An Investigation of the Utility of Remotely Sensed Meteorological Satellite Data for Predicting the Distribution and Abundance of the Tsetse Fly (Diptera: Glossinidae).

by

Simon Iain Hay

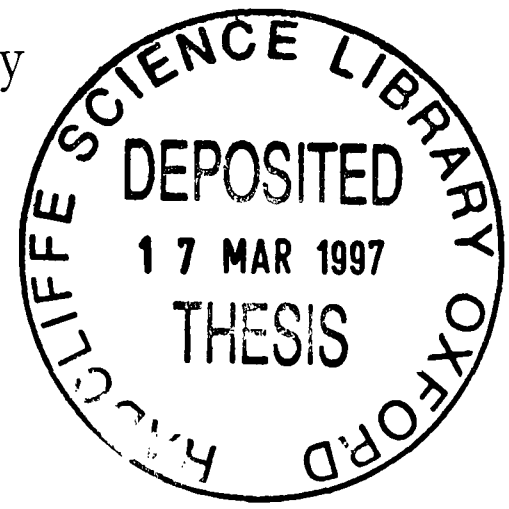
Green College

A thesis submitted in partial fulfilment
of the requirements for the degree of
Doctor of Philosophy

Department of Zoology

University of Oxford

Trinity Term 1996



An investigation of the utility of remotely sensed meteorological satellite data for predicting the distribution and abundance of the tsetse fly (Diptera: Glossinidae).

by

Simon I. Hay, Green College, University of Oxford

Submitted for the degree of Doctor of Philosophy,

Trinity Term 1996.

Abstract

This thesis investigates the potential contribution of data from the Advance Very High Resolution Radiometer (AVHRR) on-board the National Oceanic and Atmospheric Administrations (NOAA) polar-orbiting meteorological satellites and data from the High Resolution Radiometer (HRR) on-board the Meteosat geostationary meteorological satellites for predicting the distribution and abundance of the tsetse fly (Diptera: Glossinidae) in Africa. The images were processed to produce a range of monthly land surface temperature, atmospheric moisture and rainfall indices for the period 1988 to 1990. The performance of these indices, derived from several different methods, was tested using meteorological records collected during these years at stations across continental Africa and the most accurate used to form a refined dataset for subsequent analysis. The time-series of these land surface temperature, atmospheric moisture and rainfall indices and a range of Spectral Vegetation Indices (SVI) were subject to temporal Fourier analysis to parameterise the seasonal variation in these variables. These data, in combination with elevation information from a digital elevation model (DEM) were used to predict the land-cover of Nigeria determined independently by an aerial survey in 1990. The Normalised Difference Vegetation Index (NDVI) performed best and so was used in combination with the satellite proxy meteorological and DEM data to predict the distribution and abundance of eight tsetse fly species in Côte d'Ivoire and Burkina Faso, West Africa. The results are discussed in relation to the ecology of the different tsetse species. Conclusions are then drawn on the potential of such meteorological satellite data for remote tsetse fly population surveillance and, in the wider context, to the study and control of arthropod vectors of disease.

TABLE OF CONTENTS

MATERIALS AND METHODS

REMOTELY SENSED SURROGATES FOR METEOROLOGICAL VARIABLES

CONCLUSIONS AND GENERAL DISCUSSION

LIST OF FIGURES

LIST OF TABLES

Chapter 1

INTRODUCTION

" ... A bad beginning makes a bad ending ... "

Euripides.

1.1 General introduction

This thesis investigates the potential contribution of multitemporal coarse spatial resolution meteorological satellite imagery to epidemiological mapping, with particular reference to the distribution of tsetse fly species in West Africa. The hypothesis is that satellite systems can provide proxy meteorological variables of importance in describing disease vector distributions and abundances. The advantage of obtaining such surrogate environmental data from satellites is that their collection is on a daily basis, at the continental scale and is independent of the infrastructure of the region concerned. This is of particular advantage to those developing countries where vector-borne diseases are most prevalent and the resources to deal with the problem the least, making the need to target resources paramount. The work is important, because should these remotely sensed data prove useful in the mapping of disease vector distributions and abundances, it will add a powerful and cost effective new dimension to disease vector control campaigns.

The introduction to this thesis is presented in six parts. A knowledge of remote sensing is not assumed and so detailed information is presented on this discipline from the outset. The second section (1.2), provides an overview of remote sensing from its physical basis, through the interaction of electromagnetic radiation with the atmosphere, to details of the meteorological satellite platforms that have been used in epidemiological studies. Particular emphasis is placed on the satellite systems used in this work and illustrative examples from these systems are used throughout. Preliminary justification for the choice of these satellite systems is also provided.

Information is then presented on some of the standard procedures required to turn data derived from satellite sensors into information meaningful to ecologists and epidemiologists. This is given in section (1.3), dealing with how raw digital data are collected by satellite sensors, quality controlled and converted into geophysical values, and in section (1.4) describing how vegetation, temperature, atmospheric moisture and rainfall indices can be derived from these base geophysical data. The potential sources of error in these processes are also highlighted. It is intended for these sections to be in sufficient detail for those unfamiliar with remote sensing techniques to understand the following review (1.5) of their application to epidemiology and disease control.

The review in section 1.5 has been restricted to mosquito, tick and tsetse vectors, as most of the research has been directed toward these organisms. The section details application of both high and low-spatial resolution imagery to each of the vectors described and concludes with a critique of the present status of the work. This leads on to the final section (1.6) which describes the aims, objectives and scope of this study and how it might be seen to improve on previous applications.

Large parts of this document have been published or are in press. It was decided not to present these papers verbatim however, as the result would have been a less coherent document, with many repetitions. The details of what information has been published or is in press, with notes on authorship are given in Appendix I. Appendix II contains the papers that have been published and Appendix III provides a list of acronyms that are used in the thesis.

1.2 Remote sensing

Remote sensing is simply the process of acquiring data about an object from a distance. In the context of this work we consider the amount of radiant energy *i.e.* the magnitude of electromagnetic radiation (EMR), measured by satellite-borne radiometers, that arises from the Sun and is reflected and radiated from the Earth's surface and atmosphere. For a general background to remote sensing and its applications see, among others, Swain and Davis (1978), Colwell (1983a, b), Curran (1985), Sabins (1987), Asrar (1989), Cracknell and Hayes (1991), Barrett and Curtis (1992) and Lillesand and Kieffer (1994).

1.2.1 Electromagnetic radiation

Electromagnetic radiation is emitted by all objects above absolute zero (0 K, - 273 °C). The total amount of energy an object emits is expressed by the *Stefan-Boltzman law* which states that;

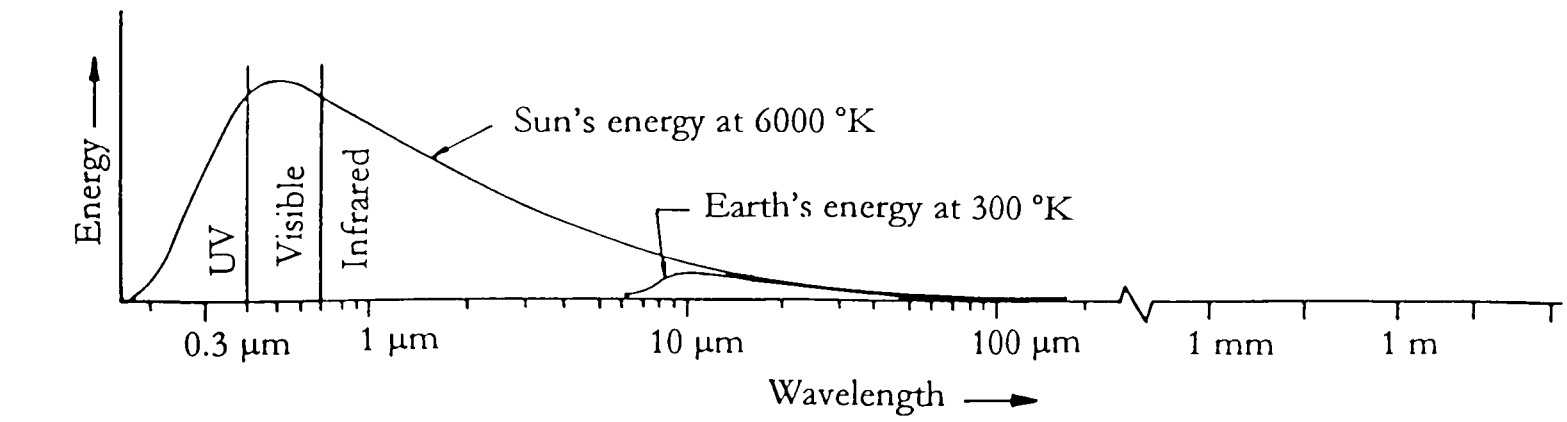
$$M = \sigma T^4$$

where, M, is the total exitance (emitted radiant flux per unit area) from the surface of the material (Wm^{-2}), σ, the *Stefan-Boltzman* constant (5.6×10^{-8} $Wm^{-2}K^{-4}$) and T, the absolute temperature of the emitting material (K). It can be seen therefore, that the total amount of energy emitted by an object increases rapidly with temperature. This phenomenon is demonstrated graphically in figure 1.1.1 by the larger area under the curve representing the electromagnetic spectrum (EMS) emitted by the Sun (at approximately 6000 K) than the corresponding area under the curve representing the EMS of the Earth (at approximately 300 K). The figure also demonstrates the *Wiens displacement law*, or the shift towards emission of shorter wavelengths by an object at higher temperatures (Colwell *et al.* 1963, Montieth and Unsworth 1990).

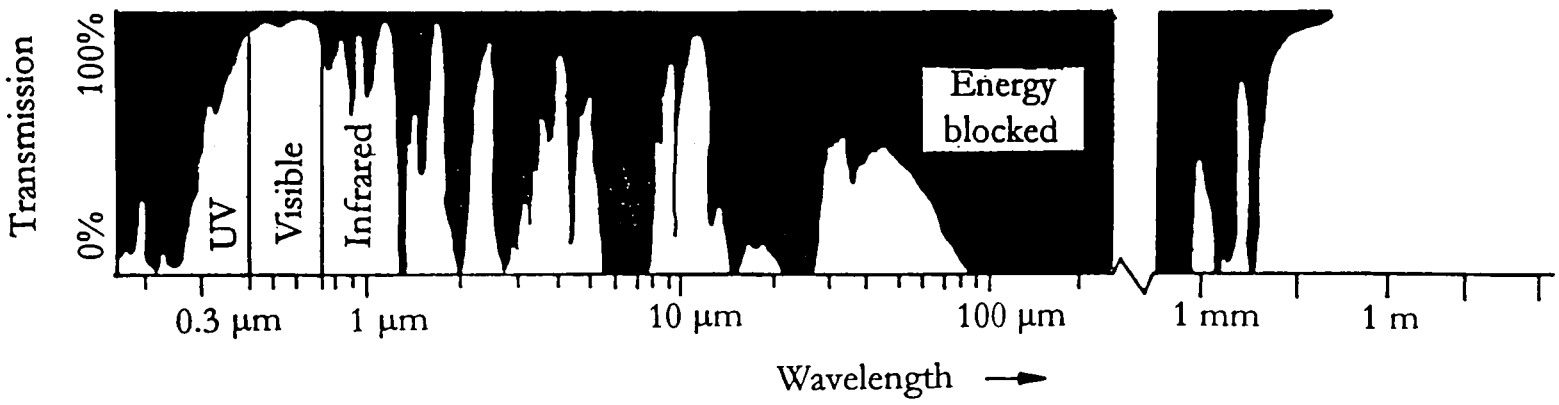
1.2.2 Atmospheric transmittance and radiometer design

Satellite-borne radiometers or sensors are instruments for measuring the intensity of EMR within a certain "band" or narrow range of wavelengths. The EMR they measure travels through the atmosphere which both scatters and absorbs EMR. These interactions are most significant close to the Earth's surface (~ 1 - 5 km) in the atmospheric boundary layer (ABL). Figure 1.1.2 shows that, due to such interactions, atmospheric transmission of EMR is wavelength dependent. Consequently, radiometers

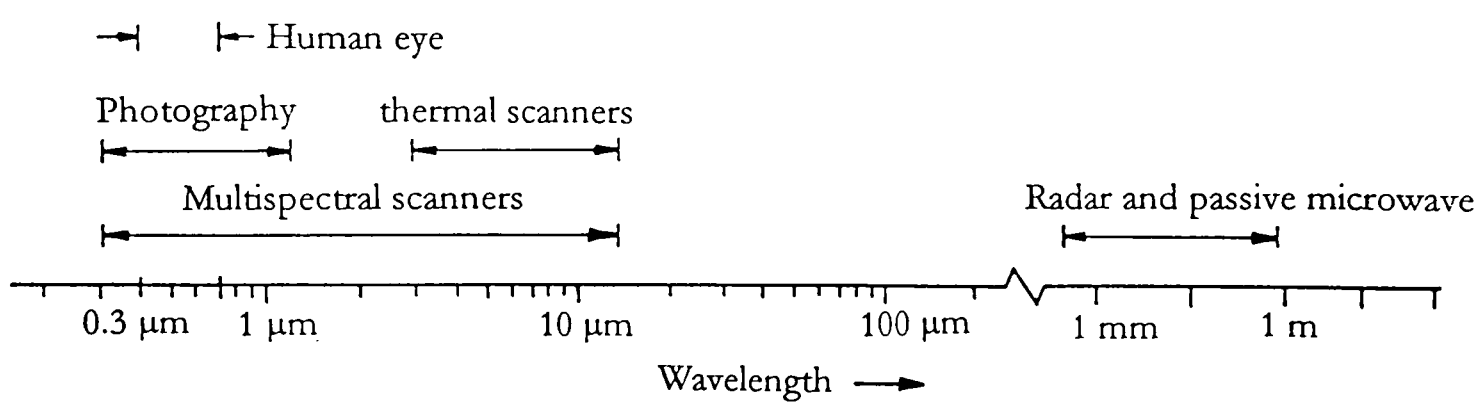
are often designed to maximise the information content of the signal received by operating in "atmospheric windows" of maximal EMR transmission, thus reducing the effect of atmospheric attenuation. The spectral location in the EMS of some familiar remote sensing systems is shown in figure 1.1.3. Figure 1.1.4 shows the spectral positioning of the channels of the Advanced Very High Resolution Radiometer (AVHRR) carried on the National Oceanic and Atmospheric Administrations' (NOAA) meteorological satellites and the High Resolution Radiometer (HRR) carried on the Meteosat satellites in the EMS. Further details on these satellite systems is provided in sections 1.2.3.1 and 1.2.3.2 with justification as to why these satellite sensor data are used in this work.



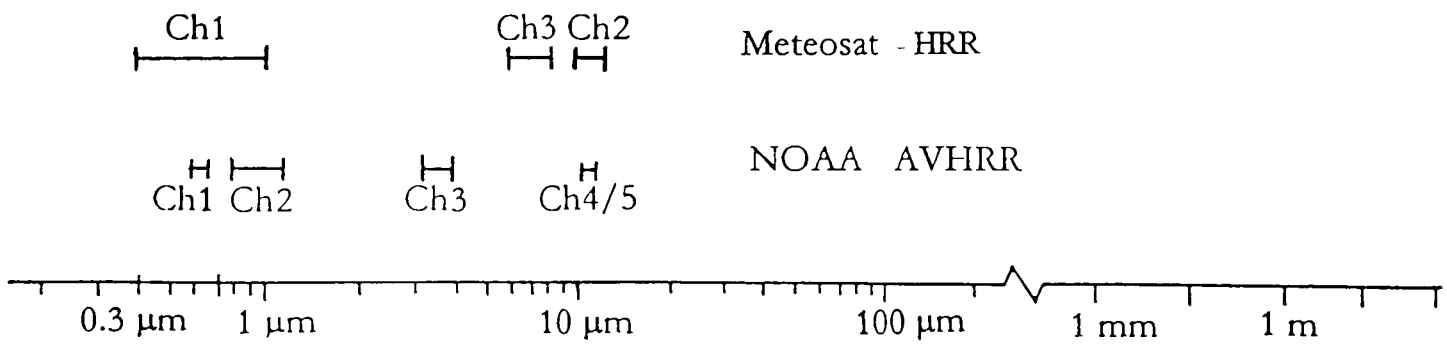
(1.1.1) Energy sources



(1.1.2) Atmospheric transmittance



(1.1.3) Generic remote sensing systems



(1.1.4) The NOAA - AVHRR and the Meteosat HRR.

Figure 1.1. The spectral characteristics of EMR sources, sinks and sensors; energy sources (1.1.1), atmospheric transmittance (1.1.2), generic remote sensing systems (1.1.3), the NOAA - AVHRR and the Meteosat - HRR (1.1.4). UV: Ultraviolet; Ch: channel. Note that the wavelength scale is logarithmic. Redrawn from Lillesand and Kieffer (1994).

1.2.3. Satellite systems

Radiometers can be carried on two broad categories of satellite; geostationary and polar-orbiting. Geostationary satellites are put into a high altitude (~ 23,000 - 40,000 km) orbit at the equator with a speed equal to that of the Earth's rotation, so that they remain above a fixed point on Earth. Polar-orbiting satellites circle the globe repeatedly at a much lower altitude (~ 600 - 900 km), so that successive orbits pass over a different section of the Earth as it rotates (Lillesand and Kieffer 1994).

At any instant a sensor will receive EMR from the cone within which energy is focused on the detector, as is shown in figure 1.2. The relationship between the angle of this cone at the sensor in radians, also called the instantaneous field of view (IFOV), β, the height of the sensor above the Earth, H, and the resulting diameter of the viewing area, D, both in metres, often referred to as the spatial resolution is given by;

$$D = H\beta.$$

For example, the spatial resolution of the AVHRR with an average IFOV of 1.4 milliradians and an orbiting altitude of approximately 833 km can be found from the above equation as D = 833,000 x (1.4×10^{-3}) = 1166 m. This is close to the 1.1 km often quoted (Kidwell 1995).

The data from satellites are stored and transmitted as digital numbers, with each value referring to the smallest area for which the satellite sensor can record data. When viewed on a computer monitor these areas are often called picture elements or pixels. The assumption that for each pixel the digital number represents the mean spectral signal from all objects within the IFOV is generally correct (Price 1982).

In polar-orbiting satellites the sensor scans across the track of the satellite as the orbit progresses to generate a series of contiguous scan lines which, when combined, form a two dimensional image or scene. In geostationary satellites the radiometer itself must move perpendicularly to the plane of the scan line at regular intervals to generate the image. The swath width (breadth of the area over which data are recorded by a sensor) is determined by both the satellite altitude and sensor characteristics. For example the AVHRR scans to ± 55.4° from the point of Earth directly under the satellite (nadir) which, at an altitude of approximately 833 km, results in a swath width of approximately 2700 km. Due to the curvature of the Earth the effective distance

between the Earth and the sensor increases with the scan angle so that the spatial resolution decreases to approximately 4 km toward the edge of the swath.

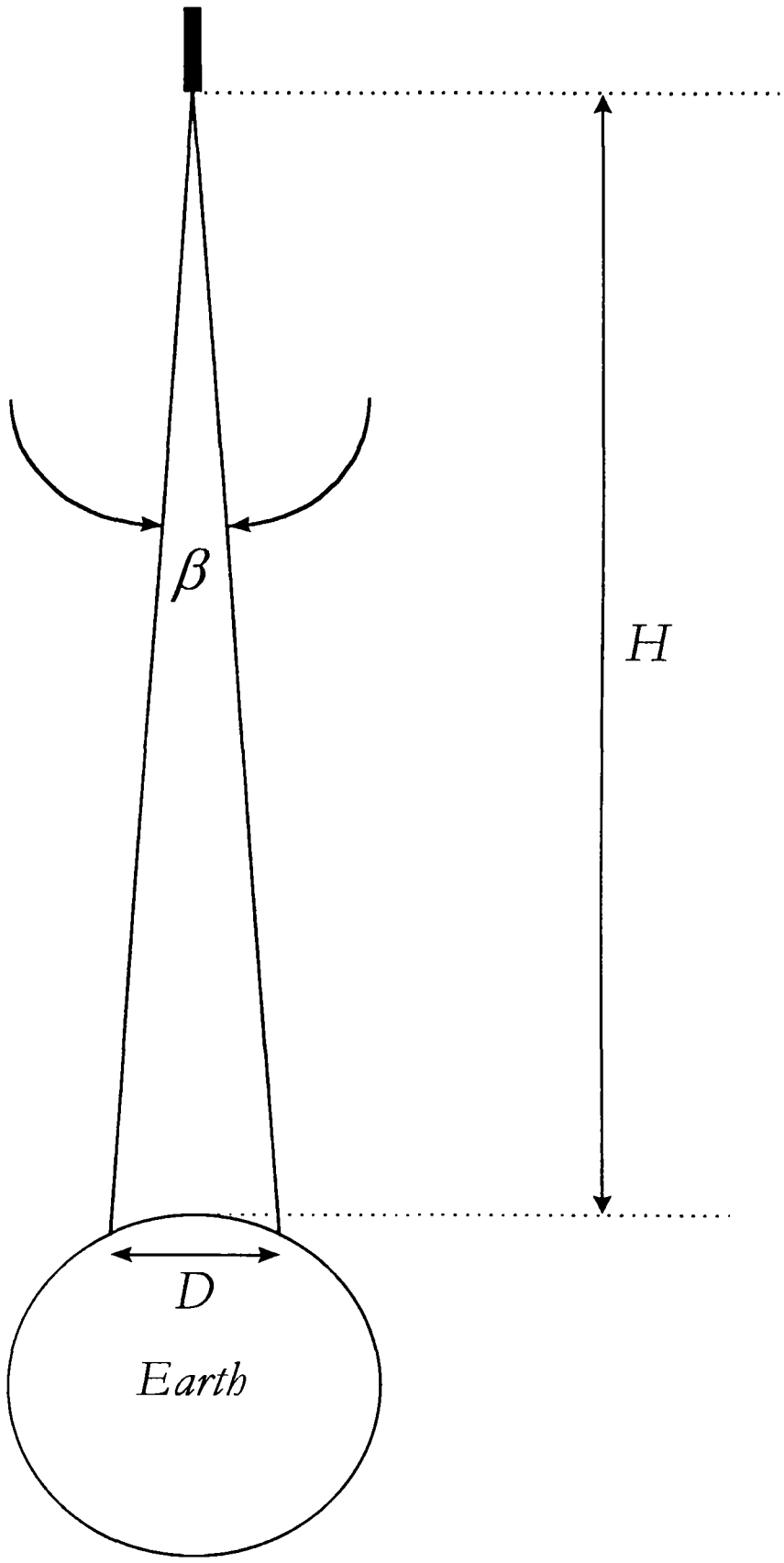


Figure 1.2. A schematic diagram of the factors affecting the spatial resolution of a radiometer. The instantaneous field of view (IFOV), β, is measured in radians; the height of the sensor above the Earth, H, and the resulting diameter of the viewing of the Earth, D, are measured in metres.

The "repeat time", *i.e.* the time taken between viewing the same part of the Earth's surface, varies between satellite systems and is determined by a combination of the swath width and orbital characteristics. Furthermore, data volume of satellite sensors is often constrained by on-board storage media (especially in older satellites) and the limited opportunity for telemetry (transmission of data between satellite and receiving station) during the satellite overpass, so that remotely sensed images tend to have either a high-temporal resolution or a high-spatial resolution, but not both. Satellite data therefore are limited in their spectral, spatial and temporal resolution by a variety of factors which reflect the compromise between the constraints of atmospheric effects, engineering limitations and the desired application. The compromises reached between the spectral, spatial and temporal resolutions by four current satellites often used in epidemiological applications are detailed in table 1.1.

In general the high-spatial resolution satellites of the Landsat Multi Spectral Scanner (MSS) and Thematic Mapper (TM) series, and the Satellite Pour l'Observation de la Terre (SPOT) series, give finely detailed images with spatial resolutions down to 30 m or 10 m (Landsat TM and SPOT High Resolution Visible (HRV) panchromatic respectively), but with repeat-times of approximately 16 and 26 days respectively. Frequent cloud contamination of the images means that such satellites sensors give few clear images of the Earth's surface per year, especially over tropical regions. This, together with the high cost of such imagery, generally limits the biological application of high-spatial resolution imagery to the production of habitat maps for relatively small areas and for a particular time in the season. For these reasons such data are not used in the present study.

In contrast to the sensors onboard Landsat and SPOT, the sensors onboard NOAA series of polar-orbiting meteorological satellites and Meteosat series of geostationary satellites have relatively high-temporal and low-spatial resolutions. The advantages of these features are detailed extensively in the following sections. Furthermore these data are available free to research institutes and developing countries, will be collected well into the next century and will soon have increased spectral, temporal and spatial coverage (Hay *et al.* 1996b). Considerable detail follows on the specifications of the NOAA and Meteosat satellite sensor systems, as data from both these platforms are used in the present work.

TABLE 1.1.

The spatial, temporal and spectral resolution of the sensors carried by the SPOT, Landsat, NOAA and Meteosat satellites.

Resolution	*Spectral*[a]	*Spatial*[b]	*Temporal*
Satellite Pour l'Observation de la Terre (SPOT)			
High Resolution Visible (HRV)			
Panchromatic mode	Ch 1 (0.51 0.73)	10 m	26 days[c]
Multispectral mode	Ch 2 (0.50 - 0.59)	20 m	
	Ch 3 (0.61 0.68)	20 m	
	Ch 4 (0.79 0.89)	20 m	
Landsat - 1, - 2, - 3, - 4, - 5			
Multispectral Scanner - MSS	Ch 4 (0.5 - 0.6)	79/82 m[d]	16 18
	Ch 5 (0.6 - 0.7)		days
	Ch 6 (0.7 - 0.8)		
	Ch 7 (0.8 - 1.1)		
Landsat - 4, 5			
Thematic Mapper TM	Ch 1 (0.45 - 0.52)	30 m	16 - 18
	Ch 2 (0.52 - 0.60)	30 m	days
	Ch 3 (0.63 - 0.69)	30 m	
	Ch 4 (0.76 - 0.90)	30 m	
	Ch 5 (1.55 - 1.75)	30 m	
	Ch 6 (10.40 12.50)[e]	120 m	
	Ch 7 (2.10 - 2.35)	30 m	
National Oceanic and Atmospheric Administration (NOAA)			
Advanced Very High Resolution	Ch 1 (0.58 - 0.68)	1.1 km	12 hours
Radiometer (AVHRR)	Ch 2 (0.72 - 1.10)		
	Ch 3 (3.55 - 3.93)		
	Ch 4 (10.30 11.30)		
	Ch 5 (10.50 - 11.50)		
Meteosat - 4, - 5, 6			
High Resolution Radiometer (HRR)	Ch 1 (0.40 - 1.10)	km	0.5 hours
	Ch 2 (10.50 - 12.50)	5 km	
	Ch 3 (5.70 - 7.10)	5 km	

a: The spectral resolutions are the electromagnetic wavelengths in μm, where 0.4 μm is at the blue end of the spectrum and 0.6 μm at the red, see figure 1.1.1.
b: The spatial resolution is given as diameter of the viewing area of the sensor, *D*, at nadir.
c: A pointing facility can improve the frequency of coverage.
d: The spatial resolution is 79 m for Landsat - 1 to - 3 and 82 m for Landsat - 4 and 5.
e: Channel 6 is out of wavelength sequence as it was added late to the TM in the original system design process.

1.2.3.1 The NOAA satellite series

The NOAA series of polar-orbiting Television Infrared Observation Satellites (TIROS) has been operational since 1978 (Hastings and Emery 1992). TIROS N (later renamed NOAA - 6) was the first satellite to carry the Advanced Very High Resolution Radiometer and has been followed by seven satellites each achieving an operational lifetime of between 2 and 4 years. The "very high resolution" refers to the 10 bit radiometric resolution of the sensor which therefore has the ability to store numbers in the zero to $2^{10} = 1024$ range. The definitive description of the NOAA polar-orbiting satellites, their radiometer payloads and the data they generate is given in Kidwell (1995).

The NOAA satellites complete 14.1 near-polar, Sun-synchronous orbits per day at an altitude of 833 - 870 km. Since the number of orbits is not an integer the orbital track over the Earth does not repeat on a daily basis. The even-numbered satellites have an ascending node with a north-bound equatorial crossing during the evening (19:30) and a descending node with a south-bound equatorial crossing in the morning (07:30), whereas the odd-numbered satellites have an ascending node in the afternoon (14:30) and a descending node at night (02:30). At present the exception is the latest NOAA - 14 satellite, launched in December 1994, which replaces NOAA 13 in functionality and thus has a daytime ascending node.

The AVHRR can view a 2400 km swath of the Earth and, at this orbital frequency, daily data are recorded for the entire Earth surface. Radiation is measured in five distinct bands of the EMS so that five separate waveband images are recorded for each orbit. The visible channel 1 and near infrared (NIR) channel 2 measure reflected solar radiation whereas the thermal channels 4 and 5 measure emitted thermal infrared (TIR). Channel 3 in the mid infrared (MIR) is a hybrid and sensitive to a combination of both reflected and emitted radiances. The exact specifications of the individual channels are detailed in table 1.1.

The spatial resolution of the AVHRR is approximately 1.1 km beneath the track of the orbiting satellite (see section 1.2.4). These nominal 1.1 km AVHRR data are continuously transmitted and may be received by stations along or near to the satellite's path, where they are referred to as High Resolution Picture Transmission (HRPT) data. On request to NOAA these data may also be recorded on an on-board tape storage system and later transmitted to Earth as the satellites pass over a network of receiving

stations. The data are then referred to as Local Area Coverage (LAC) data. These 1.1 km data have found application to a very wide range of disciplines (Huh 1991) and are reviewed by Ehrlich *et al.* (1994) and Cracknell (1996).

Two processing steps further reduce the spatial resolution of most of the AVHRR data available to the user community. The on-board tape system is incapable of holding global coverage data at 1.1 km resolution. Instead the information from each area of five (across-track) by three (along-track) pixels is stored as a single value, the average of the first four pixels only of the first row of the 5 by 3 block. The resulting imagery is referred to as Global Area Coverage or GAC data. GAC data, with a stated nominal spatial resolution of 4 x 4 km, are obviously far from ideal representations of the raw data (Justice *et al.* 1989) and their method of sub-sampling has consequences for environmental modelling (Belward and Lambin 1990, Belward 1992). Nevertheless, GAC data are the form in which most of the AVHRR archive was collected and reasonable quality global datasets are available at a variety of spatial resolutions (4 x 4 km or coarser) from the early 1980s to date (Townshend 1994).

1.2.3.2 The Meteosat satellite series

The EUMETSAT (European Organisation for the Exploitation of Meteorological Satellites) geostationary Meteosat satellite series began with the launch of Meteosat - 1 in 1977. Experimental satellites were used until the launch of the first operational satellite, Meteosat - 4, in June 1989. The spin-stabilised satellites are put into orbit at an altitude of 35,800 km over the Gulf of Guinea, at the crossing of the Equator and the Greenwich meridian (0° N, 0° E). In this position images are captured for the full Earth's disc including Africa, Europe and the Middle East. A reserve satellite operates nearby in a standby condition (WMO 1994).

The principal payload of the satellite is a High Resolution Radiometer (HRR). The radiometer operates in a broad visible waveband (channel 1), a thermal infrared waveband (channel 2) and a water vapour absorption infrared waveband (channel 3), see table 1.1. The Meteosat satellites were designed for meteorological applications so that channel 3 is located in the thermal infrared area of maximal water vapour absorption (see figure 1.1.4) and hence ideal for monitoring clouds. At nadir the spatial resolution is 2.5 km for the visible images and 5 km for the thermal infrared and water vapour images. Further from the equator the spatial resolution decreases so that over northern

Europe the spatial resolution is 4 km in the visible wavebands and 8 km in the thermal infrared and water vapour wavebands. Each image is transmitted to the Earth in real time as each scan line is completed and new full Earth disc images are generated at 30-minute intervals (Guyenne 1987).

1.3 Turning meteorological satellite data into geophysical data

This section details some of the fundamental problems experienced in measuring reflected and radiated EMR from a curved surface, through an atmosphere of spatially heterogeneous composition by a moving satellite. It then deals with how such data are converted into calibrated geophysical variables of known geographical location.

1.3.1 Image registration to base maps

Raw digital data derived from satellites sensors need to be pre-processed geometrically to rectify or register them to a base map at a particular scale and in a particular map projection (Snyder 1987). Registration is generally an automated process that uses an ephemeris model of the orbital parameters of the satellite and a time signal sent down with the satellite sensor imagery to predict the satellite's position relative to the Earth at the time of image capture (Brush 1988, Emery *et al.* 1989, Baldwin and Emery 1993). Algorithms for positional calculations originally assumed that the satellites were in their correct attitude and were following precisely their intended orbits. This was unfortunately not the case, since satellites varied considerably in both their orbit and attitude (McGregor and Gorman 1994). An extreme example is that the equator crossing time for NOAA - 11, which was 14:20 when the satellite was launched in February 1985, but had drifted to 16:07 by November 1988 (Kidwell 1995). These deviations from design values, and variations in these deviations, mean that some of the resulting imagery is not accurately registered to the appropriate base maps.

Satellite sensor images also suffer from geometric distortions due to other factors which include; panoramic distortion, Earth curvature, atmospheric refraction, relief displacement and non-linearities in the sensor's field of view (Lillesand and Kieffer 1994). These errors may be systematic or random. Geometric correction of systematic errors is usually done by modelling the sources of errors mathematically and applying resulting corrective formulas.

Random distortions are overcome by measuring the shift of ground control points (GCP), distinctive geographical features of known location on the image, and resampling or reforming the original image to a new one accordingly. The functional relationship (f) between the X and Y file co-ordinates of the original satellite sensor image and the known latitude (x) and longitude (y) are determined by a least squares regression to determine the coefficients for two co-ordinate transform equations;

$$x = f_1 (X, Y)$$

$$y = f_2 (X, Y).$$

After a geometrically correct geographical grid is defined in terms of the longitude and latitude, each cell in this grid is given values of x and y according to the co-ordinate transform equations above. The computer then maps the digital number from the pixel closest to this address in the raw image to the geometrically correct geographical grid. This last step can be done in a simple way, such as to the nearest neighbouring pixel, or by using more sophisticated and computer intensive methods, such as bilinear and cubic spline interpolations (Khan *et al.* 1992, 1995). The spatial interpolation process may substantially alter the radiometric fidelity of the data (Goward *et al.* 1991).

The on-board sensors of geostationary satellites view the Earth as a disc, so that apart from spherical distortion there are few, if any, problems of geo-registering such images. The raw data from polar-orbiting satellites however, are a series of strips which must be co-registered and geometrically corrected before successive images can be joined together. The effects of these various stages of image resampling on AVHRR data are considered in Khan *et al.* (1995).

Image registration to a map tends to involve a loss of spatial resolution, the extent of which is usually increased to allow for the coarsest spatial resolution of the data rather than oversampling at the nadir pixel size. Final registration to a base map frequently has to be performed by visual inspection of the image with a map overlay (Krasnopolsky and Breaker 1994).

1.3.2 Removal of cloud contamination

Each of the high-temporal resolution images from the NOAA (and Meteosat) satellite sensors is as affected by cloud contamination as any single Landsat or SPOT sensor image but their much higher frequency means that data quality can be improved by combining images over a relatively short period of time by compositing. The aim of compositing is to choose the most cloud free and /or least atmospherically contaminated radiance value within the compositing period.

Most compositing algorithms rely on the fact that one common image product from the AVHRR indices, the Normalized Difference Vegetation Index or NDVI,

produced from channels 1 and 2 (see section 1.5.1 for details), has values that are generally reduced by cloud and other atmospheric contamination (Holben 1986, Kaufman and Tanré 1992). The highest NDVI values recorded during any relatively short time period are therefore thought to occur when cloud cover is least and such values are taken to represent the least attenuated pixel value for the period. This method of image production is called Maximum Value Compositing (MVC), and is usually carried out over a ten-day (decadal) compositing period. It has the important consequence that maximum image values for adjacent pixels in a single image may have been collected on different days during the compositing period. MVC methods tend to degrade still further the spatial resolution of the final image product, primarily for geo-referencing reasons (Meyer 1996, Robinson 1996), so that the recorded value for any nominal 8 x 8 km pixel may in fact have been drawn from an area as large as 20 x 20 km.

Other compositing regimes include taking the root mean square value for a given pixel time-series (Kineman *et al.* 1990) and more sophisticated methods, such as best index slope extraction (BISE) (Viovy *et al.* 1992). This technique assumes that vegetation grows and senesces over weeks, so that the evolution and decay of the resulting NDVI signal is gradual. Values which lie outside a user determined threshold from the previous signal are interpolated, as they are probably due to short-term effects of atmospheric contamination.

Selection of the least cloud-contaminated images in the AVHRR channels usually depends upon the selection of the NDVI date chosen by MVC. The same image that is used to generate the NDVI for any pixel is also taken as the source of information for the other AVHRR channels for that pixel and period. Increasingly in land applications over the tropics however, AVHRR channels 4 and 5 are composited separately *i.e.* without reference to the NDVI, since the overlying clouds are generally colder than the land so that the highest thermal value in the series will probably be the least cloud contaminated (Lambin and Ehrlich 1995, 1996).

1.3.3 Removal of other atmospheric effects

During image registration to a base map other corrections for atmospheric effects such as Rayleigh scattering caused by aerosols and absorption by water vapour, carbon dioxide and ozone, may also be applied, using ancillary information in the satellite data stream (Vermote *et al.* 1990, Tanré *et al.* 1992). This is very important because

atmospheric aerosols, which are highly spatially and temporally variable in the atmosphere (Holben *et al.* 1991), scatter light particularly in channel 1, whilst atmospheric water vapour absorbs particularly in channel 2 (Kaufman and Tanré 1992). If corrections for atmospheric effects are not made at the time of image registration they generally cannot be made so accurately at a later stage. Instead corrections are based on average values for a "standard" atmosphere within a region (Hannan *et al.* 1995). Despite attempts to remove the effects of clouds and other contaminants by MVC of AVHRR imagery, continuous total or sub-pixel cloud cover and haze may still affect MVC images.

1.3.4 Satellite sensor drift

During their operational life-time the AVHRR sensor characteristics change with use and as their components age (Gorman and McGregor 1994) so that external or "vicarious" calibration is required. The TIR channels are continuously calibrated against the 4 K space background temperature, measured by a thermistor on the baseplate of the satellite. Channels 1 and 2 are calibrated by making simultaneous measurements with aircraft underflights (Smith *et al.* 1988) or by examining the change in signal from relatively invariant reflectors such as deserts and high clouds, or from invariant reflective phenomena such as the molecular scattering of the visible signal over oceans and areas of Sun glint (Che and Price 1992, Kaufman and Holben 1993). Correction factors can also be added to spectral vegetation indices (SVIs) without recourse to the visible channel data by assuming a linear degradation in sensor response, examples of which are given by Los (1993) for the NDVI.

