

FINAL TECHNICAL REPORT

Identifying the constraints on livestock productivity and land-use in Africa imposed by trypanosomosis

NRI Extra Mural Contract: X0239

Collaborating Institution: Department of Zoology, Oxford University, South Parks Rd., Oxford OX1 3PS

PI: Dr D.J. Rogers
Applicants: P.I. + Drs D. Bourn, G.R.W. Wint and S.E. Randolph

Personnel: Dr M.J. Packer and Mr (Dr) Simon Hay

Project Location: Department of Zoology, Oxford

NRI Liaison Officers: Mr R. Allsopp, Professor M. Gill and Dr J. Wadsworth

NRI Strategy Area: Resource Assessment and Farming Systems

Programme Area: Livestock Production

Problem Area: Livestock Protection

Date: September 1993 - September 1996

Total financial support: £284,000

EXECUTIVE SUMMARY

The aims of the TALA project were:

- 1) *to identify determinants of the local distributional limits of tsetse vectors and the diseases they transmit,*
- 2) *to identify determinants of the local abundance of vectors or prevalence of diseases within these distributional limits,*
- 3) *to establish the distribution of cattle and seasonal changes in their distributions,*
- 4) *to determine the links between the vector, disease and livestock distributions that may be used to predict the one from the other or each from other variables,*
- 5) *to define the effects of agriculturalists and patterns of land-use on the relationship between cattle and disease,*
- 6) *to describe the likely changes that might take place if the vector or disease were controlled or eradicated and*
- 7) *to assess the difficulty with which the vector/disease may be controlled in different areas.*

Project activities were focussed upon statistical analyses of the patterns of distribution and abundance of vegetation, tsetse and disease, using sets of independent map, table and satellite data as predictor variables within a discriminant analytical framework. Field data were derived through collaborative links with the Environmental Research Group Oxford (ERGO), with a number of institutes in Ghana and with Project GCP/TOG/013/BEL in Togo (Dr. Guy Hendrickx). Multi-temporal satellite data were derived from collaborative links with the ARTEMIS program, FAO, Rome, from the EU Joint Research Centre, Ispra, or from NASA's Pathfinder program, and were first subjected to temporal Fourier analysis which extracts important measures of seasonality (*i. e.* means, amplitudes and phases) from each multi-temporal data stream. The field application of Fourier processed imagery and predictions derived from such imagery was tested during several ground-truthing trips to Ghana.

Correlation of satellite data processed in various ways with ground-based meteorological observations confirms that selected information in the AVHRR and Meteosat data streams may be used as surrogates of vapour pressure deficit, rainfall and temperature normally measured at ground stations. Satellite data therefore can give a more complete and uniform coverage than is possible with the rather sparse network of meteorological stations in Africa.

A variety of satellite data layers for the whole of Africa were used to produce maps of areas of the continent showing similar seasonal characteristics, using unsupervised classification methods. The maps show many similarities with published vegetation maps of Africa, but also differences at local levels

Using the satellite-derived surrogate variables, land-cover types in Nigeria were described with accuracies ranging from 26% (for scrub) to 100% (for mangrove) (average 70.2%). Tsetse presence/absence was described with accuracies from 67% to 100% (average 82.3%) and abundance categories with accuracies of 30% to 100% (average 73%). Finally cattle trypanosomosis prevalence in Togo was described with accuracies of 78% to 83% (average 80%). In each case the descriptions of the training set data is extended to predictions of land-cover, tsetse and disease over much wider areas in West Africa. The approach adopted is also

able to define areas of high and low suitability for tsetse, the implication being that flies will be more easily eradicated from areas of low suitability that are often near the margins of their local distribution limits.

Historical archives of satellite data may be used to monitor habitat changes over time, though shorter term dry/wet periods, and the vegetation's delayed responses to these periods, make the interpretation of such changes rather difficult.

The understanding of land-surface processes that was developed from the statistical relationships explored in the project, and further analysis of the ERGO Nigerian data set, revealed that although tsetse and trypanosomosis appear to have strong seasonal effects on livestock distributions, the longer-term impact of fly suppression or removal on agricultural activity is less easily detected, although in places it appears to be very significant.

Work by the project has led to commissioned research with FAO, the Climatic Research Unit, UEA, Norwich, U.K. (as part of a WWF-funded study on the impact of Climate Change in the SADC region of Africa), the award of Darwin Initiative funding to Dr. Mike Packer and undergraduate projects on the distributions of wild vertebrates and ticks in Africa. The project outputs to date include 5 articles in independently reviewed journals, 5 consultant's reports, one D. Phil. (Simon Hay), three undergraduate student projects and presentations at 10 national and international meetings. Two articles are in press and at least 8 more have been submitted or are being prepared for submission.

0) A are interested - specifies of what
 FAO have taken up etc. What is
 impact likely to be Who in NARS
 have shown interest. How are
 the outputs going to contribute to
 the Purpose of the Prod. Syst. > All
 very nice scientific analysis but
 how do we show it was of use
 use? Frankly I'm disappointed
 as I discussed the need to talk
 to the NARS & to show updates
 with Rogers et al

BACKGROUND

African animal trypanosomosis remains one of the most serious constraints on African agriculture and rural development and affects an area on the continent of approximately 10 million km². The ubiquity of the animal trypanosomoses, however, makes it difficult to quantify their impact, and therefore the benefits to be gained by eradicating them. Historically this group of diseases has been heavily implicated in both the pattern and processes of human and cattle distributions in Africa, in trans-humance, in arable agriculture and finally in the seasonal grazing of crop residues by cattle. Currently, human population pressure on the land is forcing people towards a changing lifestyle (*e. g.* the sedentarisation of nomadic pastoralists) and, in places, into closer contact with tsetse and diseases. Both will lead to changes in land-use and in the distribution of human and natural resources. In extreme cases, land degradation may follow, though the rôle of people and cattle in this process is still disputed, as is the resilience of natural ecosystems to such pressures. From the perspective of development, what is required at present is a means of rapidly assessing the distribution of people, their cattle and the natural resource base; the pressures for change acting on each of them; and the opportunities for development that mitigate the harmful effects of habitat over-exploitation.

Development does not take place in a sociological or economic vacuum, either at the individual or societal level, so it is also necessary to include socio-economic assessments of the alternative recommendations for development. These involve not just the benefits and costs to the individual cattle owner (the basis of the assessment of many previous tsetse control projects) but also the wider costs to society of, for example, the loss of wild-life areas to cattle rearing, or the local change of habitats from woodland to grassland, to scrub and even, perhaps, to desert. Finally the opportunity costs of funds devoted to tsetse control (and, in a wider context, veterinary services) are rarely considered by either the tsetse control or veterinary care communities, but it is increasingly clear that hard-pressed Governments are no longer able to sustain the levels of veterinary services they once supported. Annual budgets are unable to keep up with annual expenditures, and what begins as a cost-cutting exercise eventually requires a more radical assessment of the aims and objectives of Government involvement in veterinary services in general, and in tsetse and trypanosomosis control in particular.

For many years, high resolution LANDSAT and SPOT images have found wide application in development project areas, often providing the only maps available for remote sites. Their high cost and relatively small coverage per scene, however, precludes their use over wider areas, as does the sheer volume of data involved. Low resolution meteorological satellite images have been used in the last 10 years as inputs to famine and drought early warning systems but, until relatively recently, were not readily available to the wider research community. Drought and famine forecasting were based on work showing how meteorological satellite data are related to plant photosynthetic activity, biomass and plant phenology (*e. g.* the 'greening up' of vegetation after seasonal rains). FAO's ARTEMIS program is centrally involved in such forecasting systems for Africa and continues to provide a welcome source of processed meteorological satellite imagery free to *bona fide* research projects. When correlations were discovered between the satellite image values and tsetse mortality rates, tsetse abundance, trypanosomosis incidence and prevalence at sites throughout

Africa (Rogers & Randolph 1991, Rogers 1991, Rogers and Williams 1993) the potential for satellite imagery to describe and monitor key features of tsetse, trypanosomosis and land-use patterns was clearly established. The present project was designed to explore this potential further, and to begin to work towards a set of guidelines by which disease situations could be rapidly assessed, control programs could be designed with more complete environmental information to hand, and the impacts of suppression and eradication programs could be monitored through time. The project's title, 'Trypanosomosis and Land-use in Africa' also gave the project an acronym - TALA - by which it has become known.

During the lifetime of this project, NASA has implemented a bold decision to re-process the entire AVHRR data archive using the most up-to-date algorithms, and to make these data free to the user community under the PATHFINDER program. This has involved the present project in a significant amount of data processing, but has also allowed us to explore and assess a variety of remotely sensed meteorological and environmental indicators.

PROJECT PURPOSE

The project falls under the Livestock Production Programme Area, within the Livestock Protection Problem Area which has identified the African Animal Trypanosomoses as one of the most important constraints on livestock productivity in sub-Saharan Africa.

The project was set a number of tasks, as follows:

- 1) to identify determinants of the local distributional limits of tsetse vectors and the diseases they transmit,
- 2) to identify determinants of the local abundance of vectors or prevalence of diseases within these distributional limits,
- 3) to establish the distribution of cattle and seasonal changes in their distributions,
- 4) to determine the links between the vector, disease and livestock distributions that may be used to predict the one from the other or each from other variables,
- 5) to define the effects of agriculturalists and patterns of land-use on the relationship between cattle and disease,
- 6) to describe the likely changes that might take place if the vector or disease were controlled or eradicated and
- 7) to assess the difficulty with which the vector/disease may be controlled in different areas.

RESEARCH ACTIVITIES

Overview and Summary

Many of the questions in the project were addressed using statistical methods that seek correlations between the data of interest (fly distributions, disease prevalence etc.) and sets of predictor variables drawn either from published tables or maps (*e. g.* meteorological data, soil fertility) or from a rapidly expanding archive of satellite data. In the latter case, research concentrated upon the use of NOAA AVHRR and Meteosat multi-temporal satellite imagery, both of which provide uniform over-views of large areas that cannot be easily obtained by other methods. Such over-views allow the extension of conclusions drawn from small study areas to much wider regions. During the lifetime of the project, we have used several sources of satellite data, each more detailed or more accurate than the previous ones. This has often entailed re-analysis of field data sets using the more precise satellite imagery. As mentioned above, the most complete and carefully calibrated AVHRR data set is available from NASA's PATHFINDER program (James and Kalluri 1994) and we now have the complete Pathfinder archive of daily data from 1981 (part) to 1993. Daily data have allowed us to calculate a variety of vegetation and thermal indices that could not be calculated from composited 10-day or monthly data (see the later Section on data processing), and we were thus able to examine the predictive power of a variety of raw satellite channels and the vegetation, vapour pressure and thermal indices that may be derived from them.

Multi-temporal satellite data were first subjected to temporal Fourier analysis that describes each data stream in terms of its mean value and the amplitude and phases (*i. e.* timing) of the annual, bi-annual, tri-annual cycles that may be used to describe it. This processing achieves the important aim of data reduction without much loss of information, with the additional bonus that the Fourier descriptions have an obvious biological interpretation (something which is not possible in alternative data reduction techniques such as principal components analysis). Fourier analysis provides information that may be directly related to land-surface phenomena (*e. g.* the amplitude of the annual cycle of the NDVI signal is correlated with the percentage cover by deciduous savannah woodland) but also partitions the variance of the signal orthogonally, so that the contribution of annual, bi-annual and tri-annual components to the overall annual variation may be easily examined. This allows the production of seasonality maps, providing important insights into the spatial arrangement of habitat types in Africa. Fourier analysis also readily lends itself to studies of environmental change over time. Each of the obvious changes in habitat type (*e. g.* gradual de-forestation as land is brought into the cultivation cycle; degradation of vegetated habitats to bare soil; re-forestation of annual grassland areas with perennial woodlands) is associated with obvious and predictable changes in Fourier components. It follows from this that judicious use of maps of the difference in key Fourier components (*e. g.* amplitudes, phases) over the time-span of the data archive should be able to reveal longer term changes in habitat use. Our initial application of this approach suggests that whilst this overall objective is probably attainable, the images are also strongly affected by shorter-term climatic cycles (dry/wet periods), the impact of which varies spatially (see also Goward and Prince 1995).

The raw and processed satellite imagery and other data derived from maps and tables were used within the project as predictor variable data-sets to describe the distribution of vegetation, tsetse and trypanosomosis in Africa. In each case a set of training data (land-cover types, fly distribution, disease prevalence) were available through collaborative links with other projects, or from published maps, and these were used to define the predictor variable values characteristic of each class or category under study. In the case of land-cover types, for example, collaborative links with the Environmental Research Group Oxford (ERGO) provided aerial survey data for Nigeria and in parts of Tchad and Mali that allowed us to define up to 10 characteristic vegetation types (and mixtures of these types). Aerial survey grid squares in which one of these 10 types predominated (*i. e.* had > 60% coverage) were used to define the satellite and other data characteristics of each vegetation type in turn, and these were then used to make predictions for Nigeria and elsewhere. Analysis and predictions were based on various modifications of discriminant analytical techniques that appear to be extremely powerful in describing biological data. As in the case of Fourier analysis, we have used discriminant techniques because they give results that are easy to interpret biologically, in preference to the more 'black box' techniques such as neural network analysis. Dominant vegetation types in Nigeria are described with varying accuracy, from 26% (for scrub) to 100% (for mangrove). Examination of the confusion matrix of this classification exercise showed that misclassifications could be easily understood. For example active cultivation was often confused with various forms of woodland, a habitat in which much cultivation actually takes place; different forms of woodland and forest were also confused.

In the case of tsetse distributions we have used mainly historical maps as our training set data (Ford and Katondo 1977, Laveissiere and Challier 1977, 1981), with additional data for Togo from our links with GCP-TOG-013-BEL. When large areas of Africa were being considered, it soon became clear that the very wide range of conditions, especially in areas of tsetse absence, required the specification of more than one area each of presence and absence. The set of predictor variables for sites of presence or absence were therefore first subjected to cluster analysis (the 'k-means cluster' option in SPSS) and subsequent discriminant analysis then dealt with a variable number of clusters for both presence and absence sites. In general, a relatively modest number of clusters (up to three or four) improved predictions considerably.

The impact of trypanosomosis on cattle distributions were explored using ERGO's Nigerian data set in which the wet/dry season numbers of cattle were recorded in approximately 20 km² grid squares throughout the country. From this analysis it is clear that fly presence continues to have an impact on cattle numbers locally in Nigeria, and that this impact depends upon the fly species involved; *G. morsitans* has a greater impact than *G. palpalis*. Whether or not fly presence actually retards agricultural development was examined by investigating the rate of agricultural expansion in parts of northern Nigeria subjected to tsetse control. This rate was determined from the percentage agricultural activity recorded by two surveys carried out in 1976 and 1990 and presented by ERGO in several reports (*e. g.* ERGO 1994). As the original reports explain, the two surveys used different methods (the first using side-looking airborne radar, or SLAR, images and the second using aerial surveys carried out by ERGO) but, with some adjustments, could be made comparable by combining SLAR categories sensibly. The comparison of the percentage of land under cultivation at these two points in time allowed the intrinsic rate of increase of agricultural activity to be calculated (assuming that this follows a logistic rise to some saturation level below 100%). These rates were then related to the

historical presence of tsetse in parts of northern Nigeria subjected to tsetse control in the years preceding the first survey. There are signs that the intrinsic rates of increase are, on average, higher in areas cleared of tsetse, though there is considerable variability in these rates throughout the region examined.

Details of Research Activity

The following sections provide details of the research activities within the project, and the theoretical background that underpins much of the analysis.

Data sets used in the study

Meteorological data sets

Climate data for the world are prepared by the World Meteorological Organization (WMO) for selected meteorological stations around the globe and these data are published each month by the National Oceanographic and Atmospheric Administration (NOAA) National Climatic Data Centre (NCDC) (NOAA 1990). Data are available as monthly summaries for 250 named meteorological stations throughout Africa and the distribution of the 207 stations used in this study is shown in Fig. 1. Meteorological stations located within 20 km of the coast or large inland lakes and rivers were excluded from the analysis because the corresponding 8 x 8 km pixel in the satellite data could have been contaminated by the signal from these water bodies.

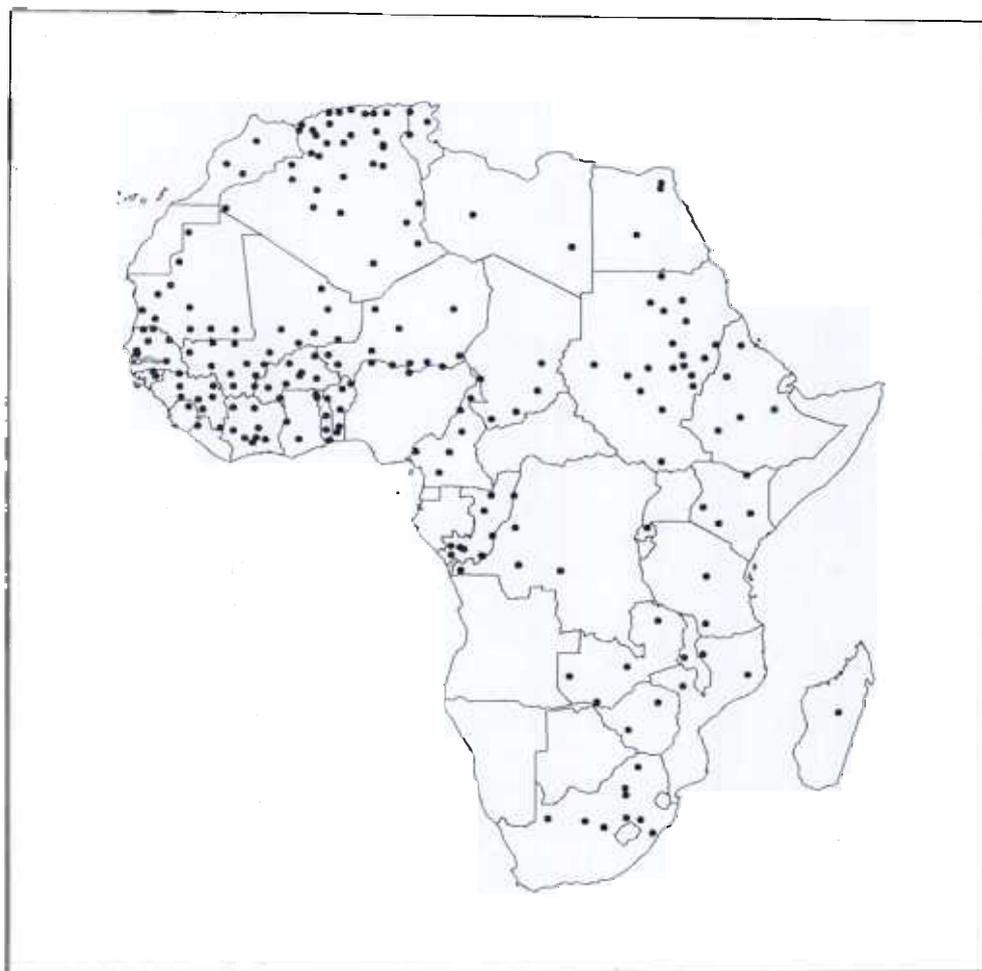


Figure 1 *The distribution of the WMO meteorological stations used in this study (from Hay 1996).*

Temperatures are recorded from standard maximum and minimum thermometers (NOAA 1972) fitted inside a Stevenson screen. All the meteorological measurements are taken at noon local time.

The NOAA - NCDC publications provide for each meteorological station the WMO standard name, country of origin, geographical location (in degrees of latitude and longitude) and elevation in meters above mean sea level. The land surface data recorded that were used in this investigation were the mean monthly temperature ($^{\circ}\text{C}$), the mean monthly vapour pressure (mb) and total monthly precipitation (mm). Humidities and vapour pressure deficits (sometimes called saturation deficits) were calculated from the mean temperature, T (K), and the mean vapour pressure, Vp (mb), using formulas provided by Unwin (1980).

Pre-PATHFINDER satellite data sets

The project began by using the standard measure of vegetation activity, the Normalised Difference Vegetation Index (NDVI) (data supplied by ARTEMIS); a measure of ground surface temperature, derived from one of the thermal infra-red channels of the same AVHRR instrument that produces the NDVI data (data supplied by the NASA Global Inventory Monitoring and Modelling Systems (GIMMS) group); and a measure of surface rainfall, the Cold Cloud Duration (CCD), derived from the METEOSAT satellite (also from ARTEMIS). For more details of the performance of each satellite series, the reader is referred to Hay *et al* (1996).

Normalised Difference Vegetation Indices (NDVIs) are derived from readings in Channels 1 and 2 (Ch_1 and Ch_2 respectively) of the Advanced Very High Resolution Radiometer on board the NOAA series of meteorological satellites and are calculated from the following formula:

$$NDVI = \frac{(Ch_2 - Ch_1)}{(Ch_2 + Ch_1)} \quad 1$$

Ten-day ('decadal') NDVI were obtained from the Food and Agriculture Organisation's (FAO) African Real Time Environmental Monitoring and Information Systems (ARTEMIS) program and their registration to a reference image was checked and, where necessary, corrected (this involved shifting images by from 0 to 3 pixels in an East-West or North-South direction, concentrating on the West African region). The raw imagery was then adjusted using calibration coefficients derived by Los (1993), which corrects for satellite sensor drift, and then maximum value composited (MVC) (by selecting the highest value of the decadal pixels for each site within each month; Holben 1986) to produce a set of monthly images for further analysis.

AVHRR channel 4 radiometric brightness temperature imagery correlates with thermodynamic air temperature at the Earth's surface (Hay *et al.*, 1996). Decadal data were generated by MVC for the period 1987-1992 from the archives of the GIMMS group at the NASA Goddard Space Flight Centre (GSFC), Maryland, USA. Monthly imagery was later produced by further MVC of the three decadal images each month.

Cold Cloud Duration (CCD) imagery was obtained from the FAO ARTEMIS program, initially as the 5-year monthly averages for the period 1988-1992, but later as monthly images. The

CCD imagery has been correlated with surface rainfall measurements as part of the Tropical Applications in Meteorology of Satellite and other data (TAMSAT) program (Snijders, 1991) over the whole of the region from which came the tsetse and land-cover data used in this project.

Digital Elevation Model (DEM) data were obtained from a 0.083 degree resolution elevation surface for Africa, produced by the Global Land Information System (GLIS) of the United States Geological Survey, Earth Resources Observation Systems (USGS, EROS) data centre. The original files were resampled to a 7.6 x 7.6 km resolution image for compatibility with the other data layers.

PATHFINDER satellite data sets

The PATHFINDER AVHRR Land (PAL) data set was obtained from the Earth Observing System Data and Information System (EOSDIS) Distributed Active Archive Centre (DAAC) at the Goddard Space Flight Center (GSFC) (Asrar and Greenstone 1995). The PAL Global Area Coverage (GAC) data were derived from visible and infrared radiance imagery collected by the afternoon ascending node satellites, namely NOAA - 7, NOAA - 9 and NOAA - 11 (Kidwell 1995) for the period July 1981 to date. The definitive description of the PAL processing chain and the data set it generates is given in Agbu and James (1994) and James and Kalluri (1994), and shown here diagrammatically in Fig. 2. For further details the reader is referred to Hay (1996).

PATHFINDER products are mapped to a global 8 x 8 km equal area grid using the Goode's Interrupted Homolosine projection (Steinwand *et al.* 1992, Steinwand 1994) and are distributed from the GSFC DAAC as global coverages containing the 12 bands of information detailed in Table 1. They are stored as scaled 8 and 16 bit binary information in Hierarchical Data Format (HDF) (NCSA 1990, Brown *et al.* 1993). The global image is 5004 pixels by 2168 lines corresponding to an image size of approximately 228 Mb (approximately 35 Mb when UNIX compressed). A separate ancillary HDF file of approximately 55 Mb, contains the latitude and longitude, elevation, and a land - sea mask.

Within the project, the HDF files were uncompressed, and the African continent was extracted from the global data.. All the daily data for 1988, 1989 and 1990 were processed, a total of $(228 \times 365 \times 3) \approx 250$ Gb of information. These years were chosen as they were average climatic years for Africa and subject to few extreme conditions, determined by reference to the NOAA meteorological data, and were prior to the Mount Pinatubo eruption of June 1991 which had significant effects on AVHRR data quality (Stowe *et al.* 1992, Jeyaseelan and Thiruvengadachari 1993). The final year (1990) also coincided with the date of ERGO's aerial survey in Nigeria.

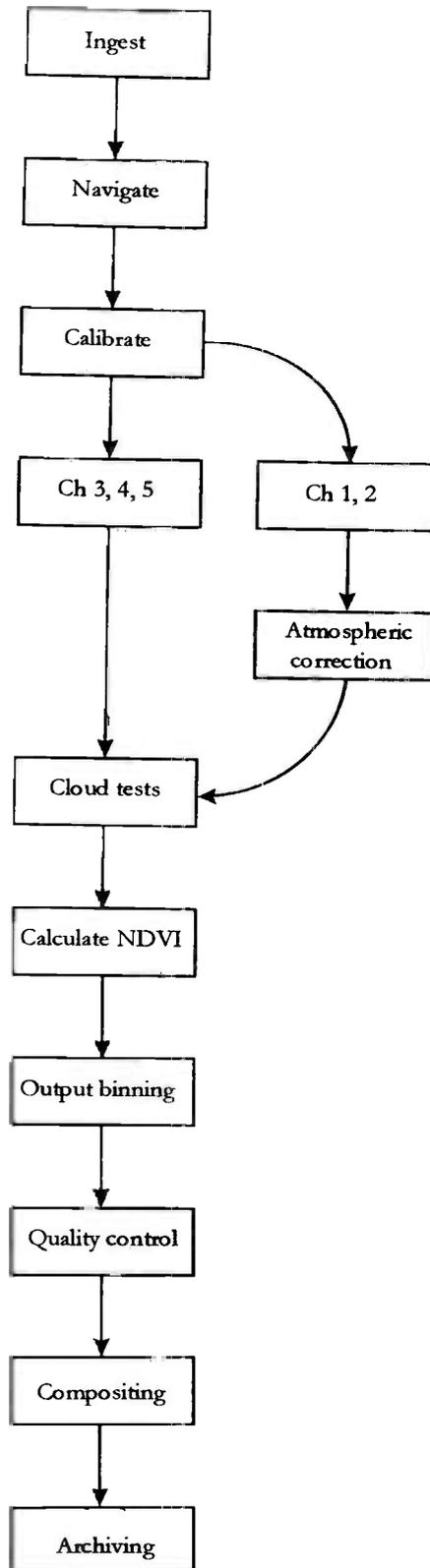


Figure 2 A diagram of the Pathfinder AVHRR Land data processing chain.

The monthly data were then sorted into two 10 day (*decade_01* and *decade_02*) and a variable third 8, 9, 10 or 11 day decade (*decade_03*) directories. There remained therefore approximately 100 files per decade (bands 6 and 12, Table 1, were discarded). Programs were written using the ERDAS Imagine 8.2. spatial modeller to convert the scaled binary PAL data into real geophysical values (with the exceptions of bands 2 and 3, which are thematic data not requiring re-scaling).

Cloudy and mixed pixels as determined from the CLAVR data layer were eliminated at this stage and such pixels given a mask value to indicate this status during later maximum value compositing.

TABLE

The information stored in the 12 bands of Pathfinder AVHRR Land (PAL) data, including scaling details.

