INT. J. GEOGRAPHICAL INFORMATION SCIENCE, 2002
VOL. 16, NO. 4, 345–360

Research Article

An example of decision support for trypanosomiasis control using a geographical information system in eastern Zambia.

T. P. ROBINSON[1], R. S. HARRIS[2], J. S. HOPKINS[3] and B. G. WILLIAMS[4]

[1]International Livestock Research Institute, P.O. Box 30709, Nairobi, Kenya
[2]MRC Institute for Environment and Health, University of Leicester, Leicester, England, UK
[3]116 Cappel Lane, Standstead Abbots, Ware, SG12 8BY, England, UK
[4]CSIR, P.O. Box 91230, Auckland Park 2006, Johannesburg, South Africa

(*Received 14 December 2000; accepted 2 June 2001*)

Abstract. In many African countries where both Government resources and donor aid for the control of tsetse-transmitted trypanosomiasis are declining, there is an increasing need to identify areas where intervention is most likely to be technically, economically, socially and environmentally sustainable. Activities then can be focused so that the maximum benefits are obtained from limited resources. We describe a decision-support framework based on a geographical information system to identify areas of high priority for the control of tsetse and trypanosomiasis in the common fly belt of eastern Zambia. Digital coverages were generated for six environmental variables: (1) cattle density, (2) human density, (3) land designation, (4) relative arable potential, (5) crop-use intensity and (6) proximity to existing control operations. The distribution of tsetse in the area was predicted using a multivariate (maximum likelihood) analysis of areas of known presence and absence and a series of environmental data. Experienced Zambian veterinarians and biologists working in the region established criteria weights for the input variables and the data were integrated in a geographical information system (GIS), using weighted linear combinations to prioritize areas for trypanosomiasis control. The results of this exercise and estimates of the errors involved are discussed.

1. Introduction

1.1. *Tsetse-transmitted trypanosomiasis in Eastern Zambia*

Tsetse (Diptera: Glossinidae) are the vectors of the trypanosomes that cause ‘nagana’ in cattle and sleeping sickness in people. Kristjanson *et al.* (1999) estimated some 47.75 million cattle to be at risk from trypanosomiasis in sub-Saharan Africa. They further estimate that 38% of these are treated with trypanocidal drugs each year, at an annual cost of some US$35 million. In southern Africa, a supposedly discrete infestation of tsetse, covering an estimated 322 000 km^2 in Malawi, Mozambique, Zambia and Zimbabwe, is known as the ‘common fly belt’; the highest densities of tsetse are centred on the drainage systems of the Luangwa and Zambezi rivers. The geographical focus of this analysis was the Zambian part of the common

International Journal of Geographical Information Science
ISSN 1365-8816 print/ISSN 1362-3087 online © 2002 Taylor & Francis Ltd
http://www.tandf.co.uk/journals
DOI: 10.1080/13658810110095057

fly belt, given in figure 1, which shows the distribution of *Glossina morsitans morsitans* Westwood (Ford and Katondo 1977). The distribution of *G. pallidipes* Austen is a subset of that. Districts included in the analysis are labelled in figure 1.

In this part of Zambia, most agricultural land is devoted to rain-fed arable crops; maize is the staple diet. Oxen and ploughs are used to till the land where livestock can be kept; in other areas, hand-held hoes are used. Agricultural activity is generally far greater outside the bounds of the tsetse distribution where there is no restriction on draft power and much-larger areas can be cultivated. Household surveys (Regional Tsetse and Trypanosomiasis Control Programme (RTTCP), unpublished data) have shown that, in tsetse-free areas, more households own cattle (49% versus 33%), the average herd size is larger (5 versus 2), and the calving and weaning rates are

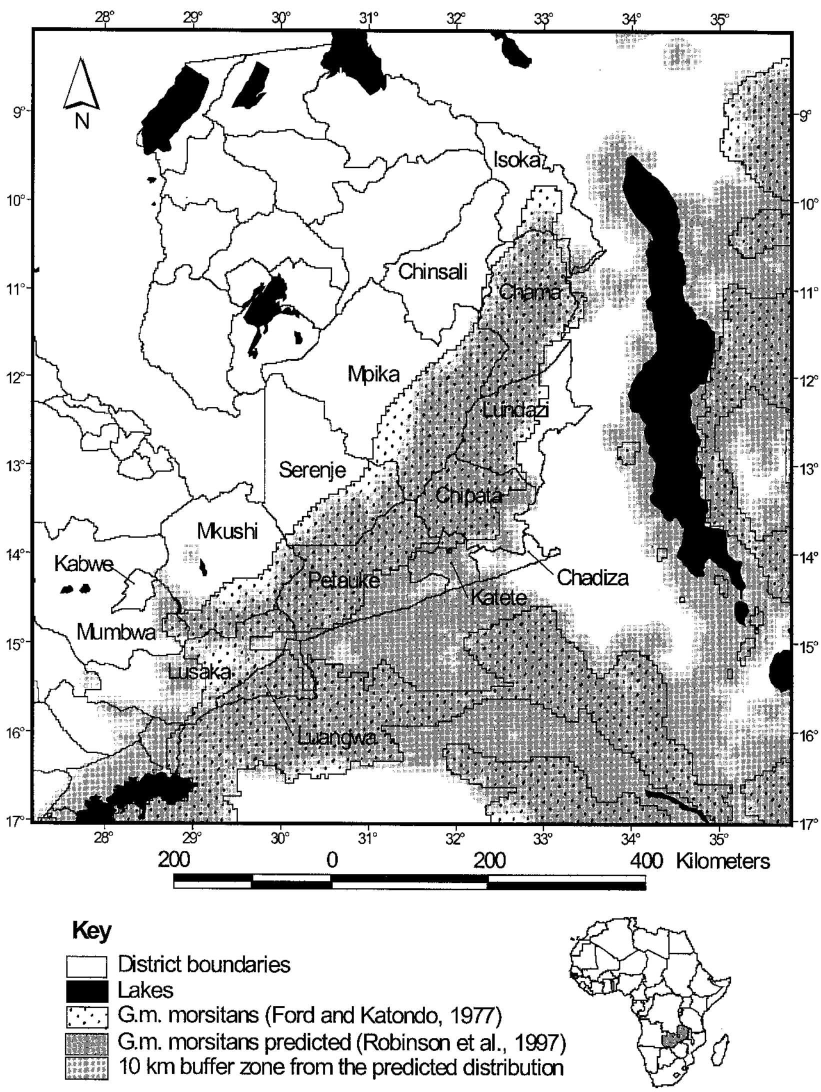


Figure 1. Map of eastern Zambia. The districts labeled are those included in the analysis. The predicted distribution plus the 10 km buffer zone were combined for use as a constraint. Further details are given in Appendix A.

significantly higher, as compared to infested areas. In tsetse-free areas, emphasis is given to herd growth and accumulation of breeding stock, as opposed to the infested areas where oxen constitute a major proportion of the average herd. They estimated that mean stocking densities in the areas free of trypanosomiasis could increase by 50% above present levels, and that in the infested areas, densities could increase by 400–500%.

