

Monitoring trypanosomiasis in space and time

D. J. ROGERS and B. G. WILLIAMS*

Department of Zoology, South Parks Road, Oxford OX1 3PS

SUMMARY

The paper examines the possible contributions to be made by Geographic Information Systems (GIS) to studies on human and animal trypanosomiasis in Africa. The epidemiological characteristics of trypanosomiasis are reviewed in the light of the formula for the basic reproductive rate or number of vector-borne diseases. The paper then describes how important biological characteristics of the vectors of trypanosomiasis in West Africa may be monitored using data from the NOAA series of meteorological satellites. This will lead to an understanding of the spatial distribution of both vectors and disease. An alternative, statistical approach to understanding the spatial distribution of tsetse, based on linear discriminant analysis, is illustrated with the example of *Glossina morsitans* in Zimbabwe, Kenya and Tanzania. In the case of Zimbabwe, a single climatic variable, the maximum of the mean monthly temperature, correctly predicts the pre-rinderpest distribution of tsetse over 82% of the country; additional climatic and vegetation variables do not improve considerably on this figure. In the cases of Kenya and Tanzania, however, another variable, the maximum of the mean monthly Normalized Difference Vegetation Index, is the single most important variable, giving correct predictions over 69% of the area; the other climatic and vegetation variables improve this to 82% overall. Such statistical analyses can guide field work towards the correct biological interpretation of the distributional limits of vectors and may also be used to make predictions about the impact of global change on vector ranges. Examples are given of the areas of Zimbabwe which would become climatically suitable for tsetse given mean temperature increases of 1, 2 and 3 °Centigrade. Five possible causes for sleeping sickness outbreaks are given, illustrated by the analysis of field data or from the output of mathematical models. One cause is abiotic (variation in rainfall), three are biotic (variation in vectorial potential, host immunity, or parasite virulence) and one is historical (the impact of explorers, colonizers and dictators). The implications for disease monitoring, in order to anticipate sleeping sickness outbreaks, are briefly discussed. It is concluded that present data are inadequate to distinguish between these hypotheses. The idea that sleeping sickness outbreaks are periodic (i.e. cyclical) is only barely supported by hard data. Hence it is even difficult to conclude whether the major cause of sleeping sickness outbreaks is biotic (which, in model situations, tends to produce cyclical epidemics) or abiotic. The conclusions emphasize that until we understand more about the variation in space and time of tsetse and trypanosomiasis distribution and abundance we shall not be in a position to benefit from the advances made by GIS. The potential is there, however, to re-introduce the spatial and temporal elements into epidemiological studies that are currently often neglected.

Key words: tsetse, trypanosomiasis, GIS, NDVI, discriminant analysis.

INTRODUCTION

In 1865 Sir John Kirk reported that the Makalolo people of the Victoria Falls region of Africa understood the risks to their cattle posed by the presence of the tsetse fly. They had learnt to avoid certain areas where the flies were abundant, or to pass through them only at night, when flies are inactive, if the areas were 'sufficiently narrow to allow the cattle being driven through before sunrise' (Kirk, 1865 reported in Austen, 1903). On a much larger spatial and historical scale, the patterns of transhumance by the nomadic pastoralists of West Africa were, and to a large extent still are, determined by the presence of tsetse flies in the more humid, southerly regions of this part of the continent. Fly populations fall in the dry season, thus presenting a lower risk of trypanosomiasis which allows the cattle to take advantage of dry season grazing in the south. Almost one hundred years of Western science

applied to the problem of African trypanosomiasis resulted in the production of graphs such as Fig. 1, which relates trypanosomiasis risk to tsetse challenge. Unfortunately much information on the spatial and temporal variation in trypanosomiasis has been lost in our search for patterns of this sort. Such variation appears as noise in the statistical relationship of Fig. 1 and is often conveniently ignored. The increasing availability of Geographical Information Systems (GIS) presents an opportunity to re-introduce into our epidemiological studies the spatial and temporal elements that are currently neglected. This paper explores the possible increases in understanding of vector biology and disease dynamics that could result from using this new technology.

VARIATION IN SPACE

The importance of vector biology in determining the spatial variation of the diseases transmitted by vectors is illustrated by the formula for their basic rate of reproduction, Ro , defined as:

$$Ro = a^2 mcbe^{-uT}/ur$$

* Present address: London School of Hygiene and Tropical Medicine, Keppel street, London WC1E 7HT.

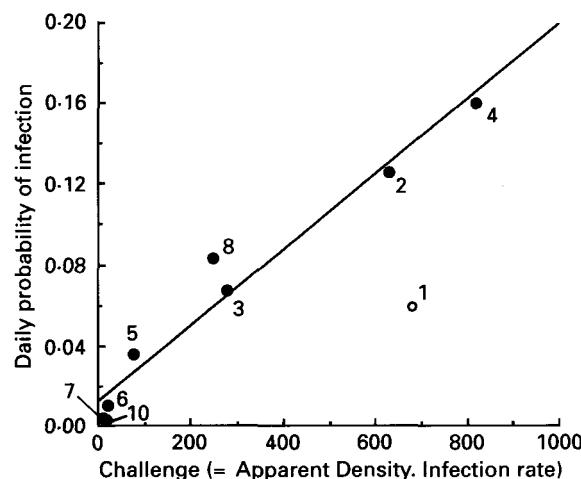


Fig. 1. Relationship between the daily probability of trypanosome infection (all species) in zebu cattle and the challenge presented by the local tsetse populations, the product of tsetse Apparent Density (i.e. the mean numbers of flies caught on a standard 10000 yard transect) and infection rate. 1, Shinyanga, Tanzania; 2 & 3, Lugala, Uganda; 4, Ankole, Uganda; 5 & 6, Kenya; 7, Tanzania; 8, Uganda; 10, Simba, Kenya. The somewhat anomalous point 1 is omitted from the regression calculations. Re-drawn from Rogers, 1985. ($y = 0.012 + 0.000186x$; $r = 0.978$, $P < 0.001$.)

where a = the biting rate of the vector; m = the ratio of vectors to hosts; c = the proportion of infective bites by the vector on uninfected hosts that give rise to infection in those hosts; b = the proportion of infected blood meals that give rise to infections in uninfected vectors; u = the vector mortality rate; T = the incubation period of the parasite within the vector; and r = the rate of recovery of the host from infection (Rogers, 1988a). Many of the components of Ro are therefore vector-related and a study of vector-borne diseases must involve a study of the biology and ecology of the vectors concerned.

African animal trypanosomiasis occurs throughout the range of tsetse flies in Africa, whilst human trypanosomiasis is generally restricted to rather well-defined zones or foci within the distributional range of flies (Apted, 1970; Scott, 1970). One explanation for this difference is that the threshold for transmission of the human disease is much greater than that for the animal disease, mainly because flies are much more reluctant to feed on humans than on other animals, and because humans outnumber other tsetse hosts even in many rural situations (Rogers, 1988b). In the case of vector-borne diseases the threshold (obtained by solving the above equation for $Ro = 1$) depends on the ratio of vectors to hosts. For one of the African animal trypanosomiases, caused by *Duttonella* (previously *Trypanosoma*) *vivax* and assuming realistic values for the variables and parameters involved in the equation, the calculated threshold is around 0.26 flies per host, whilst for the human infective *T. brucei* it is 150 flies per host

(Rogers, 1988a). Human sleeping sickness may therefore be restricted to foci where flies are particularly abundant relative to the human host, or where the degree of man-fly contact is especially high. Very few quantitative data have been gathered on this point, but Morris working in Northern Ghana showed how the fall in the percentage prevalence of human sleeping sickness with distance to the nearest fly-belt was paralleled by a similar fall in the average catch of the tsetse vectors, *Glossina palpalis* and *G. tachinoides*, after clearing of the riverine vegetation in which they live (Morris, 1946, 1952: Fig. 2). He also showed that when sleeping sickness prevalence in a village reaches around 3% there is a net decrease in the human population with time, which often leads to the abandonment of the villages concerned (Morris, 1952). Such a low prevalence has this dramatic effect because the disease, if left untreated, is fatal in most people. Those most affected tend not to be the children, but the younger or middle-aged people, reflecting the fact either that this segment of the population spends much time in the fields and therefore in contact with the flies, or that the disease itself is relatively difficult to catch (or some combination of the two). Age prevalence curves show an increase in infection with age, but only in the 50–60-year age class does trypanosomiasis prevalence usually exceed 10% (Rogers, 1989). In comparison with other vector-borne disease such as malaria, therefore, sleeping sickness is less prevalent but more deadly and the focal nature of the human disease has been appreciated for some time.

Studies in West Africa have linked the biological characteristics of populations of *G. palpalis* with regional variation in disease prevalence through the use of satellite-derived data on the photosynthetic activity of the natural vegetation (Rogers & Randolph, 1991). Tsetse are unusual insects in that the female nourishes each larva within her body through all three larval instars; when deposited in shady places the larvae burrow into the soil and pupariate within a few minutes. Nutritional stress on the parental female translates to under-sized larvae, pupae and thence the next generation of adults. In East Africa, the physical size of *G. pallidipes* (measured as the length of a vein within the wing) was found to be negatively correlated with fly mortality rates estimated from changes in population size or age structure (Dransfield *et al.*, 1989), and the same standard measure of fly size was used in the West African study. The size of *G. palpalis* along a 700 km transect from the coast of Côte d'Ivoire to near Bobo Dioulasso in Burkina Faso was constant in the wet season, but significantly smaller in the north than in the south in the dry season. This regional variation in fly size was related to the Normalized Difference Vegetation Index derived from remotely sensed data of the NOAA series of meteorological

