Satellites, Space, Time and the African Trypanosomiases

D.J. Rogers

Trypanosomiasis and Land Use in Africa (TALA) Research Group, Department of Zoology, University of Oxford, South Parks Road, Oxford OX1 3PS, UK

ADVANCES IN PARASITOLOGY VOL 47
0065-308X $ 30.00

Copyright © 2000 Academic Press
All rights of reproduction in any form reserved

ABSTRACT

The human and animal trypanosomiases of Africa provide unique challenges to epidemiologists because of the spatial and temporal scales over which variation in transmission takes place. This chapter describes how our descriptions of the different components of transmission, from the parasites to the affected hosts, eventually developed to include geographical dimensions. It then briefly mentions two key analytical techniques used in the application of multi-temporal remotely sensed imagery to the interpretation of field data; temporal Fourier analysis for data reduction, and a variety of discriminant analytical techniques to describe the distribution and abundance of vectors and diseases. Satellite data may be used both for biological, process-based models and for statistical descriptions of vector populations and disease transmission. Examples are given of models for the tsetse *Glossina morsitans* in the Yankari Game Reserve, Nigeria, and in The Gambia. In both sites the satellite derived index of Land Surface Temperature (LST) is the best correlate of monthly mortality rates and is used to drive tsetse population models. The Gambia model is then supplemented with a disease transmission component; the mean infection rates of the vectors and of local cattle are satisfactorily described by the model, as are the seasonal variations of infection in the cattle.

High and low spatial resolution satellite data have been used in a number of statistical studies of land cover types and tsetse habitats. In addition multi-temporal data may be related to both the incidence and prevalence of trypanosomiasis. Analysis of past and recent animal and human trypanosomiasis data from south-east Uganda supports the suggestion of the importance of cattle as a reservoir of the human disease in this area; mean infection prevalences in both human and animal hosts rise and fall in a similar fashion over the same range of increasing vegetation index values.

Monthly sleeping sickness case data from the districts and counties of south-east Uganda are analysed and often show significant correlations with local LST. Case numbers increase with LST in areas that are relatively cooler than average for this part of Uganda, but decrease with LST in areas that are on average warmer. This indicates different seasonal cycles of risk across the region, and may be related to the differing vectorial roles of the two local tsetse, *G. fuscipes* and *G. pallidipes*.

Finally, the increasing pace of change, and the likelihood of new or re-emerging vector-borne diseases, highlight the need for accurate and timely information on habitat changes and the impacts these will have on disease transmission. The next generation of satellites will have significantly more spectral and spatial resolution than the current satellites, and will enable us to refine both statistical and biological predictions of trypanosomiasis and other vector-borne diseases within disease early warning systems.

1. INTRODUCTION

1.1. The Geography of the African Trypanosomiases

The African trypanosomiases affect humans, domestic animals and wildlife over *c*.10 million km^2 of sub-Saharan Africa (Murray and Gray, 1984). The 30 species and subspecies of the tsetse vectors of these diseases are adapted to forest (*Glossina fusca* group), forest and riverine (*G. palpalis* group) and savannah (*G. morsitans* group) vegetation types and each species has unique habitat requirements. Human African trypanosomiasis (HAT), 'sleeping sickness', is a zoonosis with domestic and wild vertebrates recognized as competent hosts for both of the human-infective strains, *Trypanosoma* (*Trypanozoon*) *brucei gambiense* in West Africa and *T. b. rhodesiense* in East Africa. Domestic animal trypanosomiasis, 'nagana', involves mostly *T.* (*Duttonella*) *vivax*, *T.* (*Nannomonas*) *congolense* and, to a much lesser extent, *T. b. brucei*, none of which infects humans (with very rare exceptions) (Hoare, 1972).

Putting together the pieces of this particular epidemiological jigsaw has exercised generations of scientists and administrators in Africa. If the three steps of science are to describe, to explain and then to predict, we began the first step more than a century ago, with Bruce (1895) first associating the symptoms of nagana with trypanosome infection and tsetse with its transmission and Kleine (1909) later showing the cyclical development of the trypanosomes within these insect vectors (Vickerman, 1997). The second step was taken by colonial administrators, scientists and historians, not always holding the same view of this complex problem (Ford, 1971). The colonial administrators saw tsetse and trypanosomiasis as a long-standing scourge on the continent, and devoted a disproportionate amount of colonial research funding (approximately one quarter) to this single group of diseases (Farley, 1991). Scientists saw a more detailed picture of the interaction between three very different systems, those of the parasites, the vectors and the hosts, but their studies initially emphasized one or other element of this triumvirate and only rarely considered the totality, i.e. epidemiology.

1.2. The Trypanosome Parasites

Once sufficiently precise diagnostic techniques were established, the parasitologists were generally struck by the profusion of trypanosome strains circulating in areas with tsetse vectors, and by the antigenic variation of each strain over time within each individual host's bloodstream (Barry, 1997). Trypanosome strains were initially collected and cryopreserved in

parasitological 'stamp collections', which were drawn upon for laboratory experiments, but otherwise unexploited. A question asked in the 1980s of the scientists engaged in the feverish search for yet more strains to add to such unorganized collections was 'When is more, enough?' The answer came when strain variation was fitted into its correct geographical context. We are now able to put geographical pages in our stamp albums; some trypanosome strains are restricted to quite small areas; others are mostly restricted to small areas with occasional outliers hinting at a more widespread distribution, and other strains are already known to be quite widespread (Hide, 1997; Komba *et al.*, 1997).

Individual strains retain their infectivity to host species such as humans over periods of more than 20 years spent in alternative hosts (Ashcroft, 1959), and modern molecular techniques confirm the persistence of relatively homogenous strain groups of trypanosomes in the same geographical areas for periods of at least 35 years (Hide, 1997). This part of the trypanosomiasis story is now near the completion of its descriptive stage: explanations will almost certainly involve elements of space and time that are implicit in the term 'landscape epidemiology' (Kitron, 1998).

1.3. The Tsetse Vectors

Whilst the parasitological work continued, the entomologists were engaged in pioneering work on tsetse ecology. The distribution of the different tsetse species was captured in a series of maps based on point captures of flies and judicious interpolation, often based on elevation that assumed cold temperature limitation of tsetse breeding (Newstead *et al.*, 1924; Ford and Katondo, 1977). Excellent but generally forgotten work in West Africa, following Bodenheimer's pioneering studies of climatic constraints on animal species distributions (Bodenheimer, 1938), began to relate tsetse distributions to a combination of temperature and humidity conditions, captured by climatographs (Gaschen, 1945). Much later this approach was given a quantitative treatment that allowed predictions to be made of continental species' distributions, based on temperature and saturation deficit surfaces (Rogers, 1979; Rogers and Randolph, 1986). Within tsetse distributional limits, a number of studies investigated seasonal variations in fly numbers (Glasgow and Welch, 1962). Jackson's mark–release–recapture studies in the 1930s were advised by Ronald Fisher in Cambridge, UK, and were the first of their kind for any insect species (Jackson, 1930, 1937). It was quickly appreciated that the numbers of tsetse captured by a variety of methods (a relative estimate of population size) was only indirectly related to the numbers of flies present in an area (the absolute population size), and the links between relative and absolute population size estimates continue to be

investigated (Hargrove and Borland, 1994). Both fly distribution and fly abundance have environmental correlates and must therefore be studied in a geographical framework.

1.4. Human Trypanosomiasis

Both Jackson and Nash examined the impact of the environment on tsetse populations and Nash was amongst the first to appreciate that the degree of contact between humans and flies ('man–fly contact') could locally be more important than fly abundance in the transmission of HAT (Nash, 1944, 1948). A small population of flies persisting around a watering point can present a significant risk to humans coming to collect water. The link between vectors and the human host was, however, most critically investigated by Morris in both West and East Africa (Morris and Morris, 1949; Morris, 1951, 1952, 1959, 1960a,b, 1963). Morris showed first how the risk of human disease could be related to the distance of human habitation from the nearest tsetse habitat. Second, he showed that a reduction of this habitat, especially by the removal of riverine forest, could bring about a sustained reduction of both the fly population at watering points and the risk of human infection in such places. Thus both the broad geographical distribution of flies and the fine-scale interactions of flies with habitat features are vital ingredients of our present, rather poor, understanding of HAT.

1.5. Animal Trypanosomiasis

Veterinary research on the African trypanosomiases involved pathology, parasitology and epizootiology. It is clear that the major cattle-rearing regions of Africa are constrained by the presence of tsetse, which exclude the relatively susceptible zebu (*Bos indicus*) races from potentially productive areas. The relatively more trypanotolerant (*Bos taurus*) races are widespread in parts of West Africa, especially in areas of low to intermediate tsetse challenge (neither type of animal can survive in the areas of highest challenge). The impact of trypanosomiasis on much of Africa can be appreciated by examining places where it does not occur, such as the tsetse-free highland areas of Ethiopia and Kenya, and the colder regions of southern Africa; in such places, livestock are integrated into farming systems and contribute to significant increases in productivity in terms of meat, milk, ploughing and manuring (Swallow, 1997). These comparisons led major United Nations agencies to estimate that continental meat production could

be doubled if tsetse could be eradicated from Africa. Maps may be produced showing the potential productivity of tsetse-infested areas, to inform the debate between scientists, politicians and development specialists about how best to tackle the African trypanosomiasis problem.

1.6. Colonial Impacts and HAT Epidemics

The parasitological, entomological and veterinary research outlined above was prompted by political interest in, and concern about, devastating epidemics of human sleeping sickness that swept through Africa at the end of the nineteenth and beginning of the twentieth centuries. Unlike the ubiquitous animal trypanosomiases, HAT is often confined to relatively small foci that appear to have persisted for many centuries (Ford, 1971). European exploration followed rapidly by colonial expansion in Africa at the end of the nineteenth century appear to have spread the disease from these persistent foci into new areas. Thus, beginning in 1912, a wave of HAT outbreaks began in what is now the Democratic Republic of the Congo and spread through West Africa taking 30 years to reach Sierra Leone about 3000 km away (Scott, 1965); prevalences commonly exceeded 5%, and locally reached 50%. At more or less the same time, East Africa began to experience hitherto unknown epidemics of HAT in Uganda and Tanzania, with peak numbers of cases in 1902 and 1929 respectively. Several major expeditions, including those of H.M. Stanley accompanied by thousands of African porters, passed through known disease foci in the basin of the River Zaire (Dutton and Todd, 1906) before returning to East Africa, perhaps carrying the infections that later caused HAT outbreaks in new areas with many susceptible humans (Ford, 1971; McLynn, 1991).

Whilst modern epidemiological theory can explain why the human disease is so focal (Rogers, 1988), historical analysis is required to provide an explanation for the place and timing of HAT outbreaks.

1.7. Alternative Views of the Tsetse/Trypanosomiasis Problem

The impacts of both human and animal trypanosomiases are undoubtedly severe but, towards the end of the last century, arguments were put forward suggesting a beneficial effect of both tsetse and trypanosomiasis: that of preserving habitats from human destruction by overgrazing and desertification (Ormerod, 1976, 1986, 1990; Ormerod and Rickman, 1988). Such problems are apparent in the drier regions of West Africa, where animal production is limited on its southerly borders by tsetse and in the

north by the Sahara desert; such marginal areas appear to show dramatic short-term impacts of heavy use by cattle, although longer term damage is limited, because the local vegetation has a remarkable capacity to recover when rain eventually falls (Hiernaux and Justice, 1986). Once again, therefore, the importance of trypanosomiasis must be seen in a geographical framework.

2. DISEASE TRANSMISSION MODELS

The third step of science, that of prediction, can be made successfully only once all the above influences have been fitted into a quantitative framework. Models for the African trypanosomiases, as for most vector-borne diseases, owe much to Ross's pioneer modelling of malaria, again at the beginning of the twentieth century (Ross, 1909, 1911). Macdonald's development of this model (Macdonald, 1957) provided the stimulus to modelling the more complex African trypanosomiases, with multiple vector and host species (Milligan and Baker, 1988; Rogers, 1988). Because of the complexity of these models, they were generally solved only for their equilibrium predictions, and were used to estimate the basic reproductive number of the trypanosomiases, R_0 (Anderson and May, 1991). The spatial nature of disease risk was not explicitly modelled and the seasonality of transmission was modelled only in a very general way, by imagining simple monthly (sinusoidal) variations in vector numbers (Rogers, 1988). The result of such mathematical exercises was therefore not so much the production of a model for the African trypanosomiases, but rather a model of a model for these diseases. These proto-models told us what we still needed to know about the trypanosomiases, and many of these 'needs' had the dimensions of both space and time.

