classification on the ground and is the proportion of predicted pixels that are classified correctly (Story & Congalton, 1986).

RESULTS

The malaria prevalence data

The data search identified 330 parasite survey data points that fulfilled the inclusion criteria. The spatial distribution of these data is shown in Figure 1. 217 studies were from Kenya, 86 in Tanzania and 27 in Uganda (for details of the PR data see Omumbo & Snow, 2004). The mean sample size was 375 children with a median (interquartile range) of 204 (120; 427). Thirty five surveys described a PR 0-<5% (N=10 for PR=0), 80 surveys recorded a prevalence between 5-<25%, 177 surveys recorded a prevalence between 25-<75% and 38 surveys reported prevalence's in the childhood populations surveyed greater than or equal to 75%.

The malaria transmission predictive model

DA was performed initially without controlling for ecological zone or urbanization and the accuracy of the prediction tested. OA (the proportion of correctly classified pixels over all prevalence categories), was 72.4% (K = 0.502, T = 0.494). On visual comparison with historical (Government of Tanganyika, 1956) and contemporary (Craig *et al.*, 1999; Omumbo *et al.*, 2002) modelled malaria risk maps, the resulting predictive map was found to be significantly anomalous in southern Tanzania. In this area, the model over-predicted areas of low malaria prevalence (PR=5-<25%). The results were improved by stratifying the analysis according to two ecozone classes and by forcing the inclusion of urbanization as a predictor (Figure 2, Area A). These modifications marginally reduced OA in both Ecozone 1 [OA =64.0%; K = 0.483; T = 0.478 (Table 1a)] and Ecozone 2 [OA = 61.4 %; K = 0.45; T= 0.308 (Table 1b)] but provided an output with fewer large-area anomalies when compared with historical (Figure 3b) and more recent climate-driven maps (Figures 3c & d). The overall accuracy for the combined ecozone/urban adjusted map shown in Figure 2 was 62.1% (K =0.477, T = 0.495).

PA and CA values by ecological zone class are also provided in Tables 1a & 1b. In both Ecozones 1 & 2, training data predictions were poorest for the PR category 25- <75% (PA=53.6% & PA=45.6% respectively). CA was poorest for the PR category $\geq 75\%$ also for in both ecozones (CA=42.9% (Ecozones 1) and 332.7% (Ecozone 2)).

The top ten most discriminating predictor variables selected during the DA and subsequently used in the final prediction were: NDVI a1, LST a2, LST p2, CCD min, Africover water body %, DEM, MIR p1, CCD mean, MIR p2, percent urban (Africover classification). Vegetation, rainfall and 'water body area' were chosen as significant predictors in the "dry" Ecozone 1. In Ecozone 2 where we assume water was not generally limiting, temperature variables were most abundant among the predictor variables selected: LST mean, LST p1, MIR min, LST p2, MIR p1, CCD max, NDVI p1, DEM, LST a2, Africover urban %.

DISCUSSION

We have developed a predictive model of four categories of malaria prevalence in East Africa. The model is driven by empirical data and fully exploits the potential of currently available satellite-derived climate surrogate and other digital environmental data. The significance of non-climatic determinants of malaria transmission, such as urbanization, in determining the intensity of malaria is highlighted. The application of an ecological zone stratification increased overall accuracy by 6.1% and increased *kappa* values from 0.394 to 0.477, values. Similar improved results have been seen with the use of ecozone stratification for other models of malaria in West Africa (Kleinschmidt *et al.*, 2001) and for schistosomiasis (Brooker *et al.*, 2002).

The modelling approach used here differs in a number of important respects to previous empirical malaria mapping efforts (Snow *et al.*, 1998; Kleinschmidt *et al.*, 2000; 2001). Firstly, rigorous efforts have been made to ensure that the training data were of highest quality. To this end, historical, and possibly outdated, data have been excluded while only samples with large sample sizes that were randomly selected were included. For studies that overlapped spatially, an attempt was made to select the 'best' estimate for the study site. Secondly, much higher spatial resolution satellite data than have been utilised previously are used here. The environmental data, furthermore, have been summarised using TFA and so relate to the seasonal cycles of transmission. Thirdly, the influences of non-climatic factors such as urbanization have been allowed for. Urbanization has been conclusively shown to reduce *P. falciparum* entomological inoculation rates (Hay *et al.*, 2000; 2004a; Robert *et al.*, 2003) and human infection risks (Omumbo *et al.*, 2004) independent of climate and was seen to be an important influence in this study.

The methods used in this study have produced statistically accurate maps (OA>60%). The highest prevalence categories, however, were over predicted (PA>75% compared with CA<43%). The confidence of this measure is difficult to predict as there were very few training data pixels in this category (N=38). It is important to assess the consistency of this prediction with what is known of malaria transmission across the region and previous malaria maps.

Historical maps of malaria for East Africa (Figure 3b) were largely based on expert opinion and knowledge of climatic patterns. Malaria intensity is defined according to the duration of transmission seasons; at one extreme are highly seasonal areas where transmission is limited to rainy seasons and at the other extreme transmission is more stable, occurring during well-defined seasons (usually the rainy seasons) each year. The map developed in this study is compared with the historical map (Figures 3a & 3b). While there are similarities at the low extreme of transmission (Figure 3a; Area A) there is a marked anomaly in Tanzania in the region south of Lake Victoria where the model appears to have under-predicted high prevalence areas (Figure 3a; Area B). The reasons for this are not clear but may be due to the lack of training data from the region (Figure 1; Area C). It is possible that none of the available training data were able to adequately describe the environmental conditions found in the area. This area of Tanzania has been described very rarely in malaria studies since historical times (Omumbo & Snow, 2004). Climate based models (Figure 3c; Area A Craig *et al.*, 1999) suggest climatic conditions in this area are highly suitable for vector and parasite development but do not describe the variations in transmission intensity that are likely exist. Highland areas defined as malaria-free according to the historical map are, in general, predicted well as are areas in north western Kenya at the low extremes of transmission.

Despite efforts to maximise the accuracy of the predictive map developed in this study, the measures of statistical accuracy derived suggest that there are still many transmission modifying factors that are unaccounted for. A *kappa* value of 0.477, though good, is far from excellent. It is clear that clustering the training data according to appropriate ecological strata markedly improves predictions. Appropriate spatial definitions of such strata, however, are not available. Only 2 ecozones were defined for this study but it is possible that 2 classes are inadequate to describe the ecology of the study area.

The marked improvement in the visual consistency of urban-forced predictions suggests that hitherto undefined population factors may improve predictions further. The definition of 'urban' in this study was problematic. The extents of urban centres are poorly defined and a clear definition of 'urbanized' is unavailable to date. The effect of urbanization on malaria transmission is largely mediated through changes in local habit brought about by urbanization and development of infra-structure and so on. These factors are ill-defined for Africa and so their inclusion in the model was not possible. Many of the areas of poor prediction revealed here could be improved through additional targeted sampling. Such targeted sampling stratified according to relevant ecological zones have been shown to improve prevalence maps of helminths and schistosomiasis markedly (Montresor *et al.*, 1998) and could possibly benefit malaria mapping where spatially random malariometric data are not available to date.

The training data set highlights first that there is a paucity of reliable information across vast areas of East Africa (Figure 1), as such providing the impetus for modelling and mapping the risks for areas where there are no data. The most poorly represented areas are those where transmission is least intense or population densities are lowest. As mentioned above, one of the greatest limitations of the PR samples is that the study sites are not a random sample of ecological settings across East Africa and therefore do not adequately represent the spectrum of transmission conditions across the region. Instead the data used are derived inevitably from populations of "special interest", representing areas where research groups or malaria control

agencies perceive risk or serve specific research interests. The training data are biased by *a priori* knowledge of where the highest PRs are likely to exist.

The lack of more than one PR estimate for each site is an additional limitation. Aside from precluding the investigation of seasonal influences, the accuracy of any single PR estimate can not be assessed. In this study we have had to assume that each PR estimate used was measured accurately, is comparable with other estimates across the region and is a true representation of transmission in the study site. None of these assumptions can be verified without further sampling.

In using high spatial resolution environmental data we have also had to assume that the resolution of the PR data is comparable. This was not the case in practice. None of the studies described the exact spatial extent from which samples were drawn. Where administrative units were used as a sampling frame, the distribution of populations within these samples was not considered. The result is that possibly very low spatial resolution data have been used to describe transmission at high spatial resolution resulting in errors which are most evident in areas where environmental conditions are most variable. We conclude that perhaps it is not possible to address this question at this spatial resolution (particularly due to the poor resolution of the PR data) and it may be better to develop maps on a smaller scale with purposive sampling according to well-defined ecological zone strata. In this way local factors such as population distribution and effects of urbanization can be investigated concurrently. Longitudinal studies from such sites would also be more meaningful.

This area remains a challenge and the need for small area studies could be addressed within the framework of demographic surveillance sites. Maps are still needed, however, to provide a bigger picture to address the needs of intervention targeting which has been problematic to date. It has often been argued that without at least a basic knowledge of risk, efforts to control malaria will lack the ability to target limited resources to maximise coverage appropriately among those most at risk (Snow *et al.*, 1996). Such evidence-based intervention targeting has not been possible due to the lack of an appropriate tool to define areas of risk. The international community has been effective in raising the importance of malaria control as a priority health investment for Africa through the Roll Back Malaria (RBM) Partnership (WHO, 1999; WHO 2003). Nevertheless five years after the formation of RBM, countries in East Africa still have very low coverage of interventions likely to reduce the burden of malaria in this sub-region. National surveys in East Africa show that less than 5% of children less than 5 years old are protected by an insecticide-treated net ((ITN) Monasch, 2004). In addition there are large disparities in the spatial components of access to ITN in all these countries. A recent Demographic and Health Survey in Kenya showed that the highest rates of ITN use were in Nairobi, the province least likely to experience malaria risk (CBS, 2003).

The importance of very accurate high spatial resolution malaria maps is arguable given that the aim of the maps is to target interventions usually done at the district level. Thus can be achieved successfully with lower resolution data and perhaps that is the way forward for malaria mapping.

ACKNOWLEDGEMENTS

This study received financial support from The Wellcome Trust, UK; International Development Research Centre, Canada; the South African Medical Research Council, the MARA/ARMA collaboration and the Kenya Medical Research Institute. The Wellcome Trust supports JAO, SIH and RWS as part of their Prize Studentship (#060063), Research Career Development (#069045) and Senior Research Fellow (#058992) programmes respectively. The authors acknowledge the contribution of malaria control personnel and scientists in East Africa who shared unpublished data. We are grateful to the Africover team for their efficiency and advice. Carlos Guerra is thanked for help with manipulating extractions and Africover data. This paper is published with the permission of the director of KEMRI.

REFERENCES

- Beier JC, Killeen GF & Githure JI (1999). Entomologic inoculation rates and *Plasmodium falciparum* malaria prevalence in Africa. *American Journal of Tropical Medicine and Hygiene* **61**, 109-113.
- Boyd DS & Curran PJ (1998). Using remote sensing to reduce uncertainties in the global carbon budget: The potential of radiation acquired in the middle infrared wavelengths. *Remote Sensing Reviews* **16**, 293-327.
- Boyd MF (1930). An introduction to malariology. Cambridge, Massachusetts, Harvard University Press.
- Brooker S, Hay SI, Issae W *et al.* (2001). Predicting the distribution of urinary schistosomiasis in Tanzania using satellite sensor data. *Tropical Medicine and International Health* **6**, 1-10.
- Brooker S, Hay SI & Bundy DAP (2002). Tools from ecology: useful for evaluating infection risk models. *Trends in Parasitology* **18**, 70-74.
- Butler RJ (1959). Atlas of Kenya: A comprehensive series of new and authenticated maps prepared from the national survey and other government sources with gazetteer and notes on pronunciation and spelling. Nairobi, Kenya, The Survey of Kenya.
- Cohen J (1960). A co-efficient of agreement for nominal scale. *Educational and psychological measurements* **20**, 37-46.
- Cracknell AP (1997). Channel 3. The neglected channel. In: The Advanced Very High Resolution Radiometer (AVHRR). London, Taylor & Francis: 321-342.
- Craig MH, Snow RW & Le Sueur D (1999). A climate-based distribution model of malaria transmission in sub-Saharan Africa. *Parasitology Today* **15**, 105-111.
- Directorate of Overseas Surveys (1971). Series Y731. Nairobi, Survey of Kenya.

GDE Systems Incorporated (1995). Populated places: GeoName Digital Gazetteer v.1, Geographic Information Department (MZ 1211-M), GDE Systems Inc., P.O. Box 509009, San Diego, California.

Gill CA (1920). The rôle of meteorology in malaria. *Indian Journal of Medical Research* **8**, 633-693.

Government of Tanganyika (1956). Atlas of Tanganyika, East Africa. Dar es Salaam, Government Press, Tanganyika.

Government of Uganda (1962). Atlas of Uganda. Uganda, Department of Lands and Surveys.

Green RM & Hay SI (2002). The potential of Pathfinder AVHRR data for providing surrogate climatic variables across Africa and Europe for epidemiological applications. *Remote Sensing of Environment* **79**, 166-175.

Hastings DA & Dunbar PK (1998). Development and assessment of the Globe One-km Base Elevation Digital Elevation Model. *International Society of Photogrammetry and Remote Sensing Archives* **32**, 218-221.

Hay SI, Snow RW & Rogers DJ (1998). Predicting malaria seasons in Kenya using multi-temporal meteorological satellite sensor data. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **92**, 12-20.

Hay SI & Lennon JJ (1999). Deriving meteorological variables across Africa for the study and control of vector-borne disease: a comparison of remote sensing and spatial interpolation of climate. *Tropical Medicine and International Health* **4**, 58-71.

Hay SI (2000). An overview of remote sensing and geodesy for epidemiology and public health application. *Advances in Parasitology* **47**, 1-35.

Hay SI, Omumbo JA, Craig MH & Snow R.W. (2000). Earth observation, geographic information systems and *Plasmodium falciparum* malaria in sub-Saharan Africa. In: *Advances in Parasitology* **47**, 173-215.

Hay SI, Rogers DJ, Toomer JF & Snow RW (2000). Annual *Plasmodium falciparum* entomological inoculation rates (EIR) across Africa: literature survey, internet access and review. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **94**, 113-127.

Hay SI, Guerra CA, Tatem AJ, Noor AM & Snow RW (2004). The global distribution and population at risk of malaria: past, present, and future. *Lancet Infectious Diseases* **4**, 327-336.

Hay SI, Guerra CA, Tatem AJ, Aitkinson PM & Snow RW (2005a). Urbanization, malaria transmission and disease burden in Africa. *Nature Reviews Microbiology*, in press.

Hay SI, Noor AM, Nelson A & Tatem AJ (2005b). Demography for epidemiology: the precision of large-area human population maps. *International Journal of Epidemiology*, in press.

Jensen JR (1996). Thematic information extraction: image classification. Land-use classification map accuracy assessment. In: Introductory digital image processing. P Hall. Upper Saddle River, New Jersey, Prentice Hall. 247-256.

Kleinschmidt I, Bagayoko M, Clarke GPY, Craig MH & Le Sueur D (2000). A spatial statistical approach to malaria mapping. *International Journal of Epidemiology* **29**, 355-361.

Kleinschmidt I, Omumbo JA, Briët O *et al.* (2001). An empirical malaria distribution map for West Africa. *Tropical Medicine and International Health* **6**, 779-786.

Landis J & Koch G (1977). The measurement of observer agreement for categorical data. *Biometrics* **33**, 159-174.

Lindsay SW, Parson L & Thomas CJ (1998). Mapping the range and relative abundance of the two principal African malaria vectors, *Anopheles gambiae* sensu stricto and *An. arabiensis*, using climate data. *Proceedings of the Royal Society of London: Biological Sciences* **265**, 847-854.

Lysenko AY & Beljaev AE (1969). An analysis of the distribution of *Plasmodium ovale*. *Bulletin of the World Health Organization* **40**, 383-394.

Macdonald G (1953). The measurement of malaria transmission. *Proceedings of the Royal Society of Medicine* **48**, 295-301.

Metselaar D & Van Theil P (1959). Classification of malaria. *Tropical and Geographical Medicine* **11**, 157-161.

Omumbo J, Ouma J, Rapuoda B, Craig MH, le Sueur D & Snow RW (1998). Mapping malaria transmission intensity using geographical information systems (GIS): an example from Kenya. *Annals of Tropical Medicine and Parasitology* **92**, 7-21.

Omumbo JA, Hay SI, Goetz SJ, Snow RW & Rogers DJ (2002). Updating historical maps of malaria transmission intensity in East Africa using remote sensing. Special Issue: Remote Sensing and Human Health. *Photogrammetric Engineering and Remote Sensing* **68**, 161-166.

Omumbo JA, Guerra CA, Hay SI & Snow RW (2005). The influence of urbanization on measures of *Plasmodium falciparum* infection prevalence in East Africa. *Acta Tropica*, **93**: 11-21.

Omumbo JA & Snow RW (2004). *Plasmodium falciparum* parasite prevalence in East Africa: A review of empirical data 1927-2003. *East African Medical Journal*, **81**: 649-656.

Robert V, MacIntyre K, Keating J, *et al.* (2003). Malaria transmission in urban sub-Saharan Africa. *American Journal of Tropical Medicine and Hygiene* **68**, 169-176.