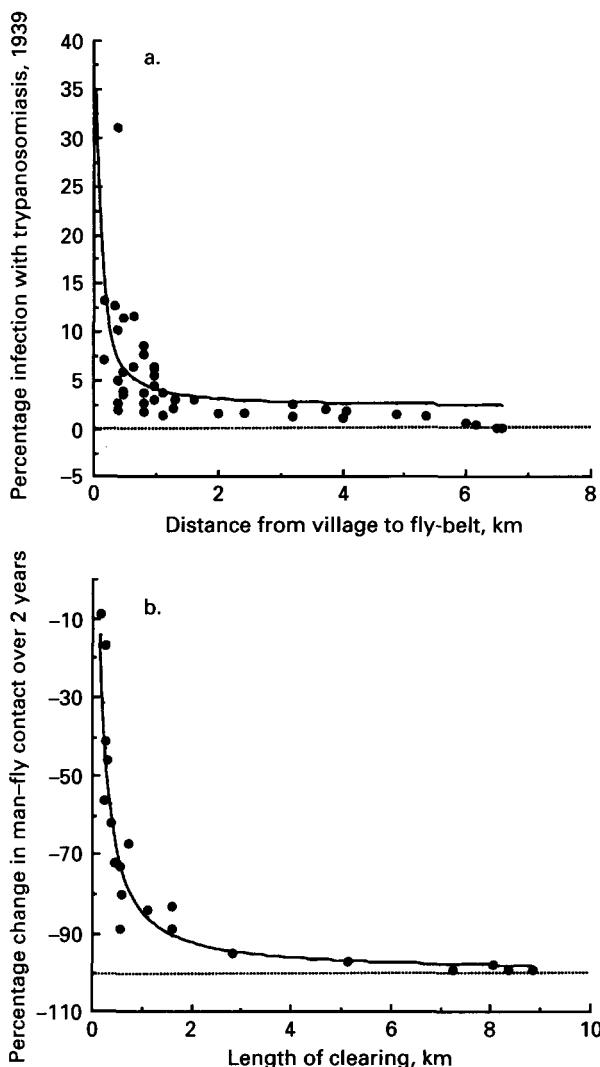


Fig. 2. (a) The relationship between the percentage infection of humans with sleeping sickness in villages in the Lawra District of Ghana and the distance of the village to the nearest permanent fly-belt of *Glossina palpalis* and *G. tachinoides*. ($y = 2.064 + 2.075/x$; $r = 0.523$, $P < 0.001$.) (Re-drawn from Morris, 1952.) (b) The relationship between the percentage change in man-fly contact (as recorded by teams of trained tsetse catchers) and the length of clearing of the riverine evergreen forest in Mamprussi, Wa and Lawra Districts of Ghana. ($y = -100.22 + 15.739/x$; $r = 0.927$, $P < 0.001$.)

satellites (see Justice, 1986 for examples of the use of NDVIs for monitoring the grasslands of Africa). The wet season NDVIs were higher than in the dry season and were associated with greater size of both male and female *G. palpalis* (Fig. 3). The reduction in the size of flies in the dry season in the north indicates considerable population mortality at this time of year. Whether or not this translates into a lower epidemiological risk depends upon the behaviour of the flies, as well as their abundance. Disease risk appears to be greatest in this area of West Africa in the middle of the transect where the average NDVIs are similar to those associated with

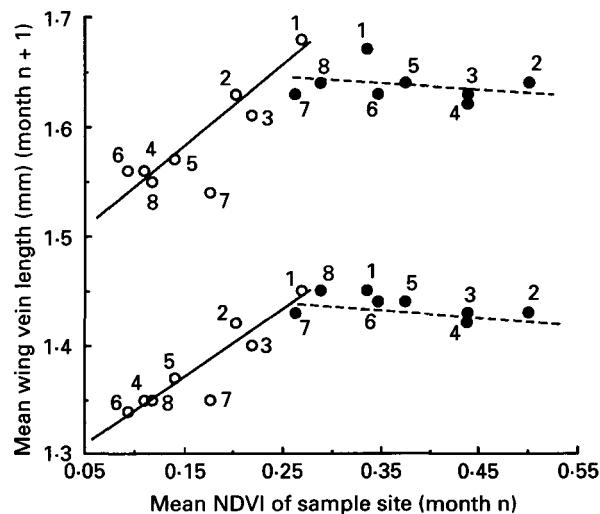


Fig. 3. The relationship between the mean lengths of the hatchet cell in the wings of *G. palpalis* (a standardized measure of fly size, related to the mortality rates of the parental female population) and the mean, satellite-derived Normalized Difference Vegetation Index (NDVI) in the wet (filled circles) and dry (open circles) seasons (1983 and 1984 respectively) of Côte d'Ivoire and Burkina Faso, West Africa. Upper lines, females (dry season, $y = 1.476 + 0.670x$, $r = 0.852$, $P < 0.01$; wet season, $y = 1.651 - 0.037x$, $r = 0.201$, n.s.); lower lines, males (dry season, $y = 1.279 + 0.598x$, $r = 0.915$, $P < 0.01$; wet season, $y = 1.467 - 0.086$, $r = 0.584$, n.s.). The sample sites, ca. 100 km apart, are numbered sequentially, south to north 1, Pauli Brousse, nr Sassandra, on the coast of Côte d'Ivoire; 2, Antonihio; 3, Degbékéré; 4, Bo'Pri; 5, River N'zi; 6, Komborodougou; 7, Oulokoussou; 8, La Guingette, near Bobo Dioulasso, Burkina Faso. Re-drawn from Rogers & Randolph, 1991.

the highest apparent densities of *G. palpalis* recorded in a different study that surveyed the fly populations over the whole of the north of Côte d'Ivoire (see Fig. 4, Rogers & Randolph, 1991). Here, therefore, as in Morris's study, it appears that the number of flies is relatively more important than fly behaviour in determining disease risk.

The studies outlined above should help develop a biological interpretation of disease risk through an understanding of vector population dynamics. Eventually we will obtain sufficient information to explain why a particular vector is in a particular place with a known level of abundance. At the moment, however, we have only a rather incomplete picture of the nature and extent of fly mortality rates, their causes and seasonal changes. Whilst it is possible to make guesses of the distributional range of flies and fly abundance (and therefore disease risk) throughout Africa, these guesses assume that, throughout the range of each species, flies respond in the same way to their abiotic environment (Rogers, 1979) and that the density-dependent processes that ultimately regulate fly population size are also geographically uniform (Rogers & Randolph, 1986). This assump-

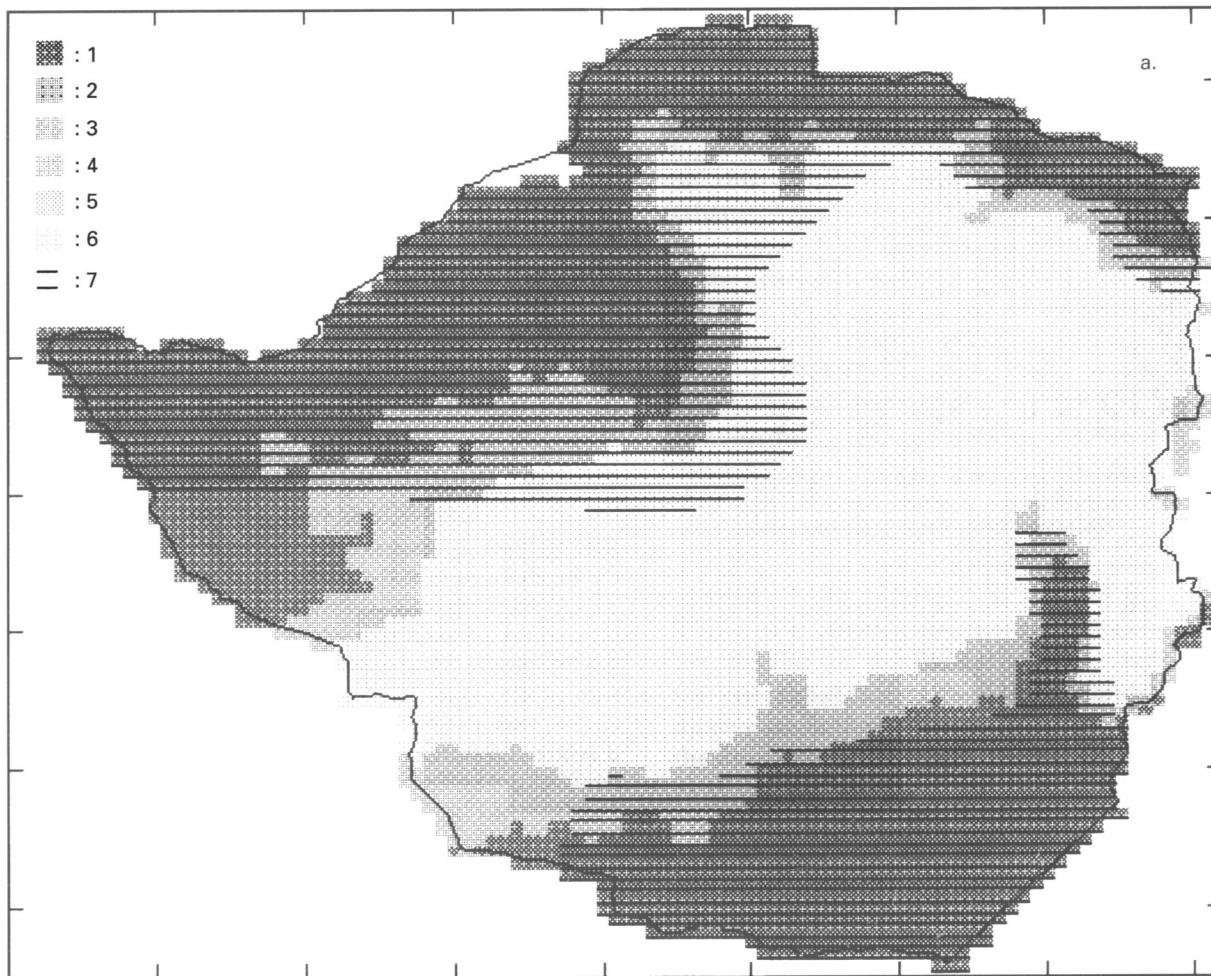


Fig. 4a. The distribution of tsetse flies, *Glossina morsitans*, in Zimbabwe, ca. 1860 (horizontal lines 1) and the output of a linear discriminant analysis applied to environmental and satellite data for Zimbabwe to predict the suitability of each area for flies. The analysis used every 50th data point as a training set to classify all areas on the probability scale 0 (i.e. predicted unsuitable for flies) to 1·0 (very suitable for flies). This probability scale is shown in shades of grey on the map (palest grey 0·0–0·349 [2]; then 0·35–0·449 [3], 0·45–0·499 [4], 0·50–0·549 [5], 0·55–0·649 [6], to darkest grey 0·65–1·0 [7]). (a) using the single predictor variable T_{mm} , the maximum of the monthly mean temperature for the year (82% correct prediction of presence and absence) (b) using all ten predictor variables listed in Table 1 (85% correct predictions).

tion may apply to species of rather restricted distribution, such as *G. austeni* along the coast of East Africa, but it is unlikely to apply to other species such as *G. morsitans* which occurs from Senegal right through the savannah regions of Africa to a latitude of 22° South in southern Mozambique.

An alternative approach to studying the distribution (and abundance) of animals relies upon statistical analysis of a known pattern of distribution of the species concerned. Correlations between the presence or absence of a species and environmental variables can be used to make predictions either about the presence or absence of the species in areas which have been inadequately surveyed in the past, or about changes in distributions that will occur with environmental changes in the future. In general, standard meteorological records are used as the predictor variables, occasionally in combination with soil or vegetation information. The key problem here is to select those variables of apparent importance in

determining the distribution pattern; the other variables can be discarded. A number of multivariate techniques appear to be appropriate for this exercise and we have explored the use of simple linear discriminant analysis (Green, 1978) to understand the distribution of *G. morsitans* in Zimbabwe, Kenya and Tanzania. Preliminary results are shown in Figs 4 and 5.

