Earth Observation, Geographic Information Systems and *Plasmodium falciparum* Malaria in Sub-Saharan Africa

S.I. Hay,¹ J.A. Omumbo,² M.H. Craig³ and R. W. Snow²

¹ Trypanosomiasis and Land Use in Africa (TALA) Research Group, Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK;
² KEMRI/Wellcome Trust Collaborative Programme, PO Box 43640, Nairobi, Kenya;
³ MARA/ARMA Investigating Centre, Medical Research Council, 771 Umbilo Road, PO Box 17120, Congella, 4013, South Africa

Abstract ................................................................. 174
1. Introduction ............................................................ 174
   1.1. The burden of *Plasmodium falciparum* malaria in sub-Saharan Africa ............................................ 174
   1.2. Defining epidemiological risk for malaria control ......................... 175
2. The Ecology of *P. falciparum* Malaria in Sub-Saharan Africa .................. 176
   2.1. Epidemiology and control .............................................. 176
   2.2. The life cycle ......................................................... 177
   2.3. Environmental determinants of disease transmission ....................... 179
3. Remote Sensing and Mosquito-borne Disease ...................................... 182
   3.1. The rationale for remote sensing ....................................... 182
   3.2. Low altitude aerial sensors ........................................... 183
   3.3. High altitude satellite sensors ...................................... 186
4. Geographic Information Systems and Mosquito-borne Disease .................. 191
   4.1. The history of malaria mapping ...................................... 191
   4.2. Continental scale GIS ................................................ 192
   4.3. Regional scale GIS .................................................. 197
5. Conclusions and the Future .................................................. 203
   Acknowledgements ...................................................... 206
   References .............................................................. 206
ABSTRACT

This review highlights the progress and current status of remote sensing (RS) and geographical information systems (GIS) as currently applied to the problem of Plasmodium falciparum malaria in sub-Saharan Africa (SSA). The burden of P. falciparum malaria in SSA is first summarized and then contrasted with the paucity of accurate and recent information on the nature and extent of the disease. This provides perspective on both the global importance of the pathogen and the potential for contribution of RS and GIS techniques. The ecology of P. falciparum malaria and its major anopheline vectors in SSA is then outlined, to provide the epidemiological background for considering disease transmission processes and their environmental correlates. Because RS and GIS are recent techniques in epidemiology, all mosquito-borne diseases are considered in this review in order to convey the range of ideas, insights and innovation provided. To conclude, the impact of these initial studies is assessed and suggestions provided on how these advances could be best used for malaria control in an appropriate and sustainable manner, with key areas for future research highlighted.

1. INTRODUCTION

1.1. The Burden of Plasmodium falciparum Malaria in Sub-Saharan Africa

It is recognized that 90% of the global malaria burden is concentrated in sub-Saharan Africa (SSA) and caused by P. falciparum. It has been estimated that approximately one million people died from the direct consequences of P. falciparum malaria infection in SSA during 1997 (Snow et al., 1999a) and that 75% of these deaths occurred among pre-school children. The malaria parasite is one of the most significant infectious agents African children encounter as they pass through childhood. Malaria not only poses a risk to survival but the repeated clinical consequences of infection during early life place a burden on households, health services and ultimately the economic development of nation states (Bloom and Sachs, 1998). Sachs and Warner (1997) have argued that the persistence of endemic malaria in the tropics, and particularly in Africa, is contributory to a perpetual state of depressed economic growth in these regions. These macro-estimates of burden and economic associations provide clear support for a renewed effort aimed at halving malaria mortality by the year 2010, referred to as the Roll Back Malaria (RBM) initiative (Nabarro and Tayler, 1998; WHO, 1998). This
optimistic goal has been conceived at a time when existing, affordable therapeutics are rapidly failing, health service provision is breaking down, vaccines seem a distant dream, and poverty, conflict and corruption continue to afflict many African states (Desowitz, 1999).

Nevertheless in the wake of this desperate position there is a renewed hope offered by new and old approaches aimed at disease management and the prevention of infection. These include (i) the delivery of services to manage clinical disease through the integrated management of childhood illnesses (IMCI) or engaging mothers and the informal drug sector for home-based management, (ii) new therapeutics and strategies involving drug combinations, (iii) approaches to preventing infection in pregnant women to reduce anaemia and the consequences of placental infection in newborns, (iv) insecticide treated bed nets provided by the informal and commercial sectors and (v) residual house spraying in areas where the interruption of transmission is a realistic objective.

Massive financial investment will be required even to begin to reduce the incidence of malaria mortality in SSA using these tools. The resources targeted at malaria control will always be limited when compared with other social sector investment. Furthermore, not all malaria interventions will be equally appropriate for every setting. The challenge for the public-health sector is to decide which interventions would be appropriate where, and how these may be tailored to the local epidemiology to achieve maximal health impact for minimal investment. Clearly not all resource allocation is evidence-based but there is an increasing recognition among the malaria research and control communities that mapping risk and the projected benefits of intervention is a fundamental monitoring and decision-informing tool. It is in this context the utility of remote sensing (RS) and geographical information systems (GIS) is evaluated.

1.2. Defining Epidemiological Risk for Malaria Control

It is only during the 1990s that the public-health risks of \( P. falciparum \) malaria in SSA have become better defined. Against a constant risk of infection throughout life, individuals develop functional immunity against the severe and fatal outcomes of infection in infancy. The risk of a self-limiting clinical attack continues through the early years of life before declining later in childhood. The speed of acquired functional immunity depends upon the frequency of parasite exposure from birth. Consequently, complicated clinical disease and mortality from malaria are concentrated amongst the very young children under conditions of intense, high transmission and those at risk include increasingly older children as the intensity of parasite transmission declines (Snow et al., 1997; Snow and Marsh, 1998). Ultimately
when parasite transmission is extremely low and subject to unstable epidemic conditions, disease and mortality risks are equally distributed across all age groups. In these conditions the lack of immunity is balanced by the low risk of infection, but any disturbance in the frequency of parasite transmission can lead to devastating effects among the population. The range of environments found in SSA permits the entire spectrum of transmission conditions. Urbanization is a rapidly proliferating demographic feature of the continent and traditionally urban settings represent areas of low parasite exposure. Unstable transmission conditions exist in the south of Africa, the highlands of East Africa and the Horn of Africa and traditionally arid areas of the Sahel and Namibia (see Plate 4). In the remainder of the stable endemic areas in SSA, conditions can vary from one infectious bite from local vectors every 3 years to one every night (Hay et al., 2000). Moreover, seasonal risks of infection vary enormously from a few weeks to constant risk throughout the year.

Facility-based health information systems or national civil registration systems are notoriously inadequate for national disease priority setting or programme planning. Consequently, little information exists on the epidemiological patterns and levels of disease risk in many areas of SSA. The location of populations in relation to health services is fundamental baseline information for health planning in the developed world. Such information, however, is often not available to those responsible for resource allocation in the developing world. Climate and its seasonality are coupled to the temporal variation in vector numbers and hence disease risk. These climate associations led to the development of crude, expert-opinion ‘seasonal risk maps’ for East Africa during the 1950s. No new map was developed until GIS rekindled interest in these types of information. Information on population distribution, health services, disease risk and seasonality should remove some of the barriers to providing a credible platform upon which to institute selected or targeted malaria control and prevention. GIS and RS provide a framework to develop high-resolution maps of risk, population and service delivery.

2. THE ECOLOGY OF P. FALCIPARUM MALARIA IN SUB-SAHARAN AFRICA

2.1. Epidemiology and Control

SSA supports a range of malaria endemicity whose geographical distribution is determined by a complex series of environmental, social, economic
and historical patterns and contingencies. A complete epidemiology of
P. falciparum malaria requires an analysis far beyond the scope of this review.
The most important epidemiological aspects are outlined however, so that
the rationale for using RS and GIS techniques can be appreciated.

The basic reproductive number, $R_0$, provides a useful structure for
thinking about environmental influences on disease. It describes the average
number of new cases of a disease that will arise from the introduction of an
infective host into a wholly susceptible population (Anderson and May,
1991a) and, for an insect vector-borne disease such as malaria, it is given by
the standard formula:

$$R_0 = \frac{a^2 m c b e^{-\mu T}}{\mu r}$$

where $a$ is the vector biting rate, $m$ is the ratio of vectors to hosts, $c$ is the
transmission coefficient from vertebrate to vector (i.e. the proportion of
bites by vectors on infected hosts that eventually give rise to mature
infections in the vectors), $b$ is the transmission coefficient from vector to
vertebrate, $\mu$ is the vector mortality rate, $T$ is the incubation period of the
parasite within the vector (sometimes referred to as the extrinsic incubation
period) and $r$ is the rate of recovery of the vertebrate from infection
(Anderson and May, 1991b). Importantly, all of the variables are associated
with the vector except $r$, which is a function of the host alone. Due to the
dominance of vector biology on the transmission dynamics of pathogens they
carry, it follows that the distribution and intensity of such diseases is
dependent upon the distribution and abundance of the vector. For the
purposes of this discussion we describe both the sensitivity of P. falciparum
and its vectors to environmental conditions.