Since 1986, donor assistance towards the control of tsetse-transmitted trypanosomiasis in the common fly belt has been directed through the RTTCP. This programme is operated under the auspices of the Southern African Development Council (SADC) and is funded by the European Commission (EC). The original objective of the regional programme was to eradicate tsetse from the common fly belt (Jordan 1985)—initially by aerial spraying and, later, using odour-baited, insecticide-impregnated target technology (Vale *et al.* 1988). The methodology behind this approach to tsetse control is that targets are progressively deployed to 'roll back the carpet' of tsetse. With this objective, a number of target control operations was started in Zambia.

In spite of the importance of trypanosomiasis, economic instability and donor-fatigue have lead to a shortage of operational funds within the Government Veterinary Department, resulting in a rapid decline of resources available for tsetse and trypanosomiasis control in Zambia. This, combined with a relative lack of success in 'area-wide' control, has resulted in a change in emphasis from widespread eradication towards smaller-scale, community-based interventions that require disease management rather than purely vector control. It is, therefore, increasingly important to identify areas that are of high priority for control.

1.2. *Application of GIS in trypanosomiasis control*

Geographical information systems (GIS) contain spatial data sets and the tools with which to display, manipulate and analyse them (Laurini and Thompson 1992; Burrough and McDonnell 1998). General reviews on their applications in epidemiology are available (Sanson *et al.* 1991, Elliott *et al.* 1992, Washino and Wood 1994, Mott *et al.* 1995, Clarke *et al.* 1996, Openshaw 1996, Hay 1997, Vine *et al.* 1997, Robinson 2000). Much research has applied these techniques to African animal trypanosomiasis; most used multivariate analysis of climate and remotely sensed data to model tsetse distributions. This has been done for Zimbabwe (Rogers and Williams 1993, 1994, Williams *et* al. 1994), Kenya and Tanzania (Rogers and Randolph 1993, Rogers and Williams 1993, 1994), West Africa (Rogers *et al.* 1996, Hendrickx 1999), and in the common fly belt of southern Africa (Robinson *et al.* 1997b).

GIS has also been used to combine data such as livestock distribution, agriculture and arable potential, at a range of spatial scales, to identify areas where tsetse flies constrain agricultural development.

At a coarse spatial resolution of c. 14 km × 14 km, trypanosomiasis risk has been linked to infrastructure and control opportunities in Togo (Hendrickx 1999). At a final spatial resolution of c. 1.25 km × 1.25 km in eastern Zambia, Robinson (1998) developed models based on decision trees to combine environmental data on the distribution of tsetse, livestock, agriculture, arable potential and protected forest and wildlife areas to produce priority maps for trypanosomiasis intervention. At a still finer resolution of c. 100 m × 100 m, De la Rocque (1997) generated a GIS model for an area in Burkina Faso, that combined biophysical and spatial elements to identify high-risk locations for trypanosome infections transmitted by riverine species of tsetse.

1.3. *Multiple-criteria decision making*

Various multiple-criteria decision models were reviewed by (Jankowski 1995) who categorized them according to whether or not they were compensatory. In non-compensatory models a higher score on one criterion cannot offset a low score on another, and options are eliminated without necessarily considering all criteria (Hwang and Yoon 1981, Minch and Sanders 1986). In compensatory models, however, a higher score on one criterion can offset a low score on another, and all criteria must be considered.

There are many ways in which decision criteria can be combined. Weighted linear combination (Rao *et al.* 1991, Eastman *et al.* 1993a, b) involves summing the criteria after multiplying each by a weight. Concordance-discordance analysis (Nijkamp and van Delft 1977, Voogd 1983, Nijkamp *et al.* 1990, Carver 1991) is a method in which each pair of alternatives (raster pixels or polygons in a GIS context) is analysed for the degree to which one outranks the other in the specified criteria. Analytical hierarchical processes (Saaty 1977, 1987) are an adaptation of weighted linear combination, and there also exist various 'multi-attribute trade-off systems' (Brown *et al.* 1986). A further group of multiple criteria models is based on specifying an 'ideal point' of suitability for an objective in an *n*-dimensional criterion space (Hwang and Yoon 1981, Voogd 1983, Lofti *et al.* 1992).

Most of the techniques described above are designed to compare relatively small numbers of alternatives and are computationally very intensive and are therefore difficult to implement in a raster GIS where each pixel represents an alternative. Weighted linear combination, however, is appropriate to implement in this context and some GIS programmes such as Idrisi (IDRISI Clarke Laboratories, USA, http://www.clarklabs.org) and SPANS (Tydac Technologies, Canada, http://tcp.ca/Aug94/SPANS.html) offer bespoke modules to conduct such analyses. Weighted linear combination is the method developed in this paper, and applied to the problem of prioritizing areas for the control of African animal trypanosomiasis.

Heywood *et al.* (1995) and Malczewski (2000) discuss some issues concerning the application of weighted linear combination in a GIS environment. They review some applications, explain the assumptions behind the methodology and highlight some common pitfalls. In the discussion section we will address some of these potential problems in relation to the present analysis.

2. Methods and results

Weighted linear combination in a GIS environment involves a number of steps: (*a*) define the objective; (*b*) identify relevant criteria; (*c*) input, geo-register and standardize the criteria; (*d*) develop and evaluate the decision rule; and (*e*) estimate the error in the output.

2.1. *Define the objective*

Whilst the general objective was to prioritize areas for tsetse and trypanosomiasis control, this had to be defined more precisely because subtle changes in the objective would be reflected by changes in the decision rule. In this example, with only one objective, we avoided the complexities that can arise through conflicting objectives in multiple objective decision making (Eastman *et al.* 1993a, b). In the present example the defined objective was:

> 'To identify areas where tsetse control can best be accomplished using odour-bated, insecticide-impregnated targets maintained by the community at minimal donor cost'.

Robinson (1998) discussed two situations where tsetse control can be prioritized. First, where the disease is constraining agricultural development directly; second, where the disease prevents expansion from areas of high land pressure, into adjacent areas that could be utilized to relieve this pressure. In this example we addressed the first of these problems, the direct disease constraint.