From the tsetse fly's point of view, in Zimbabwe two catastrophes have drastically restricted fly distribution in the last one hundred years. The first was the great rinderpest pandemic which killed around 85% of domestic stock within about ten years of its accidental introduction into N. Africa in the last decade of the last century. The impact of rinderpest on wild-life was not quantified but appears to have been equally dramatic (Ford, 1971). The disappearance of large numbers of their usual hosts caused recessions of tsetse flies in Zimbabwe (then Rhodesia), and their total disappearance from the

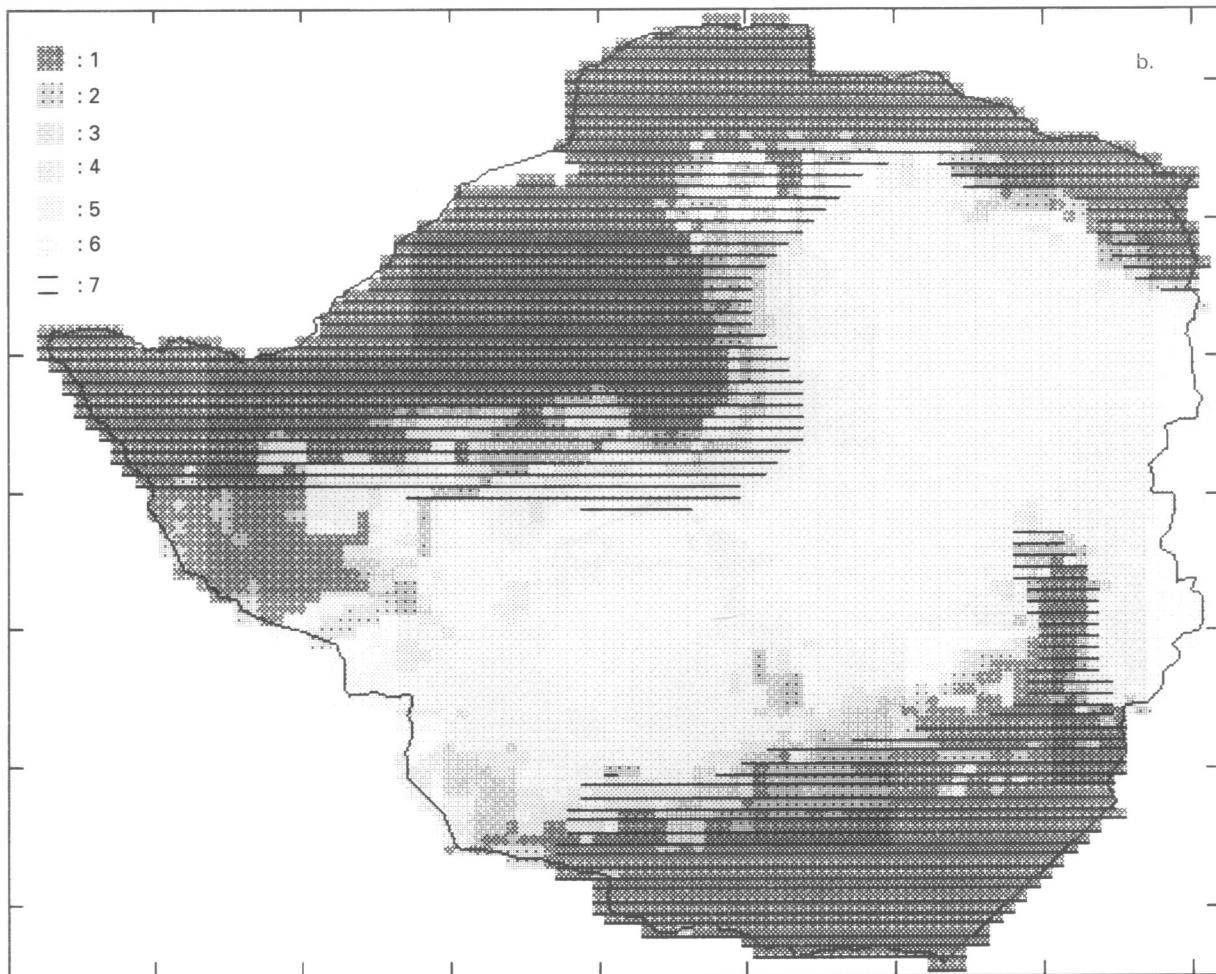


Fig. 4b. For legend see opposite.

southern part of the country. The second catastrophe was the tsetse control activities of the Veterinary Department which eradicated flies from much of the northern part of the country, whilst maintaining the southern part of the country free of re-invasion by the flies. Today flies are found only in the Zambezi Valley east of Lake Kariba. Clearly, therefore, a much larger proportion of Zimbabwe is climatically suitable for tsetse than is suggested by the present-day fly distribution maps for that country, and we take as our base map that given by Chorley (1947) which purports to show the extent of the distribution of *G. morsitans* in Zimbabwe around 1860, before the disturbance caused by the arrival of the Europeans and before rinderpest. Chorley's map is based on the reports of hunters or travellers, and it is not clear how extensive was this information: in places the limits of the fly's distribution appear to follow the contours, suggesting that the map makers interpolated on the basis of the high altitude (low temperature) fly limits known from other regions of the country. Thus before much hunting and before rinderpest in Zimbabwe tsetse flies were found throughout the north and the south of the country, but were absent from the highlands in the middle (Fig. 4).

A list of the variables used in the present analysis is given in Table 1, together with their apparent order of importance as determined by the discriminant analysis. There are several ways in which the importance of any variable might be measured in an analysis of this sort. Here we take the Mahalanobis distance, D^2 , as the best measure of separation (Marriott, 1974; Green, 1978). This is the distance in multi-variate space between the centroids of sites where flies are present and where they are absent, adjusted both for the differences in the variances of the original variables and for their co-variation. The analysis begins with the registration of the fly data to the meteorological data within a data set where geographical locations (longitude and latitude, or map co-ordinates in a particular projection) are stored as two of the variables. A subset of these data, the 'training set' (in general every n -th point, giving a manageable data set for analysis), is then selected and, if flies are present, the data for that point contribute to an estimate of the mean and covariance of sites suitable for fly occurrence and if they are absent the data contribute to an equivalent estimate for unsuitable sites. The predictor variables are initially entered singly and the one giving the greatest mean separation between sites where flies are present

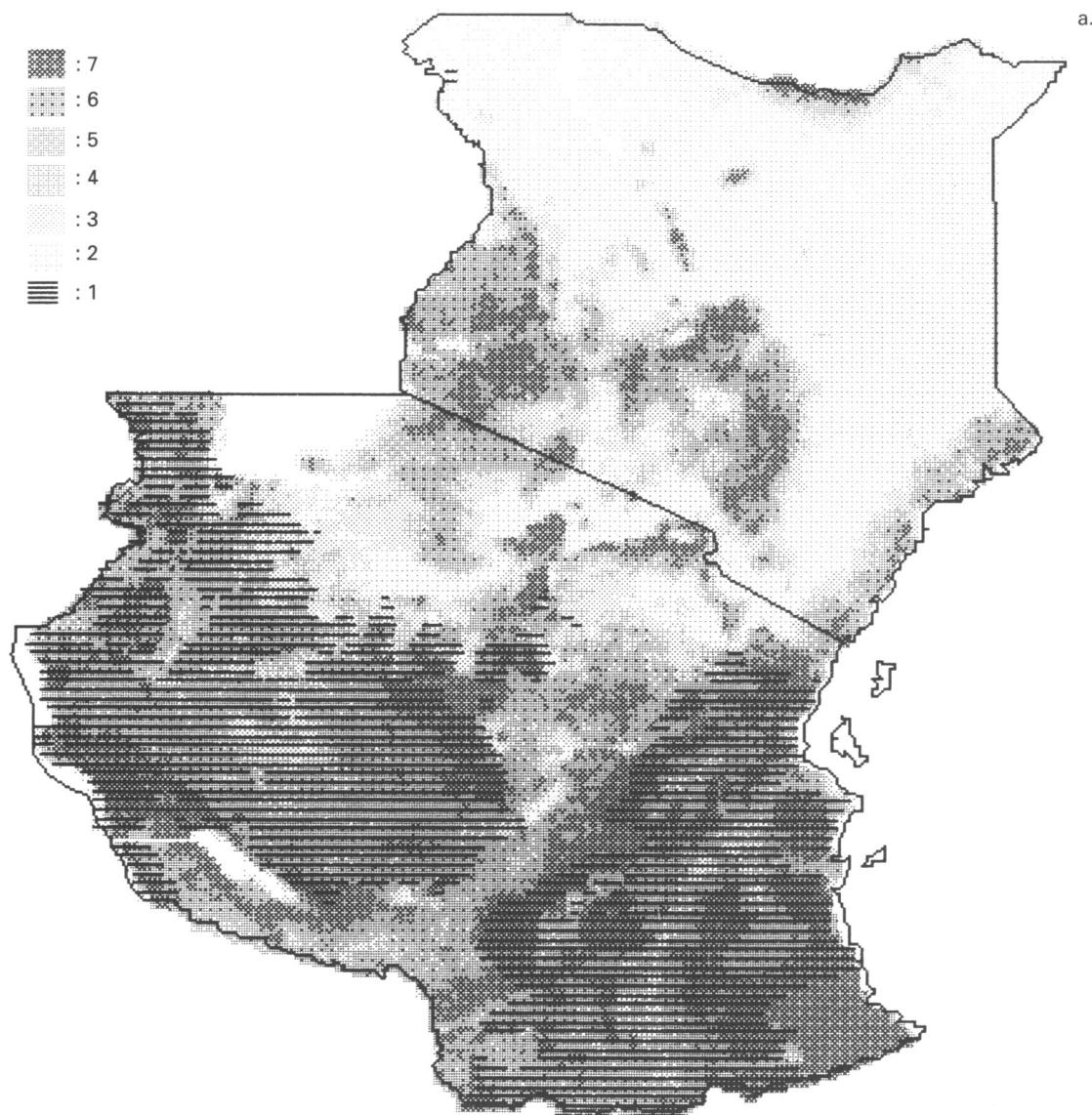


Fig. 5a. As for Fig. 4, for *G. morsitans* in Kenya and Tanzania, using every 150th data point as the training set. (a) the prediction of suitability using the most important single variable ND_{max} , the maximum of the mean monthly values of the Normalized Difference Vegetation Index (69% correct predictions) (b) using all ten variables (84% correct predictions). Probability scale as in Fig. 4. The Ford & Katondo (1977) map of fly distribution is shown by the horizontal lines. For further details see Table 1.

or absent is ranked as the most important. This variable is then in turn used in pair-wise combination with every other remaining variable, and the second most important variable is selected in the same way; and so on. At each stage of the analysis the ability of the included variables to predict the fly distribution accurately is tested by applying the results to the whole of the data set, one location (or pixel) at a time. The predictor variables are used to estimate the probability that each location is suitable for flies, on the basis of the distance of the values of the site's predictor variables from the centroids of the training set and the variance-covariance matrix for all sites combined. Clearly sites that are close to the centroid for fly presence are more likely to be suitable for flies than sites which are far from this centroid, or sites which are close to the centroid for fly absence.