<i>Band</i>	<i>Data layer</i>	<i>Units</i>	<i>Offset</i>	<i>Gain</i>	<i>Geophysical minimum</i>	<i>Geophysical maximum</i>
1	NDVI	ratio	128	0.008		1
2	CLAVR	thematic	1	1	0	31
3	QC Flags	thematic	1	1	0	31
4	Satellite Zenith	radians	10481.98	0.0001	-1.0472	1.0472
5	Solar Zenith	radians	10	0.0001	0	1.3963
6	Relative Azimuth	radians	10	0.0001	0	6.2832
7	Channel 1	% reflectance	10	0.002	0	100
8	Channel 2	% reflectance	10	0.002	0	100
9	Channel 3	Kelvin	-31990	0.005	160	340
10	Channel 4	Kelvin	-31990	0.005	160	340
11	Channel 5	Kelvin	-31990	0.005	160	340
12	Observation Data/Time	ddd.hh	10	0.01	001.00	366.23

Field data

Much of the field data used in the analyses presented here were gathered by the Environmental Research Group Oxford (ERGO) during aerial and ground surveys in West Africa. These data include land-cover type and cattle distributions in both the wet and dry seasons in Nigeria (RIM 1992). Low-level aerial survey flights were made over the whole of Nigeria in 1990 along gridded transect lines spaced at 20 km intervals (Bourn *et al.* 1994). The entire 924,000 km² area was surveyed between March and April 1990 (the end of the dry season) with a repeat survey covering the northern four fifths of the country between September and October 1990 (the end of the wet season). In the dry season, the percentage cover of cultivation, grassland, scrub, woodland, forest and bare ground were estimated from the air. Only active cultivation was recorded in the wet season, since the extent of other land-cover types did not change significantly. The reader is referred to ERGO's various reports for further details of the methodologies and results (RIM 1992).

Projet GCP/TOG/013/BEL in Togo has, over the last few years, been collecting comprehensive data on the distribution of tsetse and trypanosomosis throughout Togo, on a gridded basis. This data set ranks amongst the best available at the present time and provides a clear indication of the value of careful and extensive sampling over very large areas. We are especially grateful to Guy Hendrickx for his collaboration with our project.

In addition, the TALA project undertook field visits to Ghana, and gathered photographic and botanical data on land-use along the major road North, from Accra through Kumasi to Tamale. Information on the distribution of Forest Reserves in Ghana was obtained from the Forestry Department in Kumasi, together with copies of a complete set of LANDSAT digital data covering most of the country.

Vegetation data

Naturally vegetated, non-vegetated and human-modified landscapes of a region together constitute the land-cover of that area. Many factors and processes, natural and anthropogenic, determine the spatial pattern of land-cover and how it varies through time. An understanding of these processes may be used in studies of the distribution and abundance of animal populations, which are intimately associated with particular types of land-cover, and for resource management, planning and global change studies.

In order to be useful, land-cover/resource maps must provide reliable and precise descriptions of the land-cover of a given area at a given time. They must also be able to provide such information over large areas and frequently enough that changes in cover are effectively detected.

The development of satellite remote sensing techniques for land-cover mapping across large areas is currently hindered by a lack of suitable sample data (*i. e.* ground truth data) to use in training some classifications (*i. e.* "supervised classifications") or testing others (*i. e.* "unsupervised classifications"). Such ground-truth data are very costly to acquire, making large area surveys prohibitively expensive. The best examples we have to date include ERGO's aerial surveys of the whole of Nigeria, and parts of Tchad, Mali and elsewhere.

Despite the lack of similar data for much larger areas, it is nevertheless possible to use unsupervised classification methods with meteorological satellite data to produce whole-Africa land-cover maps that can be tested either on the basis of our collective experience in different parts of Africa, or of published vegetation maps (White 1983) or of limited ground-truth data sets gathered for this or other purposes (Millington 1995).

Previous attempts to produce continental-scale land-cover maps (*e. g.* Tucker *et al.*, 1985; Townshend *et al.*, 1987) have used time-series of normalised-difference vegetation index data. Recent work by Eric Lambin and Danielle Ehrlich (Lambin & Ehrlich, 1995; Ehrlich & Lambin, 1996) has examined the use of NDVI data alone or in combination with surface (Price 1983 split-window) temperature as a ratio ($T_s/NDVI$).

Short time series of data lack the information needed to define meaningful classes of land-cover based on the seasonal evolution of the vegetation signal from each cover type, whereas long time series' of data inevitably integrate land-cover phenomena over long periods of time during which changes in land-cover may have taken place. Ehrlich & Lambin (1996) concluded that for short time-series of data, $T_s/NDVI$ is the most useful variable since it is not as influenced by interannual climatic variation as are NDVI data alone. For long time-series, however, they found no detectable difference between the classifications produced using the ratio data or NDVI alone.

There seems to have been little attempt to use the other data, or derived products, available from the NOAA-AVHRR in land-cover classification, despite the fact that middle infra-red data appears to be potentially useful as it contains information relevant to the biophysical properties of vegetation (Boyd *et al.*, 1996, Foody *et al.* 1996).

In the project, White's map (White 1983) was taken as the 'standard', with which unsupervised classifications were compared. A digitised version of this map was obtained from NRI. Selected areas from White's map were also taken as the training sites for supervised classifications.

Assorted GIS data layers

The project also acquired data in digital or paper map formats covering a variety of subjects and areas and from a number of collaborations. The Ford and Katondo (1973) tsetse distribution maps were obtained in digital format from ILRAD, and the gridded maps for parts of West Africa (Laveissiere and Challier 1977,1981; Anon 1982) were entered into MS Excel spreadsheets for further analysis. The FAO soil map for Africa (FAO/UNESCO 1977) and its recent classification into a soil fertility map were obtained from FAO, whilst the project digitised a geological map for West Africa. Finally, assorted digital layers were produced for Togo (a vegetation map) and Ghana (a road/rail map).

Data layers produced by the project from the Pathfinder data set

Vegetation Indices

Apart from the NDVI, one of the data layers available within the Pathfinder product, the project computed a number of alternative vegetation indices suggested by the literature. The most simple is the ratio of channel 2 over channel 1 reflectances, called the Ratio Vegetation Index (RVI) or Simple Ratio Index (SRI). Other alternatives have been suggested to overcome some of these problems with NDVI (Jackson and Huete 1991), but all have been less widely applied to ecological and epidemiological problems, because they have not been shown to be generally applicable.

The Soil Adjusted Vegetation Index, or SAVI, effectively adjusts the intercept of the relationship between channel 2 and channel 1 data to minimise interference from the soil background;

$$SAVI = \frac{Ch_2 - Ch_1}{Ch_2 + Ch_1 + L} (1 + L) \quad ..2$$

where, L , is a weighting parameter that varies with vegetation coverage. The values of L recommended for sparse, intermediate and densely vegetated conditions are 1.0, 0.75 and 0.25 respectively (Huete 1988).

More recently, the Global Environment Monitoring Index (GEMI) has been proposed by Pinty and Verstraete (1992), again with the intention of reducing the variability introduced by the soil background and, in addition, of reducing atmospheric effects. Soil effects are theoretically minimised because the GEMI gives a more constant index of vegetation activity against a much wider range of soil conditions than does the NDVI. The GEMI was derived from first principles (Verstraete 1995, pers. comm.), rather than empirically, although the physical basis for the index is not fully explained in the literature, and it is defined as follows;

$$GEMI = \eta(1 - 0.25\eta) - \frac{Cb_1 - 0.125}{1 - Cb_1} \quad ..3$$

where

$$\eta = \frac{2(Cb_2^2 - Cb_1^2) + 1.5Cb_2 + 0.5Cb_1}{Cb_2 + Cb_1 + 0.5}$$

Initial application of GEMI to AVHRR data for Africa suggests a three-fold advantage over the NDVI. Firstly, the GEMI was found to be less sensitive to atmospheric variations. Secondly, it had a much enhanced ability to detect clouds. Finally, the GEMI had a higher dynamic range in sparsely vegetated, xeric environments, revealing details of features such as geological formations and land surface topology that are not visible in other imagery (Flasse and Verstraete 1994).

Within the project, the RVI, the SAVI, with the L parameter set to 0.5 for intermediate vegetation coverage, and the GEMI were calculated from the visible channel data. Vegetation Indices that required the calculation of a soil line were not included in the analysis because of the problem of obtaining sufficient 8 x 8 km pixels of bare soil over all areas of Africa.

Land surface temperature indices

The theoretical basis of determining land surface temperatures from split-window techniques is explained in Hay *et al* (1996). These techniques often rely on ancillary data to quantify atmospheric water content and surface emissivity, but any requiring other than AVHRR data were excluded from further consideration. When emissivity, ε , was a required parameter in the split-window equation it was calculated from a logarithmic relationship determined between NDVI and emissivity in Botswana by Van de Griend and Owe (1993) where;

$$\varepsilon = (A + B) \times \ln(NDVI) \quad .4$$

and $A = 1.0094$ and $B = 0.047$. The relationship was significant with a correlation coefficient of 0.941 ($r^2 = 0.89$) at the 0.01 level. Emissivity was assumed to be equal in the channel 4 and 5 wavebands, and this assumption is also made in all the applications of split window methods described below.

Of the many split-window algorithms available (Becker and Li 1995), three are suitable for extensive coverages. Price (1984) derived one of the first indices to estimate land surface temperature, T (in Kelvin), from the AVHRR channel 4, Ch_4 (K), and the AVHRR channel 5, Ch_5 (K) brightness temperatures that accounted for the emissivity of the land surface where;

$$T = Ch_4 + A(Ch_5 - Ch_4) \quad ..5$$

and, A , is a constant determined by Price to be 3.33 for channels 4 and 5 of the NOAA - 7 AVHRR. This equation has an accuracy that varies between about $\pm 2 - 3$ K and ± 4.5 K, depending on habitat type and image resolution (Cooper and Asrar 1989, Sugita and Brutsaert 1993), and is being used in the creation of global AVHRR datasets for the European community (Malingreau and Belward 1994).

A second index of land surface temperature, from Becker and Li (1990) has found preliminary application to multitemporal land surface temperature estimation of Asia (Gutman 1993). It has also been used to provide the land surface temperature parameter globally for inputs into primary production models (Prince and Goward 1995).

$$T = Ch_4 + \delta w + \delta \varepsilon$$

with

$$\delta w = 2.63((Ch_4 - Ch_5) + 1.274) \quad ..6$$

$$\delta \varepsilon = \frac{Ch_4 - Ch_5}{2} (0.156 + (3.98 \left(\frac{Ch_4 - Ch_5}{Ch_4 + Ch_5} \right))) \frac{1 - \varepsilon}{\varepsilon}$$

and

$$\varepsilon = 1.013 + 0.681 \ln(NDVI)$$

where δw is a correction for water vapour attenuation and $\delta \varepsilon$ is a correction for emissivity, ε . In the calculation of this temperature index, emissivity was estimated from the NDVI using the Van de Griend and Owe (1993) relationship described above, because it was derived for a study area more representative of the African continent than that used by Becker and Li.

Finally, the Prata and Platt (1991) estimation of land surface temperature has been shown to perform well in comparison with the above indices (Prata 1993), but has not been widely applied to regional scale data. This formulation attempts to isolate atmospheric effects by using atmospheric temperature and moisture profiles from TIROS Operational Vertical Sounder (TOVS) so that the algorithm can be applied globally, provided the emissivity of the terrain is known (Prata 1993).

$$T = 273.15 + (3.45 \times \frac{Ch_4 - 273.15}{\varepsilon_4}) - (2.45 \times \frac{Ch_5 - 273.15}{\varepsilon_5}) + (40 \times \frac{1 - \varepsilon_4}{\varepsilon_5}) \quad ..7$$

Atmospheric moisture indices

The total precipitable water content of the atmospheric column, U , is calculated using an equation given by Dalu (1986). The total precipitable water content of the atmospheric column has been estimated according to a method proposed by Dalu (1986). Similar to the split-window algorithm, this method exploits the difference in atmospheric attenuation due to atmospheric water vapour between channels 4 and 5 of the AVHRR. The algorithm was derived from atmospheric radiative transfer models over the ocean, where a surface relative humidity of 80 % was assumed due to the natural equilibrium between evaporation and diffusion and tested against measurements taken from ships. Based on a derived correction factor, a , and taking into account the changing atmospheric path length as a function of scan angle, θ , the total precipitable water content of the atmospheric column, U (kgm^{-2}), can be estimated as follows;

$$U = a \times (Ch_4 - Ch_5) \times \cos \theta \quad .8$$

The retrieved estimates were stated to have an accuracy of $\pm 5 \text{ kgm}^{-2}$ over the ocean. The accuracy of these estimates over the land surface will be influenced by varying emissivity, as well as by deviation from the assumption of 80 % relative humidity at the surface. Justice *et al.* (1991) however, have noted good agreement between values for atmospheric water content estimated using the above equation and those measured by photometers at several sites in the Sahel.

An alternative formulation was given by Eck and Holben (1994) where;

$$U = A + B(Ch_4 - Ch_5) \quad .9$$

and, A and B are constants, 1.337 and 0.837 respectively. (These were determined by a linear regression of (Ch4 - Ch5) against estimated precipitable water content of the atmospheric column using radiosonde data from the Gao meteorological station in Mali. The coefficient of determination for the relationship was 0.96.)

The total precipitable water content of the atmospheric column (U) is often expressed as kgm^2 , *i. e.* units of pressure (mass per unit area), and are converted to the amount of water that would be precipitated from the atmospheric column in cm by dividing by 10, since the density of water is 1 gcm^{-3} .

The estimated precipitable water content, U (cm), may be converted to a near surface dew point temperature, T_d ($^{\circ}\text{F}$) using the following relationship (Smith 1966);

$$T_d = \frac{\ln P_w - ((0.113 - \ln(\lambda + 1)))}{0.0393} \quad 10$$

where, λ , is a variable that is a function of the latitude and the time of the year. In this work a mean value of $\lambda = 2.99$ was calculated from the annual mean λ presented by Smith (1966) for locations between 0 and 40 degrees of latitude. It was decided not to use a seasonally and latitudinally adjusted figure as the λ values were derived specifically for data throughout the northern hemisphere (Smith 1966).

The dew point temperatures in Kelvin are finally used with the Price (1984) estimate of land surface temperature, T_p (K), to calculate the vapour pressure deficit, Vpd (KPa), using the equation provided in Prince and Goward (1995) where;

$$Vpd = 0.611 \left[\exp\left(17.27 \times \frac{T_p - 273}{T_p - 36}\right) - \exp\left(17.27 \times \frac{T_d - 273}{T_d - 36}\right) \right] \quad ..11$$

Final temporal compositing of the Pathfinder and derived data layers

The fifteen bands of satellite data used in the project, including the original AVHRR channel information and derived indices, were maximum value composited using the data values within a specific band and without reference to the NDVI. The mask values were again preserved and outputs saved to floating point accuracy.

Data Analysis

(readers uninterested in the mechanics of data analysis should skip this section and go straight to OUTPUTS)

Data reduction through temporal Fourier analysis

Data reduction of multi-temporal satellite imagery is usually achieved through principal components analysis (Townshend and Justice 1985, Townshend *et al* 1985, Eastman and Fulk 1993, Eklundh and Singh 1993). An entirely novel approach to the same problem was suggested by the literature on time series analysis (*e. g.* Chatfield, 1980). The time series {xt} may be described by a Fourier series representation where

$$x_t = a_0 + \sum_{p=1}^{N/2-1} [a_p \cos(2\pi pt / N) + b_p \sin(2\pi pt / N)] + a_{N/2} \cos \pi t, \quad (t = 1, 2, \dots, N) \quad ..12$$

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

$$a_0 = \bar{x}$$

$$a_{N/2} = \sum (-1)^t x_t / N \quad ..13$$

$$a_p = 2[\sum x_t \cos(2\pi pt / N)] / N \quad \left. \vphantom{a_p} \right\} p = 1, \dots, N/2 - 1$$

$$b_p = 2[\sum x_t \sin(2\pi pt / N)] / N$$

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

$$a_p \cos \omega_p t + b_p \sin \omega_p t = R_p \cos(\omega_p t + \phi_p) \quad ..14$$

where

R_p = the amplitude of the p th harmonic

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

and ϕ_p = the phase of the p th harmonic

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

(Chatfield 1980). The effect of Fourier analysis is to partition the variability of the time series into (orthogonal *i. e.* uncorrelated) components at frequencies of $2\pi/N, 4\pi/N, 6\pi/N, \dots, \pi$, or periods equal to $1, 1/2, 1/3 \dots 2/N$ times the duration of the observations, N . If monthly

observations are taken, Fourier analysis can partition the time series into frequencies equivalent to periods ranging from as long as the whole time series, down to two months (higher frequencies, *i. e.* shorter period cycles, cannot be distinguished by monthly data). Full Fourier analysis exactly describes the original data set (since the Fourier series in equation 2 contains N parameters to describe N observations), but not all harmonics may be contributing equally to this description. The following relationship, known as Parseval's theorem, applies to the Fourier representation of $\{x_t\}$

$$\sum (x_t - \bar{x})^2 / N = \sum_{p=1}^{N/2-1} R_p^2 / 2 + a^2_{N/2} \quad 15$$

This equation states that a quantity very similar to the variance of the original observations (the left-hand side of the equation, but with the divisor N rather than $(N-1)$) is the sum of the contributions of each of the $p = 1$ to $N/2$ harmonics, where $R_p^2/2$ is the contribution of the p th harmonic.

The combination of the orthogonality of the harmonics in the Fourier series representation of satellite data and the (perhaps illusory) biological transparency of the interpretation of these harmonics makes this approach to data reduction especially attractive to biologists (Rogers and Williams, 1994). In effect it may be possible to reduce a 10-year or longer monthly or decadal data stream to just 7 variables (the mean of the whole series and the amplitude and phases of the first three Fourier components) without a great loss of information.

For the present study the monthly satellite data channels were subjected to temporal Fourier processing and the means and amplitudes and phases of the annual, bi-annual and tri-annual cycles were calculated and stored as new image layers for analysis, at the same spatial scales as the original imagery. During the analysis the combined (*i. e.* annual + bi-annual + tri-annual cycle) Fourier description of the original signal was also calculated and its minimum, maximum and range (*i. e.* maximum - minimum) were recorded for use in the analysis. This summation essentially smooths the original data set. In addition certain combinations of the Fourier-processed signals were calculated, such as the ratio of NDVI to thermal mean values, which has been shown to be a more stable indicator of vegetation type than either variable alone (Lambin and Strahler, 1994).

Statistical treatment using discriminant analysis

The reduced-dimension data set produced by the methods outlined above form the set of predictor variables for describing land-cover types, species' distributions and abundance and disease prevalences. Here we take the simple problem of describing the distribution of a vector to illustrate the techniques of discriminant analysis.

In its simplest form, discriminant analysis assumes a multi-variate normal distribution of the predictor variables and a common within-group co-variance of the variables for all points defining vector presence and vector absence. The mean values of the predictor variables in sites of vector presence and absence, and the within-group co-variance matrix, are estimated from representative samples from reliable distribution maps, the 'training sets'. Means of multi-

variate distributions are referred to as centroids and are defined by mathematical vectors [\bar{x}_n] where n is the number of dimensions (= variables). The Mahalanobis distance, D^2 , is the distance between two multi-variate distribution centroids, or between a sample point and a centroid, and is a generalisation of the traditional squared Euclidean distance d^2 , where

$$d^2_{12} = (\bar{x}_1 - \bar{x}_2)' (\bar{x}_1 - \bar{x}_2) \\ = \mathbf{d}' \mathbf{d}$$

and

.16

$$D^2_{12} = (\bar{x}_1 - \bar{x}_2)' C_w^{-1} (\bar{x}_1 - \bar{x}_2) \\ = \mathbf{d}' C_w^{-1} \mathbf{d}$$

where d^2_{12} and D^2_{12} are the Euclidean and Mahalanobis distances between groups 1 (*e. g.* for vector absence) and 2 (*e. g.* for vector presence), $\mathbf{d} = (\bar{x}_1 - \bar{x}_2)$ with subscripts again referring to the two groups (alternatively 1 and 2 might refer to a point and a centroid) and C_w^{-1} is the inverse of the within-groups covariance (= dispersion) matrix (Green 1978) (in equation 6 the subscript n for the number of variables has been dropped for clarity). Thus the Mahalanobis distance is the distance between the sample centroids adjusted for their common co-variance. In the case of the Euclidean distance, d^2_{12} , the covariances are zero, so that the covariance matrix $C_w = C_w^{-1} = \mathbf{I}$, the identity matrix (with ones along the diagonal, and zeroes elsewhere). This reduces the equation for D^2_{12} to that for d^2_{12} in 16. If the problem is to predict only to which of the groups of 'presence' or 'absence' a new point belongs it is simply necessary to calculate the two values of D^2 between the point and each of the two centroids. The point is then assigned to the group which gives the smaller D^2 (*i. e.* to which it is closer in multi-variate space). This assignment rule is obviously an over-simplification since the values of D^2 may differ by only a little, or by a very large amount. There is always a probability, however slight, that the observation in fact belongs to the group to which it was not assigned.

The 'posterior probability' replaces the simple prediction of group membership by calculating the probability with which any observation belongs to each group as follows

$$P(1|x) = \frac{p_1 e^{-D^2_{12}/2}}{\sum_{g=1}^2 p_g e^{-D^2_g/2}}$$

and

..17

$$P(2|x) = \frac{p_2 e^{-D^2_{22}/2}}{\sum_{g=1}^2 p_g e^{-D^2_g/2}}$$

where $P(1|x)$ is the posterior probability that observation x belongs to group 1 and $P(2|x)$ the posterior probability that it belongs to group 2 (Green, 1978). p_1 and p_2 are the prior probabilities of belonging to the same two groups respectively, defined as the probabilities with which any observation might belong to either group, given prior knowledge or experience of the situation (often, when applied, based on the training set data). In the absence of any prior

experience it is usual to assume equal prior probability of belonging to any of the groups. Where there are only two groups, for absence and presence, $p_1 = p_2 = 0.5$. Equation 17 assumes that observation \mathbf{x} must come from either group 1 or group 2: the possibility it belongs to neither is discounted. Once again the assumption in equation 17 is of multi-variate normality, the other terms of the multi-variate normal equation cancelling out (Tatsuoka, 1971).

The above formulae apply only to those situations in which a common co-variance matrix can be assumed. In many cases of distribution data, however, this does not apply because animals do not live within a random subset of environmental space, but within a rather unusual subset, with particular environmental conditions which cannot be described by general environmental conditions. The result is that the co-variances of the variables within a distributional range are often different from those of the same variables outside the distributional limits. This requires a modification of equations 16 and 17, to allow for different within-group co-variance matrices. Equation 17 is then modified as follows

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

and

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

..18

where $|\mathbf{C}_1|$ and $|\mathbf{C}_2|$ are the determinants of the co-variance matrices for groups $g = 1$ and 2 respectively (the Mahalanobis distances in 18, calculated from 16, are now evaluated using the separate within-group co-variance matrices) (Tatsuoka 1971). With unequal co-variance matrices the discriminant axis (strictly speaking a plane) that separates the two groups in multi-variate space is no longer linear.

In the analyses reported here a variable proportion of the map data were used as the training sets. When data were scarce, all observations were included in the data set; when data were more abundant, a limit was placed on sample sizes, related to the overall area covered by the analysis. Predictor satellite variables were selected in a forward, step-wise manner, the criterion for inclusion being that the addition of the selected variable caused the greatest increase in the Mahalanobis distance (equation 16) compared with all other variables during that 'round' (since unequal co-variance matrices were assumed in the analysis, the Mahalanobis distance calculated for each comparison was the sum of the distance between the presence and absence category and between the absence and presence category). Thus variables were selected in order of their ability to separate the different groups either of presence/absence or of density classes. A limit of 10 predictor variables was set in all analyses, as experience showed this was sufficient to give good predictions. Generally these were then used to generate maps of posterior probabilities (equation 18) which represent the probabilities with which each grid square falls into the category of the predicted variable under study.

No transformation of the raw variables was undertaken before analysis, to make biological interpretation of the results easier. The method of variable selection, using Mahalanobis distances, overcomes the potential effect of unequal co-variances arbitrarily determining the importance of the predictor variables.

It is relatively straightforward to extend equations 16 to 18 to situations in which more than two groups (of land-cover types, or of species' absence/presence) are encountered. For example, vector abundance data may be 'binned' into more than two groups, each bin defining a range of vector densities. When distributional data are from wide geographical areas there may be regional variations that need to be represented by different co-variance matrices in different areas. There may also be different, sub-specific or strain variation responses of the vectors to environmental conditions in the different regions, again requiring different covariance matrices defining fly presence in the different areas. The statistical significance of any differences found may be tested using Bartlett's chi-square approximation (B) for testing co-variance matrix equality (Green 1978) and defined as follows:

$$B = (m - G)\ln|C_w| - \sum_{g=1}^G (m_g - 1)\ln|C_g| \quad ..19$$

where m is the total number of observations of all groups ($m = m_1 + m_2 + \dots + m_G$) and G is the total number of groups (2 in the simple case of presence/absence). B is approximately distributed as χ^2 with $\frac{1}{2}[(G-1)(n)(n+1)]$ degrees of freedom, where n is the number of variables contributing to the covariance matrices. $|C_w|$ and $|C_g|$ respectively refer to the determinants of the within-groups co-variance matrix of all groups combined or of each group, g , separately. *A priori* the best approach to analysing multiple data sets from large areas is to keep them separate initially and then to combine co-variance matrices appropriately only when they can be shown not to differ significantly. In practice, however, this may result in rather small sample sizes giving unreliable co-variance matrices (Lark, 1994). In the present project both approaches were tested. Predictions of the distribution of most of the tsetse species in West Africa were improved by keeping different surveys separate, but others (specifically for *G. palpalis*) were made worse. The whole Africa predictions of tsetse distributions were always improved when predictor variable data sets for sites of presence and absence were first divided into clusters before analysis. One particular problem with the algorithm used (SPSS's k means clustering) is that it divides data on the basis of the original data scales (*i. e.* the data are not standardised before clustering). Since elevation covers a much wider range of large numbers than does NDVI, the cluster assignments tended to be dominated by site elevation. This, possibly artificial, division at the early stage often resulted in elevation appearing high on the list of key predictor variables within the discriminant analyses, even though the choice of variables at this stage (based on the Mahalanobis Distances) is not affected by the ordinal scales used (since Mahalanobis distances are covariance adjusted).

Measures of Predictive Accuracy

The ability of the technique to describe the observed distribution and abundance or prevalence data was measured in several ways. The overall percentage correct predictions (of presence/absence, or of abundance/prevalence class) were calculated together with the percentages of false positive and false negative predictions (*i. e.* false predictions of presence or absence respectively). Finally the sensitivity (ability to predict presence correctly) and specificity (ability to predict absence correctly) were also calculated. In the case of the abundance/prevalence data the percentage correct assignment to each density class was recorded.

One of the attractions of the use of posterior probabilities (quantities that are easily calculated during maximum likelihood assignments, but rarely made available to the user) is that both the probabilities of correct assignments and those of all incorrect assignments are calculated, so that the 'certainty' with which the method makes any particular assignment can be examined. This was especially useful when examining misclassifications in the land-cover data set for Nigeria. Discriminant analysis was on average more definite about its correct assignments (*i. e.* these were made with a higher average probability) than with its incorrect assignments. Posterior probabilities also allowed us to check whether fuzzy logic rules could be used to identify mixed pixels - or 'mixels' - correctly. It has been claimed in the literature that when training sets are based on pure end members (*e. g.* 100% grassland or 100% forest) the posterior probabilities assigned to mixels represent the actual proportions of the different land cover types they contain (Foody 1992). For example in a 2-class end member case, *e. g.* of pure grassland and pure forest, it is claimed that a posterior probability of 0.3 for grassland and 0.7 for forest means that the mixel contained 30% grassland and 70% forest. Whilst such fuzzy logic appears to be useful in the examples given in the literature, we suspect that these are based on rather impoverished natural environments with few distinct end member types. In Africa, not only are there many different vegetation types, but these are often mixed together in all possible proportions. The likelihood that such fuzzy logic rules will give reliable assignments at the individual mixel level therefore seems small (this was confirmed in the Nigerian data set; whilst many correlations between the predicted posterior probabilities and the proportion of each grid square covered by different vegetation types were significant, because of large sample sizes, the residual variances were still very large).

OUTPUTS

The relationship between satellite data and meteorological data

Temperature

In all cases, the slopes of the twelve monthly regressions of satellite data values on appropriate meteorological variables for each of the years 1988, 1989 and 1990 were too variable for the data to be pooled. Instead, the mean and the standard error of the coefficient of determination (r^2) values for each monthly comparison over the 1988 - 1990 period were calculated to facilitate comparison between the data. There are positive correlations in every month ($P \ll 0.0001$, mean $n = 120$ with range of 105 to 141 observations). As an example, the data for 1990 are shown in Table 2. Non-linear alternatives to a linear regression fit did not increase coefficients of determination for the monthly comparisons.