2.1. Problems of Space and Time

Intensive studies of parasites, vectors or hosts tend to be relatively restricted geographically. Extending the results to other places involves several crucial assumptions, most of which are difficult to justify. For example, will a model for a tsetse population developed for one place apply to another, which has the same tsetse species but a different habitat and probably, therefore, a different set of natural enemies of tsetse? Can models for the natural disease situation be modified to describe the impacts of intervention with trypanocidal drugs or insecticides? These problems of time and space will be reduced by a better understanding of the dynamics of disease transmission

that will come from the development of spatially and temporally explicit disease transmission models.

2.2. Problems of Scale

Related to the problems of space and time are the problems of scale, which become especially important when disease control is being considered. Can the techniques developed for fly suppression in a small area be applied without modification to a much larger area? The answer to this question probably depends on habitat structure and whether or not habitats are scale invariant, i.e. have the same fractal dimension at all spatial scales (Bonham-Carter, 1994; Turner *et al.*, 1998). Can a transmission model for the endemic situation of the disease also describe epidemic outbreaks, i.e. longer time-scale events? The answer to this question depends upon the origin of such epidemics, and whether or not they are intrinsic to the transmission being modelled (i.e. can be described by the standard susceptible–infected–recovered/immune models (Anderson and May, 1991)) or are extrinsically driven by variation in climate or habitats. There is evidence for both in the African trypanosomiases (Rogers and Williams, 1993) and no clear idea of which is the more important. For both sorts of questions of scale, however, a system monitoring variations in both space and time will provide information to test the alternative hypotheses, or at least to rule out some of the alternatives. Satellite imagery provides the potential source of such information.

3. RELATING FIELD AND SATELLITE DATA

The two types of modelling approaches adopted here and elsewhere in this volume fall into the general categories of descriptive (statistical, descriptive) or predictive (biological, process-based) (Table 1). Descriptive models are best developed over large areas that show a variety of environmental conditions, and assume that other places are statistically similar to those for which the original models were developed. For this reason they may be called status quo models. Statistical models can give impressive descriptions of spatial variation in risk but are generally poor at predicting temporal variation in risk because the varying time delays that operate in disease transmission systems are not well captured by the statistical approach. Predictive models, on the other hand, are properly based on a description of the biological interactions between species and therefore, in theory, predict variation through time rather well. Such models, however, tend to be

Table 1 The important differences between statistical and biological approaches to vector and disease mapping.

Type	Requirements	Comments
Statistical	Large-area, multiple layer data sets sampling a wide range of ecological conditions	Extensive studies, status quo models, descriptive
Biological	Long-term study of local dynamics and transmission, measuring demographic rates and variables	Intensive studies, dynamic models, explanatory

developed for relatively small areas (because they are based on local, intensive studies), and so are poor at describing large area spatial variation. The relative strengths of statistical and biological descriptions of disease processes through space and time suggest that their combination should provide us with new and powerful tools for combating complex tropical diseases. Such a combination may be brought about by exploiting remotely sensed satellite data that provide both habitat descriptions for the descriptive approach and measures of important driving climatic variables for process-based models and the predictive approach. Very often the same satellite variable appears to be important in the two rather different approaches (Randolph, this volume).

It is clear from the above that the available satellite imagery (Hay, this volume) will find very different applications in the alternative approaches. The high spatial resolution imagery derived from sensors on the Landsat and Satellite Pour l'Observation de la Terre (SPOT) satellites are ideal for local area descriptions of the habitat (the statistical approach), at a level of detail unavailable from meteorological satellites, but it is only the latter that are capable of regularly monitoring the seasonality of the habitat that influences biological transmission variables (the biological approach). Studies at the regional and continental levels would require such large volumes of high spatial resolution Landsat TM images that meteorological satellite data may be preferred, even for statistical descriptions.

3.1. Image Preparation—Fourier Analysis

Multi-temporal data from Advanced Very High Resolution Radiometer (AVHRR) on board the National Oceanographic and Atmospheric Administration's (NOAA) series of oceanographic satellites and the High Resolution Radiometer (HRR) on the Meteosat meteorological satellites

show a repeated annual pattern of variation that may be used to describe habitat AVHRR seasonality in the different satellite sensor channels. Once the raw data are processed to provide monthly indices correlated with vegetation abundance and activity (the normalized difference vegetation index, NDVI) or meteorological variables (AVHRR—Channel 3 land surface radiance, split-window land surface temperatures, vapour pressure deficit; HRR—Cold Cloud Duration (CCD); Hay, this volume; Goetz, this volume), the data may be subjected to temporal Fourier analysis (see the Appendix) that extracts information about the seasonal cycles of these indices in terms of their annual, bi-annual, tri-annual etc. cycles (or 'harmonics'), each one described by its phase and amplitude. Once defined, the various Fourier harmonics may be recombined to provide a description of the original signal with as much detail as required by the user. Fourier analysis performs three quite separate and useful functions:

(1) It removes noise from the original satellite signal. Noise is generally at a high frequency relative to seasonal events, so the corresponding harmonics may be omitted from the Fourier description to produce a smoothed picture of seasonal change.
(2) It achieves data reduction (ordination) of monthly data sets that often show strong month-to-month correlations.
(3) It achieves data ordination in a way that has an obvious biological interpretation in terms of cycles of seasonal events. This is in marked contrast to other methods of data reduction such as principal components analysis where seasonal events may contribute in complex ways to any number of the principal component axes and images derived from them (Eastman and Fulk, 1993).

Further details of temporal Fourier processing are given in Rogers *et al.* (1996) and examples of the application of Fourier-processed imagery to studying ecological patterns and processes are given in Rogers and Williams (1994).

3.2. Distribution and Abundance Analysis—Maximum Likelihood Methods

The reduced-dimension data set produced by the methods outlined above form the set of predictor variables used to describe the field observations on vectors or diseases. As detailed elsewhere (Curran, this volume; Robinson, this volume), there is a variety of statistical methods that may be used for such descriptions, ranging from 'black-box' techniques such as neural networking, to simple thresholding on single environmental variables such as temperature or rainfall (reviewed by Williams *et al.*, 1992; see also Manel *et*

al., 1999). In general these alternatives fall into two categories; those that assume an underlying uni- or multi-variate normal distribution of the predictor variables in areas of the presence or absence of vectors or diseases, and those that do not. It is often tempting to opt for the statistical flexibility of the latter techniques but the danger of doing so is that we may, in the process, lose sight of the information about the biology of the vectors that the former methods can provide. Whilst it is unlikely that the continental distribution of widespread species such as the mosquito *Anopheles gambiae*, or the tsetse *G. morsitans*, may be described by a single set of multi-variate normal conditions, it does not seem unreasonable to imagine that their distributions are adapted to a series of multi-variate normal conditions, each differing subtly from the next. If this is the case, the application of modified multi-variate techniques will not only allow us to make more accurate statistical prediction of distributions, but will also provide information useful for regional biological models. One useful multi-variate technique, discriminant analysis, is described in the Appendix. This flexible technique can deal both with multiple categories of distribution and abundance data that arise from clustering environmental data, and with non-linearities of the response of these biological data to the environmental variables.

4. SMALL AREA, BIOLOGICAL MODELS

In this section I develop a generic approach to vector-borne disease modelling, and examine the potential contribution of satellite imagery to such studies.

4.1. Models for Tsetse Populations

Generic models for vectors involve components of birth and death, at least one of which must be density dependent. The temperature and humidity dependence of demographic rates may be investigated in the laboratory (e.g. Buxton and Lewis, 1934), but these studies provide only an approximate guide to such rates in field conditions. Tsetse apparently carefully select micro-habitats that are considerably moister (Bursell, 1959) and, at least in the hottest conditions, cooler than ambient (Hargrove and Packer, 1993). The remarkably slow rate of offspring production by tsetse has been described by simple equations involving only air temperature (Glasgow, 1963; Hargrove, 1994). These equations may be used in predictive models using either meteorological or satellite data. One problem of the latter is that satellite sensors record directly only the thermal radiance (reflected or emitted) of the

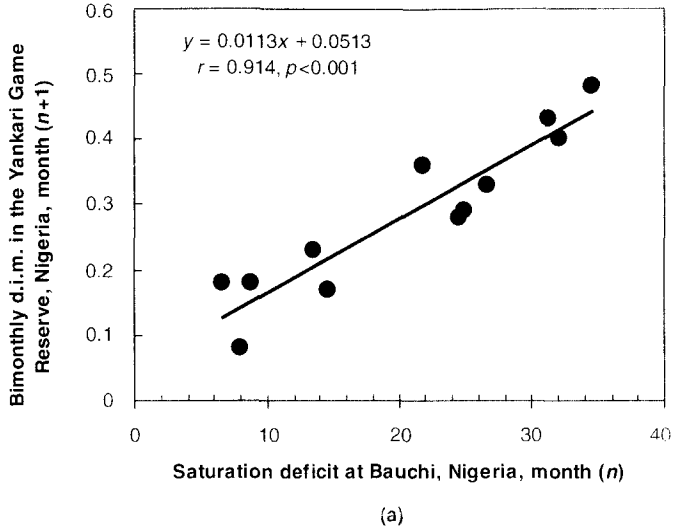
soil and vegetation cover, which is often much higher than air temperature (Hay *et al.*, 1996). Various manipulations of the satellite data can give air temperature estimates with accuracies of a few degrees Celsius (Prihodko and Goward, 1997; Prince *et al.*, 1998; Hay and Lennon, 1999; Green and Hay, 2000; Goetz *et al.*, this volume), sufficiently good for initial models.

In sharp contrast to birth rates, death rates in tsetse appear to depend on both temperature and atmospheric moisture (Rogers and Randolph, 1986), and there is also strong evidence for density dependence at both the puparial and adult stages (Rogers, 1974; Rogers and Randolph, 1984; Rogers *et al.*, 1984).

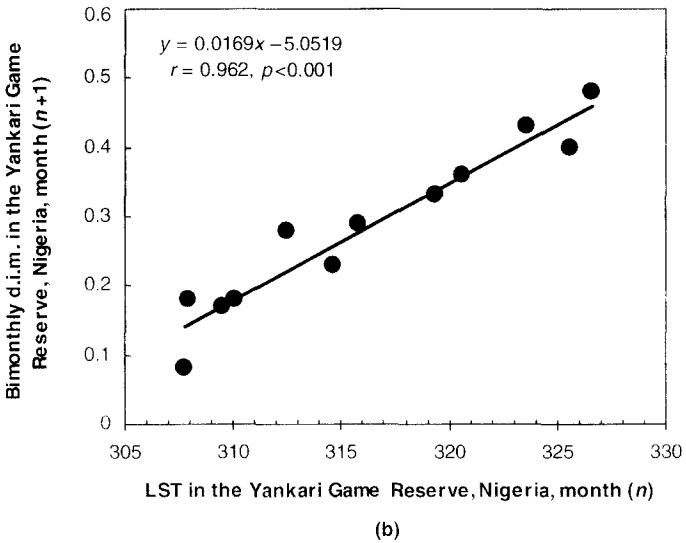
A combination of birth and death rates, each described by the locally appropriate meteorological variables, and with a variable amount of density dependence, successfully described tsetse population changes in both West and East Africa (Rogers, 1990; Rogers *et al.*, 1994). It is now possible to use satellite data as a surrogate for the standard meteorological data. For example, in the Yankari Game Reserve in Nigeria the correlation between the bi-monthly mortality rate of *G. morsitans submorsitans* and LST estimates derived from satellites orbiting more than 800 km above the earth's surface is stronger than the correlation between mortality and saturation deficit (the best ground-based correlate) calculated from meteorological data collected about 50 km from the field site (Figure 1).

A satisfactory description of seasonal changes in tsetse populations is achieved by fine tuning several critical parameter values in the biological models. This fitting process can be automated by steepest descent search methods (Hargrove and Williams, 1998), although a careful check must be kept on parameter values since there appear to be many locally stable equilibria when models are fitted to population data. One example, using the satellite data to predict the monthly mortality rate of *G. m. submorsitans* in Nigeria, is shown in Figure 2. The model also requires some estimate of air temperature for predicting inter-larval and puparial developmental periods. Rather than predicting temperature from published formulae relating land surface radiance to air temperature, the model in Figure 2 included an additional fitted parameter to relate satellite (LST) and air temperature directly. This parameter was varied along with all the others to achieve a least-squares fit of the model to the field data.

This and other fitted tsetse population models suggest that substantial density dependence operates on these vectors (Rogers and Randolph, 1984), although the agents of these mortalities have never been sufficiently investigated. Fly abundance is a product of both the density independent, abiotic mortalities (which may be predicted from satellite data) and density dependent biotic ones, and hence models developed for one area may not be extended to others unless the density dependent components are in some way described by satellite data.