2.2. The Life Cycle

2.2.1. Plasmodium falciparum

Mosquitoes of the genus Anopheles were first identified as the vectors of
human malaria in 1897 by Sir Ronald Ross (Wernsdorfer and McGregor,
1988). Today, four species of Plasmodium are known to infect humans
namely P. falciparum, P. malariae, P. vivax and P. ovale. P. falciparum is the
dominant malaria parasite found in the stable endemic areas of Africa
(Young, 1976; Gilles, 1993). The malaria parasite develops in two stages; a
sexual cycle that takes place within the mosquito vector and an asexual cycle
in the human host (Fujioka and Aikawa, 1999). The haematophagous adult
female Anopheles seek vertebrate hosts soon after emergence. The ingested
blood is used to support egg production and, following development and
subsequent oviposition, the female vector seeks further blood meals to nourish future broods. It is this repeated feeding that facilitates the transmission of parasites between hosts. Infection of the human host begins when sporozoites from an infected mosquito are injected into the blood of a susceptible human during a blood meal. It then takes 0.5–4 h for the sporozoites to invade host liver cells where they multiply and release as many as 30,000 merozoites, which, in turn, invade red blood cells. This asymptomatic period usually lasts about a week in tropical countries. The erythrocytic asexual development stage follows when the parasite develops from a ring form to a trophozoite that then becomes a schizont which multiplies to produce 4–32 merozoites. This intracellular multiplication causes red blood cells to rupture, with the resultant release of toxins into the blood, occurring in synchronized 48 h cycles for *P. falciparum*, *P. vivax* and *P. ovale* and in 72 h cycles for *P. malariae*. The bouts of fever associated with malaria correspond with these episodes of toxin release. Continued asexual multiplication with the invasion of further erythrocytes, as well as sexual differentiation, results in the production of macrogametocytes (female) and microgametocytes (male). These are the forms of the parasite infective to the mosquito.

The parasite’s sexual cycle begins when gametocytes are ingested by a mosquito vector feeding on an infected individual (Beier, 1998). Fertilization of the gametocytes to form ookinetes takes place in the midgut of the mosquito and these lodge in the midgut outer wall as oocysts. Numerous sporozoites develop within the oocysts and, as the oocysts rupture, migrate to the mosquito’s salivary gland from where they are injected into the human host during subsequent blood meals. Various aspects of this complex life cycle are affected by climate and are explored below.

2.2.2. *The Anopheles gambiae Complex*

Members of the *Anopheles* genus (subfamily Anophelinae, family Culicidae, order Diptera) are thought to be unique in being competent vectors for human malaria (Gillies and de Meillon, 1968; Service, 1985; Gillies, 1987). *Anopheles gambiae s.l.* is the dominant vector within the Afrotropical and Ethiopian faunal region (Boyd, 1930) and comprises six sibling species (see Plate 5). *An. gambiae s.s.*, *An. arabiensis* and *An. quadriannulatus* are fresh water breeders (White, 1974), *An. melas* and *An. merus* are salt water breeders (White, 1974), while *An. bwambae* breeds in mineral rich waters (Gillies and de Meillon, 1968; Service, 1985; Gillies, 1987).

*An. gambiae s.s* and *An. arabiensis* tend to occur sympatrically in approximately 70% of SSA (White et al., 1972; Lindsay et al., 1998) with their relative abundance dependent on local ecological conditions. *An. gambiae*
s.s. (formerly species A) prefers more humid and vegetated areas. It is predominantly anthropophagic and endophilic and tends to have higher vectorial capacity than other species of the gambiae complex due to its greater longevity.

*An. arabiensis* (formerly species B) predominates in the Horn of Africa and southern Arabia and is the sole vector of malaria in northern Sudan (Wernsdorfer and McGregor, 1988) (see Plate 5). This species is noted for its ecophenotypic plasticity, being predominantly exophagic and exophilic but facultatively endophagic and endophilic in adverse environmental conditions. Its occasional abundance in arid areas is due to its ability to breed in residual pools of water in dry riverbeds and rapidly increase at the onset of rains.

*Anopheles melas* and *An. merus* are adapted for salt water breeding and are hence found only along the coastlines of West and East Africa respectively (White, 1974). *An. bwambae* (formerly species D) and *An. quadriannulatus* (formerly species C) are not vectors of human malaria and therefore are not considered further.

2.2.3. Anopheles funestus

*An. funestus* is the most important vector of malaria in SSA after *An. gambiae s.s.* and *An. arabiensis* (see White *et al.*, 1972; Service, 1993). It prefers shaded, long-lived bodies of water with emergent vegetation for oviposition (Christie, 1959). This species is predominantly anthropophagic and feeds both indoors and outdoors. In areas of rice cultivation it emerges when the rice plants have grown and provide shade and therefore appears later in transmission season than *An. gambiae* (see Bradley, 1991).

### 2.3. Environmental Determinants of Disease Transmission

#### 2.3.1. Temperature

Malaria is a disease of tropical and temperate countries between the latitudinal limits of 64° North and 57° South (Gill, 1921) with prevalence increasing towards the equator. As the parasites require time to develop into infective stages, female anopheles are not immediately infective after feeding. This extrinsic incubation period (*T*) is temperature dependent and optimum conditions have been defined between 25 and 30°C, with development ceasing below 16°C and above 40°C (Russell *et al.*, 1946; Gilles, 1993). Intermittent low temperatures have also been found to delay sporogony, with the period immediately after the infective bite being most
sensitive to temperature decreases. The effects of temperature on various aspects of anopheline biology have been investigated under laboratory conditions (Jepson et al., 1947; Macdonald, 1957; Detinova, 1962; Martens, 1997) and are summarized in Table 1. High temperatures are also associated with increased frequency of feeding, since blood meals are more rapidly digested and the whole gonotrophic cycle accelerated (Boyd, 1930).

2.3.2. Rainfall

Rainfall provides surface water in which female anopheles can lay eggs. In arid areas where temperatures are usually suitable, malaria transmission occurs only when rainfall provides temporary breeding habitat for vectors. These areas are often classified as ‘malarious near water’ since transmission outside the rainy seasons typically occurs only along riverbeds, oases and other man-made surface water sites. Studies have demonstrated an association between abundance of An. gambiae s.l. and rainfall (Christie, 1959; White et al., 1972; Molineaux and Gramiccia, 1980; Charlwood et al., 1995). Rainfall effects are often most apparent during epidemics when the rise in malaria cases is often proportional to the amount of precipitation, among other factors (Christophers, 1911; Covell, 1957; Wernsdorfer and McGregor, 1988; Malakooti et al., 1998; Kilian et al., 1999). Classic epidemiological studies in Punjab compared the hospital-based monthly prevalence of fever deaths with contemporaneous total monthly rainfall data (Gill, 1920, 1921; Boyd, 1930) and found periods of increased mortality.
correlated maximally with rainfall lagged by one month. The relationship is obviously best observed where temperature is not confounding. Excessive rainfall can reduce transmission by flushing larvae out of small pools (Covell, 1957; Molineaux, 1988; Gilles, 1993). The relationship between precipitation and the development of breeding sites also depends on slope of the land, run-off and soil type leading some authors to look at the feasibility of hydrological modelling to improve rainfall–malaria associations (Patz et al., 1998).

2.3.3. Climate Seasonality

The interaction between temperature and rainfall is largely responsible for the seasonal characteristic of malaria transmission. Seasonal variation of infection risk is a common feature of malaria in SSA and is reflected in intra-annual changes in vector densities, entomological inoculation rates and malaria admissions (Christie, 1959; Julvez et al., 1992; Aniedu, 1997; Hay et al., 1998b, 2000). The map showing the heterogeneity in malaria seasonality across Uganda and Kenya (see Plate 6), based on an RS proxy of seasonality (Hay et al., 1998a) is discussed in more detail in Section 3.2.

2.3.4. Atmospheric Moisture

Early documented reports of human malaria describe its association with humid swamps and marshes (Gill, 1921) and several authors have attempted to define optimum conditions of relative humidity (RH) based on such observations (Wernsdorfer and McGregor, 1988; Gilles, 1993). Gill (1921) was the first formally to investigate the effect of changes in RH on Culex fatigans and the transmission of avian malaria. Higher relative humidities were associated with increased vector longevity and greater frequency of feeding. RH also determined the timing and duration of daytime resting behaviour (Boyd, 1930; Russell et al., 1946; Molineaux, 1988).