- Rogers DJ, Randolph SE, Snow RW, Hay SI (2002). Satellite imagery in the study and forecast of malaria. *Nature* **415**, 710-715.
- Rogers DJ & Robinson TP (2004). Chapter 8: Tsetse Distribution. In: The Trypanosomiases. I. Maudlin, P.H. Holmes & M.A. Miles (eds). CABI Publishing, CAB International, Wallingford, Oxfordshire UK. pp. 139-179.
- Smith DL & McKenzie FE (2004). Statics and dynamics of malaria infection in *Anopheles* mosquitoes. *Malaria Journal* **3**, Article 13.
- Snijders FL (1991). Rainfall monitoring based on Meteosat data - a comparison of techniques applied to the western Sahel. *International Journal of Remote Sensing* **12**, 1331-1347.
- Snow RW, Marsh K & Le Sueur D (1996). The need for maps of transmission intensity to guide malaria control in Africa. *Parasitology Today* **12**, 455-456.
- Snow RW, Omumbo JA, Lowe B, *et al.* (1997). Relation between severe malaria morbidity in children and level of *Plasmodium falciparum* transmission in Africa. *The Lancet* **349**, 1650-1654.
- Snow RW, Gouws E & Omumbo JA, *et al.* (1998). Models to predict the intensity of *Plasmodium falciparum* transmission: applications to the burden of disease in Kenya. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **92**, 601-606.
- Snow RW & Marsh K (1998). New insights into the epidemiology of malaria relevant for disease control. *British Medical Bulletin* **54**, 293-309.
- Snow RW, Craig MH, Deichmann U & Marsh K (1999). Estimating mortality and disability due to malaria among Africa's non-pregnant population. *Bulletin of the World Health Organization* **77**, 624-640.
- Snow RW & Marsh K (2002). The consequences of reducing transmission of *Plasmodium falciparum* in Africa. *Advances in Parasitology* **52**, 235-264.
- Snow RW, Craig MH, Newton CR & Steketee RW (2003). The public health burden of *Plasmodium falciparum* malaria in Africa: Deriving the numbers. Working paper 11, Disease Control Priorities Project (DCPP). Bethesda, Maryland, Fogarty International Centre, National Institutes of Health. August 2003. <http://www.fic.nih.gov/dcpp>: 1-75.
- Snow RW, Korenromp EL & Gouws E (2004). Pediatric mortality in Africa: *Plasmodium falciparum* malaria as a cause or risk? *American Journal of Tropical Medicine and Hygiene* **71**, 16-24.
- Story M & Congalton RG (1986). Accuracy assessment: a user's perspective. *Photogrammetric Engineering and Remote Sensing* **52**, 397-399.
- Tanser FC, Sharp BL & Le Sueur D (2003). Potential effect of climate change on malaria transmission in Africa. *Lancet* **362**, 1792-1798.

Thomson MC, Connor SJ, D'Alessandro U, *et al.* (1999). Predicting malaria infection in Gambian children from satellite data and bed net use surveys: The importance of spatial correlation in the interpretation of results. *American Journal of Tropical Medicine and Hygiene* 61, 2-8.

Trape J-F & Rogier C (1996). Combating malaria morbidity and mortality by reducing transmission. *Parasitology Today* 12, 236-204.

World Resources Institute (1995). Africa Data Sampler: A Geo-Referenced Database for all African Countries (CD-ROM and User's Guide), World Resources Institute, World Conservation Monitoring Center, and PADCO, Inc.

TABLES AND FIGURES

Table 1a: Classification matrix for urban-forced prediction stratified by ecozone: Ecozone 1 (OA=64%; K=0.483; T=0.478)

		Predicted prevalence (%)				N	PA%	CA%
		0-<5	5-<25	25-<75	>=75			
Observed prevalence	0-<5	18	3	3	0	24	75	81.8
	5-<25	3	21	7	2	33	63.6	60
	25-<75	1	10	15	2	28	53.6	60
	>=75	0	1	0	3	4	75	42.9
	N	22	35	25	7	89		

Table 1b: Classification matrix for urban-forced prediction stratified by ecozone: Ecozone 2 (OA=61.4%; K=0.45; T=0.308)

		Predicted prevalence (%)				N	PA%	CA%
		0-<5	5-<25	25-<75	>=75			
Observed prevalence	0-<5	11	0	0	0	11	100	91.7
	5-<25	0	35	8	4	47	74.5	71.4
	25-<75	1	14	68	66	149	45.6	89.5
	>=75	0	0	0	34	34	100	32.7
	N	12	49	76	104	241		

Figure 1: Distribution of PR surveys according to ecological strata; Ecozone 1 (light blue) comprises arid and high altitude areas. Ecozone 2 (green) comprises ecologically diverse areas. Black dots show the distribution of PR studies

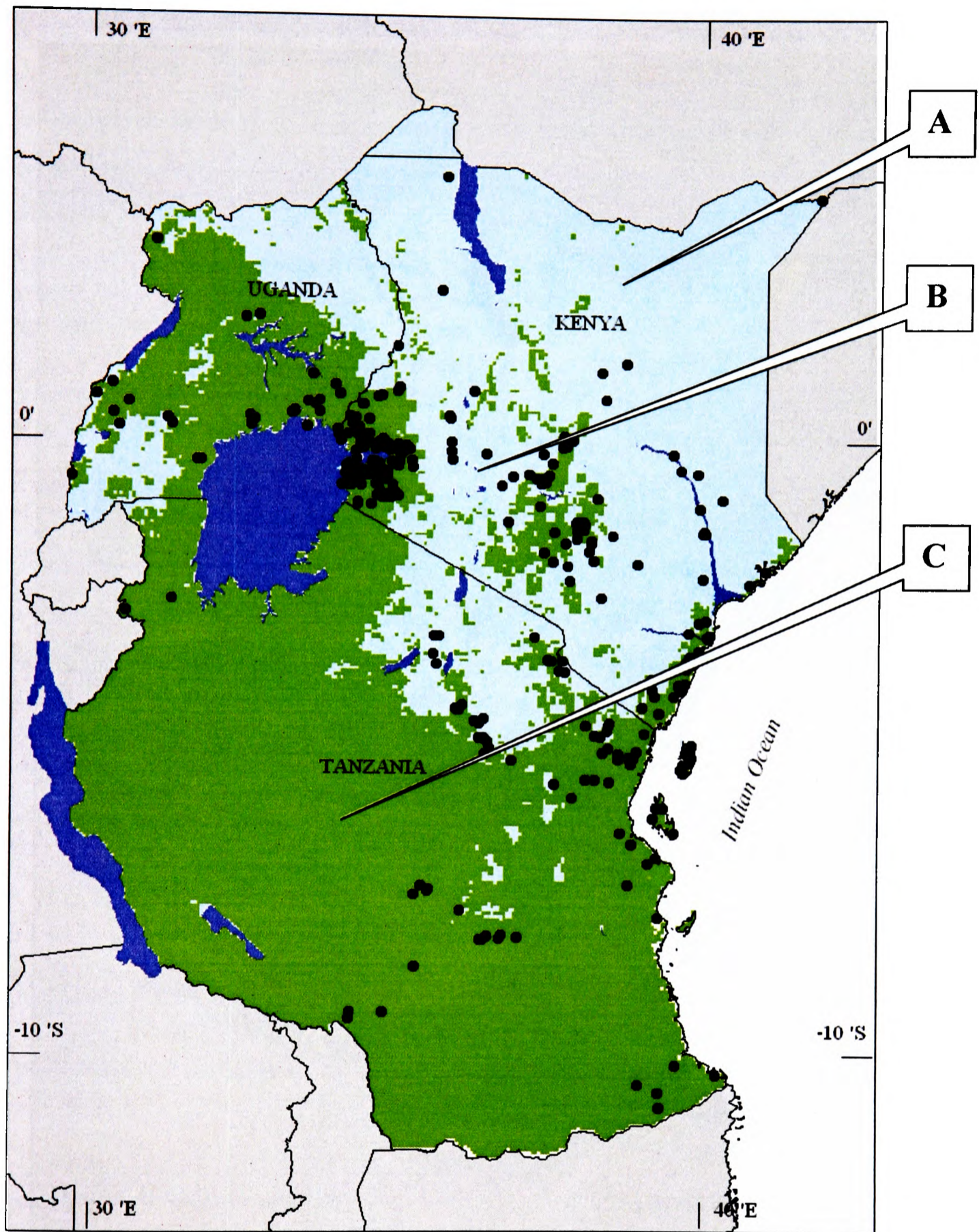


Figure 2: Urban forced prediction with ecological zone stratification. Legend: White=0- <5%, yellow=5-<25%, light green=25-<75%, dark green= >=75%PR. Blue areas are water bodies

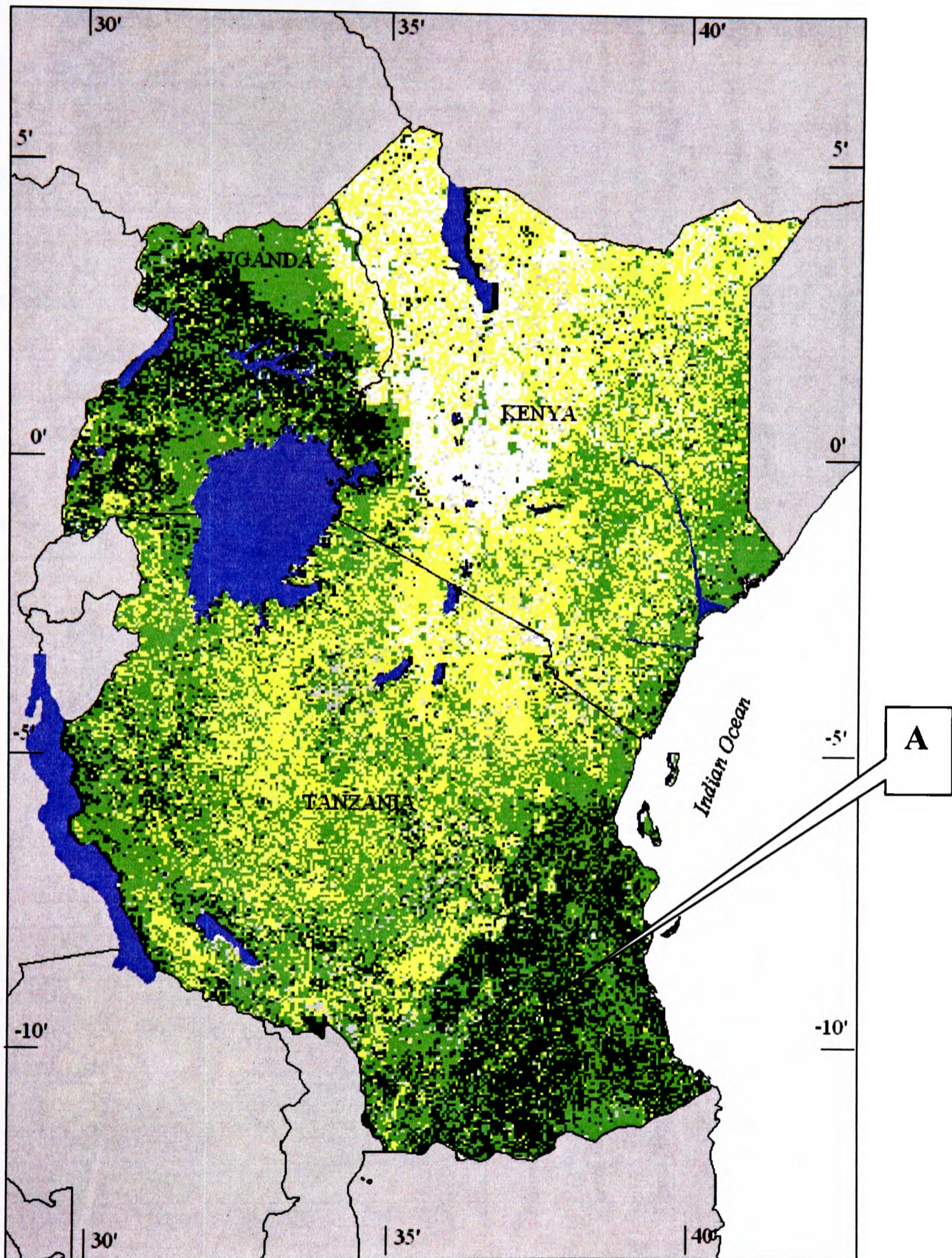
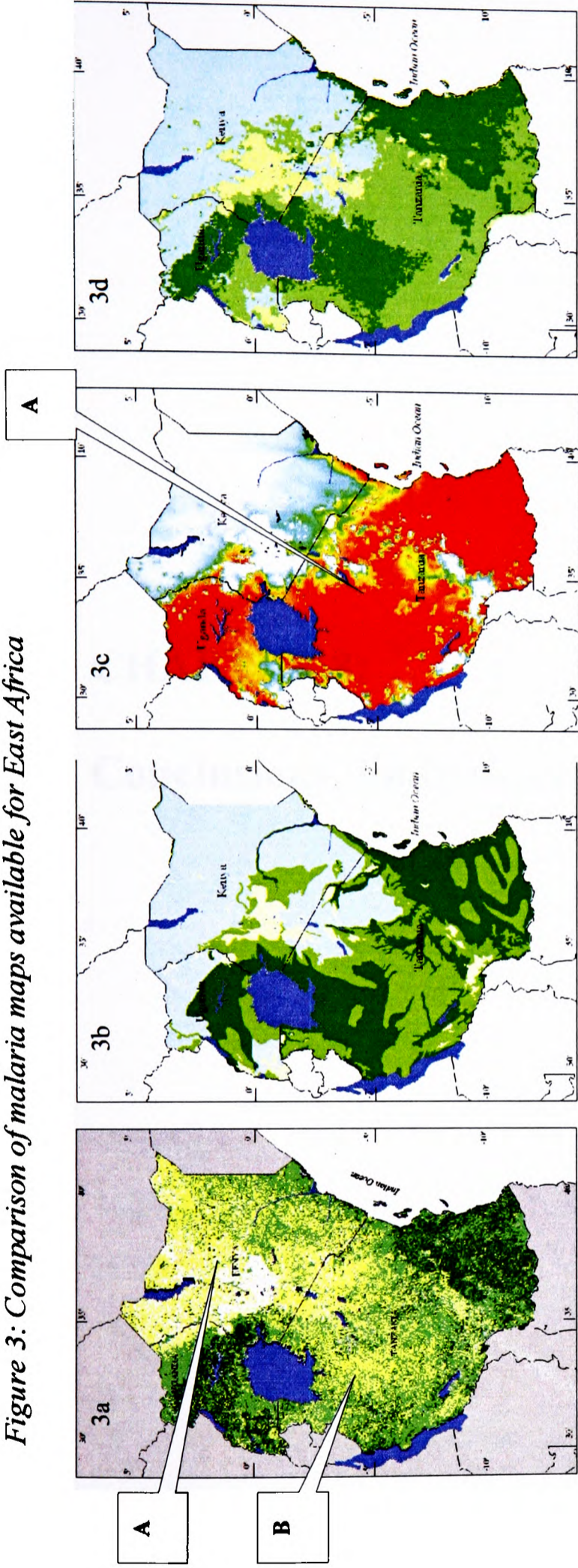


Figure 3: Comparison of malaria maps available for East Africa



- Figure 3a: Empirical model of malaria transmission in East Africa. Ecozone stratified and urban-forced prediction \square PR 0<5%, \square PR 5-<25%, \square PR 25-<75%, \square PR >=75%, \square no prediction, \square water body
- Figure 3b: Historical, expert opinion, map of malaria transmission for Kenya (Butler, 1959), Uganda (Government of Uganda, 1962) and Tanzania (Government of Tanganyika, 1956). \square = malaria free; \square = malaria near water; \square = transmission for 3-6 months in a year; \square = transmission for >=6 months in a year
- Figure 3c: MARA fuzzy climate suitability model: 0=unsuitable; 1=suitable; FCS=fuzzy climate suitability: \square FCS=0, \square FCS=<0.1, \square FCS=0.1-0.25, \square FCS=0.25-0.5, \square FCS=0.5-0.75, \square FCS=0.75-0.9, \square FCS=0.9-1.0
- Figure 3d: Discriminant analysis based predictions of historical maps of malaria in East Africa using 8x8 km satellite-derived surrogates of climate (Omumbo *et al.*, 2002). \square =malaria free; \square =malaria near water; \square =transmission 3-6 months in a year; \square =transmission >=6 months in a year.

CHAPTER 8:

Conclusions, limitations and future work

8.1. Summary

In Chapter 1 the determinants of *P. falciparum* distribution across SSA were presented emphasising both the biotic and abiotic factors. Climate, through its effect on *P. falciparum* and mosquito vector development and survival, has repeatedly been shown to be a major determinant of the limits of the distribution and intensity of *P. falciparum* transmission. As presented in Chapters 1 and 2, satellite-derived climate databases and their analysis have improved markedly over the past 5 years. The applications of these multi-temporal, high spatial resolution satellite data, with better ordination using temporal Fourier analysis techniques have been exploited in this thesis to examine the relationships between malaria risks and climate whilst retaining a basic biological framework.

The paucity of empirical data of transmission to train climate data has, to-date, been a major limitation for past efforts at malaria mapping. The thesis was structured to tackle this through an extensive data search that identified 2003 *P. falciparum* prevalence studies from across Kenya, Uganda and Tanzania (Chapters 2 & 3). Following the application of rigorous selection criteria (based on sample sizes >50, year of study post 1979, age range of sample 0-15 years old, and limiting inclusion to only randomly selected community based samples), 330 out of 2003 data points were selected as the most representative data available of the range of transmission intensities across East Africa. The study has used these data to investigate several features of malaria transmission. Firstly, they were used to perform a validation of a frequently used climate-

driven model of malaria distribution, the MARA model (Craig *et al.*, 1999) (Chapter 4), secondly to investigate an important relationship between transmission and urbanization (Chapter 5) and finally the data were used to predict *P. falciparum* prevalence for un-sampled sites across East Africa (Chapter 7).

The comparison of empirical *P. falciparum* prevalence data for Kenya with the MARA climate-driven model of the likelihood of stable transmission (Chapter 4) highlighted the limitations of this model to stratify *P. falciparum* malaria risk. The MARA model has been used previously to define areas of risk by applying cut-offs of fuzzy climate suitability criteria (FCS). An assessment of the distribution of PRs within these cut-offs showed that previous definitions of stable endemicity (FCS \geq 0.75) based on the MARA model may have been overly restrictive. A high proportion (58/137) of studies with high PR values (\geq 25%) characteristic of stable transmission areas were found in areas with FCS values $<$ 0.75.