(a)



(b)

Figure 1 Comparison of the relationship between the bimonthly density independent mortality (d.i.m.) of the tsetse *Glossina morsitans submorsitans* in the Yankari Game Reserve in Nigeria and (a) saturation deficit of the previous month derived from meteorological records from Bauchi, 50 km away and (b) land surface temperature (LST) of the previous month from the NOAA series of satellites orbiting >800 km above.

4.2. Models for Trypanosome Transmission

Once the fit of the tsetse model is validated, the disease transmission component may be added. A simple transmission model for the African trypanosomiases, based on the standard susceptible–infected—recovered/immune model (Anderson and May, 1991) is described in the Appendix. This model contains equations describing changes in the proportions of vectors and hosts that are currently incubating infections, and of hosts that have recovered and are immune to re-infection for a period of time (these proportions are usually set to their equilibrium values in models that predict only equilibrium disease prevalences (Rogers, 1988)).

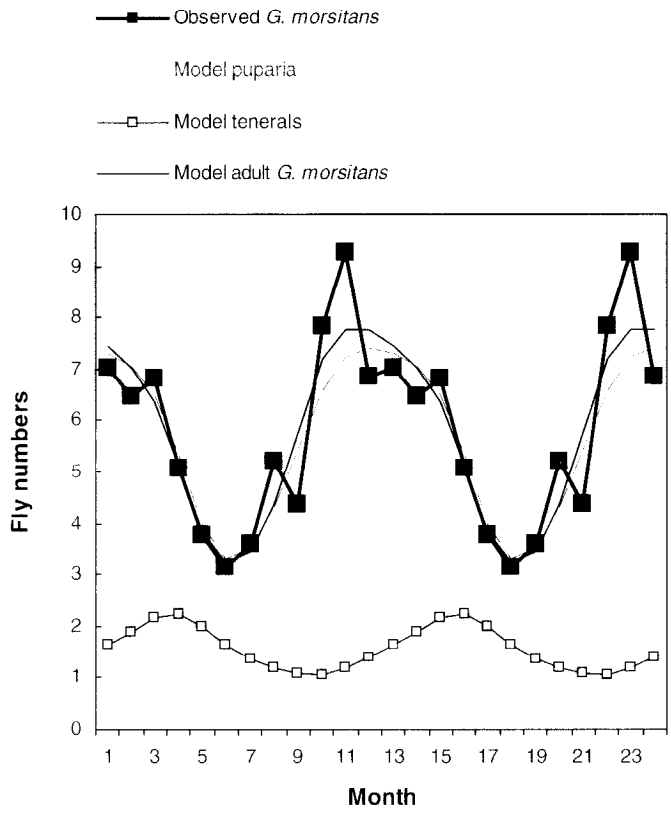


Figure 2 Satellite driven model for the tsetse *Glossina morsitans submorsitans* in the Yankari Game Reserve, Nigeria. The linear relationship shown in Figure 1(b) is used to predict monthly fly mortality rates within a generic population model for tsetse, fitted by least squares methods.

Figure 3 shows the result first of fitting the tsetse model and then of adding the trypanosomiasis component to field observations from The Gambia (data from Rawlings *et al.*, 1991). This model applies to locations where the major vector species is *G. morsitans submorsitans* and the hosts are the local trypanotolerant N'Dama cattle. The various parameters of the transmission equation can be varied to provide an excellent fit to the field data both in terms of the average level of infections in both the vectors and hosts and in the seasonal changes in infection rates. Fly infection rates are highest when

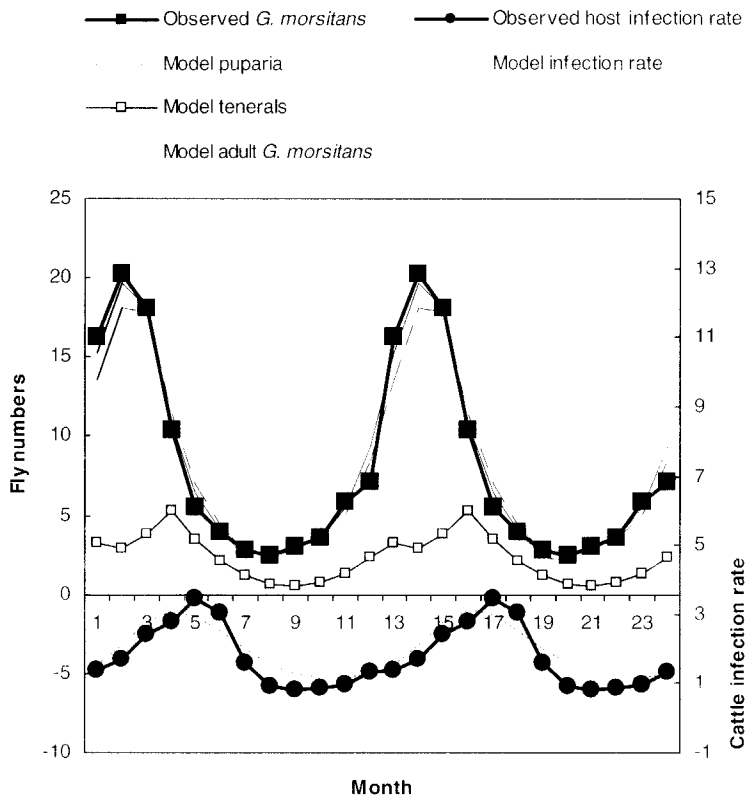


Figure 3 In The Gambia the best satellite correlate of the monthly mortality rate of the tsetse *Glossina morsitans submorsitans* is again the LST. This relationship is used in a least squares-fitted model for this species, to which are later added disease transmission equations from Appendix, C. The resulting trypanosome transmission model describes adequately the mean and seasonal variations in N'Dama cattle infection rates (lower panel; observed mean 1.77% and model mean 1.96%) and the mean fly infection rates (observed mean 2.61%, model mean 2.02%, not shown) (original fly and cattle infection data from Rawlings *et al.*, 1991, 'late dry season' data).

large proportions of the vectors are old; this does not occur at peak tsetse population levels, which include many young flies. Figure 3 shows that the model fly population reaches a peak in February, as do the flies in the field. Fly infection rates show two peaks in the model (not shown in Figure 3), in February and August, whilst field infection rates were variable and showed no significant month-to-month differences. Model host infection rates are highest in April, compared with May in the field, and show similar seasonal changes.

The surprising and encouraging result of this modelling exercise is that remotely sensed satellite data, selected on the basis of current understanding of tsetse dynamics in the field, may be used to drive a fully integrated disease transmission model. The challenge for the future is to see if such models give realistic predictions when extended to other areas, and when subjected to variations mimicking those of natural (climate) and anthropogenic (intervention) changes.

5. LARGE AREA, STATISTICAL MODELS

5.1. High Spatial Resolution Studies

5.1.1. *Vegetation Mapping*

Tsetse live in habitats that provide shade for developing puparia and resting sites for adults. If these habitats can be identified in remotely sensed imagery, not only may tsetse distributions be mapped over very wide areas, but the impact of tsetse on land use over time may be followed in a series of images taken every few years.

Giddings was apparently the first to suggest the use of Landsat Multispectral Scanner (MSC) derived imagery to identify tsetse habitats, in a feasibility exercise that was never followed up (Giddings and Naumann, 1976). A few years later, Bourn in Nigeria showed a significant positive relationship between *G. palpalis* abundance and the size of riverine forest patches as determined from Landsat MSS data (Bourn, 1983).

A recent study in Zimbabwe, using Landsat Thematic Mapper (TM) images taken between 1972 and 1993, concluded that changes in the natural vegetation cover, and in agricultural areas, could not be quantified easily because of spectral differences between similar vegetation types in the different images (Pender and Rosenberg, 1995; Pender *et al.*, 1997). The study concentrated on agricultural areas (Human Dominated Land Use, HDLU) and these were best revealed by applying an edge filter to Landsat TM band 7 and then overlaying the result on a band 3 image. This

emphasized boundaries, and agricultural areas could be relatively easily distinguished and mapped. The study concluded that tsetse-borne trypanosomiasis was only one of several reasons for changing land use patterns in the area, and seldom the dominant one. This study showed that Landsat imagery, even when taken in the same month but in different years, may not be comparable at least partly because the precise timing of seasonal events differs from one year to the next. Rainfall causes a rapid and dramatic change in habitat appearance in many tropical savannahs and is notoriously unpredictable from one year to the next.

Recently, supervised classification of a single SPOT image in Burkina Faso, West Africa, using field observations of vegetation types, was able to produce a detailed classified habitat image, to which were related the mean trap catches of both *G. palpalis* and *G. tachinoides* (La Rocque, 1997). In all, 13 separate vegetation types were involved and the resulting producer's error (i.e. the error in classifying known vegetation types; Ma and Redmond, 1995) ranged from 1% to 61% whilst the consumer's error (i.e. the error in pixels classified as a particular vegetation type) ranged from 1% to 78%.

5.1.2. *Tsetse Flies*

Kitron and colleagues (1996) carried out a detailed spatial analysis of the numbers of *G. pallidipes* caught in several blocks of traps around the Ruma National Park in the Lambwe Valley, Kenya, an area of intermittent HAT transmission. Each trap location was described in terms of vegetation type and canopy cover. Spatial autocorrelation methods were applied and demonstrated strong spatial associations between the trap catches of tsetse. Landsat TM band 7, associated with vegetation and the moisture content of the soil, was the most consistent satellite correlate of fly numbers. Spatial filtering determined the relative contribution to this relationship of spatial and aspatial effects and showed a strong spatially independent association for some of the monthly catches, and for the average catches. This suggests that a large component of the association was the result of other determinants underlying the spatial distribution of both fly numbers and spectral values.

5.2. Low Spatial Resolution Studies

The ability of low resolution multi-temporal advanced very high resolution (AVHRR) data to capture habitat seasonality makes such data ideal for vector-borne disease studies not requiring the full spatial resolution of Landsat and SPOT imagery. AVHRR data have been used in a number of statistical studies to map vegetation, tsetse and disease.

5.2.1. *Vegetation Mapping*

Aerial surveys of Nigeria were carried out in the 1990 wet and dry seasons; every 20 km grid square was surveyed with a 5% sampling intensity at the end of the dry season, and 80% of these squares were re-surveyed at the end of the wet season (Bourn *et al.*, 1994). Digital elevation and the means, maxima, minima and standard deviations of selected raw channel AVHRR data and also processed vegetation indices and thermal data were used in a step-wise, non-linear discriminant analysis of eight main land-cover types (Rogers *et al.*, 1997). At the spatial resolution of the sample grid, many grid squares contained a mixture of vegetation types (and therefore mixed spectral signals for the satellites to detect), so a training set was selected that included only grid squares containing between 50% and 70% of a single vegetation type (sample sizes were too small for higher thresholds). At the 60% threshold, producer's accuracy ranged from 48% to 100%, and consumer's accuracy from 39% to 100% (both figures omitting a small sample of 'bare ground' squares that were misclassified as 'scrub' or 'woodland'). These figures were good enough for the classification signatures to be applied to the whole imagery, and the resulting vegetation map was supported by expert opinion. Many of the classification errors could easily be explained. For example, 'open woodland' was often misclassified as 'dense woodland', and the latter in turn as 'forest'. Cultivated areas were frequently misclassified as 'scrubland', 'open woodland' or 'dense woodland' and more rarely as 'forest'. Humans exploit these land-cover types for their agricultural activities, and trees are often left in place to provide shade for young crops. Although the producer's accuracy for cultivated areas was only 48%, the consumer's accuracy was 92%. Hence we conclude that areas mapped as cultivated by supervised methods are correctly identified, but some areas identified as woodland types may be under at least partial cultivation. As with the high resolution imagery, the identification of cultivated areas is crucial in determining the impact of tsetse and trypanosomiasis on development. Unfortunately, different crops can have very different spectral signatures, making identification of a single agricultural class from satellite imagery very difficult, and it may therefore be necessary to divide study areas into zones according to one or more determinants of crop type (e.g. rainfall) before carrying out supervised classifications on the zoned imagery.

5.2.2. *Tsetse Flies*

Methods for analysing tsetse distribution using AVHRR data were developed in a series of studies of *G. morsitans* and *G. pallidipes* in Kenya, Tanzania and Zimbabwe (Rogers, 1993; Rogers and Randolph, 1993; Rogers and Williams,

Table 2 The relative importance of meteorological satellite variables for predicting the distribution of tsetse flies in Africa.

	Africa (all)	West Africa	East Africa
NDVI	22	8	14
Thermal	8	15	12
Cold cloud duration	16	6	8
Elevation	4	1	3

The table records the number of times each type of variable appeared in the top 10 predictor data sets during discriminant analysis of the distributions of *G. morsitans*, *G. longipalpis*, *G. palpalis*, *G. f. fuscipes*, *G. pallidipes* and *G. tachinoides* either in West or East Africa, or for the whole continent. Thermal variables were AVHRR Channel 4 only. Variables which were a combination of two types (e.g. thermal/NDVI) are not included.