1.3.5 Satellite precession

Even when corrections for atmospheric effects and instrumental drift have been made the resulting imagery may still show periodic changes in the signal due to a precession of the satellite's orbit known as phasing. This results in cyclical variation of over-pass times, which in the case of AVHRR imagery have a 17 day cycle (McGregor and Gorman 1994). Signal variation is then due to the changing angles between the Sun, Earth and satellite sensors.

It is clear, therefore, that images available to the user suffer from a variety of defects for many different reasons. The measures used in this work to help correct for such artefacts in the satellite sensor data are outlined in chapter 2.

1.4 Turning geophysical data into epidemiologically useful information

The previous section highlights the problems of obtaining geographically registered satellite sensor data and the sources of error involved in the processes. These percentage reflectance and brightness temperatures data distributed to epidemiologists however, cannot be immediately related to the ecology of disease vectors. This section describes how such information can be converted to vegetation, land surface temperature, atmospheric moisture and rainfall indices. The accuracy with which meteorological variables that affect disease vector population dynamics can be described is also discussed. The section concludes with information on MIR with specific details on how EMR at these wavelengths interact with and therefore what information it might contain about the land surface.

1.4.1 Spectral vegetation indices

SVIs (reviewed by Huh 1991 and Myneni *et al.* 1995b) exploit the fact that chlorophyll and carotenoid pigments in plant tissues absorb light in the visible red wavelengths (which corresponds to AVHRR channel 1), whereas mesophyll tissue reflects light in the near infrared wavelengths (which corresponds to AVHRR channel 2) (Sellers 1985, Tucker and Sellers 1986). A healthy and actively photosynthesising plant will therefore look darker in the visible, and brighter in the infrared region, than an unhealthy or dead plant. Furthermore, as vegetation coverage increases there is more absorption of red radiation due the increasing amount of pigmentation and more reflectance of near infrared due to increases in internal leaf scattering (Curran 1985). The reflectance from dry soil is less complex than that of vegetation, showing a general increase in reflectance with wavelength which is dependant on soil texture, structure, and water, organic carbon and iron oxide content (Huete and Escadafal 1991). Since soils and vegetation exhibit very different spectral properties these features are used to differentiate between the two types of surfaces. SVIs are simply designed to maximise the contrast in reflectance and thereby identify the presence of vegetation in remotely sensed images.

The most simple SVI is the ratio of AVHRR channel 2 (Ch_2) over channel 1 (Ch_1) reflectances, called the Ratio Vegetation Index (RVI) or Simple Ratio Index (SRI). Other SVIs attempted to overcome the problem of reflectance from the (usually dark or reddish) soil backgrounds by dividing the difference between these two channels by their

sum, to give the Normalized Difference Vegetation Index or NDVI (Tucker 1979) defined as follows;

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

The values of the NDVI can theoretically range from -1 to +1 but in practice usually fall within 0.2 0.8 limits (Colwell 1974, Tucker 1979, Tucker *et al.* 1991). The NDVI, in common with all red/near-infrared indices, is theoretically a specific measure of chlorophyll abundance and light absorption (Myneni *et al.* 1995a) but its use has been extended to quantify herbaceous vegetation biomass (Tucker *et al.* 1983, 1985b), vegetation primary productivity (Prince and Goward 1995), vegetation coverage (Tucker *et al.* 1985a) and phenology (Justice *et al.* 1985) in a range of ecosystems. NDVI measurements are particularly useful in areas of sparse vegetation coverage, where they have a larger dynamic range than the simpler SVIs such as RVI. The NDVI does, however, tend to saturate in areas of full coverage such as forests (Huh 1991). It is also less than ideal because of continuing problems with background soils (which are, for example, darkened by rainfall (Huete *et al.* 1985)) and differential atmospheric effects on channels 1 and 2 radiances.

Alternative indices have been suggested to overcome some of these problems (Jackson and Huete 1991) but have been less widely applied to ecological and epidemiological problems than the NDVI because they have not been shown to be generally applicable.

For example, 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;

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

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 (Verstrate 1995), rather than empirically, although the physical basis for the index is not fully explained in the literature, and it is defined as follows;

$$GEMI = \eta(1 - 0.25\eta) - \frac{Ch_1 - 0.125}{1 - Ch_1}$$

where

$$\eta = \frac{2(Ch_2^2 - Ch_1^2) + 1.5Ch_2 + 0.5Ch_1}{Ch_2 + Ch_1 + 0.5}.$$

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, Leprieur *et al.* 1996).

Vegetation indices that require the calculation of a soil line are not discussed because of the problems of finding sufficient bare-ground pixels at tropical latitudes. The relevance of SVIs to land-cover mapping are discussed in more detail in chapter 4.

1.4.2 Land surface temperature indices

The theoretical concept of a black-body is used to describe any material that absorbs and emits radiation perfectly at all wavelengths. Such a hypothetical material is described as having a spectral emissivity of one *i.e.* the ratio of emission at temperature (T) versus emission at the standard temperature (Ts = 273 K) (Montieth and Unsworth 1990). In ideal conditions therefore, the temperature of a black-body can be determined by detecting the energy it emits at a particular wavelength (see section 1.2.2). Natural

surfaces do not behave as black-bodies, however, and have emissivity values less than one; usually 0.99 for water, 0.96 - 0.99 for vegetation and lower for soils (Salisbury and Daria 1992). Furthermore, the radiometric brightness temperature (Becker and Li 1990a) measured by the satellite sensor is also affected by absorption characteristics of atmospheric constituents (particularly water vapour but also ozone, carbon dioxide and aerosols), as well as emission of radiation by the atmosphere itself (Vogt 1992). Attempts to estimate accurate surface temperatures from satellite derived brightness temperature must therefore correct for atmospheric attenuation and the spatially heterogeneous nature of land surface emissivity. These are major areas of past and current remote sensing research (Norman *et al.* 1995).

Channels 4 and 5 of the AVHRR radiometer have long been used to measure water vapour attenuation in the 10 12 μm spectral window to increase the accuracy of sea-surface temperature determination (Prabhakara *et al.* 1974). The attenuation is greater in channel 5 than in channel 4 so that the difference between the signal of these two channels can be used to estimate the amount of atmospheric water vapour attenuation and is used to reduce such effects. This simultaneous use of information from both channels to estimate surface brightness temperatures is described as a "split-window" technique, because it is performed within the same radiance window of the atmosphere.

Surface emissivity is more variable on land than over the relatively uniform sea surface and so allowance needs to be made for emissivity when comparing surface brightness temperatures of different land-surface types. Many split-window techniques have been developed, which largely rely on ancillary data to quantify atmospheric water content and surface emissivity (Prata 1993, Becker and Li 1995, Norman *et al.* 1995). Those used in the current investigation require only raw channel AVHRR data and are detailed more fully in chapter 2.

1.4.3 Atmospheric moisture indices

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;

$$U = a \times \left(Ch_4 - Ch_5\right) \times \cos\theta .$$

The estimates were stated to have an accuracy of $\pm$ 5 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.

The difference in the AVHRR brightness temperatures (channel 4 channel 5) has been shown to have a linear relationship with total precipitable water in the atmospheric column (Eck and Holben 1994) using balloon radiosonde and Sun photometer data from three meteorological stations in Mali. The standard error of the total precipitable water estimate was between 0.31 and 0.48 kgm^{-2} and found to increase for these sites when the Dalu (1986) equation was applied to the same data.

1.4.4 Rainfall indices

In tropical latitudes where diurnal heating provides large reservoirs of potential energy, weather systems are dominated by atmospheric convection processes (Martyn 1992, Emanuel 1994). The most vigorous convection currents provide the strongest updrafts which result in clouds with higher water contents that are more likely to be rain-bearing (Byers and Braham 1948). These convection currents form deep clouds with high and cold tops which emit very low radiance values in the thermal infrared. These cloud-top temperatures can be recorded by the channel 2 of the Meteosat satellite (see table 1.1). The relationship between cloud temperature and the probability of rainfall has been well established (Scofield and Oliver 1977, Burt *et al.* 1995). The particular threshold temperature associated with rain-bearing clouds and the quantity of rain they deposit varies temporally and spatially however and must be established empirically (Milford and Dugdale 1990).

More sophisticated rainfall estimation techniques that relate cloud top reflectances and the growth and decay of cloud systems to rainfall amounts are reviewed by Petty (1995). Significant advances are also being made using a combination of high spatial resolution radar, passive microwave and visible and infrared radiometer measurements by the Tropical Rainfall Measuring Mission (TRMM) (Theon 1993). These studies, however, require data not available to the current investigation and involve processing far beyond the scope of the present work.

1.4.5 Middle infrared radiation.

Land surface applications of MIR have focussed mainly on the detection of hot regions associated with forest, peat and straw fires and burn scars (Malingreau *et al.* 1985, Matson and Holben 1987, Gregoire 1990, Chuvieco and Martin 1994, Pereira and Setzer 1996). It has often been used in conjunction with visible radiation, for both surface temperature mapping and for land-cover discrimination (Kerber and Schutt 1986) where it enhances the spectral separability of land-cover classes.

Despite these studies the level of understanding and documentation regarding the interaction of MIR radiation with targets, relative to the visible and near infrared wavelengths, is limited. Moreover, the use of MIR wavelengths in land-cover mapping is at an early stage of development (Ehrlich *et al.* 1994). This is due to the hybrid nature of this spectral region (sensitive to both reflected and emitted radiation (Kidwell 1995)) which makes the interpretation of the signal returning from the target more difficult, historical difficulties of data access and that instrument noise can seriously contaminate the MIR signal in AVHRR sensors (Dudhia 1989).

There have been reported however, good reasons for using MIR radiation for land-cover discrimination in the tropics. Primarily MIR suffers less attenuation in the atmosphere (Bernstein 1982, Wooster *et al.* 1994) and can penetrate to a greater depth through smoke than the visible or NIR wavelengths (Kaufman and Remer 1994). The MIR region also suffers little attenuation due to atmospheric water (Kerber and Schutt 1986) making it particularly suitable for applications in the tropics. These factors coupled with the known interactions of MIR with vegetation (see section 4.1.2) help justify the incorporation of such data in land-cover type discrimination.

Boyd and Curran (1996) have proposed an explanation for the interaction of MIR and the biophysical properties of vegetation canopies with particular reference to

tropical forests. The primary factors they suggested to operate include the water content, surface temperature and the structure and roughness of the vegetation target. Increases in each of these factors with increasing vegetation coverage are postulated to cause a decrease in the MIR signal. Firstly, an increase in vegetation amount corresponds to an increase in liquid water that can absorb MIR (Kaufman and Remer 1994) hence reducing the signal. Secondly, there is the effects of thermal emission which dominates the response in the MIR region. The decrease in MIR emitted with increasing vegetation amount occurs due to the a decrease in the surface resistance to evapotranspiration *i.e.* greater transpiration and because canopy foliage temperatures are significantly lower than background soil surface temperatures due to their relative specific heat capacities (Lambin and Ehrlich 1996). Thirdly, an increasingly complex canopy structure which is dependent upon canopy depth, leaf orientation and distribution and has an affect on incoming MIR radiation by trapping photons and producing shadows which decrease the intensity of reflected MIR radiation (Dadhwal *et al.* 1996). This relationship is summarised in figure 1.3.

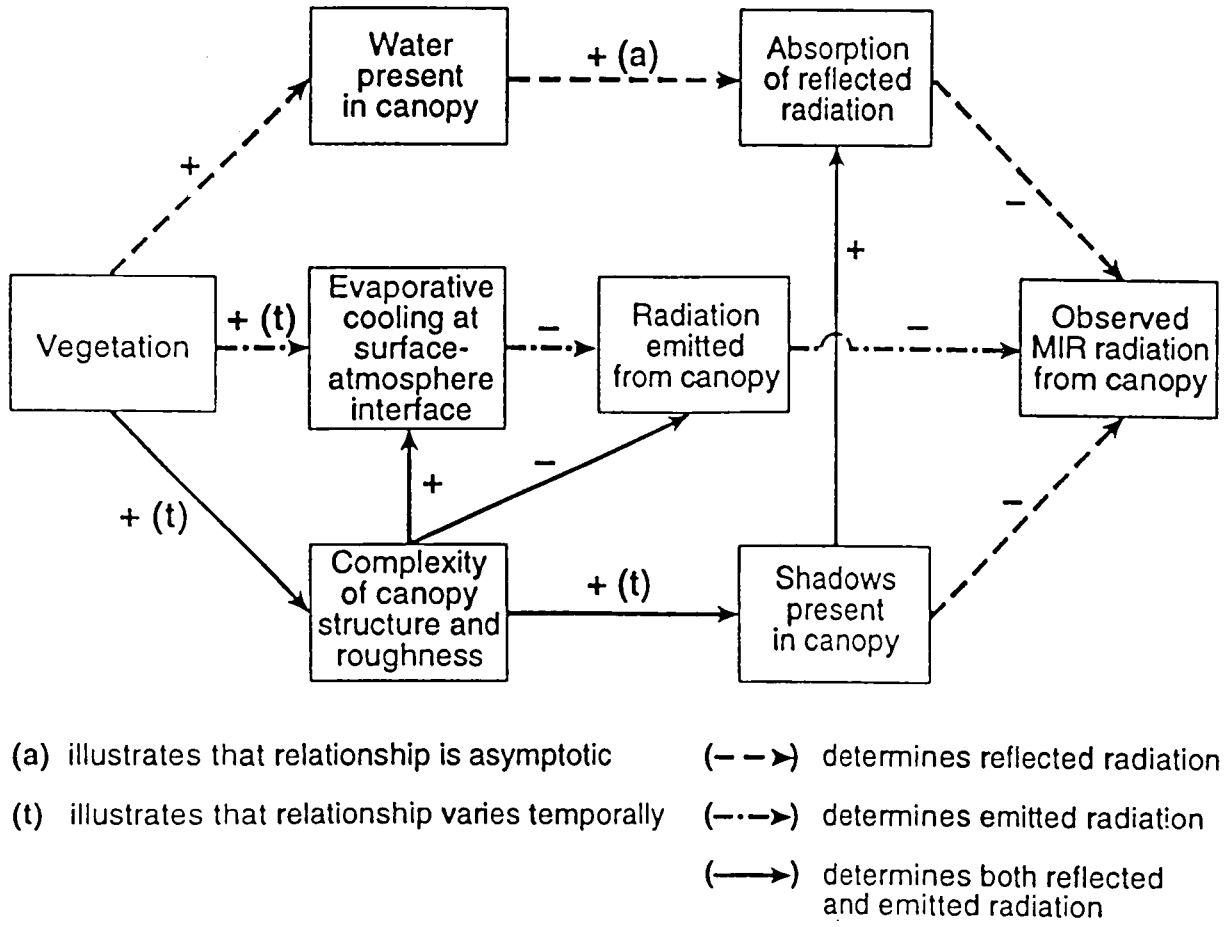


Figure 1.3. Factors that determine the relationship between middle infrared radiation (MIR) and the amount of vegetation. Redrawn from Boyd and Curran (1996).

1.5 Past application of remote sensing to vector-borne diseases

1.5.1 Introduction

Haematophagous disease vectors are organisms that introduce or inoculate pathogens directly onto or into a vertebrate through their blood-feeding habits (Molyneux 1993). The potential to use remote sensing in epidemiological studies has long been recognised (Cline 1970) and many investigations have used remote sensing techniques for the study of vector-borne diseases. Previous reviews of the subject include; Cline 1970, Rush and Vernon 1975, Jovanovic 1987a, 1987b, Hugh-Jones 1989, 1991a, Riley 1989, Wood *et al.* 1992a, Epstein *et al.* 1993, Roberts and Rodriguez 1994, Hacker and Roberts 1994, Washino and Wood 1994, Beck *et al.* 1995, Hay *et al.* 1996a.

1.5.2 Disease control

The principle goal of remote sensing in epidemiology is to map the distribution of a disease (often by mapping the distribution of its vector) so that control efforts in endemic situations and intervention strategies in epidemic situations may be most efficiently directed (Hugh-Jones 1989).

Determining the distribution of disease vectors is often the first step in an epidemiologically sound approach to disease control. This arises because the basic reproductive number (rate or ratio depending on the author) of a disease, R_o, which describes the average number of new cases of a disease that will arise from the introduction of an infective host into a susceptible population (Dietz 1988, Anderson and May 1991, Lord *et al.* 1996), is influenced primarily by factors associated with the vector. The standard formula for the basic reproductive number for insect vector-borne diseases is;

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

where, a, is the vector biting rate, m, is the ratio of vectors to hosts, c, is the transmission coefficient from vertebrate to vector (*i.e.* the proportion of bites by vectors on infected hosts that eventually give rise to mature infections in the vectors), b, is the transmission coefficient from infected vector to vertebrate, μ, is the vector mortality rate, T, is the incubation period of the infection within the vector (sometimes referred to as the

extrinsic incubation period) and, *r*, is the rate of recovery of the vertebrate from infection (Rogers 1988a). All of the parameters in this equation are related to the vector save *r*, which is a function of the host alone. Given this importance of many aspects of vector biology and behaviour to the transmission dynamics of the diseases they carry, it follows that the distribution and intensity of such diseases is dependent upon the distribution, abundance and biology of the vector. Control campaigns should therefore aim to decrease vector population numbers below the threshold of disease transmission (Rogers *et al.* 1994). This occurs when R_o falls to a value of less than one.

This survey is restricted to work on mosquito, tick and tsetse-borne diseases since most of the research effort has been focused on these vectors. A more complete review of the application of remote sensing to the understanding and control of vector borne and intermediate host transmitted diseases can be found in Hay *et al.* (1996a). For each disease a brief explanation of the transmission cycle and vector biology is given to illustrate how surveillance techniques have been of use.

1.5.3 Mosquitoes, malaria and filariaisis

1.5.3.1 Disease and vector biology

Mosquitoes (flies of the Family Culicidae, see figure 1.3.1) have a simple life-cycle with the egg, larval and pupal stages passed in still, or slow-moving fresh or brackish water. Soon after emergence the haematophagous adult females seek vertebrate hosts in order to feed. The ingested blood is used in egg production and, following oviposition, females seek further blood-meals to nourish future broods. It is this repeated feeding that facilitates the transmission of parasites between hosts (Bruce-Chwatt 1986). The organisms for which mosquitoes act as vectors include a range of arboviruses (arthropod-borne viruses) such as Yellow Fever, Dengue Fever and Rift Valley Fever viruses; nematode worms such as *Wuchereria bancrofti* and *Brugia malayi* which cause lymphatic filariasis; and four species of the protozoan *Plasmodium* that cause malaria in humans (Mattingley 1969).

World-wide, mosquitoes are the most important vectors of disease (Godfrey 1978). It is estimated that 2280 million people are at risk from malaria (approximately 42 % of the world population) and of these it is estimated that between 300 - 500 million clinical cases are reported each year, 90 % of which occur in Africa (WHO 1994). Lymphatic filariasis affects a further 900 million people and, though generally not life-

threatening, causes chronic suffering and disability when the adult worms develop in the afferent lymphatic system, leading to painfully swollen and permanently disfigured limbs, often referred to as elephantiasis (Whitfield 1993).

It is the dependence of mosquitoes on fresh and brackish water habitats in the early stages of their life-cycle that has allowed remote sensing techniques to be exploited. This is of particular advantage because insecticide is often applied to breeding habitats to kill the larvae during control campaigns.

Figure 1.4. An adult female mosquito (Family Culicidae, Subfamily Anophelinae). Redrawn from Molyneux (1993).

1.5.3.2 Application of high spatial resolution imagery

In 1971, as part of the Health Applications Office (HAO) initiatives, NASA scientists in combination with personnel from the New Orleans Mosquito Control District (NOMCD), were the first to investigate the use of colour and colour-infrared aerial photography in mapping vegetation assemblages associated with the larval habitat of *Aedes sollicitans*. This is a saltmarsh mosquito suspected of transmitting the equine encephalitis virus (NASA 1973). Previous work by the NOMCD had shown that females would always oviposit in areas of the saltmarsh intermittently flooded by freshwater. The floral assemblage dominated by spikerush (*Spartina patens*) and wiregrass

(*Juncus roemerianus*) was known to be adapted to the same hydrological regime, and hence was a reliable indicator of *A. sollicitans* larval habitat (Bidlingmayer and Klock 1955). The report documents that such vegetation assemblages were "extremely accurately" identified at an 80 ha test site near New Orleans, although no statistics were provided.

The first operational use of colour-infrared aerial photography to map forested and open wetlands, marshes and residential areas for mosquito control took place in the Saginaw and Bay Counties of Michigan (Wagner *et al.* 1979). These habitats supported populations of nuisance *Aedes* and *Culex* species of mosquitoes that were vectors for a local epidemic of St. Louis encephalitis (SLE) in 1975. SLE is an arbovirus causing severe irritation of the central nervous system which proves fatal in approximately 10 % of cases (Leake *et al.* 1934, Shope 1980). The known flight range of each mosquito species was used in combination with information on the distance between residential areas and mosquito habitat to identify control priorities for the two counties. The authors stressed the short time in which the environmental inventory was gathered and management priorities for the control campaign identified. Furthermore, the streamlining of subsequent control efforts led to a campaign of relatively low economic and environmental cost, since the area designated for insecticide treatment was considerably reduced in comparison to more traditional approaches of broadcast spraying. For example, Hopkins *et al.* (1975) report that after an outbreak of SLE in Dallas in 1966, the entire State of Texas was sprayed aerially with malathion (an organophosphate insecticide)!

The use of colour-infrared aerial photography has been cited as a more cost-effective method than conventional ground survey techniques of obtaining information on the distribution of oviposition habitats of the mosquito *Psorophora columbiae* in Louisiana and Texas rice fields (Fleetwood *et al.* 1981, Welch *et al.* 1989a, 1989b). Improvements upon colour-infrared aerial photography were demonstrated by Cibula (1976) and Barnes and Cibula (1979) using an airborne MSS. The greater number of recorded wavelength bands enabled a more accurate identification of the *Spartina - Juncas* associations characteristic of favourable egg-laying habitats for *Aedes sollicitans* than aerial infrared photographs. In a similar study Hayes *et al.* (1985) reported that *Culex tarsalis* and *Aedes vexans* larval habitats could be identified along the Niobara river in north-east Nebraska with Landsat - MSS data, using transitional freshwater plant communities associated with seasonal flooding as an indicator. This study demonstrated that such

species-habitat correlations could be successfully "scaled-up" to regional control programmes using relatively high spatial resolution satellite data.

Under the aegis of the NASA Biospheric Monitoring and Disease Prediction Program's Disease Modelling (DI - MOD) Project, various authors have investigated the hypothesis that the spatial and temporal aspects of mosquito population dynamics are controlled by environmental factors that can be observed remotely (Wood *et al.* 1994). In phase I of this project, populations of *Anopheles freeborni* in the rice fields of northern and central California were examined (Pitcairn *et al.* 1988). This species of mosquito represented no significant health risk to the local population, but provided an accessible model for the study of Anopheline vectors in irrigated rice habitat (Wood *et al.* 1991a, 1991b, 1992b). This is important, since the cultivation of irrigated rice, a land-use occupying an area of 140 million hectares globally, provides an ideal habitat for Anopheline mosquitoes which are the vectors for human malaria in the tropics (Service 1989).

Larval mosquito populations were sampled fortnightly throughout the period of rice crop development in 1985 (Wood *et al.* 1991b). An aircraft-borne radiometer collected data simultaneously in bands designed to simulate Landsat - TM data. NDVI values were then calculated for individual fields on each of the sampling dates and the differences in the NDVI between rice fields producing high and low numbers of Anopheline larvae were followed throughout the growing season. The NDVI was considered an important variable since Rejmankova *et al.* (1988) had previously demonstrated percentage rice crop cover to be positively correlated with mosquito larval production. The results showed that higher values of NDVI in the early growing season (June) were associated with high mosquito producing rice fields, but that the spectral separation (between high and low mosquito producing fields) diminished to a minimum in mid July, when percentage rice crop cover exceeded 100 %. A discriminant analysis which incorporated information from Landsat - TM equivalent channels 1, 2, 3, 4, and 7, was able to distinguish between high and low mosquito producing fields with an overall accuracy of 75 %. The uneven distribution of high mosquito producing fields was also shown to be related to patterns of surrounding land-use, with 70 % of the high mosquito producing fields being within 1.5 km of livestock pasture (*i.e.* potential hosts). A more detailed survey in 1987 repeated the above work and included data on the distance to livestock in a Geographic Information System (GIS) (Wood *et al.* 1991b).

An identical discriminant analysis combined with cattle distance data resolved high and low producing fields with an accuracy of 90 % (Wood *et al.* 1992b).

The second phase of the project investigated the population dynamics of *Anopheles albimanus* in the tropical wetlands of Chiapas, Mexico, where malaria is endemic (Roberts *et al.* 1991, Pope *et al.* 1994). Two Landsat - TM scenes of the area (one from the dry and one from the wet season) were subject to an unsupervised classification and the resulting groups assigned to land-cover types on the basis of color infrared aerial photographs and field inspection of 30 test sites. These sites were independently sampled for mosquito density and information was collected on environmental variables affecting water and vegetation characteristics (Rejmankova *et al.* 1991). The sites were then grouped into 16 habitat types using a cluster analysis and correlations were performed between the habitat types and land-cover units (Rejmankova *et al.* 1992). The land-cover units were subsequently ranked as having high, medium or low mosquito production potential on the basis of these correlations. Incorporating this information into a GIS, sites of high mosquito production around the towns of La Victoria and Efrain Gutierrez were found to occupy only 9 % of the designated control area, allowing the potential for substantial streamlining of control campaign effort and resources. Related work later demonstrated how such data could be used effectively by identifying particular villages at high risk within the control area (Beck *et al.* 1994). This was achieved by using a stepwise discriminant analysis and linear regression to establish the relationship between vector abundance and landscape element proportions. In both cases transitional swamp and un-managed pasture habitats were identified as most important, and in combination were able to distinguish between villages with high and low vector abundance to an overall accuracy of 90 %. In addition, the authors noted that such an analysis could be extrapolated spatially, albeit with a diminished accuracy, to the whole of the Chiapas coastal region where malaria is a significant health problem.

Related work by Rejmankova *et al.* (1995) showed that the density of adult *Anopheles albimanus* mosquitoes around villages in Belize could be reliably predicted using multispectral SPOT - HRV data (see table 1.1). Productive larval habitats were first identified as marshes containing relatively few emergent aquatic plants and a high coverage of cyanobacterial mats. An unsupervised Baysian maximum likelihood classification was then applied to a single SPOT HRV scene covering a test site in northern Belize. The classes generated were subsequently assigned to individual

"landscape elements" based on field observations. Human settlements were identified with ancillary map data and located more precisely on subsequent field visits with a Global Positioning System (GPS) (Herring 1996). These settlements were divided into two groups on the basis of their distance to the larval habitat class. Group 1 was composed of settlements closer than 500 m and group 2 of settlements further than 1,500 m. Based on previous measurements, a landing rate of greater than 0.5 mosquitoes per human per minute (during the hours of maximum mosquito activity from 6:30 to 8:00 p.m.) was used as a threshold for high adult mosquito density (Rodriguez *et al.* 1996). Group 1 was predicted as having adult mosquito densities higher than this threshold and group 2 lower. These predictions were tested by collecting mosquitoes landing on humans during the hours of peak activity within each of the settlements. The resulting predictions were 100 % accurate for group 2 and 89 % accurate for group 1.

1.5.3.3 Application of low spatial resolution imagery

Work in relation to mosquito vectors has not been restricted to high spatial resolution imagery. GAC NDVI data from the NOAA AVHRR have been applied to the problem of Rift Valley Fever (RVF) epidemics in Africa (Linthicum *et al.* 1987, 1990). RVF is a viral disease transmitted by various species of *Culex* mosquitoes to domestic animals and nearby human populations (Wilson 1994). The work on a regional scale showed that high NDVI values in Kenya were good indicators of seasonally flooded linear depressions in savanna habitats, known as dambos. These habitats were highly suitable for mosquito breeding and hence closely associated with RVF epidemics. On the basis of this initial work an RVF virus epidemic was correctly predicted in central Kenya following exceptionally high NDVIs throughout 1988 and 1989 (Bailey and Linthicum 1989).

The work progressed to incorporate higher spatial resolution Landsat - TM and SPOT - HRV imagery to locate individual areas of high RVF risk determined by the NDVI (Linthicum *et al.* 1991). Operational application of the technique was hindered, however, because investigators were not able to discriminate flooded from dry dambos. Data from an airborne Synthetic Aperture Radar (SAR) were therefore incorporated to detect dambo flooding status (Pope *et al.* 1992). A significant advantage of using SAR was that data collection was independent of cloud coverage, especially important during the East African rains. The study also demonstrated that the spatial resolution of

current satellite-borne SARs was not sufficient to reveal many of the smaller dambos in the region.

Linthicum *et al.* (1994) investigated a severe and extensive outbreak of RVF in the West African Senegal river basin in 1987. The RVF outbreak was particularly unusual in that it occurred during a period of only moderate rainfall. Analysis of SPOT - HRV scenes for the period of the epidemic revealed that extensive flooding in Mauritania peaked in October 1987, as a result of the construction of the Diama and Manatelli dams on the river Senegal. This coincided exactly with the period of maximum RVF disease activity in the area. Furthermore, maximum values of LAC NDVIs were associated with increased rice production (and hence productive mosquito larva habitats) around the newly flooded regions in Daro and Rosso, the foci of RVF outbreaks. Unfortunately, no statistics were presented in the study.

Thomson *et al.* (1995b, 1996) also investigated the potential of GAC and LAC NOAA - AVHRR data, as well as cold cloud duration (CCD) data from United Nations (UN) Food and Agriculture Organization (FAO) African Real Time Environmental Monitoring and Information System (ARTEMIS) (see 2.3.2) to predict malaria epidemics in the Gambia. They concluded, however, that although there were clear relationships between satellite data and environmental variables associated with malaria transmission, it was difficult to predict how these would affect adult mosquito abundance and behaviour. For instance, along the River Gambia a decrease in rainfall may at times increase available Anopheline breeding sites by increasing the number of suitable pools in the alluvial soils at the river margin. They also noted that relationships between malaria incidence and environmental variables were complicated by sociological factors, because in areas where Anopheline abundance was greatest (and hence biting most frequent) people were more likely to protect themselves with insecticide-impregnated bed nets (Thomson *et al.* 1995a).

1.5.4 Tick-borne disease

1.5.4.1 Disease and vector biology

The biology of ticks is quite different from that of insects. Ticks of the Family Ixodidae (see figure 1.3.3) take a blood meal only once per life-cycle stage, as a larva, nymph or adult (Sonenshine 1991). To act as vectors, therefore, ticks must transmit the pathogen between stages within the same generation (trans-stadially), and sometimes

from female to the larvae of the next generation via the eggs (trans-ovarially). Between meals, ticks drop from their hosts into the vegetation where they undergo long periods of development, lasting between one and twelve months depending on the ambient temperature. The survival and development rates of ticks, and therefore the transmission dynamics of the tick-borne pathogens, are thus directly determined by environmental conditions (Sonenshine 1993a, Daniel and Dusbábek 1994).

Ticks are second only to mosquitoes in the number and diversity of viral, bacterial and protozoan pathogens they transmit to humans and livestock (Sonenshine 1993b). For example, the three genera *Rhipicephalus*, *Hyalomma* and *Amblyomma* are vectors for protozoan parasites of the genus *Theileria*, the most important being *Theileria parva*, the

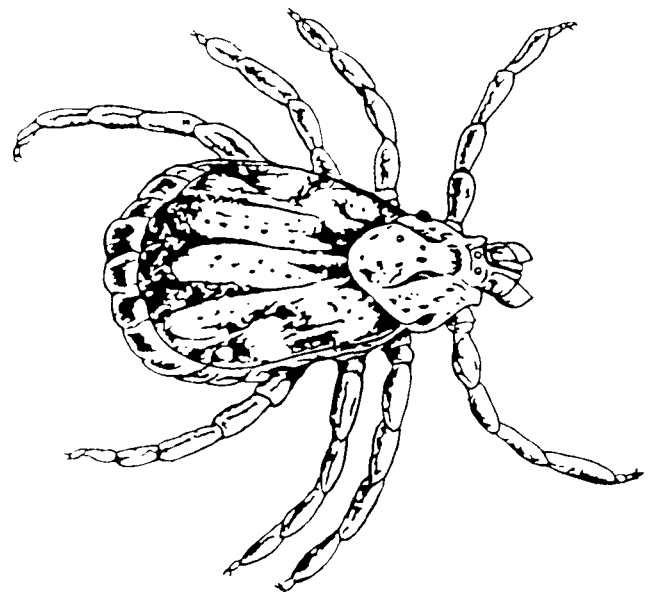


Figure 1.5. An adult female tick (Family Ixodidae). Redrawn from Molyneux (1993).

causative agent of East Coast Fever (ECF) in cattle. This disease has a major economic impact in Africa through the reduction in cattle numbers and productivity, with associated costs in human welfare (Mukhebi 1992). In temperate regions, ticks of the *Ixodes ricinus* complex are the major vectors for the bacterial spirochete *Borrelia burgdorferi*, responsible for Lyme borreliosis (Lyme disease) in humans throughout Asia, Europe and North America (Burgdorfer *et al.* 1982).

1.5.4.2 Application of high spatial resolution imagery

Amblyomma variegatum was accidentally introduced to the Caribbean on cattle imported from Africa where it is a vector for heartwater, a fatal tick-transmitted disease caused by the protozoan *Cowdria ruminantium.* Hugh-Jones and O'Neil (1986) were among the first to investigate the potential for Landsat - 1 - MSS imagery to identify tick habitats, specifically for *Amblyomma variegatum* in St. Lucia. The major regions of tick infestation were located in arid, rough grazing habitat in the north of the island, characteristically dominated by the mesquite *Prosopis juliflora.* The Landsat MSS images revealed that this habitat was expanding, reportedly as a result of poor grazing management, thus explaining the increasing incidence of heartwater in the region.

Hugh-Jones (1991b) extended this pilot study using Landsat - 5 - TM imagery to infer habitat quality for *Amblyomma variegatum* in Guadeloupe. Tick density was measured in 103 cattle herds, and associated environmental variables were recorded along transects across the surrounding grazing habitat. A discriminant analysis of these variables, which included plant composition, grazing cover, soil type and depth, slope and rainfall, distinguished four main tick habitats. These were lightly infested dry meadows, moderately infested foothills, heavily infested dry scrub and rocky grasslands. These four habitats could be discriminated remotely and were resolved by an unsupervised classification of a 1986 Landsat TM scene of Grand Terre, Guadeloupe (Hugh-Jones *et al.* 1988). Subsequently, a moisture index (MI) using bands (4-7)/(4+7) and the Perpendicular Vegetation Index (PVI) (Richardson and Weigand 1977) were derived for each pixel (Hugh-Jones *et al.* 1992). A 5 x 5 pixel array was then centred upon each visible farm and a cluster analysis performed on the tick density in each of the farms against the variance in the raw waveband data, the MI and PVI. It was found that herds in grazing areas with a high variance in the band values (indicating heterogeneous vegetation) had more ticks than areas with low variances (*i.e.* homogeneous vegetation). Furthermore, within the heterogeneous areas those with higher PVI and MI values (*i.e.* more vegetation and moisture respectively) had higher tick densities.

Daniel and Kolar (1990) investigated the possibility of forecasting the occurrence of *Ixodes ricinus* from Landsat - 5 MSS data. They worked on a scene centred on Poteply, a popular recreational area south-west of Prague and an historic focus of tick-borne encephalitis. Long-term field studies in the area enabled selection of pixels for a supervised maximum likelihood classification (incorporating bands 2, 3 and 4) separating

the scene into coniferous, broad-leaved and mixed forest classes, as well as water basins, glades and housing developments. Though no statistics were presented, the authors state that the broad-leaved and mixed vegetation classes were characteristic of high tick density areas, especially where the spatial heterogeneity of such classes was high, indicating the presence of ecotones considered particularly favourable for *I. ricinus*. This work has resulted in the production of "risk maps" for *I. ricinus* and tick-borne encephalitis in parts of the Czech Republic.

1.5.5.3 Application of low spatial resolution imagery

Predicting the distribution and abundance of the brown ear tick, *Rhipicephalus appendiculatus*, has been the focus of much recent research as it is the principle vector of ECF in Africa. The most widely used models have been based on the program CLIMEX which calculates the climatic suitability of geographical regions for arthropod species using a temperature dependent growth index, moderated by four user-defined stress indices (hot, cold, dry and wet) (Sutherst and Maywald 1985). Driven by a synoptic meteorological dataset the resulting predictions for the pan-African distribution of R. *appendiculatus* have been compared visually with the observed distribution of ticks (Lessard *et al.* 1990, Perry *et al.* 1990). Further visual comparison with mean maximum GAC NDVI values for 1987 provided a "remarkable visual correlation" between the known distribution of R. *appendiculatus* and NDVI values greater than 0.15 (Perry *et al.* 1991). Again, no statistics were presented in these studies but, using both the NDVI values and CLIMEX the authors predicted that suitable habitat for R. *appendiculatus* existed in western and central Ethiopia and therefore that the region was at risk from ECF in the future (Norval *et al.* 1991). Given that CLIMEX also predicted even greater suitability for ticks in much of the Zaire basin and many parts of West Africa, areas from which R. *appendiculatus* has never been recorded, the accuracy of the predictions from this model and the underlying methodology are open to question. The CLIMEX model has also recently been used to predict an expansion in the range of the mosquito *Anopheles farauti sensu stricto* (s.s.) vector for malaria in northern Australia, given a climate change scenario of a 1.5 °C average increase in temperature and a 10 % increase in summer rainfall (Bryan *et al.* 1996). Again no statistics were presented on the accuracy of predicting present day distributions using CLIMEX.

Randolph (1993, 1994) took an explicitly biological approach to understanding the abiotic constraints on the distribution of R. *appendiculatus*. Correlations between the

seasonally variable abundance of this tick in southern Africa and simultaneously monitored meteorological data and a more recently processed MVC GAC NDVI dataset for the region (1987 89) indicated that the major abiotic constraint was moisture availability to the desiccation-vulnerable egg and larval stages. This was confirmed when the author derived seasonal interstadial mortality indices from published datasets in both equatorial and southern Africa, and showed strong correlations between mortality at the female-to-larval stage and factors determining moisture availability. The critical climatic factor, however, that determines moisture availability to ticks varies geographically; in equatorial regions, rainfall and consequent soil moisture was most important, while in the temperate regions low daily minimum temperatures providing high condensation appear to allow good survival in the dry season. Indeed, in certain southern sites, conditions that are too wet result in high mortality, perhaps because of fungal infections of the eggs (Randolph 1996). The influence of these variable conditions presumably underlies the observed correlations between tick mortality rates and the NDVI at the time of larval emergence; the correlation was negative in equatorial Africa, where high rainfall and thus high NDVI is favourable to tick survival, but positive in Zimbabwe, where high minimum temperatures and related high NDVI are unfavourable.