The mean value of r^2 for the correlations between surface temperature and AVHRR channels 3, 4 and 5 over the three years were relatively low at 0.31, 0.38 and 0.32. The split-window corrections explained more of the variance in the air temperature data however, with mean r^2 values of 0.54 for Price (1984), 0.43 for Becker and Li (1990) and 0.38 for Prata (1991) formulations. This trend in the performance of the split-window equations was followed without exception in the monthly comparisons. Furthermore, the standard error (an indication of the variation in accuracy over the three years) was least for the Price (1984) equation. Examples of the best and worse fit to the ground temperature data are shown in Fig. 3.

The deviations from the regression between the Price (1984) equation and the NOAA - NCDC land surface temperature for December 1990 were plotted on a map of Africa (Fig. 4). Distinct geographical patterns were revealed in these residual values and were found to be related to altitude. Positive deviations from the regression line occurred largely in those stations above 500 m and negative deviations occurred in those below. When elevation data from the DEM were included as an additional predictor variable to the Price (1984) data in monthly multiple regression equations, the r^2 increased from a mean of 0.54 to 0.70 for the three years. This effect can be seen for each of the monthly comparisons in Fig. 5 which also reveals that there were no seasonal trends in the goodness of fit. The scatter of points in the monthly comparisons was not similarly reduced by including the NDVI in the regression equation.

TABLE 2.

Examples of the coefficients of determination obtained when the monthly maximum thermal channel and split-window corrected brightness temperatures are linearly regressed with the monthly mean NOAA - NCDC air temperatures for 1990 (from Hay 1996).

	<i>Channel 3</i>	<i>Channel 4</i>	<i>Channel 5</i>	<i>Price (1984)</i>	<i>Becker and Li (1990)</i>	<i>Prata and Platt (1991)</i>	<i>Price84 and elevation</i>
<i>Jan</i>	0.08†	0.17	0.12	0.36	0.23	0.17	0.47
<i>Feb</i>	0.07†	0.19	0.13	0.47	0.29	0.22	0.56
<i>Mar</i>	0.20	0.35	0.32	0.49	0.41	0.38	0.59
<i>Apr</i>	0.28	0.46	0.41	0.59	0.57	0.54	0.65
<i>May</i>	0.31	0.40	0.35	0.58	0.51	0.46	0.69
<i>Jun</i>	0.52	0.51	0.43	0.68	0.59	0.56	0.77
<i>Jul</i>	0.54	0.33	0.22	0.62	0.50	0.47	0.81
<i>Aug</i>	0.27	0.27	0.21	0.47	0.37	0.33	0.64
<i>Sep</i>	0.20	0.26	0.23	0.38	0.30	0.29	0.68
<i>Oct</i>	0.10	0.19	0.17	0.28	0.19	0.16	0.69
<i>Nov</i>	0.36	0.51	0.46	0.61	0.44	0.37	0.71
<i>Dec</i>	0.48	0.66	0.61	0.74	0.58	0.54	0.83
Mean	0.28	0.36	0.31	0.52	0.41	0.37	0.67
SE	0.05	0.04	0.04	0.03	0.04	0.04	0.03

† Not significant at the 0.01 level.

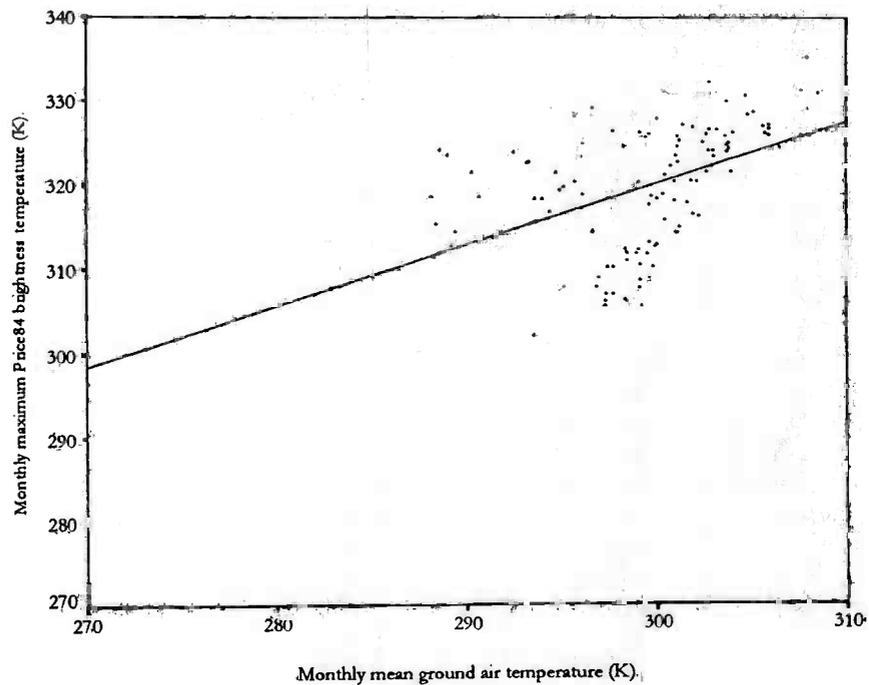
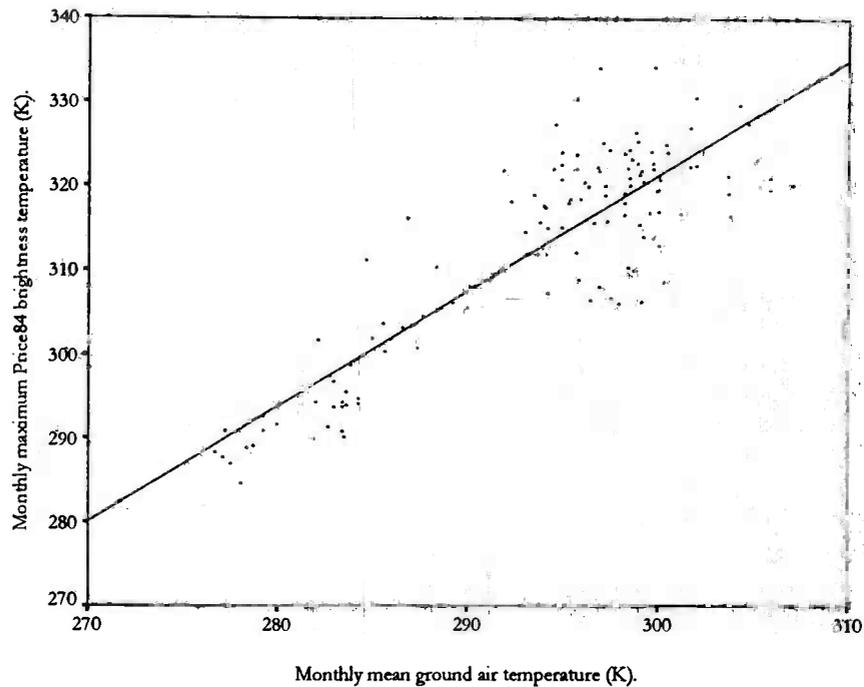


Figure 3 The comparison of mean monthly NOAA - NCDC air temperature with the maximum monthly Price (1984) split-window corrected brightness temperature. Examples are given of the best fit to the ground data (December 1990, $n = 124$, $P \ll 0.0001$, $r^2 = 0.74$, top graph) and the worst fit (September 1989, $n = 117$, $P \ll 0.0001$, $r^2 = 0.19$, lower graph) (from Hay 1996).

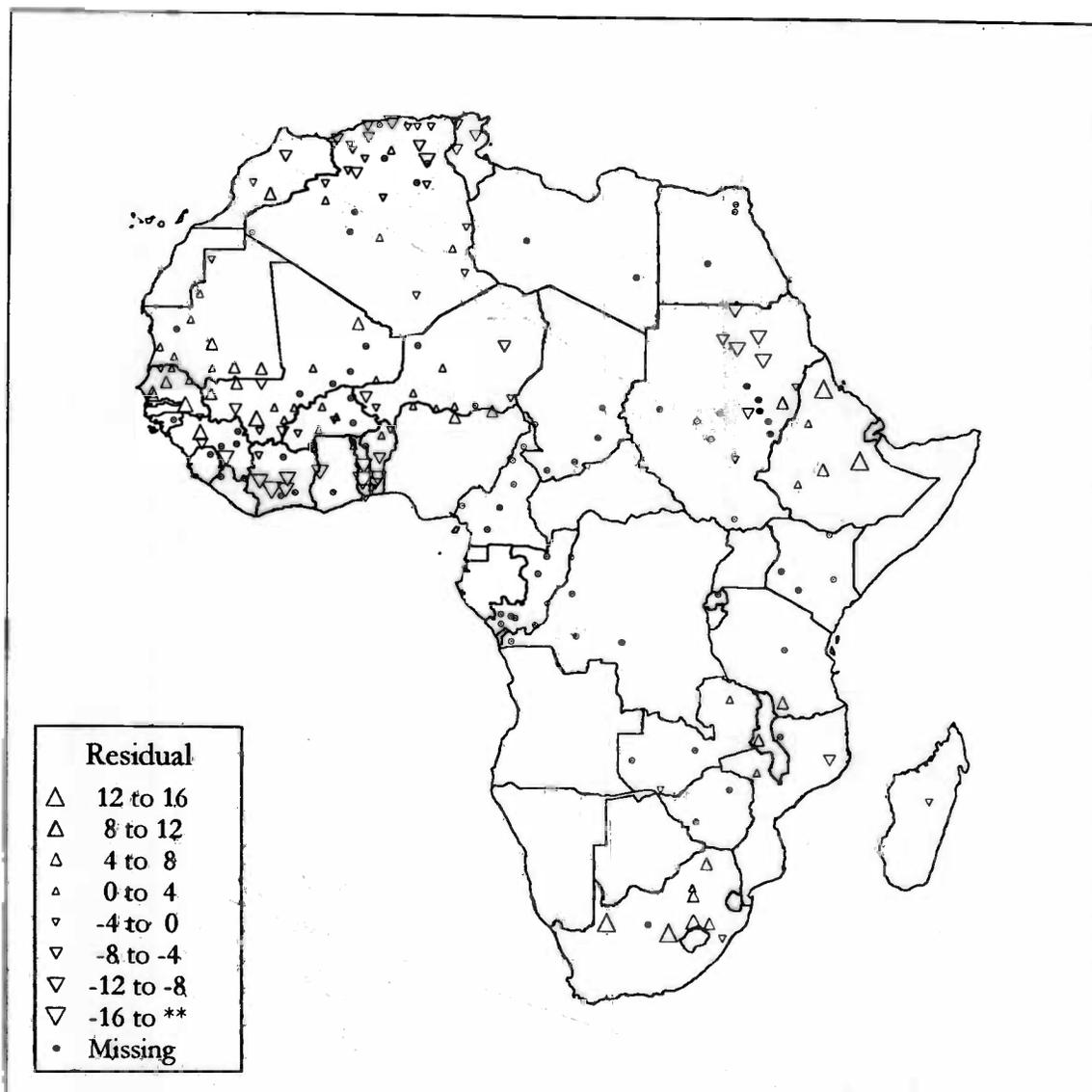


Figure 4 The distribution of the residual values in the regression of Price (1984) corrected split-window corrected brightness temperature and the mean NOAA - NCDC air temperature for December 1990. Stations for which data were missing are also indicated (from Hay 1996).

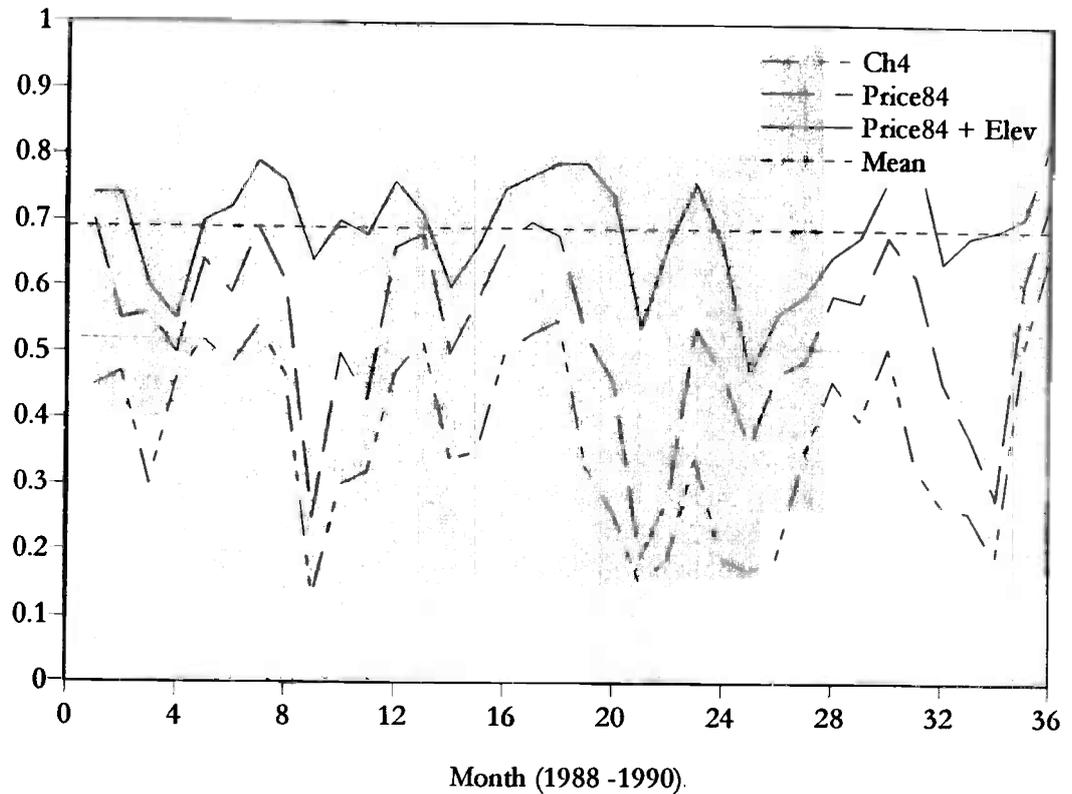


Figure 5 The adjusted r^2 of the monthly comparison of NOAA - NCDC air temperature and Price (1984) split-window corrected brightness temperatures for the period 1988 - 1990. The graph shows the improvement of fit of the Price (1984) equation over the AVHRR channel 4 data and the further improvement when elevation is used as a term in the regression equation. The mean r^2 (i. e. the proportion of the total variance explained) for the Price (1984) and elevation combination over the 36 months is also shown by the dotted horizontal line (from Hay 1996).

Atmospheric moisture

Examples of the comparison of mean monthly NOAA - NCDC saturation deficit with the maximum monthly vapour pressure deficit (VPD) derived from the satellite data are presented for the year 1990 in Table 3. The VPDs were calculated using the Dalu (1986) (VPD₁) and the Eck and Holben (1994) (VPD₂) procedures outlined previously. A highly significant positive linear relationship was observed in each month ($P \ll 0.0001$, mean $n = 121$ with range of 106 to 140 observations), except where † indicates significance below the 0.01 level. Examples of the best and worse fit to the saturation deficit data are shown in Fig. 6.

The mean r^2 for the three year period using VPD₁ and VPD₂ were 0.63 and 0.62 respectively, with the VPD₁ procedure showing a larger standard error of the mean. In contrast to the split-window data, the inclusion of elevation or the NDVI as additional predictor variables in the multiple regression equations only marginally improved the fit to the ground data (to a mean of 0.66 in both cases). There was also a small improvement to the fit (to 0.67) if an exponential relationship was used to describe the data. The time series of these monthly comparisons, shown in Fig. 7, revealed no temporal trends in the data, but demonstrates the massive increase in accuracy over a direct comparison of saturation deficit with the magnitude of the (channel 4 - channel 5) brightness temperature difference on which the technique is based.

TABLE 3.

The table shows the coefficients of determination obtained when the monthly mean NOAA - NCDC saturation deficits are linearly regressed with the monthly maximum vapour pressure deficits for 1990 (from Hay 1996).

	<i>Ch4 - Ch5</i>	<i>VPD₁</i>	<i>VPD₂</i>	<i>VPD₂ and elevation</i>	<i>VPD₂ and NDVI</i>
<i>Jan</i>	0.00†	0.40	0.41	0.54	0.52
<i>Feb</i>	0.03†	0.67	0.64	0.73	0.70
<i>Mar</i>	0.07†	0.64	0.62	0.69	0.66
<i>Apr</i>	0.10	0.69	0.64	0.65	0.69
<i>May</i>	0.05†	0.71	0.72	0.72	0.72
<i>Jun</i>	0.02†	0.67	0.63	0.63	0.64
<i>Jul</i>	0.01†	0.66	0.63	0.62	0.62
<i>Aug</i>	0.14	0.63	0.66	0.66	0.66
<i>Sep</i>	0.27	0.70	0.65	0.65	0.68
<i>Oct</i>	0.14	0.59	0.56	0.59	0.61
<i>Nov</i>	0.01†	0.63	0.64	0.70	0.65
<i>Dec</i>	0.08†	0.56	0.60	0.68	0.64
Mean	0.08	0.63	0.62	0.66	0.65
SE	0.02	0.02	0.02	0.01	0.02

† Not significant at the 0.01 level.

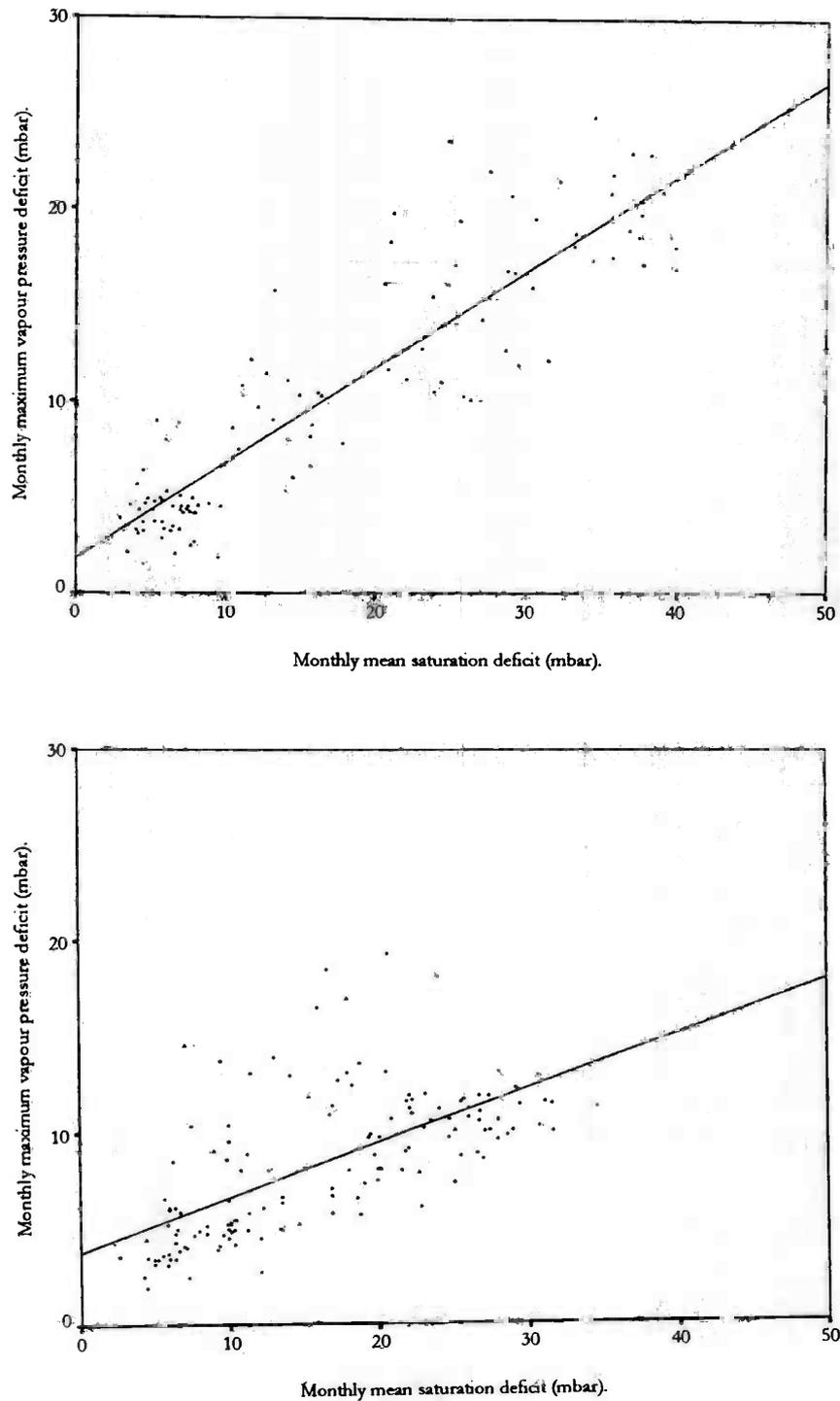
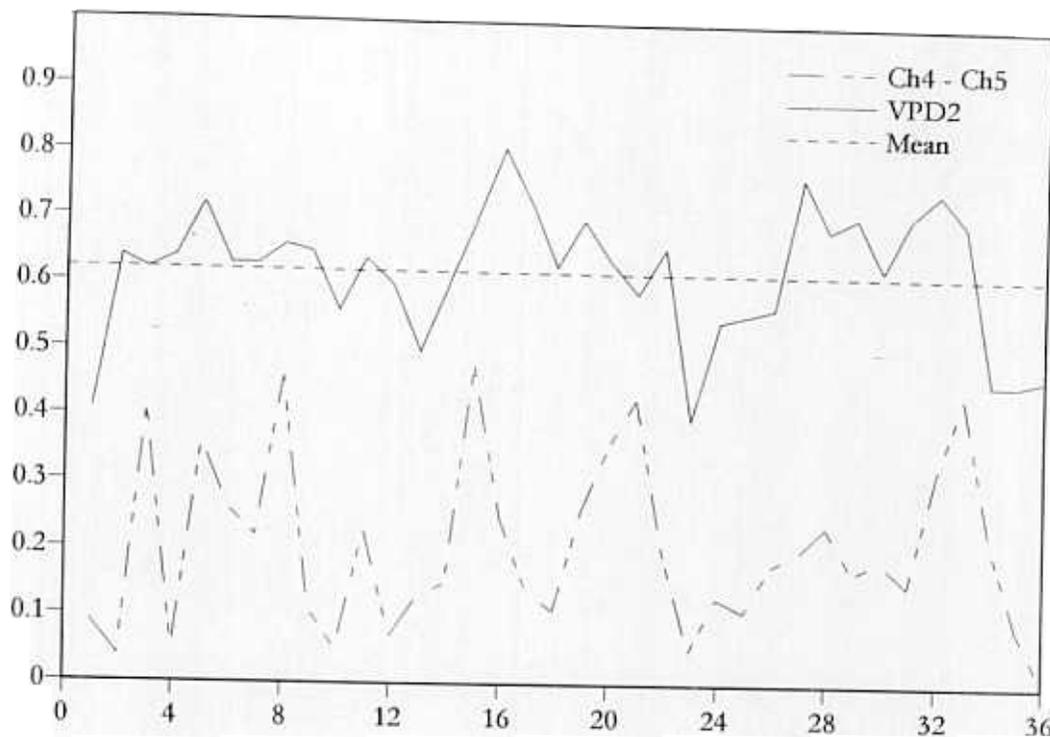


Figure 6 The comparison of mean monthly NOAA - NCDC saturation deficit with maximum monthly vapour pressure deficit₂. Examples are given of the best monthly fit (Apr 1989, $n = 107$, $P \ll 0.0001$, $r^2 = 0.81$, top graph) and of the worst fit (Nov 1989, $n = 140$, $P \ll 0.0001$, $r^2 = 0.40$, lower graph) (from Hay 1996).



Month (1988 -1990)

Figure 7 The adjusted r^2 of the comparison of mean monthly NOAA - NCDC saturation deficit and vapour pressure deficit₂ for the period 1988 - 1990. The graph shows the improvement of fit of the VPD₂ formulation over the brightness temperature difference (channel 4 - channel 5). The mean level of variance explained by the VPD₂ variable is shown by the horizontal dotted line (from Hay 1966).

Rainfall

The comparison of total monthly precipitation and total monthly CCD for 1988, 1989 and 1990 and is shown in Table 4. A highly significant positive linear relationship occurs in every month ($P \ll 0.0001$, $n = 143$, range 104 to 152). The mean r^2 for the three year period was 0.65. Examples of the best and worse fit to the precipitation data are shown in Fig. 8.

Restricting the analysis to those meteorological stations situated between 27° N and the equator (which corresponds to the TAMSAT calibration area) decreased the mean r^2 to 0.53 ($P \ll 0.0001$, $n = 84$ with a range 73 to 106 observations). The inclusion of elevation and NDVI in the regression equations did not increase the goodness of fit. The time-series of these data, plotted in Fig. 9, again revealed no seasonal relationships, but illustrated the large variability about the mean of the r^2 values.

TABLE 4.

The table shows the coefficients of determination obtained when relating total monthly NOAA - NCDC precipitation to total monthly CCD for 1988 - 1990 (from Hay 1996).

	1988			1989			1990		
	Linear	Cubic	Tamsat	Linear	Cubic	Tamsat	Linear	Cubic	Tamsat
<i>Jan</i>				0.81	0.87	0.75	0.84	0.90	0.00
<i>Feb</i>				0.75	0.78	0.50	0.72	0.74	0.56
<i>Mar</i>				0.76	0.77	0.80	0.41	0.48	0.36
<i>Apr</i>				0.50	0.65	0.49			
<i>May</i>				0.72	0.73	0.69	0.71	0.73	0.78
<i>Jun</i>				0.67	0.69	0.49	0.63	0.65	0.59
<i>Jul</i>				0.57	0.61	0.16	0.60	0.66	0.46
<i>Aug</i>				0.76	0.79	0.59	0.68	0.69	0.57
<i>Sep</i>	0.58	0.67	0.43	0.74	0.76	0.65	0.63	0.68	0.65
<i>Oct</i>	0.64	0.66	0.66	0.65	0.66	0.84	0.67	0.69	0.78
<i>Nov</i>	0.66	0.67	0.57	0.71	0.72	0.44	0.34	0.37	0.47
<i>Dec</i>	0.64	0.64	0.45	0.73	0.57	0.40	0.57	0.58	0.32
Mean	0.63	0.66	0.53	0.70	0.73	0.57	0.63	0.65	0.50
SE	0.02	0.01	0.05	0.02	0.02	0.05	0.04	0.04	0.06

missing CCD data.