2.2. *Identify relevant criteria*

Criteria are selected that are linked to the defined objective. Criteria may be continuous variables such as cattle density or discrete classes such as land use that can be scored for suitability to the objective, or they could be Boolean constraints that limit the suitability of an area for tsetse control (such as the presence or absence of tsetse).

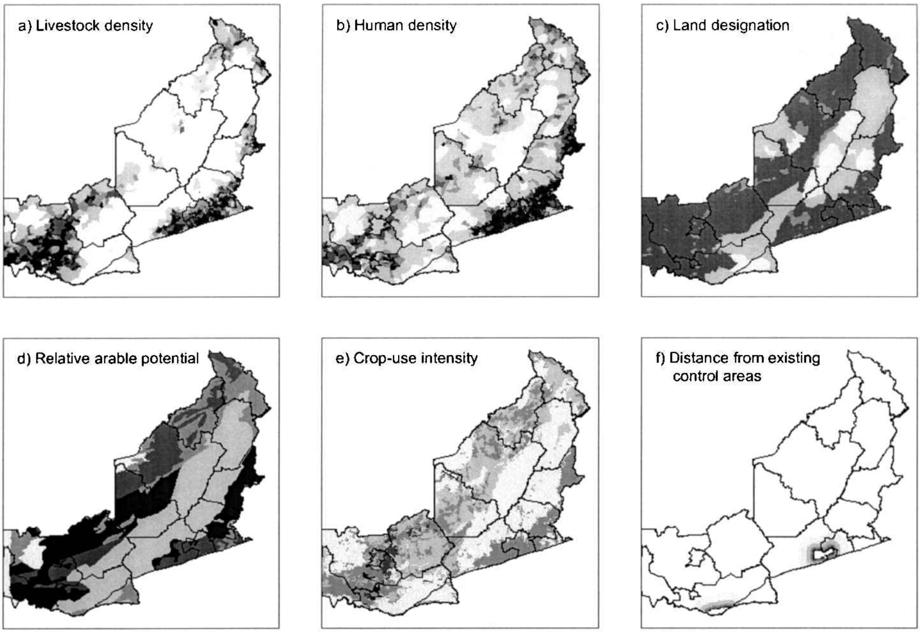
The criteria used in the present analysis were (*a*) cattle density, (*b*) human density, (*c*) crop-use intensity, (*d*) relative arable potential, (*e*) land designation, (*f*) proximity to existing control operations, and (*g*) tsetse distribution (Appendix A). In terms of the defined objective, areas of relatively high cattle and cropping density are of higher priority for tsetse control because benefits arise through intensification of agriculture. For a community-based operation the higher the density of people, the lower the per capita costs and the greater the likelihood of success and sustainability. It is clearly preferable to prioritize areas with a greater potential for arable development, and because it is important to avoid encroachment of conserved areas by clearing these of tsetse, land designation must also be considered. Distance from existing control operations is important when using targets because it is cheaper to build upon existing operations and because the fly-invasion front is minimized.

2.3. *Input, geo-register and standardize the criteria*

Selected criteria are then processed in a GIS, digitized if necessary, registered to a common scale and map projection and standardised such that (*a*) increasing suitability to the objective is associated with increasing data values and (*b*) the absolute range of data values is similar in each criterion. The six factors listed above are shown in figure 2 and the ways in which they were generated are described in Appendix A. The constraint of tsetse presence was derived from the multivariate predictions shown in figure 1 and described in Appendix A. All data were transformed to a digital raster-format (on a Platte Carrée projection) at a resolution of 80 pixels per degree; at these latitudes this corresponds to a pixel size of approximately $1.4\,\mathrm{km} \times 1.4\,\mathrm{km}$.

2.4. *Develop and evaluate the decision rule*

Weighted linear combination is a convenient way to combine continuous or graded criteria (Rao *et al.* 1991, Eastman 1992, Eastman *et al.* 1993b). In this technique, each standardized criterion is assigned a weight that reflects its relative importance in contributing to the objective. *S*, the suitability to the objective being



Key	Class	Livestock density (TLU km^{-2})	Human density (km^{-2})	Land designation	Relative arable potential	Crop-use intensity (%)	Distance from existing control (decimal °)
	1	No Cattle	No People	National Park			> 0.400
	2	< 1	< 1		Swamp	0 – 5	0.201 – 0.400
	3	1 – 5	1 – 5	Game Management Area	Low potential	6 – 30	0.101 – 0.200
	4	6 – 10	6 – 10			31 – 50	0.051 – 0.100
	5	11 – 20	11 – 20	National forest	Intermediate potential	51 – 70	0.025 – 0.050
	6	21 – 40	21 – 40			71 – 100	< 0.025
	7	> 40	> 40	Communal land	High potential		

Figure 2. Input criteria used in the weighted linear combination to prioritize areas for tsetse control. The higher the class values (darker shading) the more suitable the factor for tsetse control. Further details are given in Appendix A.

considered (tsetse control) is defined as

$$S = \sum_{i=1}^{n} w_i x_i \prod_{j=1}^{m} c_j \qquad (1)$$

where w_i is the weight of factor i, x_i is the criterion score of factor i, n is the number of factors and c_j is the criterion score (1 or 0) of constraint j and m is the number of constraints.

In the present example the criterion weights were assigned by a group of experienced Zambian veterinarians and biologists working in the region. They were assigned randomly into four groups of five or six individuals and asked to provide weights for the six criteria listed to meet the stated objective. Specifically, they were asked first to decide which criterion was least important and to give this a value of 1. They were then asked to give weights to the other criteria indicating how much more important they were, compared to the least important criterion. The weights were then normalized such that for each group $\Sigma w_i = 1$.

Table 1 shows the criterion weights for tsetse control assigned by four groups of Zambian veterinarians and biologists and their mean values. The standard error of the mean weight is also given for each criterion.

The groups generally agreed that cattle density and relative arable potential were

Table 1. Criterion weights for tsetse control in eastern Zambia assigned by four groups (Grp 1–4) of Zambian veterinarians and biologists and their mean value. The standard errors (SE) of the means are given. The errors in each criterion were estimated by the authors as plus or minus a proportion of class ($\pm a_i$) and the resulting variance (V_i) was calculated. For each group, and for the mean, the lower part of the table gives the variance of the estimate of *S* (the suitability for control) calculated using equation (2), and its standard deviation and 95% confidence limits.