1993) and of the two subspecies of *G. morsitans* in Zambia (Robinson *et al.*, 1997a,b). Linear discriminant analysis, using a single cluster each for presence and absence data, and equal prior probabilities during classification (see Appendix, B), gave predictive accuracies of 80% or greater. These studies revealed several important features of local and regional fly distribution. First, only a single climatic variable appears to determine tsetse distribution at the edge of its continental range (e.g. the maximum of the mean monthly temperature for *G. morsitans* in Zimbabwe). Second, more than one climatic variable is required to describe distributions well within the continental range (e.g. in Kenya and Tanzania). In such areas, one variable excludes the flies from some places and others are more important elsewhere. Third, the average temperature difference between areas of fly presence and absence may be less than 1°C (Rogers and Randolph, 1993). Fourth, the two southern African subspecies of *G. morsitans* appear to respond very differently to climatic variables, but the distribution of each is described well by multivariate methods (Robinson *et al.*, 1997a,b). The studies were later extended to cover the entire continental distribution of tsetse, and the analytical methods adapted to allow clustering of the environmental variables in areas of both presence and absence (see Plate 2). The analyses also allowed the use of different covariance matrices for each cluster, and variable *a priori* probabilities (see Appendix, B), and thus provided the maximum likelihood solutions to the problem of defining tsetse distributions. These further studies suggested that flies show adaptations to regional climate (deduced from the increased accuracy obtained with clustered data) and that in different regions, different sets of variables appear to be important in determining fly distributions. This is illustrated in Table 2 which records the frequency of occurrence of elevation, CCD, temperature or NDVI data in the top ten selected variables used to define the distribution of key species (*G.*

tachinoides in West and *G. pallidipes* in East Africa), or closely related subspecies (*G. morsitans* ssp. and *G. palpalis* ssp. or *G. f. fuscipes*) either at regional or continental levels. In West Africa, temperature variables are almost twice as frequently selected as are NDVI variables (West Africa is on average 2–3°C warmer than East and central southern Africa), while in East Africa NDVI variables are more important than are thermal variables, and elevation also frequently plays a role. At the continental level NDVI dominates, followed by CCD. These statistical analyses in many ways confirm and extend our biological knowledge of tsetse. Regional differences in the genetics and behaviour of *G. pallidipes* were already known (Langley *et al.*, 1984) and population analyses had also suggested that the same species of tsetse showed a different tolerance to atmospheric dryness in the Lambwe Valley of Kenya (where conditions are near to ideal for this species) than in Somalia (Rogers, 1990). A challenge for future research is to investigate whether the environmental variable clusters of a single species' range correspond geographically to genetic differences between flies from the different areas.

Endemic HAT in much of West Africa is especially common in moist savannah regions, but is rare both in the more forested South and in the drier north. In an attempt to explain this distribution, the World Health Organization carried out detailed studies in villages in the affected zone in Côte d'Ivoire and supported a transect study running from the coast more or less directly inland to Bobo Dioulasso, Burkina Faso, about 700 km away. At about 100 km intervals the local populations of *G. palpalis* (which occurs throughout the transect) were sampled in both wet and dry seasons, and various measures were made of fly population behaviour and condition (map and imagery, see Plate 3). One striking result from this study was the demonstration that the sizes of the flies along the transect were similar in the wet season but significantly different in the dry season, when flies in the north were smaller than flies in the south. This result was shown separately for both male and female tsetse and the satellite images confirmed both a lower mean NDVI in the dry season and variations in NDVI values that related to changes in fly size (Figure 4). Tsetse fly size is determined by the parent female at the time of larval development *in utero*; environmentally and nutritionally stressed flies suffer a higher death rate, and also produce smaller offspring about one month later (hence the best correlate of fly size was with satellite data of the previous month). The apparently very small, but significant, variation in fly size recorded along this transect has been associated in other studies on *G. pallidipes* with a four-fold change in fly mortality rate seasonally (Dransfield *et al.*, 1989).

Our tentative explanation for the restriction of HAT to the central zone of the transect is that only in this region are there sufficient numbers of seasonally stressed flies biting humans to maintain the human disease. Flies

in the south are not as seasonally stressed and so avoid biting humans (not a favoured host); to the north, although seasonal stress causes flies to bite humans, there are insufficient vectors to maintain the human disease endemically. This explanation is consistent with the results of large area surveys of this vector species in Côte d'Ivoire that showed the largest numbers of tsetse at intermediate NDVI levels (Rogers and Randolph, 1991).

Habitat signatures in the thermal (LST), vegetation (NDVI) and rainfall (CCD) variables were used in conjunction with ground data in the interpretation of a unique data set from Togo, that included contemporary observations on flies, disease and cattle at a spatial resolution of 0.125 degrees across the entire country (Hendrickx, 1999; Hendrickx *et al.*, 1999a,b; 2000a,b). In selected sites, fly populations and cattle disease were monitored monthly, so that the project recorded both spatial and temporal patterns of vectors and disease. The project confirmed and extended the relationships previously found elsewhere in West Africa, between mean trap catches of *G. tachinoides* and *G. palpalis* and remotely sensed AVHRR NDVI (Rogers and Randolph, 1991), with further relationships between fly abundance and both CCD and LST estimates. The country-wide distributions of these two species and of *G. m. submorsitans* and *G. longipalpis* were subjected to non-linear discriminant analysis involving Fourier processed AVHRR data and were described with accuracies exceeding 90%. Fly abundance, divided into three classes of low, medium and high, were described with accuracies greater than 70% for *G. tachinoides*, but only 56% for *G. palpalis* (the only two species for which abundance data were available).

The study also investigated the effects on accuracy of using a sub-sample of the entire data-set and of changing the numbers of predictor variables used. As might be expected, accuracy tended to diminish with a smaller proportion (and therefore number) of observations in the training set but, within limits, increased with an increasing number of predictor variables. The accuracy of predicting samples *not* included in the training set was maximized with fewer predictor variables than were required to maximize the predictions of the training set itself; this suggests there is a danger of 'over-fitting' a training set so that the results are less generalizable to predict conditions in unsampled areas (Hendrickx *et al.*, 2000a). The study also compared the fly situation in Togo with predictions previously made for Togo on the basis of less accurate tsetse distribution maps for Côte d'Ivoire and Burkina Faso and satellite data for all three countries (Rogers *et al.*, 1996). These previous predictions for Togo were rather poor; whether this was a problem of a rather unsatisfactory training set (old and possibly out-dated maps), or represents a genuine problem of extending such analyses from one place to another requires urgent investigation.

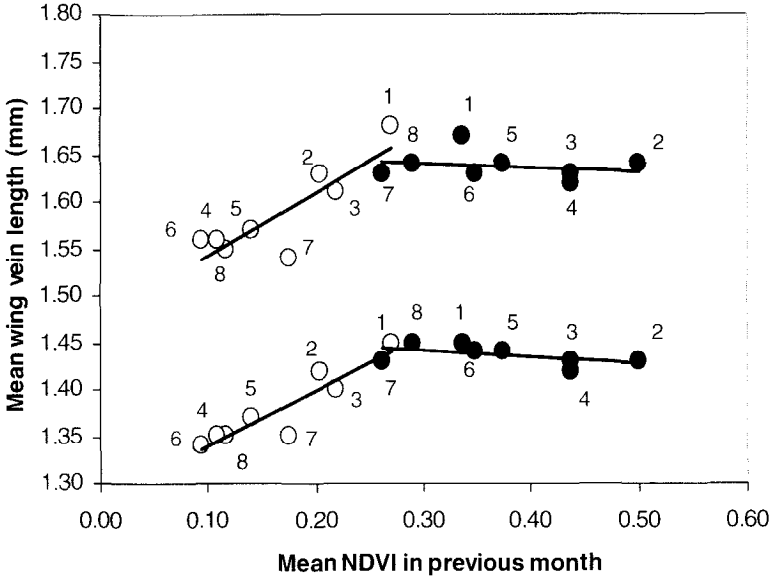


Figure 4 Relationship between the mean vein length of the hatchet cell of female (above) and male (below) *G. palpalis* collected along the transect shown in Plate 3 and satellite derived NDVI of the previous month for the dry (left, open circles) and wet (right, filled circles) season samples. All sites appear to be equally suitable in the wet season, but fly size is significantly smaller in the northern, drier sites in the dry season. Site numbers (see Plate 3 for map and images) are 1, Pauli Brousse; 2, Antonihio; 3, Degbézéré; 4, Bo'Pri; 5, River N'zi; 6, Komborodougou; 7, Oulokoussou; 8, La Guingette. Regression details: wet season, males $y = 1.463–0.072x$, $r = 0.584$, n.s.; females $y = 1.651–0.037x$, $r = 0.201$, n.s.: dry season, males $y = 1.2794 + 0.598x$, $r = 0.915$, $P < 0.01$; females $y = 1.476 + 0.670x$, $r = 0.852$, $P < 0.01$. (Redrawn from Rogers and Randolph, 1991.)

One criticism of the environmental envelope approach to mapping distributions is that it ignores the influence of biotic factors that are so important in determining the abundance of flies within their geographical limits (Davis *et al.*, 1998). We suspect that in most cases there are very few flies at the edge of distributions, and therefore that biotic factors are much less important there than abiotic ones. The fact that we can produce such good fits without explicitly modelling biotic factors supports this idea, although it is also possible that some of the satellite data channels are acting as surrogates for unquantified biotic factors.

Several problems remain for local, regional and continental mapping of tsetse distributions using satellite data, the most important being the ultimate degree of accuracy that such an approach can provide. Tsetse

control personnel want a map that is 100% reliable, but this degree of accuracy is essential only when tsetse eradication is the ultimate aim. Even ground surveys fail to detect some tsetse species (e.g. *G. austeni* which is not attracted to conventional sampling techniques) and no method is particularly efficient at detecting the presence of any all tsetse species at very low densities. The prohibitive costs of total eradication and the difficulty of maintaining this situation over time have caused many to think more realistically of fly control. Control, rather than eradication, requires the identification of tsetse 'hot-spots' which must be targeted by control measures, and of ecological corridors along which flies might move to re-invade controlled areas. Each of these tasks should be more easy to accomplish using satellite data and carefully prepared predictive risk maps.

Combining such information in a GIS and decision making framework can also reveal the desirability of such operations in many locations (Robinson, this volume).

5.2.3. *Animal Trypanosomiasis*

Relatively few studies provide area-wide estimates of the prevalence of trypanosomiasis in domestic animals. The Togo study, mentioned above, collected cattle infection data from sedentary herds in 92% of the grid squares that contained sufficient cattle (approximately two-thirds of the total squares) and thus provides another unique data set for analysis (Hendrickx, 1999; Hendrickx *et al.*, 1999a,b). Overall infection rates, divided into low, medium and high categories, were described with an accuracy of 80%; this figure improved when infections of *T. congolense* (83%) and *T. vivax* (89%) were analysed separately. In discussing these results Hendrickx emphasized that this degree of accuracy was obtained only by using as a predictor variables both the satellite data and anthropogenically determined factors such as herd management and nutrition, breed of cattle, and percentage of land under cultivation. The link between satellites and animal disease is clearly less direct than that between satellites and the tsetse vectors, because the chain of causation is longer.

The Togo study revealed close links between trypanosomiasis prevalence and an easily obtained field measure of anaemia, the packed cell volume, which future area-wide studies could use. It also showed how genes from zebu cattle (i.e. more trypanosusceptible) are introgressing into the trypanotolerant stock, especially in areas where tsetse are disappearing under human pressure on the land. This reveals an elegant degree of adaptation of local farming practices (which may of course arise through trial-and-error processes) to local epizootiological situations. Furthermore, it suggests that development should concentrate on sensible encouragement

of the evolution of the current situation rather than revolutionary, and probably inappropriate, changes to it. As satellite data improve with the next generations of satellites (Goetz *et al.*, this volume) and analytical methods are refined, landscape changes brought about by human activity will be much more easily monitored from space.

5.2.4. *Human African Trypanosomiasis in Uganda*

As in the case of animal trypanosomiasis, there are few extensive data sets for the human disease. Uganda's long history of HAT outbreaks, however, provides us with data that are both historical and contemporary with the satellite imagery. Although other areas in the country have been affected by HAT, the Busoga region that borders the north-eastern shores of Lake Victoria has had the most persistent and well documented HAT focus, with records from the early years of the twentieth century until the present day (Ford, 1971). In the 1960s, Robertson suggested that HAT in this area was an occupational disease, occurring among fishermen in the dry season (with peak incidence in January) and others, mostly agriculturalists, in the wet season (with a peak in May/June; Robertson, 1963). Robertson was never quite able to explain the differences he found, since the seasonal activities of the fishermen did not coincide with their peak risk of infection.