The investigation of the relationship between transmission and urbanization in Kenya, Uganda and Tanzania, showed that rural studies had a 10% higher infection prevalence compared with urban studies. In addition, there were no studies from urban sites that recorded a PR $>$ 75%. The classification of studies as urban *vs.* rural was challenging and had to be based on inference from an amalgamation of data sources where survey reports did not provide adequate information. There continues to be a need for better operational definitions of urbanization and geographical descriptions of urban extents and for more data from urban areas in order to stratify risk more accurately.

The utility of Fourier processed data for predicting malaria distribution was demonstrated in Chapter 6 where the data were used to improve the historical classifications of malaria. The satellite data were able to predict historical classifications of malaria with an overall accuracy of 89.9% ($kappa=0.775$) with areas of the least intense transmission being predicted with greatest accuracy (96%). In Chapter 7 the strictly selected empirical data were used to train climate data and non-climatic variables, notably urbanization, to provide a more comprehensive map of malaria transmission for East Africa. Surrogates of moisture availability (NDVI) and temperature (LST) were found to be among the principal predictors of malaria transmission with an overall prediction accuracy of 62% and mapped output was improved with the inclusion of urban criteria and ecological zonation.

8.2. The limitations of the *P. falciparum* prevalence training data

Ensuring the reliability and accuracy of PR data was perhaps the most challenging aspect of the study. Several factors affect the accuracy and reliability of individual *P. falciparum* estimates and their comparability between studies including the accuracy and reliability of microscopy, which was not possible to control for in the selection criteria. A pre-requisite for successful modelling using the statistical methods employed in this study is that training data capture the entire range of environmental conditions under which malaria occurs. As outlined in Chapter 3, the prevalence studies in this study were highly clustered, being concentrated in high transmission, densely populated areas while sparsely populated areas such as NE Kenya and south and western Tanzania were poorly

represented. The selection of survey sites is likely to have been based on convenience, the need to answer specific research questions or on proximity to medical or malaria research centres. National surveys would be a more appropriate alternative in order to randomly select data that represent the true range of ecological conditions in East Africa and so avoid the selective nature of retrospective literature sourced data. Truly national sample surveys, however, have rarely been undertaken in any of the three East African countries (Chapter 3).

The clustering of the data also presented the additional problem of spatial correlation. This is important to consider as the discriminant analysis techniques used in the study assume that prevalence observations are independent and spatially correlated data can not be considered independent observations. There are several ways in which future studies can improve on the spatial limitations of the training data including conducting studies in sites that have not been sampled previously or using stratified sampling techniques based on ecological zone strata and obtaining truly spatially random community-based samples.

Very few survey reports gave information on the season during which prevalence studies were conducted. Most reports provided a month of survey and for areas where rainfall patterns are known, it was possible to deduce that most studies were conducted during peak transmission seasons. Where available, such longitudinal data have been averaged but for most studies, the season of survey was not considered due to lack of information provided in the original reports.

Very rarely were details of randomisation methods provided in the reports and the sample selection and analysis has assumed that surveys described as randomly selected were indeed random. For a proportion (88/330) of the selected surveys it was either not absolutely clear that children included were selected at random or samples included a few volunteers. Bias caused by differences in age groups sampled were minimised by restricting samples to include only children under 0-15 years old but even small differences in the weights of age-structured samples might have influenced the final childhood parasite prevalence estimate. Again it was not possible to correct every survey by individual years of age due to the paucity of information in the original reports and this might have influenced final prevalence estimates.

Despite these limitations, the selected data set probably represents the best combined data available for East Africa to-date. There are still many data gaps and although additional sampling from previously un-sampled areas would be ideal, the cost and manpower required for such an exercise makes it prohibitive. For these reasons, malaria mapping will continue to rely on modelling for data poor areas. Although beyond the scope of this study, validation is a crucial step in model development. None of the currently available models have been validated according to field studies and this should be an important step in future modelling efforts.

Figure 8.1. Summary of available malaria maps for East Africa

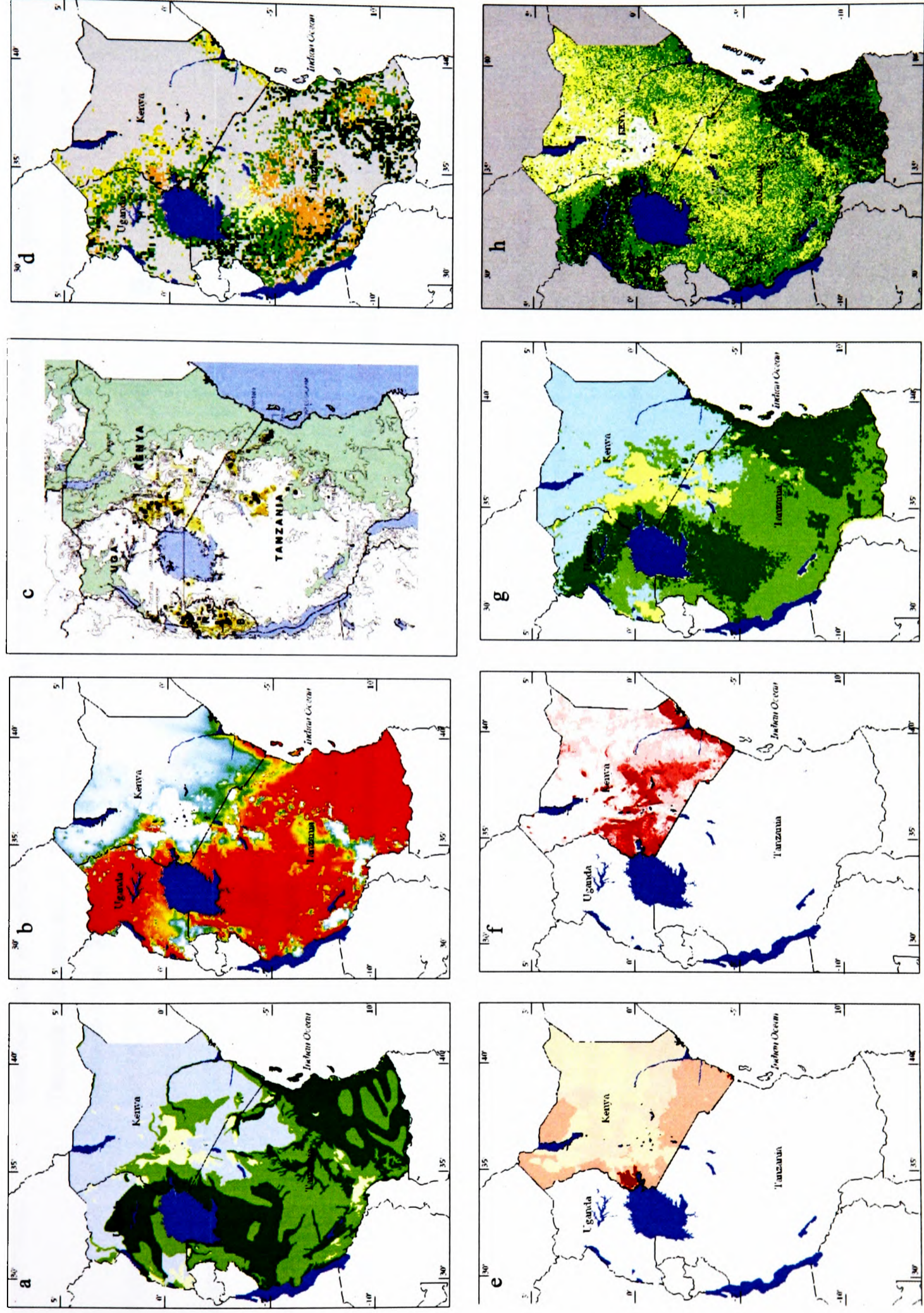


Figure Legends

- Figure 8.1a:** Historical map of malaria distribution for Kenya (Butler, 1959), Uganda (Government of Uganda, 1962) and Tanzania (Government of Tanganyika, 1956). =malaria free; =malaria near water; =transmission for 3-6 months in a year; =transmission for >=6 months in a year
- Figure 8.1b:** MARA FCS model. 0=unsuitable; 1=suitable; FCS=fuzzy climate suitability: =0, =<0.1, =0.1-0.25, =0.25-0.5, =0.5-0.75, =0.75-0.9, =0.9-1.0
- Figure 8.1c:** Epidemic risk model (Highland Malaria Risk project (Cox *et al.*, 1999)). Epidemic risk classes: =low; =moderate; =high; =non-epidemic areas <1000m ASL
- Figure 8.1d:** Satellite-derived predictions of EIR for East Africa (Rogers *et al.*, 2002). EIR range: =0-4; =5-30; =31-89; =90-259; =260-703; =no prediction
- Figure 8.1e:** Modelled predictions of stable endemicity for Kenya (Snow *et al.*, 1998). =high; =moderate; =low; =unstable endemicity
- Figure 8.1f:** Satellite-derived predictions of the malaria seasonality for Kenya (Hay *et al.*, 1998a). =>6 months; =3-6 months; =<3 months transmission; =no transmission
- Figure 8.1g:** Discriminant analysis predictions of malaria transmission intensity in East Africa using satellite-derived surrogates of climate (Omumbo *et al.*, 2002). =malaria free; =malaria near water; =transmission 3-6 months in a year; =transmission >=6 months in a year
- Figure 8.1h:** Empirical model of the PR in East Africa based on 1x1km resolution RS environmental surrogate data. 0<5%; 5-<25%; 25-<75%; >=75%; no prediction; water body.

8.3. A comparison of previous East African malaria maps

As described in Chapter 1, there are several malaria endemicity maps available for East Africa. The different measures used to describe malaria's endemicity include the duration (in months) of periods of transmission within a year or seasonality (Figures 8.1a, 8.1f & 8.1g), the suitability of climate to support transmission ranging from suitable to unsuitable (Figure 8.1b), epidemic prone areas (Figure 8.1c), climate-based predictions of the EIR (Figure 8.1d), predictions of stable endemicity for 4th level administrative units (Figure 8.1e) and the risk of *P. falciparum* infection (Figure 8.1h). A comparison of these maps reveals some striking similarities particularly at the extremes of transmission. All of the maps presented show that East Africa is characterised by a wide range of endemicities with areas of no transmission at one extreme and highly intense, stable transmission at the other extreme. The maps also show that this diversity is particularly marked in Kenya where a major proportion of surface area is represented by areas of low risk. Tanzania is the least diverse with all maps showing that most of the country experiences moderate-high, perennial transmission. Within the broad area of stable endemicity are small restricted areas of lower transmission intensity, notably highland areas in the Usambara Mountain range of North Eastern Tanzania. Similar highland areas are found in the western highlands of Kenya and in South Western Uganda.

Cox *et al.*, (1999) model a single stratum of malaria risk, *i.e.* areas potentially at risk of epidemics (brown areas in Figure 8.1c). Their distribution is similar in all the

transmission maps presented in that areas defined as 'epidemic prone' correspond well with areas defined as low risk or malaria-free on the other maps.

Historical definitions of transmission were used to compare the spatial extents of transmission intensity classes between climate based models (Figure 8.2) and empirical data driven models (Figure 8.3). Correspondence between historical malaria-free areas (Figure 8.2a; area a) and areas of zero fuzzy climate suitability was good. Similar correspondence was found between the historical map and the predictions of the historical map made using satellite-derived climate surrogate data (Figure 8.2b). Areas with a predicted EIR >90 infective bites/person/year also showed good correspondence with historical intense transmission areas (≥ 6 months in a year; Figure 8.2c). Comparison of the historical map with empirical data driven maps for Kenya, however, revealed poor agreement in most areas (Figure 8.3a; areas e) but was good within a few smaller highly intense transmission areas (Figure 8.3a & 8.3b; areas g).

Figure 8.2. Comparisons between historical classifications of malaria and available maps of malaria distribution for East Africa

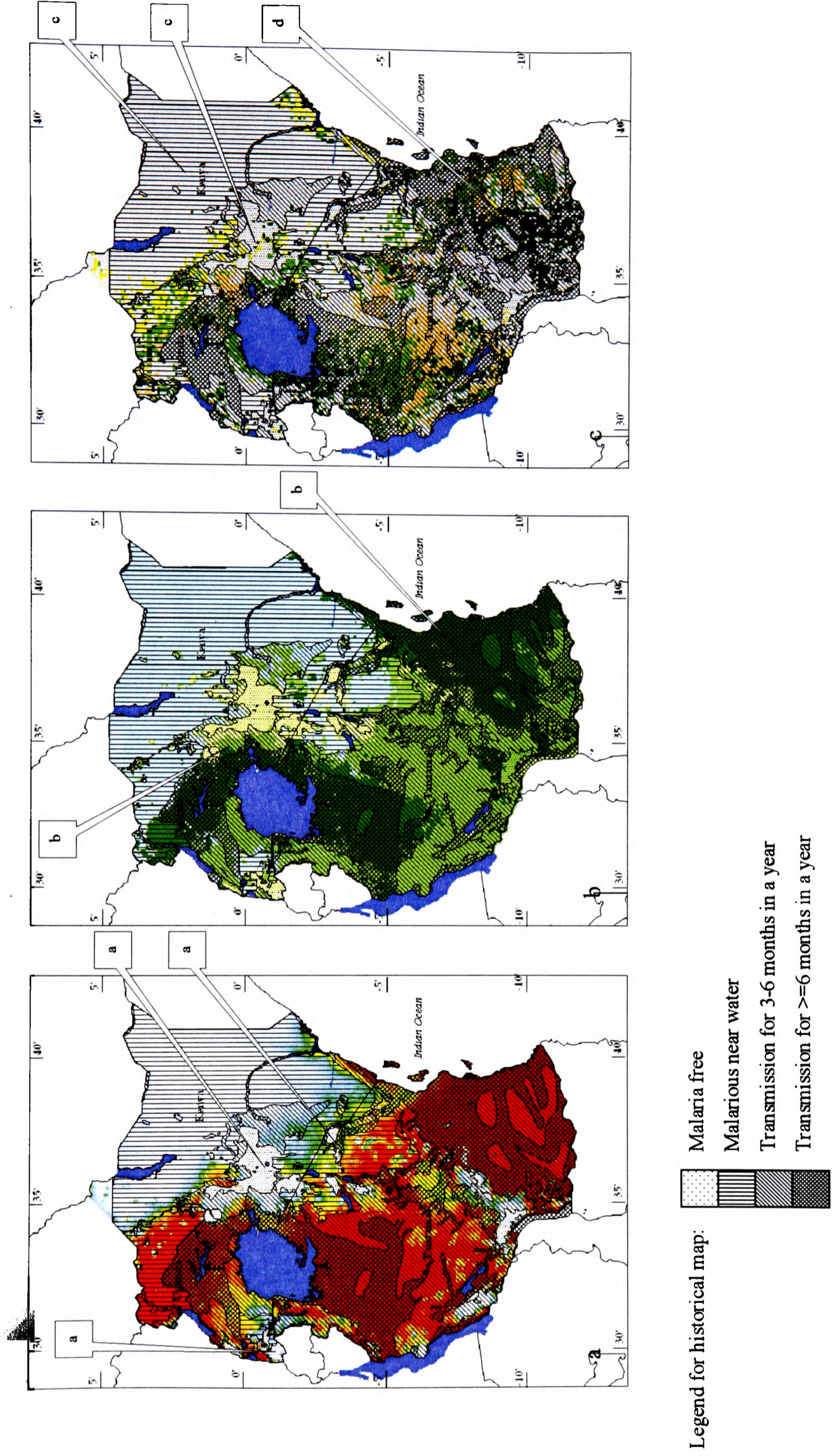
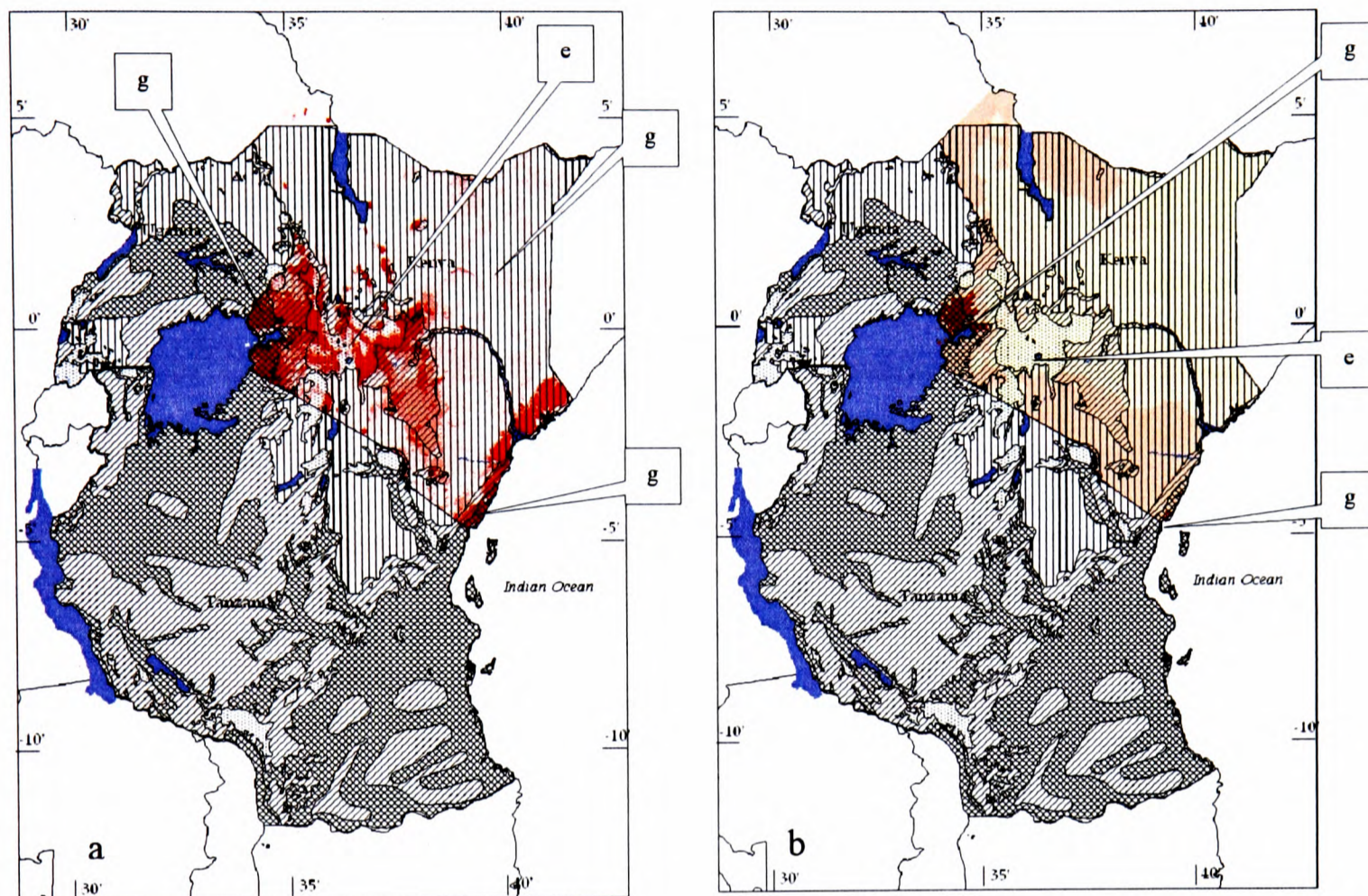
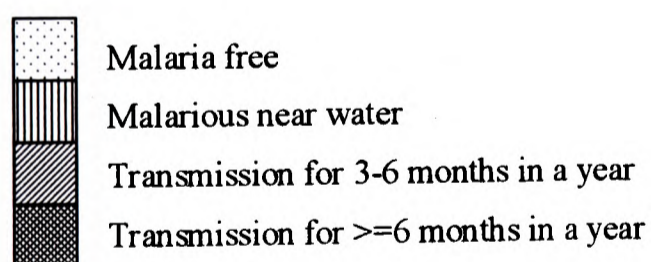