1.5.5 Tsetse flies and trypanosomiasis

1.5.5.1 Disease and vector biology

Tsetse flies (*Glossina* spp., see figure 1.3.2) are an entirely African family of haematophagous flies, the Glossinidae, in which both sexes can transmit protozoan trypanosome parasites between vertebrate hosts while feeding (Ford 1971). They are unusual in being larviviparous, with each female giving birth to a single full-grown larvae once every 7 - 10 days. The larva then rapidly burrows into the soil, pupates, and the adult emerges approximately one month later (Buxton 1955). Thus tsetse birth rates and mortality rates are extremely low when compared with other arthropod vectors of disease (Rogers 1988a).

Fifty million people in Africa are reportedly at risk from contracting human trypanosomiasis or sleeping sickness (Kuzoe 1993) which, when left untreated, proves fatal due to the invasion of the central nervous system by the trypanosomes. Tsetse flies are also vectors for trypanosome species which infect cattle, causing animal trypanosomiasis or "nagana" (Hoare 1972). The occurrence of animal trypanosomiasis

excludes livestock from an area of approximately 10 million km^2 and is therefore a major constraint on livestock productivity in Africa (Murray and Gray 1984).

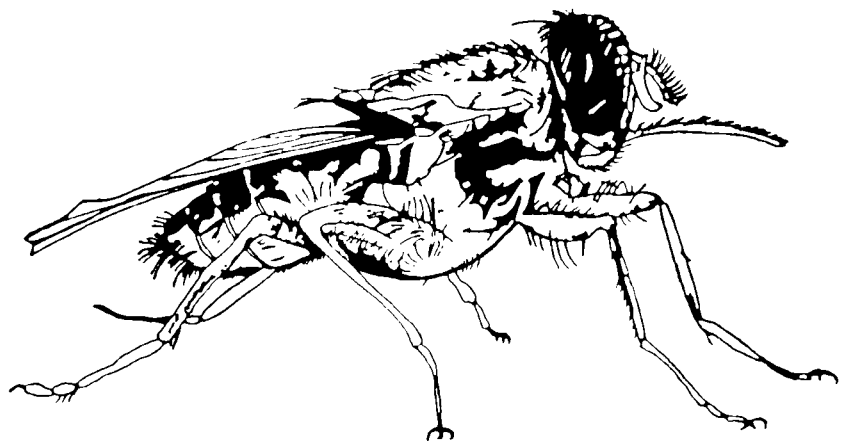


Figure 1.6. An adult female tsetse fly (Family Glossinidae). Redrawn from Molyneux (1993).

1.5.5.2 Application of high spatial resolution imagery

The NASA - HAO was again among the first to look at the potential of remote sensing for the identification of tsetse fly habitats (Giddings 1976). Near infra-red data from Landsat - 2 - MSS were used to map land-cover at a test site in Tanzania. Although such mapping was considered feasible, no conclusions were drawn on how this might help in tsetse control.

Kitron *et al.* (1996) analysed tsetse fly catches from sets of biconical traps set in the Lambwe Valley of Western Kenya during 1988 - 1990. They found that a multiple regression using the seven wavelength bands of the Landsat TM was able to explain 87 % of the variance in fly catch density. The Landsat - TM band 7 that is associated with soil water content was found to be consistently highly correlated, reflecting the importance of soil moisture in tsetse survival (Buxton and Lewis 1934, Rogers and Randolph 1986) The data were also used to calculate the NDVI, but this did not improve the relationship over that with the original bands. The spatial autocorrelation in the data was assessed using Moran's (*I*) statistic and found to explain a significant part of the association between the fly density and the spectral data. Kitron *et al.* (1996)

stressed that these spatial factors should be further investigated and on the basis of this work stated "*the incorporation of remotely sensed imagery into a GIS with ground data on fly density and environmental conditions can be used to predict favourable fly habitats in inaccessible sites, and to determine number and location of fly suppression traps in a local control programme*".

1.5.5.3 Application of low spatial resolution imagery

Rogers and Randolph (1991) explored the utility of GAC NDVI data derived from the AVHRR in a study of tsetse fly ecology and distribution in West Africa, since they considered the NDVI to integrate a variety of environmental factors of importance to tsetse survival. They found an inverse relationship between monthly NDVI and fly mortality rate in the Yankari game reserve in Nigeria and significant non-linear relationships between tsetse fly abundance and NDVI in the northern part of Côte d'Ivoire. Later, they focused on a 700 km transect running north-south through Côte d'Ivoire and Burkina Faso. This area is of particular epidemiological interest, since sleeping sickness is found only in the central region of the transect, despite the local vector (*Glossina palpalis*) occurring throughout.

The solution of this conundrum lay in the length of a particular vein in the tsetse's wing. This vein length is an indicator of overall fly size, which is determined by the environmental conditions experienced by the maternal fly while she is gestating the larva (approximately one month before the new adult fly emerges from its puparia). Mean vein lengths for flies caught at eight equi-distant sample sites along the north-south transect were related to the NDVI of the previous month for each site. During the wet season, the NDVIs across the transect were all high and fly size was uniformly large. In the dry season, however, fly size was strongly correlated with NDVI, with flies in the drier north significantly smaller than those in the wetter south. Mortality increases with decreasing fly size in tsetse and similar size differences have been shown to be related a significant increase in *Glossina palpalis* mortality in Kenya (Dransfield *et al.* 1989). These data were interpreted as indicating a geographical gradient in the degree of man-fly contact and thus trypanosome transmission potential. In the south, low mortality rates resulted in high densities of flies, but the flies were not nutritionally stressed (even seasonally) and so did not often resort to biting humans, who are not a favoured hosts. Conversely, in the north, fly populations suffered too high a mortality to pose a serious health risk. Only in the central areas was there an intermediate density of sufficiently stressed flies, resulting in a regional and seasonal focus of disease transmission.

Further work illustrated correlations between the monthly incidence of trypanosomiasis and mean monthly NDVIs (Rogers 1991). Positive correlations were shown in settlements surrounding the Lambwe Valley game reserve in Kenya and negative correlations in Kigulu County on the shores of Lake Victoria in Uganda. The differences were attributed to the contrasting ecologies of the respective tsetse vector species. In Uganda the vector *Glossina fuscipes* coexists with man throughout the year. The flies prefer other mammal hosts and only bite humans frequently, when they require more blood-meals, in the dry season. Transmission therefore increases as vegetation cover and photosynthetic activity (and hence the NDVI) decrease in the dry season, leading to a negative correlation. In contrast, the local vector *Glossina pallidipes* in Kenya spends most of its time in dense scrub thickets in the centre of the game reserve, but spreads out from the thicket, and comes into contact with humans that inhabit the reserve margins, when humid conditions allow fly dispersal in the rainy season. Transmission therefore increases as the NDVI increases in the wet season, resulting in a positive correlation. These initial studies showed that both tsetse distribution and abundance, and disease incidence and prevalence could be related to the NDVI at low spatial resolutions, although the interpretation of the variable correlations required a knowledge of local conditions and fly biology from ground studies.

Rogers and Williams (1993) describe the application of GAC NDVI data and synoptic meteorological temperature data to the problem of predicting the distribution of *Glossina morsitans* in Zimbabwe, Kenya and Tanzania. Temperature data (a critical climatic variable in determining the survival of tsetse (Bursell 1958, 1959)) were included in the analysis by interpolating data from meteorological stations to grid squares covering the whole of Zimbabwe. When these data were combined with NDVI variables in a linear discriminant analysis the historical distribution of *Glossina morsitans* in Zimbabwe, as illustrated by Ford and Katondo (1977), was predicted with an accuracy of over 80 %. Historical distributions were used because the authors considered them to reflect the original fly habitat more closely than present day distributions. This is because current ranges have been extensively modified by the elimination of game, the rinderpest panzootic at the end of the last century (that killed over 85 % of domestic stock), agricultural expansion and the activities of tsetse control campaigns (Ford 1971).

The utility of these statistical approaches was shown in two studies that investigated how tsetse distributions might be modified with global climate change

(Rogers and Randolph 1993, Rogers 1995). Rogers and Randolph (1993) looked again at the distributions of *Glossina morsitans* in Zimbabwe, Kenya and Tanzania. Using elevation, NDVI and synoptic temperature data they were able to predict the distribution to an overall accuracy of 82 %. The mean temperatures in areas suitable and unsuitable for tsetse was shown to differ be only 0 - 1 °C and hence the distributions were very sensitive to temperature changes. The scenario of increasing the average temperature by 1 °C throughout the region resulted in the predicted range of *Glossina morsitans* expanding into the presently unfavourable and livestock-productive highland areas of Zimbabwe.

Rogers (1995) continued this theme and contrasted the distribution changes predicted for *Glossina morsitans* and *Glossina pallidipes* in Kenya and Tanzania using a similar elevation, NDVI and synoptic meteorological temperature dataset. The distribution of *Glossina morsitans* was again found to expand, given scenarios of 1 - 3 °C warming, but the range of *Glossina pallidipes* contracted substantially. The sensitivity of such analyses was illustrated by taking elevation out as a predictor variable, causing the range of *Glossina pallidipes* to expand. Given such uncertainty, the author suggested that studies should consider the reasons for present and past distributions before extrapolating through time.

1.5.6 Conclusion and critique

Data from satellite remote sensing have shown potential to assist in the understanding and control of mosquito, tsetse and tick borne diseases. The degree of sophistication with which remote sensing has been applied to these diseases varies, and reflects the time for which remote sensing has been adopted by different groups of the disease vector control community. A general criticism of many of the early studies, and a number of more recent ones, is that they rarely report accuracy statistics in dissemination of work, so that the utility of the various techniques cannot be objectively compared.

Considerable advances have been made from the primary problem of habitat distribution mapping to the incorporation of aspects of vector ecology, but only a narrow range of the available satellite data and analytical techniques have been utilised. For instance, though the limitations of the NDVI have been widely acknowledged, none of the above investigations have explored the potential of alternative vegetation indices.

Historically this has been due, at least in part, to problems of obtaining access to the original satellite sensor data, although this situation has improved in recent years with much of the data now available in the public-domain (Mulchay and Clarke 1994, Justice *et al.* 1995).

The focus on high spatial resolution imagery is also inappropriate for the targeting of national and regional control programmes against vectors with continental distributions. Split-window corrected thermal channel data have not been widely investigated, nor methods for estimating surface moisture conditions and rainfall. Seasonal aspects of vector ecology have been similarly ignored and there is great potential for further investigation of such phenomena with low spatial, high temporal resolution imagery. The aim of this thesis is to explore the complete range of NOAA - AVHRR and Meteosat data which has become accessible to epidemiologists in recent years (Townshend 1994, Snijders 1995).

1.6 Scope and structure of the thesis

The information presented in this introduction is intended to provide a basic background for the ecologist and epidemiologist to understand the techniques of remote sensing and for the remote sensing specialist to understand the biological aspects of the problems addressed. The following chapters look at the central question of how accurately the range of low spatial, high temporal resolution NOAA - AVHRR and Meteosat satellite sensor data can be used to predict the spatial patterns of disease vectors, with particular reference to the tsetse fly.

The materials and methods described in chapter 2 expands on the introduction and details the datasets and analytical techniques common to the following analysis chapters. The processing applied to the NOAA - AVHRR and Meteosat data before distribution in the public-domain are discussed with an emphasis on the steps that affect data quality. The subsequent processing applied for the present work to produce temperature, atmospheric moisture, rainfall and vegetation indices are then outlined. The seasonal variation in these factors was considered to be of paramount importance in determining vector species ranges (and abundances within these limits) and the analytical technique of temporal Fourier analysis used to parameterise important aspects of this seasonality and ordinate the multitemporal data are documented.

The distributional ranges of the *Glossina* are determined by a combination of climatic factors and their affect on local vegetation (see section 3.1 for a discussion). To predict tsetse distribution therefore, information on the spatial and temporal dynamics of climate and land-cover are needed. To understand more of the information content of indices derived from satellite sensors, "verification" was performed at three major levels. The first step was to investigate how accurately these meteorological surrogates could predict climate across Africa throughout the year (chapter 3). The second step was to see how accurately the vegetation indices could be used to map land-cover at a regional scale (chapter 4). On the basis of these results the level of confidence with which these proxy data can be applied to tsetse distribution and abundance analysis can be assessed.

Chapter 3 is introduced with a brief discussion on how climate has been shown to be important in tsetse and disease vector ecology. How this understanding has enabled tsetse distributions to be predicted and interpreted is also discussed and the limitations

of using ground-based meteorological records are highlighted. Results are then presented on how accurately ground temperature, humidity and rainfall can be predicted using data derived from the NOAA - AVHRR and Meteosat - HRR radiometers. The "accuracy" was assessed by comparing monthly temperature, humidity and rainfall values published by NOAA for WMO member meteorological stations throughout continental Africa with variables derived from satellite sensor data. On the basis of these results the dataset was then refined and the best split-window and atmospheric moisture indices used in the vegetation and tsetse fly distribution and abundance work.

The next level of investigation was to determine the potential of the various published SVIs (discussed in the section 1.4) and other data for predicting regional scale vegetation coverage. This is described in chapter 4 using land-cover data for the whole of Nigeria, assessed by a low-level aerial survey in 1990. Conclusions are drawn on the relative performance of the SVIs and accuracies compared to with previous studies. Using these results several SVIs were excluded from the tsetse fly distribution and abundance analysis.

In chapter 5, the satellite sensor dataset refined on the basis of their relative performance in the previous two chapters, is used to predict the range and abundance of eight tsetse species surveyed in the late 1970s and early 1980s in Côte d'Ivoire and Burkina Faso. The relative importance of the various data layers is discussed and compared with current understanding on the factors that limit the range and abundance of tsetse species.

These ideas are then brought together in a general conclusions and discussion section in chapter 6, where the wider application of these techniques and their potential for use in disease vector control programmes are considered.

Chapter 2

MATERIALS AND METHODS

" ... The greatest invention of the nineteenth century was the invention of the method of invention ... "

A. N. Whitehead.

2.1 Introduction

The data used in this investigation are diverse in their origin, as are their methods of acquisition and the subsequent processing to which they are subject. For clarity therefore, the ground data (section 2.2) and satellite sensor derived data (section 2.3) are treated separately. Data and processing techniques used by each of the three subsequent analysis chapters are documented. The specific ground datasets used in the "verification" of these data, along with the analytical and statistical techniques applied are described, where relevant, in later chapters. For each of the "common" datasets, details are presented on where and how the data were collected, as well as information on any processing performed prior to dissemination. Section 2.4 then explains how the satellite sensor data were further manipulated to create the vegetation, land surface temperature, atmospheric moisture and rainfall indices for the purposes of this thesis. The final section (2.5) deals with the dimension reduction technique of temporal Fourier analysis used to ordinate the multitemporal satellite sensor imagery. A preliminary discussion is also included on why the technique was favoured over the canonical principal components analysis approach.

Data processing was done using the Earth Resources and Data Analysis Software (ERDAS) Imagine 8.2. software package on a Silicon Graphics Indy workstation running the Irix 5.2. operating system. Where this package was unable to perform the processing required, Quick Basic programs written by Dr David Rogers (DJR) were used.

2.2 Ground data

2.2.1 Digital elevation model

A Digital Elevation Model (DEM) of Africa was obtained from the Global Land Information System (GLIS) of the United States Geological Survey (USGS), Earth Resources Observation Systems (EROS) data centre (GLIS 1994). The DEM was produced from topographic data digitised from 1:1,000,000 scale air navigation charts from the U.S. Defense Mapping Agency (DMA) Operational Navigation Chart (ONC) series (DMA 1992). All spot heights from the maps were included, as well as selected points on contours and coastlines. These data were then spatially interpolated using the Australian National University Digital Elevation Model (ANUDEM) described in Hutchinson (1988, 1989). The hydrological network on the DMA - ONC maps was also digitised and used to spot anomalies in drainage conditions (usually arising from spurious inland sinks) created by the interpolation process.

The accuracy of the map has not been measured or calculated, but must be restricted by the ± 650 m vertical error at 90 percent confidence defined by the DMA for the ONC map series (DMA 1992). The measure of confidence in the DEM values has been assessed by comparing the elevations reported for the 207 meteorological stations used in chapter 3 (see section 3.2.1) with those extracted from the GLIS DEM for the same locations. Figure 2.1 shows the highly significant linear relationship (adjusted $r^2 = 0.95$, $n = 207$, $P << 0.0001$).

The GLIS Africa DEM was distributed as four binary files with elevation expressed in feet above mean sea level and registered to a Plate Carrée projection at a 30 arc-second (approximately 1 x 1 km) spatial resolution. The files were first imported into ERDAS Imagine 8.2. using image dimensions supplied with each of the files and mosaiced into a single whole Africa scene. The elevations were then converted into metres. Twenty five ground control points were chosen from distinctive promontories and features around the coastline and used to resample a mean of a 3 x 3 DEM pixel array to a Lambert azimuthal equal-area projection at an 8 x 8 km spatial resolution for compatibility with the satellite sensor data. The Lambert azimuthal equal-area projection was chosen because historically the GAC AVHRR dataset was in this format.

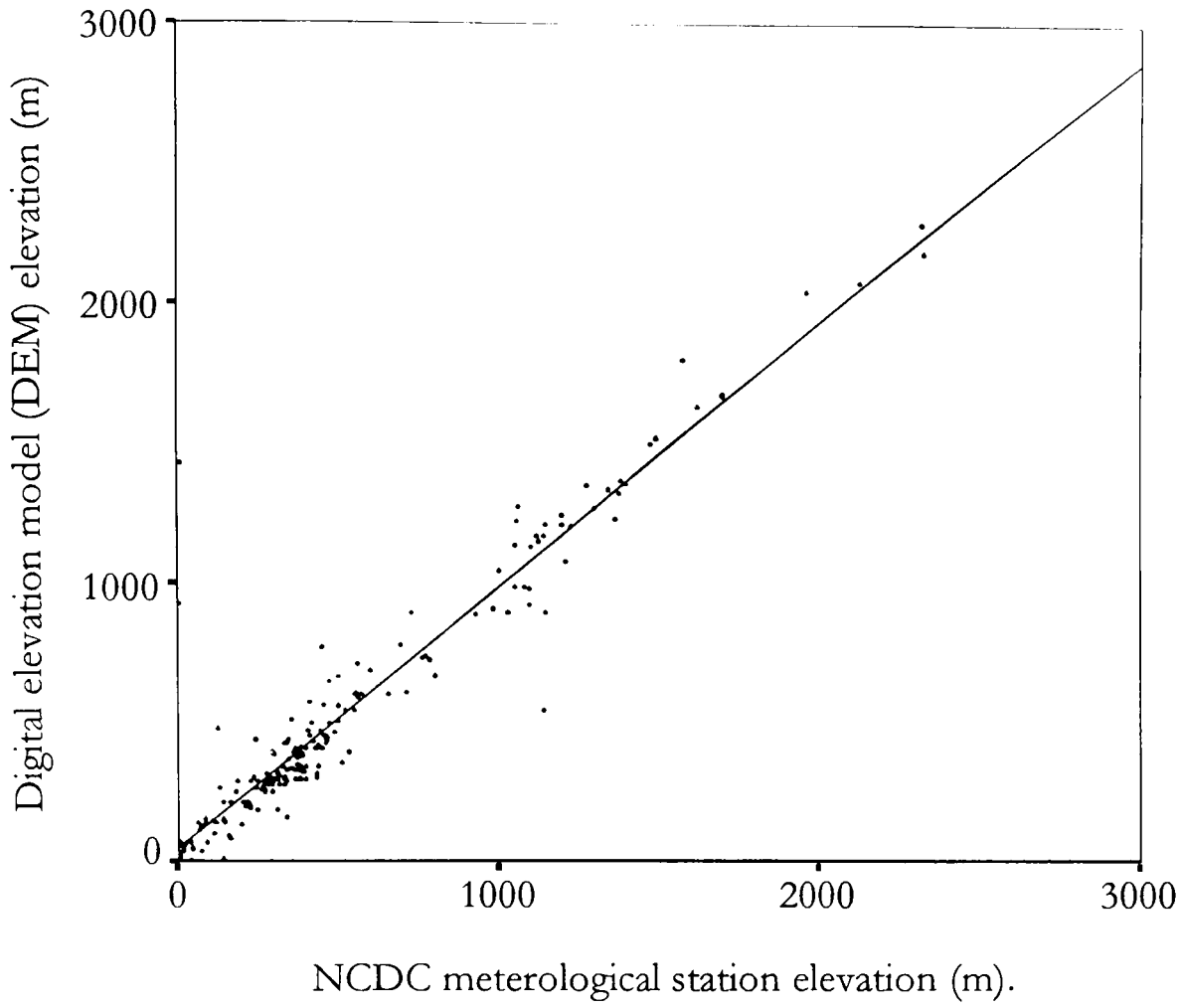


Figure 2.1. The comparison of National Climate Data Centre (NCDC) meteorological station elevations with those extracted from the GLIS DEM for the same geographic locations. The relationship is highly significant ($P << 0.0001$, adjusted $r^2 = 0.95$, $n = 207$).

2.3 Satellite sensor data sources and pre-processing

2.3.1 Meteosat HRR data source

The Meteosat HRR data were supplied by the FAO ARTEMIS in Rome as monthly CCD images from October 1988 until December 1995.

2.3.2 The Meteosat HRR pre-processing chain

Geo-correction and calibration of Meteosat HRR imagary is performed centrally by the EUropean METeorological SATellite programme (EUMETSAT) before being distributed to one of a global network of Primary Data User Stations (PDUS) (EUMETSAT 1995). It is these data that are received by FAO - ARTEMIS and converted to a Hammer-Aitoff projection and calibrated using the EUMETSAT recommendations (NRL 1988).

The particular threshold temperature associated with rain-bearing clouds and the quantity of rain they deposit varies temporally and spatially and must be established empirically (Rosema 1990, Milford and Dugdale 1990). This has been done by the Tropical Applications in Meteorology of SATellite and other data (TAMSAT) programme for the area shown in figure 2.2 (Snijders 1991). The northern and southern extent of the calibration exercise represent the northern and southern most limits of the Inter Tropical Convergence Zone (ITCZ) (Dugdale *et al.* 1995). These results are now used by the FAO - ARTEMIS project to generate decadal and monthly CCD images, where each pixel represents the number of hours for which it was covered by cold clouds during the compositing period (measured by channel 2 of the Meteosat HRR). There are set thresholds of -50 °C in the summer and -60 °C in the winter, used for areas north of the TAMSAT latitudes and -40 °C throughout the year for regions south (Snijders 1996). The images were distributed in Image Display and Analysis (IDA) format (Pfirman and Hogue 1995) and consisted of CCD data scaled in units of 0.5 hours per month to 8 - bit and registered to a modified Hammer-Aitoff projection at a 7.6 x 7.6 km spatial resolution. These data were imported into ERDAS Imagine 8.2 using image dimensions determined in IDA and converted to 16 bit as follows;

$$CCD = (DigitalNumber \times 4) + 3.$$

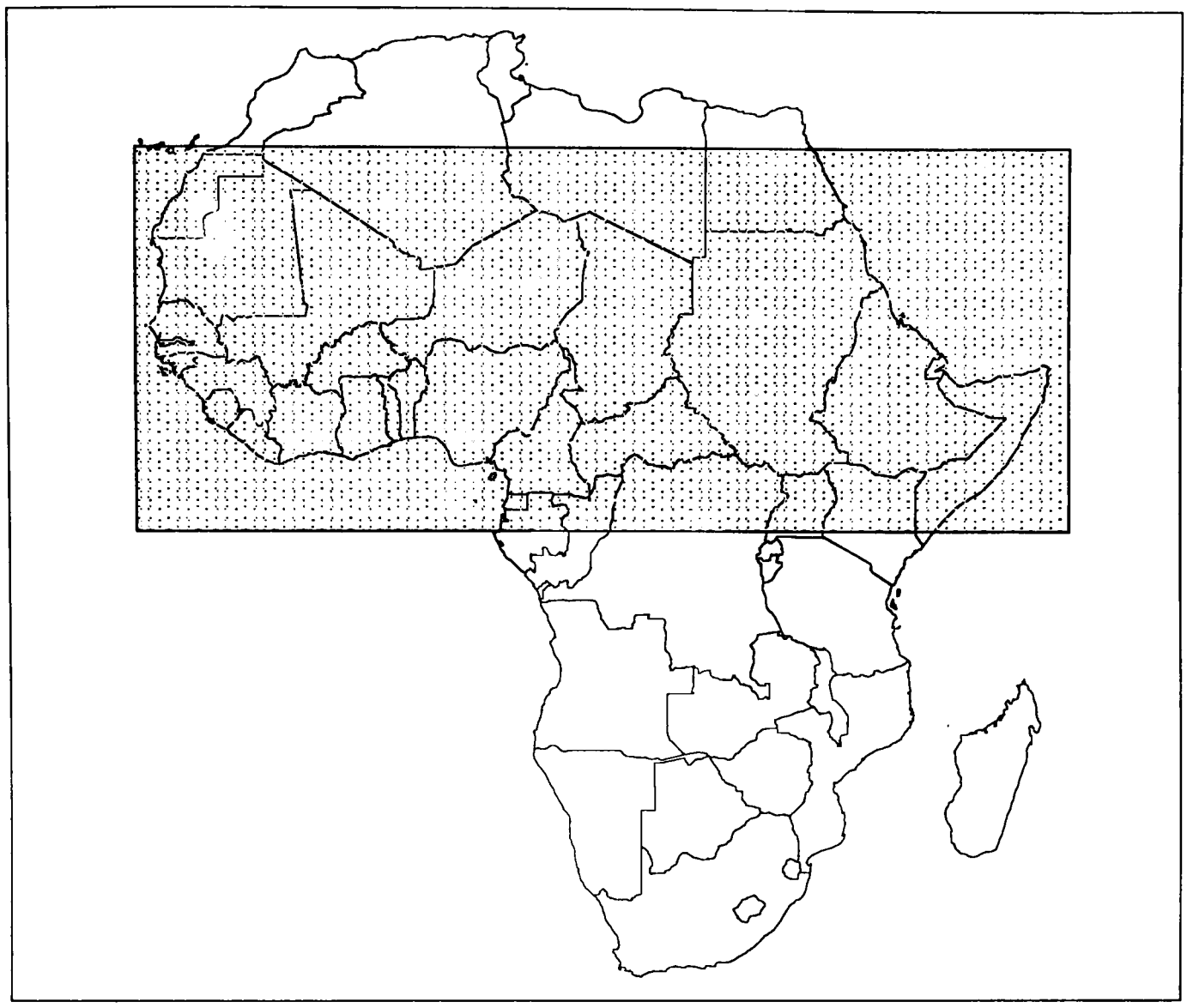


Figure 2.2. The geographical extent of the FAO - ARTEMIS calibrated area for CCD image generation. Redrawn from Dugdale *et al.* (1995). The northern limit is approximately 27° N and the southern limit the equator.

2.3.3 The NOAA - AVHRR data source

The Pathfinder AVHRR Land (PAL) dataset 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, Knox 1995). The PAL data were derived from visible and infrared radiance imagery collected by the AVHRR sensors onboard the afternoon ascending node satellites, namely NOAA - 7, NOAA - 9 and NOAA - 11 (Kidwell 1995). The dataset was produced from the historical GAC raw data spanning the period July 1981 to 1994. Subject to funding the ultimate temporal coverage will be until December 1998 when the superseding Earth Observation Satellite (EOS) AM is scheduled for launch. The definitive description of the PAL dataset is given in Agbu and James (1994) and James

and Kalluri (1994). The following section describes in detail the processing performed prior to the distribution of this dataset. This is important because many of the techniques applied can have significant affects upon data quality.

2.3.4 The PAL pre-processing chain

A schematic diagram of the PAL data processing chain is given in figure 2.3. Images from the AVHRRs onboard the NOAA 7, - 9, and 11 satellites with afternoon ascending node crossing times were taken from archives and stored on computer. Ancillary data needed in processing were retrieved and appended to these data. This included ozone data from Nimbus 7 Total Ozone Mapping Spectrometer (TOMS), land surface elevation from the Earth Topographic Five Minute Grid (ETOPO5) dataset (NGDC 1993), a land sea mask, and satellite orbit ephemeris files.

Each pixel was then navigated with an orbital model and updated ephemerides information. Time offsets were used to update the onboard clock information for each scan line and the navigation model of Baldwin and Emery (1993), stated to provide navigation accuracy to within one 4 x 4 km GAC pixel, was used to propagate ephemeris data using these updated clock values. The final geolocation of each pixel then used a closed-form geolocation algorithm developed by Patt and Greg (1994) and the resulting latitudes and longitudes were recorded for each pixel. The angular relationships between the satellite sensor, the Sun and each ground pixel were also calculated since they can have significant effects on data quality. Figure 2.4 demonstrates the angular features calculated including the solar zenith, solar azimuth, scan, and relative azimuth angles.

The variable stability and accuracy of the AVHRRs onboard the NOAA 7, - 9 and - 11 satellites are well documented (see section 1.3.4). Channels 1 and 2 were therefore calibrated following the procedures outlined in Rao (1993a) which provide offsets and gains based on calibration trends determined by U2 aircraft underflights of NOAA - 9 over the southeastern Libyan desert in 1986 and 1988. The TIR channels (channel 4 and channel 5) were calibrated using the methods described in Rao (1993b). These recommendations take into account work that has re-analysed pre-flight calibration data and that models the non-linear response of the detectors (Brown *et al.* 1985, Weinreb *et al.* 1990). The MIR channel 3 calibration was performed using the procedures described in Kidwell (1995).

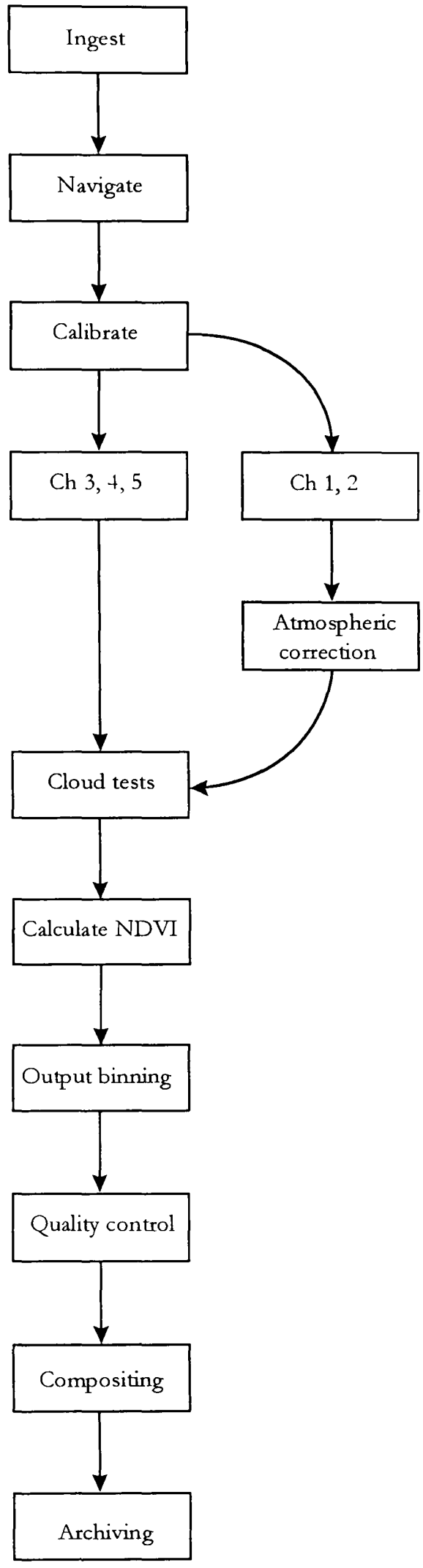


Figure 2.3. A diagram of the Pathfinder AVHRR Land data processing chain.

Atmospheric corrections were then applied to the visible channels. It was decided to correct only for Rayleigh scattering (the scattering of EMR by particles that are much smaller than the wavelength of the EMR concerned) and ozone absorption because while there had been significant advances in understanding the effects of water vapour and aerosols on AVHRR data (Vermote *et al.* 1990, Tanré *et al.* 1992), there had been little experience of applying these techniques globally. The Rayleigh correction is calculated using the standard radiative transfer equation and methodology outlined in Gordon *et al.* (1988). This also includes correction for the reduced solar irradiance resulting from the absorption of the ozone layer using daily ozone data from TOMS following Fleig *et al.* (1983). In addition, the pixel elevation as determined from the ETOPO5 dataset was used to correct the pressure level used in the calculation of the Rayleigh coefficients.

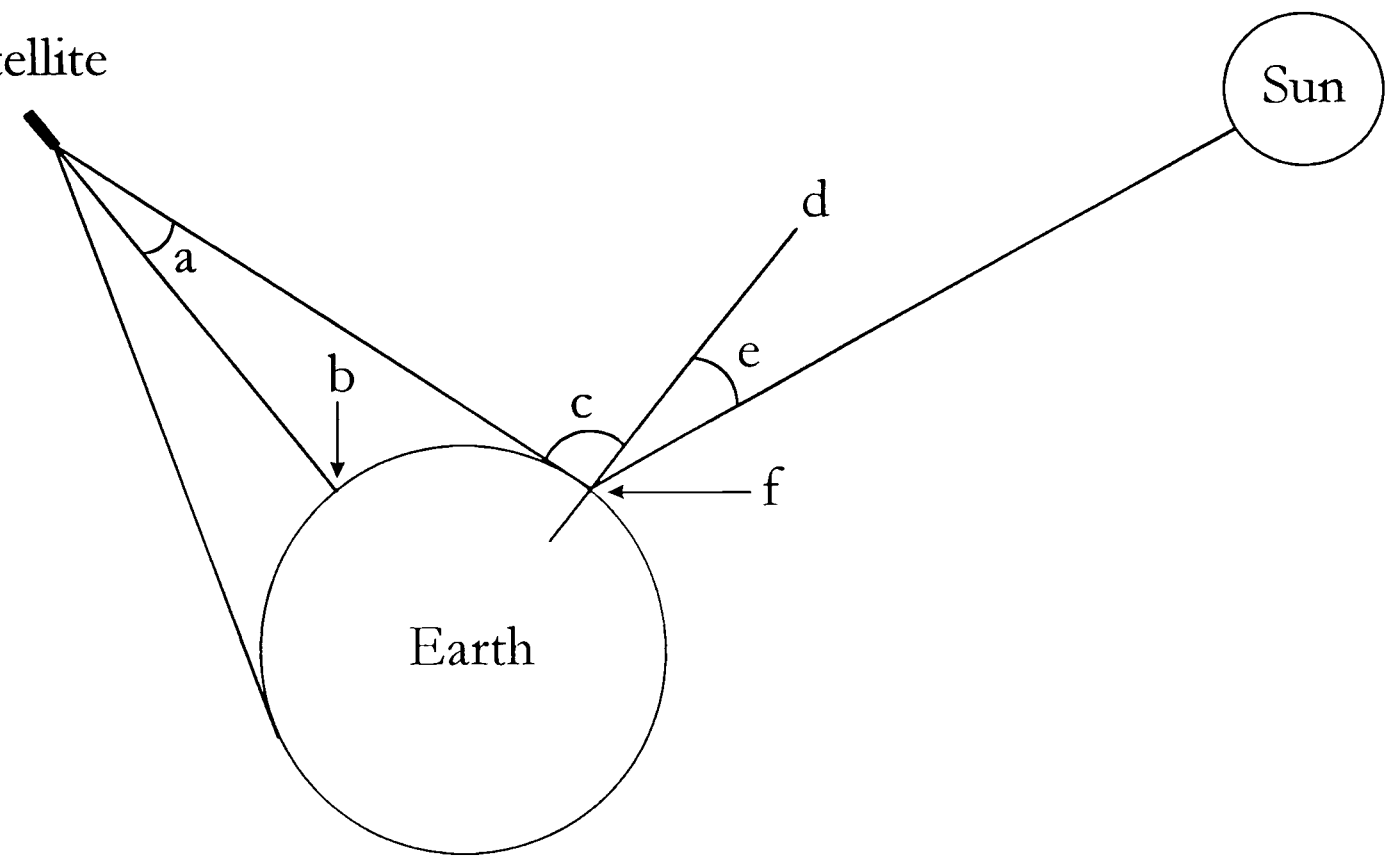


Figure 2.4. The angular relationships of a satellite to the Earth and the Sun. a: satellite scan angle, b: satellite subpoint (nadir), c: satellite of local zenith angle, d: zenith line, e: solar zenith angle, f: scan point.

An estimate of cloud contamination using the NOAA Clouds from AVHRR (CLAVR) algorithm (Stowe *et al.* 1991) was then calculated for each pixel. The CLAVR algorithm uses the five-channel multispectral information of the AVHRR to perform a series of tests on a 2 x 2 array of pixels using thresholds derived from sample data over a variety of land-cover types. The technique is based on the differences between the radiative and physical properties of clouds and the underlying surface. The specific thresholds and tests used by CLAVR are shown in table 2.1. If 1 – 3 pixels in the array are flagged as cloudy, all four pixels are flagged as mixed. Otherwise the array is flagged as clear or cloudy. Zero values mean no decision was made because of missing data. Gutman and Ignatov (1996) have shown that use of the CLAVR flags significantly improved the accuracy in each of the AVHRR channels using a time series of 10-day composit data.

The data are then mapped to a global 8 x 8 km equal area grid using the Goode Interrupted Homolosine projection (Steinwand *et al.* 1992, Steinwand 1994). The output resolution was chosen to maintain highest possible spatial resolution using data up to 42 degrees off nadir and the projection because it was good at reducing the distortion of the major land masses while maintaining equal area world-wide.

Each 4 x 4 km GAC pixel therefore had to be "binned" into one of the 8 x 8 km pixels of the output map. This was done using a nearest-neighbour, forward-binning approach where the centre latitude and longitude of each satellite sensor pixel is used to determine the output map bin. The differences in scale and orientation between the data surfaces resulted in as many as 6 pixels mapping to one output bin at nadir locations and, at the edge of swaths, output bins existing for which there were no input pixels. This results in a "watered silk" or Moiré pattern of null values and the small data gaps of two or fewer pixels are filled with the value of the pixel to the east of the gap.

The remaining consideration in the forward mapping was how to select the pixel values retained in the output bin when at high latitudes there was orbital overlap, and at latitudes near nadir the input pixels were significantly smaller than the output bins. The pixel selected as representative for that bin was the pixel with the greatest NDVI value for those pixels within 42 degrees of nadir since maximum NDVI pixels generally represent the pixels with the clearest atmosphere (Holben 1986). Data for solar zenith angles greater than 80 degrees were not used in the daily or composite data because erroneous NDVI values were often found along the edge of the swath (Holben 1986)

and the accuracy of the measured reflectances in twilight areas is very low (Goward *et al.* 1991). All data flagged as ocean were then removed.