2.3.5. Altitude

Altitude has long been a subject of interest among malariologists (Schwetz, 1942; Garnham, 1948; Heisch and Harper, 1949; Covell, 1957; Roberts, 1964; Malakooti et al., 1998). Altitude and temperature are explicitly linked, with every 100 m increase in height corresponding to an approximately 0.5°C decline in temperature. The use of altitude can be confusing, however, with the limit for malaria transmission variously reported above 2000 m in Ethiopia (Covell, 1957), 1800 m in the Congo (Schwetz, 1942) and 1950 m in
Kenya (Garnham, 1948). Use of the phrase ‘highland malaria’ continues, but it is more clearly thought of as temperature-limited unstable transmission.

2.3.6. Other Environmental Factors

Many reviews have described the effects of anthropogenic environmental change on malaria transmission (Bradley, 1992; Mouchet et al., 1998) including specifically the effects of agricultural practices (Coosemans, 1985; Robert et al., 1985; Githeko et al., 1993; Dossou-Yoyo et al., 1994; Hay et al., 2000), water resource development (El Gaddal, 1991; Hunter et al., 1993), urbanization (Trape and Zoulani, 1987; Coene, 1993), population movements (Nega, 1991), malaria control (Draper and Smith, 1957, 1960; Draper et al., 1972) and climate change (Lindsay and Birley, 1996; Lindsay and Martens, 1998). There is no scope to elaborate on these here.

3. REMOTE SENSING AND MOSQUITO-BORNE DISEASE

3.1. The Rationale for Remote Sensing

The principal goal of RS in malaria epidemiology is to help map the disease distribution so that control efforts in endemic situations and intervention strategies in epidemic situations may be most directed efficiently. Section 2 highlighted the long understood relationships between mosquito numbers and climate, so that shorter term meteorological variation could be used in predicting the onset and severity of malaria epidemics (Swaroop, 1949). Decades later, our understanding of how factors such as temperature, humidity and rainfall influence mosquito population dynamics has improved, along with the sophistication with which weather variables can be used to predict malaria transmission rates (Onori and Grab, 1980). Furthermore, many mosquito species oviposit in specific aquatic habitats which support characteristic plant communities and which are hence easily identified (Rioux et al., 1968). The use of RS techniques to investigate mosquito and malaria ecology stems from the understanding that aerial and space-borne sensors could provide relevant surrogate information relating to the spatial variation in these meteorological and vegetation variables (Cline, 1970).

Previous reviews of RS applied to the problems of mosquito and malaria control (Roberts and Rodriguez, 1994; Washino and Wood, 1994; Thomson et al., 1997; Dale et al., 1998; Hay et al., 1998a; Connor, 1999) 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. Related themes, such as the advantages of RS in terms of the rate, geographical extent and spatial and spectral resolution of data collection have also been well documented (Hugh-Jones, 1989; Hay, 1997), as have some of the disadvantages regarding the interpretation and quality of such RS data (Hay et al., 1996; Hay and Lennon, 1999). The aim of this section is not to reiterate these works, but to outline the increased epidemiological understanding and control operation efficiency that can be gained by adopting RS techniques. The extent of current research does not allow a specific focus on P. falciparum malaria in the continent of Africa. Instead we look at RS applications to all mosquito-borne diseases, from around the world, with an emphasis on the techniques applied. This background will enable us to anticipate, identify and plan for the opportunities made available by forthcoming advances in satellite-sensor technology.

Details of RS terminology, acronyms and satellite-sensor system specifications are outlined by Hay (this volume). In brief however, low altitude, high spatial resolution sensors are those that can resolve objects of less than 1\texttimes1 km at nadir and high altitude, low spatial resolution sensors those that cannot. Passive sensors derive their energy from the reflected and emitted energy of the Sun, and actively sensors generate energy themselves, usually at radar frequencies.

3.2. Low Altitude Aerial Sensors

It is the dependence of mosquitoes on fresh and brackish water habitats in the early stages of their life-cycle that first allowed RS techniques to be exploited. The following sections describes how early studies used RS techniques for the purpose of high spatial resolution habitat mapping.

3.2.1. Passive Photography

In 1971, National Aeronautics and Space Administration (NASA) scientists, in combination with personnel from the New Orleans Mosquito Control District (NOMCD), were the first to investigate the use of colour and colour–infrared (CIR) aerial photography in mapping vegetation assemblages associated with the larval habitat of Aedes sollicitans, a nuisance saltmarsh mosquito, that in addition is a competent vector for the equine encephalitis flavivirus (Anonymous, 1973). Previous work by NOMCD had shown that female mosquitoes always oviposited in areas of the saltmarsh intermittently flooded by freshwater. The floral assemblage dominated by spikerush
(Spartina patens) and wiregrass (Juncus roemerianus) was adapted to this same hydrological regime and hence was a reliable indicator of Ae. sollicitans larval habitat (Bidlingmayer and Klock, 1955). The report documents that such vegetation assemblages were extremely accurately identified at an 80 ha test site near New Orleans, although accuracy statistics were not provided.

The first operational use of CIR aerial photography was for delineating forested and open wetlands, marshes and residential areas in a new mosquito control area in the Saginaw and Bay counties of Michigan (Wagner et al., 1979). As well as nuisance floodwater Aedes mosquitoes, Culex spp. that oviposited in tree-hole pools posed a significant public-health problem causing a local epidemic of St Louis encephalitis (SLE) in 1975. The known flight range of each mosquito species was used in combination with information on the distance between residential areas and mosquito habitat to identify control priorities for the two counties. The authors stressed the short time in which an environmental inventory was gathered and management priorities identified. Furthermore, the streamlining of subsequent control efforts led to a campaign of relatively low economic and environmental cost, since the area designated for insecticide treatment was considerably reduced in comparison with the alternative approach of broadcast aerial spraying. In contrast, Hopkins et al. (1975) reported that after an outbreak of SLE in Dallas in 1966 the entire county of Texas was sprayed aerially with an organophosphate insecticide (malathion) mist. Moreover, in response to the highly publicized outbreak of a West Nile-like flavivirus in New York (Lanciotti et al., 1999), large areas of the city were sprayed using aerial insecticide. Should these flaviviruses become newly established in urban centres of North America (Anderson et al., 1999), aerial spraying guided by RS would save time, money and unnecessary environmental impact.

The purpose of these studies was to identify larval habitats of mosquito species that constituted a significant nuisance or public health risk, so as to direct larval control more efficiently. Although simple and empirical, species–habitat correlations from CIR aerial photography have been, and continue to be, cited as more cost-effective than conventional ground survey techniques for obtaining information on the distribution of potential larval habitats; for example, for Psorophora columbicae in both Louisiana (Fleetwood et al., 1981) and Texas rice fields (Welch et al., 1989a, b) and for Culex annulirostris in urban areas of Queensland (Dale and Morris, 1996).

Perhaps unsurprisingly, it has also been observed from a study in Queensland, Australia that Ae. aegypti breeding sites (usually water accumulated in discarded containers) were not easily resolved using CIR photography (Moloney et al., 1998), which was hence not thought to be of sufficient resolution to assist in mapping dengue vector habitat to assist control.
3.2.2. Passive Optical Radiometers

Investigators were quick to exploit these early successes and move from aerial photography to aerial observations with Multispectral Scanners (MSS). Improvements upon CIR aerial photography were demonstrated by Cibula (1976), and Barnes and Cibula (1979) using an airborne MSS that had a spatial resolution of 2.5 × 2.5 m when flown at an altitude of 1200 m. The greater number of wavelengths over which data were recorded (22 channels from 0.3 to 13 μm) enabled more accurate spectral identification of the Spartina–Juncus associations characteristic of favourable egg-laying habitats for *Ae. sollicitans* than by aerial infrared photographs.

Under the aegis of the NASA Global Monitoring and Disease Prediction Program (GMDPP), various authors investigated the relationship between mosquito population dynamics and environmental factors that could be observed remotely (Wood *et al.*, 1994). In phase I of this project, populations of *An. freeborni* in the rice fields of northern and central California were studied (Pitcairn *et al.*, 1988). This mosquito species represented no health risk to the local population, but provided an accessible model for the study of anopheline vectors in irrigated rice habitat (Wood *et al.*, 1991a,b, 1992).

Larval mosquito populations were sampled fortnightly throughout the period of rice crop development in 1985 (Wood *et al.*, 1991a). An airborne MSS collected data simultaneously with a spatial and spectral resolution designed to simulate the Landsat Thematic Mapper (TM). A Normalized Difference Vegetation Index (NDVI) was then calculated for individual fields on each of the sampling dates and the differences in spectral signal between rice fields producing high and low numbers of anopheline larvae were followed throughout the growing season. The NDVI was considered an important variable since Rejmankova *et al.* (1988) had previously demonstrated percentage rice crop cover to be positively correlated with mosquito larval production. The results showed that higher values of NDVI in the early growing season (June) were associated with high mosquito-producing rice fields, but that the spectral separation (between high and low mosquito-producing fields) diminished to a minimum in mid-July, when percentage rice crop cover was at the maximum. A discriminant analysis which incorporated information from Landsat-TM equivalent channels 1, 2, 3, 4, and 7 was able to distinguish between high and low mosquito-producing fields with an overall accuracy of 75%. The uneven distribution of high mosquito-producing fields was also shown to be related to patterns of surrounding land use, with 70% of the high mosquito-producing fields being within 1.5 km of livestock pasture (i.e. potential hosts). A more detailed survey in 1987 repeated the above work and included data on the distance to livestock in a GIS (Wood *et al.*, 1991b). An identical discriminant analysis combined with cattle distance data resolved high and low producing fields.
with an accuracy of 90%, with errors of omission (not identifying high producing fields) and commission (identifying as a high producing field one which is not) of 10% and 40% respectively (Wood et al., 1992).