Figure 8.3. Comparisons between historical classifications of malaria and empirical maps of malaria distribution for Kenya



Legend for historical map:



A selection of points for Kenya randomly selected according to the four strata of malaria in the historical map (Figure 8.1a) was used to make comparisons between Figures 8.1b; 8.1e & 8.1g. Each of the comparison maps was reclassified in IDRISI (The IDRISI Project, Clarke University, 1997) according to 4 strata representing malaria free/unstable, low risk, moderate and high transmission. Thresholds of 0; $>0 < 0.25$; $0.25 < 0.75$ and ≥ 0.75 FCS were selected as have been used previously for the fuzzy climate suitability map of Craig *et al.*, (1999). Using a GIS, approximately 100 points were randomly selected from each of the strata “malaria free,” “malaria near water,” “moderate” and “intense” transmission in the historical map (Table 8.1). These points were then used to extract endemicity class values from each of the comparison maps. Overall agreement between the maps varied as follows; 77.3% for Butler, 1959 vs. Omumbo *et al.*, 2002 (Table 8.1a), 64.3% for Butler, 1959 vs. Craig *et al.*, 1998 (Table 8.1b), 43.0% for Butler, 1959 vs. Snow *et al.*, 1998 (Table 8.1c) and 62.0% for Craig *et al.*, 1998 vs. Snow *et al.*, 1998 (Table 8.1d).

The high level of agreement between Butler, 1959 and Omumbo *et al.*, 2002 maps was not surprising as the latter was based on training data derived from the former. A comparison of the geographic extents of transmission categories between maps by Butler, 1959 and Snow *et al.*, 1998 is difficult as the latter was attributed to administrative boundary units and probably contributes to the low agreement between the two maps (Table 8.1c). The agreement between Butler, 1959 and Craig *et al.*, 1998 is notable though. Both maps are solely climate driven and this is perhaps the reason for their similarity (Table 8.1b).

Table 8.1a. Agreement between Butler, 1959 and Omumbo et al., 2002

		Omumbo et al., 2002				
		malaria free	malaria near water	3 - 6 months	>= 6 months	Total
Butler, 1959	malaria free	93	1	3	0	97
	malaria near water	6	92	4	1	103
	3 - 6 months	14	39	41	6	100
	>= 6 months	1	3	13	83	100
	Total	114	135	61	90	400

Overall agreement = 77.3%

Table 8.1b. Agreement between Butler, 1959 and Craig et al., 1999

		Craig et al., 1999				
		zero	>0 - 0.24	0.25 - 0.74	0.75 - 1	Total
Butler, 1959	malaria free	88	7	2	0	97
	malaria near water	2	77	24	0	103
	3 - 6 months	10	52	32	6	100
	>= 6 months	1	2	37	60	100
	Total	101	138	95	66	400

Overall agreement = 64.3%

Table 8.1c. Agreement between Butler, 1959 and Snow et al., 1998

		Snow et al., 1998				
		unstable	low	moderate	high	Total
Butler, 1959	malaria free	85	4	8	0	97
	malaria near water	43	26	34	0	103
	3 - 6 months	19	42	38	1	100
	>= 6 months	0	2	75	23	100
	Total	147	74	155	24	400

Overall agreement = 43.0%

Table 8.1d. Agreement between Craig et al., 1999 and Snow et al., 1998

		Snow et al., 1998				
		unstable	low	moderate	high	Total
Craig et al., 1999	zero	93	2	6	0	101
	>0 - 0.24	53	54	31	0	138
	0.25 - 0.74	1	15	78	1	95
	0.75 - 1	0	3	40	23	66
	Total	147	74	155	24	400

Overall agreement = 62.0%

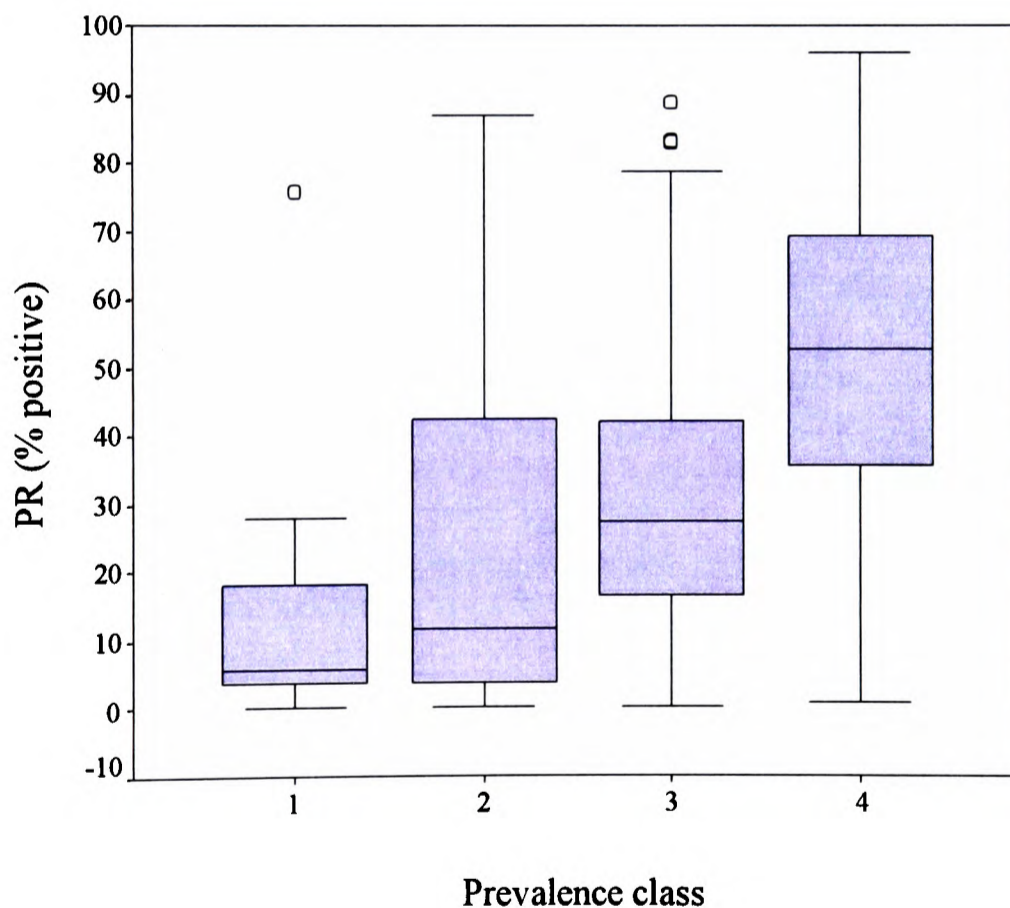
All available maps have used climate data to predict malaria but few (Figure 8.1d-h) have used empirical data to train these climate data sets. Most available empirical data are derived from moderate to intense transmission sites. The predictions of EIR (Figure 8.1d) demonstrate this. EIR data were unavailable for low malaria transmission areas (displayed in grey) resulting in a very patchy distribution map particularly for Kenya where low risk areas predominate.

The high transmission risk areas north of Lake Victoria in Uganda and Kenya as well as the southern part of the Indian Ocean coastline are similar in all four maps. Also similar are the low risk areas of NE Kenya where transmission is limited by the unavailability of mosquito breeding sites for most of the year. These areas are classified as low endemicity (Figure 8.1e), supporting transmission for <3 months in a year (Figure 8.1f), malarious near water (Figure 8.1g) and have been predicted as low transmission risk areas (PR=0- <5%) in this study (Figure 8.1h). The abiotic factors that drive transmission at its extremes are likely to be more distinct than those in areas of moderate transmission. In highland areas, for example, transmission is limited by low temperatures, and in arid areas it is limited by low rainfall. These areas are the least difficult to stratify.

There is considerably more overlap in transmission determinants in areas of intermediate transmission and the boundaries of these vary significantly between the different transmission maps. It is here that the use of empirical data is most important. Although a poor correlation ($R^2=0.28$; $p<0.001$) was found on a regression analysis between the PR data as continuous variable and the categories of satellite data based predictions of the

historical map presented in Figure 8.1g, Figure 8.3 suggests a definite signal in the distribution of the PR across differing levels of malaria transmission intensity. The signal, however is not clear and it is suggested in Chapter 7 that this may be due to the coarse resolution of the PR data used compared with the high spatial resolution of the environmental data sets.

Figure 8.3. Boxplot of PR by malaria prevalence category. 1=malaria-free (N=19), 2=transmission near water (N=56), 3=transmission for 3-<6months a year (N=102), 4=transmission for >=6months a year (N=153). Outliers (open circles) are cases with values between 1.5 and 3 box lengths from the upper edge of the box. The box length is the interquartile range and the line within the box represents the median.



In the absence of ground verification or other gold standard, it is difficult to make meaningful comparisons of these diverse maps. This analysis shows that climate data do

provide consistent information about the potential for transmission but, as revealed by the poorer agreement when empirical data are included, there are other determinants of transmission that climate either captures poorly or not at all. Much has been done in this study to address this deficiency and the result of these efforts is shown in Figure 8.1 h but there are important areas that require further attention.

8.4. Future aims for mapping malaria and suggested steps to make improvements

The significance of developing malaria models will be determined by their usefulness in the prevention and control of malaria. This study has shown that there is a wide scope for developing improved models. Investment could be made in higher resolution climate and empirical malariometric data and the range of ecological databases available expands almost daily. How does one decide what is important? For the malaria programme manager at the district level, high spatial resolution classifications may not be a priority. Information from currently available maps showing areas for which data are not available, on the other hand, may be useful for targeting surveillance studies. Additional data particularly from low transmission areas are still required as these areas remain poorly characterised.

An important application for high resolution data is in mapping urban malaria. The marked heterogeneity of transmission within urban micro-environments is a well-recognised phenomenon, however, characterising these that high resolution satellite data

derived from passive RS systems requires further investment given the dynamic nature of urbanization in SSA (Chapter 5).

The study has iteratively improved our understanding of the determinants of malaria endemicity in East Africa, how these might be modelled and the limitations of existing approaches. It is by no means comprehensive but it is a beginning. It is hoped that areas requiring improvement and validation can be addressed by future studies.

REFERENCES

Abeku T.A., van Oortmarssen G.J., Borsboom G., de Vlas S.J., Habbema J.D.F. (2003). Spatial and temporal variations of malaria epidemic risk in Ethiopia: factors involved and implications. *Acta Tropica*, **87**: 331-340.

Africover (2003). FAO - Africover Eastern Africa Project (GCP/RAF/287/ITA). Food and Agriculture Organization of the United Nations. www.africover.org.

Akogbeto M., Chippaux J.P., Coluzzi M. (1992). Le paludisme urbain cotier à Cotonou (République du Bénin). Étude entomologique. [Coastal urban malaria in Cotonou (Republic of Benin). Entomological study]. *Revue d'Epidémiologie et Santé Publique*, **40**: 233-239.

Allan R. (2003). Complex emergencies: new approaches to malaria control in extreme settings. Gordon Research Conference on the epidemiology of malaria, Queen's College, Oxford.

Alonso P.L., Lindsay S.W., Armstrong J.R., Conteh M., Hill A., David P.H., Fegan G., de Francisco A., Hall A.J., Shenton F.C., Cham K., Greenwood B.M. (1991). The effect of insecticide-treated bed nets on mortality of Gambian children. *Lancet*, **337**: 1499-1502.

Anderson R.M., May R.E. (1992). Infectious diseases of humans: Dynamics and control. Oxford UK, Oxford University Press.

Armengaud M., Louvain M., Diop Mar I. (1962). Etude portant sur 448 cas de paludisme chez l'Africain de la région Dakaraise. *Bulletin de la Société du Medecin Africaine Noire de la Langue Français*, **17**: 167-19.

Armstrong J.A., Bransby-Williams W.R. (1961). The maintenance of a colony of *Anopheles gambiae*, with observations on the effects of changes in temperature. *Bulletin of the World Health Organization*, **24**: 427-435.

Balk D., Pozzani F., Yetman G., Nelson A., Diechmann U. (2004). Methodologies to improve global population estimates in urban and rural areas. New York, U.S.A., CIESIN, Columbia University.

Balls, M.J., Bødker, R., Thomas, C.J., Kisinza, W., Msangeni, H.A. and Lindsay, S.W. (2004). Effect of topography on the risk of malaria infection in the Usambara mountains, Tanzania. *Transactions of the Royal society of Tropical Medicine and Hygiene*, **98**: 400-408.

Bayoh M.N., Thomas C.J., Lindsay S.W. (2001). Mapping distributions of chromosomal forms of *Anopheles gambiae* in West Africa using climate data. *Medical and Veterinary Entomology*, **15**: 267-274.

- Beier J.C., Killeen G.F., Githure J.I. (1999). Entomologic inoculation rates and *Plasmodium falciparum* malaria prevalence in Africa. *American Journal of Tropical Medicine and Hygiene*, **61**: 109-113.
- Beljaev A.E. (2002). Determinants of malaria in the Middle East and North Africa. In: The contextual determinants of malaria. E.A. Casman and H. Dowlatabadi. Washington DC, Resources for the future. 137-166.
- Benasseni, R., Gazin, P., Carnevale, P., Baudon, D. (1987). Le paludisme urbain à Bobo-Dioulasso. 1. Étude de la morbidité palustre. *Cahier d'O.R.S.T.O.M. Série Entomologie Médicale et Parasitologie*, **25**: 165-170.
- Binka F.N., Kubaje A., Adjuik M., Williams L.A., Lengeler C., Maude G.H., Armah G.E., Kajihara B., Adiamah J.H., Smith P.G. (1996). Impact of permethrin impregnated bednets on child mortality in Kassena-Nankana district, Ghana: a randomized controlled trial. *Tropical Medicine and International Health*, **1**: 147-154.
- Björkman A., Hedman P., Brohult J., Willcox M., Diamant I., Pehrsson P.O., Rombo L., Bengtsson E. (1985). Different malaria control activities in an area of Liberia-effects on malariometric parameters. *Annals of Tropical Medicine and Parasitology*, **79**: 239-246.
- Bødker R., Kisinza W., R M., Msangeni H., Lindsay S. (2000). Resurgence of Malaria in the Usambara Mountains, Tanzania, An Epidemic of Drug-Resistant Parasites. *Global Change and Human Health*, **1**: 134-153.
- Bouma M.J., Dye C. (1997). Cycles of malaria associated with El Niño in Venezuela. *Journal of the American Medical Association*, **278**: 1772-1774.
- Boyd D.S., Foody G.M., Curran P.J., Lucas R.M., Honzak M. (1996). An assessment of radiance in Landsat TM middle and thermal infrared wavebands for the detection of tropical forest regeneration. *International Journal of Remote Sensing*, **17**: 249-261.
- Boyd D.S., Ripple W.J. (1997). Potential vegetation indices for determining global forest cover. *International Journal of Remote Sensing*, **18**: 1395-1401.
- Boyd D.S., Curran P.J. (1998). Using remote sensing to reduce uncertainties in the global carbon budget: The potential of radiation acquired in the middle infrared wavelengths. *Remote Sensing Reviews*, **16**: 293-327.
- Boyd M.F. (1930). An introduction to malariology. Cambridge, Massachusetts, Harvard University Press.