TABLE 2.1.

Clouds from AVHRR (CLAVR) cloud detection tests with showing daytime land surface thresholds (Stowe et al. *1991).*

Cloud test	*Thresholds and values for the 2 x 2 array*
Reflectance gross cloud	If channel 1 reflectance is > 44 %, the array is flagged as mixed (1 - 3 pixels) or cloudy (all 4 pixels flagged).
Reflectance uniformity	If the difference between the minimum and maximum channel 1 of the array is > 9 %, the array is flagged as mixed.
Reflectance ratio cloud	For each pixel in the array, the ratio of channel 2 to channel 1 is calculated. If the ratio is between 0.9 and 1.1 the array is flagged as mixed or cloudy.
Channel 3 albedo	The channel 3 albedo is calculated for each pixel, and if the albedo is > 6 % the array is flagged as mixed or cloudy.
Thermal uniformity	If the difference of the minimum and maximum channel 4 in the 2 x 2 array is > 3 °K, the array is flagged as mixed.
Four minus five	A threshold, which is a function of channel 4, is calculated. If the difference between channel 4 and channel 5 is > the threshold, the array is flagged as mixed or cloudy.
Thermal gross cloud	If channel 4 is < 249 °K, the array is flagged as mixed or cloudy.
Channel 3 albedo restoral	The channel 3 albedo is calculated for each pixel, and if the albedo is < 3 % the array is restored to the clear category.
Thermal uniformity restoral	If the maximum and minimum channel 4 difference is < 1 °K, the pixels are restored to the clear category.
Thermal gross cloud restoral	If the channel 4 is > 293 °K, the pixels are restored to clear.

Once all the orbits from a single day had been processed, the PAL daily products were quality controlled for gross errors by visual inspection. Furthermore, a quality control flag was appended for each pixel to highlight potentially erroneous pixels resulting from the various conditions shown in table 2.2. When ten days of daily data had been processed, they were composited by choosing values for each bin, based on the day which had the highest NDVI value. Only those pixels within 42 degrees of nadir

were used in the production of composites. These data were then archived and made available for distribution in the public domain.

TABLE 2.2.

Pathfinder AVHRR Land (PAL) processing data quality flags.

Status	*Example condition*
Channel 1, 2 processing non-standard	Ozone value not available so average value used.
Channel 3, 4, 5 processing non-standard	Calibration coefficients unavailable.
Interpolated data	Filled gaps resulting from forward transform used in binning.
Range check failure	Calculated values were outside the range of acceptable data values.
NOAA Quality control flag	Quality control flags added by NOAA (see Kidwell 1995).

2.3.5 Potential limitations of the PAL data product

Prince and Goward (1995) attempted for the first time to estimate global net primary productivity using satellite sensor data alone as inputs to a GLObal Production Efficiency Model (GLO-PEM). During this effort they became aware of several problems with the Pathfinder dataset (Prince and Goward 1996). Those problems specific to the composited datasets are not further discussed as only daily data were used in this work. Two minor errors were that the Rayleigh and ozone corrections were underestimated and that solar radiation was not normalised for solar zenith angle, before being used to calculate the reflectances from the channel 1 and 2 radiances. Software were later provided to correct for these mistakes, but due to additional problems discovered with the coding of the solar zenith angle have been withdrawn (Kalluri 1996). A more fundamental problem involved the failure to flag radiometric temperatures data above 320 K at which the AVHRR sensor saturates. These problems are considered more fully in later chapters.

2.4 Satellite sensor data processing

This section deals with the processing applied to manipulate the satellite sensor data for the purposes of this investigation.

2.4.1 Meteosat HRR data processing

There was no subsequent processing required for the CCD data. Each of the monthly images was checked for gross errors and none were found.

2.4.2 PAL data processing

2.4.2.1 Stage I: extracting African data

The PAL images are shipped from the GSFC DAAC as global coverages containing the 12 bands of information detailed in table 2.3. 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.

A month of daily data were read from the Digital Data Storage (DDS) cartridges supplied by the GSFC DAAC into a single directory on the computer. Batch files were then run to uncompress the HDF files and extract the African continent from the global data. The 12 bands of PAL data (see table 2.3) for each day of the month were then extracted from the daily Africa continent files. The remaining global data were then deleted and the monthly data for Africa compressed and archived.

These steps were repeated until all the daily data for 1988, 1989 and 1990 were processed; a total of (228 x 365 x 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 the aerial survey of land-cover data for Nigeria used in the analysis in chapter 4.

2.4.2.2 Stage II: conversion of binary data into geophysical values

The observation date and relative azimuth angle data layers were deleted to save space, since they were not required in further analysis. The monthly data were then sorted into two 10 day (*decade_01* and *decade_02*) and depending on the month, a variable third 8, 9, 10 or 11 day decade (*decade_03*) directories. The approximately 100 files per decade formed a manageable unit for subsequent processing. Daily binary files were then imported into ERDAS Imagine 8.2. using image dimensions provided in Agbu and James (1994).

TABLE 2.3.

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

Band	*Data layer*	*Units*	*Offset*	*Gain*	*Geophysical minimum*	*Geophysical maximum*
1	NDVI	ratio	128	0.008	-1	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

Programs were written using the ERDAS Imagine 8.2. spatial modeller utility to convert the scaled binary PAL data into real geophysical values. The band 2 CLAVR data and band 3 quality control (QC) flags were exceptions however, as they were thematic data and not scaled prior to distribution. The scaling values are detailed for each band in table 2.3. All output data were stored to 32 bit single precision floating point accuracy (hereafter referred to simply as floating point accuracy).

Band 4 containing the satellite zenith or scan angle (SCA) data was re-scaled and converted from radians to decimal degrees using;

$$SCA = (((x - 10481.98) \times 0.0001) \times 57.3)$$

and band 5, containing the solar zenith angle (SZA) data, was re-scaled and converted from radians to decimal degrees using;

$$SZA = (((x - 10) \times 0.0001) \times 57.3).$$

The PAL mask values of 0 for missing data over land, 1 for missing data over oceans or lakes and 3 for Goode's Interrupted space were recoded to -101, -102 and -103 respectively. The visible channel data (bands 7 and 8) were then converted into percentage radiances values using;

$$Ch_1, Ch_2 = ((x - 10) \times 0.002).$$

Pathfinder mask values were recoded in the same way as the angular data. Cloudy and mixed pixels as determined from the CLAVR data layer were also eliminated and such pixels given a mask value of -105. This was a conservative threshold as highly reflective clouds with high radiance values would be preserved in subsequent MVC of channel 1 and 2 data. The angular data were then used to exclude pixels viewed at over 42° from nadir or at solar zenith angles of greater than 80° (see section 2.2.1.2) and were recoded to -106 and -107 respectively. The thermal channel data (bands 9, 10, 11) were then converted into brightness temperatures in degrees Kelvin using;

$$Ch_3, Ch_4, Ch_5 = ((x - (-31990)) \times 0.005).$$

Further processing for these data was identical to that used for the visible channels except that a less conservative threshold from the CLAVR algorithm was used, allowing mixed pixels to remain in the images. This was done because cloud contaminated pixels

were largely eliminated in the subsequent MVC and there was a danger of excluding uncontaminated data at this stage. Finally, binary NDVI data were converted to geophysical values using;

$$NDVI = ((x - 128) \times 0.008)$$

and the subsequent processing was identical to that applied to the thermal data.

The spatial models were used to generate scripts which were then iterated using batch files to complete a month of data processing at a time. These data were then compressed and archived and the process repeated until the data for 1988, 1989 and 1990 were completed. This method of processing monthly batches of data was also applied to Stage III of the processing.

2.4.2.3 Stage III: deriving epidemiologically useful products

This section describes the critical stage of turning the NOAA - AVHRR data into vegetation, temperature and atmospheric moisture indices.

2.4.2.3.1 Spectral vegetation indices

The RVI, the SAVI with the L variable set to 0.5 for intermediate vegetation coverage and the GEMI were calculated from the visible channel data (see section 1.4.1 for details). It was not necessary to calculate the NDVI as it was supplied as one of the bands of the original PAL dataset. The mask values were maintained for each data layer and outputs saved to floating point accuracy. SVIs that required the calculation of a soil line were not included in the analysis due to the problem of obtaining sufficient 8 x 8 km pixels of bare soil over all areas of Africa.

2.4.2.3.2. Land surface temperature indices

Section 1.4.2 explained the theoretical basis of determining land surface temperatures from split-window techniques. These techniques often rely on ancillary data to quantify atmospheric water content and surface emissivity and those that required anything other than AVHRR data were excluded. In this analysis when emissivity, ε, was a required variable 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;

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

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. Where emissivity values were required at both the channel 4 and 5 wavelengths in the equations they were assumed to be equal.

Of the many split-window algorithms available (Becker and Li 1995), three were applied to the thermal channels data because application to geographically extensive satellite image datasets has been shown or the equations were derived with consideration for global application. Price (1984) derived from radiative transfer theory one of the first indices to estimate land surface temperature, T (K), 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;

$$T = Ch_4 + A(Ch_5 - Ch_4)$$

and, A, is a constant determined by Price to be 3.33 for channels 4 and 5 of the NOAA 7 AVHRR when channel 4 and 5 emissivities were assumed equal. The abbreviations and units defined for the terms in this equation are identical where they occur in following equations and so are not defined repeatedly. This equation was stated to provide land surface temperature estimates accurate to ± 2 - 3 K after modelling potential error sources. This algorithm was later found to be accurate to ± 3 K using LAC data for a uniform tallgrass prairie habitat in Kansas when a constant emissivity was assumed (Cooper and Asrar 1989) and subsequently by Sugita and Brutsaert (1993) to be accurate to ± 4.5 K using LAC data from a similar habitat. This algorithm has furthermore been adopted in the creation of global AVHRR datasets for the European community (Malingreau and Belward 1994).

The following land surface temperature index has found preliminary application to multitemporal land surface temperature estimation of Asia (Gutman 1993). It is based on the formulation of Becker and Li (1990b) and again assumes equal emissivities at the channel 4 and 5 wavelengths;

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

where, δw is a correction for water vapour attenuation;

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

and, $\delta\varepsilon$, is a correction for emissivity;

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

where

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

In this analysis emissivity is 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 here. The Becker and Li (1990b) algorithm has also been used to provide land surface temperature estimates globally for inputs into primary production models (Prince and Goward 1995).

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. In its formulation however, an effort was made to isolate atmospheric effects by using atmospheric temperature and moisture profiles from TIROS Operational Vertical Sounder (TOVS) so that the algorithm could be applied globally provided the emissivity of the terrain was known (Prata 1993). The equation is as follows;

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

Again in the calculation of this index the emissivity at the channel 4 wavelength, ε_4, was assumed equal to the emissivity at the channel 5 wavelength, ε_5.

2.4.2.3.3 Atmospheric moisture indices

The total precipitable water content of the atmospheric column, U, was calculated using the Dalu (1986) equation presented in section 1.4.3. Total precipitable water was also calculated using the equation from Eck and Holben (1994) where;

$$U = A + B(Ch_4 - Ch_5)$$

and, A and B are constants, 1.337 and 0.837 respectively. These were determined by a linear regression of (channel 4 - channel 5) 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 perceptible water content of the atmospheric column (U) is often expressed as kgm^{-2}. These are units of pressure (*i.e.* 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), was then converted to a near surface dew point temperature, Td (°F), or the temperature to which a sample of air must be cooled for it to become saturated and condense using the following relationship (Smith 1966);

$$Td = \frac{\ln Pw - ((0.113 - \ln(\lambda + 1))}{0.0393}$$

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 λ = 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 values were then converted into Kelvin and used with the Price (1984) estimate of land surface temperature, Tp (K), to calculate the vapour pressure deficit, Vpd (KPa), using the equation provided in Prince and Goward (1995) where;

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

2.4.2.4 Stage IV: temporal compositing

The fifteen bands of data, including the original AVHRR channel information and derived indices, were maximum value composited (see table 2.4). Note that the band numbers are reassigned as some data from the original PAL dataset are not used in the analysis. The MVC was performed using the data values within a specific band and without reference to the NDVI (see section 1.3.2). These decadal data were then composited into maximum monthly files by taking the maximum value for a given pixel over the three decade files. The mask values were again preserved and outputs saved to floating point accuracy.

2.5. Satellite sensor data ordination

The multitemporal satellite sensor derived dataset generated (see table 2.4) results in multivariate data for each unit area (pixel) within an image. For example in a single band (of the fifteen available to the analysis) 3 years of data result in 36 monthly data points, each of which forms a single axis in a multivariate space. Many of these axes will be strongly correlated with each other because, for example, a pixel with a high NDVI in one month is likely to have a high NDVI in other months. This suggests that data reduction (*i.e.* ordination) could be achieved, without a significant loss of information, by replacing the monthly data with a combined signal derived from these highly correlated values. This of course would be a considerable advantage since 3 years of monthly data in fifteen potential bands create a dataset of (3 x 12 x 15) = 540 dimensions.

2.5.1 Principal components analysis

The simplest combination of these data would be the arithmetic mean, and seasonal variability could be captured by the variance or standard deviation of the mean. More complex techniques generally involve ordination projecting the data onto a rotated (usually orthogonal) set of axes called "principal components", such that the first new axis captures the largest proportion of data variance, the second captures the largest proportion of the remaining variance, and so on. Principal components analysis (PCA) retains the same number of axes as the original dataset, but the sequential partitioning of the variance often means that many of the axes in principal component space explain only a very small proportion of the variance in the original dataset (Green 1978).

Projection of the original dataset onto the principal component axes involves applying a series of coefficients or weights to the raw data, effectively to achieve the desired axis rotation (the weights are the cosines of the angles between the original and rotated co-ordinate axes). Data values in the original co-ordinate system all contribute (*via* their weighted values) to each principal component axis score. Thus, for example, a series of 12 monthly images from January to December may be subjected to principal components analysis and every month would then contribute to each of the 12 resulting principal component axes. If, however, only the first few principal component axes explain the great majority of the variance in the original data, only these need be used in further analysis.

PCA has been widely applied to reduce the data handling demands of multitemporal satellite imagery (Eklundh and Singh 1993). For example, analysis of an NDVI time-series usually gives a first component correlated with the mean vegetation index for the year, with the second and third components related to aspects of seasonality (Eastman and Fulk 1993). These components have in turn been used for land-cover classification (*e.g.* Townshend *et al.* 1985, Townshend and Justice 1986) although their biological interpretation is less than transparent.

A further problem is that PCA is not independent of the scale of the original axes and it is generally necessary to standardise (or transform) the raw variables to roughly similar variances before analysis, or else to use the correlation matrix (Marriott 1974). This makes it difficult to extend the results of PCA to other times and places because principal component axis rotation is uniquely determined by the original set of observational data (the "training set"). Finally, it is not always correct to assume that the higher principle component axes, that explain only a small proportion of the total variance, contain little relevant information (Eastman and Fulk 1993, Anyamba and Eastman 1996).

2.5.2 Temporal Fourier analysis

An entirely different approach to the problem of data reduction is time-series analysis (Chatfield 1980). The time-series {xt} may be described by a Fourier series representation where;

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

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

$$a_0 = \bar{x}$$

$$a_{N/2} = \Sigma(-1)^t x_t / N$$

$$\left.\begin{aligned} a_p &= 2[\Sigma x_t \cos(2\pi pt / N)] / N \\ b_p &= 2[\Sigma x_t \sin(2\pi pt / N)] / N \end{aligned}\right\} p = \mathrm{I}, \ldots\ldots N/2-1$$

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

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

where R_p, is the amplitude of the pth harmonic

$$= \sqrt{(a^2_p + b^2_p)}$$

and ϕ_p, is the phase of the pth harmonic (Chatfield 1980)

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

The effect of Fourier analysis is to partition the variability of the time-series into (orthogonal and thus uncorrelated) components at frequencies of $2\pi/N$, $4\pi/N$, $6\pi/N$ ... , π, or periods equal to 1, $1/2$, $1/3$ ... $2/N$ 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 the "Nyquist frequency" (higher frequencies, *i.e.* shorter period cycles, cannot be distinguished by monthly data). A full Fourier analysis exactly describes the original dataset (since the Fourier series contains N variables to describe N observations) but, as with principal components, not all harmonics may be contributing equally to this description. The following relationship, known as Parseval's theorem, applies to the Fourier representation of $\{xt\}$ where;

$$\Sigma(x_t - \bar{x})^2 / N = \sum_{p=1}^{N/2-1} R^2{}_p / 2 + a^2{}_{N/2}$$

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 $R^2_p/2$ is the contribution of the pth harmonic.

The analysis is demonstrated graphically in figure 2.5 for an NDVI time-series from an example pixel located near Ouagadougou in Burkina Faso. The lower graph shows the three derived Fourier "components" *i.e.* the annual, bi-annual and tri-annual cycles in black, green and blue respectively. The top graph shows the sum of these first

three components in red and illustrates its fit to the NDVI signal in black. The outputs of temporal Fourier analysis are expressed as an overall average, and an amplitude (extent) and phase (timing) of the signal maximum of each component for the duration of the analysis. In effect Fourier analysis makes it possible to reduce the three year (or longer) data stream to just 7 variables (the mean of the whole series and the amplitude and phases of the first three cycles) without a significant loss of information.

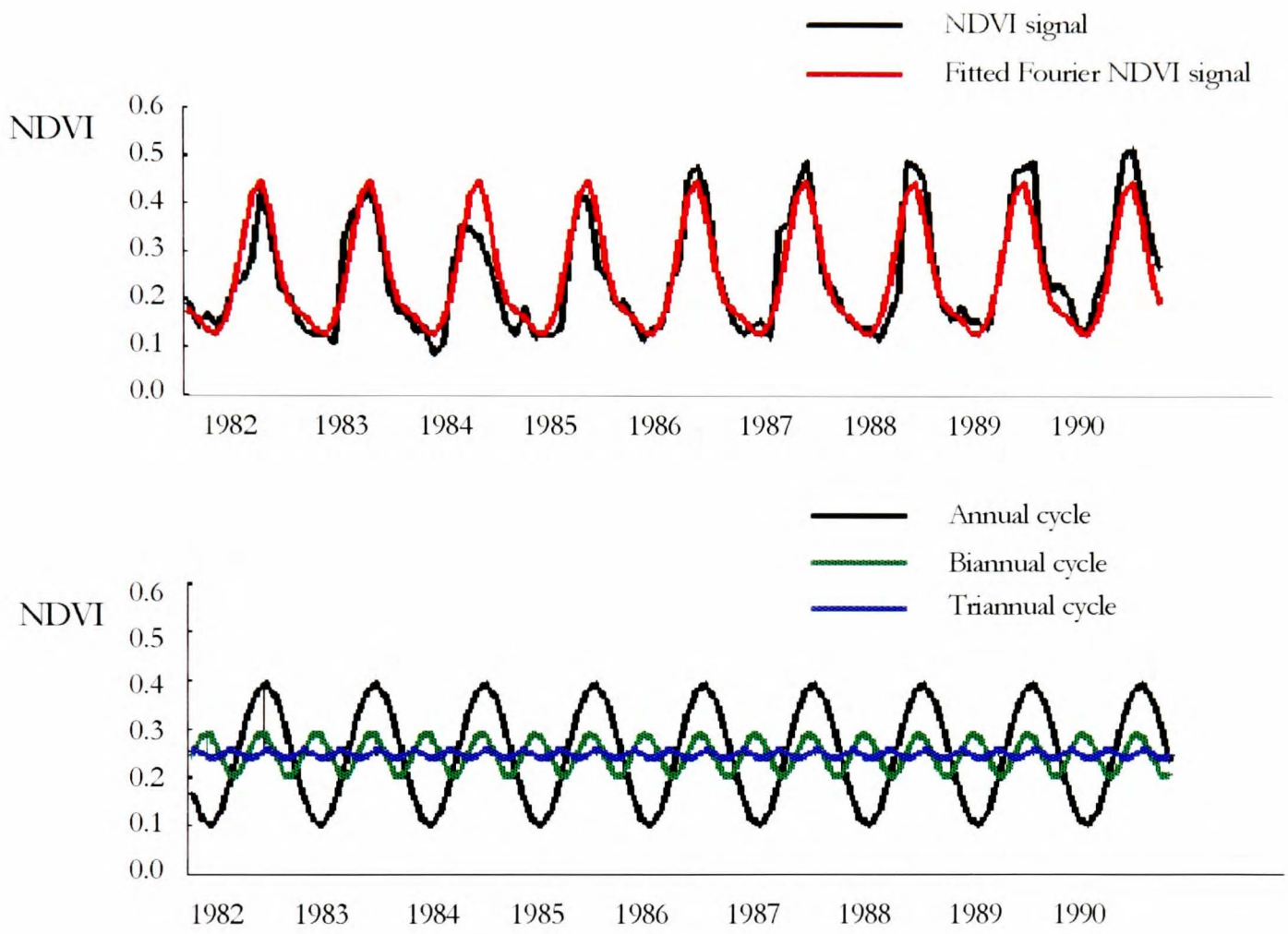


Figure 2.5. Fourier analysis of the NDVI time-series for a pixel located over Ouagadougou in Burkina Faso (1982 - 1990). The lower graph shows the annual, bi-annual and tri-annual cycle in black, green and blue respectively. The upper graph shows the combined signal in red and its fit to the true NDVI signal in black.

For the present study, each of the fifteen bands of image data were subjected to temporal Fourier processing and the means, amplitudes and phases of the annual, bi-annual and tri-annual cycles calculated. For the PAL data this was done from 1988 - 1990 and for the CCD data form 1989 - 1991. In temperate regions we are immediately familiar with annual cycling of biological processes, but for many species from tropical latitudes (especially near the equator where two rainy seasons can occur) bi-annual and tri-annual cycles in flowering, leaf-flushing and leaf-fall are common (Ewusie 1968, 1969 and 1992, Goward and Prince 1995). These variables do therefore have potential phenological significance and were maintained in the analysis.

The temporal Fourier processing was also exploited to correct obvious irregularities in the time-series, such as mask values preserved in the monthly MVC, because in certain locations and seasons the duration of cloud coverage at the time of the satellite overpass can be longer than the compositing period. These problem pixels were corrected by interpolating between the pixel values for the previous and subsequent months and the Fourier analysis reapplied to the corrected data. The range of threshold deviations from the fitted Fourier signal used in the satellite sensor derived bands are detailed in table 2.4.

The Fourier variables were stored as new image layers for analysis, at the same spatial resolution as the original imagery. The combined (*i.e.* annual + bi-annual + tri-annual cycle) Fourier description of the original signal was also calculated (a summation that essentially smoothes the original dataset) and its minimum, maximum and range were also recorded.

In conclusion, the combination of the orthogonality of the harmonics (*i.e.* removal of temporal autocorrelation) in the Fourier series representation of multitemporal satellite sensor data and the ease of the interpretation of these harmonics (Rogers and Williams 1994) made this the technique of choice for the present analysis. Furthermore, this technique is becoming more widely accepted as a dimension reduction technique in remote sensing, as many recent studies have related Fourier processed AVHRR data to regional and continental scale biological processes (Menenti *et al.* 1993, Andres *et al.* 1994, Olsson and Eklundh 1994, Fuller and Prince 1996, Verhoef *et al.* 1996).

TABLE 2.4.

The fifteen bands of satellite sensor data and derived indices.

Band	*Data layer*	*Units*	*Threshold 1*	*Threshold 2*
	AVHRR raw channel data			
1	Channel 1	% reflectance	100	10
2	Channel 2	% reflectance	100	10
3	Channel 3	Kelvin	333	10
4	Channel 4	Kelvin	333	10
5	Channel 5	Kelvin	333	10
	Land surface temperature indices			
6	Price (1984)	Kelvin	360	10
7	Becker and Li (1990b)	Kelvin	360	10
8	Prata and Platt (1991)	Kelvin	360	10
	Vegetation indices			
9	RVI	ratio	2	0.2
10	NDVI	ratio	2	0.2
11	SAVI (L = 0.50)	ratio	2	0.2
12	GEMI	ratio	2	0.2
	Atmospheric moisture indices			
13	Dalu (1986)	kgm^{-2}	30	15
14	Eck and Holben (1994)	kgm^{-2}	30	15
	Rainfall indices			
15	CCD	hours	-	-

Threshold 1: the absolute maximum allowed for an image value.
Threshold 2: the maximum departure allowed from the fitted Fourier signal.

Chapter 3

REMOTELY SENSED SURROGATES FOR METEOROLOGICAL VARIABLES

" ... The great tragedy of science the slaying of a beautiful hypothesis with an ugly fact ... "

T. H. Huxley.

3.1 Introduction

This chapter presents an investigation of the utility of remotely sensed data from meteorological satellite sensors for predicting ground meteorological variables at the broad spatial scale. The following introduction develops substantially on rudiments of tsetse ecology provided in section 1.5.5.1 and highlights where climate is important to tsetse fly population dynamics.

3.1.1 Tsetse ecology and climate

Tsetse flies vary between 5 - 12 mm in length and are generally brown in colour. At rest they are characterised by a proboscis which extends in front of the fly and by wings that are folded flat over the dorsal surface of the abdomen. There are two unique and commonly used diagnostic features which separate the *Glossina* from other diptera. They have a hatchet-shaped discal cell on each wing defined by the configuration of the wing veins, and the presence of secondary branches in the hairs on the arista of the antennae. Individual species are identified by the structure of the male genitalia, the colour and patterns of the dorsal surface of the abdomen and the colour of the leg tarsi.

All tsetse flies belong to the genus *Glossina* which is monotypic to the family Glossinidae. Twenty-three tsetse species are known, as well as several well defined and geographically restricted sub-species (Potts 1970). The genus is divided into three ecologically distinct sub-genera (Ford 1970). The *morsitans* group (subgenus *Glossina S. str.*) occurs mainly in semi-arid savanna and savanna-woodland landscapes. Blood-meal analysis has shown that they feed largely on bovids and suids (Moloo 1993) and are

therefore significant vectors of animal trypanosomiasis. Second, the *palpalis* group (subgenus *Nemorhina*) which most commonly inhabit riparian and lakeside vegetation, where the humidity is high, but is also capable of invading peri-domestic habitats in West Africa and *Lantana* thickets in Uganda. The frequency of contact with humans can be high in such areas, especially so when water is limiting in the dry season and human and fly populations concentrate around water resources (Nash 1969, Frezil *et al.* 1990, D'Amico *et al.* 1992). The *palpalis* group are therefore important vectors for human trypanosomiasis. Third, the ancestral *fusca* group (subgenus *Austenina*) that lives mostly in high forest habitats or in more humid relict forested on the periphery of rain forest belts and are of little epidemiological significance. All species can transmit the pathogenic trypanosomes.

Tsetse flies can be caught in a variety of trapping devices which have largely replaced the use of fly-rounds (in which humans or cattle are used as bait for tsetse sampled for standard time periods along known routes) and are reviewed by Green (1994). These traps incorporate information from a large body of research on the olfactory and visual bases of host seeking behaviour in tsetse (Colvin and Gibson 1992), which has culminated in the use of odour-baited box traps for the *morsitans* group and biconical or pyrimidal traps for the *palpalis*.

Male and female tsetse flies are haematophagous and vector to trypanosomes which undergo temperature dependent cycles of development within the flies (Hoare 1972). As has been mentioned, the analysis of fly blood-meals has demonstrated that the various species of *Glossina* have different food preferences (Moloo 1993). The most abundant host in a particular locality is often the one from which most of the blood meals are derived. In general terms the *morsitans* and *fusca* groups take their meals largely from bovids and suids whereas the *palpalis* group is capable of feeding on most available hosts including reptiles. Tsetse take a blood meal approximately every 3 days although the timing is strongly temperature dependent. Having obtained a meal, flies rest in well defined sites of more equitable temperature and humidity (see section 4.1.1).

Mating in tsetse often occurs near the host (Buxton 1955). Females mate only once, or on very few occasions, and store sperm transferred by the males in spermathecae. The female tsetse is larviviparous so that after mating the fertilised egg is nourished, for a temperature dependent 7 14 day period, by milk glands within the uterus. The fully developed larva can be of greater mass than the adult female.

Larviposition is at favoured humid sites (see section 4.1.1). The deposited larva burrows rapidly into the moist soil to a depth of approximately 1 - 2 cm whereupon it hardens to form a puparium, although occasionally it may pupate on the soil surface. The puparia are black, barrel shaped and have two characteristic polypnuestic lobes for respiration. Puparial life can last for 25 60 days, again depending on temperature. Young adults before the first blood meal are referred to as "teneral" and are particularly susceptible to infection with *Trypanosoma brucei* and *T. congolense* type trypanosomes (Maudlin 1991).

There are many stages in the tsetse life-cycle where climatic factors can be of influence. Roubaud (1909) was one of the first to recognise the importance of climate, and its short-term manifestation, weather, to the ecology of tsètse. He contrasted the thermal autecology of *G. morsitans* with that of *G. palpalis* and conducted some preliminary laboratory experiments on the temperature and humidity tolerances of the two species. A large body of work followed this example and related local temperature, humidity and rainfall measurements to the life-history variables of tsetse fly species. This extensive literature has been reviewed by Gaschen (1945) and Buxton (1955).

Notable among these studies is the work of Nash (1933) who showed that the numbers of *G. morsitans* in Tanzania were influenced by the incidence of rainfall, although primarily by the combination of temperature and humidity. Buxton and Lewis (1934) added to this understanding when they found that temperatures and humidities experienced by *G. tachinoides* in West Africa were far from the optimum conditions for survival and reproduction established in the laboratory. A series of experiments pursued this line of study by investigating the effect of combinations of temperature and humidity, on the water balance of different stages in the life-cycle of eight tsetse species (Bursell 1957, 1958, 1959). A close correlation was found between the resistance to desiccation of the pupae and the habitat in which the particular species occurred. A similar trend was not observed in adults and it was suggested that it was the water balance of the puparial stages of the life-cycle that was limiting tsetse species' distributions.

These investigations showed that abiotic climatic variables were of primary importance in determining the distribution of tsetse species. Rogers (1979) and Rogers and Randolph (1986) realised that with a detailed knowledge of the climatic tolerance of a given tsetse species from one location, potentially favourable habitats could be identified over wide areas using extensive meteorological data. This was done by relating

monthly fly mortality rates to climograms combining temperature and saturation deficit for *G. morsitans*, *G. palpalis*, *G. fuscipes* and *G. tachinoides* from three sites in West Africa. Bioclimatic tolerances were predicted for *G. morsitans* using the Moran curve technique outlined by Rogers (1979) and were interpolated across Africa using meteorological data. Ranges were defined on the basis of these interpolated tolerances and found to coincide well with historical distribution maps. Despite the complication of density-dependent factors affecting abundance, the analysis of the bioclimatic optima for *G. morsitans* revealed strong correlations with areas of known disease foci. Although these results were promising, it was noted that the potential for meteorological data in forecasting the changing risks of trypanosomiasis transmission (largely a combined function of host and tsetse abundance (Rogers 1985)), was severely restricted by the infrastructure for meteorological data collection in Africa. A density of approximately one meteorological station per 28,000 km^2 imposed considerable limits on the spatial resolution and thus utility of any derived distribution and abundance information.

The geographically extensive and repeated collection of data by meteorological satellite sensors was recognized as a means of providing surrogate information for these meteorological data. Rogers and Randolph (1991) showed that the correlation between mean monthly mortality rates for *G. morsitans submorsitans* and monthly MVC GAC NDVI data was at least as good as that between the same mortality rates and monthly saturation deficit calculated from meteorological records near the Yankari game reserve of Nigeria. The authors did not consider it surprising to find such a relationship, since the vegetation was influenced and in turn affected both by local macro- and by micro-climatic conditions. The major advantage of this technique was that such remotely sensed surrogate variables could be recorded on a daily basis for every 8 x 8 km pixel on the African continent by the NOAA - AVHRR.

3.1.2 Objectives and epidemiological perspectives

The work of Rogers and Randolph (1991) was the first step in attempting to provide remotely sensed surrogates of meteorological variables for the purposes of mapping tsetse distribution and abundance. The accuracy to which this can be achieved using the full range of information available from the NOAA - AVHRR and CCD data from the Meteosat - HRR is investigated in this chapter. This work has only become possible since the relevant satellite sensor data became freely available to the research community in 1994 (see section 2.3.3).

Temperature, humidity and rainfall have also been shown to be critical to the population dynamics of other arthropod vectors of disease, including mosquitoes (Platt *et al.* 1958, Service 1978) and ticks (Randolph 1993, Daniel and Dusbábek 1994). Furthermore, these climatic factors have been related directly to the diseases transmitted by such vectors (Epstein and Chikwenhere 1994) including trypanosome infection rates (Ford and Leggate 1961) and the prevalence of human sleeping sickness (Rogers and Williams 1993). They are also known to have a major affect on malaria transmission rates (Fialho and Schall 1995, Freeman and Bradley 1996, Mouchet and Carnevale 1991) and thus malaria epidemics (Macdonald 1953). The determination of accurate and geographically extensive surrogates for meteorological variables by remote means could therefore find very wide application in epidemiology, far beyond the present objective of tsetse population surveillance.

3.2 Methods

3.2.1 Ground meteorological data

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 NOAA National Climatic Data Centre (NCDC) (NOAA 1990). Data are available as monthly summaries for 250 named meteorological stations throughout Africa. The distribution of the 207 stations used in this study is shown in figure 3.1. 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 sensor data could have been contaminated by the signal from these water bodies.

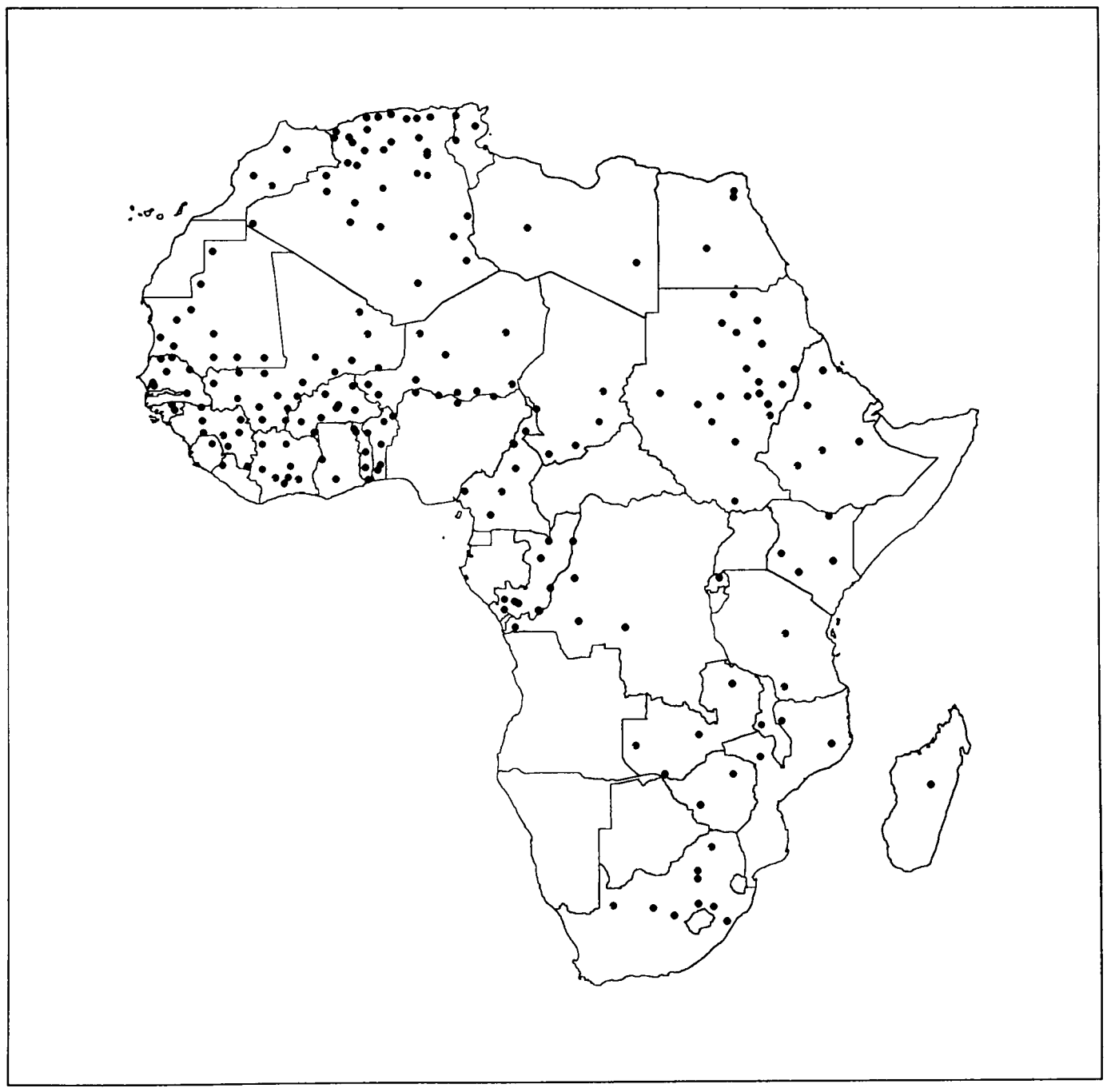


Figure 3.1. The distribution of the WMO meteorological stations used in this study.

There are strict standards governing the collection of meteorological data from WMO stations, so that measurements collected across the globe can be compared (WMO 1971). The general principle behind these standards is that the measurements should be representative of air circulating freely in the locality. To facilitate this it is recommended that the meteorological station should be sited on level ground, covering an area of at least 10 x 7 m and be situated at a "significant distance" from any surrounding buildings or trees.

Temperature (strictly the thermodynamic air temperature (Becker and Li 1990a)) is recorded from standard maximum and minimum thermometers (NOAA 1972) fitted inside a white-painted, double-louvered wooden box, often referred to as a Stevenson screen. The louvering prevents direct insolation whilst enabling free vertical ventilation by convection within the box. The base must be 1.25 - 2 m above a well-mown lawn or the natural vegetation cover of the region. In the Northern hemisphere the door must face north so that the thermometers are not exposed to the Sun while measurements are taken. Precipitation is measured using copper gauges (calibrated in mm) and situated at a distance beyond four times the height of the nearest object. Snow and sleet deposition are not differentiated from rainfall. Vapour pressure is measured with a standard sling or whirled psychrometer (NOAA 1972). This consists of two identical mounted thermometers; one with an exposed (or dry) bulb and one with its bulb covered in water saturated muslin (the wet bulb). These are swung through the air for a set period of time and the difference in temperature between the thermometers used to calculate the vapour pressure from empirical tables (WMO 1971). 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 metres above mean sea level. The land surface data used in this investigation were the mean monthly temperature (°C), the mean monthly vapour pressure (mb) and total monthly precipitation (mm). The SI unit for pressure is the pascal, but the millibar (1 newton per square metre; 1 millibar = 100 pascals) is commonly used in meteorology. This is because a major user of meteorological information is the aviation industry for which the unit is standard. The departures of each of the monthly readings from the long-term average (established from the records

collected between 1951 and 1980) were listed with each of the measurements. Late reports and corrections are also published on receipt.