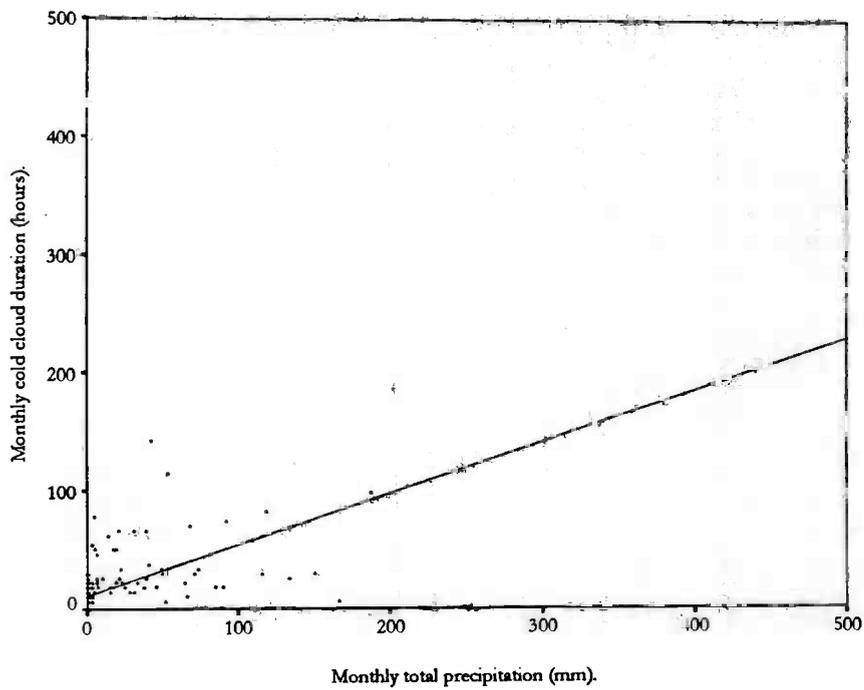
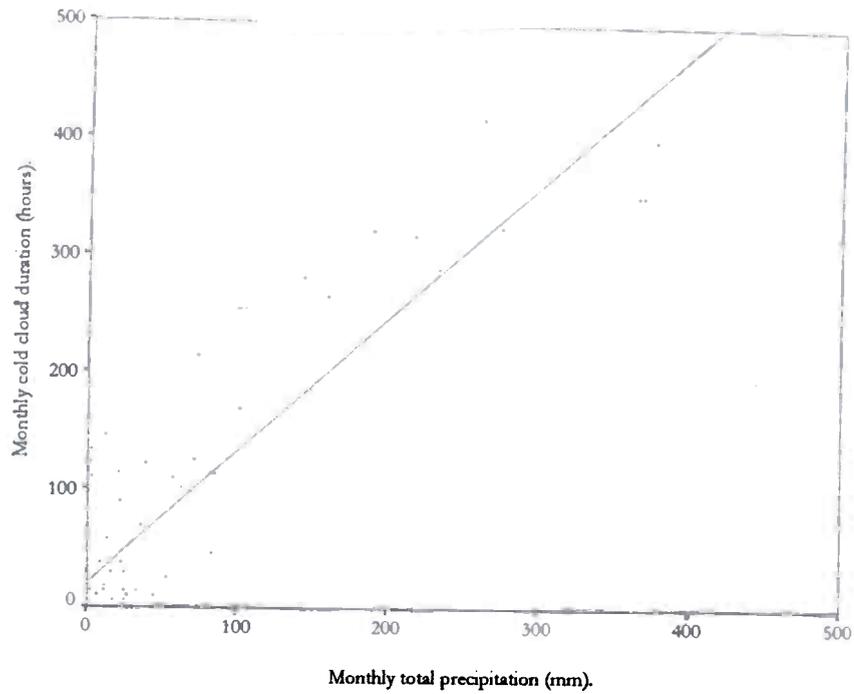


Figure 8. *The comparison of total monthly precipitation with total monthly CCD. Examples are given in the top graph (Jan 1990, $n = 104$, $P << 0.0001$, $r^2 = 0.84$) for the best monthly fit and in the lower graph (Nov 1990, $n = 125$, $P << 0.0001$, $r^2 = 0.34$) for the worst fit to the ground data.*

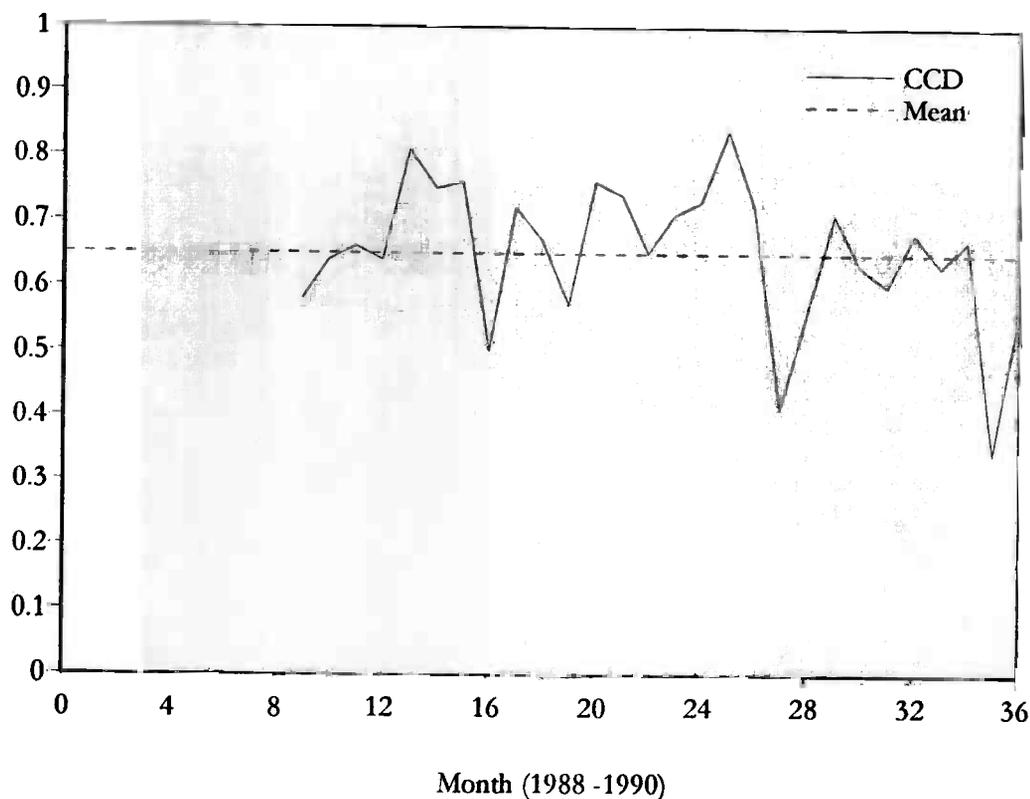


Figure 9 The adjusted r^2 of the monthly comparison of total monthly precipitation and total monthly CCD for 1988 - 1990. The mean level of variance explained is also shown as the horizontal dotted line.

Land-cover

Nigeria

Early analysis of the Nigerian vegetation data concentrated on predicting the percentage of each grid square covered with a particular vegetation type (e. g. forest or woodland etc.). This exercise showed that whilst the levels of predictive accuracy were high for grid squares with a high percentage coverage, those for grid squares with low percentage coverages were generally lower. On inspection of a series of bi-variate plots relating key predictor variables, it became clear that the bi-variate scatter of points around the mean was very much less for grid squares dominated by a particular vegetation type than it was for grid squares in which the same vegetation type was scarcer. As a result, we decided to concentrate on the dominant vegetation type within each grid square, and to carry out further analyses using all dominant vegetation types simultaneously.

The Nigerian vegetation analysis was carried out when only the 1990 satellite data had been processed, and before Fourier analysis of the Pathfinder products (subsequent analysis using

the 3-year Fourier processed data set confirms most of the conclusions that follow). The RVI, NDVI, SAVI and GEMI, and the Price (1984) and Gutman (1993) ground surface temperature estimates were calculated from the Pathfinder data, as described previously. Yearly maximum, minimum and mean values were calculated from the maximum value decadal for each variable, along with standard deviations to obtain an indication of the annual signal variability.

Three land-cover datasets were generated from the Nigerian data containing grid-squares where a given land-cover class exceeded threshold coverages of 50, 60 and 70 %. They were analysed independently in a discriminant analysis exercise where the nearest 2 x 2 satellite pixel array was averaged for each sample grid-square. Different covariance matrices were calculated for each vegetation type and the observed sample prior probabilities were assumed in assigning grid cells to land-cover classes (since these more fairly represented the actual proportions of the different land-cover types). The ten variables giving the greatest separation between vegetation classes in multivariate space (measured by the summed Mahalanobis distances) were chosen from the 44 satellite variables and DEM data and were then used to classify the training set data, pixels being assigned to that vegetation type giving them the highest posterior probability of class membership.

The ability of the discriminant analysis to separate vegetation classes was measured by the “producer accuracy” or the percentage of grid-squares within the training set known to have been classified correctly and “consumer accuracy” or the percentage of grid-squares correctly assigned to a class, for each vegetation class in turn. The Kappa and Tau statistics (Ma and Redmond 1995) were also calculated from the marginal totals of the rows and columns of the classification matrix.

Table 5 shows the ten predictor variables which gave the best description of land-cover for the 50, 60 and 70% threshold datasets. Although elevation ranks as the most important variable in each case, raw waveband information, surface temperature estimates and the alternative vegetation indices were all featured. Accuracy increased with threshold coverage as the grid-squares became increasingly dominated by signals from the vegetation class to which they were assigned.

Fig. 10 shows maps of the observed and predicted land-cover at the 60 % threshold and Table 6 presents the classification matrix for these data. A statistically perfect description of the ground data is indicated when all the numbers are on the diagonal of this matrix. Fig. 11 shows the same 60 % threshold dataset with the remaining grid-squares assigned to their dominant land-cover class as predicted by the analysis.

Table 7 shows the producer and consumer accuracies derived from Table 6. These results are encouraging in terms of the Anderson criteria for evaluating land-cover classification accuracy (Anderson *et al.* 1976) especially since these were formulated for use with maps generated from high spatial resolution satellite imagery. Although interpretation accuracy falls below 85 % in all but two classes, the overall level of prediction is high. Accuracies are approximately equal across classes and differences seem to be related to the sample size (*e. g.* compare open woodland and cultivation with forest and scrubland). In addition, misclassifications are understandable: open woodland is misclassified as dense woodland; dense woodland as forest.

The levels of accuracy with which class estimates were obtained using low-spatial resolution satellite imagery suggests that spatial extrapolation to the whole of West Africa for ecological and environmental studies is feasible. Such a prediction is shown in Fig. 12.

TABLE 5

The top predictor variables for Nigeria land-cover at the 50, 60 and 70 % thresholds, and accuracy measures for the predictions.

Rank	Variable (50 %)	Variable (60 %)	Variable (70 %)
2	Elevation	Elevation	Elevation
3	Ch1 SD	Ch1 SD	Ch2 max
4	Ch3 max	Ch1 min	Gutman max
5	Ch1 min	RVI max	GEMI mean
6	RVI mean	NDVI max	GEMI min
7	NDVI mean	SAVI max	Ch1 min
8	Ch2 SD	Price mean	Price min
9	Ch1 mean	GEMI max	RVI min
10	RVI min	Price min	Ch3 min
	GEMI max	Ch2 mean	Price mean
Kappa	0.468	0.541	0.640
Tau	0.467	0.525	0.635

TABLE 6

Classification matrix of land-cover classes for the 60 % threshold coverage showing the number of Nigerian sample grid-squares in each category.

Observed		2	3	4	5	6	7	8	Total
Predicted									
1. Bare ground	0	0	5			0	0		8
2. Grassland	0	19	3	2	1	0	0	0	25
3. Scrubland	0	0	41	2	4	5	0	1	53
4. Open woodland	0	7	21	89	34	6	0	5	162
5. Dense woodland	0		9	29	168	40	0	3	250
6. Forest	0	0		0		86	0	0	88
7. Mangrove	0	0	0	0	0	0	12	0	12
8. All cultivation	0	3	26	42	47	10	0	120	248
Total	0	30	106	165	256	147	12	130	846

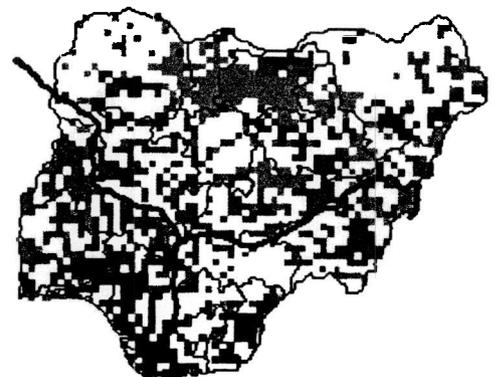
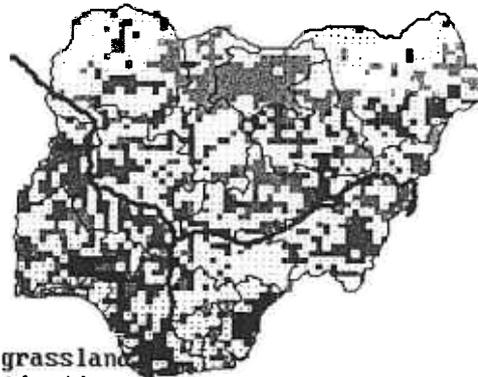
Predictions of All vegetation types in Nigeria

Using variables (S+pp)

- ELEU
- Tha1
- THp2
- NDmax
- Tha3
- NDp3
- THp3
- NDa0
- CCDp3
- NDp2

Observed

Predicted



Scale

- Wooded grassland
- All cultivation
- Mangrove
- Forest
- Dense woodland
- Open woodland
- Scrub
- Grass
- Bare ground

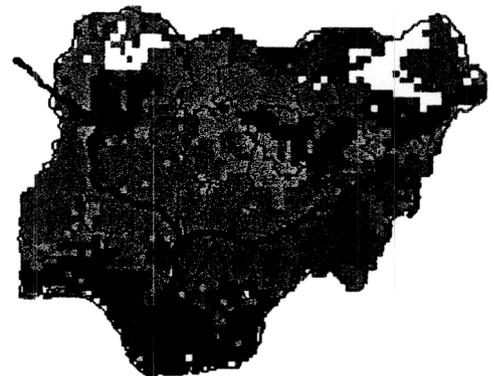
Predictions of All vegetation types in Nigeria.

Using variables (S+pp)

- ELEU
- Tha1
- THp2
- NDmax
- Tha3
- NDp3
- THp3
- NDa0
- CCDp3
- NDp2

Observed

Predicted



Scale

- Wooded grassland
- All cultivation
- Mangrove
- Forest
- Dense woodland
- Open woodland
- Scrub
- Grass
- Bare ground

Fig. 10 (upper). Observed land-cover types in Nigeria (only those grid squares with >60% coverage of a single land-cover type are shown) and the cover predicted from the set of satellite variables listed (details in Table 6).

Fig. 11 (lower). As for Fig. 10, but with the predicted dominant vegetation cover for all grid squares, using the same set of satellite predictor variables.

Predictions of all vegetation types in West Africa.

■ = No prediction, ■ = Bare ground, ■ = Grass, ■ = Scrub, ■ = Open woodland
■ = Dense woodland, ■ = Forest, ■ = Mangrove, ■ = All cultivation,
■ = Wooded grassland.

Variables (S+pp)

ELEU
Tha1
THp2
NDmax
Tha3
NDp3
THp3
NDa0
CCDp3
NDp2

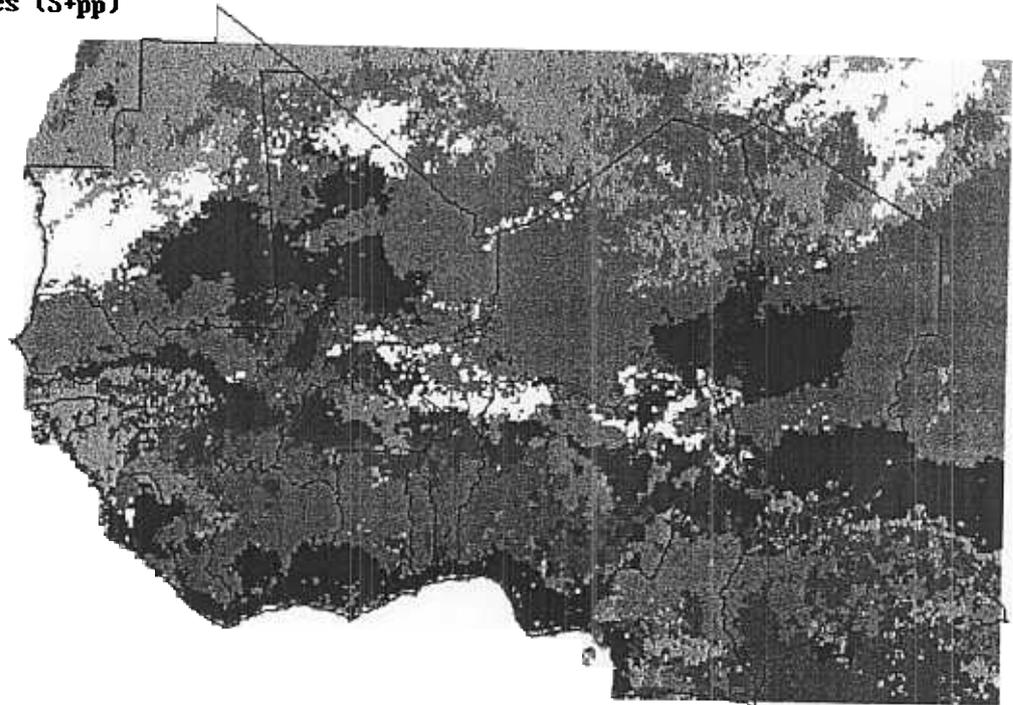


Fig. 12. Predicted dominant land-cover types for West Africa, using the ERGO training set data for Nigeria (Figs 10 and 11) and parts of Mali and Tchad and satellite data for the whole region.

TABLE 7

Classification accuracy of the land-cover classes in Table 6

Class	Producer accuracy (%)	Consumer accuracy (%)
1. Bare ground	0.0	N.A.
2. Grassland	76.0	63.3
3. Scrubland	77.4	38.7
4. Open woodland	54.9	53.9
5. Dense woodland	67.2	65.6
6. Forest	97.7	58.5
7. Mangrove	100.0	100.0
8. All cultivation	48.4	92.3

Africa

Unsupervised classification methods were applied to the satellite data layers, and the results later compared with White's vegetation map for Africa (White 1983). This map of vegetation coverage for the continent, although a current standard, has several apparent contradictions. For example, the phytochoria given by White (his Fig. 4) appear to cut across some of the vegetation bands of his more detailed sheet maps. The reason for this appears to be that phytochoria are based on endemic species, whilst the detailed sheet maps reflect actual vegetation assemblages.

Satellite data were prepared in the same way as for Nigeria, but for the 3-year period 1988 - 1990 (or 1989 - 1991 uniquely for the CCD data). Similar spectral vegetation indices were used as before (NDVI, GEMI, SAVI) with a few additional data layers such as for vapour pressure deficit (VPD) and channel-3 waveband data. The monthly data were temporal Fourier processed as described previously. On the basis of preliminary analyses, the maximum, minimum and range of the Fourier-transformed data were not included in the final classification.

Image processing was conducted using ERDAS IMAGINE 8.2 software. Images consisting of various combinations of Fourier components were subjected to an unsupervised classification process (ISODATA) and pixels were assigned to 9 or 15 clusters, one of which always represented water. This resulted in a number of statistically defined (and therefore spectrally-separate) classes, each one potentially representing a different land-cover class. Given ground-data or expert knowledge, cluster signatures can be assigned to different land-cover classes and so can be later used in supervised classification of the original images.

The results of the ISODATA classification (a minimum distance classification procedure) were compared to determine:

- i) how the number of classes requested in ISODATA affected the final classification (NDVI classifications only).
- ii) whether GEMI and SAVI perform as well as or better than NDVI and

iii) whether the inclusion of vapour pressure deficit (VPD) layer affects the classification.

A combination of temporal Fourier (amp0, amp1, amp2, amp3, phs1, phs2, phs3 only) processed normalised-difference vegetation index (NDVI), Price split-window thermal (T_s) and Channel-3 (Ch-3) data produced unsupervised classification results which matched much of the detail of the phytocoria map of vegetation in Africa (White, 1983). Comparison of Figs. 13 and 14 with Fig. 15 (a vegetation map of 17 classes derived by amalgamating White's vegetation groups - see map legend) indicates that there is much detail in the unsupervised output that is not represented in the vegetation map. This is also the case with the 9-class (8 land-cover classes) classifications. In making these comparisons it should be considered that White's map represents the areal extent of different vegetation types and is not strictly a land-cover map (*i. e.* it does not take account of anthropogenic influences). Comparison of Fig. 14 with Figs 17 and 18 indicate that NDVI provides a more acceptable classification of land-cover than does either GEMI or SAVI. Indeed, the use of GEMI and SAVI produced results which appear spurious when compared with White's map. Comparison of Figs. 14 and 16 also suggests that inclusion of VPD data does not improve the classification. These results are broadly in line with those of the more detailed analysis of Nigerian land-cover using supervised classification methods applied to similar satellite data sets.

Further progress is dependent on the availability of representative, land-cover data for large areas. Digital elevation information will also be included in future classifications.

Land cover Change Detection

The 10-year data set for both NDVI and the Price (1984) split window thermal index from the Joint Research Centre, Ispra, were accurately co-registered within the JRC (Malingreau and Belward 1994) so that the time series of changes at the level of the pixel can be taken as a measure of real changes in the recorded variables, rather than noise due to misaligned pixels. This data set has been used by Eric Lambin in several papers on change detection analysis, using a technique first suggested by Singh (1989) (Lambin and Strahler 1994 a, b). Briefly change vector analysis places each pixel for each year within a 12-dimensional Euclidean space (1 dimension per month) and compares years over time. The difference between a pixel's position from one point in time (*i. e.* one year) to the next (*i. e.* a different year) has both magnitude and direction. The magnitude contains information about the extent of change, whilst the direction indicates the type of change. Given the inter-year variability of NDVI and other signals it is usual for these comparisons to be made in relation to the long-term average values for each pixel. A consistent change over time with reference to this standard image is taken to indicate a real effect in need of an explanation, whilst random changes are taken to indicate noise. It is also usual to examine the direction of any changes seen within principal component space as determined by the long-term average image.

An obvious and perhaps clearer means of detecting change is to examine changes in the Fourier components. Most of the obvious changes that take place in vegetation type over time will result in fairly obvious and predictable changes in the Fourier components. For example, the change from a forest to a more open woodland causes a reduction in the mean vegetation index but an increase in the amplitude of the seasonal vegetation cycle. Differences between the phase timings of the annual signals from CCD (*i. e.* rainfall) and NDVI (vegetation growth)

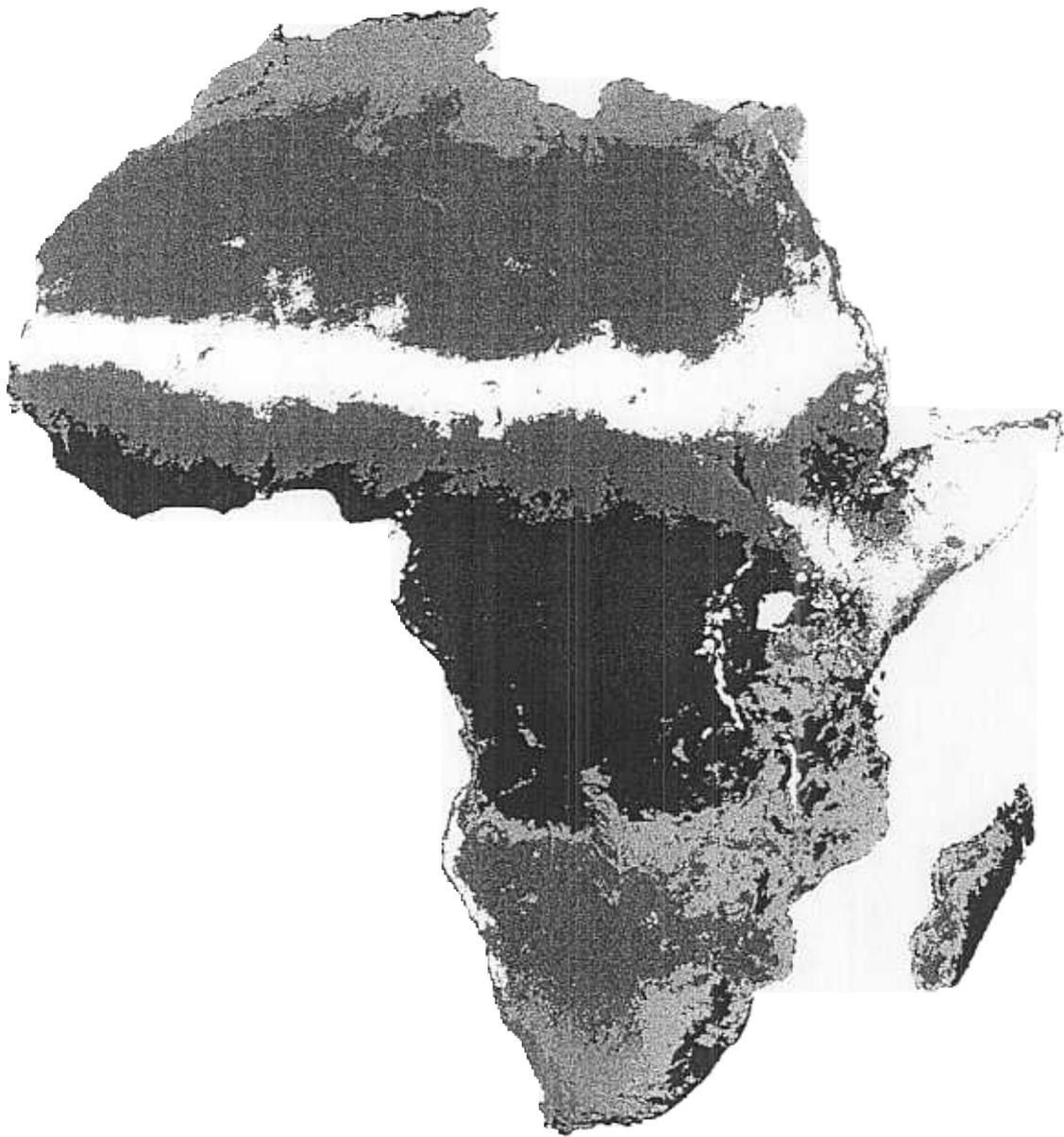


Fig. 13. ISODATA classification of Fourier-processed satellite data - NDVI 9 classes



Fig. 14. ISODATA classification of Fourier-processed satellite data - NDVI 15 classes