Soon after this, Mwambu and Odhiambo surveyed trypanosomiasis in cattle in the Tororo District of Uganda along a transect from Mjanji, about 8 km from Lake Victoria, inland to about 45 km from the lake (Mwambu and Odhiambo, 1967). *T. vivax* and *T. congolense* were found at all distances from the lake, although *T. congolense* infections fell rapidly beyond about 13 km from the lake, coinciding with the local limits of *G. pallidipes*. *T. brucei* infections (not further identified) were restricted to the lowland forests and thickets favoured by this species of tsetse, and thus did not occur beyond the 13 km limit. These observations tended to confirm the earlier impressions that *G. pallidipes*, a species that readily feeds on domestic and wild animals, is the major source of introduction of *T. b. rhodesiense* into human populations from these alternative hosts. Once introduced into the human population, infections are then rapidly transmitted by *G. fuscipes*, which more readily bites humans and other primates.

Although there are no contemporary satellite data, we can relate the point prevalences of Mwambu and Odhiambo's cattle survey to satellite NDVI for the same sites recorded in 1984–88 (Figure 5a). These relationships were confirmed by two Cambridge undergraduate expeditions to the area in 1991 and 1992, which examined infection rates in cattle along transects determined by the NDVI satellite imagery (J. Hodgson, *pers. comm.*). A new epidemic of HAT began in this area in the 1970s. In many

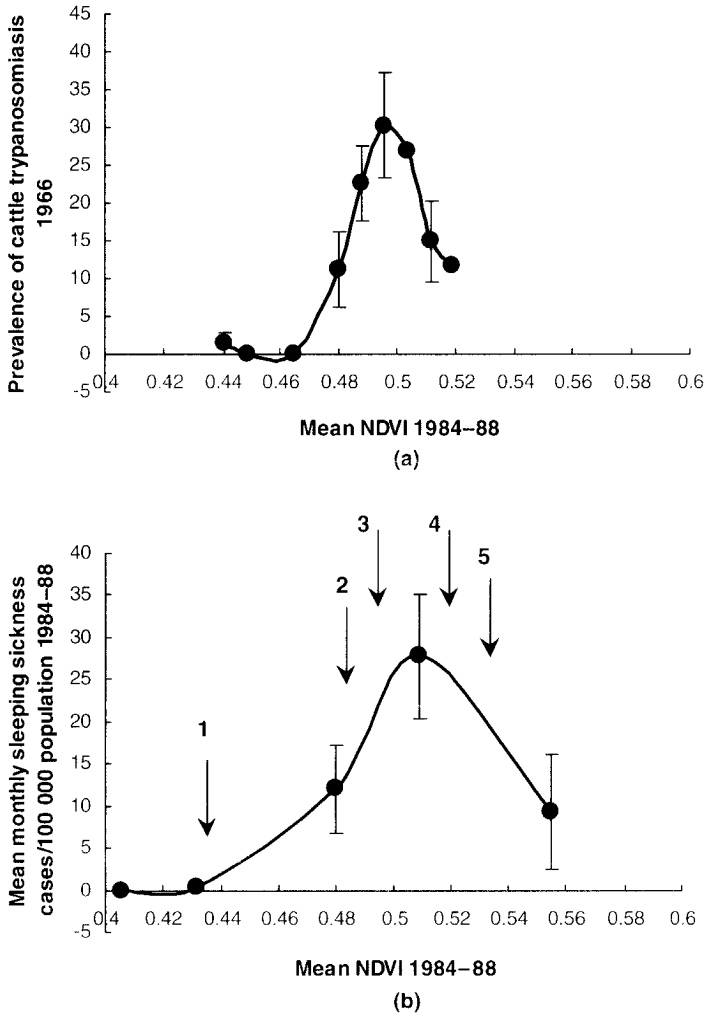


Figure 5 (a) Relationship between the mean trypanosome prevalence in groups of village cattle in the Busoga region of south-eastern Uganda in 1960s (data from Mwambu and Odhiambo, 1967) and mean annual satellite NDVI 1984–88 for the sample sites. (b) Relationship between the mean monthly prevalence of human sleeping sickness (cases per 100 000) in the same region of Uganda between 1984 and 1988 and mean NDVI from each county, for the same period. The similarity of these two curves lends support to the idea that cattle are quantitative reservoirs of the human disease in this area (Hide, 1997). The arrows in (b) indicate the mean NDVI of the local districts; 1, Tororo; 2, Kamuli; 3, Iganga; 4, Jinja; 5, Mukono. The figures also show ± 1 standard error of the means.

ways it repeated the pattern shown during the previous outbreak in the 1940s (Ford, 1971). Starting near Jinja and the source of the White Nile, the epidemic spread slowly east, towards Tororo and the border with Kenya. In the 5 year period 1984–88, the mean prevalence of HAT in the sub-counties of the region showed a remarkably similar pattern of variation to that of the cattle surveys, across approximately the same range of NDVI values (Figure 5b). This is evidence, albeit indirect, that cattle are a potentially important reservoir of the human disease in this region; this idea originated with the isolation of human-infective *T. brucei* from cattle in an ecologically similar lake-side site in Kenya in 1966 (Onyango *et al.*, 1966) and was more recently confirmed by parasitological surveys in this area (Hide, 1997). It is likely that cattle are particularly important as a reservoir of HAT in this part of Uganda because they are relatively resistant to the local strains of *T. vivax* (*T. congolense* is rare in the area). In the absence of any pathogenic infections in their animals, local cattle owners rarely use trypanocides, which would otherwise kill off both the pathogenic trypanosomes and the (to cattle) non-pathogenic *T. brucei* spp. among which are human-infective forms.

The relatively fine temporal and spatial resolution of the HAT data from the Busoga region (cases recorded monthly at sub-county level) allows us to investigate the relationships between the incidence of disease in humans and the satellite data. Before analysis, the monthly case numbers were expressed as a percentage of each year's total, to remove the effects of variation in the annual totals as the epidemic moved slowly through the area. Counties were excluded from the analysis in years when they had fewer than 12 cases. Finally the monthly percentage figures were averaged for the 5 years 1984–88 inclusive, for comparison with the AVHRR data for the same years. Correlations of case numbers with LST were significant at the 5% level or better in 12 out of 20 counties, with a lag of 0, 1 or 2 months (the approximate delay between a person receiving an infective fly bite and presenting with HAT at a local clinic) (Figure 6). The slopes of the significant relationships were significantly inversely related to the mean land surface temperatures for each county, and positively related to the respective mean NDVIs; they could therefore be positive in some areas and negative in others (Figure 7). Where temperatures are relatively cool, and NDVIs relatively high, HAT case numbers increase with monthly temperatures and are most frequent in the hottest months, at the start of the year. Where temperatures are higher, by an average of only 1 or 2°C, in the east of the area near Tororo, HAT case numbers decrease with monthly temperatures and cases peak in June, the coolest time of the year. We can thus reinterpret Robertson's observations by suggesting that HAT in this area is not so much a disease of occupation (fishermen vs. non-fishermen) as a disease of location.

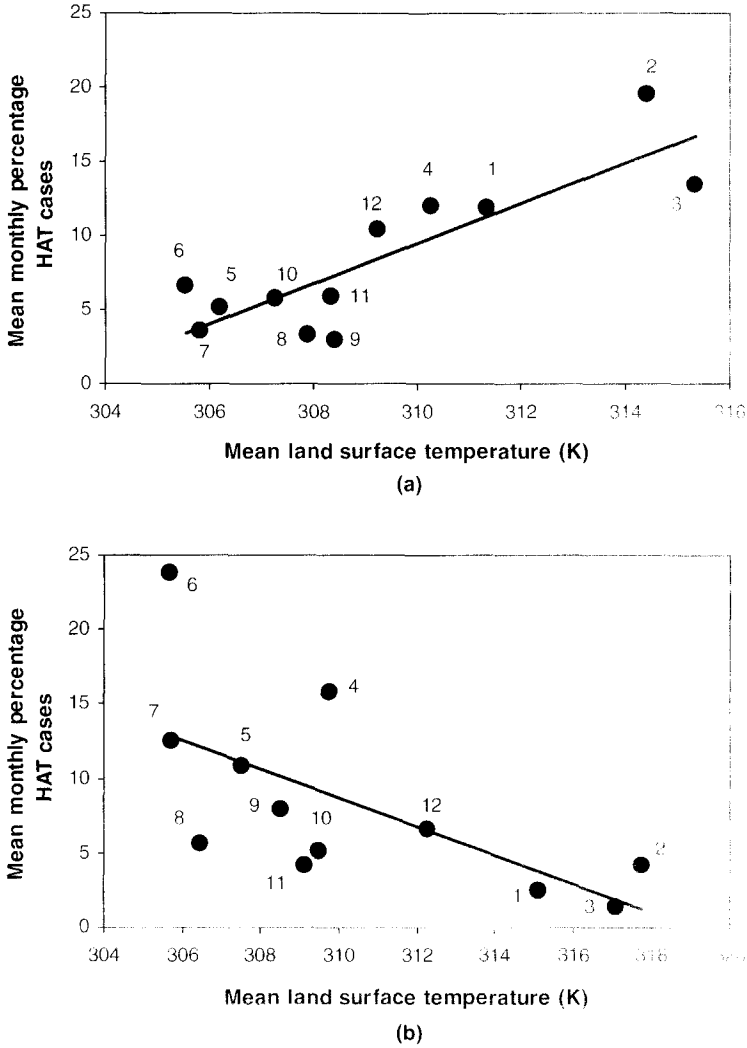


Figure 6 Relationship between the mean monthly numbers of sleeping sickness cases in the Busoga region of Uganda and satellite derived land surface temperature (LST) for the same month and district, 1984–88. (a) Kamuli District, Bulamogi County ($y = 1.361x - 412.3$, $r = 0.847$, $P < 0.001$); (b) Tororo District, Tororo County ($y = -0.972x + 310.05$, $r = 0.641$, $P < 0.05$). Case numbers are expressed as a percentage of the annual totals for each county; months are indicated by numbers 1 = January etc.

Analysis of tsetse fly catches made in this area show that the mean monthly mortality of one of the local vector species, *G. fuscipes*, is highest in the locally hotter season, whilst that of the other, *G. pallidipes*, is highest in the locally cooler season (D.J.Rogers, unpublished observation). The different responses of these two species to effectively the same climate, suggests that their relative

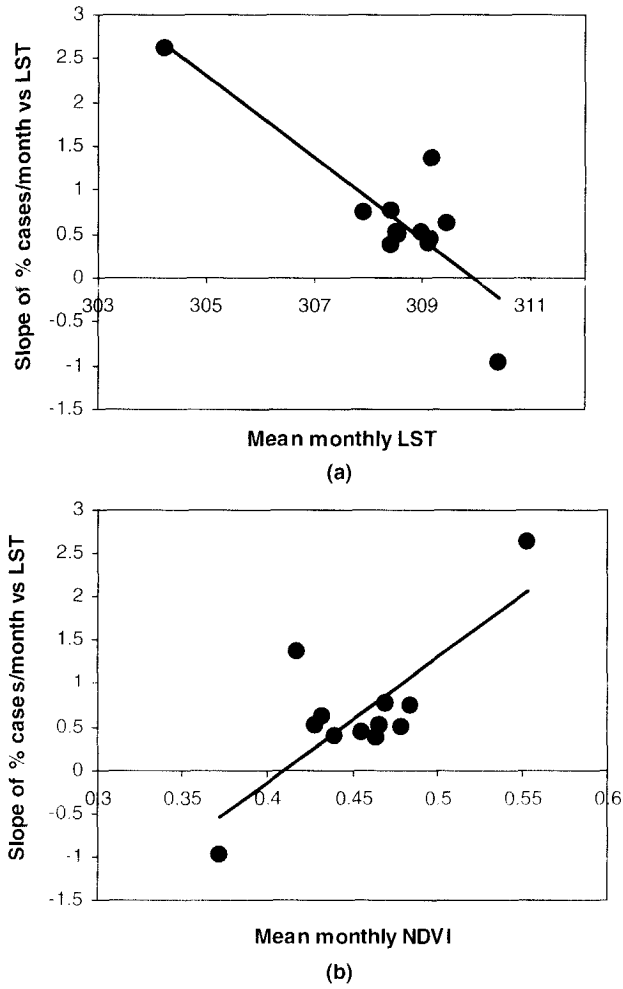


Figure 7 The regression coefficients shown in Figure 6 are significant in 12 of the 20 counties of south-eastern Uganda affected by HAT during the epidemic of the 1970s and 80s. Here these significant coefficients are related to (a) the mean land surface temperature (LST) ($y = -0.467x + 144.69$, $r = 0.855$, $P < 0.001$) and (b) the mean NDVI ($y = 14.344x - 5.876$, $r = 0.777$, $P < 0.01$) for each county for the period 1984–88.

importance as HAT vectors will change seasonally, and perhaps spatially. As in the case of the West African study (see above), tsetse tend to bite humans only when especially stressed and the seasonal risk to humans in this part of Uganda will therefore depend upon the local mix of these two vector species. There is anecdotal evidence that *G. pallidipes* numbers in the eastern end of this fly belt have fallen considerably in recent years, as a result of human impacts, although it survives in the lakeside vegetation nearer to Jinja.