- Bouma M.J., Vanderkaay H.J. (1996). The El Niño Southern Oscillation and the historic malaria epidemics on the Indian subcontinent and Sri Lanka: An early warning system for future epidemics? *Tropical Medicine and International Health*, **1**: 86-96.
- Boyd M.F. (1930). An introduction to malariology. Cambridge, Massachusetts, Harvard University Press.
- Boyd M.F. (1949). Epidemiology of malaria: factors related to the intermediate host. In: *Malariology*. Philadelphia, Saunders: 551-607.
- Brabin B.J. (1983). An analysis of malaria in pregnancy in Africa. *Bulletin of the World Health Organization*, **61**: 1005-1016.
- Brabin B., Premji Z., Verhoeff F. (2001). Iron-deficiency anemia: re-examining the nature and magnitude of the public health problem: An analysis of anemia and child mortality. *American Society of Nutritional Sciences*, 636S-648S.
- Brewster D.R., Greenwood B.M. (1993). Seasonal variation of paediatric diseases in the Gambia, West Africa. *Annals of Tropical Paediatrics*, **13**: 133-146.
- Briët O.J.T., Gunawardena D.M., van der Hoek W., Amerasinghe F.P. (2003). Sri Lanka malaria maps. *Malaria Journal*, **2**: 22-32.
- Brooker S., Hay S.I., Issae W., Hall A., Kihamia C., M., Lwambo N., J.S., Wint W., Rogers D.J., Bundy D.A.P. (2001). Predicting the distribution of urinary schistosomiasis in Tanzania using satellite sensor data. *Tropical Medicine and International Health*, **6**: 1-10.
- Brooker S., Hay S.I., Bundy D.A.P. (2002). Tools from ecology: useful for evaluating infection risk models. *Trends in Parasitology*, **18**: 70-74.
- Brooker S., Beasley M., Ndinaromtan M., Madjiouroum E. M., Baboguel M., Djenguinabe E., Hay S. I., Bundy D.A.P. (2002). Use of remote sensing and a geographical information system in a national helminth control programme in Chad. *Bulletin of the World Health Organization*, **80**: 783-789.
- Brown V., Isaak M.A., Rossi M., Barboza P., Paugam A. (1998). Epidemic of malaria in north-eastern Kenya. *Lancet*, **352**: 1356-1357.
- Bruce-Chwatt L.J. (1952). Malaria in Nigeria: Morbidity and mortality due to malaria. *West African Medical Journal*, **1**: 102-110.
- Bruce-Chwatt L.J. (1984). Lessons learned from applied field research activities in Africa during the malaria eradication era. *Bulletin of the World Health Organization*, **62**: 19-29.

Bureau of Statistics Tanzania (2002). 2002 population and housing census. <http://www.tanzania.go.tz/census/index.html>.

Butler R.J. (1959). Atlas of Kenya: A comprehensive series of new and authenticated maps prepared from the national survey and other government sources with gazetteer and notes on pronunciation and spelling. Nairobi, Kenya, The Survey of Kenya.

Carme B., Guillo du Bodan H., Lalleman M. (1992). Infant and child mortality and malaria in the Congo: The trend in the suburbs of Brazzaville between 1981 and 1998. *Tropical Medicine and Parasitology*, **43**: 177-180.

Carme B. (1996). Low malaria mortality among children and high rates of *Plasmodium falciparum* inoculation: A Congolese reality in the 1980s. *Parasitology Today*, **12**: 206-207.

Cattani J.A., Moir J.S., Gibson F.D., Ginny M., Paino J., Davidson W., Alpers M.P. (1986) Small-area variations in the epidemiology of malaria in Madang Province. *Papua New Guinea Medical Journal*, **29**:11-17.

Central Bureau of Statistics, 1989. Kenya population census – 1989. Volume 1. Central Bureau of Statistics, Ministry of Economic Planning and Development, Government of Kenya, Nairobi, 1989.

Central Bureau of Statistics (2001). 1999 population and housing census: counting our people for development. Volume 1.1 Population distribution by administrative and urban centres. Ministry of Finance & Planning, Nairobi, Kenya. 2001.

Central Bureau of Statistics (2003). *Kenya Demographic & Health Survey 2003: Preliminary Report*. Central Bureau of Statistics, Nairobi, Kenya.

CIESIN/CIAT (2004). Gridded Population of the World (GPW), Version 3. (<http://beta.sedac.ciesin.columbia.edu/gpw>), Columbia University, Centre for International Earth Science Information Network (CIESIN), Columbia University Palisades, New York; Centro Internacional de Agricultura Tropical (CIAT).

CIESIN/IPFRI/CIAT (2004). Global Rural-Urban Mapping Project (GRUMP): Gridded Population of the World, version 3, with urban reallocation (GPW-UR). Palisades, New York, CIESIN, Columbia University, Centre for International Earth Science Information Network (CIESIN), Columbia University; International Food Policy Research Institute (IPFRI), the World Bank; and Centro Internacional de Agricultura Tropical (CIAT). 2004.

Charlwood J.D., Kihonda J., Sama S., Billingsley P.F., Hadji H., Verhave J., Lyimo E., Luttukhuizen P., Smith T. (1995). The rise and fall of *Anopheles*

arabiensis (Diptera: *Culicidae*) in a Tanzanian village. *Bulletin of Entomological Research*, **85**: 37-44.

Chinery W.A. (1984). Effects of ecological changes on the malaria vectors *Anopheles funestus* and *Anopheles gambiae* complex of mosquitoes in Accra, Ghana. *Journal of Tropical Medicine and Hygiene*, **87**: 75-81.

Chinery W.A. (1997). An ecological observation on malaria parasite prevalence in Ghana. *Malaria and Infectious Diseases in Africa*, **4**: 34-40.

Clarke K., Osleeb J., Sherry J., Meert J., Larsson R. (1991). The use of remote sensing and geographic information systems in UNICEF's dracunculiasis (Guinea worm) eradication effort. *Preventive Veterinary Medicine*, **11**: 229-235.

Clarke University Graduate School of Geography (1997). IDRISI Version 2.0 for Windows, Worcester, Massachusetts, USA. <http://www.idrisi.clarku.edu>.

Clyde D.F. (1967). *Malaria in Tanzania*. Oxford University Press, London, New York.

Coene J. (1993). Malaria in urban and rural Kinshasa: the entomological input. *Medical and Veterinary Entomology*, **7**: 127-137.

Coetzee M., Hunt R.H., Braack L.E.O., Davidson D. (1993). Distribution of mosquitoes belonging to the *Anopheles gambiae* complex, including malaria vectors, south of latitude 15°S. *South African Journal of Science*, **89**: 227-231.

Coetzee M., Craig M.H., le Sueur D. (2000). Distribution of African malaria mosquitoes belonging to the *Anopheles gambiae* complex. *Parasitology Today*, **16**: 74-77.

Cohen J. (1960). A co-efficient of agreement for nominal scale. *Educational and Psychological Measurements*, **20**: 37-46.

Coluzzi M., Sabatini A., Petrarca V., Di Deco M.A. (1979). Chromosomal differentiation and adaptation to human environments in the *Anopheles gambiae* complex. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, **73**: 483-497.

Coluzzi M. (1984). Heterogeneities of the malaria vectorial system in tropical Africa and their significance in malaria epidemiology and control. *Bulletin of the World Health Organization*, **62**:107-113.

Coluzzi M. (1994). Malaria and the Afrotropical ecosystems: Impact of man-made environmental changes. *Parassitologia*, **36**: 223-227.

Coluzzi M., Bradley D. (1999). The malaria challenge after one hundred years of malariology: Papers from the Malariology Centenary Conference, Accademia Nazionale del Lincei, Roma, 16-19 November, 1998. *Parassitologia*, **41**.

Connor S.J. (1999). Malaria in Africa: the view from space. *Biologist*, **46**: 22-25.

Cooper D.I., Asrar G. (1989). Evaluating atmospheric correction models for retrieving surface temperatures from the AVHRR over a tall grass prairie. *Remote Sensing of the Environment*, **27**: 93-102.

Couto P. (2003). Assessing the accuracy of spatial simulation models. *Ecological Modelling*, **167**: 181-198.

Cox J., Craig M.H., Le Sueur D., Sharp B. (1999). MARA/HIMAL technical report 1999: Mapping malaria risk in the highlands of Africa, London School of Hygiene and Tropical Medicine, UK and Medical Research Council, Durban South Africa.

Cox J.S.H., Mouchet J., Bradley D.J. (2002). Determinants of malaria in sub-Saharan Africa. In: The contextual determinants of malaria. EA Casman and H Dowlatabadi. Washington D.C., Resources for the Future. 167-186.

Cracknell A.P., Hayes L.W.B. (1991). Introduction to Remote Sensing. London, Taylor and Francis.

Cracknell A.P. (1997). Channel 3. The neglected channel. The Advanced Very High Resolution Radiometer (AVHRR). London, Taylor & Francis: 321-342.

Craig M.H., Snow R.W., Le Sueur D. (1999). A climate-based distribution model of malaria transmission in sub-Saharan Africa. *Parasitology Today*, **15**: 105-111.

Crowe S. (1997). Malaria outbreak hits refugees in Tanzania. *Lancet*, **350**: 41.

Dale P.E.R., Ritchie S.A., Territo B.M., Morris C.D., Muhar A., Kay B.H. (1998). An overview of remote sensing and GIS for surveillance of mosquito vector habitats and risk assessment. *Journal of Vector Ecology*, **23**: 54-61.

D'Alessandro U., Olayeye B.O., McGuire W., Langerock P., Bennett S., Aikins M.K., Thomson M.C., Cham M.K., Cham B.A., Greenwood B.M. (1995). Mortality and morbidity from malaria in Gambian children after introduction of an impregnated bednet programme. *Lancet*, **345**: 479-483.

Davidson G., Lane J. (1981). Distribution maps for the *Anopheles gambiae* complex. In: Mosquito studies at the London School of Hygiene and Tropical Medicine. Progress report No. 40. London, London School of Hygiene and Tropical Medicine.

Detinova T.S. (1962). Age grouping methods in Diptera of medical importance, with special reference to some vectors of malaria. Geneva, World Health Organization. 122-150.

Directorate of Overseas Surveys, 1971. *1:50,000 topographical maps. Series Y731. Survey of Kenya, Nairobi, Kenya.*

Dolo G., Sissoko M.S., Dao A., Sagara I., Traoré S.F., Sissoko M., Bouaré M., Dicko A., Sogoba N., Dembélé O., Niar, Bagayoko M. (1997). Impact of irrigated rice cultivation on malaria transmission in Niono, Mali. *American Journal of Tropical Medicine and Hygiene*, **57**(Supplement): 183-184.

Drakeley C & Reyburn H (2002). Unpublished data. *Personal communication.*

Draper C.C. & Smith A. (1957). Malaria in the Pare area of N.E. Tanganyika. Part I: Epidemiology. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, **51**: 137-151.

Draper C.C. & Smith A. (1960). Malaria in the Pare area of Tanganyika. Part II: Effects of three years' spraying of huts with Dieldrin. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, **54**: 342-357.

Draper C.C., Lelijveld J.L.M., Matola Y.G., White G.B. (1972). Malaria in the Pare area of Tanzania. IV: Malaria in the human population 11 years after the suspension of residual insecticide spraying with special reference to the serological findings. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, **66**: 905-912.

Dugdale G., McDougall V., Thorne V. (1995). The TAMSAT method for estimating rainfall over Africa and its implications. RSS95: Proceedings of the 21st annual conference of the Remote Sensing Society, 11-14 September, Southampton, England: 773-780.

Dutta H.M. & Dutt A.K. (1978). Malaria ecology: a global perspective. *Social Science in Medicine Part D (Medical Geography)*, **12**: 69-84.

Dye C. & Reiter P. (2000). Climate change and malaria - Temperatures without fevers? *Science*, **289**: 1697-1698.

Eidenshink J.C., Faundeen J.L. (1994). 1-Km AVHRR global land dataset: first stages in implementation. *International Journal of Remote Sensing*, **15**: 3443-3462.

El Gaddal A.A. (1990). Malaria control in the Gezira irrigated area. Wad Medani, Sudan, Blue Nile Health Project.

El Gaddal A.A. (1991). The experience of the Blue Nile Health Project in the control of malaria and other water associated diseases. In: *Malaria and development in Africa: A cross sectional approach*, American Association for the Advancement of Science, Sub-Saharan Africa Programme.

Ellman R., Maxwell C.A., Finch R., Shayo D. (1998). Malaria and anaemia at different altitudes in the Muheza district of Tanzania: childhood morbidity in relation to level of exposure to infection. *Annals of Tropical Medicine and Parasitology*, **92**: 741-753.

El Sayed B.B., Arnot D.E., Mukhtar M.M., Baraka O.Z., Dafalla A.A., Elnaiem D.E. A., Nugud A.H.D. (2000). A study of the urban malaria problem in Khartoum. *Acta Tropica*, **75**: 163-171.

EROS Data Center (2002). Land Processes Distributed Active Archive Center, <http://edcdaac.usgs.gov/gtopo30/README/html>.

Eyles D.E. & Young M.D. (1951). The duration of untreated or inadequately treated *Plasmodium falciparum* infections in the human host. *Journal of the National Malaria Society*, **10**: 327-326.

Fantini B. (1999). The concept of specificity and the Italian contribution to the discovery of the malaria transmission cycle. *Parassitologia*, **41**: 39-47.

Fondjo, E., Robert, V., Le Goff, G., Toto, J.C., Carnevale, P. (1992). Le paludisme urbain à Yaoundé (Cameroun). 2. Étude entomologique dans deux quartiers peu urbanisés. *Bulletin de la Société de Pathologie Exotique*, **85**: 57-63.

Fontaine R.E., Najjar A.E., Prince J.S. (1961). The 1958 malaria epidemic in Ethiopia. *American Journal of Tropical Medicine and Hygiene*, **10**: 795-803.

Food and Agriculture Organization (FAO) (1978). Report on the agro-ecological zones project, Volumel: methodology and results for Africa, World Soil Resources Report N° 48: 32-41.

Gallup J., Sachs J. (2001). The economic burden of malaria. *American Journal of Tropical Medicine and Hygiene*, **64**: 85-96.

Gardiner C.N., Biggar R.J., W. C., Nkrumah F. (1984). Malaria in urban and rural areas of southern Ghana: a survey of parasitaemia and of anti-malarial practice. *Journal of Tropical Pediatrics*, **30**: 296-299.

Garrett-Jones C. (1964). Prognosis for interruption of malaria transmission through assessment of the mosquito's vectorial capacity. *Nature*, **204**: 1173-1175.

GDE Systems Inc. (1995). Populated places: GeoName Digital Gazetteer v.1 (CD-ROM). Geographic Information Department (MZ 1211-M), GDE Systems Inc., P.O. Box 509009, San Diego, California.

Gernaat H.B., Dechering W.H., Voorhoeve H.W. (1998). Clinical epidemiology of paediatric disease at Nchelenge, north-east Zambia. *Annals of Tropical Paediatrics*, **18**: 129-38.

Gesch D.B., Larson K.S. (2000). Techniques for development of global 1-kilometer digital elevation models, United States Geological Survey - EROS data centre, Sioux Falls, South Dakota. (<http://ededaac.usgs.gov/gtopo30/papers/geschtxt.html>).

Ghebreyesus T.A., Haile M., Getachew A.M., Alemayehu T., Witten K.H., Medhin A., Yohannes M., Asgedom Y., Yeebiyo Y., Lindsay S.W., Byass P. (1998). Pilot studies on the possible effects on malaria of small-scale irrigation dams in Tigray regional state, Ethiopia. *Journal of Public Health Medicine*, **20**: 238-240.

Ghebreyesus T.A., Haile M., Witten K.H., Getachew A., Yohannes A.M., Yohannes M., Teklehaimanot H.D., Lindsay S.W., Byass P. (1999). Incidence of malaria among children living near dams in northern Ethiopia: community based incidence study. *British Medical Journal*, **319**: 663-666.

Gill C.A. (1914). The history and geography of fulminant malaria. *Indian Journal of Medical Research*, **2**: 268-314.

Gill C.A. (1920). The rôle of meteorology in malaria. *Indian Journal of Medical Research*, **8**: 633-693.

Gill C.A. (1936). Some points in the epidemiology of malaria arising out of the study of malaria epidemic in Ceylon in 1934-35. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, **29**: 427-480.

Gill C.A. (1938). The seasonal periodicity of malaria and the mechanism of the epidemic wave. London, UK, J & A Churchill.

Gilles H.M. (1993a). Epidemiology of malaria. In: Bruce-Chwatt's essential malariology. Oxford University Press., London, Boston, Melbourne, Auckland, Edward Arnold: 124-163.

Gilles H.M. (1993b). Diagnostic methods in malaria. In: Bruce Chwatt's Essential Malariology. HM Gilles HM and DA Warrell Eds. 78-95.

Gillies M.T., de Meillon B. (1968). The Anophelinae of Africa south of the Sahara. Johannesburg, South Africa, *Publications of the South African Institute for Medical Research*, **54**: 213.

Gillies M.T., Coetzee M. (1987). A supplement to the Anophelinae of Africa south of The Sahara (Afrotropical region). Johannesburg, *Publications of the South African Institute for Medical Research*, **55**: 1-43.

Githeko, A.K., Mbogo C.N.M., Curtis C.F., Lines J., Lengeler C. (1996). Entomological monitoring of large-scale vector-control interventions. *Parasitology Today*, **12**: 127-128.

Githeko A.K., Service M.W., Mbogo C.M., Atieli F., Juma F. (1993). *Plasmodium falciparum* sporozoite and entomological inoculation rates at the Ahero rice irrigation scheme and the Miwani sugar-belt in Western Kenya. *Annals of Tropical Medicine and Parasitology*, **87**: 379-391.

Goetz S., Prince S., Small J. (2000). Advances in satellite remote sensing of environmental variables for epidemiological applications. In: *Advances in Parasitology*. London, Academic Press Ltd. **47**: 289-307.

Government of Kenya (1970). National Atlas of Kenya. Nairobi, Survey of Kenya: 44.

Government of Tanganyika (1956). Atlas of Tanganyika, East Africa. Dar es Salaam, Government Press, Tanganyika.

Government of Uganda (1962). Atlas of Uganda. Uganda, Department of Lands and Surveys.

Granja A.C., Machungo, F., Gomes A., Bergstorm, S. Brabin B. (1998). Malaria related maternal mortality in urban Mozambique. *Annals of Tropical Medicine and Parasitology*, **52**: 257-263.

Green P.E. (2002). Two-group discrimination and classification. In: *Analysing multi-variate data*. Hinsdale, Illinois, Dryden press. 142-173.

Green R.M., Hay S.I. (2002). The potential of Pathfinder AVHRR data for providing surrogate climatic variables across Africa and Europe for epidemiological applications. *Remote Sensing of Environment*, **79**: 166-175.

Greenwood B.M. (1990). Populations at risk. *Parasitology Today*, **6**: 188.

Greenwood A.M., Armstrong J.R.M., Byass P., Snow R.W., Greenwood B.M. (1992). Malaria chemoprophylaxis, birth weight and child survival. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, **86**: 483-485.