3.2.3. **Active Radar Radiometers**

A recent approach has demonstrated that land-based radar may be capable of predicting habitat flooding status (Ritchie, 1993). Radar was used in this study to estimate rainfall in Collier county, Florida. This information, in combination with data from tide tables, was used to guide surveys for recently inundated saltmarsh habitat suitable for oviposition by *Ae. taeniorhynchus*. Rain, tides and rain plus tide events triggered 48%, 26% and 26%, respectively, of proposed inspections and this information system located seven of the eight habitats of mosquitoes found during the study and control period.

3.3. **High Altitude Satellite Sensors**

The increasing number of successful applications of RS techniques to mosquito habitat mapping in North America provided sufficient impetus for researchers to transfer these approaches to satellite-borne sensor mapping of tropical areas, predominantly in meso-America, where both the logistical and public health problems were significantly greater. This move from analogue photographic to digital MSS techniques was also important, as it facilitated the move to quantitative analyses which, in turn, enabled the investigation of more subtle variation in environmental variables and thus mosquito habitat and ultimately disease suitability.

3.3.1. **Passive Optical Radiometers**

(a) **High Spatial, Low Temporal Resolution Sensors.** Landsat-1 and -2 MSS data were then used for the first time to identify freshwater plant communities associated with the larval habitats of *Ae. vexans* and *Culex tarsalis* in riparian habitat bordering the Niobara river between South Dakota and Nebraska (Hayes et al., 1985). A supervised classification using multiple-date Landsat-MSS scenes identified these aquatic habitats with 95% accuracy. This study demonstrated that such species–habitat correlations could be successfully ‘scaled-up’ to regional control programmes using relatively high spatial resolution satellite sensor data.

The second phase of the previously cited NASA-GMDPP project investigated the population dynamics of *An. albimanus* and *An. pseudopunctipennis*, malaria vectors in the tropical wetlands of Chiapas,
Mexico (Roberts et al., 1991). Pope et al. (1994a) used two Landsat-TM scenes of the area, one from the dry and one from the wet season, to provide an unsupervised classification (Curran et al., this volume) of the region. The resulting clusters were assigned to land-cover types on the basis of CIR aerial photographs and field inspection of 30 test sites. These sites were independently sampled for mosquito density and information was collected on environmental variables affecting water and vegetation characteristics (Rejmankova et al., 1991). The sites were then grouped into 16 habitat types using cluster analysis and correlations were performed between the habitat types and land-cover units (Rejmankova et al., 1992). The land-cover units were subsequently ranked as having high, medium or low mosquito production potential on the basis of these correlations. Incorporating this information into a GIS, sites of high mosquito production around the towns of La Victoria and Efrain Gutierrez were found to occupy only 9% of the designated control area, allowing the potential for substantial streamlining of control campaign effort and resources.

An independent study also demonstrated that Landsat-TM data could be used to identify villages at high risk of malaria transmission within the Tapachula region of Chiapas, Mexico (Beck et al., 1994, 1995). Dry and wet season Landsat-TM scenes were again subject to an unsupervised classification. A stepwise discriminant analysis and linear regression were then used to establish the relationship between vector abundance and land-cover. Land-cover in a 1 km buffer area surrounding the perimeter of the village was analysed since this corresponded to the known flight range of An. albimanus. Transitional swamp and unmanaged pasture were found to be the most important land-covers for An. albimanus, and their combined area in the buffer zone was sufficient to predict high and low vector abundance in 40 villages to an overall accuracy of 90%. Furthermore, these relationships were sufficiently robust to predict the abundance of adult An. albimanus in a further 40 randomly selected villages from the neighbouring Huixtla region to an overall accuracy of 70% using more multi-date Landsat-TM scenes (Beck et al., 1997).

Rejmankova et al. (1995) have also shown that the density of adult An. albimanus mosquitoes around villages in Belize could be reliably predicted using multispectral Satellite pour l’Observation de la Terre (SPOT) High Resolution Visible (HRV) data. Productive larval habitats were first identified as marshes containing relatively few emergent aquatic plants and a high coverage of cyanobacterial mats. An unsupervised Bayesian maximum likelihood classification was then applied to a single SPOT scene covering a test site occupying northern Belize. The classes generated were subsequently assigned to individual ‘landscape elements’ based on field observations. Human settlements were identified with ancillary map data and located more precisely on subsequent field visits with a global positioning system.
(GPS). These settlements were divided into two groups according to their distance to the larval habitat class. Group 1 was composed of settlements closer than 500 m and group 2 of settlements further than 1 500 m. Based on independent measurements, a landing rate greater than 0.5 mosquitoes per human per minute (during the hours of maximum mosquito activity from 18:30 to 20:00) was used as a threshold for high adult mosquito density (Rodriguez et al., 1996). Group 1 was predicted as having adult mosquito densities higher than this threshold and group 2 lower. These predictions were tested by collecting mosquitoes landing on humans during the hours of peak activity within each of the settlements. The resulting predictions were 100% accurate for group 2 and 89% accurate for group 1.

Roberts et al. (1996) investigated the utility of multispectral SPOT-HRV data in predicting the distribution of the malaria vector An. pseudopunctipennis in central Belize. Previous investigations had shown that altitude and the presence of filamentous algae in sun-exposed pools were critical determinants for the presence of An. pseudopunctipennis larvae (Rejmankova et al., 1993). Using the SPOT and cartographic data, 49 sites were chosen and predicted to have a high or low probability for An. pseudopunctipennis presence. The criteria used in site selection were the distance of the houses from waterways and their altitude above such waterways, as well as the amount of forest between the houses and the waterway. The SPOT data provided more contemporary information than the cartographic data and hence allowed map errors to be corrected. They also gave information on important aspects of environmental suitability which could not be attained from the maps, such as the size of a waterway or the degree to which it was open to sunlight. Collections of mosquitoes were then made in a sample of these sites to test the predictions. Four of the eight sites that were predicted as high probability locations for presence of An. pseudopunctipennis were positive and all 12 low probability sites were negative. The absence of An. pseudopunctipennis at four high probability locations was thought to be due to field population densities of the species being below the threshold of the sampling effort.

Multi-date Indian RS (IRS) satellite Linear Imaging and Self Scanning (LISS) II sensor data have been used to identify ‘mosquitogenic conditions’ for maps to guide local mosquito control activities in Delhi, northern India (Sharma et al., 1996). Water bodies and marsh areas associated with housing were identified in a supervised classification and then changes in the area of these habitats between images correlated with adult and larval mosquito numbers. Statistics were not provided, making more detailed comment difficult.

Thomas and Lindsay (1999) conducted a retrospective spatial analysis of malaria cases in children from a series of villages in the centre of The Gambia. SPOT HR imagery and a GPS-based supervised classification were used to map exposed alluvial beds saturated with fresh water, that larval-dip
surveys had shown to be the dominant mosquito larval producing habitat. A significant, spatially corrected correlation between the area of RS derived breeding habitat within a 2 km radius of a village and entomological inoculation rate (EIR) ($r = 0.85$, $n = 10$, $P = 0.002$), allowed EIR to be extrapolated over a wider area, including hundreds of villages where clinical surveillance was being conducted. From this larger sample it was seen that both asymptomatic and clinically manifested malaria cases were negatively correlated with risk of exposure to malaria. This had been found before in The Gambia and attributed to bed net use (Thomson et al., 1996a), but this could not explain the relationship in this study. The authors concluded that higher malaria incidence in areas of lower exposure was due to a lack of challenge while young, resulting in low population immunity (Snow and Marsh, 1995; Snow et al., 1997).

(b) Low Spatial, High Temporal Resolution Sensors. Multi-temporal NDVI data at $8 \times 8$ km spatial resolution from the National Oceanic and Atmospheric Administration’s (NOAA) Advanced Very High Resolution Radiometer (AVHRR) have been applied to the problem of Rift Valley fever (RVF) epidemics (Linthicum et al., 1987, 1990). The work showed that high NDVI values in Kenya were good indicators of seasonally flooded linear depressions, known as dambos. These habitats were highly suitable for *Aedes* spp. mosquito breeding and were hence closely associated with RVF epidemics. The work progressed to incorporate higher spatial resolution Landsat-TM and multispectral SPOT-HRV imagery to locate individual areas of high RVF risk determined by the NDVI (Linthicum et al., 1991). Operational application of the technique, however, was hindered because investigators were not able to discriminate flooded from dry dambos in such images (see Section 3.3.2). Further refinement of this work has shown that a combination of NOAA-AVHRR NDVI anomaly data and Indian Ocean sea surface temperature data could predict RVF epidemics in Kenya 2–5 months before outbreaks, allowing sufficient time for public health service intervention (Linthicum et al., 1999).