The output of the analysis is a map of the probability of occurrence of the species concerned. The percentage correct classification of sites is also calculated, together with the percentage of false positives (an incorrect prediction of presence) and false negatives (an incorrect prediction of absence). In theory, as more variables are included in the analysis, there should be greater separation of the presence and absence centroids, and a more accurate prediction of the distribution of the species concerned. In practice, relatively few variables make a major contribution to the predicted distributions, and the other variables can be ignored (many environmental variables are correlated with each other and so do not increase the overall performance of the analysis). In the case of Zimbabwe, for example, the single variable Tx_{mm} , the maximum of the monthly mean

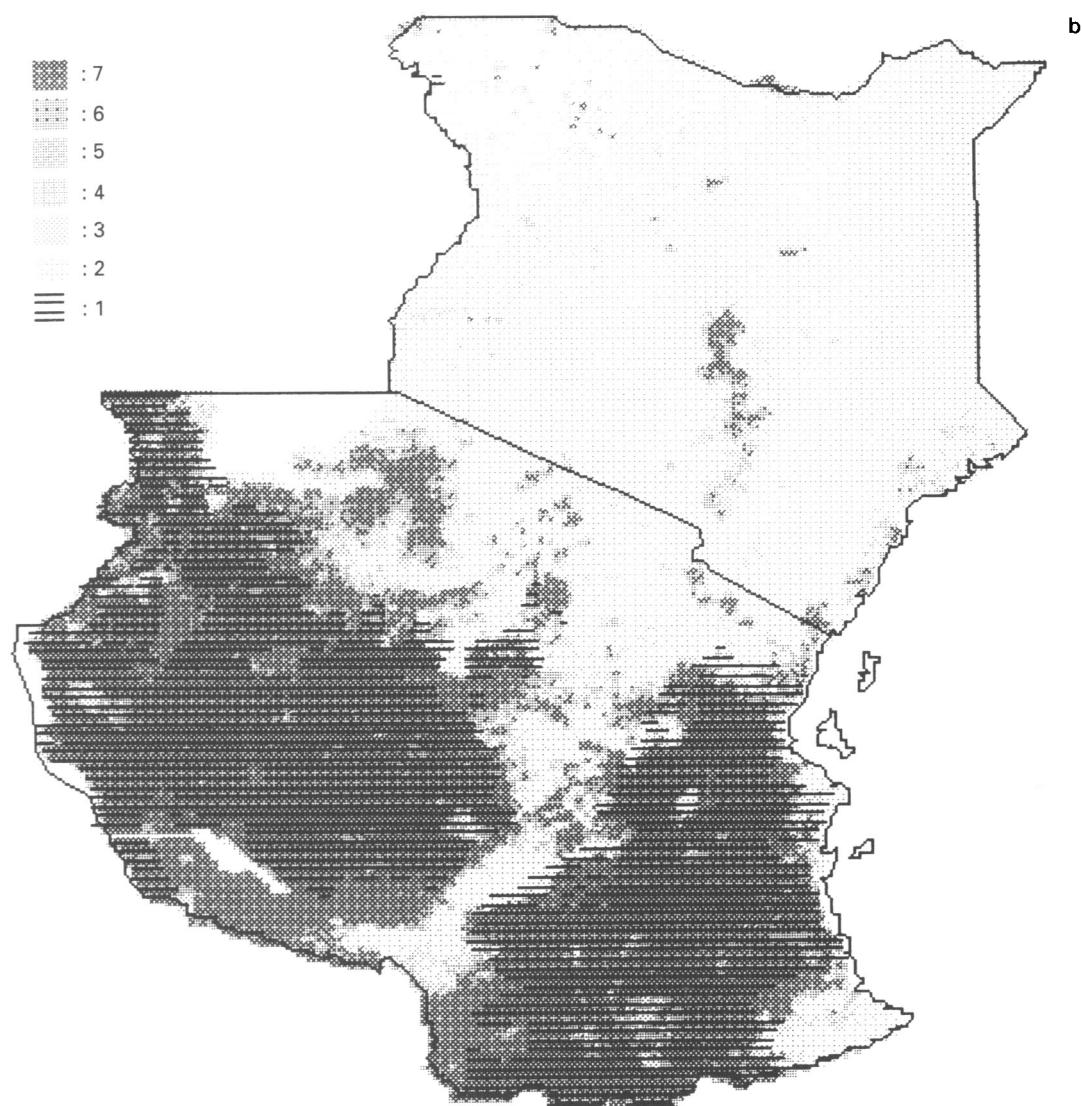


Fig. 5b. For legend see opposite.

temperature, gives an 82% correct prediction of presence and absence (with 11% of false positives and 6% of false negatives) and the other nine variables do not improve substantially on these figures (Fig. 4 and Table 1, correct predictions rising to 85%). Here, then, the analysis has picked out a temperature variable that is stressed in much of the literature on tsetse in Zimbabwe as being important in this part of the continent (e.g. Chorley, 1929). Flies are absent from the highlands of Zimbabwe because it is too cold for them there (the values of T_{xmm} and the minimum mean temperature, T_{nmm} , are both higher in tsetse than non-tsetse areas).

In the case of the same species of fly in Kenya and Tanzania, however, a similar analysis using the same ten variables suggests that a different key variable (in this case the maximum NDVI) is most important (69% correct predictions with 28% false positives and 3% false negatives, Fig. 5a). Also, additional variables make a significant contribution to predicting the distribution of flies (correct predictions rising to 84%, Fig. 5b and Table 1). Thus a single

key variable tends to set the very edge of the distributional range of a species (as in Zimbabwe, which is close to the southern limits of this species in Africa), but within this range the local occurrence of the species appears to be determined by a combination of factors (as in Kenya and Tanzania).

Statistical analysis provides both a challenge to fly ecologists to explain in biological terms what is happening and an opportunity to use a monitoring scheme, based on measuring identified variables, to predict how distributions of vectors (and presumably the diseases they transmit) will change with environmental change. Fig. 6 shows predictions of how the area of Zimbabwe climatically suitable for *G. morsitans* would increase with mean increases in temperature of 1–3 °C. Much of the highlands (which are at present below the low-temperature limits of fly distribution) become suitable for this species. Clearly fly distribution, in the absence of effective vector control services, could begin to change dramatically with an increase in global temperature.

Table 1A. Variables used in the discriminant analysis of the distribution of *G. morsitans* in Zimbabwe, Kenya and Tanzania and their relative contribution (D^2 , Mahanobis distance) to predictive mapping of the species concerned (Figs. 4 & 5). (Values of D^2 are given assuming all variables of higher predictive value are already in the analysis)

Variable	Description	Zimbabwe	D^2	Kenya and Tanzania	D^2
<i>Txmm</i>	Max. of monthly mean temperature	1	3.62	6	2.91
<i>Tnmm</i>	Min. of monthly mean temperature	2	3.93	9	3.88
<i>NDSep/min</i>	NDVI for September (Z) or annual minimum	3	4.10	10	3.86
<i>NDFeb/max</i>	NDVI for February (Z) or annual maximum	4	4.24	1	1.04
<i>Elev.</i>	Elevation, m	5	4.29	4	2.32
<i>Tmmm</i>	Mean of monthly mean temperature	6	4.34	2	1.32
<i>Tnmn</i>	Minimum of monthly min. temperature	7	4.35	3	1.83
<i>NDran</i>	NDVI range (i.e. max-min.)	8	4.35	8	3.75
<i>Txmx</i>	Maximum of monthly max. temperature	9	4.36	5	2.83
<i>NDav</i>	NDVI average	10	4.36	7	2.99

Table 1B. Progressive improvement in the % fit of the discriminant analysis with the variables added in order of importance (see above). % correct (both +ve and -ve) and % false positives (incorrect prediction of presence) and false negatives (incorrect prediction of absence)

Rank	% corr.	Zimbabwe		Kenya and Tanzania	
		False		False	
		+ve	-ve	% corr.	+ve
1	82	11	6	69	28
2	84	6	10	75	20
3	83	8	9	77	17
4	85	6	9	78	17
5	84	6	10	81	17
6	86	6	9	82	16
7	86	6	9	83	15
8	86	6	9	84	13
9	85	6	9	84	14
10	85	6	9	84	14

Thus far the biological and statistical approaches have been applied to an understanding of the distribution and abundance of the tsetse vectors, using a variety of environmental factors as explanatory or predictor variables. The same factors may also be useful in predicting the incidence and prevalence of the trypanosomiases. The south-eastern corner of Uganda has experienced three major epidemics of sleeping sickness (and several minor ones) in the last one hundred years, with the last epidemic, which began in the late 1970s (Mbulamperi, 1989) and still continuing, though apparently on the decline. The recent outbreak was unusual in that it was most serious outside the historical limits of the distribution of the local vector species, *G. fuscipes*, as recorded in the tsetse distribution maps (Ford & Katondo, 1977). The reason for this appears to be that the cotton growing, that used to occur between Lakes Victoria and Kyoga, involved the control of local vegetation

which excluded the flies. Cotton growing was abandoned during the turmoil following General Amin's rule in Uganda, and the fields were invaded first by the shrubby weed *Lantana* and later by *G. fuscipes*. (A similar growth of *Lantana* led to an outbreak of sleeping sickness, involving the same vector species, in the Alego district of Central Nyanza, Kenya in the 1960s; Willett, 1965). The recent epidemic in Uganda appeared first in Mukono district, west of Jinja, and then spread slowly towards the east, eventually reaching Tororo district. Those areas away from the shores of Lake Victoria were more affected than those next to the lake. Overall Kigulu district, between Jinja and Tororo, and Jinja district itself appear to have had the greatest number of sleeping sickness cases. Unfortunately very few studies were made of the vectors during this recent outbreak in Uganda, and the only extensive data available are for the incidence and prevalence of the human disease. One striking relationship to emerge