Guerin P.J., Olliaro P., Nosten F., Druilhe P., Laxminarayan R., Binka F., Kilama W.L., Ford N., White N.J. (2002). Malaria: current status of control, diagnosis, treatment, and a proposed agenda for research and development. *Lancet Infectious Diseases*, **2**: 564-573.

Guyatt H.L., Snow R.W. (2001a). Malaria in pregnancy as an indirect cause of infant mortality in sub-Saharan Africa. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, **95**: 569-576.

Guyatt H.L., Snow R.W. (2001b). The epidemiology and burden of *Plasmodium falciparum* - related anaemia among pregnant women in sub-Saharan Africa. *American Journal of Tropical Medicine and Hygiene*, **64**: 36-44.

Hagmann R., Charlwood J.D., Gil V., Ferreira C., do Rosário V., Smith T.A. (2003). Malaria and its possible control on the island of Príncipe. *Malaria Journal*, **2**: 15-23.

Halloran M.E. (2001). Concepts of transmission and dynamics. In: Epidemiologic methods for the study of infectious diseases. JC Thomas and DJ Weber Eds. New York, Oxford University Press.

Hamad A.A., Nugud A.E.H.D., Giha H.A., Abdel-Muhsin A.-M.A., Satti G.M.H., Theander T.G., Creasey A.M., Babiker H.A., Elnaiem D.A. (2002). A marked seasonality of malaria transmission in two rural sites in eastern Sudan. *Acta Tropica*, **83**: 71-82.

Harpham, T., Odero W., Walker A.R.P., Sareli P. (1997). Urbanization and health in transition. *Lancet*, **349**: 11-13.

Hastings D.A., Dunbar P.K. (1998). Development and assessment of the Globe One-km Base Elevation Digital Elevation Model. *International Society of Photogrammetry and Remote Sensing Archives*, **32**: 218-221.

Hay S.I., Tucker C.J., Rogers D.J., Packer M.J. (1996). Remotely sensed surrogates of meteorological data for the study of the distribution and abundance of arthropod vectors of disease. *Annals of Tropical Medicine and Parasitology*, **90**: 1-19.

Hay S.I., Packer M.J., Rogers D.J. (1997). The impact of remote sensing on the study and control of invertebrate intermediate hosts and vectors for disease. *International Journal of Remote Sensing*, **18**: 2899-2930.

Hay S.I., Snow R.W., Rogers D.J. (1998a). From predicting mosquito habitat to malaria seasons using remotely sensed data: Practice, problems and perspectives. *Parasitology Today*, **14**: 306-313.

Hay S.I., Snow R.W., Rogers D.J. (1998b). Predicting malaria seasons in Kenya using multi-temporal meteorological satellite sensor data. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, **92**: 12-20.

Hay S.I., Snow R.W., Rogers D.J. (1998c). From predicting mosquito habitat to malaria seasons using remotely sensed data: Practice, problems and perspectives. *Parasitology Today*, **14**: 306-313.

Hay S.I., Lennon J.J. (1999). Deriving meteorological variables across Africa for the study and control of vector-borne disease: a comparison of remote sensing and spatial interpolation of climate. *Tropical Medicine and International Health*, **4**: 58-71.

Hay S.I. (2000). An overview of remote sensing and geodesy for epidemiology and public health application. In: *Advances in Parasitology*. SI Hay, SE Randolph and DJ Rogers Eds. London, Academic Press Ltd. **47**: 1-35.

Hay S.I., Omumbo J.A., Craig M.H., Snow R.W. (2000a). Earth observation, geographic information systems and *Plasmodium falciparum* malaria in sub-Saharan Africa. In: *Advances in Parasitology*, Academic Press Ltd, London. **47**: 173-215.

Hay S.I., Rogers D.J., Toomer J.F., Snow R.W. (2000b). Annual *Plasmodium falciparum* entomological inoculation rates (EIR) across Africa: literature survey, Internet access and review. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, **94**: 113-127.

Hay S.I., Cox J., Rogers D.J., Randolph S.E., Stern D.I., Shanks D.G., Myers M.F., Snow R.W. (2002a). Climate change and the resurgence of malaria in the East African highlands. *Nature*, **415**: 905 - 909.

Hay S.I., Rogers D.J., Randolph S.E., Stern D.I., Cox J., Shanks G.D., Snow R.W. (2002b). Hot topic or hot air? Climate change and malaria resurgence in East African highlands [Opinion]. *Trends in Parasitology*, **18**: 530-534.

Hay S.I., Simba M., Busolo M., Noor A.M., Guyatt H.L., Ochola S.A., Snow R.W. (2002c). Defining and detecting malaria epidemics in the highlands of Western Kenya. *Emerging Infectious Diseases*, **8**: 555-562.

Hay S.I., Were E.C., Renshaw M., Noor A.M., Ochola S.A., Olusanmi I., Nicholas A., Snow R.W. (2003). Forecasting, warning, and detection of malaria epidemics: a case study. *Lancet*, **361**: 1705-1706.

Hay S.I., Guerra C., Tatem A., Noor A.M., Snow R.W. (2004). The global distribution and population at risk of malaria: Past, present, and future. *Lancet Infectious Diseases*, **4**: 327-336.

Hay SI, Guerra CA, Tatem AJ, Aitkinson PM & Snow RW (2005a). Urbanization, malaria transmission and disease burden in Africa. *Nature Reviews Microbiology*, in press.

Hay, S.I., Noor, A.M., Nelson, A., Tatem, A.J. (2005b). Demography for epidemiology: the precision of large-area human population maps. *International Journal of Epidemiology* (in press).

Hedberg K., Shaffer N., Davachi F., Hightower A., Lyamba B., Paluku K.M., Nguyen-Dinh P., Breman J.G. (1993). *Plasmodium falciparum* associated a anaemia in children at a large urban hospital in Zaire. *American Journal of Tropical Medicine and Hygiene*, **48**: 365-371.

Heisch R.B., Harper J.O. (1949). An epidemic of malaria in the Kenya highlands transmitted by *Anopheles funestus*. *Journal of Tropical Medicine and Hygiene*, **52**: 187-190.

Henry M., Rogier C., Nzeyimana I., Assi S., Dossou-Yovo J., Audibert M., Mathonnat J., Keundjian A., Akodo E., Teuscher T., Carnevale P. (2003). Inland valley rice production systems and malaria infection and disease in the savannah of Côte d'Ivoire. *Tropical Medicine and International Health*, **8**: 449-458.

Holding P.A. & Snow R.W. (2001). Impact of *Plasmodium falciparum* malaria on performance and learning: review of the evidence. *American Journal of Tropical Medicine and Hygiene*, **64**: 68-75.

Holding P.A & Kitsao-Wekulo P.K. (2004). Describing the burden of malaria on child development: what should we be measuring and how should we be measuring it? *American Journal of Tropical Medicine and Hygiene*, **71**: 71-79.

Holtz, T.H., Marum, L.H., Mkandala, C., Chizani, N., Roberts, J.M., Macheso, A., Parise, M.E., Kachur, S.P. (2002). Insecticide-treated bed net use, anaemia and malaria parasitaemia in Blantyre, Malawi. *Tropical Medicine & International Health*, **7**: 220-230.

Houghton J.T., Meira Filho L.G., Callander B.A., Harris N., Kattenberg A., Maskell K. (1996). Climate change 1995. The science of climate change. Contribution of WGI to the second assessment report of the Intergovernmental Panel on Climate Change, Cambridge University Press.

Howard S.C., Omumbo J.A., Nevill C.G., Some E.S., Donnelly C.A., Snow R.W. (2000). Evidence for a mass community effect of insecticide-treated bednets on the incidence of malaria on the Kenyan coast. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, **94**: 357-360.

- Hume J.C.C., Lyons E.J., Day K.P. (2003). Human migration, mosquitoes and the evolution of *Plasmodium falciparum*. *Trends in Parasitology*, **19**: 144-149.
- Hutchinson M.F., Nix H.A., McMahon J.P., Ord K.D. (1995). Africa: a topographic and climatic database for Africa, Centre for Resource and Environmental Studies, Australian National University, Canberra, Australia.
- Hutchinson M.F., Nix H.A., McMahon J.P., Ord K.D. (1996). The development of a topographic and climate database for Africa. [Available on-line from http://www.ncgia.ucsb.edu/conf/SANTA_FE_CD-ROM/sf_papers/hutchinson_michael_africa/africa.html].
- Ijumba J.N. & Lindsay S.W. (2001). Impact of irrigation on malaria in Africa: paddies paradox. *Medical and Veterinary Entomology*, **15**: 1-11.
- Ijumba J., Shenton F., Clarke S., Mosha F., Lindsay S. (2002a). Irrigated crop production is associated with less malaria than traditional agricultural practices in Tanzania. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, **96**: 476-480.
- Ijumba J.N., Mosha F.W., Lindsay S.W. (2002b). Malaria transmission risk variations derived from different agricultural practices in an irrigated area of northern Tanzania. *Medical and Veterinary Entomology*, **16**: 28-38.
- Isaaks E.H. & Srivastava R.M. (1989). An Introduction to Applied Geostatistics, Oxford University Press.
- James M.E., Kalluri S.N.V. (1994). The Pathfinder AVHRR land data set - an improved coarse resolution data set for terrestrial monitoring. *International Journal of Remote Sensing*, **15**: 3347-3363.
- Jensen JR (1996). Thematic information extraction: image classification. Land-use classification map accuracy assessment. In: Introductory digital image processing. P Hall. Upper Saddle River, New Jersey, Prentice Hall. 247-256.
- Jepson W.F., Moutia A., Courtois C. (1947). The malaria problem in Mauritius: the bionomics of Mauritian anophelines. *Bulletin of Entomological Research*, **38**: 177-208.
- Joy D.A., Feng X., Mu J., Furuya T., Chotivanich K., Kretzli A.U., Ho M., Wang A., White N.J., Suh E., Beerli P., Su X. (2003). Early origin and recent expansion of *Plasmodium falciparum*. *Science*, **300**: 318-321.
- Justice C.O., Townshend J.R.G., Holben B.N., Tucker C.J. (1985). Analysis of the phenology of global vegetation using meteorological satellite data. *International Journal of Remote Sensing*, **6**: 1271-1318.

Kerber A.G., Schutt J.B. (1986). Utility of AVHRR channel 3 and 4 in land-cover mapping. *Photogrammetric Engineering and Remote Sensing*, **52**: 1877-1883.

Kilian A.H.D (1995). Malaria control in Kabarole and Bundibugyo Districts-Western Uganda. International Support Programme for Malaria Control in the Framework of PHL-University of Munich and German Agency for Technical Co-operation (GTZ).

Kilian A.H.D., Langi P., Talisuna A., Kabagambe G. (1999). Rainfall pattern, El Niño and malaria in Uganda. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, **93**: 22-23.

Kirkwood B.R. (1988). Essentials of medical statistics. Oxford, Blackwell Science.

Kleinschmidt I., Bagayoko M., Clarke G.P.Y., Craig M.H., Le Sueur D. (2000). A spatial statistical approach to malaria mapping. *International Journal of Epidemiology*, **29**: 355-361.

Kleinschmidt I., Omumbo J.A., Briët O., van de Giesen N., Sogoba N., Mensah N., Windmeijer P., Moussa M., Teuscher T. (2001a). An empirical malaria distribution map for West Africa. *Tropical Medicine and International Health*, **6**: 779-786.

Kleinschmidt I., Sharp B.L., Clarke G., Curtis B., Fraser C. (2001b). Use of Generalised Linear Mixed Models in the spatial analysis of small area malaria incidence rates in KwaZulu Natal, South Africa. *American Journal of Epidemiology*, **153**: 1213-1221.

Kleinschmidt I., Sharp B., Mueller I., Vounatsou P. (2002). Rise in malaria incidence rates in South Africa: A small area spatial analysis of variation in time trends. *American Journal of Epidemiology*, **155**: 257-264.

Koram K.A., Owusu-Agyei S., Fryauff D.J., Anto F., Atuguba F., Hodgson A., Hoffman S.L., Nkrumah F.K. (2003). Seasonal profiles of malaria infection, anaemia, and bednet use among age groups and communities in northern Ghana. *Tropical Medicine and International Health*, **8**: 793-802.

Korenromp E.L., Williams B.G., Gouws E., Dye C., Snow R.W. (2003). Measurement of trends in childhood malaria mortality in Africa: an assessment of progress toward targets based on verbal autopsy. *Lancet Infectious Diseases*, **3**: 349-358.

Landis J., Koch G. (1977). The measurement of observer agreement for categorical data. *Biometrics*, **33**: 159-174.

- Lelijveld J.L.M. (1971). Sero-epidemiological studies of malaria in Tanzania. PhD thesis, University of Nijmegen. Cited by: Cox J., Craig M.H., Le Sueur D., Sharp B. (1999). MARA/HIMAL technical report: Mapping malaria risk in the highlands of Africa, London School of Hygiene and Tropical Medicine, UK and Medical Research Council, Durban South Africa.
- Lengeler C., Armstrong-Schellenberg J., D'Alessandro U., Binka F.N., Cattani J. (1998). Relative versus absolute risk of dying reduction after using insecticide-treated nets for malaria control in Africa. *Tropical Medicine and International Health*, **3**: 286-290.
- Lepeš T. (1974). Present status of the global malaria eradication programme and prospects for the future. *Journal of Tropical Medicine and Hygiene*, **77**: 47-53.
- Levine R.S., Peterson A.T., Benedict M.Q. (2004). Geographic and ecologic distributions of the *Anopheles gambiae* complex predicted using a genetic algorithm. *American Journal of Tropical Medicine and Hygiene*, **70**: 105-109.
- Lindblade K.A., Walker E.D., Onapa A.W., Katungu J., Wilson M.L. (1999). Highland malaria in Uganda: Prospective analysis of an epidemic associated with El Niño. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, **93**: 480-487.
- Lindblade K.A., Walker E.D., Onapa A.W., Katungu J., Wilson M.L. (2000). Land use change alters malaria transmission parameters by modifying temperature in a highland area of Uganda. *Tropical Medicine and International Health*, **5**: 263-274.
- Lindsay S.W. (1990). Malaria in a peri-urban area of the Gambia. *Annals of Tropical Medicine and Parasitology*, **84**: 553-562.
- Lindsay S.W. & Birley M.H. (1996). Climate change and malaria transmission. *Annals of Tropical Medicine and Parasitology*, **90**: 573-588.
- Lindsay S.W. & Martens W.J.M. (1998). Malaria in the African highlands: past, present and future. *Bulletin of the World Health Organization*, **76**: 33 - 45.
- Lindsay S.W., Parson L., Thomas C.J. (1998). Mapping the range and relative abundance of the two principal African malaria vectors, *Anopheles gambiae* sensu stricto and *An. arabiensis*, using climate data. *Proceedings of the Royal Society of London: Biological Sciences*, **265**: 847-854.
- Lindsay S.W., Bødker R., Malima R., Msangeni H.A., Kisinza W. (2000). Effect of 1997-98 El Niño on highland malaria in Tanzania. *Lancet*, **355**: 989-990.

- Litsios S. (2002). Malaria control and the future of international public health. In: Casman E.A., Dowlatabi H. eds. *The contextual determinants of malaria*. Washington DC: Resources for the future. Pp. 292-328.
- Loevinsohn M.E. (1994). Climatic warming and increased malaria incidence in Rwanda. *Lancet*, **343**: 714-718.
- Lysenko A.Y. & Beljaev A.E. (1969). An analysis of the distribution of *Plasmodium ovale*. *Bulletin of the World Health Organization*, **40**: 383-394.
- MacCormack C.P. (1984). Human ecology and behaviour in malaria control in tropical Africa. *Bulletin of the World Health Organization*, **62**: 81-87.
- MacDonald G. (1950a). The analysis of malaria parasite rates in infants. *Tropical Diseases Bulletin*, **47**: 915-935.
- MacDonald G. (1950b). The analysis of infection rates in diseases in which superinfection occurs. *Tropical Diseases Bulletin*, **47**: 907-914.
- Macdonald G (1953). The measurement of malaria transmission. *Proceedings of the Royal Society of Medicine* **48**: 295-301.
- MacDonald G. (1956). Epidemiological basis of malaria control. *Bulletin of the World Health Organization*, **15**: 613-626.
- MacDonald G. (1957). *The epidemiology and control of malaria*. London, Oxford University Press.
- Mahalanobis P.C. (1948). Historic note on the D-squared statistic. *Sankhya*, **9**: 237.
- Malakooti M.A., Biomondo K., Shanks G.D. (1998). Re-emergence of epidemic malaria in the highlands of western Kenya. *Emerging Infectious Diseases*, **4**: 671-676.
- Malaria Control Programme Ethiopia (1991). *The epidemiological basis for the control of malaria: an Ethiopian experience*. Malaria Control Programme, Ethiopia.
- Malone J.B., Bergquist N.R., Huh O.K., Bavia M., Bernardi M., El Bahy M., Fuentes M., Kristensen T., McCarroll J., Yilma J., Zhou X.N. (2001). A global network for the control of snail-borne disease using satellite surveillance and Geographic Information Systems. *Acta Tropica*, **79**: 7-12.
- MapInfo Corporation (1985-2000). *MapInfo Professional*. Troy New York, USA, MapInfo Corporation, One Global View, Troy NY 12180-8399, USA (software).

MARA/ARMA collaboration (1998). Towards an atlas of malaria transmission in Africa: First technical report of the MARA/ARMA. Durban, South Africa, South African Medical Research Council, IDRC, Canada and Wellcome Trust, UK.

Marsh K., Forster D., Waruiru C., Mwangi I., Winstanley M., Marsh V., Newton C., Winstanley P., Warn P., Peshu N., Pasvol G., Snow R.W. (1995). Indicators of life-threatening malaria in African children. *New England Journal of Medicine*, **332**: 1399-1404.

Marsh K., Snow R.W. (1999). Malaria transmission and morbidity. *Parassitologia*, **41**: 241-246.

Martens P., Hall L. (2000). Malaria on the move: Human population movement and malaria transmission. *Emerging Infectious Diseases*, **6**: 103-109.