	Weight (W_i)							
	Grp 1	Grp 2	Grp 3	Grp 4	Mean	SE	Error ($\pm a_i$)	Var. (V_i)
Cattle density (*a*)	0.13	0.31	0.24	0.23	0.22	0.04	1.0	0.333
Human population density (*b*)	0.13	0.15	0.19	0.23	0.17	0.02	1.0	0.333
Land designation (*c*)	0.25	0.08	0.10	0.05	0.12	0.04	0.0	0.000
Relative arable potential (*d*)	0.25	0.15	0.29	0.23	0.23	0.03	0.5	0.083
Crop use intensity (*e*)	0.13	0.04	0.05	0.09	0.08	0.02	0.5	0.083
Prox. to existing control area (*f*)	0.13	0.27	0.14	0.18	0.18	0.03	0.0	0.000
Variance of the estimate of S	0.02	0.04	0.04	0.04	0.03			
Standard deviation of S	0.13	0.20	0.20	0.20	0.18			
95% confidence limits in S	0.26	0.40	0.38	0.39	0.50			

important and these were weighted highly; land designation and crop-use intensity were generally considered relatively unimportant. Groups 2, 3 and 4 were quite consistent in their ranking of criteria, each giving similar values, whilst group 1 differed somewhat from these.

For each pixel the suitability (*S*) for tsetse control was then estimated from the mean weights using equation (1)

$$S=(a\times 0.22)+(b\times 0.17)+(c\times 0.12)+(d\times 0.23)+(e\times 0.08)+(f\times 0.18)$$

and the value of *S* was multiplied by the Boolean product of the criterion scores for the constraints (in this case 1 if tsetse is present and 0 if tsetse is absent).

The result was the classified image given in figure 3, with values ranging from 1 to 6 (the higher values indicating areas of high priority for tsetse control). The highest-priority areas (dark) appear to be influenced strongly by cattle and human densities (figure 2 (*a*) and (*b*)). Within these areas, those that are close to the existing control operations have been highlighted as being of highest priority. Parts of the Eastern Province show the highest priority rankings and Lusaka Province comes next in terms of priority ranking. Very few areas of high priority are indicated in the Central and Northern Provinces.

2.5. *Estimate the error in the output*

To explore the effect of variability in choice of weight among the different groups (given by the standard error (SE) of the mean weight for each criterion in table 1) we conducted a sensitivity analysis by creating tsetse control priority maps using each of the groups' weights individually. Figure 4 shows these maps, zooming in on the areas of higher control priority in the Eastern Province. The general pattern of priority areas is very similar among the groups, suggesting the mean to be a reasonable summary of the groups' conclusions, so we need not be too concerned about accounting for this source of variability.

Each criterion shown in figure 2 has an associated error that can be defined in

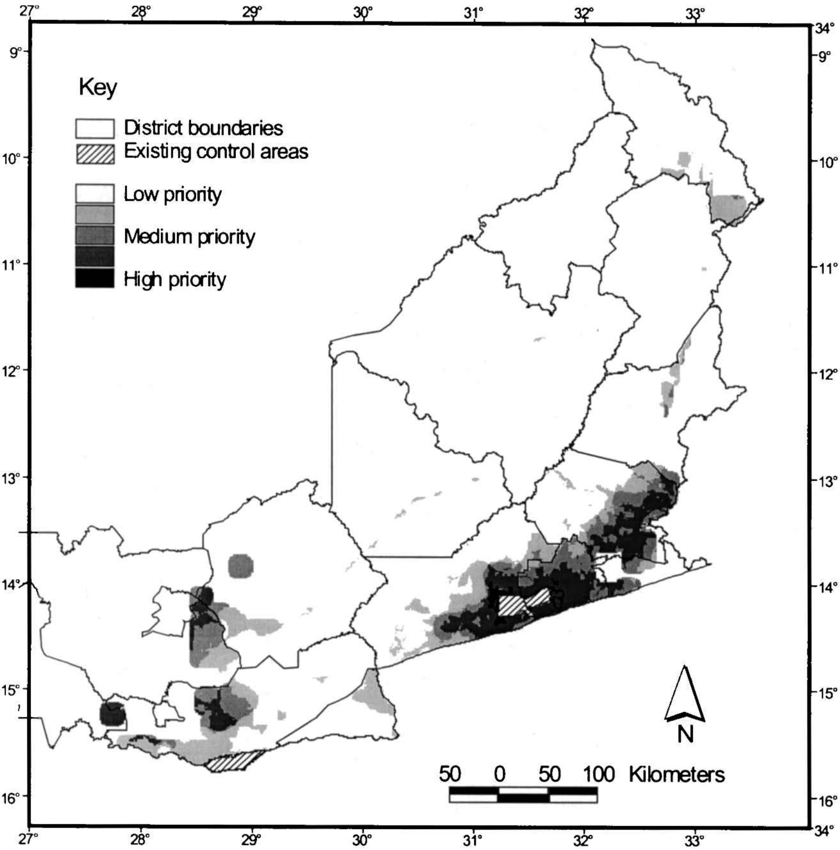


Figure 3. Result of weighted linear combination of the factors in figure 2, and the tsetse distribution constraint shown in figure 1.

terms of 95% (or other) confidence limits. In this analysis, errors were not measured but estimated by the authors, and then combined to give confidence limits to the resulting map of priority areas for tsetse control.

To estimate the sampling errors in S (our final priority map) we need to know the sampling errors in x_i, the original criteria. We assumed that there was no error in the constraint of tsetse presence, and we further assumed that the errors in the six factors (x_i) could be estimated as plus or minus some fraction of a class which we called a_i. Choosing numbers from a uniform distribution ranging from $-a_i$ to $+a_i$, the variance of the x_is is $V_i = a_i^2/3$. The variance of S was then

$$V(S) = \sum_{i=1}^{n} w_i^2 V_i \tag{2}$$

where w_i remained the weight of factors i through n. The standard deviation was then estimated as $\pm\sqrt{V(S)}$ and to get 95% confidence limits we simply multiplied the standard deviation by 1.96 (assuming that the errors in S were normally distributed). This provided the fraction of a class in the final suitability map within which we could be 95% certain that the estimated suitability value lied.