Linthicum et al. (1994) also investigated an unusually severe and extensive outbreak of RVF in the West African Senegal river basin in 1987. The RVF outbreak was particularly unusual as it occurred during a period of only moderate rainfall. Analysis of multispectral SPOT-HRV scenes for the period of the epidemic revealed extensive flooding in Mauritania to have peaked in October 1987, as a result of the construction of the Diama and Manatelli dams on the river Senegal. This coincided exactly with the period of maximum RVF disease activity in the area. Furthermore, maximum values of local area coverage (LAC) NDVIs were associated with increased rice production (and hence productive mosquito larval habitats) around the newly flooded regions in Daro and Rosso, the foci of RVF outbreaks. Unfortunately, no statistics were presented in the study.
NDVI data derived from the NOAA-AVHRR have also been used in the Mar Chiquita lake region of central Argentina to predict the abundance of *Ae. albifasciatus* an important nuisance to livestock and vector of western equine encephalitis in the region (Gleiser *et al*., 1997). The authors argued that such relationships could be used to monitor and predict future surges in the population of this floodwater mosquito.

Thomson *et al.* (1996b, 1997) investigated the potential of coarse spatial resolution GAC and LAC NOAA-AVHRR data, as well as cold cloud duration (CCD) data from the Food and Agriculture Organization’s (FAO) African Real Time Environmental Monitoring Using Imaging Satellites (ARTEMIS) programme, to predict malaria risk in The Gambia. They concluded, however, that although there were clear relationships between satellite sensor data and environmental variables associated with malaria transmission, it was difficult to predict how these would affect adult mosquito abundance and behaviour. For instance, along the river Gambia a decrease in rainfall may at times increase available anopheline breeding sites by increasing the number of suitable pools in the alluvial soils at the river margin. They also noted that relationships between malaria incidence and environmental variables were complicated by sociological factors, because in areas where anopheline abundance was greatest, and hence biting most frequent, people were more likely to protect themselves with insecticide-impregnated bed nets (Thomson *et al*., 1994). Further work, however, has demonstrated more clearly that malaria infection rates in Gambian children could be related to a parameter representing NDVI seasonality after behavioural considerations had been factored out (Thomson *et al*., 1999). The authors went on to show how these relationships could be used to predict the changes in malaria prevalence resulting from treated and untreated bed net use, as well as giving some important warnings regarding the confidence intervals surrounding relationships between spatially autocorrelated data (see also Curran *et al*., this volume; Robinson, this volume).

Predictions of malaria seasonality (the combination of disease risk in space and time) in Kenya have been achieved by establishing relationships between childhood malaria cases, with disease data collected during on-going surveillance of severe malaria morbidity in five communities in Africa (Snow *et al*., 1997), and contemporary imagery from the NOAA-AVHRR and Meteosat-HRR (Hay *et al*., 1998b). The remotely sensed data were first processed to provide surrogate information on land surface temperature (LST), rainfall (expressed as CCD or the number of hours for which a given pixel was covered by cloud below a threshold determined as rain-bearing), reflectance in the middle infrared (MIR) wavelengths and the NDVI. These variables were then subject to temporal Fourier processing (Rogers *et al*., 1996) and compared with the mean percentage of total annual malaria admissions recorded each month at three sites in Kenya. The NDVI lagged by
1 month was found to be the most significant variable consistently correlated with malaria admissions across the sites (mean adjusted $r^2 = 0.71$, range 0.61–0.79). Interestingly, Patz et al. (1998) showed that NOAA-AVHRR NDVI also explained a large amount of temporal variation in human biting rate in a 2 years study in Kisian, western Kenya. Subsequent regression analyses showed that an NDVI threshold of 0.35–0.40 is required for more than 5% of the annual malaria cases to be presented in a given month (Hay et al., 1988b). Spatial extrapolation of these thresholds allowed the number of months for which malaria admissions could be expected to be defined across Kenya. The resulting ‘malaria season’ predictions were compared with a map of malaria transmission periods compiled from expert opinion (Butler, 1959). The correspondence is remarkable when the deficiencies inherent in compiling the original map are acknowledged. The authors stressed that the maps produced did not constitute a definitive picture of malaria transmission in Kenya but were a demonstration of methodology. It was considered that much more validation of the relationship between malaria admissions and NDVI was required to check its robustness in space and time. This map for Uganda and Kenya is reproduced at $1 \times 1$ km spatial resolution (see Plate 6).

3.3.2. Active Radar Radiometers

Imagery derived from C and L bands of an airborne Synthetic Aperture Radar (SAR) (Hay, this volume) has been used to detect dambo flooding status in the search for predictive mechanisms for RVF epidemics in Kenya (Pope et al., 1992). A significant advantage of using SAR was that data collection was independent of cloud coverage, especially important during the East African rains. Pope et al. (1992) concluded that the spatial resolution of current satellite-borne SARs was not sufficient to reveal many of the smaller dambos in the region. In a similar study, Pope et al. (1994b) demonstrated that airborne SAR data collected over the wetlands of Belize were capable of mapping marshes dominated by *Eleocharis*, an important breeding habitat of *An. albimanus* (Rejmankova et al., 1993).

4. GEOGRAPHIC INFORMATION SYSTEMS AND MOSQUITO-BORNE DISEASE

4.1. The History of Malaria Mapping

Before microcomputers and the advent of GIS, few examples existed of maps defining the global distribution of malaria (Lysenko and Semashko, 1968).
More maps are found at a regional and country scale, reflecting the more local concerns and experience of research and control personnel. Common to all historical maps is their representation of expert opinion, often aided by simple climatic and/or geographical iso-lines and, more rarely, limited empirical evidence. This static information failed to reflect the great spatial and temporal heterogeneity in transmission. GIS-based mapping allows methodologies to be standardized and repeated, and data to be updated and, since it is digital, easily interrogated. Thus GIS moves the mapping of malaria from a largely subjective science to one with quantitative foundations.

4.2. Continental Scale GIS

4.2.1. Modelling Malaria Burden

A major advocate of GIS in the study of malaria at a continental level has been the Mapping Malaria Risk in Africa/Atlas du Risque de la Malaria en Afrique (MARA/ARMA) project (le Sueur et al., 1997). The project was conceived to provide comprehensive, empirical, standardized maps of malaria distribution and endemicity to help guide malaria control operations in Africa (Snow et al., 1996). The aim was to collect and geo-reference all available malaria data in the formal and ‘grey’ literature. Data collection focused initially on the parasite prevalence survey, a proxy of malaria endemicity (Metselaar and Van Theil, 1959) since it is the most commonly measured parameter. This continental database of parasite prevalence was then used to produce evidence-based maps of malaria transmission. Recent additional attempts have been initiated to collect, map and interpolate the more direct but less standardized proxy of endemicity, the EIR (Hay et al., 2000).

The second focus of the MARA/ARMA project involved defining the distribution of endemic malaria and the duration and timing of the transmission seasons. The first completed model (Craig et al., 1999) was based on the assumption that malaria transmission at the continental scale was limited primarily by climate, and could hence be inferred from the distribution of temperature and rainfall at this scale. Continuous climate surfaces of mean monthly temperature and rainfall (1920–1980) were developed from interpolated weather station data at a spatial resolution of about \(5 \times 5\) km \((0.03 \times 0.03^\circ)\) for the analysis (Hutchinson et al., 1995). The model (Craig et al., 1999) used a fuzzy logic concept to express ‘degrees of climate suitability’ in contrast to rigid Boolean limits that ignored natural gradients and uncertainty. Mean monthly temperatures of 18°C were considered too cold for endemic \(P. falciparum\) transmission, but they became
increasingly suitable to a maximum at 22°C and above (Detinova, 1962; Macdonald, 1957). Zero rainfall in any month was considered unsuitable, and again it increased to maximum suitability for stable transmission at 80 mm and above per month. Monthly temperature and rainfall surfaces were converted to fuzzy suitability maps. Suitable conditions were required to persist for five consecutive months (three in North Africa). Low temperatures affect vector distribution: *An. gambiae* was found to be limited to frost-free regions (de Mellion, 1934) or to areas where absolute minimum temperatures in winter remained above 5°C (Leeson, 1931). Thus areas with minimum winter temperatures <3°C were classified as unsuitable. These empirically derived climate limits were tested by extracting rainfall and temperature profiles in known perennial, seasonal, epidemic and non-malarial areas. The resulting model describes climate suitability for endemic malaria, in which a fuzzy value of one predicts endemic malaria and a fuzzy value of zero predicts highly unstable or no transmission (see Plate 4). The model compared well with available country-level malaria risk maps of Kenya (Nelson, 1959), Tanzania (Wilson, 1956), South Africa before control (Anonymous, 1938), Botswana (Chayabjara et al., 1975; Diseko et al., 1997) and Namibia (de Meillon, 1951).