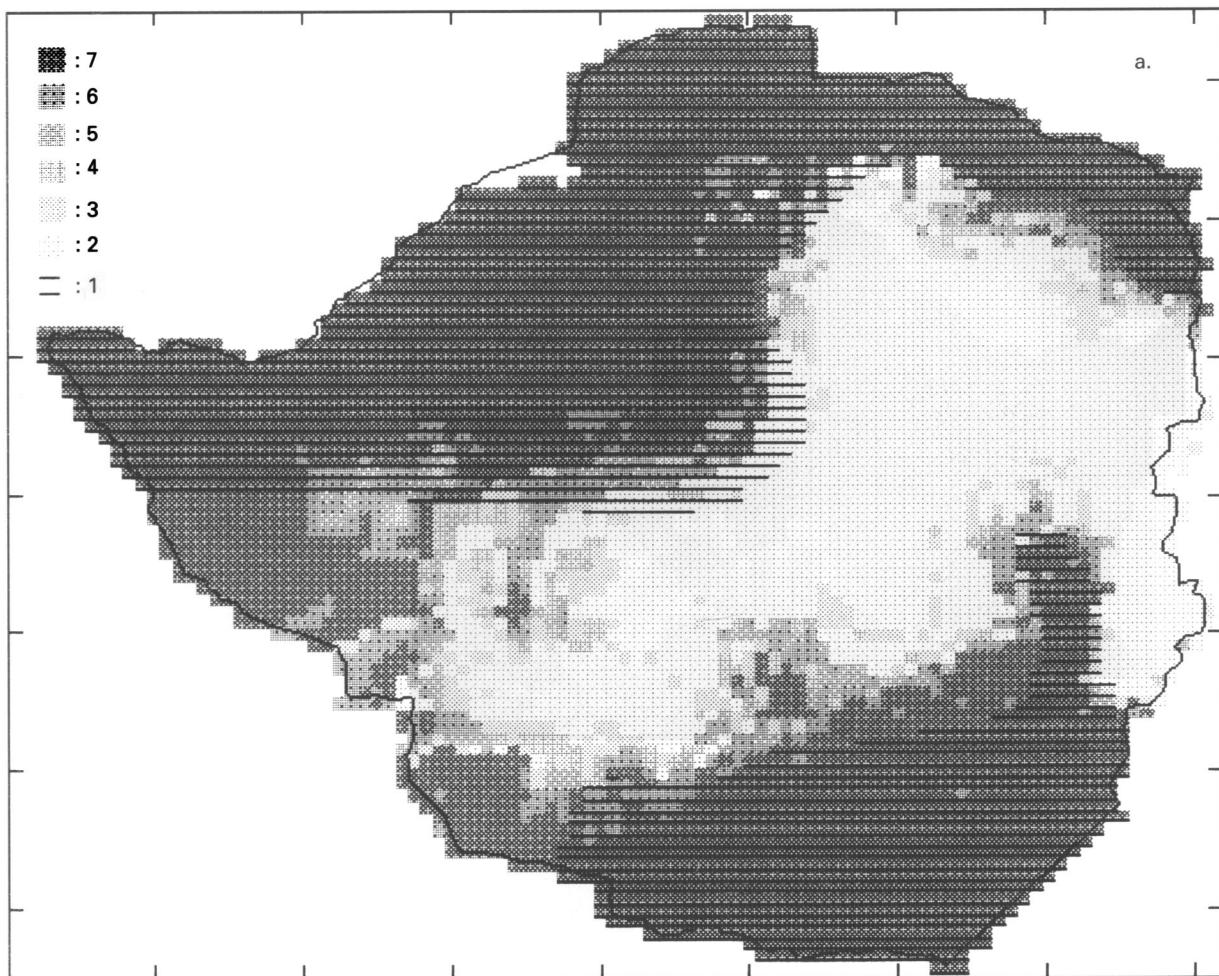


Fig. 6a. The predicted impact of global warming on the distribution of *G. morsitans* in Zimbabwe. The discriminant analysis is used to predict the suitability of different areas of Zimbabwe under scenarios of global warming of (a) 1, (b) 2 and (c) 3 °C. Probability scale as in Fig. 4.

from an analysis of these data is between the mean prevalence of sleeping sickness at the county level and the mean NDVI for each county (Fig. 7). Disease prevalence peaks at intermediate values of NDVI and declines to zero within a very narrow range of NDVI either side of the peak figure. Fig. 7 hides considerable spatial variation in the relationship between disease incidence and NDVI (which will be described elsewhere), but suggests that satellite imagery may be used to monitor spatial variation in disease risk, and may contribute towards the production of 'risk maps' of vector-borne diseases so that scarce health service resources can be targeted at areas of greatest need.

VARIATION IN TIME

On short time scales, the incidence of sleeping sickness varies seasonally. In the Lambwe Valley of Kenya, for example, most cases of human sleeping sickness are recorded in the wet season when the local vector species, *G. pallidipes*, spreads out from the Game Reserve in the Valley itself into the surrounding villages (Wellde *et al.*, 1989). Satellite

imagery is strongly correlated with mean rainfall at a continental scale in Africa (Rogers & Randolph, 1991) and monthly NDVI values for the Lambwe Valley show a significant, positive correlation with sleeping sickness case numbers of the following month (see fig. 12, in Rogers, 1991).

On longer time scales, human sleeping sickness in the drier parts of West Africa (Duggan, 1970) and throughout much of East Africa appears sporadically in epidemic form. In some areas these epidemics appear to be regular, although the inter-epidemic period, of between about 10 and 40 years, is so long that records do not exist for more than a very few 'cycles' in any one place.

There appear to be five major types of explanation for sleeping sickness outbreaks, one abiotic, three biotic and one historical, as follows.

Effect of climate

Fairbairn & Culwick (1950) showed how the apparent population size of *G. swynnertoni* changed with variation in mean soil temperature and rainfall, with lower fly numbers in years of higher rainfall.

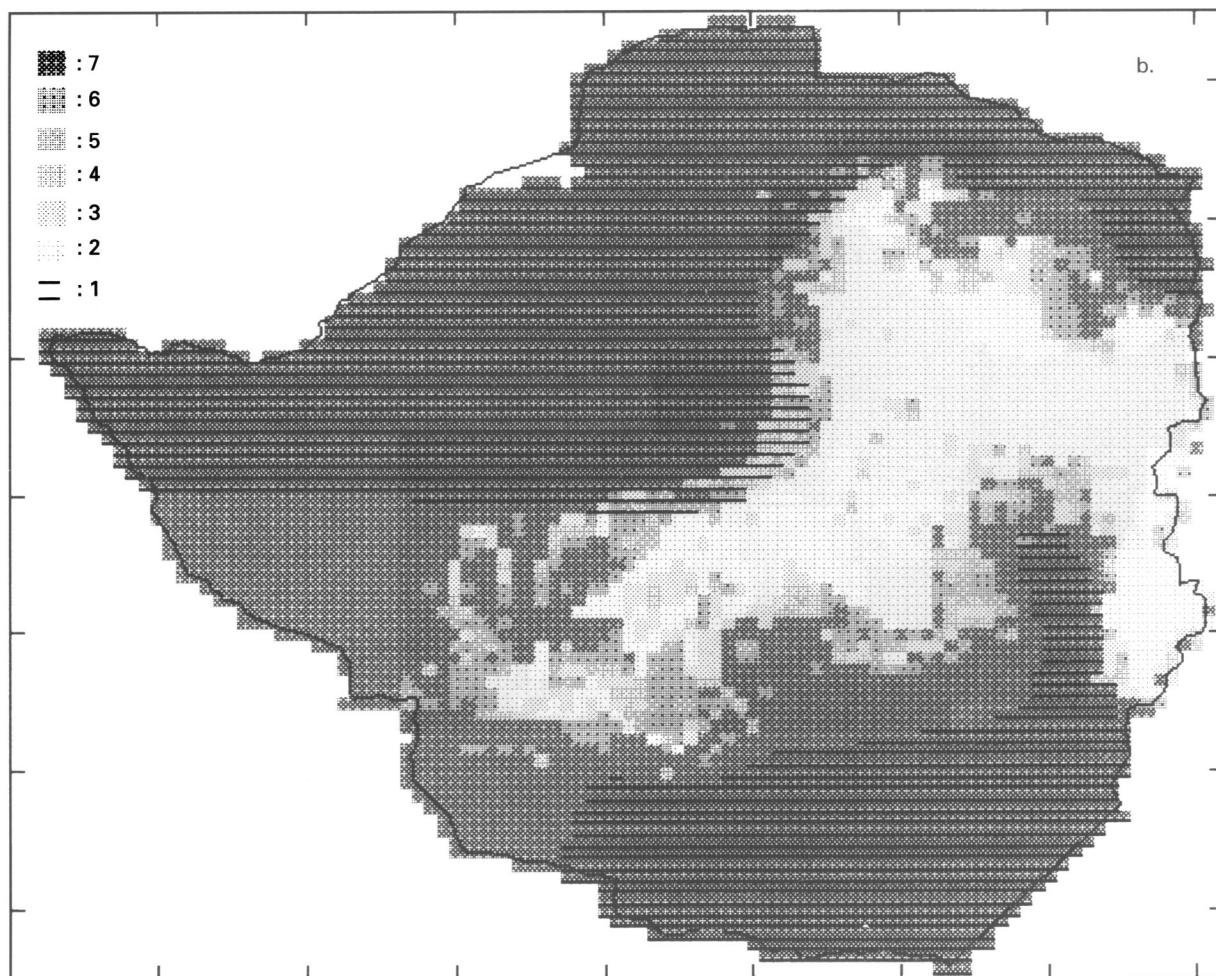


Fig. 6b. For legend see p. 85.

Fairbairn also found a positive relationship between the mean catches of *G. swynnertoni* at Shinyanga, Tanzania, and sleeping sickness case numbers in Tanzania for the period 1930–45 (Fairbairn, 1948). The form of sleeping sickness that Fairbairn reported (caused by *T. b. rhodesiense*) is thought to have arrived in Tanzania from west of Lake Tanganyika some time before 1922. This was followed by an epidemic of the disease throughout the west of the country, with country-wide case numbers peaking at over 3000 in 1929 (Kilama, Mteru & Paul, 1981), and an eventual equilibration to between 300 and 1400 cases a year from around 1940 onwards. Fig. 8 shows the relationship between these case numbers and the mean rainfall for Tabora (from Nicholson, Kim & Hoopingarner, 1988) over the same period of time. Fig. 8 divides the data into 4 groups; an initial period up to 1927 where there is no relationship between case numbers and rainfall (this was probably associated with the spread of the disease after its first introduction); a period to 1934, during the major outbreak, when there is a strong inverse relationship which is, however, not significant because of a shortage of data points (Fig. 8a); a period up to 1960 when there is a weaker but significant inverse relationship (Fig. 8b); and finally the period from

1961 when the relationship appears to break down altogether (Fig. 8c). At least for part of the time, therefore, sleeping sickness case numbers were fewer during periods of greater than average rainfall, associated with lower than average tsetse fly population sizes. The failure of the correlation in the latter part of the series (Fig. 8c) may well be due to human population pressure on the natural vegetation and therefore the vectors. In the period of 1957–78, for example, the human population in Maswa district, Shinyanga region, doubled and in parts of the district it almost quadrupled. Shinyanga, the centre of tsetse research in East Africa between the 1920s and 1950s, was considered by the 1980s ‘to be in the advanced stages of desertification’ (Gamassa, 1986).