The right-hand columns of table 1 show our estimates of the sampling errors ($\pm a_i$) and variance (V_i) in each factor based on our understanding of how the data

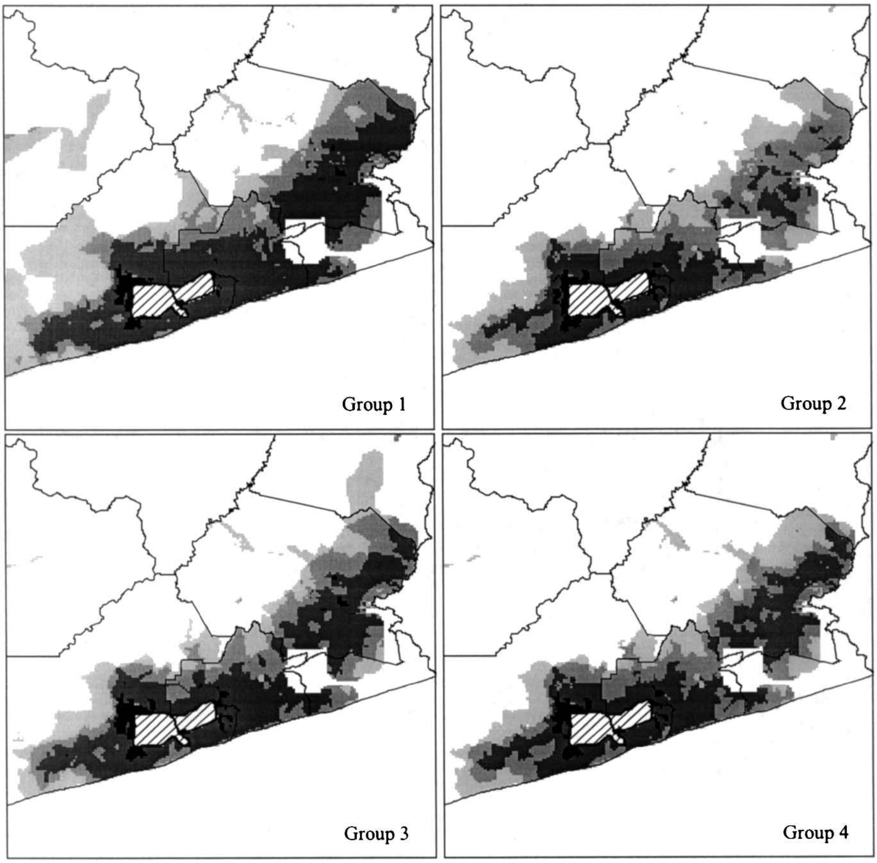


Figure 4. Result of weighted linear combinations conducted separately for the weights derived by each group. The area shown is the high-priority Eastern Province of Zambia. The key in figure 3 applies.

were generated. We used equation (2) to calculate the variance of the estimate of S to be 0.03. The standard deviation of S was then 0.18 and multiplying this value by 1.96 gave us 95% confidence limits of ± 0.35 of a class in our allocation of pixels to the priority classes for tsetse control.

3. An example from Petauke in the Eastern Province of Zambia

Figure 5 shows a detail of the mean priority map (figure 3); zooming in on Petauke, an important part of the Zambian fly belt in which tsetse control, using odour-baited targets, has been going on since 1989. Area A ($320\,km^2$) is the Chimpundu area in which tsetse control commenced in 1989 under the Belgian Animal Disease Control Project (BADCP) using locally made black-screen targets; area B ($550\,km^2$) is the Mvuvye area, in which tsetse control was initiated by the RTTCP in 1990 using black-screen targets imported from Zimbabwe. In 1992 the management of the Chimpundu area was taken over by the RTTCP. Thus, areas A and B comprise the original Petauke target control area; black-screen targets are maintained at a density of 4 per square kilometre.

In 1994 a community-based extension to the Petauke target control area was

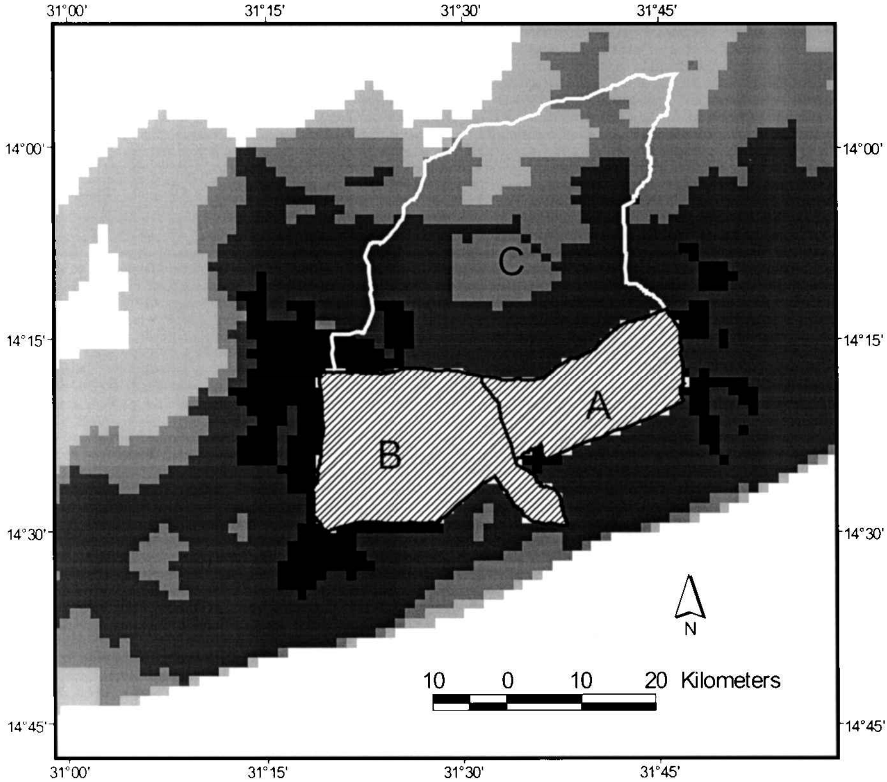


Figure 5. Detail from figure 3 (the key in figure 3 applies). Areas A and B are the existing target-control operations in Petauke District and area C, bordered by a white line, is the proposed community-based extension. The figure shows the northern half of the proposed extension to include areas of relatively low priority while, to the west of area B, areas of highest priority are not included.

proposed, the Msanzara area shown as area C in figure 4 (1272 km^2). Extensive tsetse surveys demonstrated a high infestation of *G. m. morsitans*.

The results presented in figure 5 show quite clearly that the northern part of the proposed extension to the Petauke control operation was of relative low priority because it was (*a*) of relatively low arable potential, (*b*) of a low crop-use intensity, and (*c*) relatively far from the existing control areas. Areas of higher priority are indicated to the west of area B in figure 5. Had the results of this analysis been available at the time they could have been used to help plan the extension to the Petauke control operation by optimizing its location.