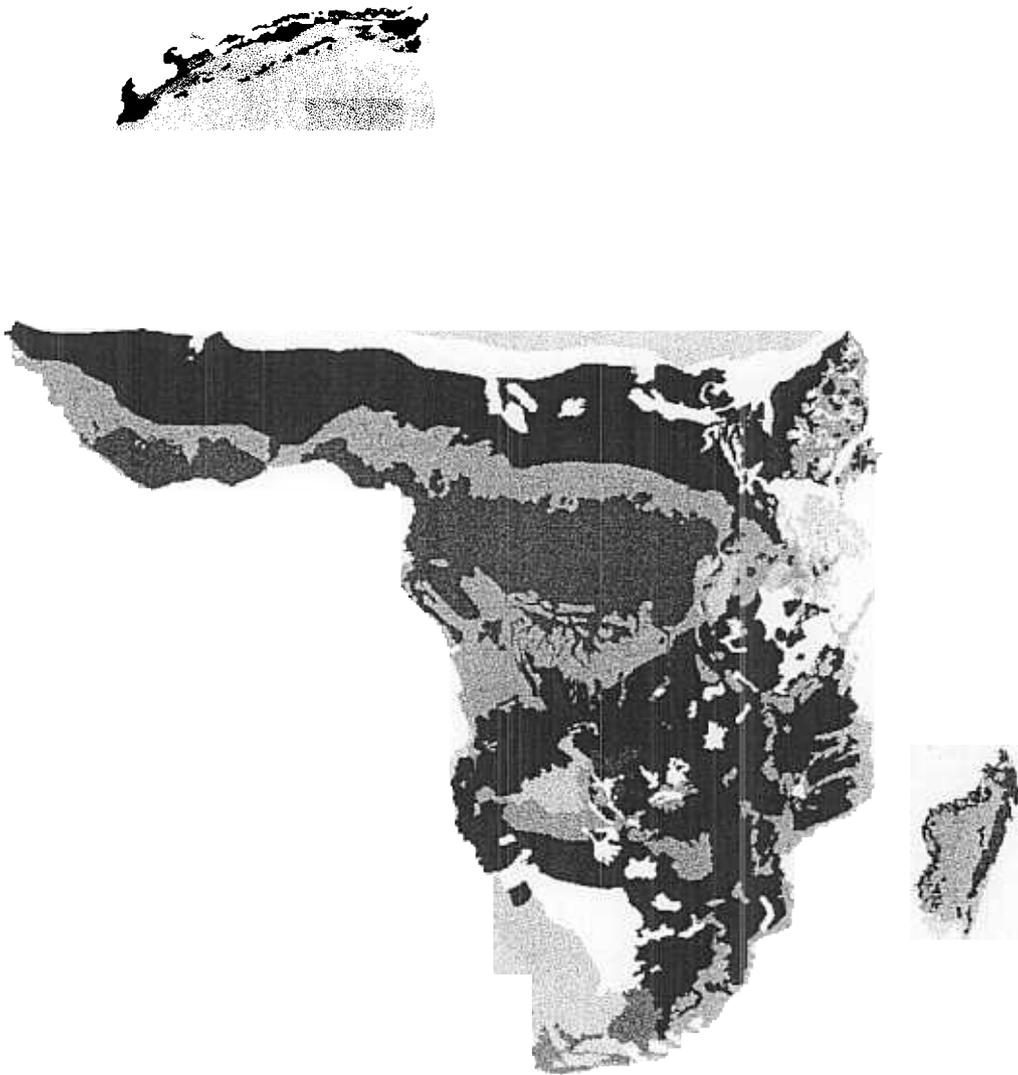


Fig. 15. White's vegetation map of Africa - 17 classes

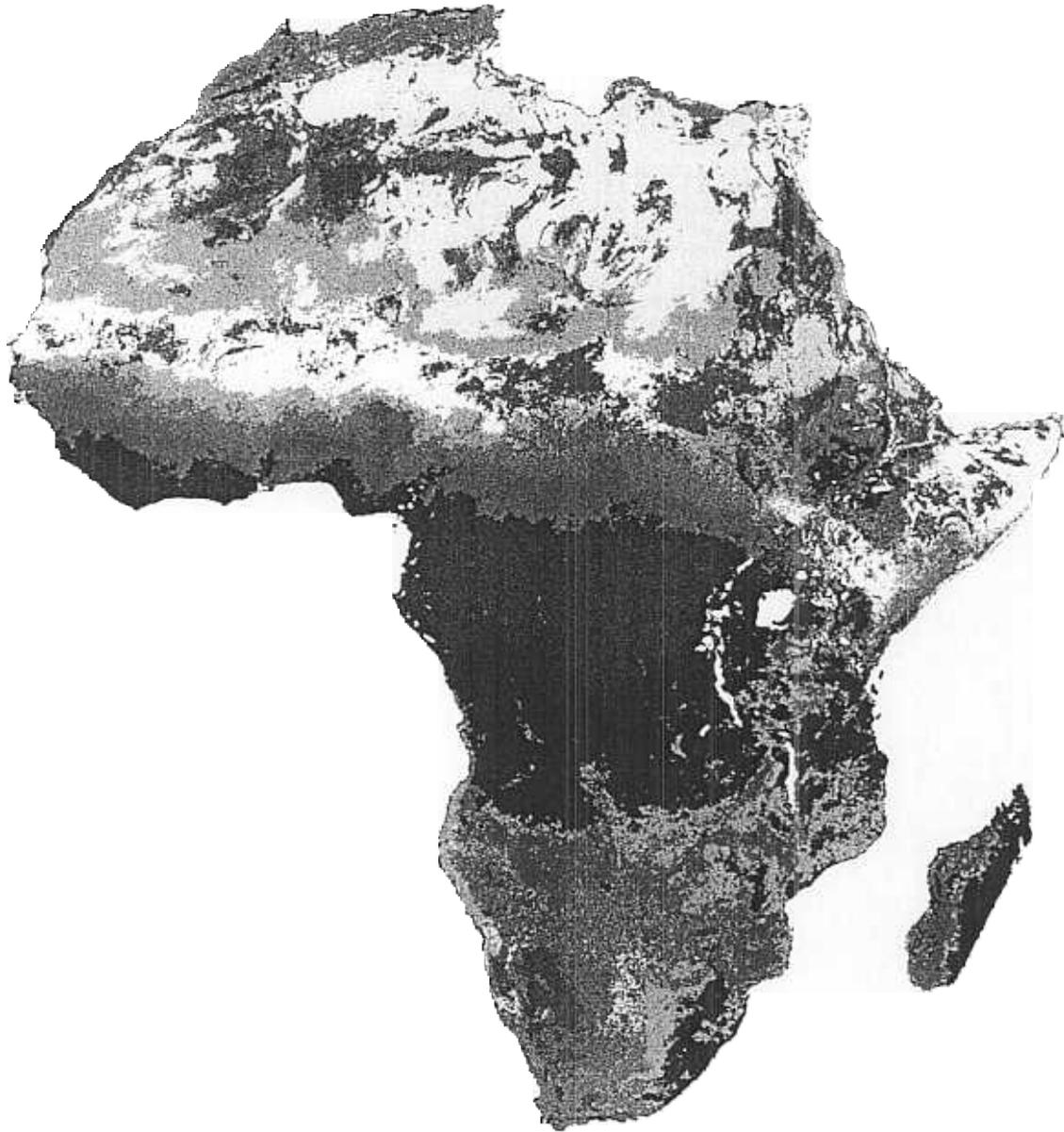


Fig. 16. ISODATA classification of Fourier-processed satellite data - VPD 15 classes



Fig. 17. ISODATA classification of Fourier-processed satellite data - GEMI 15 classes

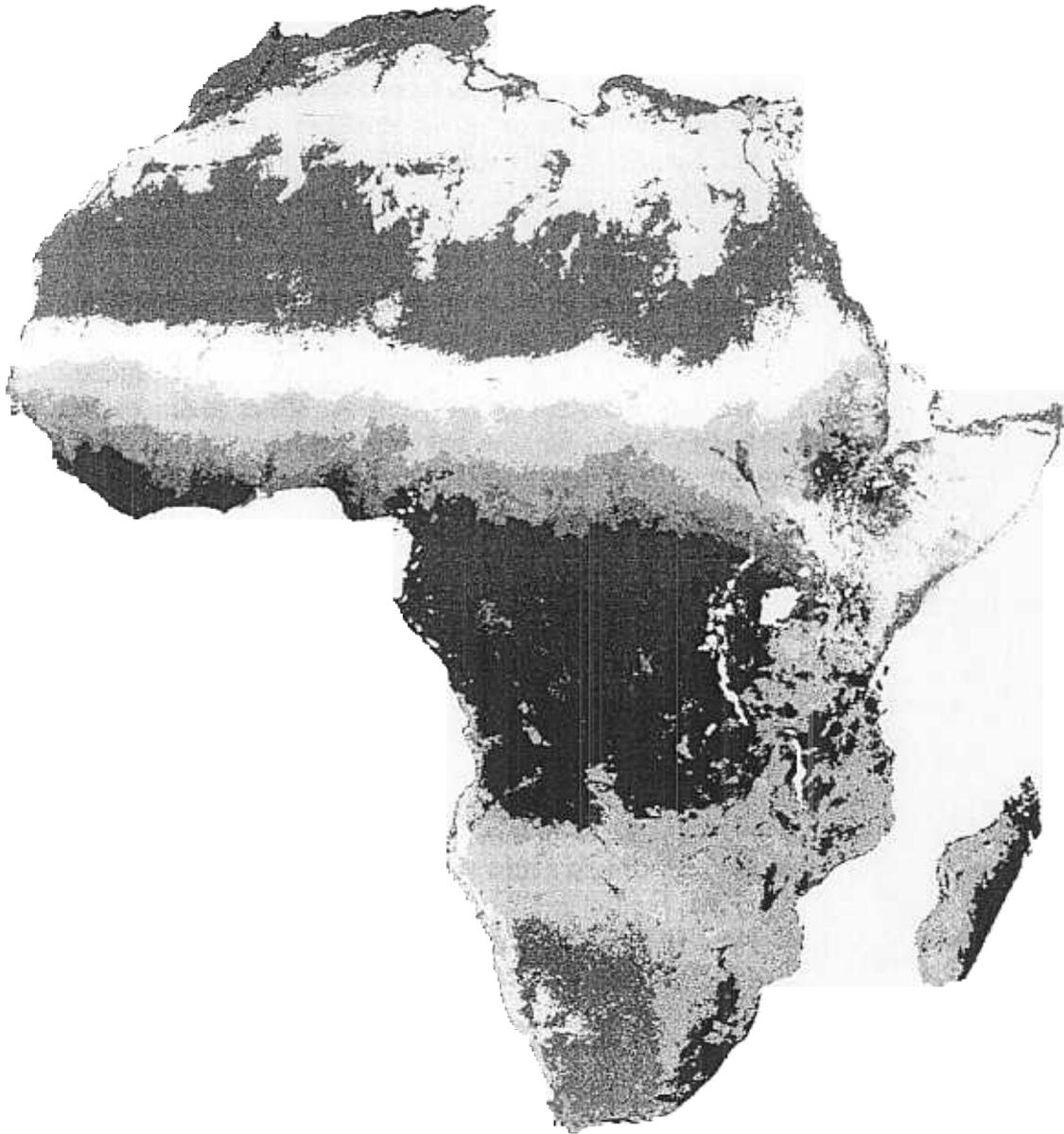


Fig. 18. ISODATA classification of Fourier-processed satellite data - SAVI 15 classes

were found to be associated with the level of agricultural activity in Nigeria. Natural vegetation is adapted to take immediate advantage of seasonal rains (in some cases even flushing before the onset of the rains) and there is little delay between the peak CCD and peak NDVI values. On the other hand, cultivated areas show a delay between these two peaks which increases with the level of cultivation (up to a limit of c. 2/3 months in West Africa), a feature probably attributable to the standard agricultural practice of delaying sowing until the first rains appear. Hence examination of the changes in key Fourier components over time should provide a useful index of environmental change. This was tested by comparing Fourier analysed NDVI and split window channels for the two, two-year periods 1982/3 and 1990/91. Examples are shown in Fig. 19 and 20 for the whole of Africa and for Nigeria alone. In these figures, increases over time in the value of the relevant Fourier component are shown in green and decreases in red.

At the continental level there is a strong spatial patterning that suggests that the technique is picking up some regionally varying signal rather than simply noise (which would give the images a more speckled appearance). The NDVI images (Fig. 19 a) - c)) show an increase in mean values in the moister parts of the continent between the two dates (perhaps a general reflection of recovery from the drought of the early 1980's?) and a general increase in the amplitude of the annual cycle of vegetation growth, though with some areas showing a marked decrease (e. g. parts of Côte d'Ivoire, Central African Republic and northern Uganda/southern Sudan). The timing of the peak of this annual cycle occurs earlier in a broad belt from northern Zaire through Uganda to parts of Kenya and SW Ethiopia, and in southern Namibia into S. Africa (and patchily along the northern limits of the savannah region, which may be artefactual - see later).

Changes in the thermal channel suggest that the average temperature has increased over the same period, and this is associated with a decreased amplitude of seasonal fluctuation in most of the vegetated parts of West Africa through to northern East Africa (and a slightly earlier timing of this peak), but an increase in amplitude in the more southerly parts of East Africa, Zambia and Zimbabwe (in places with a slightly delayed timing).

The Nigerian maps show some regional detail of these changes. The mean NDVI increased over the period in question in most of the southerly regions, but appeared to decrease slightly in the regions along and to the North of the northern border. This decrease occurs over a period that is generally acknowledged to have become wetter, and is perhaps an example of Goward and Prince's observation (Goward and Prince 1995) that whilst long-term rainfall patterns can be related to long-term vegetation patterns, shorter-term variations in each of these variables are less easily related to each other, because of presently unappreciated time-delays in the response of vegetation to rainfall. It has been suggested (Willy Wint, *pers. comm.*) that the vegetation North of Nigeria's border is dependent upon ground-water reserves. Such reserves are likely to show a delayed response to drought, and to increased rainfall following drought, and this may help to explain the patterns seen in the difference images.

Within Nigeria's borders, in the North of the country, the annual amplitude increased in the West, but decreased in the East (with a slightly earlier peak in both areas) whilst in the South there is a mixture of changes, especially pronounced in the South-West. This region has been

one of considerable change (in many cases decreases) in agricultural activity in the last 15 - 20 years, a phenomenon attributed to the impact of oil wealth. Indeed, there are strong similarities in the pattern of change in this region and the annual percentage change in cultivation activity recorded by SLAR and aerial surveys between the years 1976 and 1990 (Fig. 26), such that areas where cultivation is recorded as having decreased are associated with an earlier peak of vegetation activity in 1990/1 compared with 1982/3, *i. e.* in the direction expected from the discussion earlier in this section. There is also a broad similarity between the agricultural change map in the North of the country and changes in the annual amplitude of the NDVI signal. Increasing levels of agricultural activity are associated with increasing annual amplitudes, *i. e.* again in the direction which might be predicted.

The thermal change maps for Nigeria show a relatively uniform increase in mean temperature, with a widespread decrease in amplitude, and a seasonal peak that occurs increasingly early from South to North (the green areas North of this may be an artefact due to the way the phase information is stored in the images: once timing is earlier than the start of the year, it occurs in December of the previous year and so the difference in timing flips from being negative to positive).

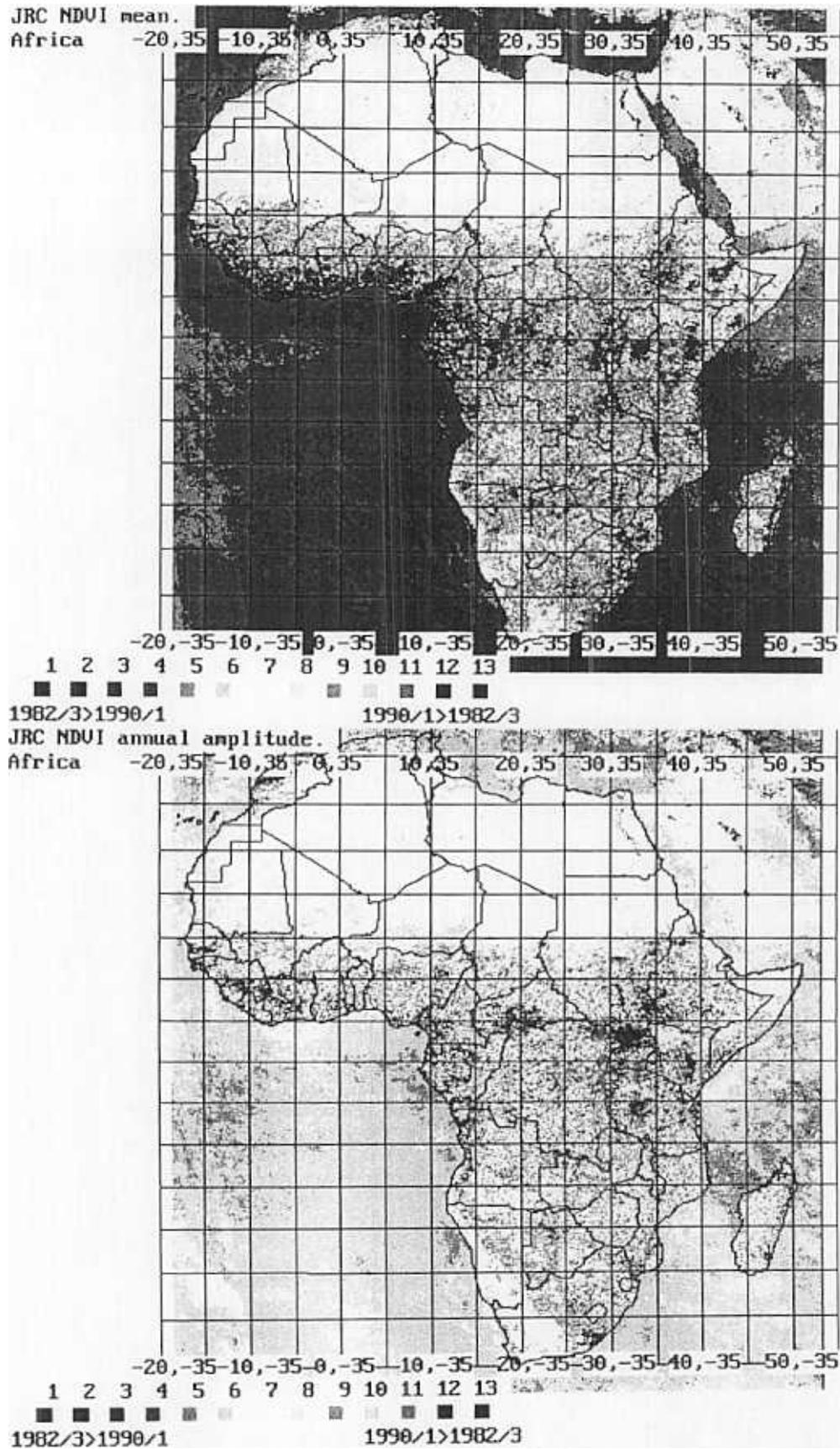
Tsetse distribution

Côte d'Ivoire and Burkina Faso

Results for *G. morsitans*, *G. tachinoides*, *G. palpalis* and *G. longipalpis* are shown in Fig. 21, using the selected satellite variables listed in Table 8 (for further details, and more examples, see Rogers, Hay and Packer 1996). Data from all three surveys in these countries were used in the analysis (Laveissiere and Challier 1977,1981; Anon 1982). No clustering was carried out of the presence/absence sites before analysis. It was found, however, that if the country surveys were kept separate in the analysis (*i. e.* if Côte d'Ivoire's results were kept separate from Burkina Faso's) the predictive accuracies increased for all but one species (the exception being *G. palpalis*).

East Africa

Subspecies of those that occur in West Africa, or related sister species, also occur in East Africa, along with species which only occur in this region. Using essentially the same predictor variable data set (in this case re-sampled to 0.05 degree grid squares, a sub-sampling which involves some repetition of the 8x8km pixels) discriminant analysis was used to describe the distributions of *G. morsitans* (all three subspecies), *G. fuscipes fuscipes* and *G. pallidipes* in East Africa, with the results shown in Fig. 22 and Table 9. For each of the maps in Fig. 22, the presence and absence data were each divided into two clusters before analysis. The training set of data for this analysis was chosen at random from the mapped distributions, 1000 points for absence and 200 for presence (excluding Arabia and Madagascar).



a

b

Fig. 19. Land-cover change detection for Africa. The difference between NDVI Fourier variables for the early and late 1980s a) mean; b) amplitude of the annual cycle. Increases in each variable over this period are indicated in green, decreases in red.

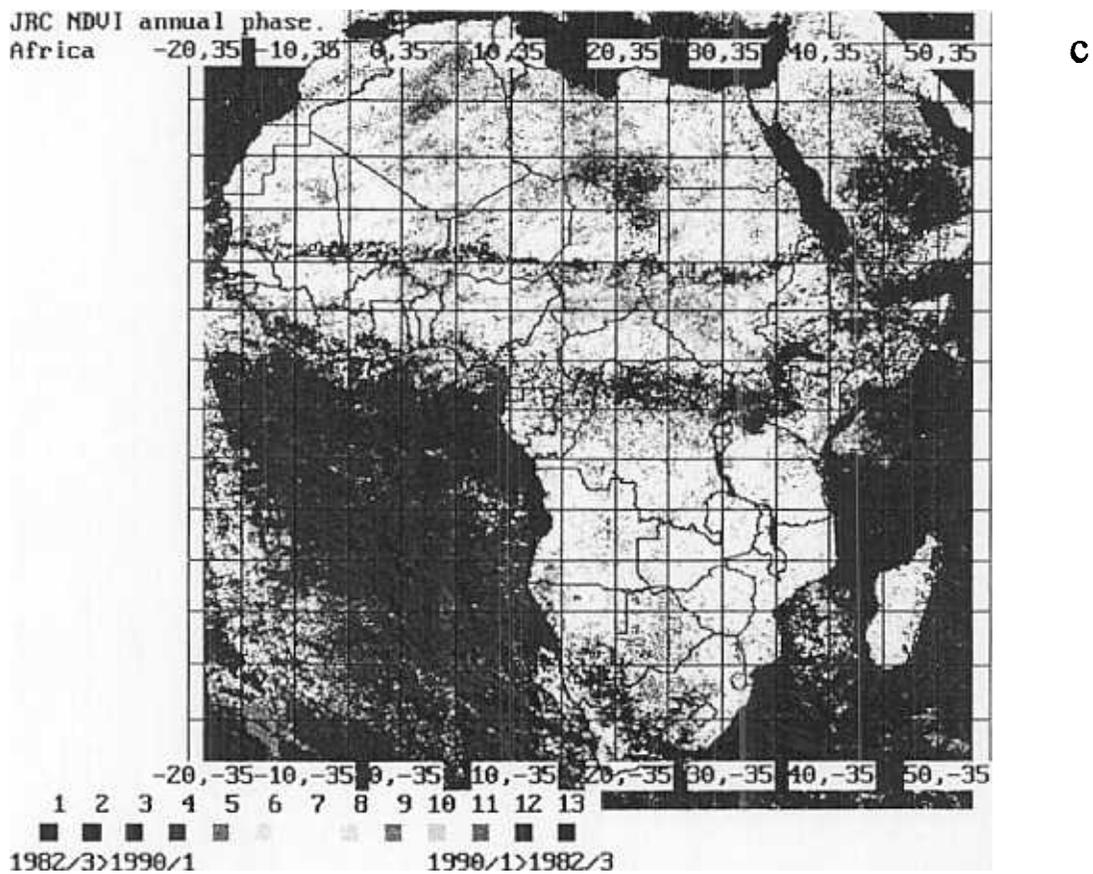


Fig. 19. Land-cover change detection for Africa. The difference between NDVI Fourier variables for the early and late 1980s c) phase (= timing) of the annual cycle. Increases over this period are indicated in green, decreases in red.

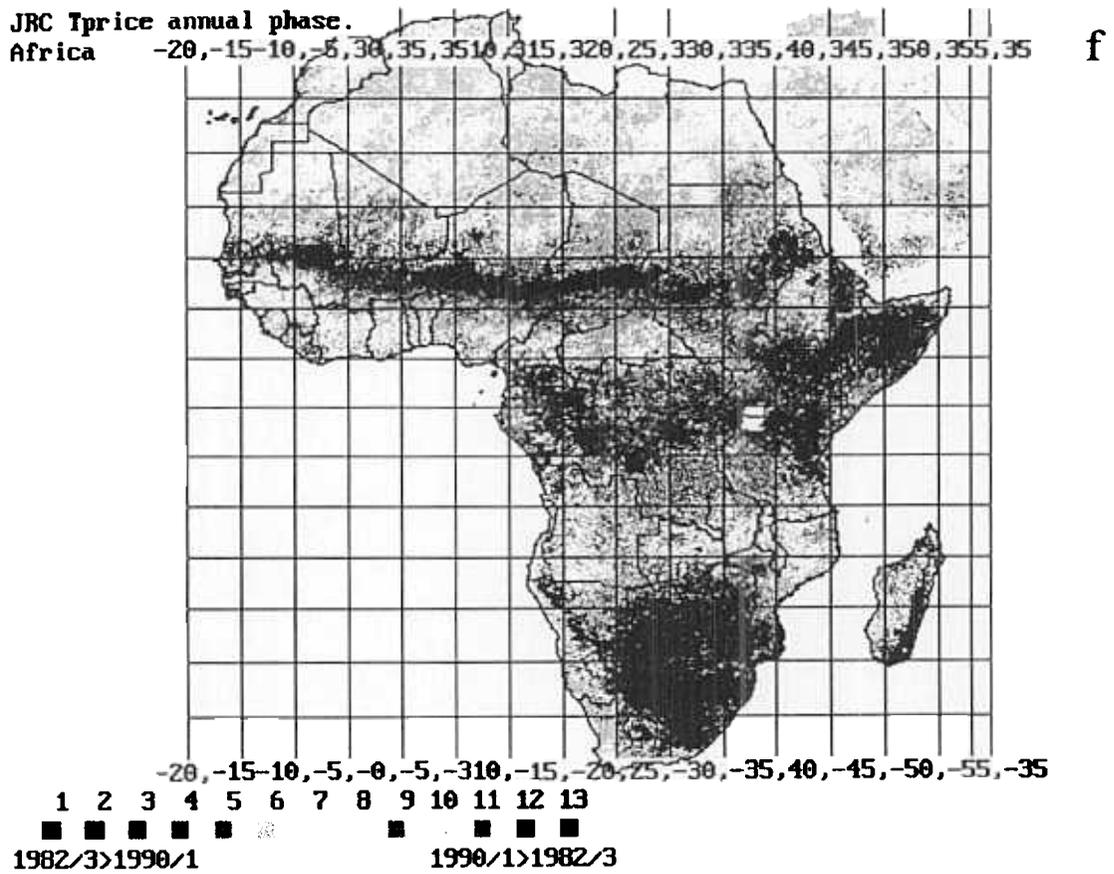


Fig. 19. Land-cover change detection for Africa. The difference between the Price (=thermal) Fourier variables for the early and late 1980s f) phase (= timing) of the annual cycle. Increases over this period are indicated in green, decreases in red.