Epidemics have raged through this area of Uganda for more than 100 years and sequential disease management strategies (resettlement, aerial spraying with insecticides, or trapping) have complicated the interpretation of the entomological, parasitological and clinical data from the region. Probably the only certainty is that there will be further epidemics in the future, and satellite imagery may be used to monitor and perhaps predict where the next epidemic is likely to begin, and to generate predictive maps to highlight populations at greatest risk of infection.

6. SATELLITE IMAGERY AND DISEASE RISK MAPPING: THE WAY FORWARD

This chapter has illustrated how remotely sensed data may be used to map habitats, vector distribution and abundance, and disease incidence and prevalence over time and space. The unprecedented rates of change imposed by dynamic and expanding human communities upon their natural habitats have several consequences. First it is less and less likely that static paper maps will be an acceptable way to store information about the distribution of natural resources, or disease risk (Myers *et al.*, this volume). Secondly, it becomes more likely that old diseases will occur in new areas, or new diseases will emerge from zoonotic reservoirs. Thirdly, more problems will be transnational, or will originate from travellers returning from further afield. Fourthly, the ability of traditional veterinary and medical services to cope with such problems will be severely limited; every delay in the response to a new disease situation will result in more people or animals eventually becoming infected, and greater strains will be placed on resource providers. Vector-borne diseases have basic reproductive numbers which are often orders of magnitude greater than those of directly transmitted diseases (Rogers and Packer, 1993) and this will result in rapid increases in case numbers when conditions are suitable. Finally, the tragedies of civil conflicts are often accompanied by epidemic disease outbreaks, most recently in Uganda, the Democratic Republic of Congo and Sudan, placing additional burdens on local health services. For these and many other reasons the world community needs to establish disease forecasting and early warning systems

(Myers *et al*, this volume.). These must couple the best statistical analyses of current situations with our developing ability to produce robust biological models (Randolph and Rogers, 1997; Randolph, this volume) in which a variety of intervention strategies may be tested before they are applied in the field. As emphasized in this chapter, remotely sensed data may be used in both the statistical and biological models and provide a vital link between the two approaches. The benefit is a more rapid advance in our understanding of variations of disease systems in both space and time.

Several problems remain to keep research agendas full. First we need to establish the levels of accuracy of the predictions of both statistical and biological models. This is difficult to do when the field data themselves contain unquantified errors: vector distribution maps drawn by interpolation using elevation contours (as seems to have been the case for Zimbabwe's historical tsetse maps) are likely to be inaccurate, at least in detail, but will be well described by environmental variables that vary with elevation. Statistical predictions of distributions using satellite and other data are probably most useful where they are least accurate; investigation of both false positive and false negative areas will give us more insight into the real determinants of local distributions, of both vectors and diseases, and thus lead to better maps.

Secondly, we need to establish generic ways of collecting field data that will be more useful for the modelling approaches. Given limited resources, is it best to distribute them thinly to produce an extensive picture of the disease situation (ideal for statistical modelling), or to concentrate them into a single area, thus producing data for detailed biological models? Satellite data may be used both to select where the extensive data should be collected (to sample as wide a variety of habitats as possible) and to identify high risk areas that would benefit most from the intensive, biological studies.

Thirdly, we urgently need to increase the spatial resolution of our multi-temporal data. It is remarkable that we can achieve quite so much with the present poor spatial resolution (generally 8×8 km pixel size) of these data: tropical habitats are extremely heterogeneous spatially, and many vectors are effectively spatial opportunists, inhabiting a very small, but precise, fraction of the entire habitat. In the short term we can use a variety of wavelet techniques (Stollnitz *et al.*, 1996; Mallat, 1999) to increase the spatial resolution of past data sets artificially: these techniques offer the best way to integrate the high spatial resolution of the Landsat and SPOT series of satellites with the high temporal resolution of the NOAA and Meteosat satellites. In the longer term, the increased spatial (and spectral) resolution of the next generation of satellites (Hay, this volume; Wood *et al.*, this volume) will improve our discrimination of vector and disease habitats.

Fourthly, temporal variation in disease risk is least amenable to analysis using the present generations of both geographical information systems and

remote sensing software. The research challenge here is to couple the increasingly sophisticated understanding we have of the intrinsic determinants of temporal variation in risk (Anderson and May, 1991) with our still relatively poor understanding of the extrinsic determinants, from seasonal weather patterns to longer time-scale climatic variation. To address this challenge we may need to adopt some of the analytical techniques of time-series analysis for intensive study data (Chatfield, 1980; Hay *et al.*, 2000) and of oceanography and meteorology for extensive study data (Storch and Zwiers, 1999).

Finally, having put all of the above together, we must concentrate on the delivery of the ideas and approaches to the health-care providers who are in the best position to benefit from them. Whilst the habitat sensing may be 'remote', product delivery must be personal, accurate and timely. The problems to be addressed here involve the ownership both of the field data and of the developed methodologies; clearly a mutually beneficial constructive dialogue must be established between the national and international partners. The development of prototype disease early warning systems (Myers *et al.*, this volume) is a first step along this exciting path to the future.

Models of one sort or another are a vital ingredient of all remote sensing activities and are especially important when satellite data are coupled to biological studies. Whilst modellers would do well to adopt as a daily mantra the warning that 'all models are wrong, but some are useful' (Anderson and May, 1991), their inspirational creed should be that of T. S. Eliot (1944):

> We shall not cease from exploration
> And the end of all our exploring
> Will be to arrive where we started
> And know the place for the first time.

ACKNOWLEDGEMENTS

I thank John Townshend, Tim Robinson, and Fred Snijders for the supply of my first AVHRR NDVI imagery in 1986, and the Pathfinder program (NASA—Goddard Space Flight Center), and Compton Tucker and the GIMMS Group at NASA-GSFC for more recent data.

Dr Mbulamberi kindly provided the HAT data for Uganda.

Sarah Randolph and Simon Hay commented on the manuscript, and research support came from the Department for International Development under Schemes R5794 and R6626, administered through the Natural Resources Institute International.

APPENDIX

A. Temporal Fourier Analysis

Temporal Fourier analysis describes variations through time of satellite signals as a series of simple sine curves with different frequencies and amplitudes. It is applicable to regularly collected data such as maximum value composited monthly AVHRR data, $\{x_t\}$.

The Fourier series representation of $\{x_t\}$ is found from the following:

$$x_t = a_0 + \sum_{p=1}^{N/2-1} [a_p \cos(2\pi pt/N) + b_p \sin(2\pi pt/N)] + a_{N/2} \cos \pi t \qquad (t = 1, 2, \ldots, N) \tag{1}$$

with coefficients $\{a_p, b_p\}$ defined as follows:

$$\left.\begin{aligned} a_0 &= \bar{x} \\ a_{N/2} &= \frac{\sum (-1)^t x_t}{N} \\ a_p &= \frac{2\left(\sum x_t \cos(2\pi pt/N)\right)}{N} \\ b_p &= \frac{2\left(\sum x_t \sin(2\pi pt/N)\right)}{N} \end{aligned}\right\} p = 1, \ldots, N/2-1 \tag{2}$$

The component at a frequency $\omega p = 2\pi p/N$ is called the pth harmonic, and for all $p \neq N/2$ these harmonics may be written in the equivalent form

$$a_p \cos \omega_p t + b_p \sin \omega_p t = R_p \cos(\omega_p t + \theta_p) \tag{3}$$

where

R_p = the amplitude of the pth harmonic

$= \sqrt{(a_p^2 + b_p^2)}$

and

θ_p = the phase of the pth harmonic

$= \tan^{-1}(-b_p/a_p)$

Thus the important elements of these equations are a_p and b_p that describe the relative mix of sine and cosine terms that together determine the timing of the peak of each harmonic, p. Further details of time series analysis are given by Chatfield (1980) and Diggle (1990).

B. Discriminant Analysis

In its simplest form, discriminant analysis assumes both multivariate normality and a common within-group covariance of the variables for all points defining vector or disease presence and absence. Covariances are estimated from representative samples from reliable distribution maps, the 'training sets'. Means of multivariate distributions are referred to as centroids and are defined by mathematical vectors $[\mathbf{x}_n]$ where n is the number of dimensions (= variables). The Mahalanobis distance, D^2, is the distance between two multivariate distribution centroids, or between a sample point and a centroid, and is defined as follows:

$$\begin{aligned} D^2_{12} &= (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)' \mathbf{C}_w^{-1} (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2) \\ &= \mathbf{d}' \mathbf{C}_w^{-1} \mathbf{d} \end{aligned} \tag{4}$$

where the subscripts refer to groups 1 (e.g. for vector absence) and 2 (e.g. for vector presence), $\mathbf{d} = (\bar{\mathbf{x}}_1 - \bar{\mathbf{x}}_2)$ and $\mathbf{C}_w^{-1}$ is the inverse of the within-groups covariance (= dispersion) matrix (Green, 1978). Thus D^2 is the distance between the sample centroids adjusted for their common covariance. Equation (4) may be used in a number of ways. First it may be used to assign new data points to one or other category (of presence or absence) by examining the value of D^2 between each point and each of the training-set defined centroids. The point is then assigned to the group for which D^2 is a minimum. Secondly, the equation may be used to calculate the probability with which each data point belongs to each of the training set groups. This involves defining the position of the point within each multivariate distribution around each centroid (most easily achieved by calculating D^2 which is distributed as χ^2 with $(g-1)$ degrees of freedom, where g is the number of variables defining each centroid). In general these measures are normalized by dividing each by the sum of all measures (i.e. the sum of the probabilities across all classes in the training set) to give posterior probabilities, defined as follows:

$$\begin{aligned} P(1 \mid \mathrm{x}) &= \frac{p_1 \mathrm{e}^{-D^2_{12}/2}}{\sum_{g=1}^{2} p_g \mathrm{e}^{-D^2_{g2}/2}} \\ P(2 \mid \mathrm{x}) &= \frac{p_2 \mathrm{e}^{-D^2_{22}/2}}{\sum_{g=1}^{2} p_g \mathrm{e}^{-D^2_{g2}/2}} \end{aligned} \tag{5}$$

where $P(1 \mid \mathrm{x})$ is the posterior probability that observation $\mathbf{x}$ belongs to group 1 and $P(2 \mid \mathrm{x})$ the posterior probability that it belongs to group 2 (Green,

1978). In equation (5), p_1 and p_2 are the prior probabilities of belonging to the same two groups respectively, defined as the probabilities with which any observation might belong to either group given prior knowledge or experience of the situation. In the absence of any prior experience it is usual to assume equal prior probability of belonging to any of the groups; in the simple case of two-group discrimination, therefore, $p_1 = p_2 = 0.5$. The normalization invoked by summing the probabilities in equation (5) is based on the assumption that observation x must come from one or other of the classes defined in the training-set data (this important assumption is frequently not tested by examining the absolute value of the Mahalanobis distances that are used in the calculations of $P(1|x)$ and $P(2|x)$, although image processing systems are usually capable of optional output of an image of these distances). Other terms of the multivariate normal distribution in equation (5) then cancel out (Tatsuoka, 1971).

As indicated earlier, equations (4) and (5) should be modified when the assumption of common covariances is obviously invalid. Not only may areas of presence and absence differ in their environmental characteristics, but different parts of a species range may also show more subtle differences, requiring separate multivariate descriptions of their climatic conditions. It is generally convenient to anticipate this need in the case of widespread species, and to define clusters of environmental similarity of areas of presence and absence before statistical analysis. Each cluster (either for presence or absence) is then treated as a separate multi-variate normal distribution, with its own covariance characteristics, and the posterior probabilities are calculated by summing across all distributions. In the case of two groups only (one for presence and one for absence), equation (5) is then modified as follows:

$$P(1|x) = \frac{p_1 |\mathbf{C}_1|^{-1/2} e^{-D_1^2/2}}{\sum_{g=1}^{2} p_g |\mathbf{C}_g|^{-1/2} e^{-D_g^2/2}}$$

$$P(2|x) = \frac{p_2 |\mathbf{C}_2|^{-1/2} e^{-D_2^2/2}}{\sum_{g=1}^{2} p_g |\mathbf{C}_g|^{-1/2} e^{-D_g^2/2}} \qquad (6)$$

where $|\mathbf{C}_1|$ and $|\mathbf{C}_2|$ are the determinants of the covariance matrices for groups $g = 1$ and 2 respectively. The Mahalanobis distances in equation (6), calculated from equation (4), are evaluated using the separate within-group covariance matrices $\mathbf{C}_1$ and $\mathbf{C}_2$ (Tatsuoka, 1971). When there is more than a single class of presence or absence data (e.g. multiple clusters) the summation in the denominators of equation 6 covers the entire set of $g > 2$ groups and there are as many posterior probability equations as there are

groups. With unequal covariance matrices the discriminant axis (strictly speaking a plane) that separates the two groups in multivariate space is no longer linear, and equation (6) then effectively defines the maximum likelihood solution to the problem (Swain, 1978).