Effect of tsetse susceptibility

Recent work has shown that the susceptibility of tsetse to infection with trypanosomes is a function of the activity of mid-gut lectins, which have a trypanocidal effect. Lectins are at low concentrations in newly emerged flies, which are more susceptible than older flies to infection with *T. brucei* and *Nannomonas congolense* (Maudlin, 1991). As lectin

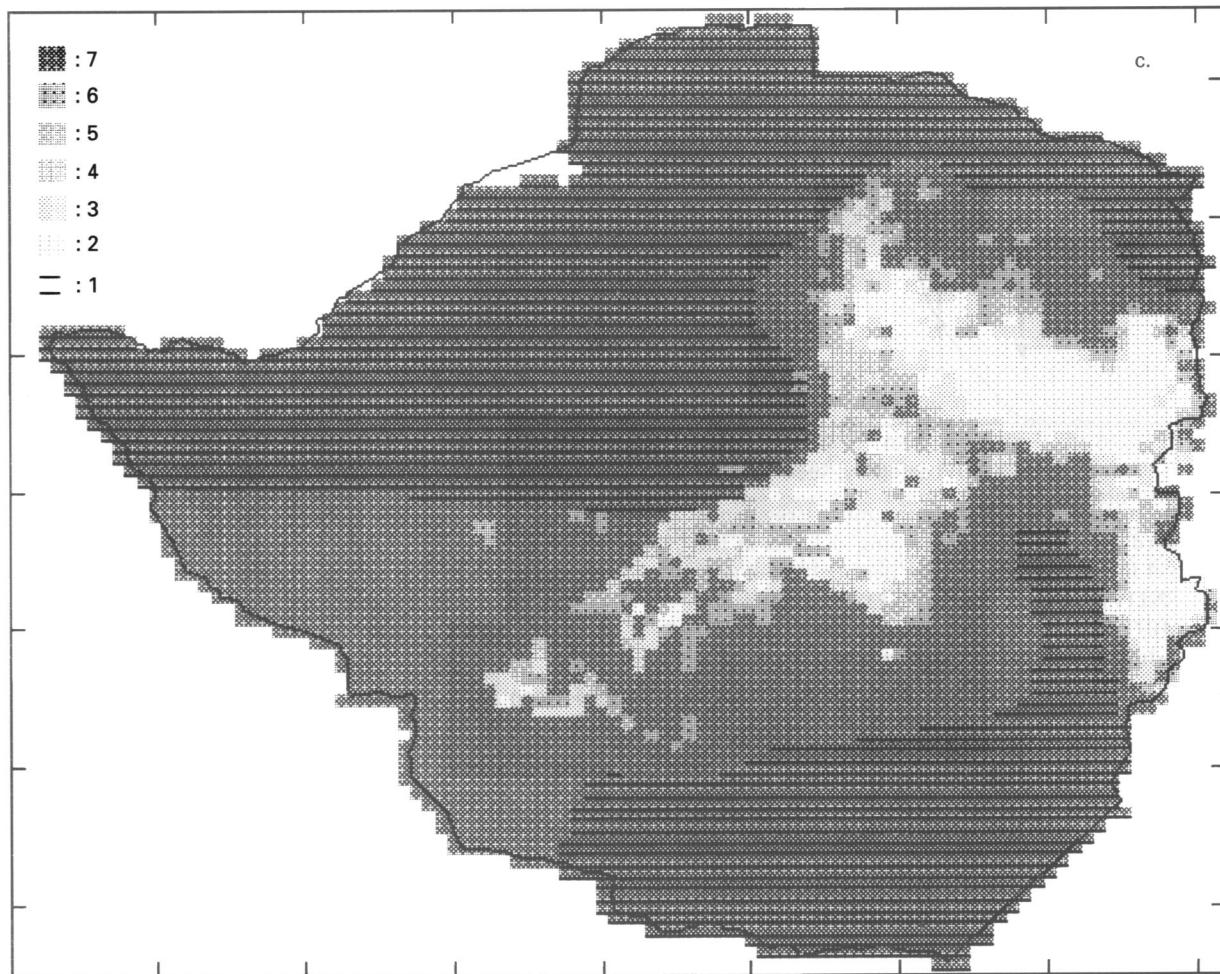


Fig. 6c. For legend see p. 85.

levels increase with age, fly susceptibility to infection decreases. Tsetse often contain rickettsia-like organisms (rlos) in various parts of the gut and other tissues. These rlos cause a reduction in the levels of mid-gut lectins in the flies, and thus render them more susceptible to trypanosome infection. If trypanosome-infected flies suffer a higher natural mortality rate (a somewhat controversial point for which there is little hard evidence), such flies will leave fewer offspring for future generations and thus the rlo prevalence in the fly population will decline to extinction. This will be prevented, however, if the rlos increase the survival prospects of flies in other ways, which it is claimed they do through an increase in puparial survival rates. Thus in the absence of trypanosome infection the rlo prevalence in a fly population would increase to 100 %. There is a net balance of advantage (to rlo+ve, trypanosome-ve flies) and disadvantage (to rlo+ve, trypanosome+ve flies) which can, in a model situation, result in a cycle of rlo prevalence, and therefore trypanosome prevalence, with a periodicity of many years. It has been claimed that sleeping sickness cycles are driven by this mechanism (Baker *et al.* 1990), which need not involve any changes in fly population size or the vector/host ratio.

Effect of host population susceptibility

Host populations which experience infection and develop immunity after recovery from infection support periodic epidemics, rather than stable disease endemism, when the basic rate of reproduction of the disease is low (Anderson & May, 1991). Whilst these conclusions are drawn from models of directly transmitted diseases they may also apply to some vector-borne diseases, including trypanosomiasis. Although there is little evidence for an effective immune status in the human host, wild-life rarely show the high levels of trypanosome infection expected from tsetse blood meal feeding patterns and would appear to be at least partly immune (Mulla & Rickman, 1988). Simple catalytic models of human sleeping sickness incorporating an alternative host population with a fixed level of infection predict epidemics of human sleeping sickness with periods of at least 40 years (determined by human demographic rates which are low in comparison with those of wild-life); the larger the reservoir, the more damped are the sleeping sickness cycles (Rogers, 1988a). A slightly different model, in which the animal hosts pass from a susceptible to infected and thence life-long immune category, generates damped

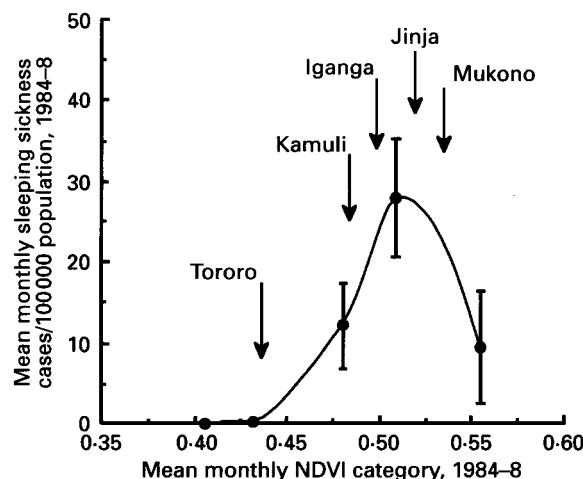


Fig. 7. The relationship between the mean prevalence (± 1 standard error) of human sleeping sickness at the county level in SE Uganda for the period 1984–8 and the mean NDVI for each county (NDVI categories aggregated). The graph also shows the mean district NDVIs for the study period (arrows above the figure, with district names).

cycles of infection with very much shorter periods (again determined by birth and death rates, which are higher in the non-human hosts), and these cycles could spill over into the human host population via the tsetse vectors (Rogers, 1988a). A key feature in both models is that the tsetse population size is related to host population size, so that the ratio of vectors to hosts stays relatively constant as host populations diminish (through death or emigration in the first model) or become immune (in the second). This makes these vector-borne disease models more similar to the models of directly transmitted disease previously referred to in which epidemic behaviour is well understood. The assumption of a constant vector-host ratio is much more likely to be correct in the case of the African trypanosomiases, where the viviparous tsetse obtain all their larval nutrients directly from the host, than in the case of other vector-borne diseases such as malaria, with free-living larval stages. Field studies relating fly to host abundance support this assumption for tsetse (see fig. 4 in Rogers, Randolph & Kuzoe, 1984). Under this hypothesis, therefore, cycles of human-infective *T. brucei* occur in the non-human hosts; if fly numbers change, they do so only in response to changes in host abundance over time.