3.2.2 Data analysis

The data for 1988 - 1990 were entered into a Microsoft Excel 7.0 database and were augmented with any corrections or late records published between 1991 and 1995. Humidity variables were then calculated from the mean temperature, T (K), and the mean vapour pressure, Vp (mb), using formulas provided by Unwin (1980). First, the saturation vapour pressure, Svp (mb), or the amount of water vapour that would be contained in the same volume of saturated air at that temperature was calculated using;

$$\log_{10} Svp = 9.24349 - \frac{2305}{T} - \frac{500}{T^2} - \frac{100000}{T^3}.$$

Second, the relative humidity, Rh (%), or the amount of water vapour in the air divided by the saturation vapour pressure was calculated;

$$Rh = \frac{Vp \times 100}{Svp}.$$

Third, the saturation deficit, Sd (mb), sometimes referred to as the vapour pressure deficit, and defined as the difference between the recorded vapour pressure and the saturation vapour pressure at the same temperature, was given by;

$$Sd = Svp - Vp.$$

This final variable is a measure of the "drying power" of the air or the lack of moisture equilibrium between an object and its surrounding atmosphere. The higher the saturation deficit the more rapidly desiccation will occur.

The satellite sensor data used in this chapter, with details of their conversion to proxy meteorological data, have been described in section 2.4. Extraction programs (written by DJR) used the latitude and longitude of each meteorological station to extract the corresponding pixel value from the relevant monthly MVC satellite sensor band for each month during 1988 - 1990. These values were then transferred to the database and statistical associations explored using the Statistical Package for the Social Sciences (SPSS) version 7.0.

Comparing the slopes of the twelve regressions for each of the monthly comparisons in 1988, 1989 and 1990, using Bartlett's test (Snedecor and Cochran 1967), showed that the monthly data could not be pooled since there was significant differences in the residual variances between samples. As an alternative, 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. The problem of missing reports (only 28 of the 250 stations published a complete three year series of data) and the masking of cloud-affected data in the MVC imagery led to the number of observations and the subset of sites investigated varying between months. For this reason the population adjusted coefficient of determination or the adjusted r^2 (Williams 1993) are shown throughout.

3.3 Results

3.3.1 Prediction of land surface temperature

The comparison of the NOAA - NCDC mean monthly air temperature with the AVHRR thermal channel and split-window corrected brightness temperatures are presented for 1988, 1989 and 1990 in tables 3.1, 3.2, and 3.3 respectively. A positive linear relationship is observed in every month ($P << 0.0001$, mean $n = 120$ with range of 105 to 141 observations). Non-linear alternatives to a linear regression fit did not increase coefficients of determination for the monthly comparisons.

The mean annual r^2 for 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 (1990b) and 0.38 for Prata and Platt (1991) formulations. This trend in the performance of the split-window equations was followed without exception in the monthly comparisons (see tables 3.1, 3.2 and 3.3). 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 figure 3.2.

The residuals 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 (see figure 3.3). Distinct geographical patterns were revealed in these residual values that were apparently related to altitude. Positive deviations from the regression line occurred largely in those stations above 500 m and negative deviations occurred in those below. This relationship was found to be highly significant when the standardised residuals were plotted against elevation; an example is given in figure 3.4 for December 1990. When elevation data from the DEM (see section 2.2) 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 figure 3.5, which also reveals that there were fairly well defined seasonal trends; with peaks and troughs around months 8 and 4 respectively. The scatter of points in the monthly comparisons was not similarly reduced by including the NDVI in the regression equation which might have been expected as vegetation coverage has a major affect on local evapotranspiration and temperature regimes (see section 3.4.1).

TABLE 3.1.

The table shows 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 1988.

	Channel 3	*Channel 4*	*Channel 5*	*Price (1984)*	*Becker and Li (1990b)*	*Prata and Platt (1991)*	*Price (1984) and elevation*
Jan	0.29	0.45	0.35	0.70	0.54	0.47	0.74
Feb	0.35	0.47	0.39	0.70	0.56	0.50	0.74
Mar	0.22	0.29	0.24	0.56	0.45	0.40	0.60
Apr	0.36	0.47	0.39	0.50	0.42	0.43	0.55
May	0.52	0.52	0.48	0.64	0.55	0.53	0.70
Jun	0.60	0.48	0.45	0.60	0.50	0.50	0.72
Jul	0.66	0.54	0.49	0.69	0.54	0.50	0.80
Aug	0.60	0.47	0.42	0.61	0.43	0.38	0.76
Sep	0.27	0.13	0.11	0.25	0.15	0.10	0.64
Oct	0.25	0.30	0.19	0.50	0.35	0.30	0.70
Nov	0.17	0.32	0.28	0.43	0.28	0.22	0.68
Dec	0.28	0.47	0.38	0.65	0.5	0.43	0.76
Mean	**0.38**	**0.41**	**0.35**	**0.57**	**0.44**	**0.40**	**0.70**
SE	**0.05**	**0.03**	**0.03**	**0.04**	**0.04**	**0.04**	**0.02**

TABLE 3.2.

The table shows 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 1989.

	Channel 3	*Channel 4*	*Channel 5*	*Price (1984)*	*Becker and Li (1990b)*	*Prata and Platt (1991)*	*Price (1984) and elevation*
Jan	0.32	0.52	0.43	0.68	0.55	0.48	0.72
Feb	0.14	0.34	0.30	0.50	0.48	0.38	0.60
Mar	0.12	0.35	0.29	0.59	0.44	0.37	0.66
Apr	0.24	0.50	0.44	0.67	0.61	0.55	0.75
May	0.48	0.53	0.47	0.70	0.66	0.61	0.77
Jun	0.62	0.55	0.49	0.68	0.64	0.61	0.79
Jul	0.48	0.33	0.27	0.53	0.45	0.41	0.79
Aug	0.29	0.26	0.20	0.46	0.36	0.33	0.74
Sep	0.06†	0.15	0.14	0.19	0.14	0.14	0.58
Oct	0.09†	0.19	0.17	0.28	0.17	0.15	0.67
Nov	0.20	0.34	0.26	0.54	0.39	0.31	0.76
Dec	0.11	0.26	0.19	0.47	0.27	0.19	0.66
Mean	**0.26**	**0.36**	**0.30**	**0.52**	**0.43**	**0.38**	**0.71**
SE	**0.05**	**0.04**	**0.03**	**0.04**	**0.05**	**0.05**	**0.02**

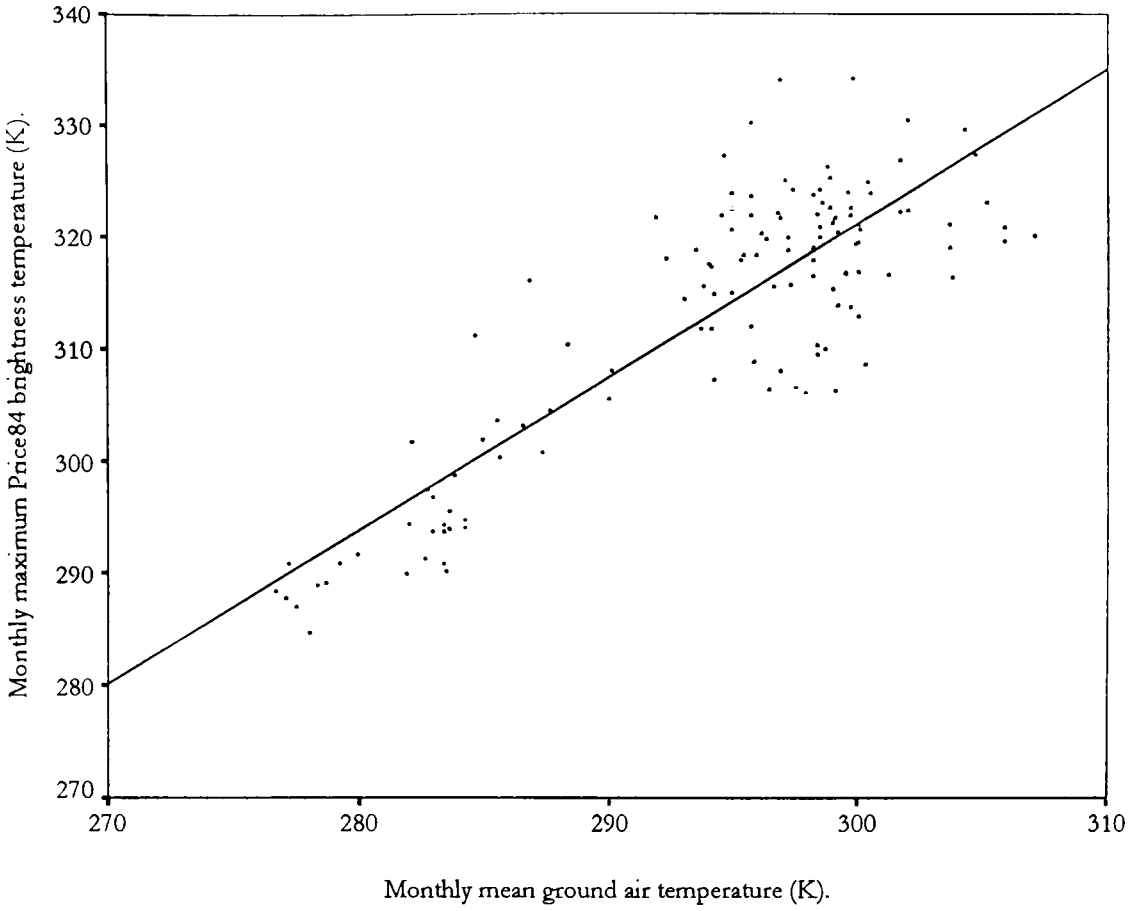
† Not significant at the 0.01 level.

TABLE 3.3.

The table shows 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.

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



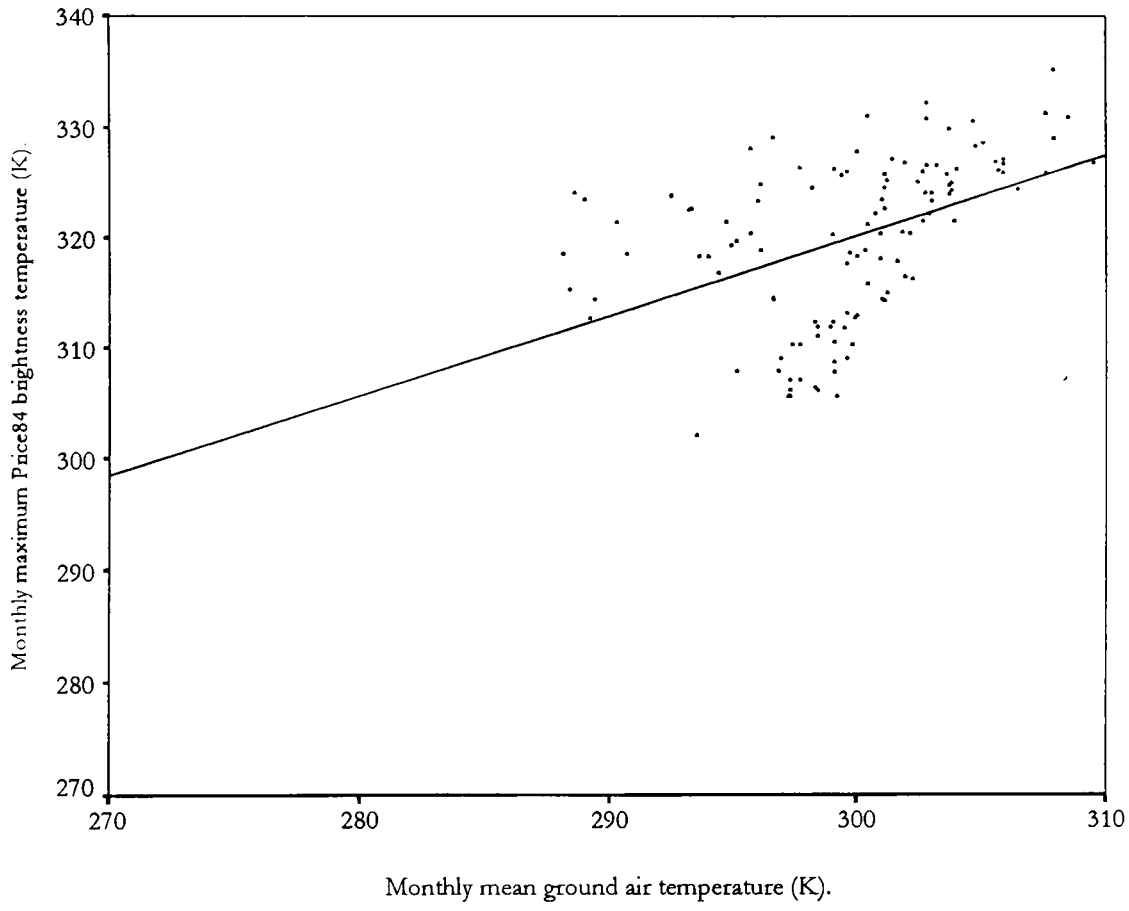


Figure 3.2. The comparison of mean monthly NOAA - NCDC air temperature with the maximum monthly Price (1984) split-window corrected brightness temperature. Examples are given in the top graph (December 1990, n = 124, $P << 0.0001$, $r^2 = 0.74$) for the best fit to the ground data and in the lower graph the worst (September 1989, n = 117, $P << 0.0001$, $r^2 = 0.19$).

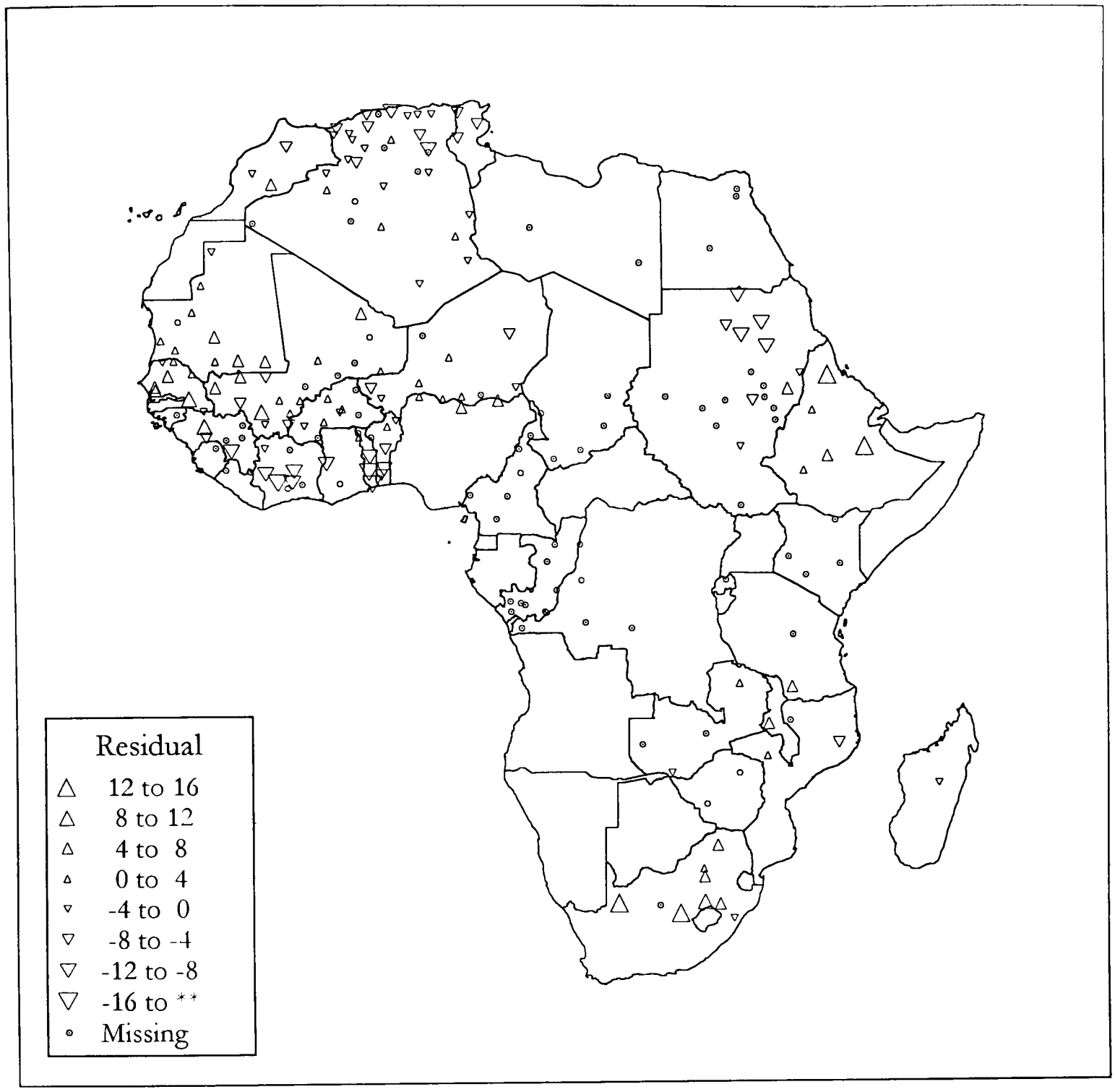


Figure 3.3. The distribution of the residual values (K) in the regression of Price (1984) corrected split-window corrected brightness temperature and the mean NOAA - NCDC air temperature for December 1990. Stations for which data were missing are also indicated.

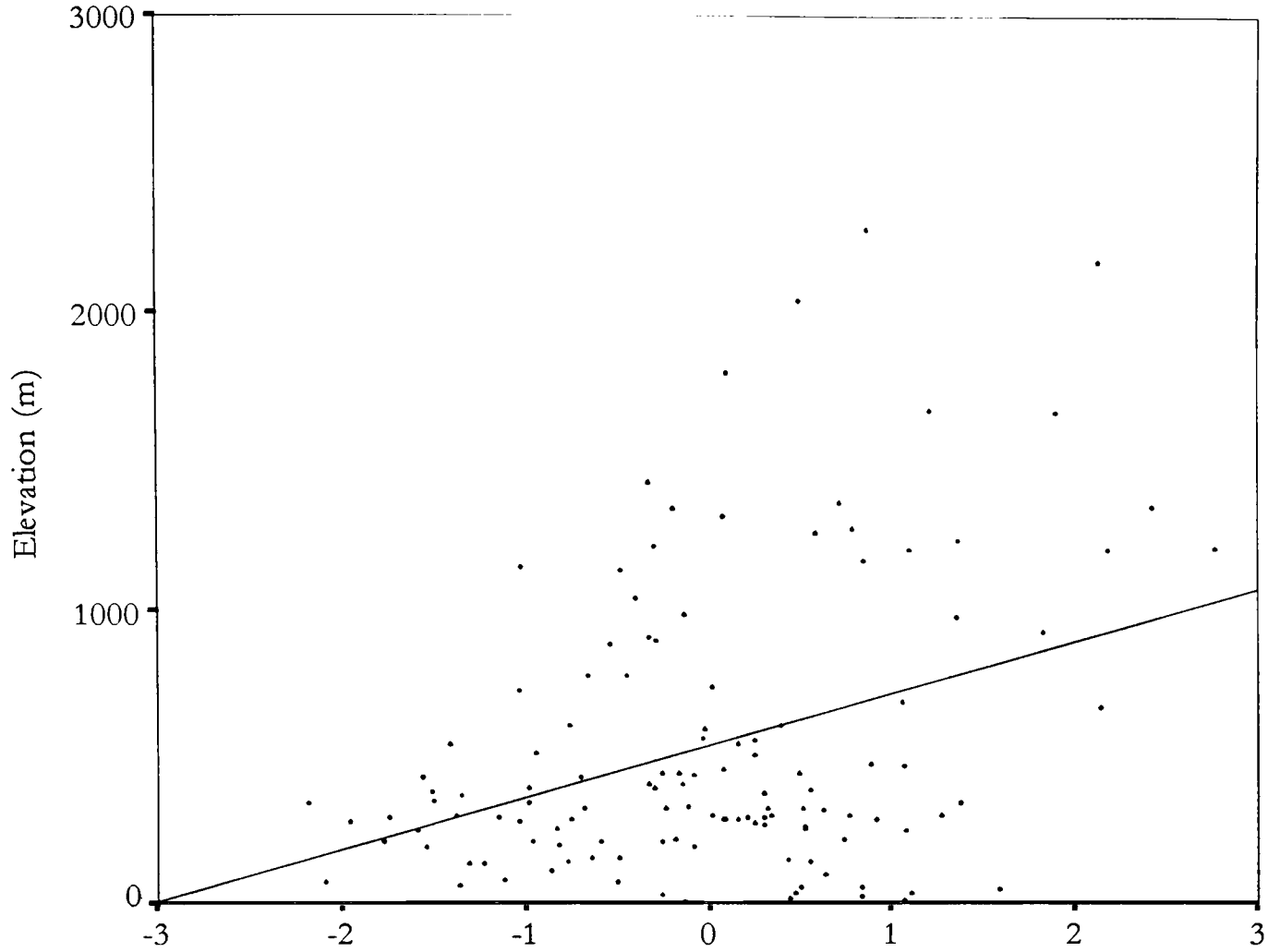


Figure 3.4. The comparison of the standardised residual values (K) and elevation resulting from the regression of Price (1984) corrected split-window brightness temperature and the mean NOAA - NCDC air temperature (December 1990, n = 124, $P << 0.0001$, $r^2 = 0.13$).

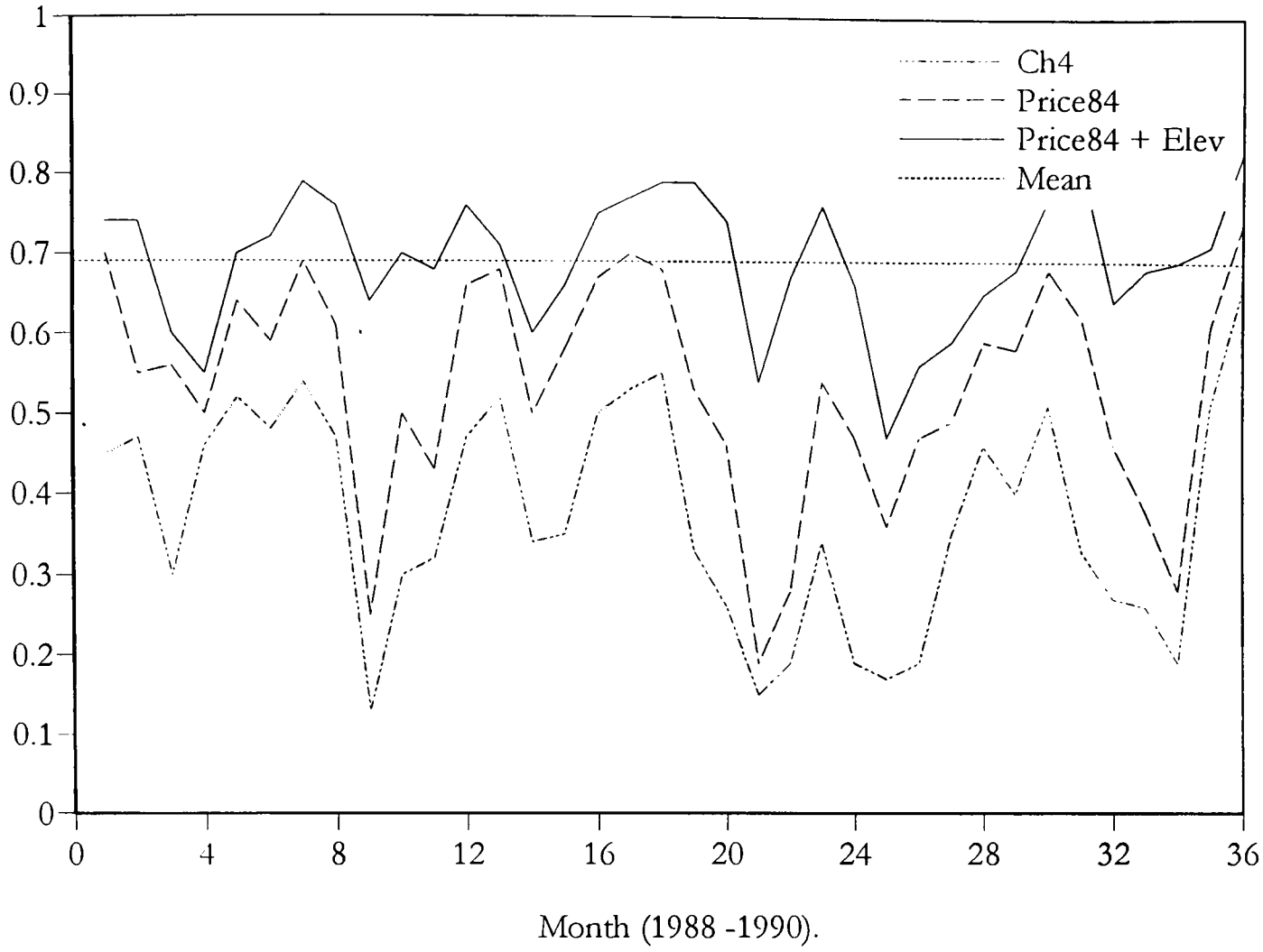


Figure 3.5. The adjusted r^2 of the monthly comparison of NOAA - NCDC air temperature and Price (1984) split-window corrected brightness temperatures for the period 1988 1990. The graph shows the improvement of fit of the Price (1984) equation over the AVHRR channel 4 data and the further improvement 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.

3.3.2 Prediction of atmospheric moisture

The comparison of mean monthly NOAA - NCDC saturation deficit with the maximum monthly vapour pressure deficit derived from the satellite sensor data are presented by year for 1988 - 1990 in tables 3.4, 3.5, and 3.6. The VPD is calculated using the Dalu (1986) (VPD_1) and the Eck and Holben (1994) (VPD_2) procedures for deriving total precipitable water in the atmospheric column. A highly significant positive linear relationship was observed in each month ($P << 0.0001$, 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 figure 3.6.

The mean r^2 for the three year period using VPD_1 and VPD_2 were 0.63 and 0.62 respectively, with the VPD_1 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 figure 3.7 revealed no temporal trends in the VPD data and 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 3.4.

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 1988.

	Ch4 Ch5	*VPD_1*	*VPD_2*	*VPD_2 and elevation*	*VPD_2 and NDVI*
Jan	0.01†	0.57	0.56	0.64	0.68
Feb	0.01†	0.53	0.57	0.64	0.63
Mar	0.15	0.69	0.77	0.78	0.79
Apr	0.01†	0.74	0.69	0.69	0.73
May	0.21	0.76	0.71	0.72	0.74
Jun	0.15	0.59	0.63	0.63	0.64
Jul	0.13	0.72	0.71	0.71	0.73
Aug	0.29	0.78	0.75	0.76	0.76
Sep	0.10	0.67	0.70	0.70	0.72
Oct	0.02†	0.52	0.46	0.45	0.50
Nov	0.10	0.44	0.46	0.51	0.50
Dec	0.00†	0.41	0.47	0.59	0.50
Mean	**0.10**	**0.62**	**0.62**	**0.65**	**0.66**
SE	**0.03**	**0.04**	**0.03**	**0.03**	**0.03**

† Not significant at the 0.01 level.

TABLE 3.5.

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 1989.

	Ch4 Ch5	*VPD_1*	*VPD_2*	*VPD_2 and elevation*	*VPD_2 and NDVI*
Jan	0.00†	0.50	0.50	0.63	0.57
Feb	0.01†	0.59	0.60	0.69	0.66
Mar	0.16	0.73	0.70	0.73	0.73
Apr	0.06†	0.78	0.81	0.82	0.82
May	0.01†	0.73	0.73	0.73	0.73
Jun	0.01†	0.62	0.63	0.62	0.66
Jul	0.10	0.78	0.70	0.70	0.70
Aug	0.15	0.72	0.64	0.64	0.67
Sep	0.34	0.58	0.59	0.60	0.60
Oct	0.15	0.64	0.66	0.67	0.67
Nov	0.00†	0.37	0.40	0.54	0.44
Dec	0.01†	0.47	0.55	0.64	0.62
Mean	**0.08**	**0.63**	**0.63**	**0.67**	**0.66**
SE	**0.03**	**0.04**	**0.03**	**0.02**	**0.03**

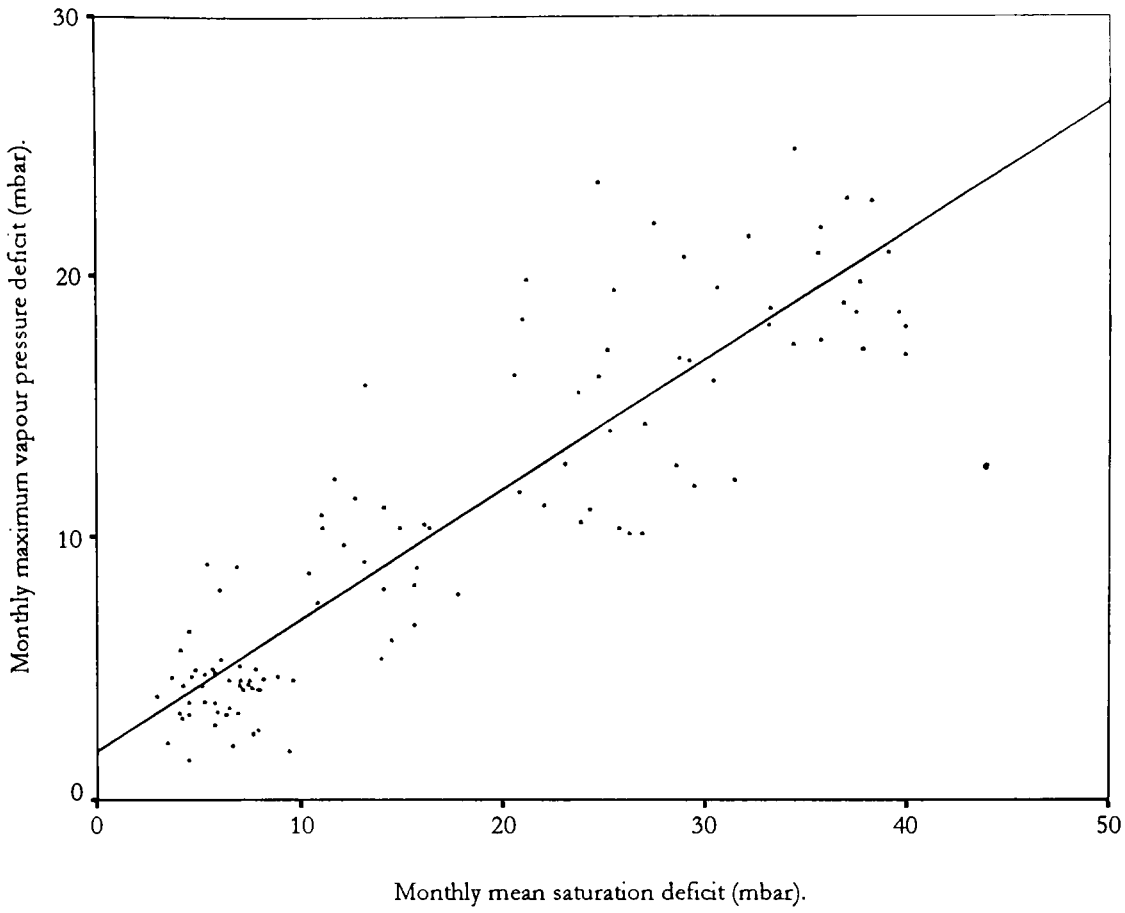
† Not significant at the 0.01 level.

TABLE 3.6.

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.

	Ch4 Ch5	*VPD_1*	*VPD_2*	*VPD_2 and elevation*	*VPD_2 and NDVI*
Jan	0.00†	0.40	0.41	0.54	0.52
Feb	0.03†	0.67	0.64	0.73	0.70
Mar	0.07†	0.64	0.62	0.69	0.66
Apr	0.10	0.69	0.64	0.65	0.69
May	0.05†	0.71	0.72	0.72	0.72
Jun	0.02†	0.67	0.63	0.63	0.64
Jul	0.01†	0.66	0.63	0.62	0.62
Aug	0.14	0.63	0.66	0.66	0.66
Sep	0.27	0.70	0.65	0.65	0.68
Oct	0.14	0.59	0.56	0.59	0.61
Nov	0.01†	0.63	0.64	0.70	0.65
Dec	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.



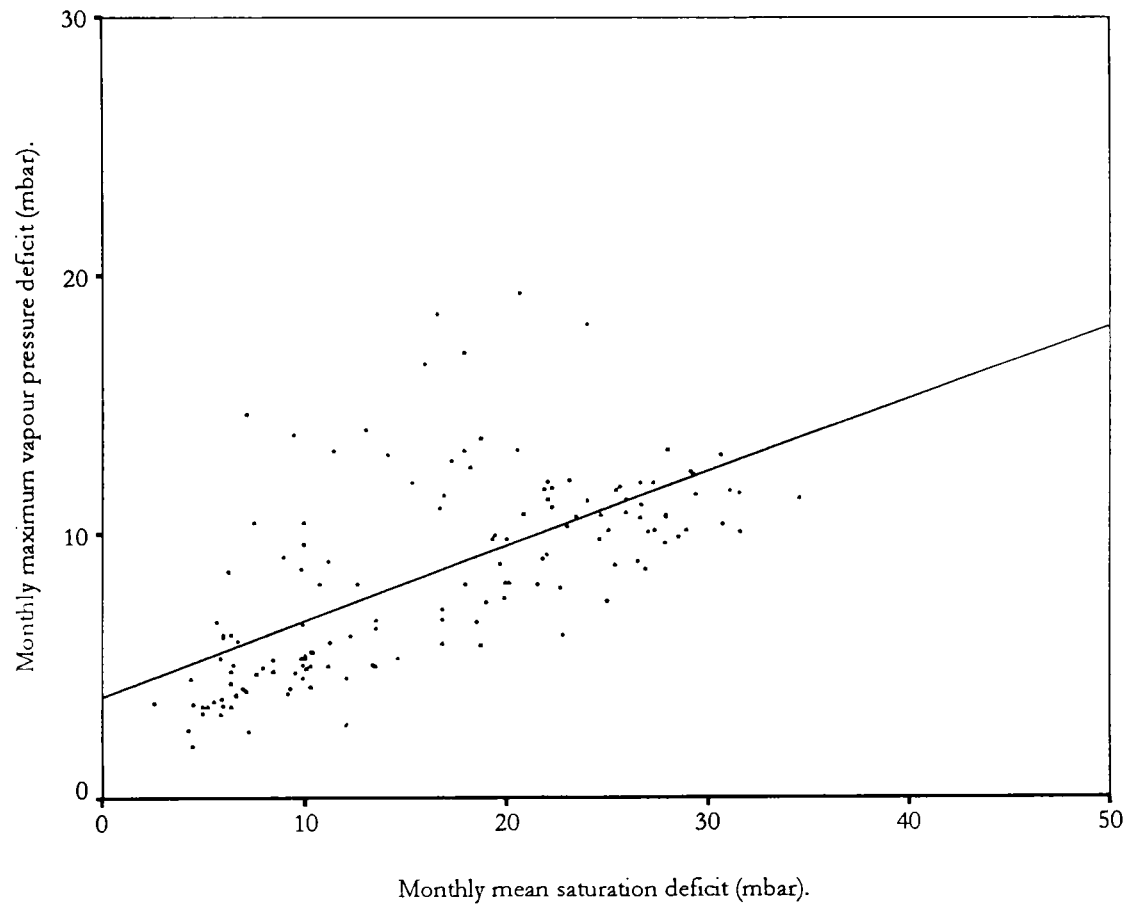


Figure 3.6. The comparison of mean monthly NOAA - NCDC saturation deficit with maximum monthly vapour pressure deficit. Examples are given in the top graph (Apr 1989, n = 107, $P << 0.0001$, $r^2 = 0.81$) for the best monthly fit and in the lower graph (Nov 1989, n = 140, $P << 0.0001$, $r^2 = 0.40$) for the worst fit to the ground data.

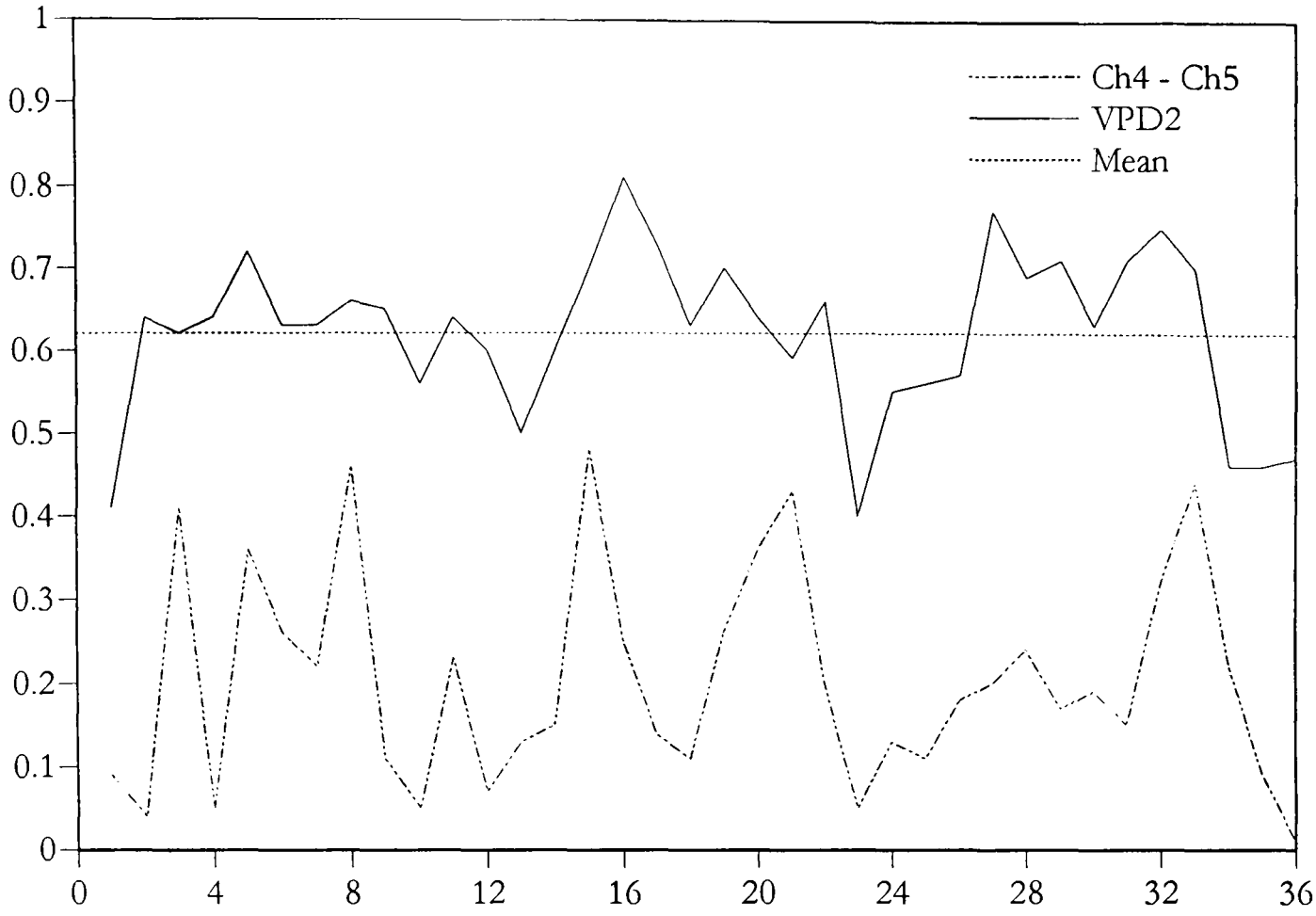


Figure 3.7. The adjusted r^2 of the comparison of mean monthly NOAA - NCDC saturation deficit and vapour pressure deficit for the period 1988 - 1990. The graph shows the improvement of fit of the VPD_2 formulation over the brightness temperature difference (channel 4 - channel 5). The mean level of variance explained by the VPD_2 variable is shown by the horizontal dotted line.