Training-set data are generally limited in size, and it is unwise to subdivide them by clustering if this results in too few observations in some of the clusters. This will cause the covariance matrices to be ill-characterized, and their inversions (equation (4)) will either be impossible to calculate or will perform badly in predicting presence and absence through use of equation (6). Similarly there is no obvious rule about the use of expected or observed prior probabilities in equations (5) or (6). Use of observed (generally training-set) prior probabilities shifts the equi-probability contours towards the smaller groups, resulting in a larger proportion of assignments to the classes with larger group sizes. This shift, however, may occasionally be large enough to reduce the accuracy of describing even the training-set data.

Ideally the training set should be divided, with half used to develop the covariance matrices and the other half used to test the accuracy of the predictions. Frequently, however, training data are scarce and the entire data set must be used in the training exercise. The resulting predictions will tend to inflate estimates of the accuracy of the techniques, although perhaps only modestly (Randolph, this volume). A further problem is that environmental variables are often spatially auto-correlated, whilst the above multivariate methods assume no such auto-correlation, i.e. statistical independence of training-set observations. The importance of these effects may be judged by including some measure of spatial covariation in the analyses. This is done by generating new variables of the auto-correlations, and by making these variables available for inclusion during step-wise discriminant analysis. If they are selected, the spatial effects contribute significantly to the distributions being modelled. In some analyses spatial covariation affects the variance of the resulting estimates but not their means (Thomson *et al.*, 1999), whilst in others both means and variances appear to be affected (Augustin *et al.*, 1996). It is a curious feature of satellite data sets that spatial covariation creates problems for some sorts of analyses but is actively exploited by other methods of analysis, such as co-kriging (refer to Robinson, this volume). Further details of multivariate analysis may be found in Tatsuoka (1971), Green (1978) and Krzanowski and Marriott (1995).

C. A Simple Model for Trypanosome Transmission

Changes in the proportions of the vertebrate hosts that are susceptible (s), infected but not yet infectious (f), infectious (x) and recovered, immune (i) are described by the following set of equations:

$$\begin{aligned}
\frac{ds}{dt} &= -abmx's + tN + wi - \mu s \\
\frac{df}{dt} &= abmx's - vf - \mu f \\
\frac{dx}{dt} &= vf - rx - \mu x \\
\frac{di}{dt} &= rx - wi - \mu i
\end{aligned} \tag{7}$$

where

a is the biting rate of vectors on hosts,
b is the transmission coefficient from vector to vertebrate,
m is the ratio of vectors to hosts ($= M/N$),
x' is the proportion of infected vectors,
t is the birth rate of the hosts,
N is host population size (the present equations describe proportions, so that $N = s + f + x + i = 1$),
w is the rate of loss of immunity, returning immune animals to the susceptible category, s,
μ is the host's natural death rate,
v is the incubation rate of the disease in the vertebrate hosts, and
r is the rate of recovery of the vertebrates from infection.

The equivalent equations for the vectors, continuing to use prime to indicate vector parameters and variables analagous to those of the hosts, are as follows:

$$\begin{aligned}
\frac{ds'}{dt} &= -acxs' + t'M - \mu's' \\
\frac{df'}{dt} &= acxs' - v'f' - \mu'f' \\
\frac{dx'}{dt} &= v'f' - \mu'x'
\end{aligned} \tag{8}$$

where, in addition,

c is the transmission coefficient from vertebrate to vector,
t' is the birth rate of the tsetse population,
M is the tsetse population size (as in the case of the hosts, $M = s' + f' + x' = 1$), and
μ' is the tsetse mortality rate ($= t'$ at equilibium).

Vector mortality rate appears explicitly in these equations since it is assumed that fly infections can be lost only when infected flies die. These

losses are balanced by births, which introduce new, susceptible flies into s'. In the case of the vertebrate hosts, animals which lose their immunity recycle into the susceptible category.

The sum of each set of equations, for both hosts and vectors, is zero, indicating that although each subpopulation (susceptible, infected etc.) may change over time there is no net change in the summed proportions, which must always be equal to 1.0.

The above equations, which apply to the simple situation of a single-vector, single-host disease may be applied to the African trypanosomiases assuming all other hosts fed upon by flies are negligible sources of infection compared to the modelled hosts. The equations are therefore simplified versions of those that have been written for the African trypanosomiases (Rogers, 1988) and may be taken to apply to the situation of trypanosomiasis in domestic animals in areas with few alternative hosts; this simplification allows us to model disease transmission quickly, and to estimate the need for more complex transmission models in field situations. HAT involving domestic and wild reservoir hosts will not be satisfactorily modelled in this simple way.

In making the output of the tsetse population model one of the inputs into the disease transmission model, additional scaling parameters are required to relate the tsetse numbers to the vertebrate host numbers. Since the vector/host ratio (m) always and only appears in the above equations with the transmission coefficient b, estimates of these two quantities will vary together so that their product remains the same.

REFERENCES

Anderson, R.M. and May, R.M. (1991). *Infectious Diseases of Humans: Dynamics and Control*. Oxford: Oxford University Press.

Ashcroft, M.T. (1959). A critical review of the epidemiology of human trypanosomiasis in Africa. *Tropical Diseases Bulletin* **56**, 1073–1093.

Augustin, N.H., Mugglestone, M.A. and Buckland, S.T. (1996). An autologistic model for the spatial distribution of wildlife. *Journal of Applied Ecology* **33**, 339–347.

Barry, J. (1997). The biology of antigenic variation in African trypanosomes. In: *Trypanosomiasis and Leishmaniasis: Biology and Control* (G. Hide, J.C. Mottram, G.H. Coombs and P.H. Holmes, eds), pp. 89–107. Wallingford: CAB International.

Bodenheimer, F.S. (1938). *Problems of Animal Ecology*. Oxford: Clarendon Press.

Bonham-Carter, G.F. (1994). *Geographical Information Systems for Geoscientists*. Oxford: Pergamon/Elsevier.

Bourn, D. (1983). Tsetse control, agricultural expansion and environmental change in Nigeria. Ph.D. thesis, University of Oxford.

Bourn, D., Wint, W., Blench, R. and Wolley, E. (1994). Nigerian livestock resources survey. *World Animal Review* **78**, 49–58.

Bruce, D. (1895). *Preliminary Report on The Tsetse Fly Disease or Nagana in Zululand*. Durban: Bennet and David.

Bursell, E. (1959). The water balance of tsetse flies. *Transactions of the Royal Entomological Society of London* **3**, 205–235.

Buxton, P.A. and Lewis, D.J. (1934). Climate and tsetse flies: laboratory studies upon *Glossina submorsitans* and *tachinoides*. *Philosophical Transactions of the Royal Society of London, Series B* **224**, 175–240.

Chatfield, C. (1980). *The Analysis of Time Series: An Introduction*. London: Chapman and Hall.

Davis, A.J., Jenkinson, L.S., Lawton, J.H., Shorrocks, B. and Wood, S. (1998). Making mistakes when predicting shifts in species range in response to global warming. *Nature* **391**, 783–786.

Diggle, P.J. (1990). *Time Series: A Biostatistical Introduction*. Oxford: Clarendon Press.

Dransfield, R.D., Brightwell, R., Kiilu, J., Chaudhury, M.F. and Adabie, D.A. (1989). Size and mortality rates of *Glossina pallidipes* in the semi-arid zone of southwestern Kenya. *Medical and Veterinary Entomology* **3**, 83–95.

Dutton, J.E. and Todd, J.L. (1906). *The Distribution and Spread of Sleeping Sickness in the Congo Free State with Suggestions on Prophylaxis*. Liverpool: Liverpool School of Tropical Medicine.

Eastman, J.R. and Fulk, M. (1993). Long sequence time series evaluation using standard principal components. *Photogrammetric Engineering & Remote Sensing* **59**, 991–996.

Eliot, T.S. (1944). *Four Quartets*. London: Faber and Faber.

Farley, J. (1991). *Bilharzia: A History of Imperial Tropical Medicine*. Cambridge: Cambridge University Press.

Ford, J. (1971). *The role of the Trypanosomiases in African Ecology: A Study of the Tsetse Fly Problem*. Oxford: Clarendon Press.

Ford, J. and Katondo, K.M. (1977). *The Distribution of Tsetse Flies in Africa*. Nairobi, OAU: Cook, Hammond & Kell.

Gaschen, H. (1945). Les glossines de l'Afrique Occidentale Francaise. *Acta Tropica* (Supplement) **2**, 1–131.

Giddings, L.E. and Naumann, A. (1976). Remote sensing for the control of tsetse flies. *Technical Memorandum Prepared by Lockheed Electronics Company Inc. for Bioengineering Systems Division, NASA*. Houston, TX: Johnson Space Flight Center.

Glasgow, J.P. (1963). *The Distribution and Abundance of Tsetse*. Oxford: Pergamon Press.

Glasgow, J.P. and Welch, J.R. (1962). Long-term fluctuations in numbers of tsetse fly *G. swynnertoni* Austen. *Bulletin of Entomological Research* **53**, 129–137.

Green, P.E. (1978). *Analyzing Multivariate Data*. Hinsdale, IL: Dryden Press.

Green, R.M. and Hay, S.I. (2000). Mapping of climate variables across tropical Africa and temperate Europe using meteorological satellite sensor data. *Remote Sensing of Environment*, in press.

Hargrove, J.W. (1994). Reproductive rates of tsetse flies in the field in Zimbabawe. *Physiolgical Entomology* **19**, 307–318.

Hargrove, J.W. and Borland, C.H. (1994). Pooled population parameter estimates from mark–recapture data. *Biometrics* **50**, 1129–1141.

Hargrove, J.W. and Packer, M.J. (1993). Nutritional states of male tsetse flies (*Glossina* spp.) (Diptera: Glossinidae) caught in odour-baited traps and artificial refuges: models for feeding and digestion. *Bulletin of Entomological Research* **83**, 29–46.

Hargrove, J.W. and Williams, B.G. (1998). Optimized simulation as an aid to modelling, with an application to the study of a population of tsetse flies, *Glossina morsitans morsitans* (Diptera: Glossinidae). *Bulletin of Entomological Research* **88**, 425–435.

Hay, S.I. and 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 & International Health* **4**, 58–71.

Hay, S.I., Tucker, C.J., Rogers, D.J. and 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., Myers, M.F., Burke, D.S. *et al.* (2000). Etiology of mosquito-borne disease epidemics. *Proceedings of the National Academy of Sciences of the USA*, in press.

Hendrickx, G. (1999). *Georeferenced Decision Support Methodology Towards Trypanosomosis Management in West Africa*. Merelbeke: Universiteit Ghent.

Hendrickx, G., Napala, A., Dao, B. *et al.* (1999a). The area-wide epidemiology of bovine trypanosomosis and its impact on mixed farming in subhumid West Africa: a case study in Togo. *Veterinary Parasitology* **84**, 13–31.

Hendrickx, G., Napala, A., Dao, B. *et al.* (1999b). A systematic approach to area-wide tsetse distribution and abundance maps. *Bulletin of Entomological Research*, in press.

Hendrickx, G., Napala, A., Slingenbergh, J.H.W., De Deken, R. and Rogers, D.J. (2000a). The contribution of remote sensing to area-wide tsetse surveys. *Bulletin of Entomological Research*, in press.

Hendrickx, G., Napala, A., Slingenbergh, J.H.W., Dr Deken, R., Vercruysse, J. and Rogers, D.J. (2000b). The spatial patterns of trypanosomosis predicted with the aid of satellite imagery. *Parasitology* **120**, 121–134.

Hide, G. (1997). The molecular epidemiology of trypanosomatids. In: *Trypanosomiasis and Leishmaniasis: Biology and Control* (G. Hide, J.C. Mottram, G.H. Coombs and P.H. Holmes, eds), pp. 289–303. Wallingford: CAB International.