Effect of trypanosome strain variation

Trypanosomes are extremely complex antigenically, and recombination during sexual reproduction (Jenni *et al.* 1986) could produce new strains of increased virulence, thus possibly starting an epidemic event. Whilst this theory for human sleeping sickness outbreaks is the one favoured by many, there is no field evidence either for or against it. It would appear

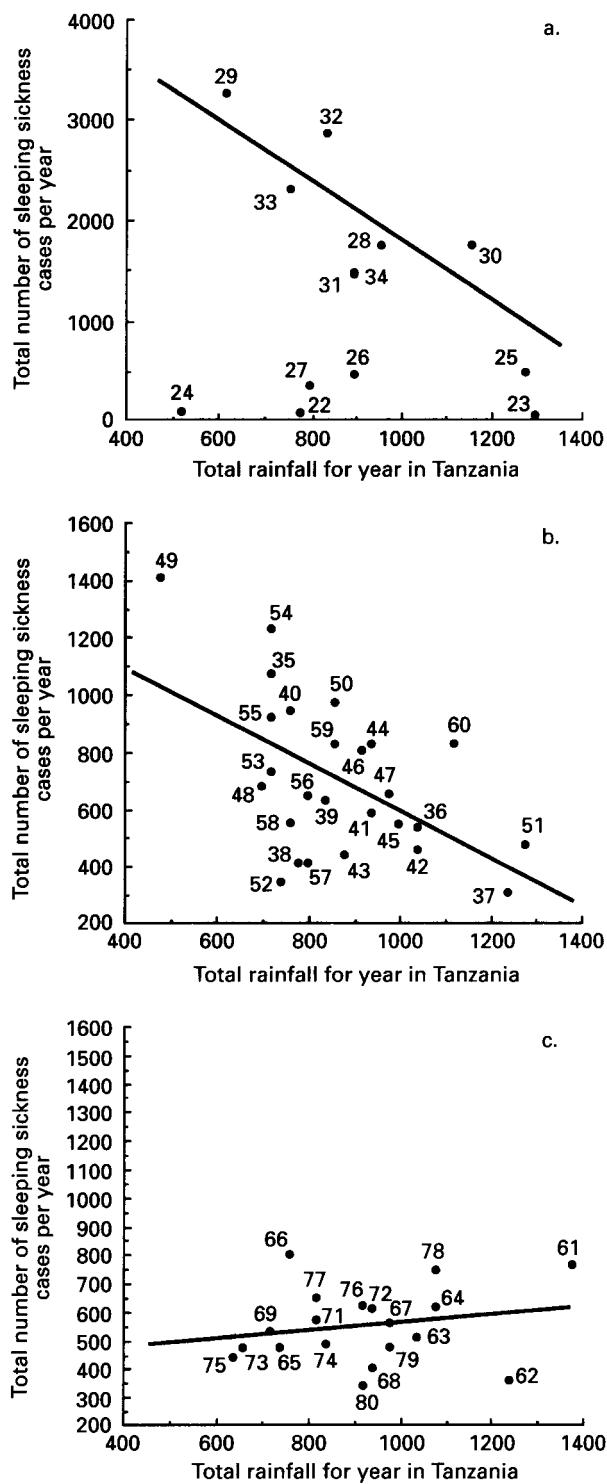


Fig. 8. The relationship between the numbers of cases of human sleeping sickness throughout Tanzania (data from Kilama, Mteru & Paul, 1981) and the annual rainfall recorded at Tabora, Tanzania (data from Nicholson, Kim & Hoopingarner, 1988) for the period (a) 1922–34, following the introduction of the disease in the early 1920s (the regression covers the period 1928–34) ($y = 4758.7 - 3.015x$; $r = 0.709$, d.f. = 5, n.s. (0.754).) (b) 1935–60 ($y = 1424.9 - 0.833x$; $r = 0.539$, d.f. = 24, $P < 0.01$.) and (c) 1961–80 ($y = 424.6 + 0.136x$; $r = 0.198$, d.f. = 18, n.s.)

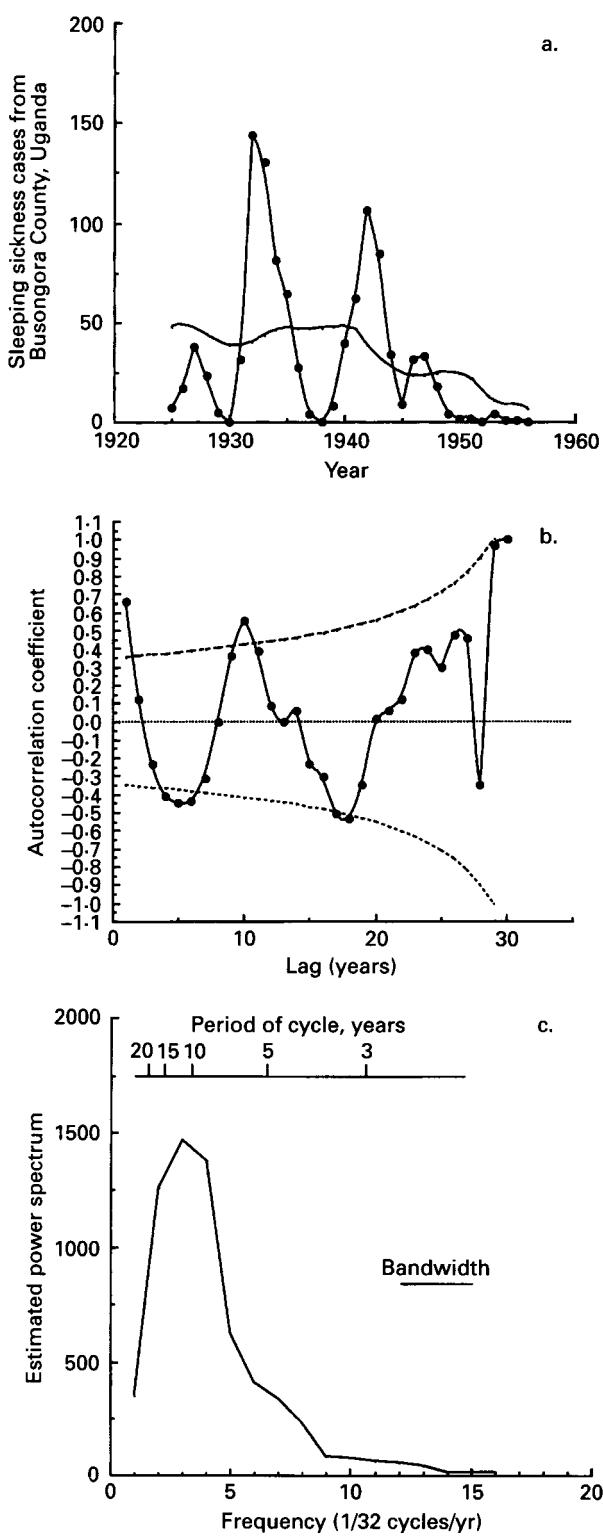


Fig. 9. Spectral analysis of the numbers of cases of sleeping sickness in the Busongora County region of W. Uganda, 1925–56. (a) raw data with 18-point running mean (autocorrelation and spectral analysis were carried out on the deviations from this mean line), (b) autocorrelation of residuals from mean trend line at lags of 1 to 30 years ($\pm 95\%$ confidence intervals of the autocorrelation coefficient). The autocorrelations suggest a mean periodicity of c. 11/12 years. (c) power spectrum (i.e. smoothed periodogram) of the data in (a), showing a peak at a frequency corresponding to a cycle period of 11/12 years (original data from Morris, 1960).

to require that the 'new' strains of trypanosomes are temporarily more transmissible than the old strains from which they arise and that, if there is no long-term change in prevalence, the new strains rather rapidly lose their increased transmissibility/virulence. It is unlikely that these events would occur regularly, to cause periodic outbreaks of sleeping sickness.

Effect of history

The first great epidemic wave of human sleeping sickness occurred in the Busoga region in Uganda at the end of the last century and the beginning of the present one. It seems that this was a new disease from the point of view of the Baganda and Busoga peoples most affected, although some form of sleeping sickness may have been present at very low levels before this time (Ford, 1971). One explanation put forward for the outbreak was that H. M. Stanley's 1889 expedition for the relief of Emin Pasha (who had retreated from southern Sudan into Uganda after the fall of Khartoum) had passed through the rather well-defined foci of human sleeping sickness along the River Congo (now Zaire) before arriving in Uganda and had brought the disease as infections among the Zanzibarian porters. After the meeting between Stanley and Emin Pasha near Lake Albert, where local species of tsetse could have transmitted the disease to Emin Pasha's followers, 7000–8000 of these were later marched south-east and settled in the fertile area north of Lake Victoria, where the epidemic of human sleeping sickness broke out in 1896 (Duggan, 1970). Stanley's was the first European-led expedition in Africa involving the exposure of thousands of people to what must have been for them entirely new strains of an African disease. Stanley's stated reason for his unusual route into Uganda was that he felt his porters, recruited in east Africa, would be keen to get home via Uganda if deposited on the west coast of the continent; in fact his hidden agenda involved King Leopold's wish for a Belgian territory in Africa (Pakenham, 1991). Catastrophic epidemics are often associated with the arrival of a conquering army closely followed by the parasites to which the army has already adapted (as in the case of the strains of plague probably introduced during the Mongol penetration of Europe, giving rise to the Black Death, and the equally devastating impact of European diseases on the Aztec and Inca cultures of America (McNeill, 1976)). Single historical events such as these have been termed 'contingencies' (Gould, 1989) and a case can be made that they, more than any other, have been responsible for sleeping sickness outbreaks in Africa. (It is equally challenging to imagine what might have happened had Stanley been murdered by his two disaffected European fellow travellers during an earlier ex-

pedition in 1871 between Bagamayo and Tabora in Tanzania. A bullet fired at him through his tent at night by one of these men missed him by inches (McLynn, 1989). Had Stanley been killed, would the Uganda sleeping sickness epidemic have occurred 600 miles away and 25 years later?

The five explanations given above for variation in trypanosome prevalence over time involve different mechanisms, different symptoms and different end results. Two of the three biotic mechanisms appear to generate periodic epidemics, and much of the literature on sleeping sickness in Africa suggests periodic recurrence of the disease in particular locations (Duggan, 1962; Morris, 1960). The first, abiotic explanation for sleeping sickness would also generate epidemic cycles if the driving variable, rainfall, showed cyclical changes of approximately fixed period. Autocorrelation and spectral analysis of the rainfall data for Tabora, however, show no evidence for regular cycles of rainfall. Close inspection of the field data on sleeping sickness cases also casts doubt on the occurrence of regular periodicity in sleeping sickness outbreaks. For example, the autocorrelation analysis of the Tanzanian sleeping sickness data suggesting an approximately 15–20 year period cycle (Rogers, 1988a) fails to reveal any convincing periodicity when data for the first epidemic outbreak (1925–35) are removed from the analysis. On a country-wide basis, sleeping sickness in Tanzania has shown a gradual decline since 1940 and has not obviously been cyclic. As mentioned earlier, the inter-epidemic period in the Busoga district of Uganda is so long that it is impossible to say whether there are regular cycles of sleeping sickness there. The historical contingencies of Stanley in the last century and Amin in the present may provide all the explanation needed for sleeping sickness outbreaks in this part of Africa. The most convincing example of sleeping sickness cycles appears to be that of the Lake Edward and Lake George region in the Toro district of Uganda from the 1920s to the 1950s (Morris, 1960), presented in Fig. 9, but even here the cycles appear to have died away by the end of the sequence. We are left, therefore, with the phenomenon of epidemic outbreaks of human sleeping sickness which may or may not be cyclic, brought about by five possible mechanisms which themselves may or may not generate periodic cycles. What should we be monitoring to anticipate future epidemic outbreaks we have so far failed to comprehend?

DISCUSSION AND CONCLUSIONS

At the present time we seem to be quite close to an understanding of the determinants of the distribution of tsetse at continental and regional scales through a combination of biological and statistical approaches. The biological approach requires de-

tailed population measurements and so is often restricted in the area that can be covered. The statistical approach, on the other hand, generally works best with extensive data sets covering a wide range of eco-climatic conditions. The analysis picks out the variables with the best predictive (i.e. discriminating) powers, and these should be closely monitored during precisely targeted biological studies arising from the statistical analyses. Through the rather simple relationship between challenge and risk, it should be possible to turn this understanding of fly distribution into a prediction of the spatial variation in disease risk for both humans and animals.