3.3.3 Prediction of rainfall

The comparison of total monthly precipitation and total monthly CCD for 1988, 1989 and 1990 is shown in table 3.7. A highly significant positive linear relationship occurs in every month ($P << 0.0001$, $n = 143$, 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 figure 3.8.

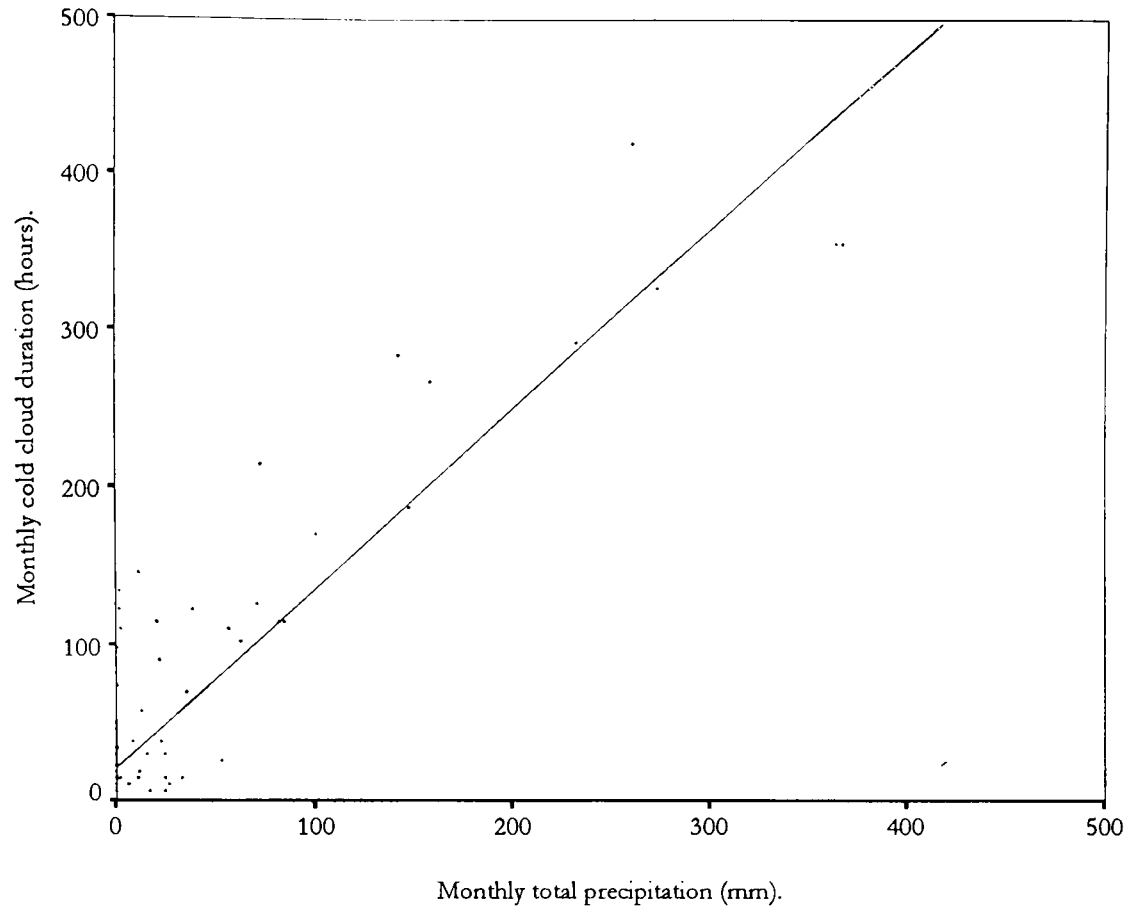
Restricting the analysis to those meteorological stations situated between 27° N and the equator (which corresponds to the TAMSAT calibration area) decreased the mean r^2 to 0.53 ($P << 0.0001$, $n = 84$ with a range 73 to 106 observations). The inclusion of elevation or NDVI in the regression equations did not increase the goodness of fit. The use of alternative curve fitting equations also had little effect, with the largest increase in the three year mean (to 0.67) arising from a cubic form expression. The time-series of these data plotted in figure 3.9 again revealed no consistent temporal relationships in the data, but illustrated the large variability about the mean of the r^2 values.

TABLE 3.7.

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

	1988			*1989*			*1990*		
	Linear	*Cubic*	*Tamsat*	*Linear*	*Cubic*	*Tamsat*	*Linear*	*Cubic*	*Tamsat*
Jan				0.81	0.87	0.75	0.84	0.90	0.00
Feb				0.75	0.78	0.50	0.72	0.74	0.56
Mar				0.76	0.77	0.80	0.41	0.48	0.36
Apr		-		0.50	0.65	0.49			
May				0.72	0.73	0.69	0.71	0.73	0.78
Jun				0.67	0.69	0.49	0.63	0.65	0.59
Jul				0.57	0.61	0.16	0.60	0.66	0.46
Aug				0.76	0.79	0.59	0.68	0.69	0.57
Sep	0.58	0.67	0.43	0.74	0.76	0.65	0.63	0.68	0.65
Oct	0.64	0.66	0.66	0.65	0.66	0.84	0.67	0.69	0.78
Nov	0.66	0.67	0.57	0.71	0.72	0.44	0.34	0.37	0.47
Dec	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.



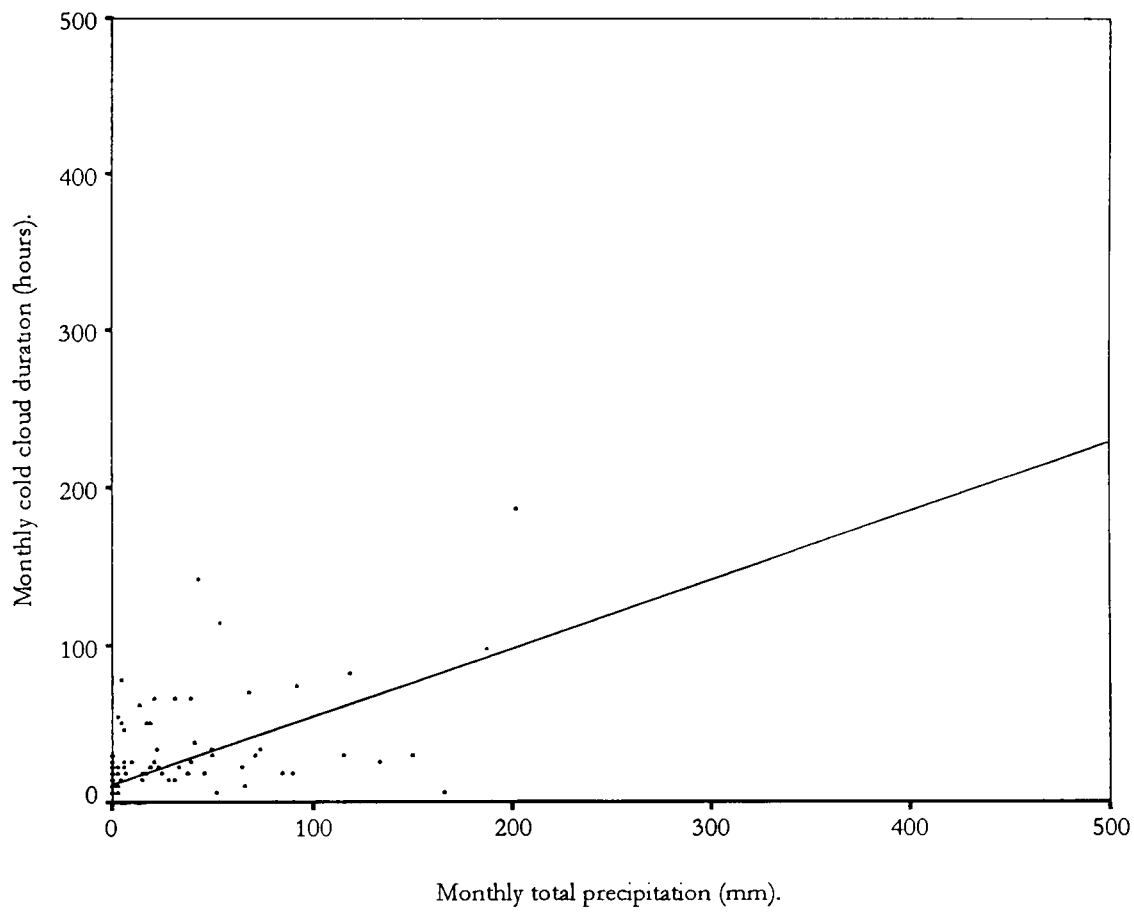


Figure 3.8. The comparison of total monthly precipitation with total monthly CCD. Examples are given in the top graph (Jan 1990, n = 104, P << 0.0001, r2 = 0.84) for the best monthly fit and in the lower graph (Nov 1990, n = 125, P << 0.0001, r2 = 0.34) for the worst fit to the ground data.

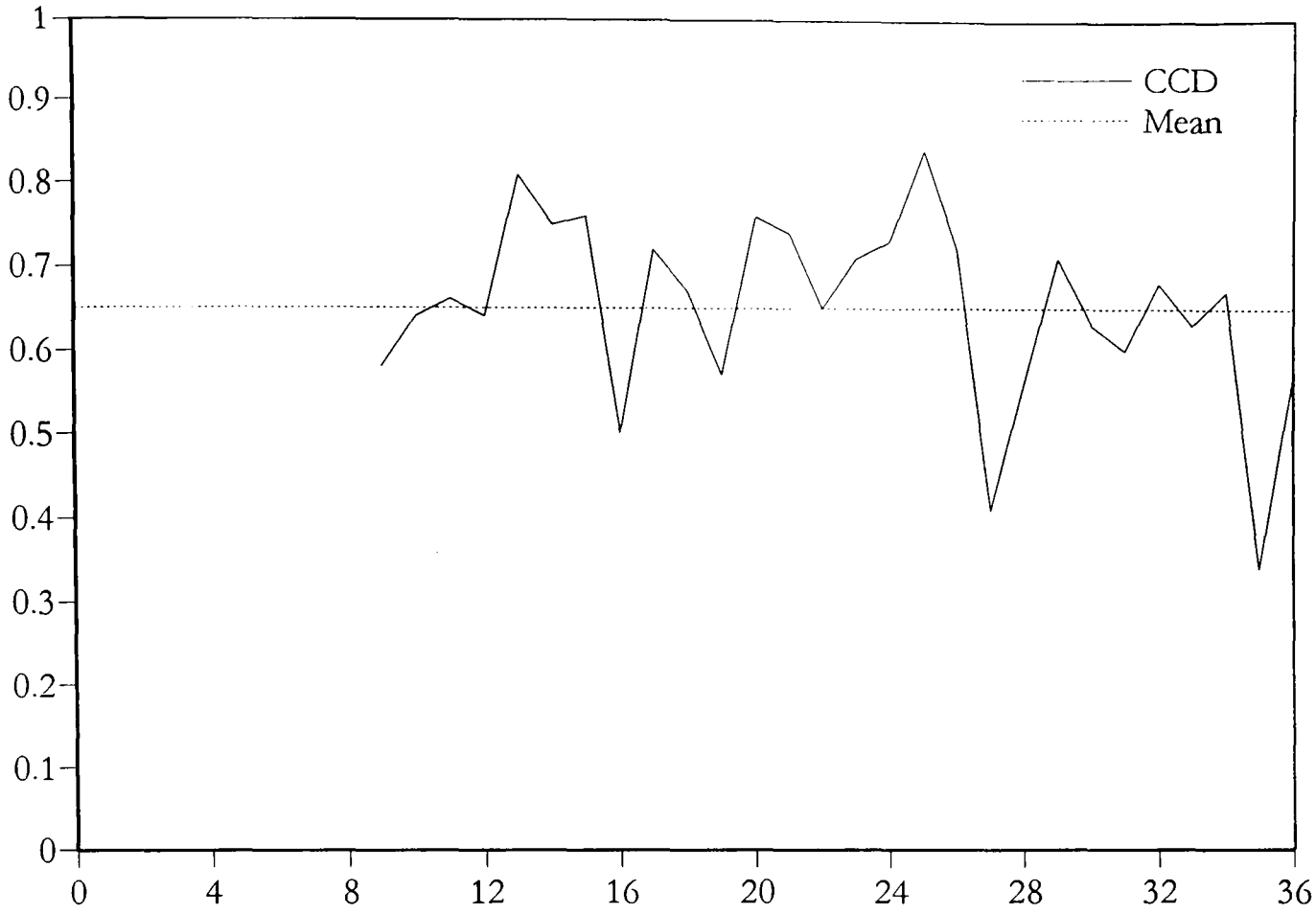


Figure 3.9. The adjusted r^2 of the monthly comparison of total monthly precipitation and total monthly CCD for 1988 - 1990. The mean level of variance explained is also shown as the horizontal dotted line.

3.4 Discussion

The comparison of point meteorological readings with satellite sensor derived data values arising from an 8 x 8 km pixel sample area will be subject to a considerable amount of noise due to the spatial heterogeneity of the ground data, and this effect is unlikely to be uniform on the pan-African scale. For this reason assessment of absolute accuracies for satellite sensor surrogate predictions of meteorological data in this work was not appropriate. What we can learn from these comparisons, however, is the relative predictive power of the satellite sensor data layers for describing a particular meteorological variable. These results can then be used to refine the data used for subsequent land-cover mapping, and ultimately be applied to the central problem of tsetse fly distribution and abundance predictions.

3.4.1 Prediction of land surface temperature

The data compared in this part of the analysis were the monthly mean NOAA - NCDC air temperatures and monthly maximum land surface brightness temperatures. The relative performance of the AVHRR thermal channels in describing the NOAA - NCDC air temperatures might be expected to follow from the relative rates of atmospheric attenuation experienced at these wavelengths (Kidwell 1995); see figure 1.1. The positive linear nature of the relationships and the improvement in the correlations brought about by the split-window corrections to the brightness temperatures were also anticipated. A more detailed analysis of these relationships, including a consideration of the factors affecting the slopes and intercepts of the regression lines (for a detailed discussion of the possible sources of error see Prince and Goward. (1996)), was complicated by the continually changing monthly sample of meteorological stations. In addition, among any given monthly subset, the data could be recorded from sites located in habitats ranging from desert to rainforest, potentially differing by up to 2000 m in elevation and experiencing different seasons. Despite these complicating factors two observations were consistent throughout the comparisons, both of which can be seen in figure 3.2. The first was that the temperature predictions based on the satellite sensor variables were consistently high. This is to be expected where areas are dry and hence have low evapotranspiration regimes, as there is no cooling effect from the evaporation of water (Price 1983). This evapotranspiration affect is likely to be exacerbated in these comparisons as the brightness temperature measurements are taken during an afternoon satellite overpass, when the land (of high specific heat capacity) retains energy from

morning and early afternoon heating. The second was that the regression coefficients were affected markedly by the range of values in NOAA - NCDC air temperature. This sensitivity to range in dependent variable data is a feature of all linear regression comparisons (Williams 1993).

The Price (1984) split-window technique was consistently better at describing NOAA - NCDC air temperature than the Becker and Li (1990b) equation, which in turn was always a better predictor than the Prata and Platt (1991) formulation. This result was somewhat surprising considering the simplicity of the Price (1984) algorithm. The explanation for this observation is likely to be related to the errors involved in estimating emissivity, since it was required by both of the poorest performing techniques. The assumption of equivalence in emissivity at the channel 4 and 5 wavelengths would also therefore be likely to compound errors. The sensitivity of split-window equations to emissivity was highlighted by Becker (1987) who noted that an error in the assumption that $\varepsilon_4 = \varepsilon_5$ of only 0.0007 would lead to errors of more than 0.5 K in the resulting temperature. An alternative to the above procedure would be to incorporate an emissivity map (*e.g.* Prabhakara and Dalu 1976) as Prince and Goward (1995) used when determining land surface temperature for inputs to global primary production models. Ideally however, the NDVI and emissivity relationship would be parameterised over a range of land-covers in African habitats throughout the season, so that its application could be adjusted temporally and spatially. Many methods have been suggested to determine emissivity (Li and Becker 1993) and might also be explored. These methods range from laboratory radiometer measurements of emissivity for different surfaces extrapolated to the field (Nerry *et al.* 1990, Salisbury and Daria 1992) to comparisons of the day and night brightness temperatures from AVHRR channel 3 that is used to eliminate the emitted thermal signal and leave only the reflected emissivity component of the reading in the day (Becker and Li 1993).

Elevation was found to account for a considerable amount of the residual variance in the relationship between split-window brightness temperature and the NOAA - NCDC air temperature (figure 3.5). This trend was revealed by the geographical distribution of the residual values (figure 3.3), with positive deviations from the regression line occurring at altitudes above 500 m and negative deviations for those below. Furthermore, the largest of these deviations were associated with the extremes of the elevation range. It is evident that at an altitude of 2000 m atmospheric

attenuation of the TIR will be significantly less than for a site at sea level as the signal must travel through 2 km less of moist atmosphere. Elevation has not been explicitly parameterised in any split-window equations but should be considered where application is required at the broad spatial scale or where the range in local elevation is large. The less satisfactory approach, followed in this work, was to use these split-window data in combination with a DEM.

It was surprising that the NDVI did not affect the correlations since strong relationships between brightness temperature and NDVI have been demonstrated at the local (Nemani and Running 1988) and global scale (Schultz and Halpert 1993). Typically the TIR signal decreases as the amount of vegetation in the field of view of the sensor increases and this is thought to be related to evapotranspiration regimes *i.e.* greater latent than sensible heat. This relationship has been exploited in a technique to derive air temperature by assuming that as vegetation coverage increases, so does the proportion of the TIR signal that can be attributed to it (Nemani *et al.* 1993). Since vegetation has a low specific heat capacity it rarely deviates more than 1 - 2 K from the surrounding air in a dense canopy, so contextural approaches (using a pixel array) can then be used to regress NDVI and brightness temperature and extrapolate the regression line to the air temperature of an "infinite canopy" *i.e.* NDVI = 1.0 (Goward *et al.* 1994). This technique was not used in the present analysis, but might be considered in future applications as both land surface temperature (during the puparial stage) and air temperature (as flying adults) are important in tsetse ecology (see section 3.1).

Seasonal trends in the data were apparent with maximal r^2 occurring around the months of April and May and minimal r^2 around the months of September and October. These relationships were very noisy however, and since seasonal trends in the data will have varied geographically across the continent these trends were difficult to explain. The largest deviations in degree of variance explained in the regressions did not coincide with changes between the AVHRRs onboard the NOAA - 9 and NOAA - 11 satellites in November 1988 (Kidwell 1995) or major adjustments to the calibration of these sensors (Rao 1993a and b). There were also no obvious trends in the data that could be related to recession in the overpass time of the NOAA satellites (Price 1991), which could have a substantial affect on the daily recorded brightness temperatures (Prince and Goward 1996). The noise in the monthly variation in the coefficient of determination is therefore, attributed largely to the variable subset of meteorological

stations used in the comparisons, as well as the inherent noise when comparing point data with values of 8 x 8 km pixels. The seasonal trends could only be resolved by stratifying the meteorological station data into seasonal zones, whereupon the samples would be very small.

A comparison of the accuracies achieved in this investigation with those detailed in the introduction is unwise, as previous studies have analysed daily comparisons of LAC AVHRR data over small areas (Cooper and Asrar 1989, Sugita and Brutsaert 1993, Becker and Li 1995). Little work has been done to validate these data at continental scales because continental meteorological datasets do not usually contain land surface temperature. Recent work using data from the UK climate records has shown that methods used to interpolate point meteorological data should not be dismissed as they may achieve accuracies approaching 95 % for points distant from the observations, when a sufficient density and spread of recording stations is available (Lennon and Turner 1995). Preliminary work using the same interpolation routines indicates similar levels of accuracy to the Price (1984) and DEM data predictions using half the 250 station NOAA NCDC dataset (Hay and Lennon 1996).

To conclude, when elevation is used as an additional predictor variable, the variance about the three year mean r^2 was small, allowing confidence in the ability of the Price (1984) equation in combination with altitude from a DEM to parameterise the land surface temperature of the African environment the thermal component of the tsetse fly niche. The data derived from the Becker and Li (1990b) and the Prata and Platt (1991) equations were not used in the subsequent analyses due to their poorer performance in predicting NOAA - NCDC air temperatures.

3.4.2 Prediction of atmospheric moisture

In the comparison of monthly mean NOAA - NCDC saturation deficit and monthly maximum vapor pressure deficit, no difference was found between using the Dalu (1986) or the Eck and Holben (1994) estimates of total precipitable atmospheric water. The overall accuracy was high and the variance low despite the many steps required in the VPD calculation. This was also surprising because the VPD incorporates the errors involved in the estimation of surface air temperature as well as total precipitable atmospheric water. A possible reason for this could be less spatial heterogeneity in the VPD, such that a point reading is on average more representative of

the corresponding 8 x 8 km remotely sensed pixel and would also corroborate the idea that the comparisons of point data with a large area pixel average accounts for a large part of the noise in the relationships.

The coefficients of determination for the monthly relationships were again affected strongly by the range in the ground data, with the highest correlations occurring in the months with the greatest range in ground data. Furthermore, the results were too noisy to determine seasonal trends in the data and again the major deviations were not related to switching between satellites, or to major calibration changes or to the delay in time of afternoon overpass. No significance is attributed to the very marginal increase in mean three-year r^2 when using an exponential fit to the data.

Absolute assessments of accuracy were again inappropriate (see 3.4.1), but the highly significant relationship between satellite sensor derived VPD estimates and NOAA - NCDC saturation deficits throughout the different habitats and elevations of Africa provides confidence that the atmospheric moisture dimension of the tsetse fly habitats can be predicted remotely.

3.4.3 Prediction of rainfall

The large spatial variation in rainfall restricts comparison of point rain-gauge data and satellite sensor estimates of rainfall, since rainfall measurements over a season can vary by a factor of two over distances of less than 10 km (Flitcroft *et al.* 1989). It follows therefore, that this should be the least robust of the relationships. Despite these considerations, however, the coefficients of determination were high, albeit with significant variation about the means (see figure 3.9). These correlations were not improved when only the TAMSAT region (figure 2.2) of CCD calibration was analysed, due in large part to the smaller sample sizes. There is a need for calibration of such data over the whole of Africa, where weather systems are more complex and are not always dominated by convective rainfall phenomena on which the CCD technique relies. As in the previous comparisons, the larger the range in NOAA - NCDC data the better the correlations. Again there seem to be no consistent temporal trends or biases that can be resolved from the noise in these data.

In sum, this was the least well predicted of the meteorological variables but it is also a climate variable which is probably of less importance than saturation deficit and temperature to tsetse ecology (see section 3.1) except when in extreme (*i.e.* drought

periods and floods). We can, however, describe the rainfall component of the tsetse niche with acceptable accuracy.

In each of these comparisons accuracy assessment is difficult. Progress would be made if a network of automated meteorological stations were established to take readings over an 8 x 8 km area at a series of test sites across Africa. This would give much better insight into the factors affecting these relationship (*e.g.* vegetation coverage and altitude) and the problems of relating two different scales of data. This will become increasingly important as these data are more widely applied to a diversity of subjects ranging from primary productivity to climate modelling (Ehrlich *et al.* 1994).

3.5 Conclusions and summary

Despite the range of factors contributing noise to the signal received by meteorological satellite sensors (see section 1.3), the correlations with ground-based meteorological data were high. This suggests that the noise, though not negligible, was small in comparison with the range of data values recorded. In addition the relative accuracy is likely to be high. Though seasonal trends could be seen to affect the temperature data they were of same magnitude as the natural variance which in turn did not seem to be related to variation in the satellite and radiometer performance. Remotely sensed meteorological data therefore present a significant advance in both temporal and spatial resolution over meteorological station records. Based on the relative performance of the indices, the Price (1984) split-window brightness temperature, Eck and Holben (1994) VPD_2 and the FAO ARTEMIS CCD data were used for the land surface temperature, atmospheric moisture and rainfall respectively, in the subsequent analysis.

Chapter 4

REMOTELY SENSED PREDICTION OF LAND-COVER

" ... *Thought is impossible without an image* ... "

Aristotle.

4.1 Introduction

4.1.1 Land-cover and tsetse ecology

The term land-cover is used to describe the naturally vegetated, non-vegetated and human affected landscapes within a region (Meyer and Turner 1992, 1994). These landscapes are themselves the product, *inter alia*, of small-scale variation in the distribution of soils, temperature, atmospheric moisture and precipitation. Natural and anthropogenic forces act continuously to modify and change land-cover. The objective of the work described in this chapter was to map this mosaic using meteorological satellite sensor data as an aid to understanding and predicting the distribution and abundance of tsetse.

The importance of habitat to tsetse ecology was recognised by those first concerned with the problem of tsetse and hence trypanosomiasis control (Swynnerton 1936), who realised that modifying vegetation through either total (early) or partial (later) bush clearance would make the environment unsuitable for tsetse (Ford 1971). The distribution of land-cover directly defines the natural barriers and hence boundaries to tsetse ranges (Ford 1968). Information on land-cover is also important because tsetse exploit different habitats to ameliorate extremes of climate both spatially and temporally, so that a detailed knowledge of meteorological factors alone is not sufficient to predict habitat suitability. An example of such behavioural adaptation to weather variation was shown by Pilson and Leggate (1962) who studied *G. pallidipes* in the Zambezi Valley. They found that after feeding the flies "perched" on the underside of the branches of small trees and shrubs. Perching sites were between 1 and 3 m above ground for most

of the year, but when saturation deficit and temperature become dangerously high in the dry season, they occupied these sites only in the afternoon and evening. Around midday they descended to the cooler air below 1 m, despite a presumed greater risk of predation by lizards. Harley (1954) also showed that the larviposition sites used by female *G. morsitans* varied throughout the year. The warmest sites were favoured during the coolest months and the often shady and hence cooler sites favoured in the dry (and hot) season.

As has already been detailed (see section 3.1.1), tsetse systematists have divided the genus *Glossina* into three subgenera, each with markedly different habitat and host preferences. It follows that an accurate description of land-cover is required to describe the distribution of suitable habitats for the different tsetse species, and thus to help understand the relative risks of transmission of the animal and human forms of trypanosomiasis within a region.

The potential usefulness of land-cover mapping in studies of tsetse ecology has been stressed. The ability to obtain this information from meteorological satellite sensors is particularly attractive because of the substantial problems that exist in acquiring and maintaining accurate information on land-cover over large areas using conventional techniques. Ground surveys are subject to interpreter bias, to potential logistical difficulties resulting from local topography and infrastructure and, not least, to the overall expense of the survey exercise. Aerial photographic techniques can cover large areas rapidly and acquire data on a systematic basis but are also often prohibitively expensive. Remote sensing by meteorological satellite sensors, in contrast, can provide a relatively inexpensive and homogenous sample over extensive geographical areas and through a range of spatial, temporal and spectral resolutions depending on the choice of satellite and sensor (see table 1.1.1).

4.1.2 Remote sensing and land-cover

Existing attempts to map land-cover at regional and global levels have concentrated on the analyses of low spatial resolution, multitemporal SVIs (*e.g.* Townshend and Tucker 1984, Justice *et al.* 1985, Tucker *et al.* 1985a, Townshend and Justice 1986, Townshend *et al.* 1987, Marsh *et al.* 1992, Andres *et al.* 1994, DeFries and Townshend 1994, Olsson and Eklundh 1994) as they offer the only practicable approach to this mapping problem at this broad spatial scale (Foody and Curran 1994). The

studies most commonly derive GAC NDVI data from the visible and NIR channels (1 and 2) of the NOAA AVHRR. Multitemporal datasets are favoured because they incorporate information on the seasonal variation of land-cover, which is the expression of constituent species phenologies at the biome scale (Fontés *et al.* 1995). In section 2.5.2 it was outlined how temporal Fourier analysis is used in this study to transform multitemporal SVI signals of vegetation amount into a phenological signature of considerable use in land-cover mapping.

The alternative SVIs such as the RVI, SAVI and GEMI detailed in section 1.4.1 have not been widely adopted for land-cover mapping. Moreover, there has been only a limited exploration of the other AVHRR channels in this context, so that the information in the MIR channel 3 and TIR channels 4 and 5 is often wasted (Ehrlich *et al.* 1994). The physical interaction of such EMR with vegetation canopies is not well understood (Goel 1988), but the use of similar spectral data from the Landsat - TM suggests these wavelengths reveal important information on the biophysical properties of vegetation (*e.g.* Horler and Ahern 1986, Curran and Williamson 1987, Boyd *et al.* 1996).

The relatively few applications of TIR - AVHRR data include mapping the seasonal evolution of vegetation and forest cover in West Africa using LAC AVHRR channel 4 (Achard and Blasco 1990) and using the Price (1984) split-window corrected brightness temperatures in combination with the NDVI to predict whole Africa land-cover (Lambin and Ehrlich 1995, 1996, Ehrlich and Lambin 1996). These investigations commonly concluded that using TIR usefully complements the information gathered from the visible and NIR channels. The MIR data have also been used to discriminate forest boundaries (Tucker *et al.* 1984, Kerber and Schutt 1986, Malingreau *et al.* 1989) and the suitability of these wavelengths in land-cover mapping is discussed in section 1.4.2. Despite being strongly correlated with visible radiances, the MIR suffers less atmospheric attenuation due to the smaller size of the ratio between aerosol size and radiation wavelength and this results in better haze penetration which is particularly important in the tropics (Bird 1984). MIR is also thought to provide some of the only information on the drought adapted microphyllous vegetation that is relatively abundant in semi-arid environments and that does not respond well to conventional SVIs (Kaufman and Remer 1994, Ringrose and Matheson 1995). A recent systematic evaluation of the information content of the five AVHRR channels was provided by

Foody *et al.* (1996), where the authors found that the correlations between remotely sensed radiation and the biophysical properties of tropical forest in Ghana were stronger in the MIR and TIR wavelengths than those from the visible and near infrared.

To date there has been no comparison of the discriminatory ability of the alternative SVIs in land-cover mapping. The MIR and TIR radiances have also been under-used (Ehrlich *et al.* 1994, Boyd and Curran 1997). In this chapter the potential for land-cover discrimination of the range of SVIs detailed in the introduction is evaluated. The DEM data are also included in the analysis, as well as the refined set of land surface temperature, atmospheric moisture and rainfall indices defined in chapter 3. In this approach, the different environmental factors that are known to affect land-cover (and can be parameterised with meteorological satellite sensor data) are investigated as an alternative to looking for direct relationships between biophysical variables and radiance data from satellite sensors. The advantage of this approach is the potential for increased ecological and epidemiological insight.

A further problem with existing land-cover studies at the regional scale is an objective assessment of product accuracy (Trodd 1995). This is generally due to a lack of ground data over such large areas to test predictions. In this work, accuracy is assessed using a variety of statistical measures that compare the predictions against an extensive Nigerian land-cover dataset. The results are then evaluated using the Anderson criteria for land-cover map assessment (Anderson *et al.* 1976).

There are many advantages to establishing a reliable technique to map the distribution of land-cover beyond the present aim of defining tsetse fly habitats. Primarily the establishment of a baseline inventory of land-cover resources is the basis from which a better understanding of the factors that control land-cover can develop and change be monitored. This information facilitates better informed management of the present situation and should enable future changes to be more effectively predicted. These benefits are especially important in developing regions, where land-cover information is often inadequate and where ever-increasing human populations effect environmental change.

4.2 Methods

This section details the specific datasets used in the land-cover analysis and the statistical techniques employed to establish and interrogate the relationships. The land-cover data are first described and the satellite sensor derived "predictor" variables summarised. The statistical techniques used are then outlined with some preliminary justification for their adoption.

4.2.1 The land-cover data

As part of a World Bank funded national livestock survey, a series of low-level systematic reconnaissance flights were flown over Nigeria in 1990 (RIM 1992, Bourn *et al.* 1994). A uniform sampling regime was achieved by flying at a nominal altitude of 245 m above ground level along parallel transect lines spaced at 20 km intervals (as demonstrated in figure 4.1). Each line was divided into sectors of 20 km which resulted in data being recorded for 2,280 grid cells (each measuring 20 x 20 km) for the whole 924,000 km^2 land area of Nigeria. The first survey was conducted between March and April 1990 at the end of the dry season. The northern 80 % of the country was surveyed again between September and October 1990 at 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. In the wet season only active cultivation was recorded, since the extent of other land-cover types did not change significantly *i.e.* the majority of the land-cover was perennial.

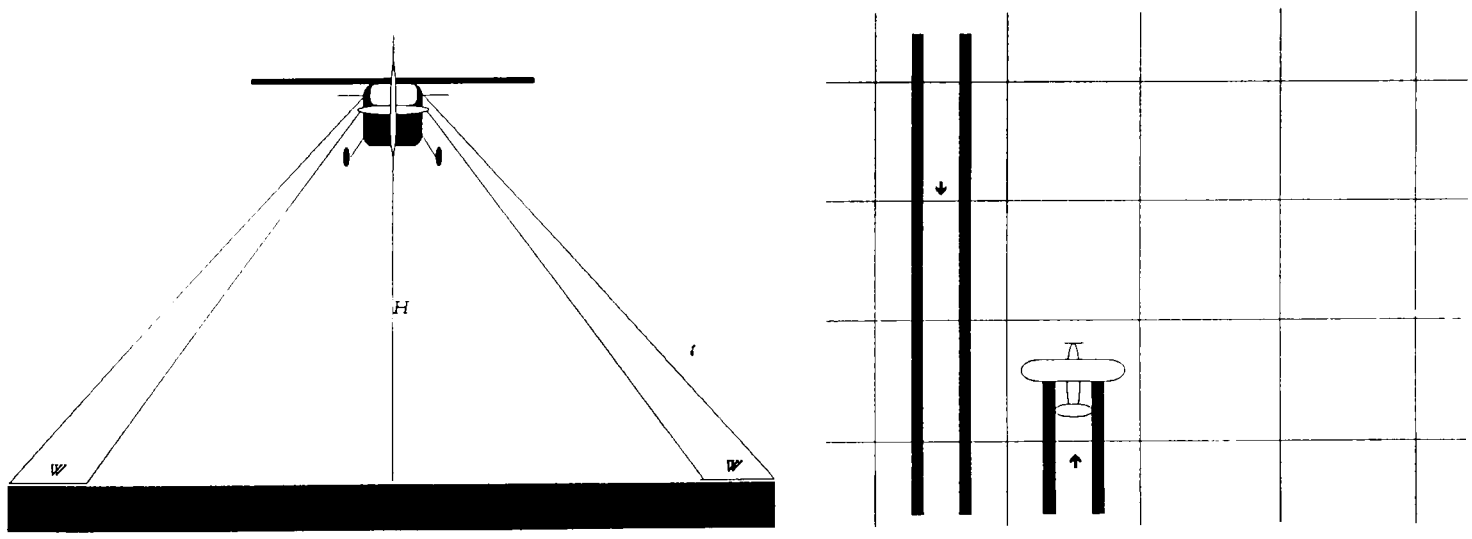


Figure 4.1. A diagram of the aerial survey sample strips. The survey aircraft (to the left), has a height, *H*, of approximately 245 m resulting in the width of the viewing strip, *W*, of approximately 500 m. The aerial survey grid sampling pattern is shown to the right. Redrawn from RIM (1992).

The sampling density of the aerial survey can be calculated as 5 % (two 0.5 km wide strips viewed by the observers across a 20 x 20 km grid cell) assuming a constant altitude. A measure of the accuracy of these survey observations can be gauged from a comparison of the visual survey estimates with those from aerial photograph interpretation for the active cultivation class in northern Nigeria. The techniques resulted independently in mean percentage cover estimates of 32.5 and 32.05 % respectively (ERGO 1994).

Three datasets were created from these aerial survey data. They consisted of a sub-sample of those survey grid-squares in which a land-cover class dominated more than 50, 60 and 70 % of the area of the grid-square. These 50, 60 and 70 % "threshold" datasets contained 44, 38 and 31 % of the original 2280 surveyed grid-squares respectively. Dominant vegetation classes only were used because a grid-square with only 30 % of a particular land-cover consists of a 70 % mixture of other land-cover classes and would present obvious difficulties for the analysis.

4.2.2 The satellite sensor data

The relative performances of the RVI, NDVI, SAVI and GEMI SVIs in predicting the land-cover of Nigeria were assessed using a discriminant analysis procedure described below. The refined dataset arising from the comparisons of prediction "accuracy" for the meteorological variables described in chapter 3 was also presented to the analysis. Furthermore, data from the MIR channel 3 were included as they contain spectral information not present in any of the other variables. For each of the variables, the Fourier mean and the amplitude and phase of the annual, bi-annual and tri-annual signal were calculated (see section 2.5.2). The maximum, minimum and range of the Fourier series were also provided, so that for each satellite sensor derived data band there were ten Fourier derived variables presented to the analysis. In addition to these data, elevation information from the GLIS DEM was also included as a predictor variable. The satellite sensor derived data are summarised in table 4.1.

4.2.3 Analysis and statistical procedure

The "predictor" data were subject to a linear discriminant analysis performed using Quick Basic programs written by DJR. The reader is referred to section 5.2.3 where justification for the use of discriminant analysis is given and the technique

TABLE 4.1.

A summary of the eight bands of satellite sensor derived data used to predict Nigerian land-cover. The maximum and minimum values of the Fourier mean in the dataset are presented with the abbreviations used in following tables expanded.