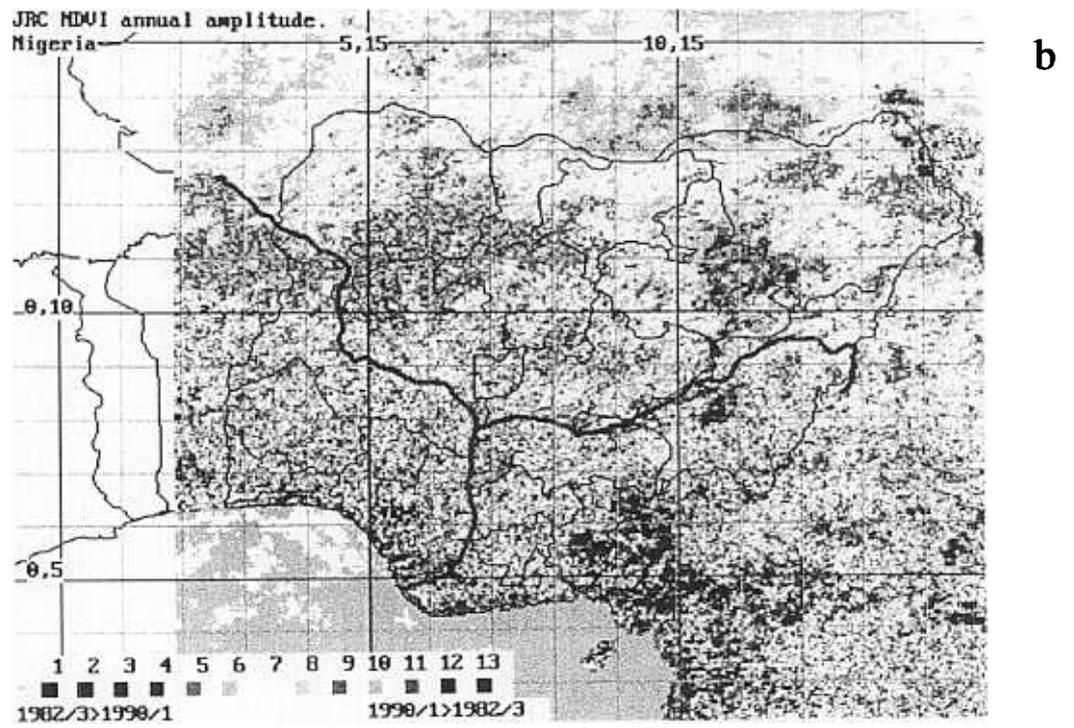
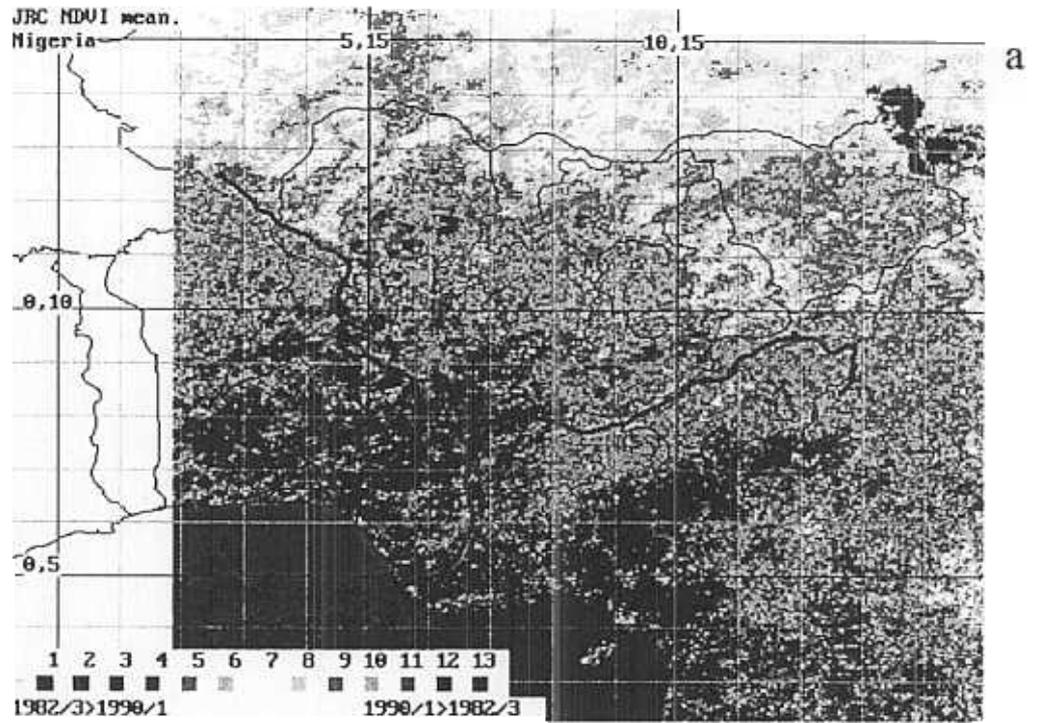
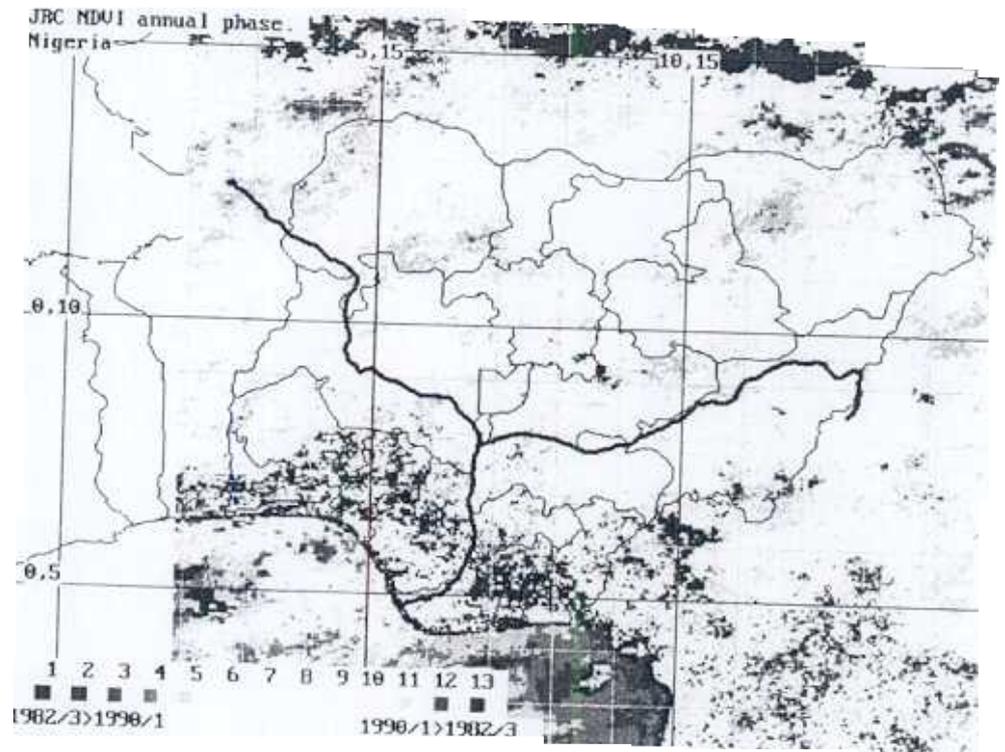
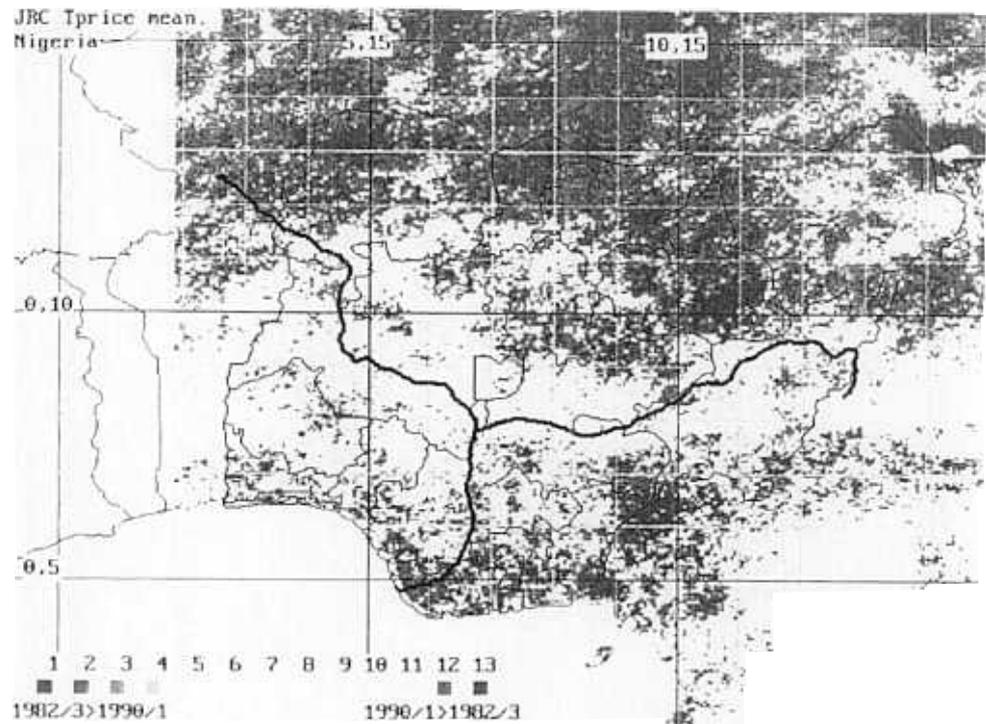


Fig. 20. Land-cover change detection for Nigeria. The difference between NDVI Fourier variables for the early and late 1980s a), mean; b) amplitude of the annual cycle. Increases in each variable over this period are indicated in green, decreases in red.

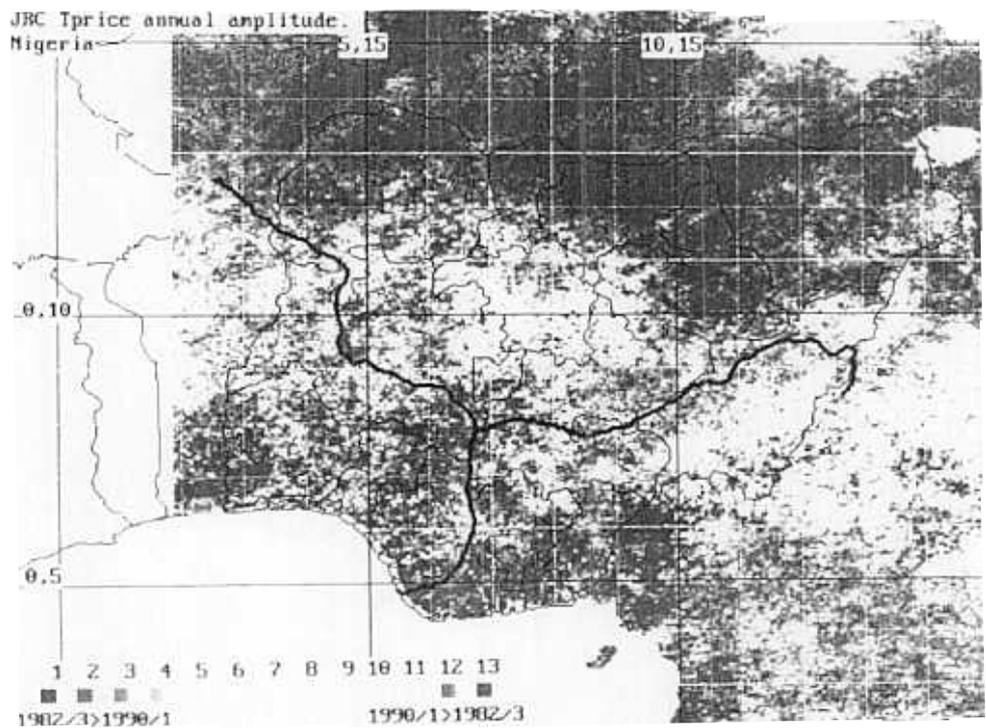


c

Fig. 20. Land-cover change detection for Nigeria. The difference between NDVI Fourier variables for the early and late 1980s c) phase (= timing) of the annual cycle. Increases over this period are indicated in green, decreases in red.

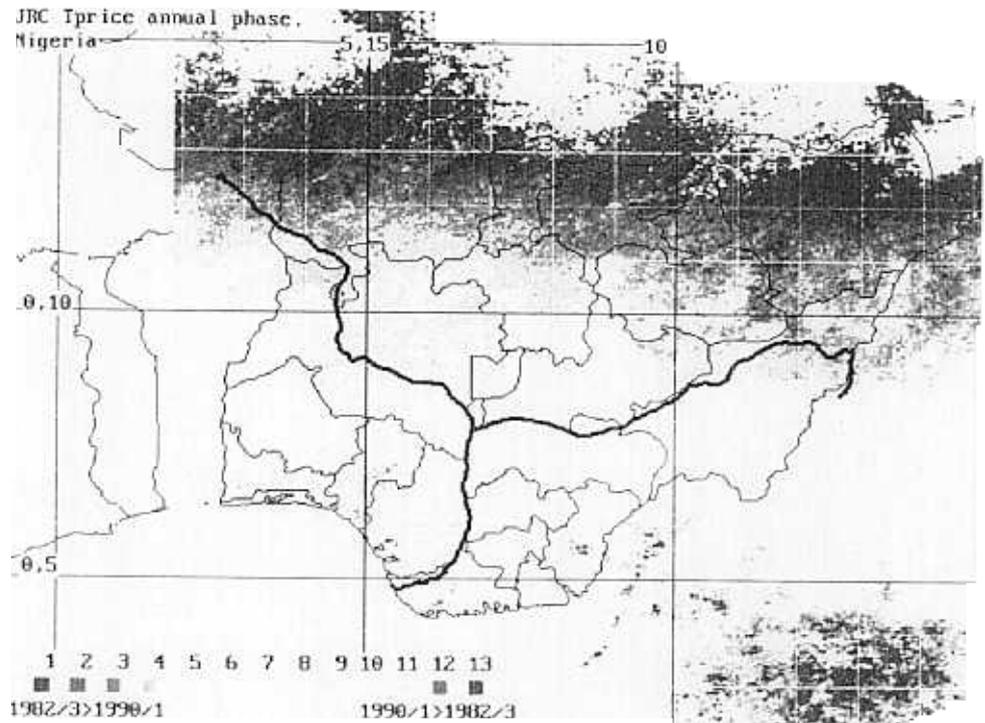


d



e

Fig. 20. Land-cover change detection for Nigeria. The difference between the Price (=thermal) Fourier variables for the early and late 1980s d), mean; e) amplitude of the annual cycle. Increases in each variable over this period are indicated in green, decreases in red.



f

Fig. 20. Land-cover change detection for Nigeria. The difference between the Price (=thermal) Fourier variables for the early and late 1980s f) phase (= timing) of the annual cycle. Increases over this period are indicated in green, decreases in red.

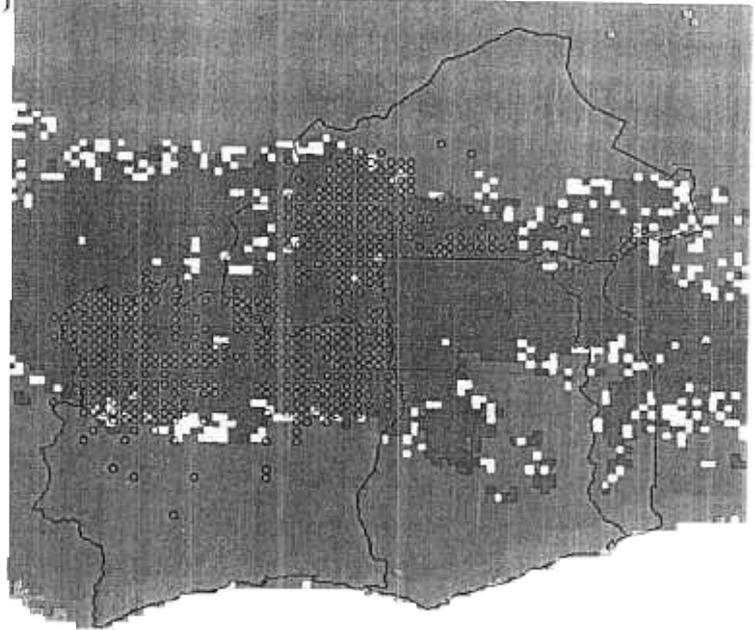
Predictions of areas in Cote d'Ivoire, Burkina Faso, Ghana and Togo with *G. morsitans*

Using variables (S+p)

CH4range
CCDmin
NDp-CDp
NDa1
NDa1/CH4a1
CCDa2
NDa3
NDa0
Tha2
NDm/CH4m

Probability scale

■ = 0.65 - 1.0
■ = 0.55 - 0.649
■ = 0.50 - 0.549
■ = 0.45 - 0.499
■ = 0.35 - 0.449
■ = 0.0 - 0.349
○ = Actual
■ = No prediction



a

Training set results: % Correct = 67 % False +ve = 29 % False -ve = 4
Sensitivity = .916 Specificity = .485

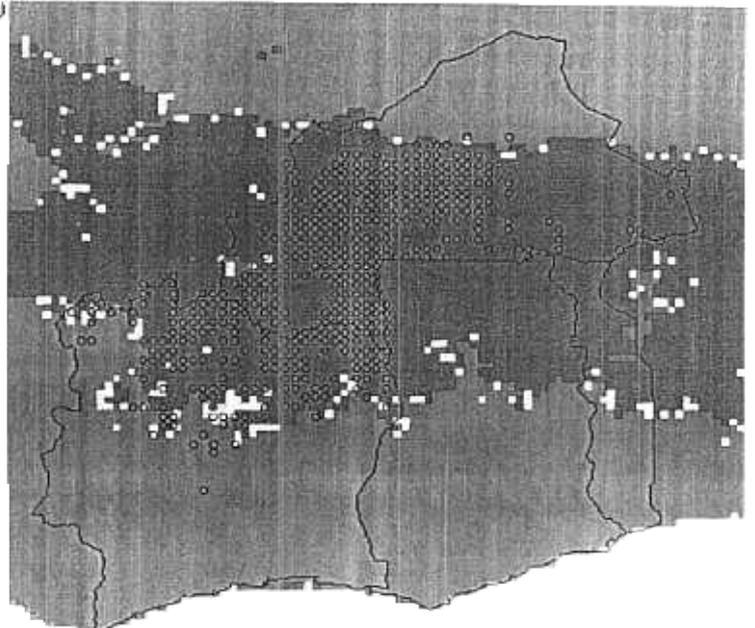
Predictions of areas in Cote d'Ivoire, Burkina Faso, Ghana and Togo with *G. tachinoides*

Using variables (S+p)

CH4range
NDm/CH4m
Thp2
NDa0
Tha0
Tha2
NDa2
ELEV
Tha1
CCDmin

Probability scale

■ = 0.65 - 1.0
■ = 0.55 - 0.649
■ = 0.50 - 0.549
■ = 0.45 - 0.499
■ = 0.35 - 0.449
■ = 0.0 - 0.349
○ = Actual
■ = No prediction



b

Training set results: % Correct = 74 % False +ve = 24 % False -ve = 2
Sensitivity = .946 Specificity = .629

21) Observed (black circles) and predicted distributions of a) *G. morsitans* and b) *G. tachinoides* in Côte d'Ivoire and Burkina Faso. The predicted distributions are based on the satellite data listed in Table 8, and are on a colour scale from red (low probability) to green (high probability).

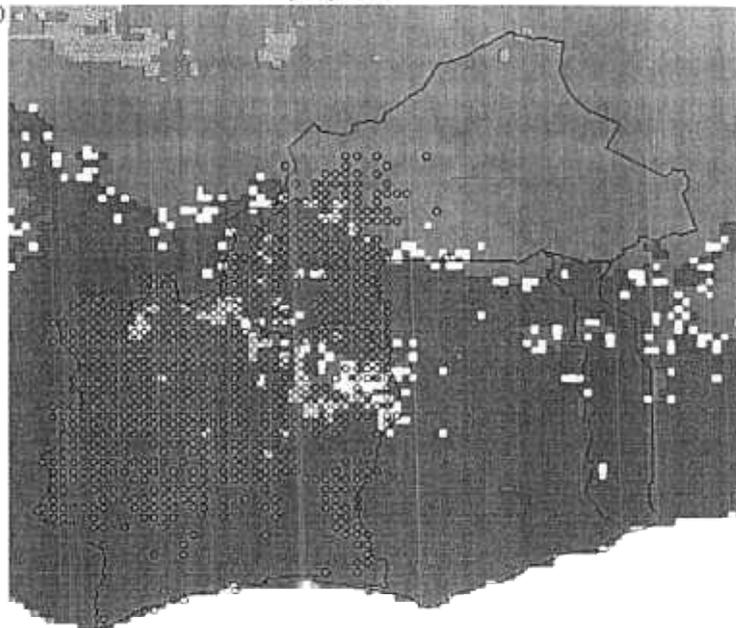
Predictions of areas in Cote d'Ivoire, Burkina Faso, Ghana and Togo with *G. palpalis*

Using variables (S+p)

NDmax
Tha2
Tha1
Tha0
CH4min
NDm/CH4m
Tha3
CCDp2
CCDa3
THp2

Probability scale

■ = 0.65 - 1.0
 ■ = 0.55 - 0.649
 ■ = 0.50 - 0.549
 ■ = 0.45 - 0.499
 ■ = 0.35 - 0.449
 ■ = 0.0 - 0.349
 ○ = Actual
 ■ = No prediction



Training set results: % Correct = 87 % False +ve = 4 % False -ve = 8
 Sensitivity = .901 Specificity = .744

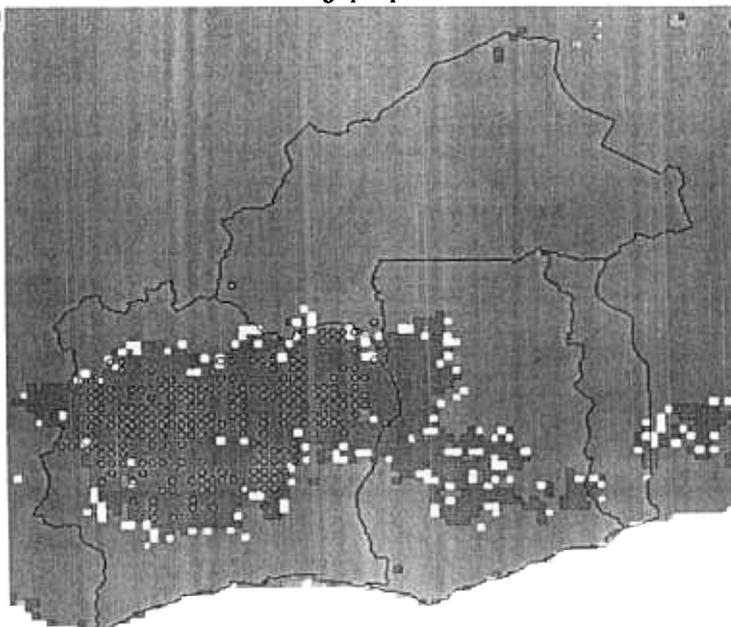
Predictions of areas in Cote d'Ivoire, Burkina Faso, Ghana and Togo with *G. longipalpis*

Using variables (S+p)

CCDp2
NDmax
Tha1
NDp1
NDa2
CCDa3
CCDa1
NDp3
Tha0
Tha3

Probability scale

■ = 0.65 - 1.0
 ■ = 0.55 - 0.649
 ■ = 0.50 - 0.549
 ■ = 0.45 - 0.499
 ■ = 0.35 - 0.449
 ■ = 0.0 - 0.349
 ○ = Actual
 ■ = No prediction



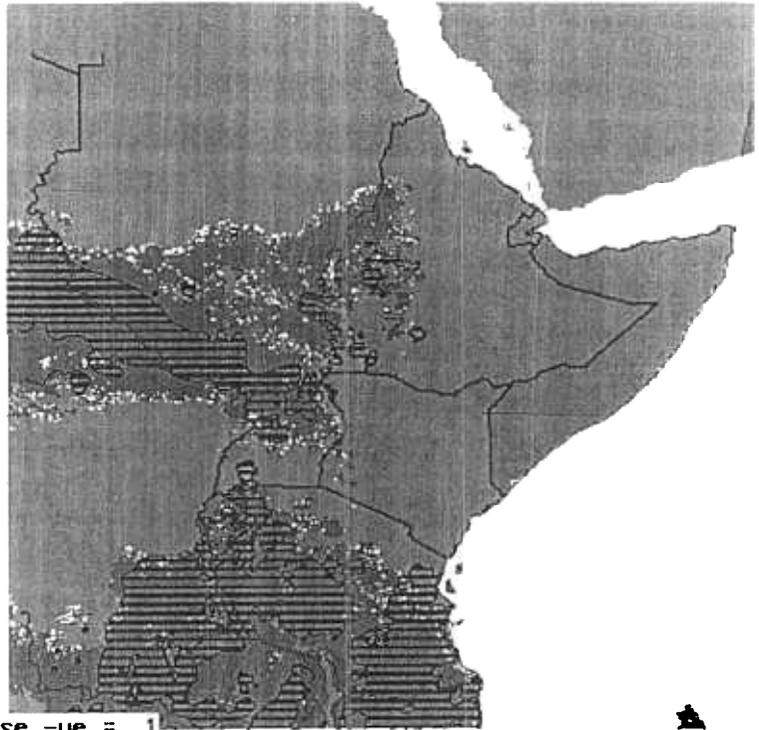
Training set results: % Correct = 71 % False +ve = 28 % False -ve = 1
 Sensitivity = .976 Specificity = .640

21) Observed (black circles) and predicted distributions of c) *G. palpalis* and d) *G. longipalpis* in Côte d'Ivoire and Burkina Faso. The predicted distributions are based on the satellite data listed in Table 8, and are on a colour scale from red (low probability) to green (high probability).

Species: *G. morsitans*
 Probability of occurrence
 ■ = 0.65 - 1.0
 ■ = 0.55 - 0.649
 ■ = 0.50 - 0.549
 ■ = 0.45 - 0.499
 ■ = 0.35 - 0.449
 ■ = 0.0 - 0.349
 ≡ = Actual

Vars (S,pp,2/2)
 NDx
 DEM
 Nm/Cm
 CCa0
 Np-Ch4p
 NDa0
 Ch4x
 ATN(N/Ch4)
 Ch4n
 Ch4a0

Training data results:
 % Correct = 82
 % False +ve = 18 % False -ve = 1
 Sensitivity = .960 Specificity = .786

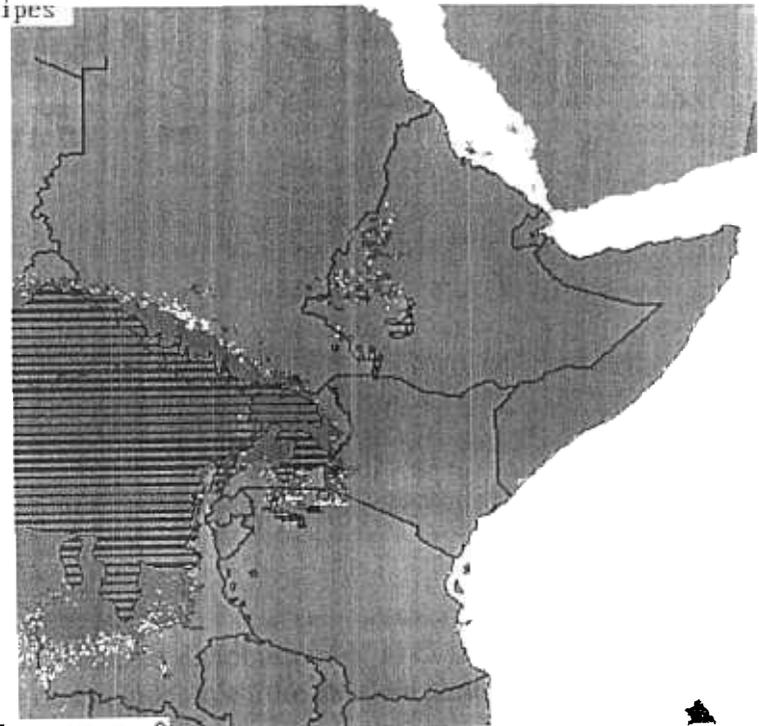


a

Species: *G. fuscipes fuscipes*
 Probability of occurrence
 ■ = 0.65 - 1.0
 ■ = 0.55 - 0.649
 ■ = 0.50 - 0.549
 ■ = 0.45 - 0.499
 ■ = 0.35 - 0.449
 ■ = 0.0 - 0.349
 ≡ = Actual

Vars (S,pp,2/2)
 Nm/Cm
 DEM
 NDx
 CCa0
 NDa0
 Ch4a0
 Ch4a1
 ATN(N/Ch4)
 CDa2
 CDr

Training data results
 % Correct = 94
 % False +ve = 6 % False -ve = 0
 Sensitivity = .975 Specificity = .931



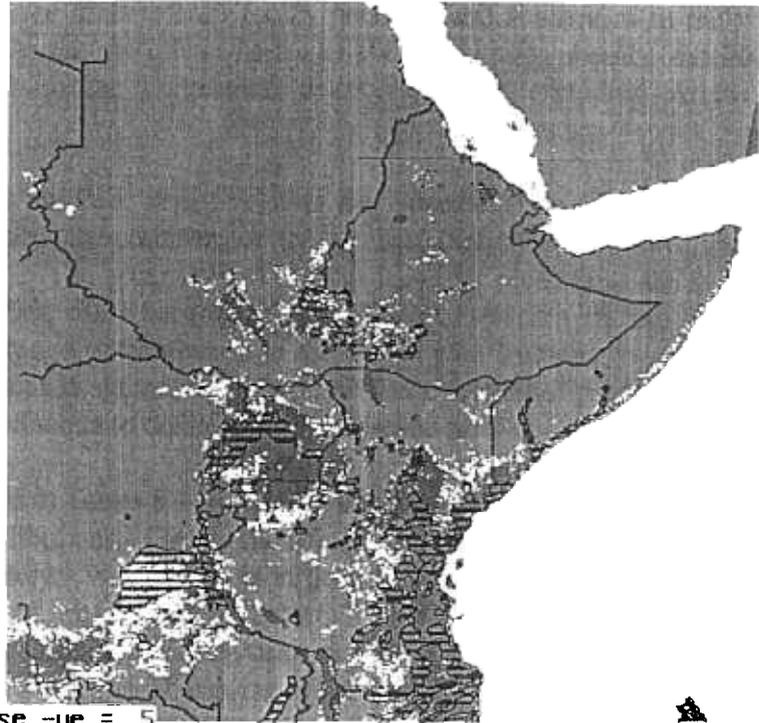
b

22) Observed (black lines) and predicted distributions of a) *G. morsitans* and b) *G. fuscipes fuscipes*, and in East Africa. The predicted distributions are on a colour scale from red (low probability) to green (high probability).