Hiernaux, P.H.Y. and Justice, C.O. (1986). Suivie du développement végetal au cours de l'été 1984 dans le Sahel Malien. *International Journal of Remote Sensing* **7**, 1515–1531.

Hoare, C.A. (1972). *The Trypanosomes of Mammals: A Zoological Monograph*. Oxford: Blackwell Scientific.

Jackson, C.H.N. (1930). Contributions to the bionomics of *Glossina morsitans*. *Bulletin of Entomological Research* **21**, 491–527.

Jackson, C.H.N. (1937). Some new methods in the study of *Glossina morsitans*. *Proceedings of the Zoological Society of London 1936*, 811–895.

Kitron, U. (1998). Landscape ecology and epidemiology of vector-borne diseases: tools for spatial analysis. *Journal of Medical Entomology* **35**, 435–445.

Kitron, U., Otieno, L.H., Hungerford, L.L. *et al.* (1996). Spatial analysis of the distribution of tsetse flies in the Lambwe Valley, Kenya, using Landsat TM satellite imagery and GIS. *Journal of Animal Ecology* **65**, 371–380.

Kleine, (1909). Weitere wissenschafteliche beobachtungen über die Entwicklung von Trypanosomen in Glossinen. *Deutsche Medizinische Wochenschrift*, 469–470.

Komba, E.K., Kibona, S.N., Ambwene, A.K., Stevens, J.R. and Gibson, W.C. (1997). Genetic diversity among *Trypanosoma brucei rhodesiense* isolates from Tanzania. *Parasitology* **115**, 571–579.

Krzanowski, W.J. and Marriott, F.H.C. (1995). *Multivariate Analysis Part 2. Classification, Covariance Structures and Repeated Measurements*. London: Arnold.

Langley, P.A., Maudlin, I. and Leedham, M.P. (1984). Genetic and behavioural differences between *Glossina pallidipes* from Uganda and Zimbabwe. *Entomologia Experimentalis et Applicata* **35**, 55–60.

La Rocque, S. de (1997). *Identification des Facteurs Discriminants Majeurs de la Présence des Glossines dans une Zone Agro-pastorale du Burkina Faso. Intéret pour la Prévision du Risque Trypanosomien*. Montpellier: Université Montpellier II.

Ma, Z.K. and Redmond, R.L. (1995). TAU coefficients for accuracy assessment of classification of remote sensing data. *Photogrammetric Engineering & Remote Sensing* **61**, 435–439.

Macdonald, G. (1957). *The Epidemiology and Control of Malaria*. London: Oxford University Press.

Manel, S., Dias, J.M., Buckton, S.T. and Ormerod, S.J. (1999). Alternative methods for predicting species distributions: an illustration with Himalayan river birds. *Journal of Applied Ecology* **36**, 734–747.

McLynn, F. (1991). *Stanley. The Making of an African Explorer*. Oxford: Oxford University Press.

Milligan, P.J.M. and Baker, R.D. (1988). A model of tsetse-transmitted animal trypanosomiasis. *Parasitology* **96**, 211–239.

Morris, K.R.S. (1951). The ecology of epidemic sleeping sickness. I. The significance of location. *Bulletin of Entomological Research* **42**, 427–443.

Morris, K.R.S. (1952). The ecology of epidemic sleeping sickness. II. The effects of an epidemic. *Bulletin of Entomological Research* **43**, 375–396.

Morris, K.R.S. (1959). The epidemiology of sleeping sickness in East Africa. Part I. A sleeping sickness outbreak in Uganda in 1957. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **53**, 384–393.

Morris, K.R.S. (1960a). Studies on the epidemiology of sleeping sickness in Africa. II. Sleeping sickness in Kenya. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **54**, 71–86.

Morris, K.R.S. (1960b). Studies on the epidemiology of sleeping sickness in East Africa. III. The endemic area of Lakes Edward and George in Uganda. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **54**, 212–224.

Morris, K.R.S. (1963). The movement of sleeping sickness across central Africa. *Journal of Tropical Medicine and Hygiene* **66**, 59–76.

Morris, K.R.S. and Morris, M.G. (1949). The use of traps against tsetse in West Africa. *Bulletin of Entomological Research* **39**, 491–528.

Murray, M. and Gray, A.R. (1984). The current situation on animal trypanosomiasis in Africa. *Preventive Veterinary Medicine* **2**, 23–30.

Mwambu, P.M. and Odhiambo, J.O. (1967). Cattle trypanosomiasis in the area adjoining the South Busoga fly-belt. In: *East African Trypanosomiasis Research Organization Annual Report 1966*, pp. 56–59. Tororo: EATRO.

Nash, T.A.M. (1944). A low density of tsetse associated with a high incidence of sleeping sickness. *Bulletin of Entomological Research* **35**, 51–59.

Nash, T.A.M. (1948). *Tsetse Flies in British West Africa*. London: HMSO.

Newstead, R., Evans, A.M. and Potts, W.H. (1924). *Guide to the Study of Tsetse Flies*. Liverpool: Liverpool School of Tropical Medicine.

Onyango, R.J., van Hoeve, K. and de Raadt, P. (1966). The epidemiology of *T.*

rhodesiense sleeping sickness in Alego Location, Central Nyanza, Kenya. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **60**, 175–182.

Ormerod, W.E. (1976). Ecological effect of control of African trypanosomiasis. *Science* **191**, 815–821.

Ormerod, W.E. (1986). A critical study of the policy of tsetse eradication. *Land Use Policy* **3**, 85–99.

Ormerod, W.E. (1990). Africa with and without tsetse. *Insect Science and its Application* **11**, 455–461.

Ormerod, W.E. and Rickman, L.R. (1988). Sleeping sickness control–how wildlife and man could benefit. *Oryx* **22**, 36–40.

Pender, J. and Rosenberg, L.J. (1995). *Impact of Tsetse Control on Land Use in the Semi-arid Zone of Zimbabwe. Phase 1: Classification of Land Use by Remote Sensing Imagery*. Chatham: Natural Resources Institute.

Pender, J., Mills, A.P. and Rosenberg, J. (1997). *Impact of Tsetse Control on Land Use in the Semi-arid Zone of Zimbabwe. Phase 2: Analysis of Land Use Change by Remote Sensing Imagery*. Chatham: Natural Resources Institute.

Prihodko, L. and Goward, S.N. (1997). Estimation of air temperature from remotely sensed surface observations. *Remote Sensing of the Environment* **60**, 335–346.

Prince, S.D., Goetz, S.J., Dubayah, R.O., Czajkowski, K.P. and Thawley, M. (1998). Inference of surface and air temperature, atmospheric precipitable water and vapour pressure deficit using advanced very high-resolution radiometer satellite observations: comparison with field observations. *Journal of Hydrology* **212–213**, 230–249.

Randolph, S.E. and Rogers, D.J. (1997). A generic population model for the African tick *Rhipicephalus appendiculatus*. *Parasitology* **115**, 265–279.

Rawlings, P., Dwinger, R.H. and Snow, W.F. (1991). An analysis of survey measurements of tsetse challenge to trypanotolerant cattle in relation to aspects of analytical models of transmission. *Parasitology* **102**, 371–377.

Robertson, D.H.H. (1963). Human trypanosomiasis in South-East Uganda: a further study of the epidemiology of the disease among fishermen and peasant cultivators. *Bulletin of the World Health Organization* **28**, 627–643.

Robinson, T., Rogers, D.J. and Williams, B. (1997a). Mapping tsetse habitat suitability in the common fly belt of Southern Africa using multivariate analysis of climate and remotely sensed vegetation data. *Medical and Veterinary Entomology* **11**, 235–245.

Robinson, T., Rogers, D.J. and Williams, B. (1997b). Univariate analysis of tsetse habitat in the common fly belt of Southern Africa using climate and remotely sensed vegetation data. *Medical and Veterinary Entomology* **11**, 223–234.

Rogers, D. (1974). *Ecology of Glossina: Natural Regulation and Movement of Tsetse Fly Populations*. Paris: Institut d'Élevage et de Médecine Vétérinaire des Pays Tropicaux.

Rogers, D. (1979). Tsetse population dynamics and distribution: a new analytical approach. *Journal of Animal Ecology* **48**, 825–849.

Rogers, D.J. (1988). A general model for the African trypanosomiases. *Parasitology* **97**, 193–212.

Rogers, D.J. (1990). A general model for tsetse populations. *Insect Science and its Application* **11**, 331–346.

Rogers, D.J. (1993). Remote sensing and the changing distribution of tsetse flies in Africa. In: *Insects in a Changing Environment* (R. Harrington and N.E. Stork, eds), pp. 167–183. London: Royal Entomological Society.

Rogers, D.J. and Packer, M.J. (1993). Vector-borne diseases, models, and global change. *Lancet* **342**, 1282–1284.

Rogers, D.J. and Randolph, S.E. (1984). A review of the density-dependent processes in tsetse populations. *Insect Science and its Application* **5**, 397–402.

Rogers, D.J. and Randolph, S.E. (1986). Distribution and abundance of tsetse flies (*Glossina* spp.). *Journal of Animal Ecology* **55**, 1007–1025.

Rogers, D.J. and Randolph, S.E. (1991). Mortality rates and population density of tsetse flies correlated with satellite imagery. *Nature* **351**, 739–741.

Rogers, D.J. and Randolph, S.E. (1993). Distribution of tsetse and ticks in Africa: past, present and future. *Parasitology Today* **9**, 266–271.

Rogers, D.J. and Williams, B.G. (1993). Monitoring trypanosomiasis in space and time. *Parasitology* **106**, S77-S92.

Rogers, D.J. and Williams, B.G. (1994). Tsetse distribution in Africa: seeing the wood and the trees. In: *Large-scale Ecology and Conservation Biology* (P.J. Edwards, R.M. May and N.R. Webb, eds), pp. 249–273. Oxford: Blackwell Scientific.

Rogers, D.J., Randolph, S.E. and Kuzoe, F.A.S. (1984). Local variation in the population dynamics of *Glossina palpalis palpalis* (Robineau-Desvoidy) (Diptera: Glossinidae). I. Natural population regulation. *Bulletin of Entomological Research* **74**, 403–423.

Rogers, D.J., Hendrickx, G. and Slingenbergh, J. (1994). Tsetse flies and their control. *Revue Scientifique et Technique de l'Office International des Epizooties* **13**, 1075–1124.

Rogers, D.J., Hay, S.I. and 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., Hay, S.I., Packer, M.J. and Wint, G.R.W. (1997). Mapping land-cover over large areas using multispectral data derived from the NOAA-AVHRR: a case study of Nigeria. *International Journal of Remote Sensing* **18**, 3297–3303.

Ross, R. (1909). *Report on the Prevention of Malaria in Mauritius*. London: Churchill.

Ross, R. (1911). *The Prevention of Malaria*. London: Murray.

Scott, D. (1965). *Epidemic Disease in Ghana 1901–1960*. London: Oxford University Press.

Stollnitz, E.J., Derose, T.D. and Salesin, D.H. (1996). *Wavelets for Computer Graphics*. San Francisco: Morgan Kaufmann.

Storch, H. von and Zwiers, F.W. (1999). *Statistical Analysis in Climate Research*. Cambridge: Cambridge University Press.

Swain, P.H. (1978). Remote sensing: the quantitative approach. In: *Remote Sensing: the Quantitative Approach* (P.H. Swain and S.M. Davis, eds), pp. 136–187 and 221–223. New York: McGraw-Hill.

Swallow, B.M. (1997). Impacts of trypanosomosis on African agriculture. In: 24th Meeting of the International Scientific Council for Trypanosomiasis Research and Control, Maputo, p. 607. Nairobi: Organisation of African Unity.

Tatsuoka, M.M. (1971). *Multivariate Analysis: Techniques for Educational and Psychological Research*. New York: Wiley & Sons.

Thomson, M.C., Connor, S.J., D'Alessandro, U., Rowlinson, B., Diggle, P., Cresswell, M. and Greenwood, B. (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.

Turner, M.J., Blackledge, J.M. and Andrews, P.R. (1998). *Fractal Geometry in Digital Imaging*. San Diego: Academic Press.

Vickerman, K. (1997). Landmarks in trypanosome research. In: *Trypanosomiasis and Leishmaniasis: Biology and Control*, (G. Hide, J.C. Mottram, G.H. Coombs and P.H. Holmes, eds), pp. 1–37. Willingford: CAB International.

Williams, B., Rogers, D., Staton, G., Ripley, B. and Booth, T. (1992). Statistical modelling of georeferenced data: mapping tsetse distributions in Zimbabwe using climate and vegetation data. In: *Modelling Vector Borne and Other Parasitic Diseases* (B.D. Perry and J.W. Hansen, eds), pp. 267–280. Nairobi: International Laboratory for Research on Animal Diseases.