Soon after the initiation of the community control programme the RTTCP activities in the region began to decline as the programme eventually closed down. It has not, therefore, been possible to assess the impact of this control operation, nor to evaluate the outputs of this analysis in the field.

4. Discussion

The results presented demonstrate the application of weighted linear combination, within a GIS, to integrate relevant criteria for decision making in trypanosomiasis control. Heywood *et al.* (1995) and Malczewski (2000) highlight potential problems with using weighted linear combination in a GIS environment. Here we combine the most important of these with our own observations, under seven categories.

a) *Objective definition*. Defining objectives is central to multi criteria evaluation: the more clearly the problem is defined the more specific the decision rule can be. Heywood *et al.* (1995) discuss some issues surrounding problem definition. In the present example we started with the broad objective of prioritizing areas for trypanosomiasis control, and then made this more specific by stating that it would be a community operation, using insecticide-impregnated targets.

b) *Attribute completeness*. Malczewski (2000) highlights an assumption behind weighted linear combination that the criteria must be complete, i.e. they should cover all aspects of the decision problem. Selection of attributes is necessarily driven by data availability but it is important to ensure that no key attributes are missing from the analysis as this could completely invalidate the result. In our analysis we have incorporated the important variables, though it would be preferable to have more detail in some of these. For example estimates of disease risk would be better than simple estimates of tsetse presence and economic data on potential benefits of control would be better than estimates of livestock numbers, crop-use intensity and arable potential.

c) *Attribute independence*. The issue of attribute independence is highlighted both by Heywood *et al.* (1995) and Malczewski (2000) and clearly needs further research; the risk being that the influence of highly correlated variables is overstated in the application of the decision rule. There is a high level of cross-correlation between some of the variables used in our analysis, particularly between livestock and human densities. We cannot simply drop one of a pair of correlated variables such as these since it is often in areas where the correlation breaks down that are of particular importance, for example where people occur but cannot keep livestock due to the trypanosomiasis constraint. Perhaps to retain one of these variables and a composite (difference or ratio) of the two might be the way forward, though we then run the risk of deriving complex attributes that become difficult to scale and assign weights to.

d) *Attribute linearity*. A further assumption behind weighted linear combination (and one that we have made in our example) is that of linearity in the input criteria, in terms of suitability to the specified objective. For simple cases of non-linearity in attribute suitability (e.g. exponential increase in suitability) some simple arithmetic transformation may be used to scale the data. Malczewski (2000) explains a further solution to non-linearity in the 'value function approach'. In the commonly used 'midvalue method' (e.g. Lai and Hopkins 1989) the decision maker estimates the midpoint between the minimum (0) and maximum (1) value of a scaled attribute, based not on its absolute value but on its suitability to the objective, and assigns that a value of 0.5. Quartiles are then defined in the same way, etc. until the required resolution is obtained.

In situations of non-linearity where the optimal value of a criterion occurs within a distribution rather than at one extreme, or indeed there is more than one optimal value, the problem is more complex. The lies in the group of decision support models in which an ideal point (or points) of suitability for an objective is (are) located in an *n*-dimensional criterion space, and each pixel is then classified according to its statistical proximity to that point (or to those points).

e) *Scale and aggregation*. Malczewski (2000) describes at length the risks associated with varying spatial scale and levels of aggregation in spatial applications of weighted linear combination; the best alternative at one spatial scale does not necessarily hold at another. In the present analysis we were fortunate to have data of comparable and relatively high spatial resolutions so these problems did not affect us.

f) *Weighting*. Incorrect specification of weights is a common error in weighted linear combination, as highlighted by Hobbs (1980) and Lai and Hopkins (1989). Weights are often assigned without due consideration of the scaling or measurement units of the criteria; the influence of a weight is clearly dependent on the data range. Malczewski (2000) describes the 'swing weights technique' where the decision maker compares a change from the least to the most preferable value in one criterion with the same change in another; ranks the criteria depending on which of these 'swings' has the biggest influence on the decision; and weights criteria relative to their performance compared to the highest ranking. Some analysts (e.g. Hobbs 1980) advocate the 'analytical hierarchy process' developed by Saty (1977), in which a pairwise comparison technique is used to ensure consistency in the attribute weights. In the present example we adopted an approach similar to Malczewski's (2000) 'swing weights technique', but assigned weights relative to the least important criterion. Here, the knowledge and experience of local experts was harnessed, though a more rigorous approach might be to estimate weights statistically, using the evidence on the efficacy of control strategies in various environments and on economic models of the costs and benefits of different control options. For the present analysis such evidence was not available but this is an important area for future research.

g) *Error assessment*. Assessment of errors in weighted linear combination is a further matter for consideration, and one too frequently ignored. There are essentially two types of error that are quite different: errors caused by inaccuracies in the estimated criteria, database uncertainty, and those introduced in the development of the decision rule, decision rule uncertainty. In the present analysis we estimated database uncertainty by defining confidence intervals for the priority classes, based on out confidence in the attribute values. We assesssed decision rule uncertainty by conducting a sensitivity analysis of the different weightings provided by different groups. Various other methods have been reported to address decision rule uncertainty, including Bayesian probability theory (Lee *et al.* 1997), fuzzy set theory (Fisher 1991, Lee *et al.* 1997, Zadeh 1965) and Dempster-Shafer theory (Lee *et al.* 1997). Monte Carlo simulation has also been used to investigate sensitivity to changes in the importance of criteria within the decision rule (Janssen 1992).

The next step that needs to be taken in developing decision support models for trypanosomiasis control is to adapt to a GIS environment the group of models mentioned above, in which an ideal point of suitability for an objective is located in n-dimensional criterion space. That approach has the potential to overcome many of the potential problems listed above, in particular that of attribute linearity.

Moreover, this approach has the advantage that it tends towards multiple objective decision making, which may be helpful in deciding which control methods should be recommended for intervention. The options for trypanosomiasis control include disease control using trypanocidal drugs, adoption of trypanotolerant livestock, vector control using insecticide-treated animals or artificial bait technology and vector eradication using methods such as ground or aerial spraying or the sterile insect technique. These vary in cost, in whether the benefits are private (e.g. trypanocidal drugs) or public (e.g. bait technology) and also in whether the onus is on the individual, the community or some external body to conduct and finance them.