Martens W.J.M., Niessen L.W., Rotmans J., Jetten T.H., McMichael A.J. (1995). Potential impact of global climate change on malaria risk. *Environmental Health Perspectives*, **103**: 458-464.

Martens W.J.M., Jetten T.H., Focks D.A. (1997). Sensitivity of malaria, schistosomiasis and dengue to global warming. *Climatic Change*, **35**: 145-156.

Martens P., Kovats R.S., Nijhof S., de Vries P., Livermore M.T.J., Bradley D.J., Cox J., McMichael A.J. (1999). Climate change and future populations at risk of malaria. *Global Environmental Change-Human and Policy Dimensions*, **9**: S89-S107.

Martin P.H., Lefebvre M.G. (1995). Malaria and climate: Sensitivity of malaria potential transmission to climate. *Ambio*, **24**: 200-207.

Matheron G. (1963). Principles of geostatistics. *Economic Geology*, **58**: 1246-1266.

Matola Y., White G., Magayuka S. (1987). The changed pattern of malaria endemicity and transmission at Amani in the eastern Usambara mountains, north-eastern Tanzania. *Journal of Tropical Medicine and Hygiene*, **90**: 127-134.

Matteelli A., Donato F., Shein A., Muchi J.A., Abass A.K., Mariani M., Leopardi O., Maxwell C. A., Carosi G. (1996). Malarial infection and birthweight in urban Zanzibar, Tanzania. *Annals of Tropical Medicine and Parasitology*, **90**: 125-134.

Mbogo C., Mwangangi J., Nzovu J., Gu W., Yan G., Gunter J., Swalm C., Keating J., Regens J., Shililu J., Githure J., Beier J. (2003). Spatial and temporal heterogeneity of *Anopheles* mosquitoes and *Plasmodium falciparum* transmission along the Kenyan coast. *American Journal of Tropical Medicine and Hygiene*, **68**: 734-742.

McGregor I.A. (1965). The passive transfer of human immunity. *American Journal of Tropical Medicine and Hygiene*, **13**: 237-239.

McKenzie F.E., Killeen G.F., Beier J.C., Bossert W.H. (2001). Seasonality, parasite diversity, and local extinctions in *Plasmodium falciparum* malaria. *Ecology*, **82**: 2673-2681.

McMichael A.J., Martens W.J.M. (1995). The health impacts of global climate change: grappling with scenarios, predictive models, and multiple uncertainties. *Ecosystem Health*, **1**: 23-33.

MEASURE Program. Demographic and Health Surveys for sub-Saharan Africa databases derived using STATCompiler, <http://www.measuredhs.org/> accessed 30.4.2004.

Menendez C (1995). Malaria during pregnancy: A priority area of malaria research and control. *Parasitology Today*, **2**: 178-183.

Metselaar D., Van Theil P. (1959). Classification of malaria. *Tropical and Geographical Medicine*, **11**: 157-161.

Mika A.M. (1997). Three decades of Landsat instruments. *Photogrammetric Engineering and Remote Sensing*, **63**: 839-852.

Ministry of Health, Uganda/World Health Organization (1959-1962). Annual reports of the Uganda Malaria Eradication Pilot Project.

Modiano D., Sirima B., Sawadogo A., Sanou I., Pare J., Konate A., Pagnoni F. (1999). Severe malaria in Burkina Faso: urban and rural environment. *Parassitologia*, **41**: 251-254.

Molineaux L., Cornille-Brogger R., Mathews H.M., Storey J. (1978). Longitudinal serological study of malaria in infants in the West African savanna. Comparisons in infants exposed to, or protected from, transmission from birth. *Bulletin of the World Health Organization*, **56**: 573-578.

Molineaux L. (1988). The epidemiology of human malaria as an explanation for its distribution, including some implications for its control. In: *Malaria: Principles and practice of malarology*. Eds W.H. Wernsdorfer & I.A. McGregor. Edinburgh, London, Melbourne, New York, Churchill Livingstone. 913-998.

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

- Monasch R, Reinisch A, Steketee RW, Korenromp EL, Alnwick D, Bergevin Y. (2004). Child coverage with mosquito nets and malaria treatment from population-based surveys in African countries: a baseline for monitoring progress in roll back malaria. *American Journal of Tropical Medicine and Hygiene*. **71**: 232-238.
- Morgan H.G. (1994). Placental malaria and low birthweight neonates in urban Sierra Leone. *Annals of Tropical Medicine and Parasitology*, **88**: 575-580.
- Mouchet J., Sircoulon J., Onapa A.W., Manguin S., Laventure S. (1997). Recrudescence du paludisme dans les hautes terres d'Afrique et de Madagascar. Paris, Institut Santé et Développement/ORSTOM.
- Mouchet J., Manguin S., Sircoulon J., Laventure S., Faye O., Onapa A.W., Carnevale P., Julvez P., Fontenille D. (1998). Evolution of malaria in Africa for the past 40 years: impact of climatic and human factors. *Journal of the American Mosquito Control Association*, **14**: 121-130.
- Mugisha, F. and Arinatwe, J. (2003). Sleeping arrangements and mosquito net use among under fives: results from the Uganda Demographic & Health survey. *Malaria Journal*, **2**: 40-50.
- Mung'ala-Odera V, Snow RW, Newton CRJC (2004). The burden of the neuro-cognitive impairment associated with *Falciparum* malaria in sub-Saharan Africa. *American Journal of Tropical Medicine & Hygiene*, **71**: 64-70.
- Murray, C.J.L. and Lopez, A.D. (1997). Mortality by cause for eight regions of the world: global burden of disease study. *Lancet*, **349**: 1269-1276.
- Myneni R.B., Maggion S., Iaquina J., Privette J.L., Gobron N., Pinty B., Kimes D.S., Verstraete M.M., Williams D.L. (1995). Optical remote sensing of vegetation: modeling, caveats and algorithms. *Remote Sensing of the Environment*, **51**: 169-188.
- Nabarro D., Tayler E. (1998). The Roll Back Malaria campaign. *Science*, **280**: 265-272.
- Nájera J.A., Kouznetzov R.L., Delacollette C. (1998). Malaria Epidemics: detection and control, forecasting and prevention. World Health Organization. Report No. WHO/MAL/98.1084.
- Nevill C.G., Some E.S., Mung'ala V.O., Mutemi W., New L., Marsh K., Lengeler C., Snow R.W. (1996). Insecticide-treated bednets reduce mortality and severe morbidity from malaria among children on the Kenyan coast. *Tropical Medicine and International Health*, **1**: 139-146.

Ngimbi N.P., Beckers A., Wery M. (1982). Aperçu de la situation épidémiologique du paludisme à Kinshasa (République du Zaïre) en 1980. [Survey of the epidemiological status of malaria in Kinshasa (Republic of Zaïre) in 1980]. *Annales de la Société Belge Médecin Tropicale*, **62**: 121-137.

Nicholson S.E., Davenport M.L., Malo A.R. (1990). A comparison of vegetation response to rainfall in the Sahel and East Africa using Normal Difference Vegetation Index from NOAA-AVHRR. *Climate Change*, **17**: 209-241.

Noor A.M., Gikandi P.W., Hay S.I., Muga R.O., Snow R.W. (2004). Creating spatially defined databases for equitable health service planning in low-income countries: the example of Kenya. *Acta Tropica*, **91**: 239-251.

Oliver M.A. (1996). Kriging: a method of estimation for environmental and rare disease data. In: Environmental Geochemistry and Health. D Appleton, R Fuge and GJH McCall, Geological Society Special Publication, **113**: 245-254.

Omumbo J.A., Ouma J., Rapuoda B., Craig M.H., le Sueur D., Snow R.W. (1998). Mapping malaria transmission intensity using geographical information systems (GIS): an example from Kenya. *Annals of Tropical Medicine and Parasitology*, **92**: 7-21.

Omumbo, J.A., Hay, S.I., Goetz, S.J., Snow, R.W., Rogers, D.J. (2002). Updating historical maps of malaria transmission intensity in East Africa using remote sensing. *Photogrammetric Engineering and Remote Sensing*, **68**: 161-166.

Omumbo J.A., Snow R.W. (2004). *Plasmodium falciparum* parasite prevalence in East Africa: A review of empirical data 1927-2003. *East African Medical Journal*, **81**: 649-656.

Omumbo J.A., Hay S.I., Guerra C.A., Snow R.W. (2004). The relationship between the *Plasmodium falciparum* parasite ratio in childhood and climate estimates of malaria transmission in Kenya. *Malaria Journal*, **3** Article number 17.

Omumbo J.A., Guerra C.A., Hay S.I., Snow R.W. (2005). The influence of urbanization on measures of *Plasmodium falciparum* infection prevalence in East Africa. *Acta Tropica*, **93**: 11-21.

Onori E. (1967). Distribution of *Plasmodium ovale* in the Eastern, Western and Northern regions of Uganda. *Bulletin of the World Health Organization*, **37**: 665-668.

Onori E., Grab B. (1980). Indicators for the forecasting of malaria epidemics. *Bulletin of the World Health Organization*, **58**: 91-98.

PAHO (2003). Status of malaria programmes in the Americas (based on 2002 data). Report of the 44th directing council, 55th session of the regional committee. Pan American Health Organization (PAHO) / World Health Organization (WHO). Washington D.C.

Pampana E.J., Russell P.F. (1955). Malaria: a world problem. Geneva, Switzerland, WHO.

Patz J.A., Strzepek K., Lele S., Hedden M., Greene S., Noden B., Hay S.I., Kalkstein L., Beier J.C. (1998). Predicting key malaria transmission factors, biting and entomological inoculation rates, using modelled soil moisture in Kenya. *Tropical Medicine and International Health*, **3**: 818-827.

Population Division of the Department of Economic and Social Affairs of the United Nations Secretariat (2002). World urbanization prospects: The 2001 revision data tables and highlights. New York, United Nations. 1-173.

Port G.R., Boreham P.F.L. (1980). The relationship of host size to feeding by mosquitoes of the *Anopheles gambiae* Giles complex (Diptera: Culicidae). *Bulletin of Entomological Research*, **70**: 133-144.

Price J.C. (1984). Land surface temperature measurements from the split window channels of the NOAA Advanced Very High Resolution Radiometer. *Journal of Geophysical Research*, **89**: 7231-7237.

Pringle G., Matola Y.G. (1967). Report on the Pare-Taveta vital statistics survey, 1962-1966. Nairobi, East African High Commission. East African Institute of Malaria and Vector-borne Diseases, East African Common Services Organisation. Printed by the East African Common Services Organisation Printer. 47.

Prothero R.M. (1961). Population movements and problems of malaria eradication in Africa. *Bulletin of the World Health Organization*, **24**: 24, 405.

Prothero R.M. (1977). Disease and mobility: A neglected factor in epidemiology. *International Journal of Epidemiology*, **6**: 259-267.

Rapuoda B.A., Achola P. (1984). Studies on malaria and its vectors in Nairobi: A review of the distribution of the vectors and the prevalence of the disease. In: 5th Annual Scientific Conference, 1984. 1984. Nairobi, Kenya: KEMRI/KETRI.

Rashed, S., Johnston, H., Dongier, P., Moreau, R., Lee, C., Lambert, J., Schaefer, C. (2000). Economic impact of febrile morbidity and use of permethrin-impregnated bed nets in a malarious area. I: Study of demographics, morbidity and household expenditures associated with febrile morbidity in the Republic of Benin. *American Journal of Tropical Medicine & Hygiene*, **62**: 173-180.

- Reiter P. (1998). Global warming and mosquito-borne disease. *Lancet*, **348**: 839-840.
- Reiter P. (2000a). Malaria and global warming in perspective? *Emerging Infectious Diseases*, **6**: 438-439.
- Reiter P. (2000b). From Shakespeare to Defoe: Malaria in England in the Little Ice Age. *Emerging Infectious Diseases*, **6**: 1-11.
- Reiter P., Thomas C.J., Atkinson P.M., Hay S.I., Randolph S.E., Rogers D.J., Shanks G.D., Snow R.W., Spielman A. (2004). Global warming and malaria: a call for accuracy. *Lancet Infectious Diseases*, **4**: 323-324.
- Remme J., Binka E., Nabarro D. (2001). Towards a framework and indicators for monitoring Roll Back Malaria. *American Journal of Tropical Medicine and Hygiene*, **64**: 76-84.
- Republic of Kenya Government Printer (1989). Kenya Population Census 1989. Volume I. Nairobi, Kenya, Office of the Vice President and Ministry of Planning and National Development / Central Bureau of Statistics.
- Robert V., Gazin P., Boudin C., Molez J., Ouedraogo V., Carnevale P. (1985). La transmission du paludisme en zone de savane arborée et en zone rizicole des environs de Bobo Dioulasso (Burkina Faso). *Annales de la Société Belge de Médecine Tropicale*, **65**(Suppl. 2): 201-214.
- Robert, V., Gazin, P., Ouédraogo, V., Carnevale, P. (1986). Le paludisme urbain à Bobo-Dioulasso (Burkina Faso). 1. Étude entomologique de la transmission. Cahiers O.R.S.T.O.M. *Série Entomologie Médicale et Parasitologie*, **24**: 121-128.
- Robert V., MacIntyre K., Keating J., Trape J-F., Duchemin J-B., Warren M., Beier J.C. (2003). Malaria transmission in urban sub-Saharan Africa. *American Journal of Tropical Medicine and Hygiene*, **68**: 169-176.
- Roberts D.R., Rodriguez M.H. (1994). The environment, remote sensing, and malaria control. *Annals of the New York Academy of Sciences*, **740**: 396-402.
- Rogers D.J. (1988). The dynamics of vector-transmitted diseases in human communities. *Philosophical Transactions of the Royal Society London B*, **321**: 513-539.
- Rogers D.J., Packer M.J. (1993). Vector-borne diseases, models and global change. *Lancet*, **342**: 1282-1284.

Rogers D.J., Williams B.G. (1994). Tsetse distribution in Africa: seeing the wood and the trees. In: Large scale ecology and conservation biology. PJ Edwards, RM May and NR Webb. Oxford, Blackwell Scientific Publications: 249-273.

Rogers D.J., Hay S.I., Packer M.J. (1996). Predicting the distribution of tsetse flies in West Africa using temporal Fourier processed meteorological satellite data. *Annals of Tropical Medicine and Parasitology*, **90**: 225-241.

Rogers, D.J. (1998). Satellite imagery and the prediction of tsetse distributions in East Africa. Proceedings of Symposium on *Diagnosis and Control of Livestock Diseases Using Nuclear and Related Techniques: Towards Disease Control in the 21st Century*. IAEA, Vienna 1997. IAEA pp 397-420.

Rogers D.J. (2000). Satellites, space, time and the African trypanosomiasis. In: *Advances in Parasitology: Remote Sensing and Geographical Information Systems in Epidemiology*. SI Hay, SE Randolph and DJ Rogers. London, UK, Academic Press. **47**: 129-171.

Rogers D.J., Randolph S.E. (2000). The global spread of malaria in a future, warmer world. *Science*, **289**: 1763-1766.

Rogers D.J., Randolph S.E., Snow R.W., Hay S.I. (2002). Satellite imagery in the study and forecast of malaria. *Nature*, **415**: 710-715.

Rogers DJ & Robinson TP (2004). Chapter 8: Tsetse Distribution. In: *The Trypanosomiasis*. I. Maudlin, P.H. Holmes & M.A. Miles (eds). CABI Publishing, CAB International, Wallingford, Oxfordshire UK pp. 139-179.

Ross R. (1911). *The prevention of malaria*, 2nd Edition. London, John Murray.

Ross R. (1928). *Studies on malaria*. London, Murray.

Rossi P., Belli A., Mancini L., Sabatinelli G. (1986). [A longitudinal entomologic survey on the transmission of malaria in Ouagadougou (Burkina Faso)]. *Parassitologia*, **28**: 1-15.

SALB, 2004. Second Administrative Level Boundaries (SALB) Project overviews: Concepts, progress and future. http://www3.who.int/whosis/gis/salb/salb_po.htm Accessed 29th April 2004.

Schneider, A., Friedl, M.A., McIver, D.K. and Woodcock, C.E. (2003). Mapping urban areas by fusing multiple sources of coarse resolution remotely sensed data. *Photogrammetric Engineering and Remote Sensing*, **69**: 1377-1386.

Schute G.T. (1988). The microscopic diagnosis of malaria. In: *Principles and Practice of Malariology*. Eds W.H. Wernsdorfer & I.A. McGregor. 781-814.

- Schwartzlander B. (1997). Global burden of disease. *Lancet*, **350**: 141-142.
- Schwetz J. (1942). Recherches sur la limite altimétrique du paludisme dans le Congo orientale et sur la cause de cette limite. *Annales de la Société Belge de Médecin Tropicale*, **22**: 183-208.
- Seboxa C. & Snow R.W. (1997). Epidemiological features of severe paediatric malaria in North Western Ethiopia. *East African Medical Journal*, **74**: 780-783.
- Service M.W. (1997). Mosquito (Diptera: Culicidae) dispersal - the long and short of it. *Journal of Medical Entomology*, **34**: 579-588.
- Service M.W., Townson H. (2002). The *Anopheles* vector. In: Essential Malariology. DA Warrell and HM Gilles Eds. New York, Arnold Publishers. 59-84.
- Shanks GD (1999). Unpublished data. *Personal communication*.
- Shanks G.D., Biomndo K., Hay S.I., Snow R.W. (2000). Changing patterns of clinical malaria since 1965 among a tea estate population located in the Kenyan highlands. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, **94**: 253-255.
- Shililu J., Ghebremeskel T., Seulu F., Mengistu S., Fekadu H., Zerom M., Asmelash G.E., Sintasath D., Mbogo C., Githure J., Brantly E., Beier J.C., Novak R.J. (2004). Seasonal abundance, vector behavior, and malaria parasite transmission in Eritrea. *Journal of the American Mosquito Control Association*, **20**: 155-164.
- Small J., Goetz S.J., Hay S.I. (2003). Climatic suitability for malaria transmission in Africa, 1911-1995. *Proceedings of the National Academy of Science. USA*, **100**: 15341-15345.
- Smith D.L., McKenzie F.E. (2004). Statics and dynamics of malaria infection in *Anopheles* mosquitoes. *Malaria Journal* **3**, Article 13.
- Smith T., Charlwood J.D., Kihonda J., Mwankusye S., Billingsley P.F., Meuwissen J., Lyimo E., Takken W., Teuscher T., Tanner M. (1993). Absence of seasonal variation in malaria parasitaemia in an area of intense seasonal transmission. *Acta Tropica*, **54**: 55-72.
- Smith T., Leuenberger R., Lengeler C. (2001a). Child mortality and malaria transmission intensity in Africa. *Trends in Parasitology*, **17**: 145-149.
- Smith T., Hii J.L.K., Genton B., Müller I., Booth M., Gibson N. Narara A., Alpers M.P. (2001b). Associations of peak shifts in age-prevalence for human

malarias with bednet coverage. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, **95**: 1-6.