In contrast, our present understanding of temporal variation in disease risk is very poor. There are remarkably few data on the temporal variation in cattle trypanosomiasis and none at all for variation in the wild animal hosts. Epidemics of human sleeping sickness are a dramatic example of temporal variation in the risk to humans, but these must be considered against the background of the rather stable endemic pattern of the disease in many of the moister parts of west and central Africa. The most consistent of the five hypotheses advanced for the initial appearance of sleeping sickness in many areas would appear to be the most untestable—that of historical contingency (e.g. Dutton & Todd, 1960; Scott, 1965). After the initial introductions, which are often associated with epidemic outbreaks, the disease patterns may either stabilize (as they appear to have done in Tanzania) or continue to give periodic outbreaks (as in Uganda and the drier parts of West Africa). In the years between major epidemic outbreaks virtually no monitoring occurs. Even during such outbreaks little attention is paid to measuring precisely those variables and parameters that may ultimately be responsible for the outbreak. The eventual decline of the epidemic is usually attributed to the intervention of control services, but these may simply only hasten the inevitable decline that follows all epidemic outbreaks.

The data that should be collected to monitor changes in trypanosomiasis in space and time can now be stored in the spatially referenced data-bases of Geographical Information Systems. GISs provide the tools to manipulate the data within a geographical framework and the techniques for simple analysis of the stored data. As in the case of the related subject of image processing, however, most GISs provide little insight for the untrained research worker with no field experience of the problem concerned and are best suited to answer precisely targeted questions set by epidemiologists with experience of the problems in the field. They play a role similar to that of mathematical models in other fields of epidemiology, of exploring ideas rather than generating them. In gathering together the information to build a GIS, the research worker is forced to think about those

variables which are likely to be important in determining epidemiological patterns, and the relationships between the variables that can be explored within the GIS. The spatial element of many epidemiological problems is often currently ignored and the existence of GIS will not of itself remedy this. When epidemiologists pose their questions in spatial terms, however, GIS techniques will be available to help them.

ACKNOWLEDGEMENTS

We thank Dr D. Mbulamberi, the Director of the Ugandan Sleeping Sickness Control Programme, Jinja, Uganda and Dr J. Hielkema of the ARTEMIS Programme, FAO, Rome for the supply of sleeping sickness and NDVI data for Figure 7. The registration of NDVI images to the climatic data-base for Kenya and Tanzania was done by Dr Tim Robinson, to whom we are most grateful. Dr S. E. Randolph kindly read and commented upon the manuscript. The work described here was carried out with financial support from the UNDP/World Bank/WHO Special Programme for Research and Training in Tropical Diseases (Project No. 860010).

REFERENCES

- ANDERSON, R. M. & MAY, R. M. (1991). *Infectious Diseases of Humans: Dynamics and Control*. Oxford: Oxford University Press.
- APTED, F. I. C. (1970). The epidemiology of Rhodesian sleeping sickness. In *The African Trypanosomiases* (ed. Mulligan, H. W.), pp. 645–60. London: George Allen & Unwin.
- AUSTEN, E.E. (1903). *A Monograph of Tsetse-Flies*. London: British Museum (Natural History).
- BAKER, R. D., MAUDLIN, I., MILLIGAN, P. J. M., MOLYNEUX, D. H. & WELBURN, S. C. (1990). The possible role of Rickettsia-like organisms in trypanosomiasis epidemiology. *Parasitology* **100**, 209–17.
- CHORLEY, J. K. (1929). The bionomics of *Glossina morsitans* in the Umnati fly belt, Southern Rhodesia, 1922–1923. *Bulletin of Entomological Research* **20**, 279–301.
- CHORLEY, J. K. (1947). Distribution of tsetse in southern Rhodesia. *Conferencia Intercolonial sobre Tripanossomiasis, Lourenço Marques*, vol. 1, 63–6. Imprensa Nacional de Moçambique 1947.
- DRANSFIELD, R. D., BRIGHTWELL, R., KILU, J., CHAUDHURY, M. F. & ADABIE, D. A. (1989). Size and mortality rates of *Glossina pallidipes* in the semi-arid zone of southwestern Kenya. *Medical & Veterinary Entomology* **3**, 83–95.
- DUGGAN, A. J. (1962). A survey of sleeping sickness in Northern Nigeria from the earliest times to the present day. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **65**, 439–80.
- DUGGAN, A. J. (1970). An historical perspective. In *The African Trypanosomiases* (ed. Mulligan, H. W.), pp. xli–lxxxviii. London: George Allen & Unwin.
- DUTTON, J. E. & TODD, J. L. (1906). The distribution and spread of sleeping sickness in the Congo Free State with suggestions on prophylaxis. *Liverpool School of Tropical Medicine Memoir* **18**, 25–38.
- FAIRBAIRN, H. (1948). Sleeping sickness in Tanganyika Territory, 1922–1946. *Tropical Diseases Bulletin* **45**, 1–17.
- FAIRBAIRN, H. & CULWICK, A. T. (1950). Some climatic factors influencing populations of *Glossina swynnertoni*. *Annals of Tropical Medicine & Parasitology* **44**, 27–33.
- FORD, J. (1971). *The Role of Trypanosomiasis in African Ecology: a Study of the Tsetse-fly Problem*. Oxford: Clarendon Press.
- FORD, J. & KATONDO, K. M. (1977). *The Distribution of Tsetse Flies in Africa*. London: OAU/Cook, Hammond & Kell.
- GAMISSA, D.-G. (1986). The Shinyanga tsetse fly eradication campaign, *Swara* **9**, 11–13.
- GOULD, S. J. (1989). *Wonderful Life. The Burgess Shales and the Nature of History*. London: Hutchinson Radius.
- GREEN, P. E. (1978). *Analyzing Multivariate Data*. Hinsdale, Illinois: The Dryden Press.
- JENNI, L., MARTI, S., SCHWEITZER, J., BETSCHART, B., LE PAGE, R. W. F., WELLS, J. M., TAIT, A., PAINDOVINE, P., PAYS, E. & STEINERT, M. (1986). Hybrid formation between African trypanosomes during cyclical transmission. *Nature* **322**, 173–5.
- JUSTICE, C. O. (ed.) (1986). Monitoring the grasslands of semi-arid Africa using NOAA-AVHRR data. *International Journal of Remote Sensing* **7**, issue 11. London: Taylor & Francis.
- KILAMA, W. L., MTERA, K. N. M. & PAUL, R. K. (1981). Epidemiology of human trypanosomiasis in Tanzania. *International Scientific Council for Trypanosomiasis Research & Control, 17th Meeting, Arusha (Tanzania) 1981*, 187–93. Nairobi, Kenya: OAU/STRC.
- MCLYNN, F. (1989). *Stanley: The Making of an African Explorer*. Oxford: Oxford University Press.
- MCNEILL, W. H. (1976). *Plagues and People*. New York: Doubleday.
- MARRIOTT, F. H. C. (1974). *The Interpretation of Multiple Observations*. London: Academic Press.
- MAUDLIN, I. (1991). Transmission of African trypanosomiasis: interactions among tsetse immune system, symbionts and parasites. *Advances in Disease Vector Research* **7**, 117–48.
- MBULAMBERI, D. B. (1989). Possible causes leading to an epidemic outbreak of sleeping sickness: facts and hypotheses. *Annales de la Société Belge de Médecine Tropicale* **69**, Supplément 1, S173–9.
- MORRIS, K. R. S. (1946). The control of trypanosomiasis by entomological means. *Bulletin of Entomological Research* **37**, 201–50.
- MORRIS, K. R. S. (1952). The ecology of epidemic sleeping sickness II. The effects of an epidemic. *Bulletin of Entomological Research* **43**, 375–96.
- MORRIS, K. R. S. (1960). 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**, 212–24.
- MULLA, A. F. & RICKMAN, L. R. (1988). How do African game animals control trypanosomiasis infections? *Parasitology Today* **4**, 352–4.
- NICHOLSON, S. E., KIM, J. & HOOPINGARNER, J. (1988).

- Atlas of African Rainfall and its Interannual Variability.* Tallahassee, Florida: Department of Meteorology, Florida State University.
- PAKENHAM, T. (1991). *The Scramble for Africa.* London: Weidenfeld & Nicholson.
- ROGERS, D. J. (1979). Tsetse population dynamics and distribution: a new analytical approach. *Journal of Animal Ecology* **48**, 825–9.
- ROGERS, D. J. (1985). Trypanosomiasis 'risk' or 'challenge': a review. *Acta Tropica* **42**, 5–23.
- ROGERS, D. J. (1988a). The dynamics of vector-transmitted diseases in human communities. *Philosophical Transactions of the Royal Society of London, Series B*, **321**, 513–39.
- ROGERS, D. J. (1988b). A general model for the African trypanosomiases. *Parasitology* **97**, 193–212.
- ROGERS, D. J. (1989). The development of analytical models for human trypanosomiasis. *Annales de la Société Belge de Médecine Tropicale* **69**, Supplément 1, S73–88.
- ROGERS, D. J. (1991). Satellite imagery, tsetse and trypanosomiasis in Africa. *Preventive Veterinary Medicine* **11**, 201–20.
- ROGERS, D. J. & RANDOLPH, S. E. (1986). Distribution and abundance of tsetse flies (*Glossina* spp.). *Journal of Animal Ecology* **55**, 1007–25.
- ROGERS, D. J. & RANDOLPH, S. E. (1991). Mortality rates and population density of tsetse flies correlated with satellite imagery. *Nature* **351**, 739–41.
- ROGERS, D. J., RANDOLPH, S. E. & 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**, 403–23.
- SCOTT, D. (1965). *Epidemic Disease in Ghana 1901–1960.* London: Oxford University Press.
- SCOTT, D. (1970). The epidemiology of Gambian sleeping sickness. In *The African Trypanosomiases* (ed. Mulligan, H. W.), pp. 614–44. London: George Allen & Unwin.
- WELLDE, B. T., CHUMO, D. A., REARDON, M. J. & WAEMA, D. (1989). Epidemiology of Rhodesian sleeping sickness in the Lambwe Valley, Kenya. *Annals of Tropical Medicine & Parasitology* **83**, Supplement 1, S43–62.
- WILLETT, K. C. (1965). Some observations on the recent epidemiology of sleeping sickness in Nyanza Region and its relation to the general epidemiology of Gambian and Rhodesian sleeping sickness in Africa. *Transactions of the Royal Society of Tropical Medicine and Hygiene* **59**, 374–86.