The ecological, epidemiological, social and economic environment will have a very strong bearing on which alternative, or combination of alternatives, is most appropriate to a particular area. This choice will depend on criteria such as infection rate, risk and rate of re-invasion of flies to a cleared or controlled areas, potential benefits (including factors such as market access and potential for intensification), costs (including infrastructural considerations), and a range of social factors. One can visualize a multidimensional criterion space in which ideal points are located for different control approaches and, based on these criteria, pixels would be classified according to which point or points they were statistically closest to. In some cases a combined approach might be the recommended output of the decision support model.

There is clearly enormous potential for decision support tools to be applied to the problem of trypanosomiasis, and to other livestock diseases, within a spatial context. The next task is to identify suitable case studies, where detailed data are available or can be collected, with which to derive more sophisticated decision rules that can be adapted to a GIS environment.

Appendix A. Origins and accuracy of the criteria used in the analysis

A.1. *Cattle density*

Cattle density was derived from the 1990 census data for Zambia. These were summarized as tropical livestock units (TLU) per census standardized area (CSA). One TLU is approximately equivalent to 250 kg of live weight; the average unit value for cattle is 0.7 TLU. Spatial errors arise in the mapping, digitizing and projection transformation of the original CSA maps. These should be relatively small. Greater errors occur during collection of statistics for each CSA. When the original data were considered in detail, these errors were in some cases clearly very large; for example some CSAs had been assigned impossibly high livestock densities. By classifying the livestock densities into relatively broad classes, although we lost precision, the importance of these errors was greatly reduced. Further errors are inherent in the assignment of average density classes throughout each CSA; these errors of generalization are greater in larger CSAs.

A.2. *Human population density*

Human density classes were generated in the same way from the 1990 census data. Errors are likely to be similar to those in cattle density, and originate in the same way.

A.3. *Land designation*

The locations of national parks and game-management areas on the 1:250 000 scale topographic maps were digitized. A map of national forests was also digitized and geo-registered with the map of national parks. Areas falling into none of the three classes above were designated ‘common land’. Errors in the land designation criterion arise only from mapping, digitizing and projection transformation. These should be relatively small.

A.4. *Relative arable potential*

A map of relative arable potential, based on soil, relief and length of growing season data was obtained. This map divides the common fly belt of Zambia into five relative arable potential classes (indicating the suitability for rain-fed crops with a moderate level of management). Errors in estimating relative arable potential are likely to be quite large. Additional to the mapping errors are those from the source data used to produce this coverage (soil, relief and length of growing season) and errors resulting from the subjective process of assigning relative arable potential classes.

A.5. *Crop-use intensity*

The World Food Programme/Famine Early Warning System's (WFP/FEWS) crop-use intensity map for Zambia was used to estimate the amount of cropping. Landsat imagery was used to divide the area into five classes: <5%; 5–30%; 31–50%; 51–70% and 71–100% of land under cultivation. The methodology is described in (Westin and Brandner 1980) and uses a combination of colour, tone, pattern, evidence of wetness, aridity and saline conditions and, where available, soil, climate, geological, geomorphic and topographic information. The coverage originated as a raster-format GIS file, on a Plate Carrée projection, at a resolution of 333.367 pixels per degree. This was re-sampled to the base grid using the nearest-neighbour method. Spatial errors in the crop-use intensity criterion arise through the projection and scale transformations of the original Landsat satellite imagery. Value errors are introduced that are inherent in the source data, and there will be additional errors resulting from the subjective process of assigning crop-use intensity classes.

A.6. *Distance from existing control operations*

The existing target control operations in Zambia were drawn onto the 1:250 000 scale topographic maps, and digitized. The GIS function of buffering was used to assign each pixel a value corresponding to the shortest distance to the nearest target control area. Because of the use of the Platte Carrée projection these distances are expressed in decimal degrees. In this region, one decimal degree is approximately 100 km in both the north-south and east-west directions. We decided that proximity to existing control areas was no longer of any importance at distances of greater than 0.400 decimal degrees (c. 40 km). Areas further than this (and within the control operations themselves) were given a value of zero in the class map. Errors in the criterion of proximity to the existing target control area should be small, resulting only from the cartographic and projection transformation errors, and those introduced in the buffering process.

A.7. *Tsetse distribution*

The distribution of *Glossina morsitans morsitans* was predicted using maximum-likelihood classification of a range of remotely sensed environmental variables, based upon the distribution maps of (Ford and Katondo 1977). The method is described fully in (Robinson *et al*. 1997a). The results of the prediction gave an overall correspondence of 85.1% with the training data; correctly predicted pixels of suitability for tsetse (sensitivity) were 83.2% and correctly predicted pixels of non-suitability for tsetse (specificity) were 83.2 %. The Kappa index of agreement between the prediction and the training data was high (0.641) indicating a good model fit. For this exercise, the predicted tsetse distribution was increased by a 10 km buffer zone.

Acknowledgments

This work was conducted and written under ODA grants IPMI T00224 and R6558, and IFAD grant TAG 284. We are grateful to Simon Hay, Mike Packer, Willy Wint, David Rogers, Mark Eisler and three independent referees for their useful comments on the manuscript.

References

Brown, C. A., Stinson, D. P., and Grant, R. W., 1986, Multi-attribute tradeoff systems: Personal computer version users' manual. Public and Social Evaluation Office, Division of Planning Technical Services. Engineering Research Center, Bureau of Reclamation.

Burrough, P. A., and McDonnell, R. A., 1998, *Principles of Geographical Information Systems* (Oxford: Oxford University Press).

Carver, S. J., 1991, Integrating multi-criteria evaluation with Geographic Information Systems. *International Journal of Geographic Information Systems*, **5**, 321–339.

Clarke, K. C., McLafferty, S. L., and Tempalski, B. J., 1996, On epidemiology and geographic information systems: a review and discussion of future directions. *Emerging Infectious Diseases*, **2**, 85–92.

De la Rocque, S., 1997, Identification des facteurs discriminants majeurs de la presence des glossines dans une zone agro-pastorale du Burkina Faso. Interet pour la prevision du risque trypanosomien. PhD thesis, Universite Montpellier, Montpellier.

Eastman, J. R., 1992, *IDRISI: a grid based geographic analysis system. Version 4.0*. (Worcester, Massachusetts: Clarke University Graduate School of Geography).

Eastman, J. R., Jin, W., Kyem, P. A. K., and Toledano, J., 1993a, Participatory multi-objective decision-making in GIS. In *Autocarto,* pp. 33–43.

Eastman, J. R., Kymen, P. A. K., and Toledano, J., 1993b, A procedure for multi-objective decision making in GIS under conditions of conflicting objectives. In *European Conference on Geographical Information Systems,* pp. 438–447.