The malaria distribution model was subsequently used to estimate the number of people at risk of malaria in SSA (Snow et al., 1999b). The fuzzy distribution model was reclassified so that values of 0.5 and above were regarded as malarious. By overlaying this Boolean malaria distribution map with a 1990 population distribution model (Deichmann, 1996), the total population and the number of children aged below 5 years living in predicted malaria-endemic areas were estimated to be 360 million and 66 million, respectively. A more detailed study classified malaria risk differently for southern Africa (fuzzy values <0.5 = no malaria, 0.5 = stable malaria) than for the rest of Africa (where 0 = no malaria, >0 and <0.2 = epidemic risk, 0.2 = stable malaria) (Snow et al., 1999a) (Figure 1). In southern Africa, mortality (0.104 per 1000) and morbidity (11 per 1000) were estimated from surveillance data. For other endemic countries annual mortality estimates were based on 76 independent childhood mortality studies. Since very little is known about the mortality pattern in older children and adults, severe malaria admission rates in children under 10 years old, by year of age, were used to derive a model describing the decline in mortality risk with age, and from this to predict mortality risk for ages above 10 years by extrapolation. Thus the estimated mortality figures by age category were 9.4 per 1000 for 0–4 years, 2.17 for 5–9 years, 0.8 for 10–14 years and 0.13 per 1000 for 15 years and older. Morbidity estimates were based on 51 studies of malarial fever events in children: 976 cases per 1000 for all age groups in epidemic areas, and 999, 239 and 400 per 1000 for 0–4, 5–14 and 15-year-olds, respectively, in stable areas. Finally, the rate of severe malaria (cerebral,
severe anaemia), the proportion of people surviving severe malaria, and the risk of longer-term effects (neurological sequela and human immunodeficiency virus (HIV) infection following blood transfusion) were estimated from further studies. The overall estimates indicate that there may have been over 200 million clinical cases of malaria in Africa in 1995 and close to 1 million deaths. Furthermore, each year 3000 children may be left with lifelong disability (e.g. spasticity or epilepsy) and 19,000 may have contracted HIV after receiving contaminated blood for severe anaemia. These studies illustrate the rigorous way in which GIS techniques can integrate data from various origins. Whilst the final estimates do not differ significantly from previous expert guesses (Greenwood, 1990; Murray and Lopez, 1997), such

Figure 1 Zones defining different levels of risk of malaria according to the malaria distribution model (Craig et al., 1999): endemic areas in southern Africa (> 0.5), stable transmission areas in the rest of sub-Saharan Africa (>0.2), epidemic-prone areas (>0, <0.2), and negligible malaria risk (<0.5 in southern Africa, 0.0 elsewhere).
estimates are achieved using a rational and transparent methodology and are easily updated as new population and disease risk data become available.

4.2.2. Mapping Vector Distribution

Following the mapping of African *Anopheles* species collections by Gillies and de Meillon (1968), the records for the *An. gambiae* complex were updated and re-mapped by Davidson and Lane (1981). Coetzee et al. (2000) transferred these data to GIS and further updated the collection to contain 2538 records of six of the seven presently recognized species of the *An. gambiae* complex, including surveys from 1944 to the present (see Plate 5).

Lindsay et al. (1998) showed that the ranges of *An. gambiae s.s.* and *An. arabiensis* published by Davidson and Lane (1981) were described well by an array of climatic variables (Corbett and Kruska, 1994). Temperature ranges favoured by the two species were similar, but *An. gambiae* occurred in areas with higher annual precipitation (330–3224 mm) than *An. arabiensis* (237–415 mm), which corroborates previous observations; see Section 2.3.2. Lindsay et al. (1998) further analysed the relative abundance and found the distribution to be best described by the ratio of precipitation to potential evapotranspiration (ppn/pet) with *An. gambiae* dominating more humid, and *An. arabiensis* dominating more arid, areas respectively. The regression equations were fed back into the climatic data to produce a continent-wide model of proportional abundance of the two species. The expected values correlated well with the available data ($r^2 = 0.745$, $n = 14$, $P = 0.002$).

4.2.3. Climate Change

There have been several attempts to model changing malaria risk in relation to various scenarios of climate change. Inherent in these studies is the determination of baseline risk against which to measure change and some degree of spatial sophistication in data analysis.

Martin and Lefebvre (1995) used simple Boolean thresholds for temperature and moisture conditions suitable for transmission, with the requirement that these should be temporally coincident for *P. malariae*, *P. ovale*, *P. vivax* and *P. falciparum*, and iterated these monthly using 0.5 × 0.5° climate data to map areas of seasonal and perennial transmission globally. Temperatures from 19 to 32°C were considered suitable for *P. falciparum* and between 16 and 32°C for the other three *Plasmodium* species. A moisture index (ppn/pet; see above) of 0.7 defined the lower limit of suitable conditions for all four species. Malaria was considered seasonal if the suitable transmission period lasted 1–7 months and perennial if greater. They then investigated how the outputs from five global circulation models affected their
index of annual potential malaria transmission. With some notable exceptions, present day prediction agreed with the maps published previously (Lysenko and Semashko, 1968), although no accuracy statistics were provided. The future scenarios all predicted change in malaria distribution and seasonality, but did not concur on the nature and extent of that change. The five future models predicted 7–28% expansion of total malarious areas, and an increase in seasonal transmission areas of 12–55%, while all except one predicted a reduction in perennial zones by infringement of seasonal zones.

A more biological approach was adopted by Martens et al. (1995a,b), Martens (1997, 1999), and Lindsay and Martens (1998). The critical threshold density for malaria transmission (the number of infective mosquito bites per human host) were derived by re-arranging the equation for the basic reproductive number ($R_0$) for malaria (Anderson and May, 1991a). The reciprocal of this figure was then defined as the ‘epidemic potential’. Importantly, due to the complexity of estimating the change in vector abundance resulting from changing meteorological conditions, this variable was held constant (i.e. = 1). Risk for the various Plasmodium species were then derived using a $5 \times 7.5^\circ$ (latitude × longitude) global climate surface. The changes in epidemic potential in combination with a world population map were then used to make educated guesses about the changes in the total population exposed under various climate change scenarios. There was again no attempt at statistical validation of predictions of the present against current conditions. These studies contributed to the climate change debate by illustrating the sensitivity of malaria transmission to climate perturbations but were inadequate to define the current risk accurately.

The potential effect of predicted climate change on the distribution of An. gambiae s.l. has also been investigated for areas south of the equator (Rogers, 1996). The range published by Gillies and de Meillon (1968) was analysed against annual summary climate variables for 1961–1990. An. gambiae s.l. was specifically considered present in grid cells where it had been recorded, and absent in grid cells where other Anopheles species, but not An. gambiae, had been collected. Present distribution was predicted and compared with future global climate change scenarios. Interestingly, habitat suitability for the An. gambiae complex showed a net increase under all change scenarios.

A more extensive and global analysis of current and potential future P. falciparum malaria distribution by Rogers and Randolph (2000) resulted in similar conclusions. The recorded present day distribution of P. falciparum malaria was used to establish the current multivariate climatic constraints, using maximum likelihood methods based on the mean, max and minimum values of interpolated long-term climate records of temperature, rainfall and saturation vapour pressure. The prediction of the present day distribution was significantly better than previous attempts (Martin and Lefebvre, 1995;
Martens et al., 1995a,b; Martens, 1997, 1999; Lindsay and Martens, 1998) and this was attributed to the incomplete biological modelling approaches adopted in these studies. These results were applied to future climate scenarios (HadCM2 ‘medium-high’ and HadCM2 ‘high’) to predict future distributions. The potential future distributions showed remarkably few changes, even under the most extreme scenario, due to co-varying rainfall and moisture variables limiting any temperature driven expansion.

4.3. Regional Scale GIS

4.3.1. Surveillance

In South Africa malaria has been notifiable at the district level since 1956. Simple mapping of these malaria case data to administrative boundaries (see Plate 7) highlighted the lowlands adjacent to Mozambique as consistently high-risk districts (Sharp and le Sueur, 1996) and indicated expansion of malaria in the northern districts. le Sueur et al. (1995) and Sharp and le Sueur (1996) divided the northern high-risk districts of Kwazulu–Natal province into 20 control areas, and each area was subdivided into 10 sectors of approximately 30 × 30 km to investigate these trends. Annual malaria incidence varied widely from more than 300 per 1000 on the Mozambique border to less than 2 per 1000 in coastal regions from which An. funestus had been eradicated. An irrigation system around the inland town of Jozini was also implicated in the local rise of cases allowing control operations to be focused accordingly. The surveillance was further expanded to include the high-risk districts of Mpumalanga province (Booman et al., 2000). Malaria case data were entered within one week of collection, and the malaria control programme managers were able to generate near real-time reports and malaria distribution maps. Guided by the GIS surveillance system, indoor insecticide spraying is now limited to villages with incidence exceeding 8 cases per 1000. This is especially important as major redistribution in the health budget towards the management of the HIV/acquired immune deficiency syndrome (AIDS) epidemic in South Africa has increased pressure for efficient use of limited resources for malaria control.