Data type	*Abbreviation*	*Maximum*	*Minimum*
Satellite data layer			
Channel 3	Ch 3	320.7 K	302.7 K
Price (1984)	Price	326.1 K	304.3 K
Vapour Pressure Deficit	VPD	8.76 mb	1.93 mb
Cold Cloud Duration	CCD	510 hr	10 hr
Ratio Vegetation Index	RVI	1.86[a]	0.34
Normalised Difference Vegetation Index	NDVI	0.43	0.05
Soil Adjusted Vegetation Index (L = 0.50)	SAVI	0.65	0.09
Global Environmental Monitoring Index	GEMI	1.73	-0.38
Fourier variables			
annual cycle amplitude	amp 1	-	-
annual cycle phase	phs 1	-	-
bi-annual cycle amplitude	amp 2	-	
bi-annual cycle phase	phs 2		
tri-annual cycle amplitude	amp 3		-
tri-annual cycle phase	phs 3	-	-
Fourier mean	Fmean	-	-
Fourier maximum	Fmax		-
Fourier minimum	Fmin		-
Fourier range	Frange		-

a: SVIs have no units as they are ratios.

explained in detail with respect to the simple case of separating a tsetse presence from tsetse absence class. The 50, 60 and 70 % threshold subsets were analysed independently in a supervised classification exercise in which a different covariance matrix was calculated and used for each vegetation type. Prior probabilities were weighted to take account of the sample size of a particular land-cover class in the dataset (Green 1978, Tatsuoka 1971). The ten variables which gave the greatest separation in multivariate space between land-cover classes (measured by the summed Mahalanobis distances, called the Mahalanobis Index within the analysis) were chosen from the 80 processed satellite sensor derived variables and DEM data. These ten variables were then used to describe the land-cover data in multivariate space. The predicted classification of each pixel was defined by its closeness, again measured by the Mahalanobis distance, to each one of the "clouds" of pixels representing the different land-cover types. Posterior probabilities were calculated to give some indication of the likelihood of misclassification.

The ability of the discriminant analysis to separate vegetation classes, or the accuracy with which the generated maps predicted survey data, was measured in several ways. First "producer" and "consumer" accuracies were calculated for each vegetation type as the percentage of grid-squares within the training set known to have been classified correctly (the producer's accuracy) and the percentage of grid-squares assigned to a particular class that actually belonged to that class (the consumer's accuracy). Secondly the Kappa (κ) and Tau (τ) statistics were calculated from the marginal totals of the rows and columns of the classification matrices (Ma and Redmond 1995). These statistics compare the accuracy of the predicted classifications with those expected on the basis of random assignments of each grid-square, given either the final proportions of grid-squares belonging or assigned to a particular vegetation type (*i.e.* using the *a posteriori* probabilities κ) or only the proportions of grid-squares belonging to a particular vegetation type (*i.e.* using only the observed *a priori* probabilities - τ). If the observed agreement between the satellite sensor and aerial survey data equals that which could have been expected by chance then κ and τ equal zero and if the agreement was perfect then κ and τ equal one.

4.3 Results

4.3.1 The relative importance of predictor variables.

The top ten variables chosen by the discriminant analysis from the 80 Fourier variables and the DEM for the 50, 60 and 70 % dominant land-cover class subsets are listed in table 4.2. Elevation was always the most significant predictor, followed in each case by the phase of the bi-annual cycle of the MIR (Ch 3 phs 2). The remaining variables for the three threshold datasets varied considerably in their type and ranking, consisting of SVIs (54 %), TIR data (29 %), MIR data (13 %) and VPD (4 %). The CCD variables were never chosen by the analysis. Consideration of all the rankings showed the Fourier phase components to dominate the variables (46 %), followed by the amplitudes (30 %) and then by any of the Fourier range, maximum, minimum or means (14 %) (the remaining 10 % were elevation). Among the 27 Fourier variables bi-annual cycles predominated (41 %) followed by the annual cycle variables (33 %) with the "range" values (15 %) and tri-annual cycles (11 %) of less importance in prediction. The overall accuracy of the analyses was found to increase with the higher threshold vegetation coverages.

The predictive ability of the four SVIs was further resolved by presenting them independently to the discriminant analysis and the results of these tests are displayed in table 4.3. These comparisons were performed using the 60 % land-cover threshold dataset because at higher thresholds the sample sizes in some of the land-cover classes became too small to make reliable predictions. The elevation, the phase of the bi-annual cycle of the MIR (Ch 3 phs 2) and the amplitude of the annual TIR signal (Price amp 1) were also presented to the analysis on the basis on the previous 60 % threshold tests.

The overall accuracies of each SVI were again evaluated and a further indication of utility of the Fourier variables was gauged from how often they were chosen as predictor variables and where they occurred in the rankings. These were qualitative comparisons between the different Fourier components within a test for a given SVI and between the alternative SVIs in successive tests. Elevation, MIR (Ch 3 phs 2) and TIR (Price amp 1) were consistently the first, second and third ranked variables. The phase of the annual cycle of each SVI was always fourth in the rankings. The NDVI was the best predictor variable among the SVIs with 78 and 76 % overall producer and consumer accuracy respectively, with the SAVI (74 and 74 %), the RVI (74 and 70 %) and GEMI (67 and 63 %) being less accurate. The relative frequency of the phase (39 %), amplitude (36 %),

and any of the range, maximum, minimum and mean variables (25 %) in rankings was more even than with the 50, 60 and 70 % threshold tests. Furthermore, the annual (31 %) and bi-annual cycles (31 %) were of equal importance among the 36 Fourier variables and were again followed by the "range" variables (25 %) and the tri-annual cycle (13 %) in predictive importance.

TABLE 4.2.

The top ten predictor variables for Nigerian land-cover at the 50, 60 and 70 % dominance thresholds.

Rank	*Threshold*		
	50 %	*60 %*	*70 %*
1	Elevation	Elevation	Elevation
2	Ch 3 phs 2	Ch 3 phs 2	Ch 3 phs 2
3	VPD Fmin	Price amp 1	SAVI phs 1
4	NDVI phs 2	NDVI phs 1	GEMI amp 2
5	NDVI amp 2	NDVI amp 2	Ch 3 phs 3
6	Price amp 1	NDVI phs 2	GEMI phs 2
7	NDVI amp 1	RVI phs 3	GEMI phs 1
8	Price Fmean	Price phs 2	Ch 3 amp 2
9	Price Fmin	Ch 3 phs 1	GEMI Fmin
10	Price amp 3	GEMI phs 1	Price amp 1
Producer accuracy (%)	65.1	73.4	74.9
Consumer accuracy (%)	68.1	72.3	80.7
Kappa	0.499 (0.461 0.536)[a]	0.550 (0.509 - 0.592)	0.637 (0.591 0.684)
Tau	0.494 (1.456 0.532)	0.547 (0.506 - 0.589)	0.636 (0.589 0.683)

a: the brackets indicate the 95 % confidence intervals of the Kappa and Tau statistics

TABLE 4.3.

A comparison of the relative performance of the SVIs in predicting the 60 % dominant land-cover class threshold for Nigeria.

Rank	*Spectral Vegetation Indices*			
	RVI	*NDVI*	*SAVI*	*GEMI*
1	Elevation	Elevation	Elevation	Elevation
2	Ch 3 phs 2	Ch 3 phs 2	Ch 3 phs 2	Ch 3 phs 2
3	Price amp 1	Price amp 1	Price amp 1	Price amp 1
4	RVI phs 1	NDVI phs 1	SAVI phs 1	GEMI phs 1
5	RVI phs 3	NDVI amp 2	SAVI Frange	GEMI amp 1
6	RVI Fmax	NDVI phs 2	SAVI phs 3	GEMI phs 2
7	RVI amp 3	NDVI phs 3	SAVI amp 2	GEMI Fmean
8	RVI Fmin	NDVI Fmax	SAVI Fmean	GEMI Fmin
9	RVI amp 1	NDVI Fmean	SAVI amp 1	GEMI amp 3
10	RVI amp 2	NDVI Fmin	SAVI phs 2	GEMI amp 2
Producer accuracy (%)	74.0	77.8	73.6	67.0
Consumer accuracy (%)	69.9	76.0	73.5	62.5
Kappa	0.59 (0.527 - 0.607)[a]	0.61 (0.578 - 0.656)	0.61 (0.555 - 0.634)	0.60 (0.469 0.554)
Tau	0.59 (0.523 - 0.604)	0.61 (0.577 - 0.655)	0.61 (0.552 0.631)	0.60 (0.404 - 0.549)

a: the brackets indicate the 95 % confidence intervals of the Kappa and Tau statistics

4.3.2 Classification accuracy

The classification accuracy of the Nigerian land-cover prediction results are again discussed with reference to the 60 % dominant land-cover data subset because it represents a compromise between grid-squares of relatively homogenous land-cover composition and having enough samples in each land-cover class to make realistic predictions. The observed (or training set) data and the prediction based on the linear discriminant analyses of the NDVI, DEM, MIR, TIR and CCD are shown in figure 4.2. Table 4.4 provides an indication of the accuracy with which the component land-cover types were predicted and can be interpreted more readily in conjunction with the classification matrix given in table 4.5. The classification or confusion matrix shows the predicted (columns) against the observed (rows) pixels for each of the land-cover classes marked in table 4.4. If the statistical description of the ground data were perfect all the classification numbers would occur on the diagonal.

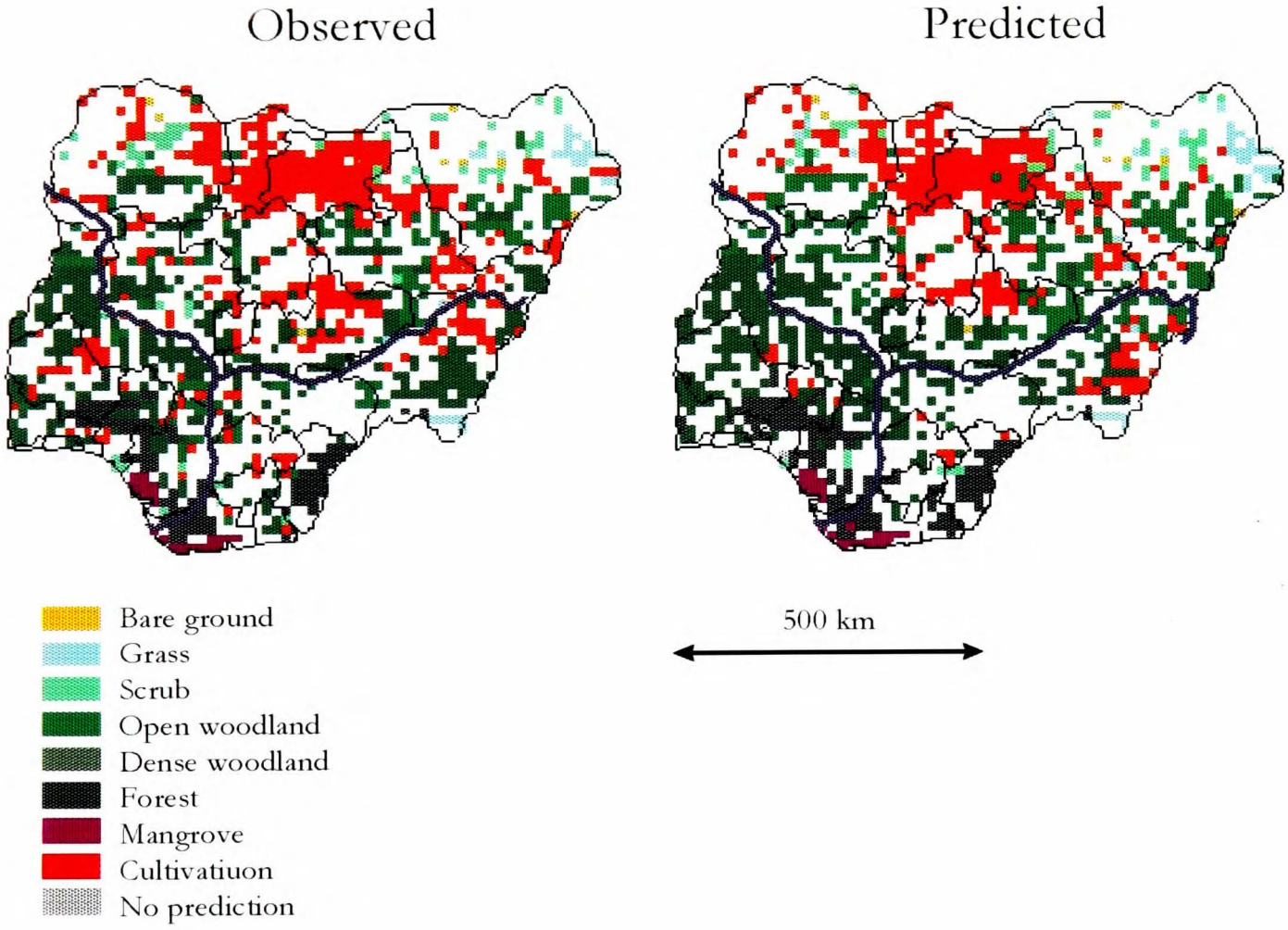


Figure 4.2. A map of observed and predicted land-cover classes for Nigeria at the 60 % dominant land-cover threshold. "No prediction" indicates a grid-square with values of predictor variables more extreme than any in the training set grid-squares. North is to the top of the page.

TABLE 4.4.

Classification accuracy of the land-cover classes at 60 % coverage of dominant land-cover.

Category	*Type*	*Producers' accuracy (%)*	*Consumers' accuracy (%)*
1	Bare ground	100.0	100
2	Grass	73.1	70.4
3	Scrub	42.1	55.2
4	Open woodland	51.9	54.2
5	Dense woodland	70.0	60.8
6	Forest	92.6	66.9
7	Mangrove	100.0	95.2
8	All cultivation	57.6	75.8
Mean		73.4	72.3

TABLE 4.5.

A classification matrix of the land-cover classes at 60 % coverage of dominant land-cover.

				Predicted					*Total*
Observed	*1*	*2*	*3*	*4*	*5*	*6*	*7*	*8*	
1	**4**	0	0	0	0	0	0	0	4
2	0	**19**	2	3	1	0	0	1	26
3	0	1	**16**	2	3	5	0	11	38
4	0	4	6	**83**	45	3	0	19	160
5	0	0	1	30	**175**	29	0	15	250
6	0	0	2	0	4	**87**	1	0	94
7	0	0	0	0	0	0	**20**	0	20
8	0	3	2	35	60	6	0	**144**	250
Total	4	27	29	153	288	130	21	190	**842**

In terms of the Anderson criteria for evaluating land-cover classification accuracy and utility (Anderson *et al.* 1976) the results are encouraging. (i) Though minimum interpretation accuracy falls below 85 % in most classes, the overall level of prediction is high (73 % for both the consumer and producer accuracies) and where misclassifications occur these seem to be understandable. For example, open woodland is misclassified as dense woodland, dense woodland as forest *etc.* (ii) Furthermore, there are approximately equal accuracies across all the classes and where there are differences, these seem to be related to the sample size (*e.g.* compare bare ground and mangrove with open woodland). (iii) Finally, the reasons for the errors in the land-cover predictions are reasonably well understood and so the present results can be used as a baseline for comparison with future data.

4.4 Discussion

It was not surprising to find that when the threshold coverage for defining a land-cover class increased, so did the ability to predict its occurrence correctly. This is because the satellite sensor data were, presumably, increasingly dominated by the signal from the chosen land-cover classes. The variation between the accuracy of prediction using the threshold datasets highlights the problem of performing such an analysis on spatially heterogeneous land-cover and the resulting mixed composition pixels. The problem of attempting to define a class on the basis of a pixel whose signal may constitute 50 % of that class, but by definition therefore comprises 50 % of others, needs to be addressed.

One method proposed for this is to use the posterior probabilities since it has been claimed that when training sets are based on pure end members (*e.g.* 100 % grassland or 100 % forest) the posterior probabilities assigned to mixed pixels (mixels) represent the actual proportions of the different land cover types they contain (Foody and Cox 1994, Gopal and Woodcock 1994, Foody 1996). For example for an end member case of two classes (*e.g.* of pure grassland and pure forest) it is suggested 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 above 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 and has been confirmed (by DJR) with the Nigerian dataset. 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. Alternative methods of sub-pixel linear mixture modelling approaches (Holben and Shimabukuro 1993) have not been assessed and the effects of spatial autocorrelation have not been investigated.

Elevation was always the most important predictor variable, presumably because it was highly correlated with other predictor variables such as temperature and rainfall. It was more surprising to find that the MIR and TIR EMR were better at separating land-cover classes than the SVIs and this corroborates the findings of several of the studies outlined in section 4.1.2. Of the SVIs the most commonly used, NDVI gave the highest

predictive accuracy and was therefore used as the SVI in the final tsetse distribution and abundance analysis. The claims of decreased sensitivity to background soil effects and atmospheric attenuation for the "improved" SVIs must be further evaluated particularly with respect to the particularly poor performance of the GEMI.

The phase of the annual cycle in the SVI was consistently the most important predictor variable, demonstrating the importance of seasonality in defining land-cover types. Among these "seasonality" variables, the phase or timing of events and the amplitudes were most important, but the annual summary variables (range, maximum, minimum and mean) were also frequently chosen. The annual and bi-annual variables were the most common predictors among the Fourier variables which can be related to seasonal aspects of land-cover at low latitudes (Njoku 1963, 1964). For example Ewusie (1968, 1992) showed that, of a hundred tree species in West Africa, 48 flowered and fruited only once, 44 twice and 6 three times annually (the others displayed apparently continuous fruiting and flowering). The importance of bi-annual phenological patterns is related to the two peaks of rainfall and periods of maximum insolation experienced at these latitudes throughout the year. It was of interest therefore, to note that a mix of annual, biannual and triannual Fourier variables were chosen in similar proportions by the discriminant analyses (see table 4.2).

In a review of accuracy assessment in land-cover mapping and modelling studies, Trodd (1995) surveyed 84 land-cover classifications based on remotely sensed data published between 1994 and 1995. The mean overall percentage of correctly classified pixels was 79.1 %, a statistic to which the present results (73 % producers' accuracy with the NDVI dataset) compare favourably, when it is considered that many of the studies investigated used high spatial resolution satellite sensor imagery to make land-cover predictions at the local scale. This is especially important as widely used maps, such as the United Nations Educational, Scientific and Cultural Organisation (UNESCO) vegetation map of Africa (White 1983), become increasingly dated and environmental change continues apace.

The vegetation of the tsetse fly's niche can be predicted relatively accurately at the broad spatial scale. This, in combination with the results of chapter 3, provides confidence that Fourier processed meteorological satellite sensor variables are providing surrogate information about the environment, of use to ecological and epidemiological

studies and allows increased confidence in their application to tsetse distribution and abundance prediction. This is the subject of the final analysis chapter.

4.5 Conclusions and summary

The objectives of the present study, described in this chapter, were to investigate the extent to which meteorological were useful in predicting land-cover, to determine whether the accuracies achieved were acceptable and to consider the utility of the various SVIs and the MIR, TIR and DEM data in such predictions. The results indicated that, given a detailed ground dataset on which to "train" the satellite sensor data, current public-domain information from meteorological satellite sensors could predict land-cover to a useful level of accuracy. The NDVI was the best predictor of land cover, both quantitatively and qualitatively among the SVIs, but land-cover prediction was only feasible when the complete range of meteorological satellite sensor data were used in the predictions. These results were of considerable interest in defining the distribution of suitable habitat for tsetse species by extending these land-cover predictions in both space and time.

Chapter 5

REMOTELY SENSED PREDICTION OF TSETSE FLY DISTRIBUTION AND ABUNDANCE

"... No more things should be presumed to exist than are absolutely necessary ..."

William of Occam.

5.1 Introduction

The impact of climate and vegetation on tsetse ecology has been discussed (see section 3.1.1 and 4.1.1 respectively) and the results of the previous chapters suggest that data from meteorological satellite sensors can be used as reliable surrogates for meteorological variables and land-cover over wide geographical areas. In this chapter how these same Fourier-processed meteorological satellite sensor data perform in predicting the distribution and abundance of tsetse is investigated. The problems caused in sub-Saharan Africa by the tsetse-transmitted trypanosomiases are discussed and the limitations of traditional approaches to addressing these problems are outlined. The potential advantages of remote sensing in this regard are then highlighted, so that the rationale for this work can be understood. Perspectives on the long-standing debate regarding tsetse control and African land degradation are also provided.

5.1.1 The trypanosomiasis problem

The human and animal trypanosomiases occur between the latitudes of 14° north and 29° south in Africa, following the geographic distribution of the tsetse fly. Within this region 50 million people are at risk from contracting sleeping sickness (De Muynck and Rogers 1989) and some 25,000 new cases are reported annually (Kuzoe 1993). The vast majority of human cases prove fatal if untreated. In addition, the occurrence of nagana excludes non-trypanotolerant livestock from an area of approximately 10 million km^2 of Africa (Steelman 1976) affecting the supplies of meat and milk, as well as manure and draught power. Nagana is therefore a major constraint on basic agricultural

development (Murray and Gray 1984, Jordan 1986) and was estimated to have cost the African economy 5 billion US dollars in 1973 (Goodwin 1973).

5.1.2 The constraints of traditional intervention techniques

The historical and prevailing strategy for trypanosomiasis control is medical and veterinary surveillance, coupled with tsetse fly control or eradication (WHO 1986). Reliable surveillance requires trained personnel equipped with adequate resources who are, in turn, facilitated by an efficient internal infrastructure. These criteria are rarely satisfied in Africa so that epidemics can develop which affect large numbers of individuals, involve high case mortalities and are difficult and expensive to control (Kuzoe 1993). Tsetse fly control campaigns are subject to similar logistical limitations and it is salutary to note that there has been no significant decrease in the continental area covered by the fly since efforts to control and limit tsetse numbers began in colonial times (Rogers and Randolph 1988).

The specificity of tsetse as vectors for sleeping sickness and nagana, coupled with the continued difficulties of trypanosomiasis chemotherapy and vaccine development (Hide 1994) have resulted in a return to an emphasis on the reduction of vector populations for trypanosomiasis control. Eradication has not proved feasible on the regional scale (Rogers 1985), so that tsetse control campaigns aim to decrease fly numbers below a threshold where the disease cannot be sustained in host populations (Rogers 1988b; see section 1.5.2).

Understanding the factors that regulate the distribution and abundance of tsetse flies is central to promoting a scientific basis for vector control. Essential to this "scientific" approach is gathering information on environmental factors of importance in tsetse ecology at the national and regional scales, which is only feasible in Africa using remotely sensed data from satellite-borne sensors. This will contribute to the production of trypanosomiasis risk maps which would help provide a focus for existing veterinary and medical surveillance, and tsetse control efforts. Furthermore, with an increasingly detailed understanding of how environmental factors control tsetse populations today, comes the possibility of building models which might attempt to predict tsetse fly population and trypanosome disease dynamics in real time.

5.1.3 Tsetse control and land degradation

It has been argued that the control of animal trypanosomiasis can result in the degradation of the lands if effective control of stocking levels is not maintained (Ormerod 1976, 1986, Ormerod and Rickman 1988). This has resulted in concern that the overall effect of controlling nagana can be detrimental. This is despite inconclusive assessments of tsetse control as a potential factor in land degradation (Jordan 1992, Pender and Rosenberg 1995) and the demonstration that land degradation has arisen from a variety of other factors (Olsson 1993, Hiernaux 1993, Dougill and Cox 1995). The inference that an important disease of livestock and its consequent economic burden should be tolerated by rural Africans as a "natural" check on environmental degradation is morally repugnant. The whole debate would be more positively addressed if the control of nagana was seen as an essential component of the activities directed to achieving sustainable rural development (Rogers and Randolph 1988). This is of particular relevance as there is considerable and mounting evidence that keeping livestock in sub-Saharan Africa is evolving toward productive mixed farming, increasingly involving the use of animals for traction (and especially so in the more densely inhabited areas) (Bourn and Wint 1994, Bourn *et al.* 1994). This process involves the settling of many former pastoralists in sub-arid regions (although this is not necessarily voluntary (Monbiot 1994)) and an increasing appreciation of the economic attraction of mixed farming in sub-humid and humid regions (McIntyre *et al.* 1992).

5.1.4 Chapter objectives

That there is a need for reliable information on tsetse fly distributions across Africa is evident, as is the fact that the resources to provide this information are scarce. This chapter investigates how the information in the Fourier-processed meteorological satellite sensor dataset, refined in the previous two chapters, performs in predicting the distribution and abundance of tsetse fly species in West Africa.

Much of this chapter is based on Rogers *et al.* (1996). The present analysis expands on that work to see if a wider range of meteorological satellite sensor variables (which have been demonstrated to be of use in predicting meteorological variables and land-cover type over wide areas) improve the fit to the statistical models and also whether the analysis highlights any of the additional variables as useful predictors of tsetse distribution and abundance.

5.2 Methods

5.2.1 Tsetse fly species distribution data

The tsetse distribution data originate from two surveys of tsetse fly populations in West Africa conducted by the Office de la Recherché Scientific et Technique Outre-Mer (ORSTOM) (Laveissiere and Challier 1977, 1981). The first survey covered the majority of Burkina Faso and the second the whole of Côte d'Ivoire. Together these data span the complete ecological gradient in West Africa from the southern Atlantic coastline through forest, woodland and savannah to the southern edge of the Sahara. The resulting distribution maps of tsetse in Côte d'Ivoire and Burkina Faso are compendia of information gathered over the preceding decades and record species' presence at a spatial resolution of 0.167 degrees or approximately 19 x 19 km. The original data sources do not give complete spatial coverage and the maps do not necessarily record fly presence in areas where a species was thought by the compiling authors to be ubiquitous (*e.g. G. palpalis* Robineau - Desvoidy in the southern part of Côte d'Ivoire). Thus whilst records of fly presence on these maps are historically accurate, records of absence are occasionally misleading. The distributions of the following eight species of tsetse were used in the analysis; *G. morsitans submorsitans* Newstead, *G. longipalpis* Wiedmann, *G. palpalis s.l.*, *G. tachinoides* Westwood, *G. pallicera* Bigot, *G. fusca* Walker, *G. nigrofusca* Newstead and *G. medicorum* Austen.

The abundance of flies in the northern part of Côte d'Ivoire was monitored by a joint FAO - Deutsche Gesellschaft für Technische Zusammenarbeit (GTZ) project that ran from 1979 to 1980 and produced detailed maps of fly distributions at scales of 1:1,000,000 and 1:200,000 (IEMVT 1982). The data at a spatial resolution of 0.250 degrees or approximately 28 x 28 km are used in the subsequent abundance analysis. Flies were sampled using Challier Laveissiere (blue biconical) traps (Challier and Laveissiere 1973) placed in suitable habitats by the survey teams and left for short periods (usually 6 hours) before collection and removal. Given the very large area sampled and the short sampling time in each habitat, these data are likely to be affected by a number of confounding effects such as sampling errors, poor weather at the time of sampling and seasonality, so only the average values (flies per trap per nominal 6 - hour trapping session) were analysed. Some of the species present in the region are inadequately sampled by the biconical traps used in these surveys. The *morsitans* group for example is not especially attracted to biconical traps with or without odour baits and

is better sampled with hand-nets and standardised fly-rounds (Takken 1984), see section 3.1.1. Supplementary catches using hand-nets were recorded separately on the maps, but the coverage using this method was relatively poor and the analysis of tsetse abundance was based entirely upon the 6 - hour trap catch data of *G. morsitans*, *G. longipalpis*, *G. palpalis*, *G. tachinoides* and *G. fusca.*

5.2.2 The satellite sensor dataset

The reduced-dimension satellite sensor dataset described in section 2.5 and subsequently refined, through verification against meteorological and land-cover data in chapters 3 and 4 respectively, form the set of predictor variables for describing species' distributions and abundance. The satellite sensor imagery was further processed by selecting a block of 2 x 2 pixels positioned at the centre of each grid-square on the fly distribution, and the mean values for each square were used in the subsequent analyses. These predictor variables are shown in detail in table 5.1 with the maximum and minimum values for each in the training dataset.

TABLE 5.1.

Predictor variables used in the analyses of tsetse fly distributions in Côte d'Ivoire and Burkina Faso, and their observed maximum and minimum values in the training set data.

Abbreviation	*Name*	*Maximum*	*Minimum*
Altitude			
Elev	Elevation	829.8	0.0
Channel 3			
Ch 3 Fmean	Channel 3 Fourier mean	321.15	302.35
Ch 3 phs 1	Channel 3 phase 1	4.25_a	1.38
Ch 3 amp 1	Channel 3 amplitude 1	9.85	0.45
Ch 3 phs 2	Channel 3 phase 2	5.68	0.83
Ch 3 amp 2	Channel 3 amplitude 2	3.63	0.23
Ch 3 phs 3	Channel 3 phase 3	3.90	0.15
Ch 3 amp 3	Channel 3 amplitude 3	1.80	0.13
Ch 3 Fmax	Channel 3 Fourier maximum	327.76	305.54
Ch 3 Fmin	Channel 3 Fourier minimum	320.34	297.30
Ch 3 Frange	Channel 3 Fourier range	21.05	1.7

Temperature			
Price Fmean	Price (1984) land surface temperature Fourier mean	324.93	303.40
Price phs1	Price (1984) phase 1	5.25	1.00
Price amp 1	Price (1984) amplitude 1	11.90	1.53
Price phs 2	Price (1984) phase 2	4.60	0.83
Price amp 2	Price (1984) amplitude 2	5.80	0.35
Price phs 3	Price (1984) phase 3	3.85	0.63
Price amp 3	Price (1984) amplitude 3	3.00	0.18
Price Fmax	Price (1984) Fourier maximum	336.99	306.37
Price Fmin	Price (1984) Fourier minimum	319.37	300.52
Price Frange	Price (1984) Fourier range	25.35	4.18
Atmospheric moisture			
VPD Fmean	Vapour pressure deficit (VPD) Fourier mean	8.30	2.61
VPD phs1	VPD phase 1	6.33	0.45
VPD amp 1	VPD amplitude 1	3.19	0.39
VPD phs 2	VPD phase 2	5.18	0.30
VPD amp 2	VPD amplitude 2	1.16	0.13
VPD phs 3	VPD phase 3	3.65	0.38
VPD amp 3	VPD amplitude 3	6.92	0.07
VPD Fmax	VPD Fourier maximum	11.43	3.89
VPD Fmin	VPD Fourier minimum	6.52	0.70
VPD Frange	VPD Fourier range	7.61	1.10
Rainfall			
CCD Fmean	Cold cloud duration (CCD) Fourier mean	115.00	29.00
CCD phs 1	CCD phase 1	6.90	2.78
CCD amp 1	CCD amplitude 1	288.00	119.00
CCD phs 2	CCD phase 2	5.28	0.50
CCD amp 2	CCD amplitude 2	278.00	3.00
CCD phs 3	CCD phase 3	4.00	0.00
CCD amp 3	CCD amplitude 3	109.00	5.00
CCD Fmax	CCD Fourier maximum	551.49	215.13
CCD Fmin	CCD Fourier minimum	0.00	-399.2
CCD Frange	CCD Fourier range	855.00	348.16
Vegetation			
ND Fmean	Normalised Difference Vegetation Index (NDVI) Fourier mean	0.49	0.07
ND phs 1	NDVI phase 1	10.03	4.38
ND amp 1	NDVI amplitude 1	0.21	0.01
ND phs 2	NDVI phase 2	5.25	1.35
ND amp 2	NDVI amplitude 2	0.18	0.01
ND phs 3	NDVI phase 3	3.93	0.10
ND amp 3	NDVI amplitude 3	0.05	0.003
ND Fmax	NDVI Fourier maximum	0.64	0.12
ND Fmin	NDVI Fourier minimum	0.40	-0.04
ND Frange	NDVI Fourier range	0.47	0.10

a: values are the timing of the maxima of the bi-annual or tri-annual cycles.

5.2.3 Statistical procedure

A variety of statistical methods are available which may be applied to predicting species distributions with multivariate datasets, such as correspondence analysis (Ter Braak 1986, Hill 1991), projection pursuit, nearest neighbour and neural network analysis and are reviewed (Williams *et al.* 1992) with respect to determining tsetse distributions in Zimbabwe using climate and weather data. In this analysis, various forms and modifications of discriminant analysis (Swain and Davies 1978, Tom and Miller 1984) are favoured because the technique is easy to apply, provides biological insight into the nature of the limits to the distribution and abundance of vector species and in relative terms is not computationally intensive. The following paragraphs summarise the main features of discriminant analysis. This is outlined in considerable detail in Rogers *et al.* (1996) (see appendix II) where the example of describing the distribution of a vector is used to illustrate the application of the technique.

Discriminant analysis assumes a normal distribution of the meteorological satellite sensor predictor variables and a common within-group co-variance of these variables for each pixel defining vector presence and vector absence. The vector presence and absence groups are defined from reliable distribution maps that form the "training set" for the satellite sensor data layers. The means of these multi-variate distributions are referred to as centroids and are defined by mathematical vectors $[\bar{x}_n]$ where n is the number of dimensions or variables. The Mahalanobis distance, D^2, is the distance between two centroids, or between a sample point and a centroid, where;

$$D^2_{12} = (\overline{\mathrm{x}}_a - \overline{\mathrm{x}}_b)' \mathrm{C}_w^{-1} (\overline{\mathrm{x}}_a - \overline{\mathrm{x}}_b)$$

and D^2_{12} is the Mahalanobis distances between group a (vector presence) and group b (vector absence), and C^{-1}_{w} is the inverse of the within-groups covariance (or dispersion) matrix (Green 1978, Tatsuoka 1971) (in this equation the subscript n for the number of variables has been dropped for clarity). The equation is further modified to allow for different within group co-variance matrices (or a non-linear discriminant plane between the presence and absence groups) since animals live in specific set of environmental conditions so that the co-variances of the variables within a distributional range (the presence group) are often different from those of the same variables outside the distributional limits (the absence group). To determine if a vector is present or absent from a sample point temporally or spatially removed from the training set, the values of

D^2 between the sample point and the centroids for the presence and absence groups are calculated. The point can then be assigned to a group by simply including it in the group which gives the smallest D^2 or assigning it a probability of group membership on the basis of posterior probabilities (Green 1978).

In the present analyses the tsetse distribution data were used as the training sets for species and the predictor satellite sensor variables selected in a forward, step-wise manner, with the criterion for inclusion being that the addition of the selected variable caused the greatest increase in the Mahalanobis distance compared with all other variables during that step of the analysis. 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. The use of Mahalanobis distances for variable selection overcomes the potential problem of unequal co-variances arbitrarily determining the importance of the predictor variables.

A total of 10 variables (out of the 51 available) were selected and used to produce maps of posterior probabilities which represent the probabilities with which each grid square falls into the category of fly presence or absence. These predicted maps cover the region from Côte d'Ivoire in the West to Togo in the East and are based on sampling the satellite sensor and other data files at a spatial resolution of 0.125 degrees. To make the biological interpretation of the results more straightforward no transformation of the meteorological satellite sensor variables was undertaken. The ability of the technique to describe the observed distribution and abundance data was measured in several ways. The overall percentage correct predictions 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 as these terms are common in the veterinary medicine literature.

The tsetse abundance data were grouped into "bins", with each bin defining a range of vector densities, and discriminant analysis similarly applied. Examples are given of binning the abundance data for the five species of tsetse in northern Côte d'Ivoire into three or five abundance classes with approximately equal sample sizes (using the rule that no abundance level appeared in more than one class). In the case of the abundance data the percentage correct assignment to each density class was recorded.

5.3 Results

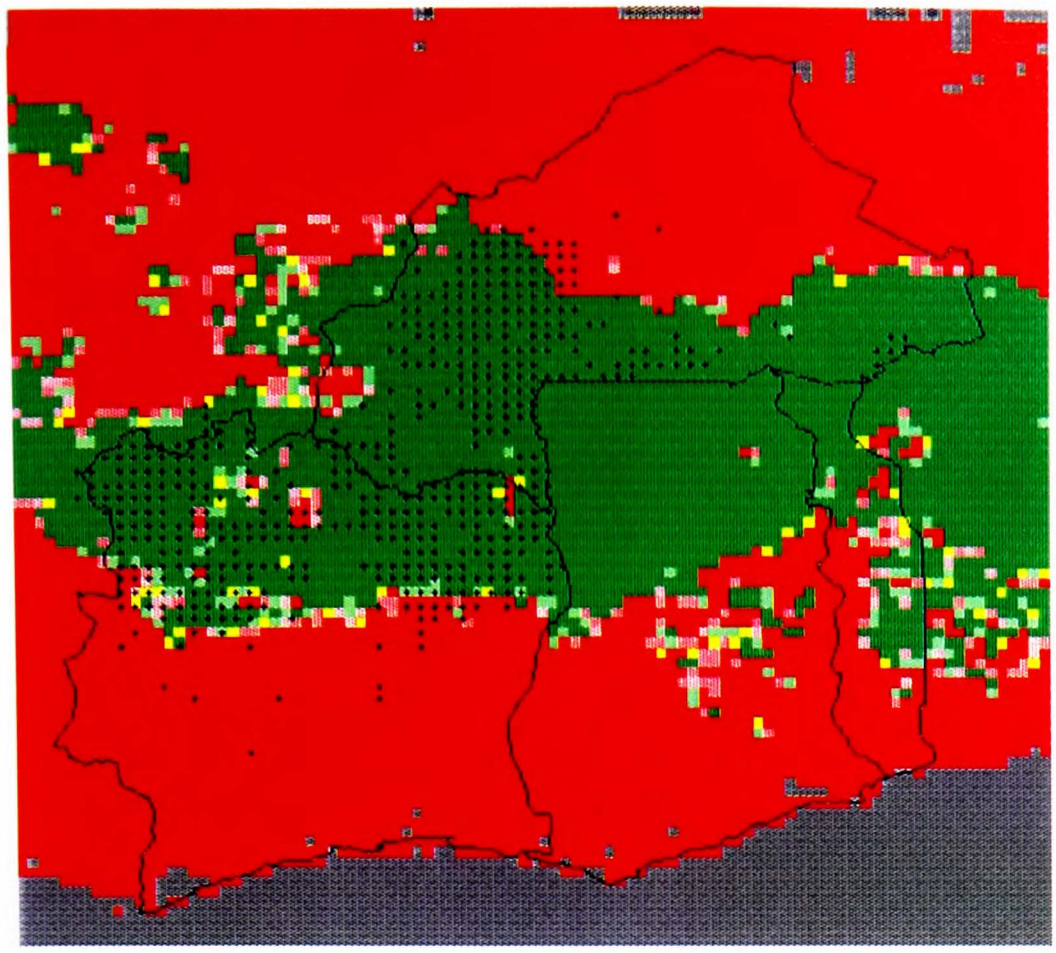
5.3.1 Tsetse distribution predictions

Figures 5.1.1 - 5.1.8 show the observed and predicted distributions for each of the eight species surveyed in Côte d'Ivoire and Burkina Faso and the accuracy achieved for each using the ten most important predictor variables. The recorded distribution of each species is indicated and suitability shown on a colour scale from red (low suitability) to green (high suitability). The term "no prediction" refers to areas where satellite sensor data were more extreme than any in the training set. The accuracies of the predictions (shown beside each figure) refer to the training set data which is not spatially coincident with the gridded-square map output. Table 5.2 lists the predictor variables in order of importance for each species and lists the accuracy of predictions when one, five or all ten predictor variables were used. Of the data derived from satellite sensors used in the present analysis, the MIR channel 3 data appear most frequently (30 % of all variables) in the predictor variables for fly distribution followed by the CCD (22 %), TIR (18 %), VPD (13 %), NDVI (11 %) and elevation (6 %). On 14 occasions, one or other MIR variables is in the top five predictor variables for the eight species of tsetse considered, whilst CCD, TIR, VPD, NDVI and DEM variables appear 8, 6, 5, 3 and 3 times respectively.