Species: *G. pallidipes*
 Probability of occurrence
 ■ = 0.65 - 1.0
 ■ = 0.55 - 0.649
 ■ = 0.50 - 0.549
 ■ = 0.45 - 0.499
 ■ = 0.35 - 0.449
 ■ = 0.0 - 0.349
 ≡ = Actual

Vars (S,pp,2/2)
 DEM
 Nm/Cm
 NDa2
 Ch4n
 Ch4a3
 NDx
 ATN(N/Ch4)
 CCDa0
 NDa0
 Ch4a0

Training data results
 % Correct = 86
 % False +ve = 9 % False -ve = 5
 Sensitivity = .715 Specificity = .889



C

22) Observed (black lines) and predicted distributions of c) *G. pallidipes* in East Africa
 The predicted distributions are on a colour scale from red (low probability) to green (high probability)

Africa

Finally, the continent-wide distributions of *G. morsitans*, *G. palpalis*, *G. tachinoides* and *G. pallidipes* were analysed, using three clusters each for fly presence and absence, with the results shown in Fig. 23 and Table 10. Again, the training set was 1000 points for absence and 200 for presence, selected at random from the continental distribution and excluding Arabia and Madagascar.

A comparison of the results from the three regions, either for 'species' (*i. e.* their subspecies or close relatives) that occur throughout the continent or for 'species' that occur in two of the three regions only, is shown in Table 11. The accuracies of the whole-Africa predictions are at least as good as those of the regional studies, either because of the splitting of the data set into clusters or (more likely) because the wider range of conditions in the larger area includes more which are unsuitable for the species. This should automatically increase the accuracy of predicting areas of absence, and may also allow the technique to define presence/absence sites more distinctly.

Table 12 brings together the results in Tables 8 - 10 to show the importance of the different satellite data channels in predicting fly distributions. Clearly the relative importance of vegetation, thermal and rainfall channels varies regionally. West Africa is in general hotter than East Africa, and the thermal variables in West Africa are relatively more important than is the case elsewhere. NDVI tends to dominate in East Africa and for the whole-Africa predictions. Finally rainfall is relatively un-important on the West/East Africa regional basis, but is much more important at the continental scale. Elevation is also a key variable at the continental scale for all four species chosen (although, as reported earlier, this may be an artefact of the clustering techniques employed to divide the data before analysis).

TABLE 8

The key predictor variables for the distribution of four species of tsetse in West Africa (from Rogers, Hay and Packer 1996).

Rank	<i>G. morsitans</i>	<i>G. longipalpis</i>	<i>G. palpalis</i>	<i>G. tachinoides</i>
1	CH4range	CCDp2	Ndmax	CH4range
2	CCDmin	NDmax	CH4a2	NDm/CH4m
3	NDp-CCDp	CH4a1	CH4a1	CH4p2
4	NDa1	NDp1	CH4mean	NDmean
5	NDa1/CH4a1	NDa2	CH4min	CH4mean
6	CCDa2	CCDmean	NDm/CH4m	CH4a2
7	NDa3	CCDa1	CH4a3	NDa2
8	NDmean	NDp3	CCDp2	Elev
9	CH4a2	CH4mean	CCDa3	CH4a1
10	NDm/CH4m	CH4a3	CH4p2	CCDmin
Result/n	10	10	10	10
%correct	67	71	87	74
%false +	29	28	4	24
%false -	4	1	8	2
Sensitivity	0.92	0.98	0.90	0.95

Specificity	0.49	0.64	0.74	0.63
-------------	------	------	------	------

TABLE 9

The key predictor variables for the distribution of three species of tsetse in East Africa.

Rank	<i>G. morsitans</i>	<i>G. fuscipes fuscipes</i>	<i>G. pallidipes</i>
1	Ndmax	NDm/CH3m	DEM
2	DEM	DEM	NDm/CH4m
3	NDm/CH4m	Ndmax	Nda2
4	CCDa0	CCDa0	CH4min
5	NDp1-CH4p1	Ndmean	CH4a3
6	Ndmean	CH4mean	Ndmax
7	CH4max	CH4a1	ATN(NDm/CH4m)
8	ATN(NDm/CH4m)	ATN(NDm/CH4m)	CCDmean
9	CH4min	CCDa2	Ndmean
10	CH4mean	CCDrange	CH4mean
Result/n	10	10	10
%correct	82	94	86
%false +	18	6	9
%false -	1	0	5
Sensitivity	0.96	0.98	0.72
Specificity	0.79	0.93	0.89

TABLE 10

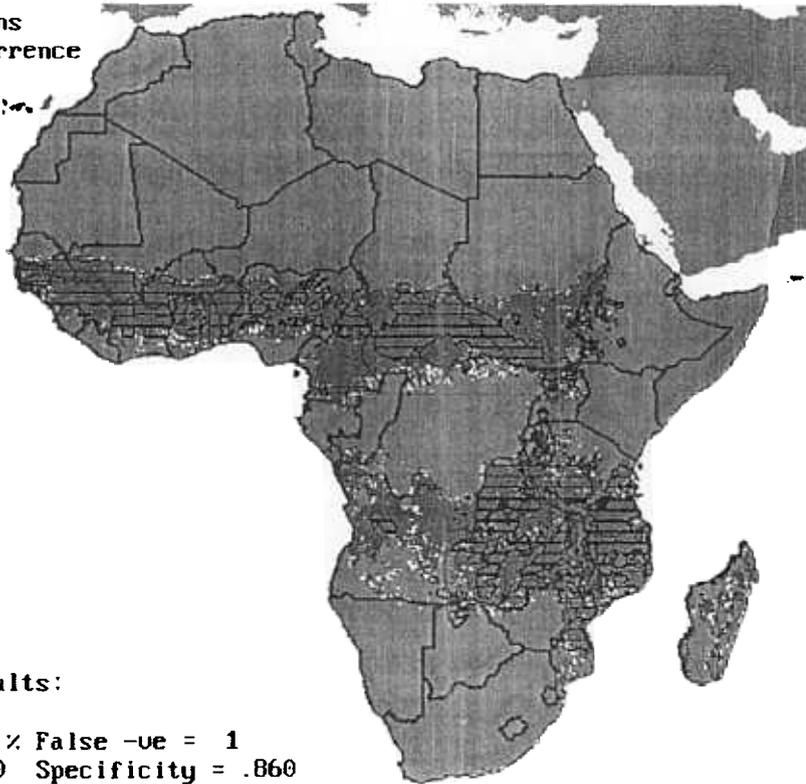
The key predictor variables for the distribution of four species of tsetse throughout Africa.

Rank	<i>G. morsitans</i>	<i>G. palpalis</i>	<i>G. tachinoides</i>	<i>G. pallidipes</i>
1	DEM	DEM	DEM	DEM
2	Ndmax	NDm/CH4m	NDm/CH4m	Ndrange
3	NDm/CH4m	Ndmax	CCDmean	CH4a1
4	CCDmean	CCDmax	Ndmean	NDp1-CCDp1
5	Ndmean	Nda2	CH4min	Nda1/CH4a1
6	NDp1-CH4p1	Nda1/CH4a1	CCDmin	Nda1
7	CH4min	CH4min	Nda1/CH4a1	Nda2
8	CCDa2	NDp1-CH4p1	CCDa2	CCDa3
9	Nda3	CCDa1	CCDa1	Ndmean
10	Nda1	CCDa3	Ndrange	NDm/CCDm
Result/n	10	10	10	10
%correct	88	94	96	89
%false +	12	5	4	9
%false -	1	0	0	2
Sensitivity	0.97	0.98	0.96	0.87
Specificity	0.86	0.935	0.96	0.89

Species: *G. morsitans*
 Probability of occurrence
 ■ = 0.65 - 1.0
 ■ = 0.55 - 0.649
 ■ = 0.50 - 0.549
 ■ = 0.45 - 0.499
 ■ = 0.35 - 0.449
 ■ = 0.0 - 0.349
 ≡ = Actual

Vars (S,pp,3/3)
 DEM
 NDx
 Nm/Cm
 CCDa0
 NDa0
 Np-Ch4p
 Ch4n
 CDa2
 NDa3
 NDa1

Training data results:
 % Correct = 88
 % False +ve = 12 % False -ve = 1
 Sensitivity = .970 Specificity = .860

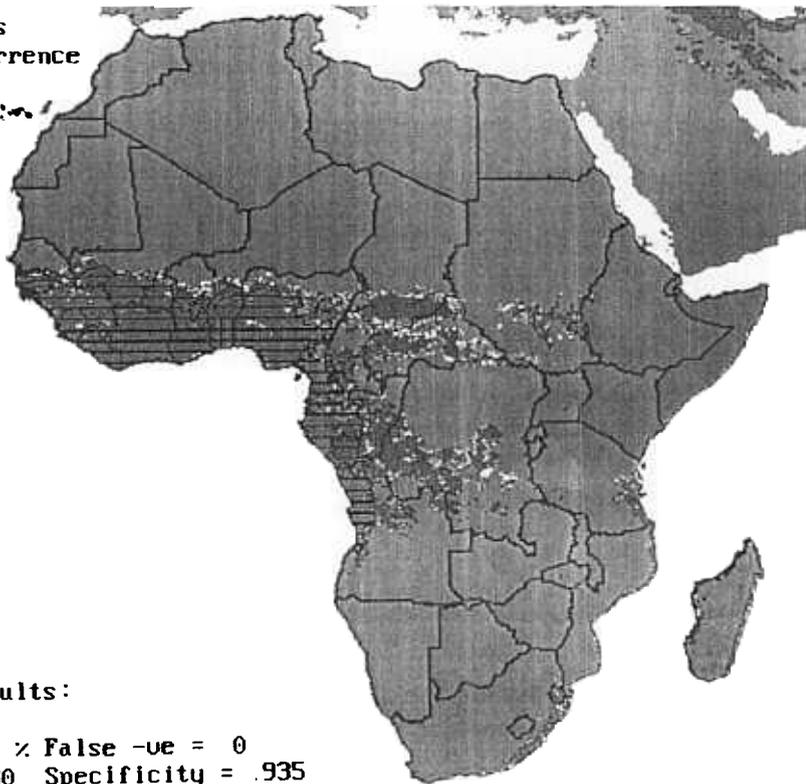


a

Species: *G. palpalis*
 Probability of occurrence
 ■ = 0.65 - 1.0
 ■ = 0.55 - 0.649
 ■ = 0.50 - 0.549
 ■ = 0.45 - 0.499
 ■ = 0.35 - 0.449
 ■ = 0.0 - 0.349
 ≡ = Actual

Vars (S,pp,3/3)
 DEM
 Nm/Cm
 NDx
 CDx
 NDa2
 Na/Ch4a
 Ch4n
 Np-Cp
 CDa1
 CDa3

Training data results:
 % Correct = 94
 % False +ve = 5 % False -ve = 0
 Sensitivity = .980 Specificity = .935



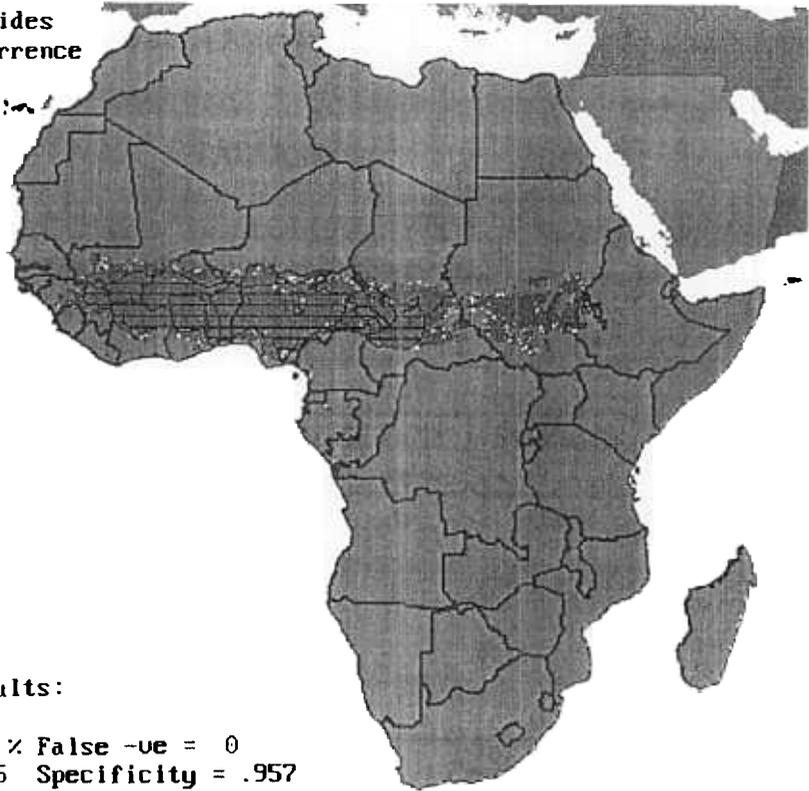
b

23) Observed (black lines) and predicted distributions of a) *G. morsitans* and b) *G. palpalis* (*s.l.*) in Africa. The predicted distributions are on a colour scale from red (low probability) to green (high probability).

Species: *G. tachinoides*
 Probability of occurrence
 ■ = 0.65 - 1.0
 ■ = 0.55 - 0.649
 ■ = 0.50 - 0.549
 ■ = 0.45 - 0.499
 ■ = 0.35 - 0.449
 ■ = 0.0 - 0.349
 ≡ = Actual

Vars (S,pp,3/3)
 DEM
 Nm/Cm
 CCDa0
 NDa0
 Ch4n
 CDn
 Na/Ch4a
 CDa2
 CDa1
 NDr

Training data results:
 % Correct = 96
 % False +ve = 4 % False -ve = 0
 Sensitivity = .985 Specificity = .957

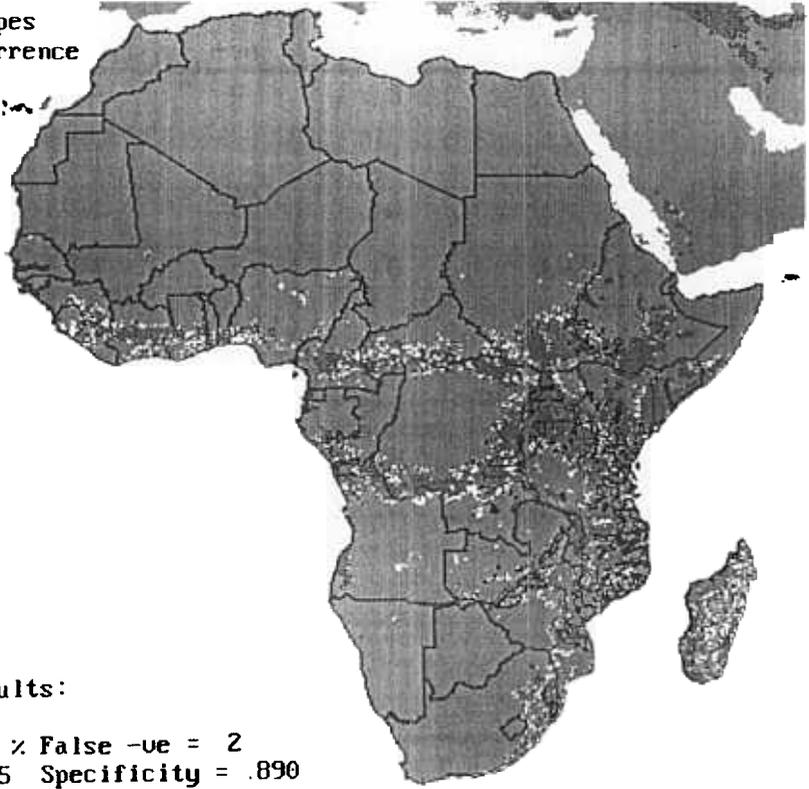


c

Species: *G. pallidipes*
 Probability of occurrence
 ■ = 0.65 - 1.0
 ■ = 0.55 - 0.649
 ■ = 0.50 - 0.549
 ■ = 0.45 - 0.499
 ■ = 0.35 - 0.449
 ■ = 0.0 - 0.349
 ≡ = Actual

Vars (S,pp,3/3)
 DEM
 NDr
 Ch4a1
 Np-Cp
 Na/Ch4a
 NDa1
 NDa2
 CDa3
 NDa0
 Nm/Cm

Training data results:
 % Correct = 89
 % False +ve = 9 % False -ve = 2
 Sensitivity = .865 Specificity = .890



d

23) Observed (black lines) and predicted distributions of c) *G. tachinoides* and d) *G. pallidipes* in Africa. The predicted distributions are on a colour scale from red (low probability) to green (high probability).

TABLE 11

Summary of the predictive accuracies of the discriminant analytical techniques applied to four tsetse species in West and East Africa and throughout the continent.

Glossina morsitans

	% Correct	% False +ve	% False -ve	Sensitivity	Specificity
WEST(1,1) G. m. submorsitans	68	28	4	0.89	0.55
EAST(2,2) G. m. submorsitans, G. m. centralis and G. m. morsitans	81	18	1	0.96	0.79
ALL(3,3)	87	12	1	0.97	0.86

Glossina palpalis

	% Correct	% False +ve	% False -ve	Sensitivity	Specificity
WEST(1,1) G. p. palpalis G. p. gambiensis	86	9	5	0.94	0.58
EAST(2,2)	94	6	0	0.98	0.93
ALL(3,3)	94	6	0	0.98	0.94

Glossina pallidipes

	% Correct	% False +ve	% False -ve	Sensitivity	Specificity
WEST(1,1)	NP	NP	NP	NP	NP
EAST(2,2)	86	9	5	0.72	0.89
ALL(3,3)	89	9	2	0.87	0.89

Glossina tachinoides

	% Correct	% False +ve	% False -ve	Sensitivity	Specificity
WEST(1,1)	77	18	5	0.88	0.69
EAST(2,2)	NP	NP	NP	NP	NP
ALL(3,3)	96	4	0	0.99	0.96

Predictions of *G. morsitans* density in Cote d'Ivoire

Using variables (S+=p)

- NDp-CDp
- ND/CD
- Tha1
- CCDp2
- THp2
- CH4range
- NDa2
- CCDmax
- CCDa2
- NDp2



Scale

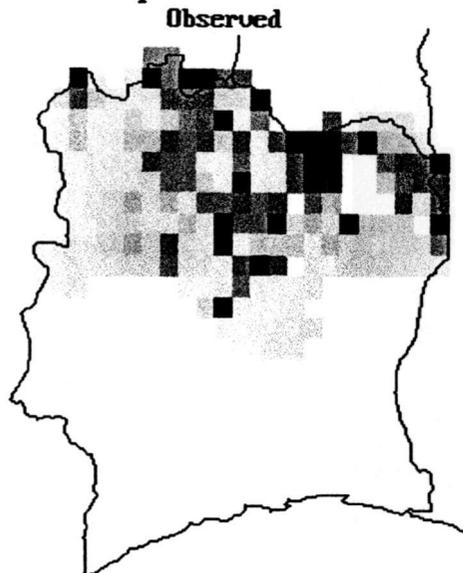
- = 0.51 ~ 12.18
- = 0.08 ~ 0.50
- = 0.02 ~ 0.07
- = 0.01 ~ 0.01
- = 0.00 ~ 0.00

a

Predictions of *G. tachinoides* density in Cote d'Ivoire.

Using variables (S+=p)

- NDm/CH4m
- ELEV
- Tha2
- NDp2
- CCDmin
- ND/CD
- CCDrange
- NDrange
- NDa3
- CH4range



Scale

- = 0.31 ~ 8.53
- = 0.10 ~ 0.30
- = 0.03 ~ 0.09
- = 0.01 ~ 0.02
- = 0.00 ~ 0.00

b

24) Observed and predicted abundances of a) *G. morsitans* and b) *G. tachinoides* in northern Côte d'Ivoire. The scale is Apparent Density, flies per trap per 'day'. The satellite predictor variables are listed in each figure.

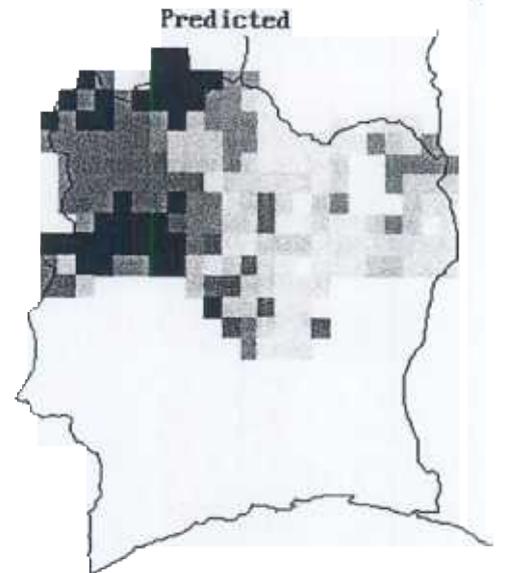
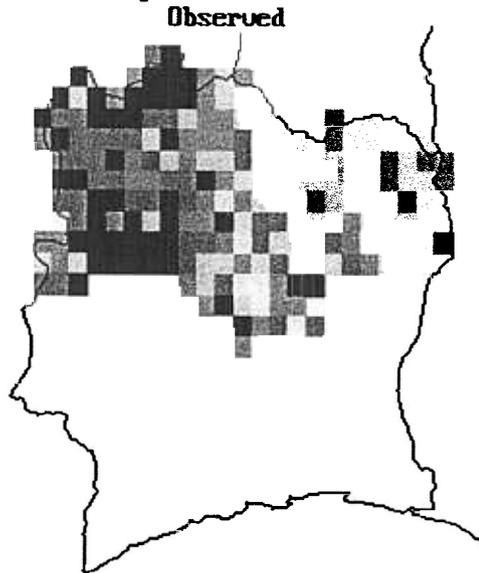
Predictions of *G. palpalis* density in Cote d'Ivoire

Using variables (S+=p)

THp3
Tha2
NDmax
CCDa3
CH4min
CH4Max
CCDp1
CCDa1
THp2
NDa3

Scale

■ = 2.57 ~ 7.45
■ = 1.53 ~ 2.53
■ = 0.86 ~ 1.49
■ = 0.39 ~ 0.85
■ = 0.00 ~ 0.38



c

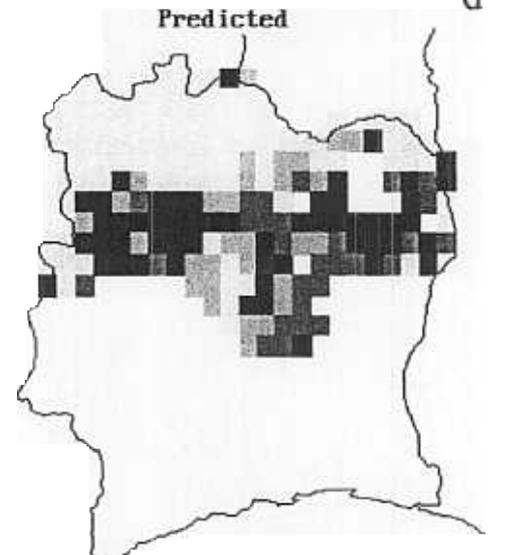
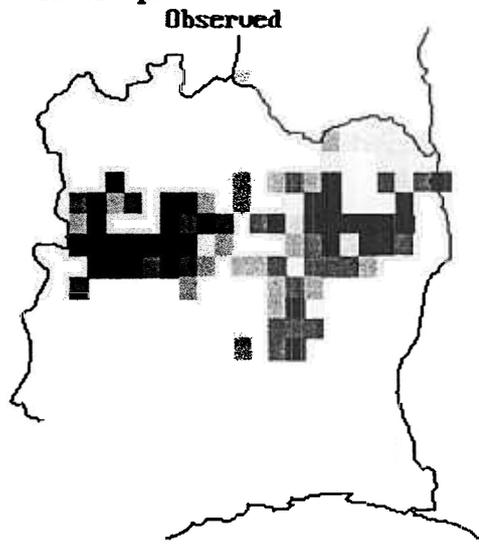
Predictions of *G. longipalpis* density in Cote d'Ivoire.

Using variables (S+=p)

CCDa2
CCDa1
NDp-CDp
Tha2
CCDp2
NDa1/CH4a1
CH4min
CCDa3
Tha0
CCDp3

Scale

■ = 0.25 ~ 2.62
■ = 0.09 ~ 0.20
■ = 0.03 ~ 0.08
■ = 0.01 ~ 0.02
■ = 0.00 ~ 0.00



d

24) Observed and predicted abundances of c) *G. palpalis* and d) *G. longipalpis* in northern Côte d'Ivoire. The scale is Apparent Density, flies per trap per 'day'. The satellite predictor variables are listed in each figure.

TABLE 12

Tsetse abundance

TABLE 13

The 10 most important predictor variables used to describe the abundance of tsetse (flies per trap per day) in northern Côte d'Ivoire, and the accuracy of the predictions of the abundance classes (five for all species) (from Rogers, Hay and Packer 1996).

Rank	<i>G. morsitans</i>	<i>G. longipalpis</i>	<i>G. palpalis</i>	<i>G. tachinoides</i>				
1	NDp-CCDp	CCDa2	CH4p3	NDm/CH4m				
2	NDm/CCDm	CCDa1	CH4a2	DEM				
3	CH4a1	NDp-CCDp	NDmax	CH4a2				
4	CCDp2	CH4a2	CCDa3	NDp2				
5	CH4p2	CCDp2	CH4min	CCDmin				
6	CH4range	NDa1/CH4a1	CH4max	NDm/CCDm				
7	NDa2	CH4min	CCDp1	CCDrange				
8	CCDmax	CCDa3	CCDa1	Ndrange				
9	CCDa2	CH4mean	CH4p2	NDa3				
10	NDp2	CCDp3	NDa3	CH4range				
Abundance	0.0-0.0	58	0.0-0.0	81	0.0-0.38	78	0.0-0.0	65
%Accuracy	0.01-0.01	100	0.01-0.02	80	0.39-0.85	50	0.01-0.02	78
	0.02-0.07	72	0.03-0.08	76	0.86-1.49	30	0.03-0.09	78
	0.08-0.5	51	0.09-0.20	94	1.53-2.53	59	0.10-0.30	62
	0.51-12.2	63	0.25-2.62	83	2.57-7.45	58	0.31-8.53	65

Disease distribution

Few data sets are available on the spatial distribution of trypanosomosis, either in cattle or humans. Our project benefitted from collaborative links with Project GCP/TOG/013/BEL and Dr. Guy Hendrickx who has sampled the prevalence of trypanosomosis on a gridded basis throughout Togo. These data, when subjected to discriminant analysis using satellite and other field data may be described with considerable accuracy, as shown in Fig. 25a. Using the same satellite variables, the predicted 'high risk' areas for Togo and adjacent countries (Ghana and Benin) are shown in Fig. 25b

The impact of tsetse and trypanosomosis on agriculture

The impact of tsetse on cattle can be found by comparing the cattle densities in areas where flies are present with those where flies are absent. Using the ERGO data set for the central band of Nigeria, it can be shown that cattle numbers in erstwhile *G. morsitans* areas are much lower than in tsetse-free areas (Rogers 1993). A number of reasons may be given for this. Either the flies have a direct impact on cattle, or the associated agricultural activities are very much less in areas known to be infested with tsetse, such activities generally being associated with increased cattle numbers. Whilst the first explanation seems reasonable (though difficult to prove within the ERGO data set) the second may be examined by comparing survey results at different points in time. The 1976 SLAR survey of vegetation cover in Nigeria has been compared with the RIM 1990 aerial survey in several reports (*e. g.* ERGO 1994). Fig. 26 a) shows the annual percentage change in agricultural activity between these two dates (the problems of relating SLAR to aerial survey data are detailed in the ERGO report). There appears to be at least a visual correlation between areas of agricultural increase and areas previously inhabited by *G. morsitans* in the middle and North of the country, although there are also fly-free areas which experienced an increase (*e. g.* in the North-East) and fly-infested areas showing a decrease (*e. g.* in the South-West). Clearly, therefore, factors in addition to tsetse have influenced the changes shown in Fig. 26a, amongst which are the increased sedenterisation in the North of Nigeria of the once nomadic Fulani pastoralists and the general abandonment of agriculture in parts of the South, associated with increased wealth derived from oil. One methodological problem with the interpretation of Fig. 26 a) is that only those areas that had a low level of agricultural activity in the 1970's could have shown a high rate of increase up to 1990, since areas already heavily cultivated in 1976 could not increase their coverage beyond 100%. Fig. 26 b) was produced in an attempt to overcome this problem, and shows the intrinsic rate of increase of agricultural activity between the two sample dates. This figure assumes that agricultural activity in an area starts from a base of zero and increases logistically up to some maximum, arbitrarily set to 80% (other maxima do not greatly affect the spatial patterning of the results). The intrinsic rate of increase should be insensitive to the actual level of cultivation seen in 1976, because it represents the maximum possible growth rate of agriculture by region, not the realised growth rate (the difference is explained in many standard ecological texts). Again, as in Fig. 26 a), there appears to be an association between high intrinsic growth rates and *G. morsitans* habitats (with the regional exceptions noted previously). Taken together, these two maps suggest that *G. morsitans* has had a significant impact on agricultural activity in Nigeria (Fig. 26 a)), and that it has done so by limiting agricultural development in potentially productive areas (Fig. 26 b)).