A community health care system has also been established in Tigray, Ethiopia (Ghebreyesus et al., 1996) with the purpose of reducing malaria morbidity and mortality through early diagnosis and treatment, vector control and preventing infection during pregnancy with prophylaxis. The programme has combined data from the Tigray malaria control programme with spatial base maps to display and survey present health service infrastructure, clinic-based reporting of malaria, health service accessibility, community-based treatment of malaria cases, and coverage of early
diagnosis and treatment at the periphery. This supports planning of activities and monitoring of success of the programme in different parts of the region (Ghebreyesus et al., 1999).

In Israel, where malaria was eradicated in 1950, a national GIS surveillance system has been used to keep track of imported malaria cases and mosquito breeding sites (Kitron et al., 1994). Depending on the proximity of the water bodies to towns where malaria cases have been reported, and depending on the known flight ranges of the vectors, breeding sites are targeted for intensive vector control, through larviciding and/or drainage and channelling of run-off water, and increased surveillance.

A similar system in Trinidad resulted in all local and imported malaria cases referred to the Insect Vector Control Division being carefully recorded since 1965. Follow-up blood surveys and notification of surrounding health facilities ensured that practically no case was missed. All cases were geo-referenced and analysed retrospectively for space–time clustering using a \( k \)-nearest neighbour analysis (Chadee and Kitron, 1999). Of a total of 213 cases over a 30 year period, 164 were imported and were concentrated around the two large port cities on the west coast. All cases of \( P. falciparum \) were imported, and no clustering was found. Some \( P. malariae \) cases were concentrated in the island interior, however, with significant clustering in time of the second (\( J = 30, n = 57, P < 0.01 \)) and seventh (\( J = 41, n = 57, P < 0.02 \)) nearest neighbours, indicating local transmission. The vectors were bromeliad breeding \( An. bellator \) and \( An. homunculus \). Most \( P. vivax \) cases were imported, but nine cases were clustered closely in time in a small port town; these were associated with local transmission, probably by \( An. aquasalis \) and \( An. albitarsis \) which breed in surrounding marsh and rice fields, following the importation of a single case. This study thus provided the confidence for the government to restrict detailed and costly surveys to rural cases of \( P. malariae \) where bromeliad malaria can occur, and in coastal areas where people live in proximity to marshes.

Su (1994) established a GIS to monitor dengue and its vectors \( Aedes \) spp. in Kaohsiung city, Taiwan. The system allowed monitoring of vector breakouts, dengue cases and distribution of breeding sites, which gave a better overview of the situation. This supported the local government in their investigations and decision making and thus contributed towards more effective control of the disease.

4.3.2. Patterns of Disease Transmission

In Kenya, malaria survey data have been aggregated and displayed at the district administrative unit level to illustrate parasite distribution and prevalence (Omumbo et al., 1998). The continental malaria distribution
model described in Section 4.2.1 was used to mask areas where malaria was predicted as unstable or absent, to avoid the misuse of surveys possibly carried out during epidemics. The survey results were then analysed against underlying environmental factors (Snow et al., 1998a). A total of 124 community-based, cross-sectional malaria surveys, representing at least 100 children 0–10 years old and conducted since 1960, was included, and prevalence ratios were classified into high (70%), intermediate (20–69%) and low (<20%) transmission categories. The best explanatory climate variables were identified in a step-wise discriminant analysis, and these were then used to allocate each sample location to one of the categories. Agreement between allocated and observed values was high: of 124 measurements, 96 were allocated to their correct category. Interestingly, environmental conditions in April (also May) contributed most strongly towards malaria prevalence. The total number of suitable months with 60 mm of rainfall and temperatures above 19°C were also correlated with prevalence (univariate $r = 0.62$, $n = 124$, $P < 0.001$). The entire country was then categorized into high, intermediate and low risk areas, based on the regression equation. Populations exposed to different levels of risk were extracted, and mortality rates estimated: 26 000 children under 5 years old were estimated to die each year in endemic areas, and a further 3300–14 500 in epidemic areas. Additionally, 145 000 children were estimated to require intensive clinical management in an average year.

A clear example of application is provided by a situation analysis of community-based health care programmes in Kenya (Snow et al., 1998b) which indicated that, apart from the Ministry of Health, 23 different non-governmental organizations were active in health care. The analysis showed that not only were several populations covered by more than one programme, which illustrates the lack of inter-sectoral collaboration due to lack of appropriate information, but malaria intervention programmes were also not ideally positioned in relation to the distribution of malaria. By comparing the map of malaria endemicity with the distribution of target populations for insecticide-treated bed net implementation, a discrepancy emerged between the people exposed to malaria and those with access to treated nets (Shretta et al., 1998). More targeted and informed intervention was made possible simply by knowing how many people are exposed to malaria and where.

The statistical method used to analyse the Kenyan data treated each survey as an independent measurement and disregarded possible spatial autocorrelation. A subsequent analysis carried out on malaria survey data collected in Mali went a step further by investigating spatial trends (Kleinschmidt et al., 1999). Using a logistic regression method, the authors modelled larger-scale risk of malaria in Mali on relationships between prevalence and underlying environmental factors, namely various measurements of climate including the number of months with more than
60 mm of rain, NDVI, and distance to water. Spatial dependence in the model was investigated through a variogram, which suggested spatial dependence over distances of 15–20 km analysis (Robinson, this volume). Local variation in malaria risk, not accounted for by the original regression model, was then modelled by kriging the residuals at every survey point (Robinson, this volume). The accuracy of the model was then improved by adjusting the predicted values to resemble the observed values more closely, based on the kriged residuals. Distance to water (categorical), NDVI during the wet season, number of months with >60 mm rain, and average maximum temperature at the end of the dry season significantly explained malaria prevalence in Mali. One initially counter-intuitive outcome was that prevalence was higher at intermediate distances from water than at very short and very long distances. The likely explanation for this was that people living close to water make frequent use of bed nets to protect themselves from the great nuisance of mosquitoes, and are thus inadvertently protected from malaria (see Section 3.3.1).

Hightower et al. (1998) gave a detailed account of their experience in mapping 1169 houses using GPS in 15 villages along the shore of Lake Victoria, as well as health facilities, vector breeding sites, roads, rivers and the lake shoreline. Weekly mosquito trap collections and malaria prevalence in children less than 5 years old, were analysed against the distance of houses from the nearest major larval breeding site. Mosquito numbers were significantly correlated ($r^2$ not given, $n = 362, P = 0.0039$) with distance to breeding sites during the dry, but not during the wet, months ($r^2$ not given, $n = 395, P = 0.1530$). Parasite prevalence was weakly correlated with distance to breeding site in the wet months ($r^2$ not given, $n = 445, P = 0.3437$), but not during the dry months. The authors reported large variation in vector abundance, between and within villages, for both An. gambiae s.s. and An. funestus and even between the species.

Smith et al. (1995) analysed this local scale variation in abundance of An. funestus and An. gambiae s.l. in Namawala village in Tanzania. The houses were geo-referenced using a GPS and mosquitoes sampled by light trap every 2 weeks in 43 houses, and every night in one sentinel house. Vector abundance was highly variable, with between-house variation as pronounced as total seasonal variation in the village as a whole. Four different models (fixed effect, random effect, mixed, and conditional auto-regressive) were used to describe catches and extrapolate temporal and spatial variation. Many factors, including effect of windows and adjoining animal sheds, position of the house (central or peripheral, proximity to rice fields), seasonality, sampling frequency and total numbers of mosquitoes all affected the outcome of the models. This study illustrated the importance of site selection when extrapolations are made from highly focal measurements, such as a few sentinel houses. Mosquito density was highest in the valleys
where breeding sites were more permanent. Strong association of mosquito density with distance from breeding sites has been demonstrated again and again (Trape et al., 1992; Ribeiro et al., 1996); see also below.

A densely-populated suburb of Maputo, Mozambique was mapped by digitizing 1:10 000 maps and geo-referencing individual homes via GPS (Thompson et al., 1997). A detailed longitudinal study of malaria prevalence, clinical incidence and vector abundance was then carried out over 2.5 years from December 1992 to June 1995. Initial results showed clearly that prevalence of *P. falciparum* decreased sharply from between 40–60% adjacent to swampy vector breeding sites to 5–11% only 500 m away. Incidence of clinical malaria also decreased markedly. In fact, the relative risk of clinical malaria, assessed by multivariate analysis using a Cox model, was 6.2 times higher for individuals living less than 200 m from the breeding sites than for individuals living 500 m or more away (confidence interval 3.5–11.1; \( n = 857; P = 0.0001 \)). EIR was estimated to be 20 infectious bites per person per year in the high-risk areas, but was too low to measure in areas where malaria prevalence was below 11%. The high variability in prevalence, EIR and incidence over very short distances was attributed partly to very high population density: mosquitoes need not fly far to find their first host, and vector dispersion, hence transmission, is thus concentrated within the immediate vicinity of the breeding sites.