Snijders FL (1991). Rainfall monitoring based on Meteosat data - a comparison of techniques applied to the western Sahel. *International Journal of Remote Sensing*, **12**: 1331-1347.

Snow R.W., Schellenberg J.R.M.A., Peshu N., Forster D., Newton C.R.J.C., Winstanley P.A., Mwangi I., Waruiru C., Warn P.A., Newbold C., Marsh K. (1993). Periodicity and space-time clustering of severe childhood malaria on the coast of Kenya. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, **87**: 386-390.

Snow R.W., de Azevedo I.B., Lowe B.S., Kabiru E.W., Nevill C.G., Mwankusye S., Kassiga G., Marsh K., Teuscher T. (1994). Severe childhood malaria in two areas of markedly different falciparum transmission in East Africa. *Acta Tropica*, **57**: 289-300.

Snow, R.W. and Marsh, K. (1995). Will reducing *Plasmodium falciparum* transmission alter mortality among African children? *Parasitology Today*, **11**: 188-190.

Snow R.W., Marsh K., Le Sueur D. (1996). The need for maps of transmission intensity to guide malaria control in Africa. *Parasitology Today*, **12**: 455-456.

Snow R.W., Omumbo J.A., Lowe B., Molyneux C.S., Obiero J.O., Palmer A., Weber M., Pinder M., Nahlen B., Obonyo C., Newbold C., Gupta S., Marsh K. (1997). Relation between severe malaria morbidity in children and level of *Plasmodium falciparum* transmission in Africa. *Lancet*, **349**: 1650-1654.

Snow R.W. & Marsh K. (1998). New insights into the epidemiology of malaria relevant for disease control. *British Medical Bulletin*, **54**: 293-309.

Snow R.W., Gouws E., Omumbo J.A., Rapuoda B.A., Craig M.H., Tanser F.C., le Sueur D., Ouma J. (1998). Models to predict the intensity of *Plasmodium falciparum* transmission: applications to the burden of disease in Kenya. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, **92**: 601-606.

Snow R.W., Craig M.H., Deichmann U., Marsh K. (1999a). Estimating mortality and disability due to malaria among Africa's non-pregnant population. *Bulletin of the World Health Organization*, **77**: 624-640.

Snow R.W., Craig M.H., Deichmann U., Le Sueur D. (1999b). A preliminary continental risk map for malaria mortality among African children. *Parasitology Today*, **15**: 99-104.

Snow R.W., Gilles H.M. (2002). The epidemiology of malaria. In: *Essential Malariology*. DA Warrell and HM Gilles Eds. New York, Arnold. 85-106.

Snow R.W., Marsh K. (2002). The consequences of reducing transmission of *Plasmodium falciparum* in Africa. *Advances in Parasitology*, **52**: 235-264.

Snow R.W., Craig M., Newton C., Steketee R. (2003). The public health burden of *Plasmodium falciparum* malaria in Africa: Deriving the numbers. Working paper 11, Disease Control Priorities Project (DCPP). Bethesda, Maryland, Fogarty International Centre, National Institutes of Health. August 2003. <http://www.fic.nih.gov/dcopp>. 1-75.

Snow RW, Korenromp EL, Gouws E (2004). Paediatric mortality in Africa: *Plasmodium falciparum* malaria as a cause or a risk? *American Journal of Tropical Medicine & Hygiene*, **71**: 16-24.

Snow, R.W., Omumbo, J.A. (2005). Malaria mortality in sub-Saharan Africa. In Eds D Jamison, R Feacham, W Makgoba, K Hofman, K Rogo. *Disease and mortality in sub-Saharan Africa*. Oxford University Press & World Bank, (in press).

Some E.S. (1994). Effects and control of highland malaria epidemic in Uasin Gishu District, Kenya. *East African Medical Journal*, **71**: 2-8.

Spielman A. (1993). Time limitation and the rôle of research in the worldwide attempt to eradicate malaria. *Journal of Medical Entomology*, **30**: 6-19.

Statistics Department (1995). The 1991 population and housing census: Final results (Main Release). Statistics Department, Ministry of Finance and Economic Planning, The Republic of Uganda, Entebbe, January 1995.

Steinwand D. (1994). Mapping raster imagery to the Interrupted Goode Homolosine projection. *International Journal of Remote Sensing*, **15**: 3463-3471.

Steketee R., Wirima J., Hightower A., Slutsker L., Heymann D., Breman J. (1996). The effect of malaria and malaria prevention in pregnancy on offspring birthweight, prematurity, and intrauterine growth retardation in rural Malawi. *American Journal of Tropical Medicine and Hygiene*, **55**: S33-41.

Steketee R., Nahlen B.L., Parise M.E., Menendez C. (2001) The burden of malaria in pregnancy in malaria endemic areas. *American Journal of Tropical Medicine and Hygiene*, **64**: 28-35.

Story M., Congalton R.G. (1986). Accuracy assessment: a user's perspective. *Photogrammetric Engineering and Remote Sensing*, **52**: 397-399.

- Struchler D. (1989). How much malaria is there world-wide? *Parasitology Today*, 5: 39-40.
- Sugita M., Brutasaert W. (1993). Comparison of land surface temperatures derived from satellite observations with ground truth during FIFE. *International Journal of Remote Sensing*, 14: 1659-1676.
- Sutton P., Roberts D., Elvidge C., Baugh K. (2001). Census from Heaven: an estimate of the global human population using nighttime satellite imagery. *International Journal of Remote Sensing*, 22: 3061-3076.
- Tanser F.C. (2000). The application of Geographical Information Systems to infectious diseases and health systems in Africa. Faculty of Medicine. Durban, South Africa, University of Natal: 168.
- Tanser F.C. (2002). The application of GIS technology to equitably distribute fieldworker workload in a large, rural South African health survey. *Tropical Medicine and International Health*, 7: 80.
- Tanser F.C., Sharp B.L., Le Sueur D (2003). Potential effect of climate change on malaria transmission in Africa. *Lancet* 362: 1792-1798.
- Tatem A., Hay S.I. (2004). Measuring urbanization pattern and extent for malaria research: A review of remote sensing approaches. *Journal of Urban Health: Bulletin of the New York Academy of Medicine*, 81: 363-375.
- Tatem, A.J., Noor, A.M., Hay, S.I. (2004). Defining approaches to settlement mapping for public health management in Kenya using medium spatial resolution satellite imagery. *Remote Sensing of Environment*, 93: 42-52.
- Thompson R., Begtrup K., Cuamba N., Dgedge M., Mendis C., Gamage-Mendis A., Enosse S., Barreto J., Sinden R.E., Hogg B. (1997). The Matola malaria project: a temporal and spatial study of malaria transmission and disease in a suburban area of Maputo, Mozambique. *American Journal of Tropical Medicine and Hygiene*, 57: 550-559.
- Thomson M.C., Connor S.J., Milligan P.J., Flasse S.P. (1996). The ecology of malaria - as seen from Earth-observation satellites. *Annals of Tropical Medicine and Parasitology*, 90: 243-264.
- Thomson M.C., Connor S.J., Milligan P., Flasse S.P. (1997). Mapping malaria risk in Africa: What can satellite data contribute? *Parasitology Today*, 13: 313-318.
- Thomson M.C., Connor S.J., D'Alessandro U., Rowlinson B., Diggle P., Cresswell M., Greenwood B.M. (1999). Predicting malaria infection in Gambian

children from satellite data and bed net use surveys: The importance of spatial correlation in the interpretation of results. *American Journal of Tropical Medicine and Hygiene*, **61**: 2-8.

Thomson M.C., Connor S.J. (2001). The development of Malaria Early Warning Systems for Africa. *Trends in Parasitology*, **17**: 438-445.

Trape J.F., Zoulani A. (1987a). Malaria and urbanization in Central Africa - The example of Brazzaville. 3. Relationships between urbanization and the intensity of malaria transmission. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, **81**: 19-25.

Trape, J.F. and Zoulani, A. (1987b). Malaria and urbanization in Central Africa: the example of Brazzaville. Part II. Results of entomological surveys and epidemiological analysis. *Transactions of the Royal Society of Tropical Medicine & Hygiene*, **81**(Supplement 2): 34-42.

Trape J.F., Zoulani A., Quinet M.C. (1987). Assessment of the incidence and prevalence of clinical malaria in semi-immune children exposed to intense and perennial transmission. *American Journal of Epidemiology*, **126**: 193-201.

Trape J-F., Lefebvre-Zante E., Legros F., Ndiaye G., Bouganali H., Druilhe P., Salem G. (1992). Vector density gradients and the epidemiology of urban malaria in Dakar, Sénégal. *American Journal of Tropical Medicine and Hygiene*, **47**: 181-189.

Trape J-F., Rogier C. (1996). Combating malaria morbidity and mortality by reducing transmission. *Parasitology Today*, **12**: 236-204.

Tucker C.J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of the Environment*, **8**: 127-150.

Tucker C.J., Sellers P. (1986). Satellite remote sensing of primary production. *International Journal of Remote Sensing*, **7**: 1395-1416.

Tulu A.N. (1993). Malaria. In: The ecology of health and disease in Ethiopia. H Kloos and ZA Zein. Colorado, Westview Press.

Uganda Bureau of Statistics (2004). <http://www.ubos.org/index.html>. Accessed in June 2004.

United Nations High Commissioner for Refugees (2001). 2000 global refugee trends. Analysis of the provisional UNHCR population statistics. Geneva, UNHCR population data unit. Population and geographic data section: 1-24.

Vlahov D., Galea S. (2002). Urbanization, urbanity and health. *Journal of Urban Health*, **79**: S1-12.

Voller A., O'Neill P., Humphrey D. (1972). Serological indices in Tanzania II. Antinuclear factor and malarial indices in populations living at different altitudes. *Journal of Tropical Medicine and Hygiene*, **75**: 136-139.

Walker, N., Schwartlander, B., Bryce, J. (2002). Meeting international goals in child survival and HIV/AIDS. *Lancet*, **360**: 284-289.

Warren M., Billig P., Bendahmane D., Wijeyaratne P. (1999). Malaria in urban and peri-urban areas in sub-Saharan Africa (Activity Report No. 71). Washington DC, Environmental Health Project (EHP). Office of Health and Nutrition, Bureau for Global Programmes, Field Support and Research, United, States Agency for International Development. 1-44.

Washino R.K., Wood B.L. (1994). Application of remote sensing to arthropod vectorsurveillance and control. *American Journal of Tropical Medicine and Hygiene*, **50**: 134-144.

Watts T.E., Wray J.R., Ngandu N.H., Draper C. (1990). Malaria in an urban and a rural area of Zambia. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, **84**: 196-200.

Wernsdorfer W.H., McGregor I.A. (1988). Malaria: Principles and practice of malariology. Edinburgh, London, Melbourne, New York, Churchill Livingstone.

White G.B. (1985). *Anopheles bwambae* sp. N., a malaria vector in the Semliki Valley, Uganda and its relationship with other species of the *An. gambiae* complex (Diptera: Culicidae). *Systematic Entomology*, **10**: 501-522.

Wilson D.B. & Wilson M.E. (1937). The manifestations and measurement of immunity to malaria in different races. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, **30**: 431-448.

World Bank (1996). Toward environmentally sustainable development in sub-Saharan Africa. Washington, The Bank.

World Health Organization (1951). World Health Organization, Technical Report Series No 38, 1951. Report on the conference on the eradication of malaria in Equatorial Africa, Kampala, Kampala, Uganda, World Health Organization (WHO).

World Health Organization (1966). Malaria eradication in 1965. *WHO Chronicles*, **20**: 286-300.

World Health Organization (1991). Basic malaria microscopy. Part 1. Learner's guide. Geneva, WHO.

World Health Organization (1994). World malaria situation in 1992. Part 1. *Weekly Epidemiological Record*, **69**: 309-314.

World Health Organization (1997). World malaria situation in 1994, Part 1. *Weekly Epidemiological Record*, **72**: 269-274.

World Health Organization (1999). Rolling back malaria. The world health report 1999: making a difference. Geneva: World Health Organization. 49-63.

World Health Organization (2002a). World Health Report 2002. Reducing risks, promoting healthy life. Geneva: World Health Organization.

World Health Organization (2002b). Global defence against the infectious disease threat: Roll Back Malaria. The stage is set for impact. Geneva, Roll Back Malaria Initiative, World Health Organization (WHO). 173-188.

World Health Organization/United Nations Children's Emergency Fund, (2003). *The Africa malaria report 2003*. Geneva/New York: World Health Organization / United Nations Children's Fund.

World Resources Institute (1995). Africa Data Sampler: A Geo-Referenced Database for all African Countries (CD-ROM and User's Guide), World Resources Institute, World Conservation Monitoring Center, and PADCO, Inc.

Yohannes M., Petros B. (1996). Urban malaria in Nazareth, Ethiopia: parasitological studies. *Ethiopian Medical Journal*, **34**: 83-90.

APPENDIX

The data collection proforma

Report number				-				
---------------	--	--	--	---	--	--	--	--

**MARA/ARMA collaboration (Mapping Malaria Risk in Africa)
Data collection proforma**

SECTION 1: DATA REFERENCE

ADMINISTRATIVE INFORMATION

Regional Centre Code	
----------------------	--

Date data transcribed (dd/mm/yyyy)			-			-		
------------------------------------	--	--	---	--	--	---	--	--

Data transcribed by (name)	
----------------------------	--

Data checked at co-ordinating centre	
--------------------------------------	--

DATA REFERENCE

1.1 Category of document

Journal		MOH report	
Book		Other report	
Thesis		Personal communication	

1.2 Reference source

Authors	

Year of publication	
---------------------	--

Title	

Origin / Journal							
Journal				Non-journal			
Volume		Issue no.		Total pages	from		to
Pages from		to		Specific pages	from		to

Does this report contain:

Parasite ratios (section 3)		Burden of disease (section 6)	
Entomology (section 4)		Incidence (section 7)	
How many data points are there in this report?			

Report number -

Number

SECTION 2: SURVEY LOCATION
2.1 Geographic location of survey

	Name	Code			
Country					
First level administrative unit (e.g. province)					
Second level administrative unit (e.g. district)					
Third level (e.g. area, division)					
Fourth level (e.g. location)					
Nearest large town					
Community level name (village, settlement)					

2.2 Survey location geo-reference

*N.B. Latitude refers to North (+) or South (-)
 Longitude refers to East (+) or West (-)*

Decimal degrees:

Point	Point name	Degrees Latitude	+/- N/S	Degrees Longitude	+/- E/W
1					
2					
3					
4					

or,

Degrees and decimal minutes:

Point	Point name	Latitude			Longitude		
		Deg	Minutes	+/- N/S	Deg	Minutes	+/- E/W
1							
2							
3							
4							

Degrees, minutes and seconds:

Point	Point name	Latitude				Longitude			
		Deg	Min	Sec	+/-	Deg	Min	Sec	+/-
1									
2									
3									
4									

Report number				-			
----------------------	--	--	--	---	--	--	--

2.3 Is the survey location a:

Point		Polygon		Multiple points (with pooled results)		Administrative unit	
-------	--	---------	--	---------------------------------------	--	---------------------	--

2.4 How did you geo-reference?

Africa Data Sampler		Scanned topographic map	
Africa Data Sampler estimate		Topographic map, manually	
Other digital map		Encarta	
HealthMap		GPS	
Geoname		Other (specify):	
Co-ordinates given in report			

If a topographic map was used, indicate its scale (e.g. 1: 5 000)	
---	--

2.5 Climatological and geographical data

<i>Is a graph, map or table available (of area, rainfall, temperature, disease profi seasonal abundance of infection or vectors, etc.)?</i>	<u>Yes</u>	<u>No</u>
---	------------	-----------

N.B. (Please attach photocopies of these to the proforma)

2.6 General information on the survey area:

<i>Is malaria transmission in the area:</i>	Yes	No	Not stated
Epidemic/unstable/highland			
Endemic/stable			
Seasonal			
Are the population studied refugees or displaced persons?			
Is rice grown in the area?			
Is there an irrigation scheme?			
Is there a swamp/wetlands/dam etc. close by?			

Is the area:	Urban		Peri-urban		Rural		Not stated	
--------------	-------	--	------------	--	-------	--	------------	--

Comments:

Data point number				-					-				
--------------------------	--	--	--	---	--	--	--	--	---	--	--	--	--

SECTION 3. THE PARASITE SURVEY

Location No.

3.1 Description of survey (Tick the relevant box/es)

<i>Survey of school children</i>	<input type="checkbox"/>
MCH/EPI survey	<input type="checkbox"/>
Drug resistance screening survey (include only initial, pre-intervention screening results)	<input type="checkbox"/>
Randomly-selected community-based survey (e.g. control populations in intervention studies)	<input type="checkbox"/>
Non-randomly-selected community-based survey (includes surveys which ask for volunteers)	<input type="checkbox"/>
Longitudinal survey of the same individuals	<input type="checkbox"/>
Longitudinal survey of randomly selected individuals	<input type="checkbox"/>
Total population survey	<input type="checkbox"/>
No details given	<input type="checkbox"/>

3.2 Survey methodology

Describe the methodology used to select subjects (*Indicate criteria for inclusion and exclusion*)