Acknowledgements

Many people contributed to the work described in this Report and all in the TALA project are grateful to them. We should particularly like to thank Brian Hursey, Jan Slingenbergh, Tony Piccolo and Fred Snijders of FAO, Guy Hendrickx and Dr. A. Napala of Project GCP/TOG/013/BEL in Togo, Professor Maggie Gill and Dr. Jonathan Wadsworth at NRI, Drs A. Belward, D. Ehrlich, E. F. Lambin and J.-P. Malingreau at the Joint Research Centre, Ispra, Italy, Drs C. O. Justice and C. J. Tucker and the GIMMS group at NASA GSFC, Ms. Kitti van Ramshorst for ARC-INFO advice and support, and the many people associated with the Pathfinder program that has made our project such an enjoyable one.

Finally we acknowledge the untiring support of Reg Allsopp, who battled to get us going in the first place.

References

- Achard, F. & Blasco, F. (1990) Analysis of vegetation seasonal evolution and mapping of forest cover in West Africa with the use of NOAA AVHRR HRPT data. *Photogrammetric Engineering and Remote Sensing* 56, 1359-1365.
- Agbu, P. A. & James, M. E. (1994). *The NOAA/NASA Pathfinder AVHRR Land Data Set User's Manual*. Version 3.1. NASA, Goddard Distributed Active Archive Centre: Greenbelt, Maryland, USA.
- Andres, L., Salas, W.A. & Skole, D. (1994) Fourier analysis of multi-temporal AVHRR data applied to a land cover classification. *International Journal of Remote Sensing* 15, 1115-1121.
- Anon (1982). *Cartographie de la repartition des Glossines*. Maisons-Alfort, France: Institut d'Élevage et de Médecine Vétérinaire des Pays Tropicaux.
- Asrar, G. & Greenstone, R. (1995). *The Mission to Planet Earth (MTPE) Earth Observing System (EOS) reference handbook*. NASA, Goddard Space Flight Centre: Greenbelt, Maryland, USA.
- Becker, F. & Z.-L. Li. (1990). Towards a local split window method over land surfaces. *International Journal of Remote Sensing* 11, 369-393.
- Becker, F. & Z.-L. Li. (1995). Surface temperature and emissivity at various scales: definition, measurement and related problems. *Remote Sensing Reviews* 12, 225-253.
- Benedetti, R., Rossini, P. & Taddei, R. (1994) Vegetation classification in the Middle Mediterranean area by satellite data. *International Journal of Remote Sensing* 15, 583-596.
- Boyd, D. S., Foody, G. M., Curran, P. J., Lucas, R. M. & Honzak, M. (1996). An assessment of radiance in Landsat TM middle and thermal infrared wavebands for the detection of tropical forest regeneration. *International Journal of Remote Sensing* 17, 249-261.
- Bourn, D. and W. Wint (1994). Livestock, land use and agricultural intensification in sub-Saharan Africa. *Pastoral Development Network* 37a, 1-22.
- Bourn, D., Wint, W., Blench, R. and Wolley, E. (1994) Nigerian livestock resources survey. *World Animal Review* 78, 49-58.

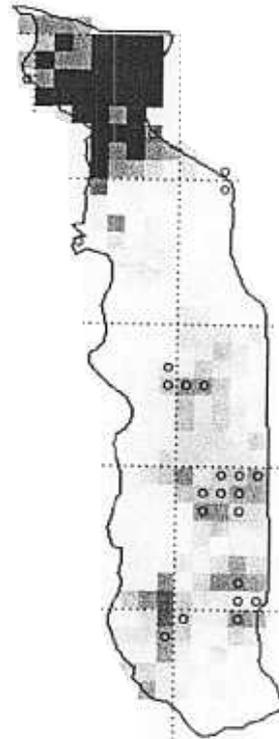
Predictions of areas in Togo with PPRUA *T. congolense* > 10

Using variables
 NDph2
 NDamp3
 Tmn
 Tmn
 NDamp1
 Tmx
 NDph3
 NDVIx
 NDVIx
 NDamp2

Probability scale

- 0.90 - 0.99
- 0.80 - 0.89
- 0.70 - 0.79
- 0.60 - 0.69
- 0.50 - 0.59
- 0.40 - 0.49
- 0.30 - 0.39
- 0.20 - 0.29
- 0.10 - 0.19
- 0.00 - 0.09
- Observed

% Correct = 81.03
 % False +ve = 14.47
 % False -ve = 4.50
 Sensitivity = .659
 Specificity = .833



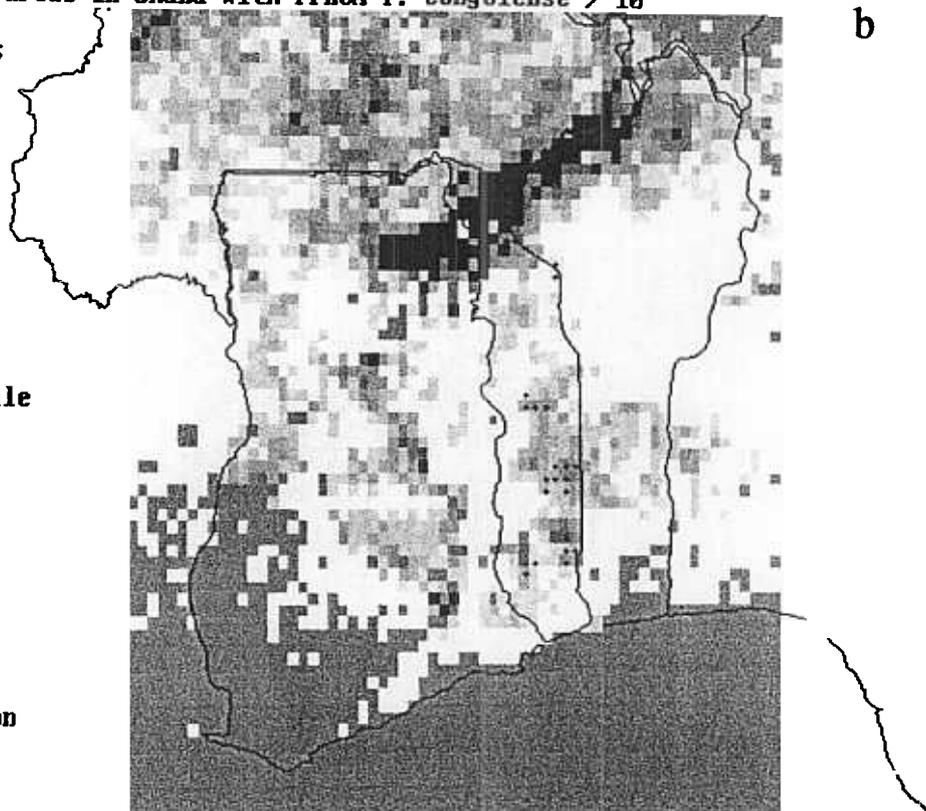
a

Predictions of areas in Ghana with PPRUA *T. congolense* > 10

Using variables
 NDph2
 NDamp3
 Tmn
 Tmn
 NDamp1
 Tmx
 NDph3
 NDVIx
 NDVIx
 NDamp2

Probability scale

- 0.90 - 0.99
- 0.80 - 0.89
- 0.70 - 0.79
- 0.60 - 0.69
- 0.50 - 0.59
- 0.40 - 0.49
- 0.30 - 0.39
- 0.20 - 0.29
- 0.10 - 0.19
- 0.00 - 0.09
- No Prediction
- Observed

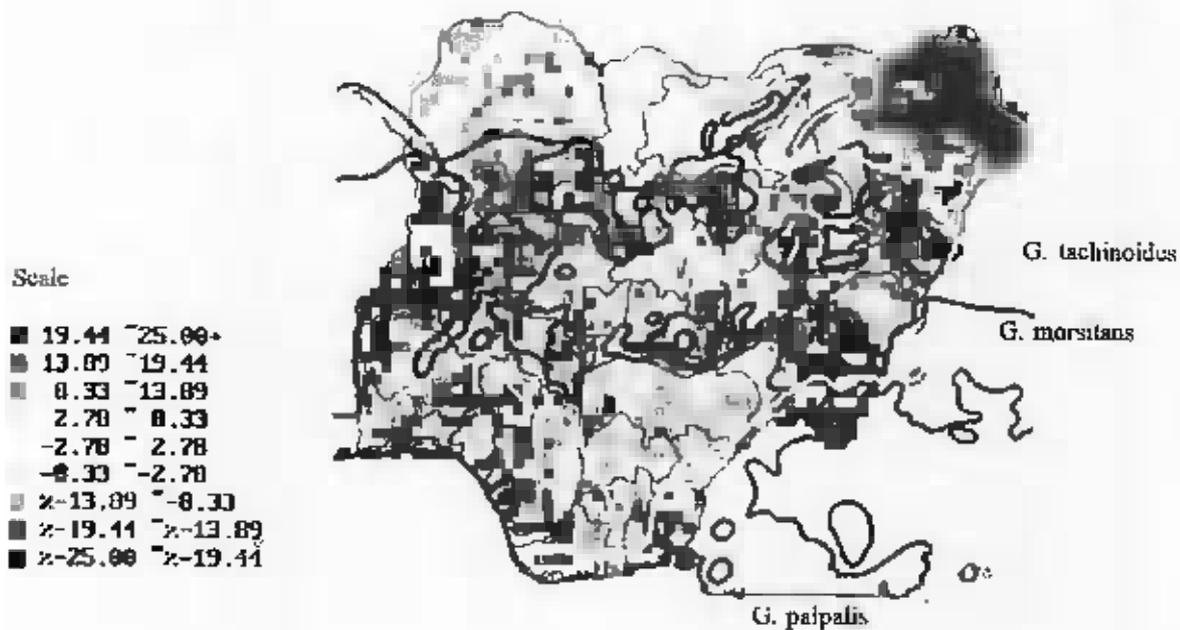


b

25) a) Observed (red circles) and predicted (green shading) of the prevalence of *Trypanosoma congolense* in cattle in Togo (data courtesy of Dr. Guy Hendrickx). b) Predicted high risk areas for *T. congolense* in Ghana, Togo and Benin using the same satellite predictor variables.

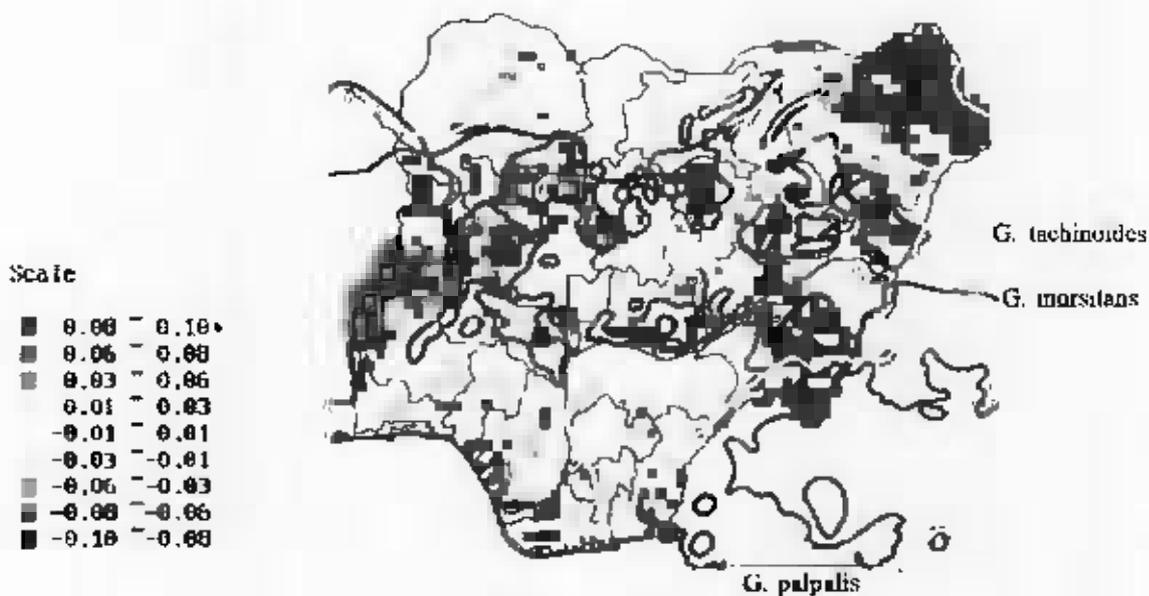
Vegetation survey results for Nigeria. Annual % change from 1976 base.
Map of Cultivation

a



Vegetation survey results for Nigeria. Intrinsic rate of increase to $K = 80\%$
Map of Cultivation

b



26) Two measures of the change in Nigerian cultivation between 1976 and 1990 (original data from ERGO) a) the annual percentage change (red = decrease, green = increase), b) the calculated intrinsic rate of increase, assuming a logistic rise in cultivation to a habitat maximum of 80%. The historical distribution limits of three tsetse species are also shown.

- Brown, S. A., Folk, M., Goucher G. & Rew, R. (1993). Software for portable scientific data management. *Computers in Physics* 7, 304-308.
- Brown, J.F., Loveland, T.R., Merchant, J.W., Reed, B.C. & Ohlen, D.O. (1993) Using multi-source data in global land-cover characterization: concepts, requirements, and methods. *Photogrammetric Engineering and Remote Sensing* 59, 977-987.
- Congalton, R.G. (1991) A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of the Environment* 37, 35-46.
- Cooper, D. I. & Asrar, G. M. (1989). Evaluating atmospheric correction models for retrieving surface temperatures from the AVHRR over a tallgrass prairie. *Remote Sensing of Environment* 27, 93-102.
- Corves, C. & Place, C.J. (1994) Mapping the reliability of satellite-derived landcover maps - an example from the Central Brazilian Amazon Basin. *International Journal of Remote Sensing* 15, 1283-1294.
- Cushnie, J.L. (1987) The interactive effect of spatial resolution and degree of internal variability within land-cover types on classification accuracies. *International Journal of Remote Sensing* 8, 15-29.
- Dalu, G. (1986). Satellite remote sensing of atmospheric water vapor. *International Journal of Remote Sensing*, 7: 1089-1097.
- DeFries, R.S. & Townshend, J.R.G. (1994) NDVI-derived land cover classifications at a global scale. *International Journal of Remote Sensing* 15, 3567-3586.
- Eastman, J. R. & Fulk, M. (1993). Long sequence time series evaluation using standard principal components. *Photogrammetric Engineering and Remote Sensing* 59, 991-996.
- Ehrlich, D. & Lambin, E.F. (1996) Broad scale land-cover classification and interannual climatic variability. *International Journal of Remote Sensing*, 17, 845-862.
- Environmental Research Group Oxford. (1994). *Land use change in Nigeria - 1976-1990*. A report prepared by ERGO for the United Nations Environmental Programme (UNEP) and the Federal Environmental Protection Agency (FEPA): Abuja, Nigeria.
- Estes, J.E. & Mooneyhan, D.W. (1994) Of maps and myths. *Photogrammetric Engineering and Remote Sensing* 60, 517-524.
- FAO/UNESCO (1977). Soil Map of the World 1:5,000,000. Volum VI, Africa. UNESCO, Paris.
- Fitzgerald, R.W. & Lees, B.G. (1994) Assessing the classification accuracy of multi-source remote sensing data. *Remote Sensing of the Environment* 47, 362-368.
- Flasse, S. & Verstraete, M. M. (1994). Monitoring the environment with vegetation indices, comparison of NDVI and GEMI using AVHRR data over Africa. In *Vegetation, Modelling and Climate Change Effects*, editors Veroustraete, F. and Ceulemans, R., pp. 107-135. SPB Academic Publishing: The Hague, The Netherlands.
- Foody, G.M. (1992). Classification accuracy assessment: some alternatives to the Kappa coefficient for nominal and ordinal level classifications. Pp 529-538 in *Remote Sensing: from Research to Operation*. Proceedings of the 18th Annual Conference of the Remote Sensing Society. (ed. A.P. Cracknell and R.A. Vaughan).
- Foody, G.M., McCulloch, M.B. & Yates, W.B. (1995) Classification of remotely sensed data by an artificial neural network: issues related to training data characteristics. *Photogrammetric Engineering and Remote Sensing* 61, 391-401.
- Foody, G. M., Boyd, D. S. & Curran, P. J. (1996). Relations between tropical forest biophysical properties and data acquired in AVHRR channels 1 - 5. *International Journal of Remote Sensing* 17, 1341-1355.

- Ford, J. & Katondo, K. M. (1977). *The distribution of tsetse flies (Glossina) in Africa in 1973*. Organization of African Unity - Scientific and Technical Research Commission. Cook, Hammond and Kell: London, UK.
- Goward, S. N. & Prince, S. D. (1995). Transient effects of climate on vegetation dynamics - satellite observations. *Journal of Biogeography* **22**, 549-564.
- Green, P. E. (1978). *Analyzing Multivariate Data*. The Dryden Press: Hinsdale, Illinois, USA.
- Gutman, G. G. (1993). Multi-annual time series of AVHRR-derived land surface temperature. *Advances in Space Research* **14**, 27-30.
- Hay, S.I. (1996). An investigation of the utility of remotely sensed meteorological satellite data for predicting the distribution and abundance of the tsetse fly (Diptera: Glossinidae). D. Phil thesis, Oxford University (submitted).
- Hay, S.I., Tucker, C.J., Rogers, D.J. & Packer, M.J. (1996) Remotely sensed surrogates of meteorological data for the study of the distribution and abundance of arthropod vectors of disease. *Annals of Tropical Medicine and Parasitology* **90**, 1-19.
- Huete, A. R. (1988). A soil adjusted vegetation index (SAVI). *Remote Sensing of Environment* **25**, 295-309.
- Jackson, R. D. & Huete, A. R. (1991). Interpreting vegetation indices. *Preventive Veterinary Medicine*, **11**: 185-200.
- Jeyaseelan, A. T. & Thiruvengadachari, S. (1993). Suspected Mt. Pinatubo aerosol impact on the NOAA AVHRR NDVI over India. *International Journal of Remote Sensing* **14**, 603-608.
- Kidwell, K. B. (1995). *NOAA polar orbiter data users guide (TIROS-N, NOAA-6, NOAA-7, NOAA-8, NOAA-9, NOAA-10, NOAA-11, NOAA-12, NOAA-13 and NOAA-14)*. National Oceanic and Atmospheric Administration: Washington DC., USA.
- Lambin, E.F. & Ehrlich, D. (1995) Combining vegetation indices and surface temperature for land-cover mapping at broad spatial scales. *International Journal of Remote Sensing*, **16**, 573-579.
- James, M. E. & Kalluri, S. N. V. (1994). The pathfinder AVHRR land data set - an improved coarse resolution data set for terrestrial monitoring. *International Journal of Remote Sensing* **15**, 3347-3363.
- Lambin, E.F. & Strahler, A.H. (1994a). Change-vector analysis: a tool to detect and categorize land-cover change processes using high temporal resolution satellite data. *Remote Sensing of the Environment* **48**, 231-244
- Lambin, E. F. & Strahler, A.H. (1994b). Indicators of land-cover change for change-vector analysis in multi-temporal space at coarse spatial scales. *International Journal of Remote Sensing* **15**, 2099-2119.
- Laveissiere, C. & Challier, A. (1977). Notice Explicative No 69. La repartition des Glossines en Haute-Volta, cartes à 1/2 000 000. Office de la Recherche Scientific et Technique Outre-Mer (ORSTOM): Paris, France.
- Laveissiere, C. & Challier, A. (1981). Notice Explicative No 89. La repartition des Glossines en Côte d'Ivoire, cartes à 1/2 000 000. Office de la Recherche Scientific et Technique Outre-Mer (ORSTOM): Paris, France.
- Lark, R.M. (1995) Components of accuracy of maps with special reference to discriminant analysis on remote sensor data. *International Journal of Remote Sensing* **16**, 1461-1480.
- Lillesand, T.M. & Kiefer, R.W. (1994) *Remote Sensing and Image Interpretation*, 3rd edition. John Wiley & Sons, New York, 750 pp.

- Los, S. O. (1993). Calibration adjustment of the NOAA AVHRR normalized difference vegetation index without recourse to component channel 1 and 2 data. *International Journal of Remote Sensing* **14**, 1907-1917.
- Ma, Z. & Redmond, R.L. (1995) Tau coefficients for accuracy assessment of classification of remote sensing data. *Photogrammetric Engineering and Remote Sensing* **61**, 435-439.
- Malingreau, J. P. & Belward A. S. (1994). Recent activities in the European Community for the creation and analysis of global AVHRR data sets. *International Journal of Remote Sensing* **15**, 3397-3416.
- Marsh, S.E., Walsh, J.L., Lee, C.T., Beck, L.R. & Hutchinson, C.F. (1992) Comparison of multi-temporal NOAA-AVHRR and SPOT-XS satellite data for mapping land-cover dynamics in the west African Sahel. *International Journal of Remote Sensing* **13**, 2997-3016.
- Millington, A.C., Critchley, R.W., Douglas, T.D. & Ryan, P. (1994) *Estimating woody biomass in sub-Saharan Africa* (Washington DC: The World Bank).
- Moody, A. & Woodcock, C.E. (1994) Scale-dependent errors in the estimation of land-cover proportions: implications for global land-cover datasets. *Photogrammetric Engineering and Remote Sensing* **60**, 585-594.
- National Oceanic and Atmospheric Administration. (1972). *Substation observations. National Weather Service (NWS) observing handbook*. No. 2. Government Printing Office: Washington DC., USA.
- National Center for Supercomputing Applications. (1990). *Hierarchical Data Format (HDF) Users Guide*. NCSA, University of Illinois: Urbana-Champaign, USA.
- National Oceanic and Atmospheric Administration. (1990). *Monthly climate data for the world. Prepared in co-operation with the World Meteorological Organisation*. Volume 41, 42 and 43; numbers 1-12. National Climate Data Centre (NCDC), NOAA: North Carolina, USA.
- Pinty, B. & Verstraete, M. M. (1992). GEMI, a non-linear index to monitor global vegetation from satellites. *Vegetatio* **101**, 15-20.
- Prata, A. J. (1993). Land surface temperatures derived from the advanced very high resolution radiometer and the along-track scanning radiometer 1. Theory. *Journal of Geophysical Research* **98**, 16,689-16,702.
- Prata, A.J. & Platt, C.M.R. (1991). Land surface temperature measurements from the AVHRR. pp 433-438 in *Proceedings of the 5th AVHRR Data Users Meeting*, EUMETSAT, Darmstadt.
- Price, J. C. (1983). Estimating surface temperatures from satellite thermal infrared data - a simple formulation for the atmospheric effect. *Remote Sensing of Environment* **13**, 353-361.
- Price, J. C. (1984). Land surface temperature measurement for the split window channels of the NOAA 7 advanced very high resolution radiometer. *Journal of Geophysical Research* **89**, 7231-7237.
- Prince, S. D. & Goward S. N. (1995). Global primary production - a remote-sensing approach. *Journal of Biogeography* **22**, 815-835.
- Resource Inventory and Management. (1992). *Nigerian livestock resources*. Four Volumes. Report prepared by RIM Ltd. for the Federal Department of Livestock and Pest Control Services, Abuja, Nigeria.
- Rogers, D. J. (1991). Satellite imagery, tsetse and trypanosomiasis in Africa. *Preventive Veterinary Medicine* **11**, 201-220.

- Rogers D.J. (1993). Consultant Report to FAO. Unpublished.
- Rogers, D.J., Hay, S.I. & Packer, M.J. (1996) Predicting the distribution of tsetse flies in West Africa using temporal Fourier processed meteorological satellite data. *Annals of Tropical Medicine and Parasitology* **90**, 225-241.
- Rogers, D. J. & Randolph, S. E. (1991). Mortality rates and population density of tsetse flies correlated with satellite imagery. *Nature* **351**, 739-741.
- Rogers, D. J. & Williams, B. G. (1994). Tsetse distribution in Africa, seeing the wood and the trees. In *Large Scale Ecology and Conservation*. Edited by P. J. Edwards., R. M. May and N. R. Webb, pp. 249-273. Blackwell Scientific Publications: Oxford, UK.
- Samson, S.A. (1993) Two indices to characterize temporal patterns in the spectral response of vegetation. *Photogrammetric Engineering and Remote Sensing* **59**, 511-517.
- Singh, A. (1989). Digital change detection techniques using remotely sensed data. *International Journal of Remote Sensing* **10**, 989-1003.
- Smith, W. L. (1966). Note on the relationship between total precipitable water and the surface dew point. *Journal of Applied Meteorology* **5**, 726-727.
- Snijders, F. L. (1991). Rainfall monitoring based on Meteosat data - a comparison of techniques applied to the Western Sahel. *International Journal of Remote Sensing* **12**, 1331-1347.
- Steinwand, D. R., Hutchinson, J. A., & Snyder, J. P. (1992). *Map projections from global and continental data sets, and an analysis of distortion caused by reprojection*. United States Geological Survey EROS Data Centre (USGS/EDC) contract report, EROS Data Center, Sioux Falls, USA.
- Steinwand, D. R. (1994). Mapping raster imagery to the Interrupted Goode Homolosine Projection. *International Journal of Remote Sensing* **15**, 3463-3471.
- Stone, T.A., Schlesinger, P., Houghton, R.A. & Woodwell, G.M. (1994) A map of the vegetation of South America based on satellite imagery. *Photogrammetric Engineering and Remote Sensing* **60**, 541-551.
- Stowe, L. L., Carey, R. M. & Pellegrino, P. P. (1992). Monitoring the Mt-Pinatubo aerosol layer with NOAA-11 AVHRR data. *Geophysical Research Letters* **19**, 159-162.
- Tatsuoka, M. M. (1971). *Multivariate Analysis: techniques for educational and psychological research*. John Wiley & Sons: New York, USA.
- Townshend, J.R.G. (1992) Land cover. *International Journal of Remote Sensing* **13**, 1319-1328.
- Sugita, M. & Brutsaert W. (1993). Comparison of land surface temperatures derived from satellite observations with ground truth during FIFE. *International Journal of Remote Sensing* **14**, 1659-1676.
- Townshend, J.R.G & Justice, C.O. (1986) Analysis of the dynamics of African vegetation using the normalised difference vegetation index. *International Journal of Remote Sensing* **7**, 1435-1445.
- Townshend, J. R. G., Justice, C. O. & Kalb, V. (1987). Characterization and classification of South American land cover types using satellite data. *International Journal of Remote Sensing* **8**, 1189-1207.
- Tucker, C. J., Townshend, J. R. G. & Goff, T. E. (1985). African land-cover classification using satellite data. *Science* **227**, 369-375.
- Unwin, D. M. (1980). *Microclimate measurement for ecologists*. Biological techniques series 3. Academic Press: London, UK.

- Van de Griend, A. A. & Owe, M. (1993). On the relationship between thermal emissivity and the normalized difference vegetation index for natural surfaces. *International Journal of Remote Sensing* **14**, 1119-1131.
- Van Deusen, P.C. (1995) Modified highest confidence first classification. *Photogrammetric Engineering and Remote Sensing* **61**, 419-425.
- White, F. (1983). The vegetation of Africa - a descriptive memoir to accompany the UNESCO/AETFAT/UNSO vegetation map of Africa. United Nations Educational, Scientific and Cultural Organisation: Paris, France.