Some examples outside Africa further illustrate the use of GIS in malaria and mosquito research and control. In Kataragama, a highly endemic region in Sri Lanka, aerial photographs were used to set up a GIS including the position 423 houses representing 1875 inhabitants (Gunawardena et al., 1995). People in houses made of mud, or with incomplete walls and thatch or palm leaf roofs, were at 2.5 times greater risk of being infected with malaria than people in houses with complete and plastered brick walls and tiled or corrugated iron roofs. Distance to water increased risk of malaria only in the case of poorly constructed houses (\( r^2 = -0.31, n = 182, P = 0.0001 \)), but not in modern style houses (\( r^2 = 0.14, n = 161, P = 0.0676 \)). The authors recommended that the government should replace all mud-and-thatch houses within 200 m of a water body (which included 76% of the people in low-quality housing) with high-quality houses. This would reduce malaria incidence by 36% and would save the government money spent on malaria control. The money savings would off-set the initial expenditure after only 7.2 years, thereafter resulting in a net return of investments (Gunawardena et al., 1998).

In a study of 100 villages in Nadiad *taluka* (i.e. district), Gujarat, India (Sharma and Srivastava, 1997) layers of soil type, water table, water quality, relief and hydro-geomorphology were reclassified into zones of high, medium and low risk of waterlogging (which provide breeding sites). Risk layers were overlaid, yielding a mosaic of zones each with a certain combination of
conditions. Each zone was then given the risk level of the dominant factor in that zone. The resulting risk map was correlated with annual malaria incidence per zone in non-irrigated, but not in irrigated, areas. Each factor on its own could not adequately explain incidence. Malaria was also more prevalent where the water table was high, and where surface water and irrigation canals were present, than in villages along natural rivers, indicating the preferred breeding habitats of the local vector (Malhotra and Srivastava, 1995).

4.3.3. Disease Clustering

Patterns of malaria morbidity have been investigated by Snow et al. (1993) in a study involving 50,000 people living in 5000 scattered homesteads throughout Kilifi district of Kenya. Hand-drawn maps were used to identify and geo-reference with GPS the homes of all children presenting with severe P. falciparum malaria at Kilifi district hospital from May 1989 to April 1992. The number of pairs of cases found to live within 1, 1.5, 2, 2.5 and 3 km from one another, and presenting within 3, 5, 7, 10 and 14 days from one another, were compared with the number of cases that could be expected to occur close in time and space by chance. A significant number of cases was found to occur within 5 days and 2 km of one another. The clustering was unexplained in this study, but microclimatic differences causing local changes in the vector populations, or local importation of different strains of the parasite, were both postulated as possible causes.

In a further study of severe childhood malaria in the same population, from May 1991 to April 1993 all homesteads in the Kilifi district were geo-referenced with, and captured in, a GIS (Armstrong Schellenberg et al., 1998). Using a simple ‘equal population area’ algorithm, households were grouped into four different sizes of artificial areas, each containing about 1000, 500, 200 and 100 children respectively. In the second study year, cases of severe malaria were distributed more or less homogeneously across these areas, but in the first year there was significant heterogeneity at every level, with more admissions from people with easy access to the hospital. Using the Monte Carlo simulation method to check for overall clustering of cases, the childhood population surrounding a circular area from which 2, 3 or 4 cases of severe malaria had been admitted within 3, 7, 14 and 31 days of one another was compared against similar figures taken from simulations that assumed a random distribution. There was evidence of clustering in both time and space, particularly in the first study year. Lack of similar clustering in all-cause mortality data suggested that the clusters were real and not due to limitations in the numerator or denominator data.

Spatial cluster analysis of cases of La Crosse encephalitis (transmitted by Aedes mosquitoes) in Illinois, USA was used to identify ‘hot spots’ of the
disease in the state and the extent of clustering. Analysis revealed that most clustering occurred at a distance 10 km at state-level, and at a distance of 3 km at city-level (Kitron et al., 1997). Local clusters of disease were concentrated in houses near breeding sites such as undeveloped wood lots, large piles of tyres, or other discarded containers, which could be targeted for clean-up and vector control.

5. CONCLUSIONS AND THE FUTURE

Mosquito and malaria control are problems that operate at a range of spatial scale, from the catchment population of rural health clinics concerned with individual patients, to the regional and continental interests of international donor agencies concerned with the continental distribution of the disease and the quantification of total malaria incidence. The variety of RS platforms and sensors provides data relevant to each of these spatial scales and GIS provides an appropriate platform for integrating and interrogating such information. We have seen a wide range of RS imagery utilized with respect to mosquito-borne diseases and the work has progressed significantly from habitat characterizations at high spatial resolution to work that has inferred habitat quality (and hence disease risk) for both larval and adult forms from a village to a continental scale. International agencies, in cooperation with developing countries, are now able to develop programmes to validate real-time predictions of malaria incidence and prevalence across continental areas. Research relating remotely sensed variables to mosquito life history parameters is an essential basis for future ‘disease forecasting’ for a malaria early warning system (MEWS). This could operate in much the same way that drought and famine conditions are now routinely monitored using satellite sensor data (Hutchinson, 1991), and this is explored further by Myers et al (this volume).

Data capture and mapping, associated with even the most rudimentary GIS analysis, has made evidence-based response by malaria control programmes possible, and many examples have been cited. Maps provide a powerful lobbying tool and contribute towards planning and monitoring malaria control operations through community health care, targeted insecticide spraying and drug distribution. Country maps of continental models of malaria distribution, duration and timing of transmission season, population distribution and distribution of surveys have served to highlight discrepancies between supply and need of resources, and have stimulated discussion about other possibilities and requirements. Equity of access to bed nets at country level, bed net coverage, services (availability of drugs, etc.); market analysis, trade tariffs, distribution network, ‘money season’ versus ‘malaria season’ for
involvement of the private sector, poverty index, and migration maps including refugee camps are all important issues that play a role in national malaria control, which would benefit from spatial representation and analysis. Finally, once certain generic data sets have been generated or gathered (e.g. boundaries, population, topography, cadastral) a GIS system can easily be adapted for further topics in malaria, such as the distribution of drug and insecticide resistance. It can be adapted equally well for other diseases and even for other sectors, such as agriculture or infrastructure planning and development.

The perception that RS and GIS are not appropriate in technologically developing regions persists and is manifest in the form of frequent objections to the cost of image processing equipment, lack of access to imagery, expertise and ground truth and the novelty of the techniques (Bos, 1990; Arambaulo and Astudillo, 1991; Barinaga, 1993; Kleiner, 1995; Oppenshaw, 1996). These problems are real concerns, but are diminishing as (i) computer processing and data storage facilities become relatively cheaper, (ii) image-processing and GIS software has become widely accessible and can automate many tasks, (iii) recent changes in the philosophy of digital image dissemination have resulted in data becoming more widely and freely available (Mulcahy and Clarke, 1994; Justice et al., 1995), and (iv) the history of successful application of RS and GIS techniques in epidemiology expands. An excellent review of many of these issues has been given by Kithsr Perera and Tateishi (1995).

The future will see a further increase in the range of data available as well as its speed of dissemination. This will present new challenges and opportunities for synergism among the epidemiological, RS and GIS research communities. The degree of sophistication with which studies are utilizing high RS data is set to advance (Goetz et al., this volume) including the way in which a variety of images is used to obtain refined information for a given area (Pohl and van Genderen, 1998), the way we deal with scale (Quattrochi and Goodchild, 1997; Walsh et al., 1999) and the more widespread use of SAR (Henderson and Lewis, 1998). We think these efforts would be best focused on addressing the following: (i) defining seasonal variation in disease risk to optimize timing of health information, insecticide treatment of bed nets, intermittent chemotherapy or chemoprophylaxis during pregnancy; (ii) identification of endemic- and epidemic-prone areas to provide information on geographic location of areas amenable to interventions aimed at eradicating infection risk (e.g. residual housespraying or mass drug administration); (iii) epidemic early warning systems to provide timely information on conditions suitable for epidemics (prior to their occurrence); and (iv) population distribution with respect to health service provision for determining optimum health service delivery for new clinical management or bed net re-treatment centres.
ACKNOWLEDGEMENTS

We are grateful to the David Rogers and Sarah Randolph for comments on this manuscript. SIH is an Advanced Training Fellow with the Wellcome Trust (No. 056642) and RWS is a Wellcome Trust Senior Research Fellow in Basic Biomedical Sciences (No. 033340).

REFERENCES

Anonymous (1973). The Use of Remote Sensing in Mosquito Control. Houston, TX: Health Applications Office, NASA.