Elliott, P., Cuzick, J., English, D., and Stern, R., 1992, *Geographical and Environmental Epidemiology* (Oxford: Oxford University Press).

Fisher, P. E., 1991, Spatial data sources and problems. In *Geographic information systems: Principles and applications*, edited by D. J. Maguire (London: Longman), pp. 175–189.

Hay, S. I., 1997, Remote sensing and disease control: past, present and future. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, **91**, 105–106.

Hendrickx, G., 1999, Georeferenced decision support methodology towards trypanosomosis management in West Africa. PhD thesis, Universiteit Gent, Gent. pp. 177.

Heywood, I., Olivier, J., and Tomlinson, S., 1995, Building an exploratory multi-criteria modelling environment for spatial decision support. In *Innovations in GIS 2,* edited by P. Fisher (London: Taylor and Francis), pp. 127–136.

Hobbs, B. F., 1980, A comparison of weighting methods in power plant sighting. *Decision Sciences*, **11**, 225–237.

Hwang, C. L., and Yoon, K., 1981, *Multiple Attribute Decision Making Methods and Applications: A State of the Art Survey* (Berlin: Springer-Verlag).

Jankowski, P., 1995, Integrating geographical information systems and multiple criteria decision-making methods. *International Journal of Geographical Information Systems*, **9**, 251–273.

Janssen, T., 1992, *Multiobjective decision support for environmental management* (Dordrecht: Kluwer Academic Press).

Jordan, A. M., 1985, Tsetse eradication plans for southern Africa. *Parasitology Today*, **1**, 121–123.

Kristjanson, P. M., Swallow, B. M., Rowlands, G. J., Kruska, R. L., and de Leeuw, P. N., 1999, Measuring the costs of african animal trypanosomosis, the potential benefits of control and returns to research. *Agricultural Systems*, **59**, 79–98.

Lai, S. K., and Hopkins, L. D., 1989, The meanings of trade-offs in multiattribute evaluation models: a comparison. *Environment and Planning B*, **16**, 155–170.

Laurini, R., and Thompson, D., 1992, *Fundamentals of Spatial Information Systems* (New York: Academic Press).

Lee, N. S., Grize, Y. L., and Dehnad, K., 1997, Quantitative models for reasoning under uncertainty in knowledge-based expert systems. *International Journal of Intelligent Systems*, **II**, 15–38.

Lofti, V., Stewart, T. J., and Zionts, S., 1992, An aspiration-level interactive model for multiple criteria decision making. *Computer Operations Research*, **19**, 671–681.

Malczewski, J., 2000, On the use of weighted linear combination method in GIS: common and best practice approaches. *Transactions in GIS*, **4**, 5–22.

Minch, R. P., and Sanders, G. L., 1986, Computerised information systems supporting multi-criteria decision making. *Decision Sciences*, **17**, 395–413.

Mott, K. E., Nuttall, I., Desjeux, P., and Cattand, P., 1995, New geographical approaches to control of some parasitic zoonoses. *Bulletin of the World Health Organisation*, **73**, 247–257.

Nijkamp, P., Rietveld, P., and Voogd, H., 1990, *Multicriteria Evaluation* (Amsterdam: North Holland).

Nijkamp, P., and van Delft, A., 1977, *Multi-criteria Analysis and Regional Decision-making* (Amsterdam: Leiden).

OPENSHAW, S., 1996, Geographical information systems and tropical diseases. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, **90**, 337–339.

RAO, M., SASTRY, S. V. C., YADAR, P. D., KHAROD, K., PATHAN, S. K., DHINWA, P. S., MAJUMDAR, K. L., SAMPAT KUMAR, D., PATKAR, V. N., and PHATAK, V. K., 1991, *A Weighted Index Model for Urban Suitability Assessment: A GIS Approach* (Bombay: Bombay Metropolitan Regional Development Authority).

ROBINSON, T. P., 1998, Geographic information systems and the selection of priority areas for control of tsetse-transmitted trypanosomiasis in Africa. *Parasitology Today*, **14**, 457–461.

ROBINSON, T. P., 2000, Spatial analysis and geographic information systems in epidemiology and public health. *Advances in Parasitology*, **47**, 81–128.

ROBINSON, T. P., 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. P., 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. 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., 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*, edited by P. J. Edwards, R. M. May, N. Webb (Oxford: Blackwell Scientific Publications), pp. 247–271.

SAATY, T. L., 1977, A scaling method for priorities in hierarchical structures. *Journal of Mathematical Psychology*, **15**, 234–281.

SAATY, T. L., 1987, Rank generation, preservation, and reversal in the analytic hierarchy decision process. *Decision Sciences*, **18**, 157–177.

SANSON, R. L., PFEIFFER, D. U., and MORRIS, R. S., 1991, Geographic information systems: their application in animal disease control. *Revue Scientifique et Technique de l'Office International des Epizooties*, **10**, 179–195.

VALE, G.A., LOVEMORE, D. F., FLINT, S., and COCKBILL, G. F., 1988, Odor-baited targets to control tsetse flies, *Glossina* spp. (Diptera: Glossinidae), in Zimbabwe. *Bulletin of Entomological Research*, **78**, 31–50.

VINE, M. F., DEGNAN, D., and HANCHETTE, C., 1997, Geographic information systems: Their use in environmental epidemiologic research. *Environmental Health Perspectives*, **105**, 598–605.

VOOGD, H., 1983, *Multicriteria Evaluation for Urban and Regional Planning* (London: Pion).

WASHINO, R. K., and WOOD, B. L., 1994, Application of remote sensing to arthropod vector surveillance and control. *American Journal of Tropical Medicine and Hygiene*, **50**, 134–144.

WESTIN, F. C., and BRANDNER, T. M., 1980, Use of Landsat data in soils and agricultural use studies. In *Proceedings of the 14th International Symposium on Remote Sensing of Environment.* (Ann Abor, Michigan: ERIM), pp. 389–397.

WILLIAMS, B., ROGERS, D., STATON, G., RIPLEY, B., and BOOTH, T., 1994, Statistical modelling of georeferenced data: mapping tsetse distributions in Zimbabwe using climate and vegetation data. In *Modelling Vector-Borne and Other Parasitic Diseases*, edited by B. D. Perry and J. W. Hansen *Proceedings of a workshop organised by ILRAD in collaboration with FAO* (Nairobi: International Laboratory for Research on Animal Diseases), pp. 267–280.

ZADEH, L. A., 1965, Fuzzy sets. *Information and Control,* **8**, 338–353.