5.3.2 Tsetse abundance predictions

Figures 5.2.1 - 5.2.5 show the observed and predicted abundance classes for the five species sampled in the north of Côte d'Ivoire and table 5.3 lists the predictor variables used and the accuracy of the predictions. The density of each species is indicated on a grey-scale where black is high density and light grey is low density. Given the relatively small ranges of density in each class the results appear good. Considering the top five variables for predicting the distribution of these five species the number of occurrences in the rankings were 8, 7, 6, 3 and 0 respectively, for the CCD, TIR, MIR, VPD and elevation data. Table 5.4 lists the mean values of the predictor variables for the five density classes for *G. palpalis,* the most widespread and well sampled of the species. In many cases there is a gradual increase or decrease in the mean values of the predictor variables across the density classes and only very slight differences between the predictor variables are found in areas of absence or presence at the different densities.

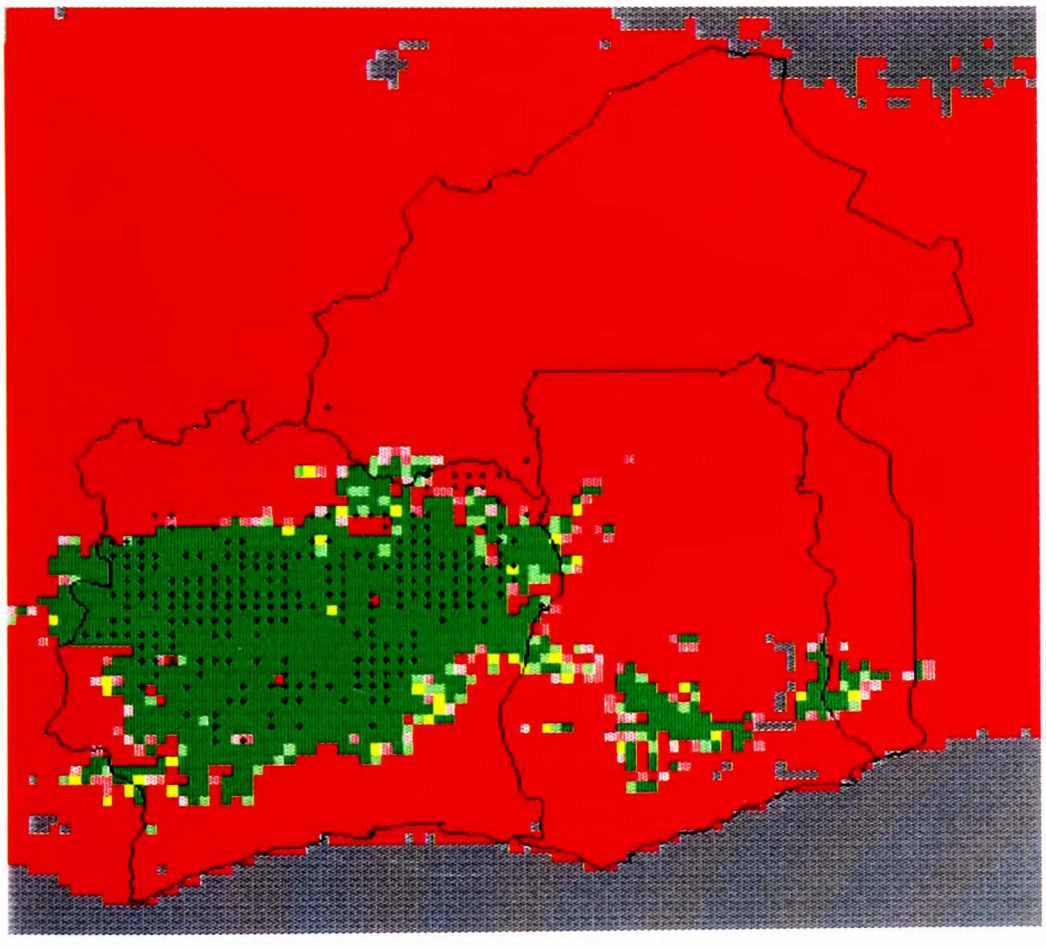
5.1.1. *Glossina morsitans.*



% correct = 70
% false positives = 25
% false negatives = 5
Sensitivity = 0.89
Specificity = 0.57

N

5.1.2. *Glossina longipalpis.*

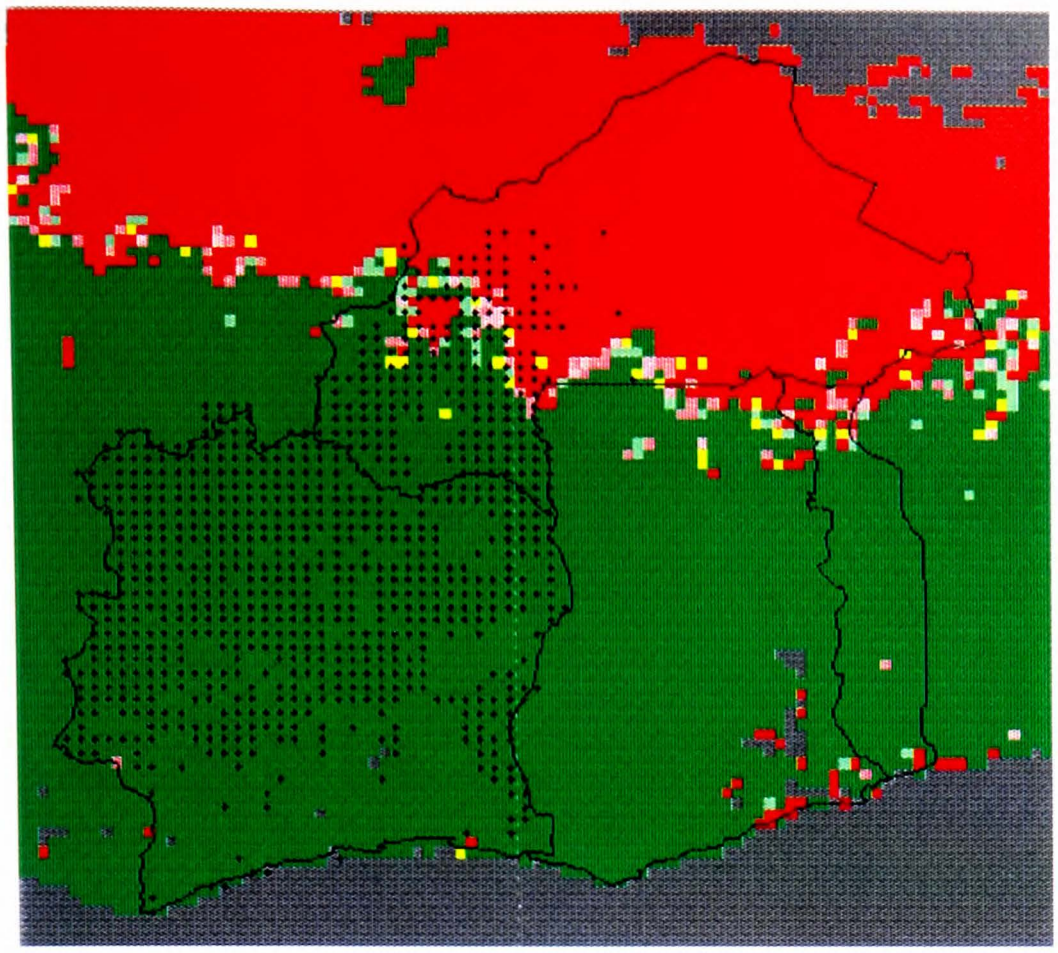


% correct = 74
% false positives = 25
% false negatives = 2
Sensitivity = 0.93
Specificity = 0.69

= 0.65 - 1.000
= 0.55 - 0.649
= 0.50 - 0.549
= 0.45 - 0.499
= 0.35 - 0.449
= 0.00 - 0.349
= no prediction
+ = species present

Figure 5.1.1 - 5.1.2. Discriminant analysis of predictions of areas of suitability for *Glossina morsitans* (5.1.1), *G. longipalpis* (5.1.2) in Côte d'Ivoire and Burkina Faso based on the satellite predictor variables listed in table 5.2.

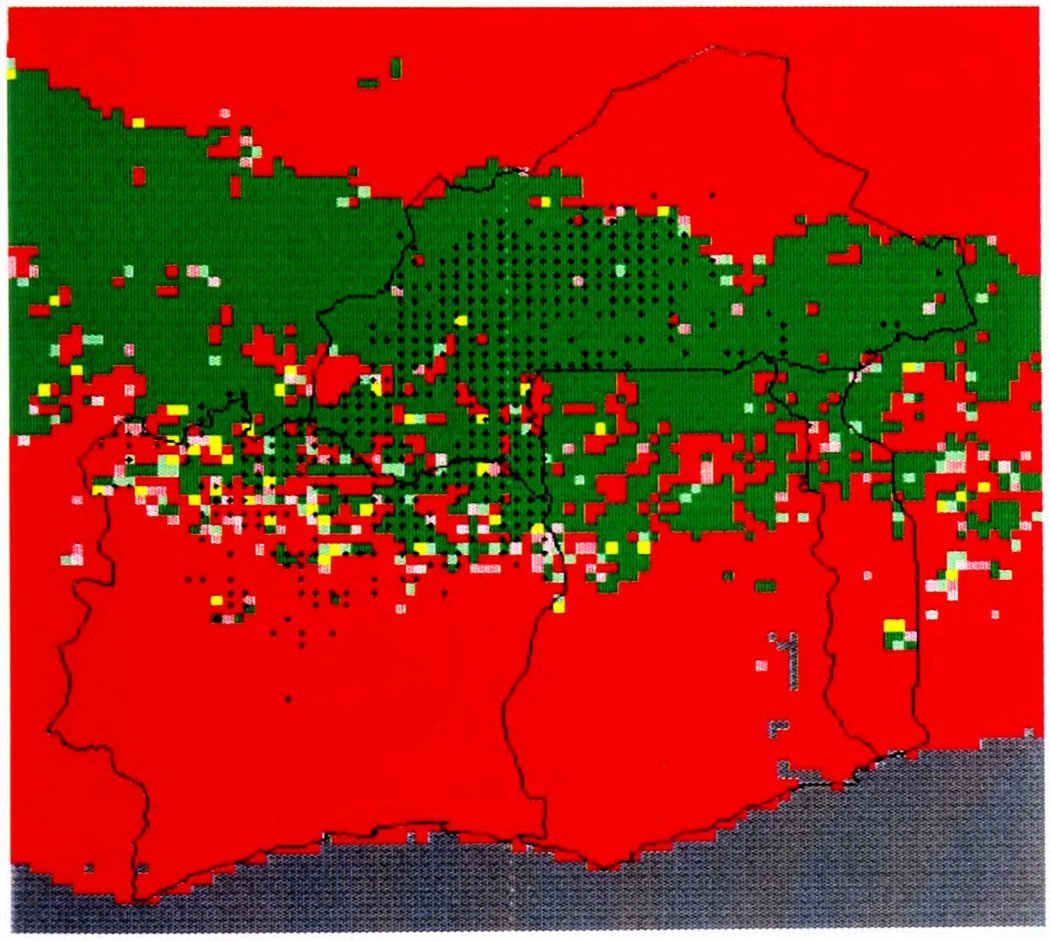
5.1.3. *Glossina palpalis.*



% correct = 90
% false positives = 6
% false negatives = 4
Sensitivity = 0.95
Specificity = 0.64

N

5.1.4. *Glossina tachinoides.*

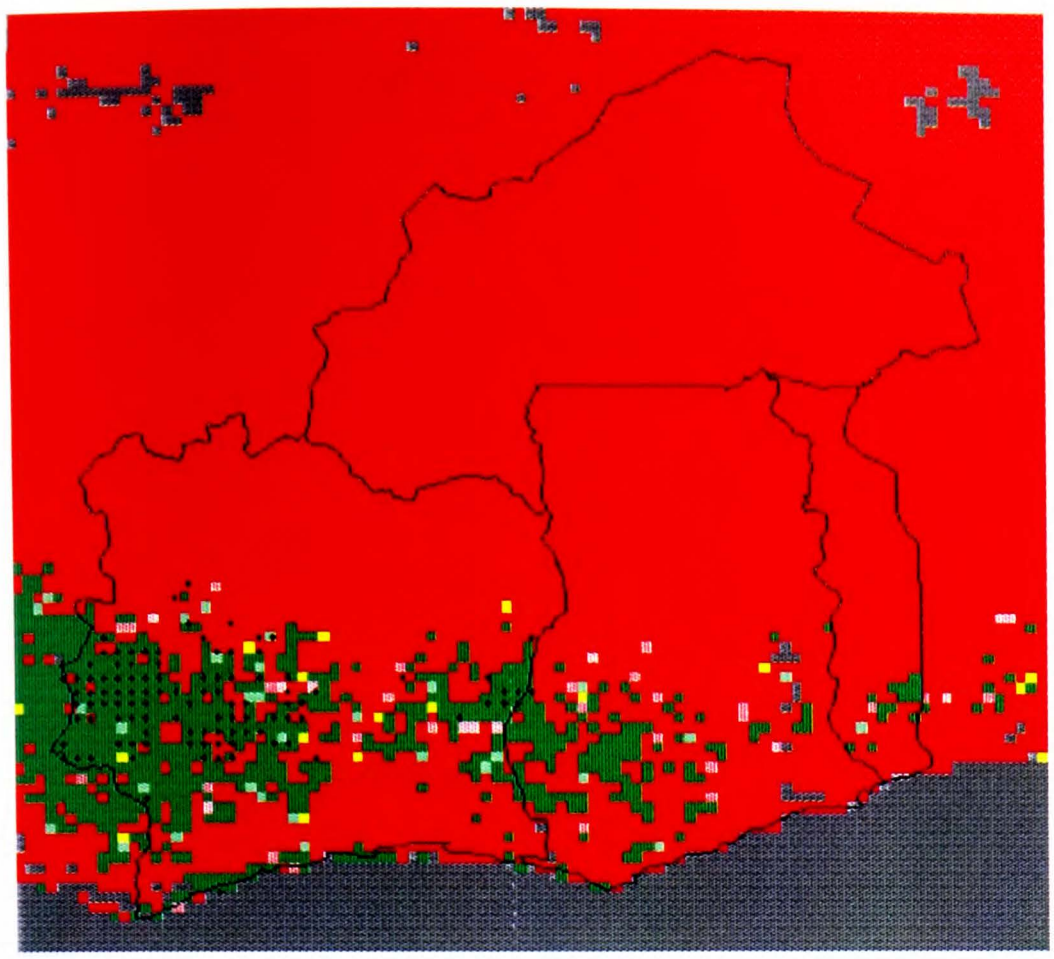


% correct = 78
% false positives = 13
% false negatives = 9
Sensitivity = 0.75
Specificity = 0.80

= 0.65 - 1.000
= 0.55 - 0.649
= 0.50 - 0.549
= 0.45 - 0.499
= 0.35 - 0.449
= 0.00 - 0.349
= no prediction
= species present

Figure 5.1.3 - 5.1.4. Discriminant analysis of predictions of areas of suitability for *Glossina palpalis* (5.1.3), *G. tachinoides* (5.1.4) in Côte d'Ivoire and Burkina Faso based on the satellite predictor variables listed in table 5.2.

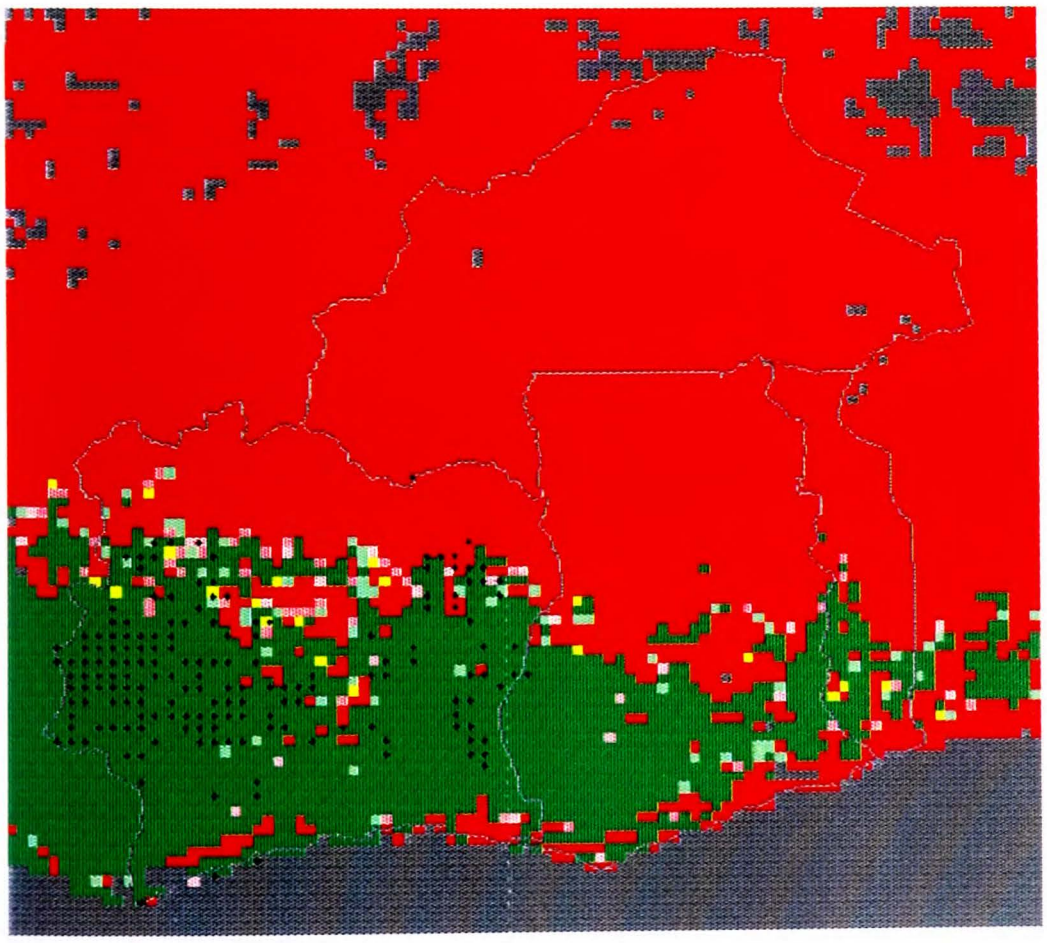
5.1.5. *Glossina pallicera.*



% correct = 92
% false positives = 6
% false negatives = 1
Sensitivity = 0.81
Specificity = 0.93

N

5.1.6. *Glossina fusca.*

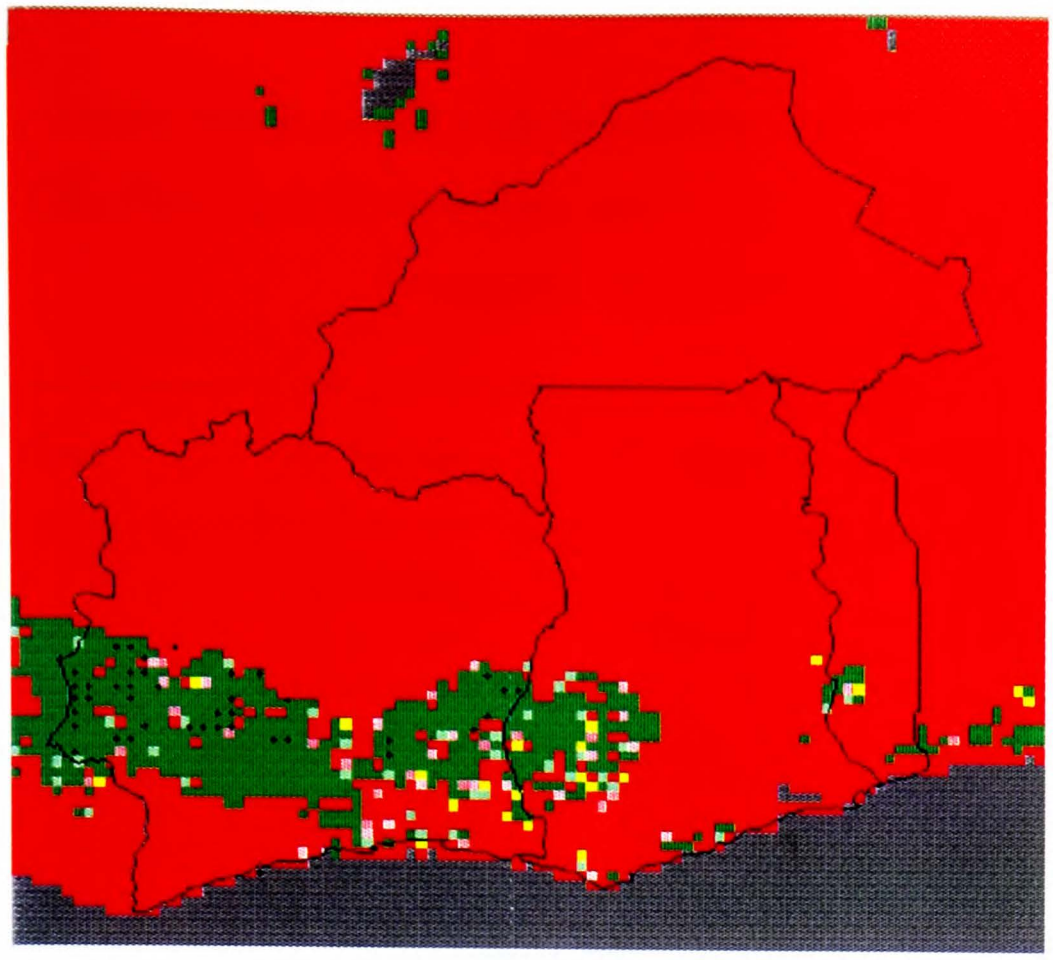


% correct = 81
% false positives = 18
% false negatives = 1
Sensitivity = 0.91
Specifitity = 0.73

= 0.65 - 1.000
= 0.55 - 0.649
= 0.50 - 0.549
= 0.45 - 0.499
= 0.35 - 0.449
= 0.00 - 0.349
= no prediction
= species present

Figure 5.1.5 - 5.1.6. Discriminant analysis of predictions of areas of suitability for *Glossina pallicera* (5.1.5), *G. fusca* (5.1.6) in Côte d'Ivoire and Burkina Faso based on the satellite predictor variables listed in table 5.2.

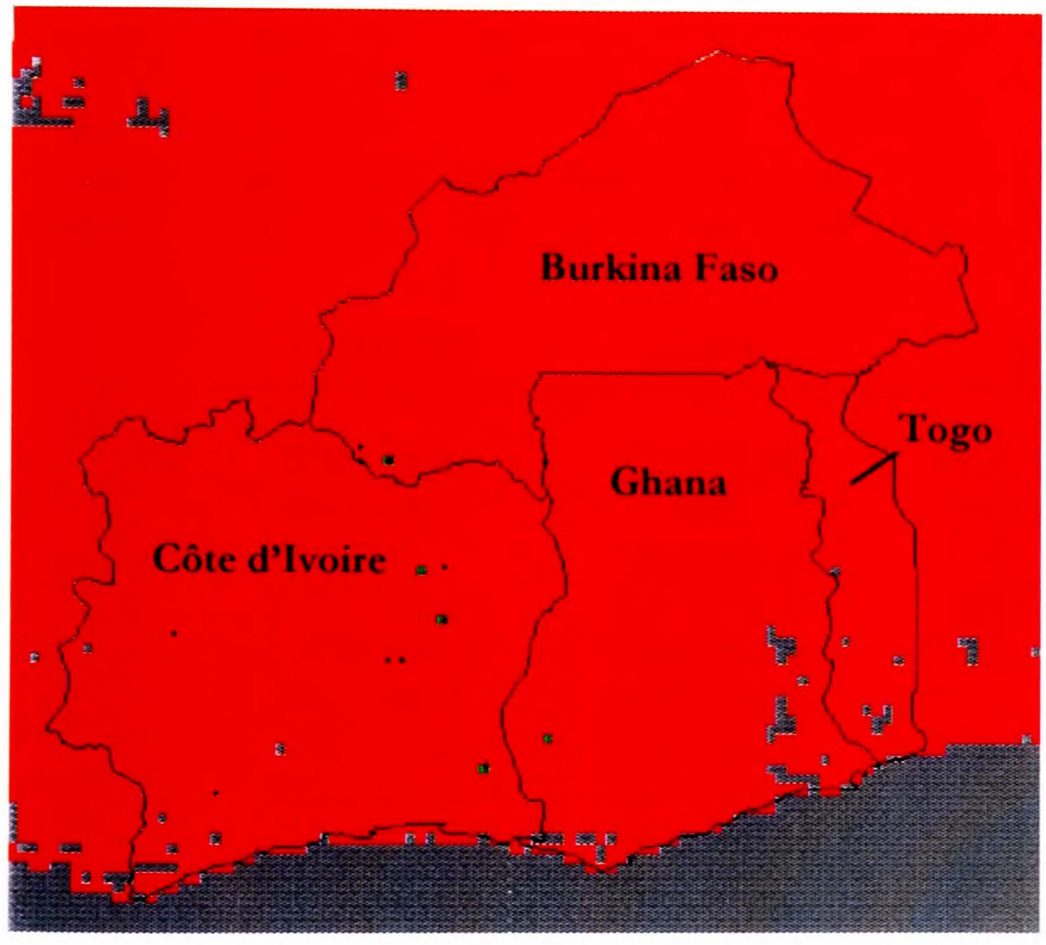
5.1.7. *Glossina nigrofusca.*



% correct = 92
% false positives = 7
% false negatives = 0
Sensitivity = 0.93
Specificity = 0.93

N

5.1.8. *Glossina medicorum.*



% correct = 100
% false positives = 0
% false negatives = 0
Sensitivity = 0.85
Specificty = 1.00

= 0.65 - 1.000
= 0.55 - 0.649
= 0.50 - 0.549
= 0.45 - 0.499
= 0.35 - 0.449
= 0.00 - 0.349
= no prediction
= species present

Figure 5.1.7 - 5.1.8. Discriminant analysis of predictions of areas of suitability for *Glossina nigrofusca* (5.1.7), *G. medicorum* (5.1.8) in Côte d'Ivoire and Burkina Faso based on the satellite predictor variables listed in table 5.2.

TABLE 5.2.

The ten most important predictor variables used to describe the distribution of tsetse in Côte d'Ivoire and Burkina Faso, and the accuracy of predictions when using n = 1, 5 or 10 variables. See table 5.1 for the full names.

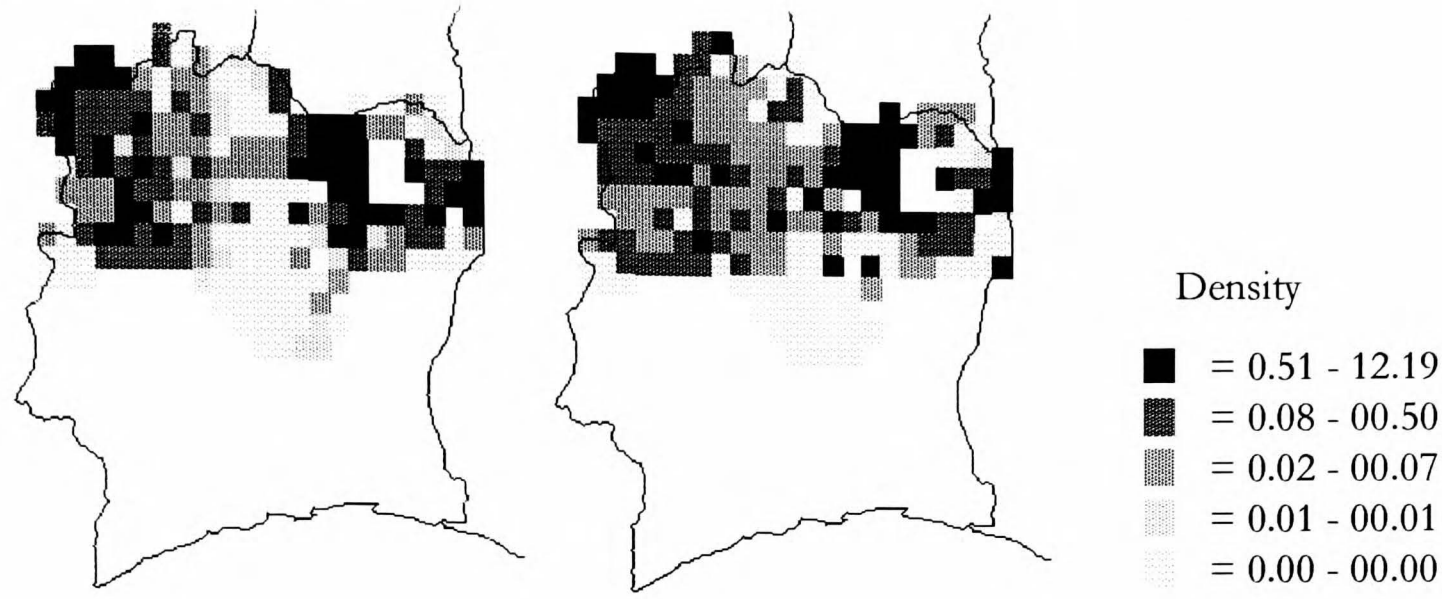
Rank	*G. morsitans*	*G. longipalpis*	*G. palpalis*	*G. tachinoides*
1	Ch 3 Frange	Ch 3 Fmin	Price phs 1	Ch 3 Fmax
2	Ch 3 amp 2	CCD phs 2	Ch 3 phs 1	VPD phs 2
3	CCD phs 1	CCD amp 3	ND amp 1	Ch 3 phs 2
4	CCD amp 2	Ch 3 amp 2	CCD amp 1	Price Frange
5	Ch 3 amp 1	CCD amp 1	Ch 3 amp 2	CCD amp 1
6	VPD Fmin	Ch 3 amp 3	ND Fmin	ND phs 2
7	Ch 3 phs 2	Price phs 1	Price amp 1	CCD phs 3
8	CCD Fmax	CCD Fmax	VPD phs 1	Ch 3 Fmean
9	CCD amp 1	ND amp2	ND Fmean	Ch 3 amp 1
10	ND phs 1	Elev	Price phs 2	ND Fmean

	1	*5*	*10*	*1*	*5*	*10*	*1*	*5*	*10*	*1*	*5*	*10*
% correct	65	70	**70**	64	70	**74**	88	90	**90**	77	78	**78**
% false +	29	23	**25**	34	28	**25**	9	7	**6**	15	17	**13**
% false -	6	7	**5**	2	2	**2**	3	3	**4**	8	6	**9**
Sensitivity	0.86	0.85	**0.89**	0.89	0.91	**0.93**	0.97	0.96	**0.95**	0.78	0.85	**0.75**
Specificity	0.49	0.59	**0.57**	0.57	0.64	**0.69**	0.50	0.60	**0.64**	0.76	0.74	**0.79**

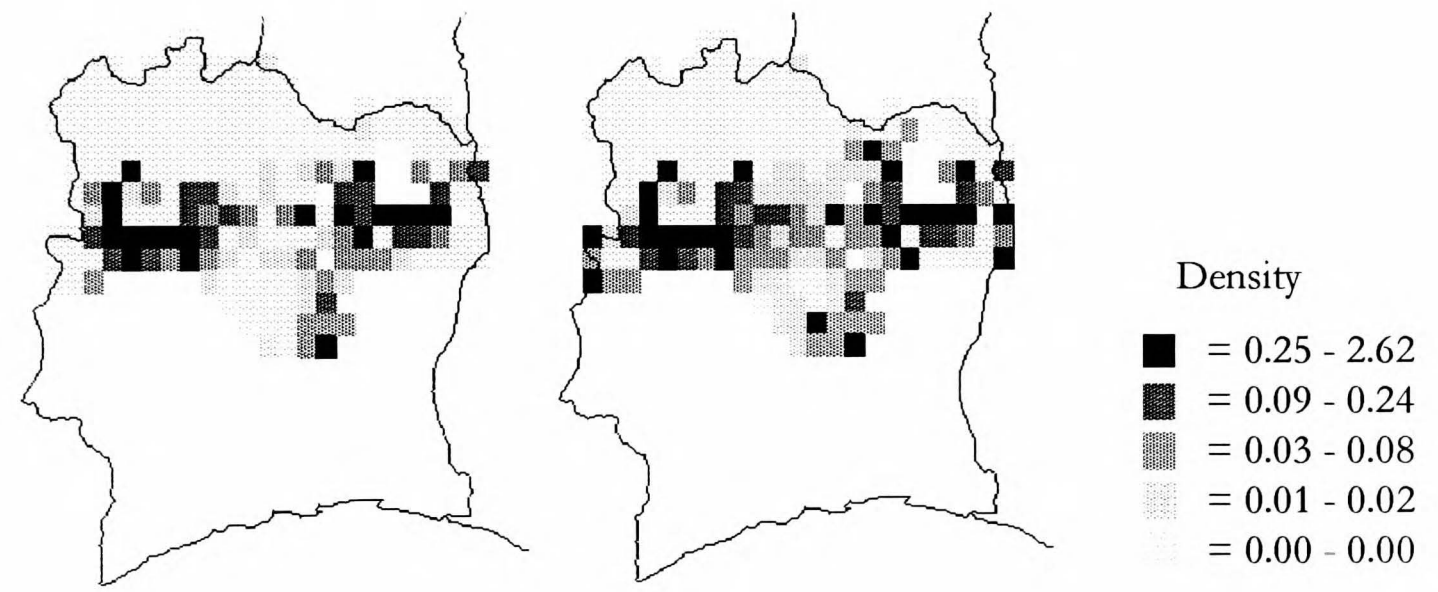
	G. pallicera	*G. fusca*	*G. nigrofusca*	*G. medicorum*
1	VPD Fmax	Ch 3 Fmin	Price Frange	Elev
2	Ch 3 amp 1	Elev	Price amp 1	Ch 3 amp 1
3	Elev	Ch 3 amp 2	CCD phs 1	Price amp 2
4	VPD Fmean	Ch 3 phs 2	VPD phs 3	ND amp 1
5	ND phs 3	Price Fmin	VPD Fmean	VPD amp 3
6	Price phs 2	CCD amp 3	CCD amp 1	Ch 3 amp 2
7	VPD Fmin	CCD phs 2	Ch 3 phs 1	Price phs 2
8	Price Fmax	Ch 3 amp 3	CCD Fmax	Ch 3 Fmin
9	CCD amp 2	Ch 3 amp 1	Elev	VPD phs 2
10	Price amp 1	Price Fmax	Ch 3 phs 2	CCD amp 2

	1	*5*	*10*	*1*	*5*	*10*	*1*	*5*	*10*	*1*	*5*	*10*
% correct	93	92	**92**	69	78	**81**	97	93	**92**	99	99	**100**
% false +	0	7	**6**	27	20	**18**	0	7	**7**	0	0	**0**
% false -	7	1	**1**	4	2	**1**	3	0	**0**	1	1	**0**
Sensitivity	0.00	0.85	**0.81**	0.69	0.86	**0.91**	0.00	0.86	**0.93**	0.00	0.15	**0.85**
Specificity	1.00	0.92	**0.93**	0.69	0.77	**0.73**	1.00	0.93	**0.93**	1.00	1.00	**1.00**

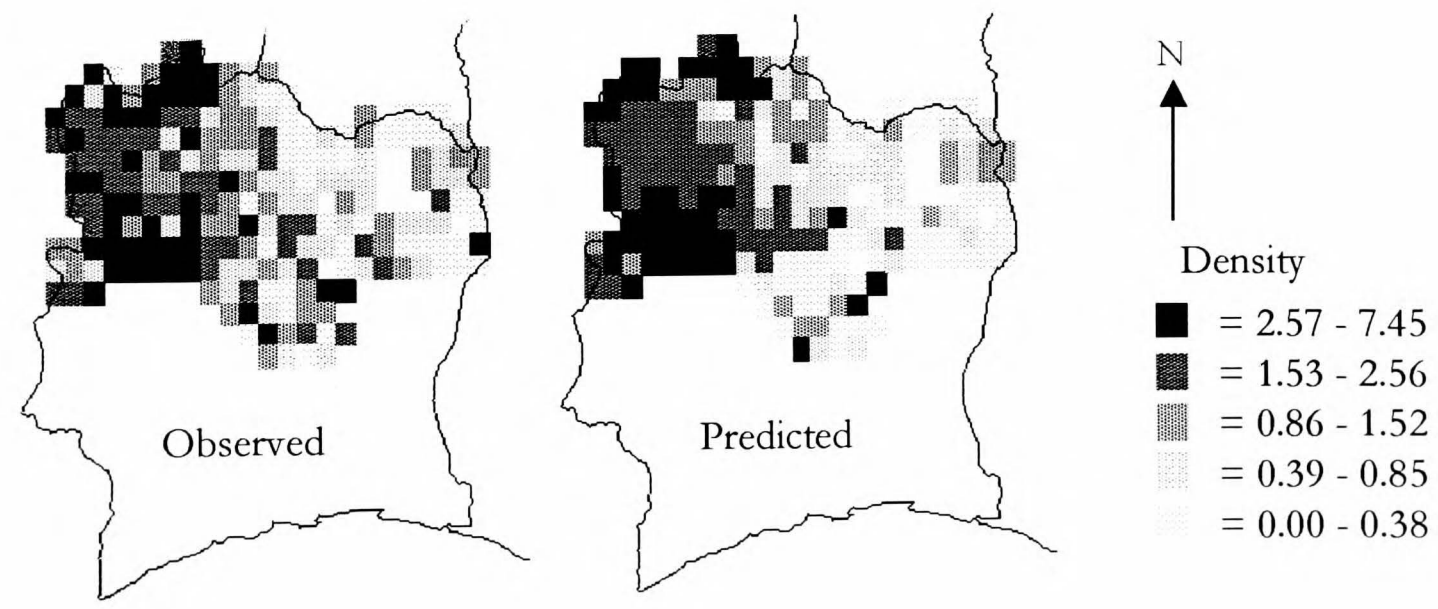
5.2.1. *Glossina morsitans.*



5.2.2. *Glossina longipalpis.*



5.2.3. *Glossina palpalis.*



5.2.4. *Glossina tachinoides.*

5.2.5. *Glossina fusca.*

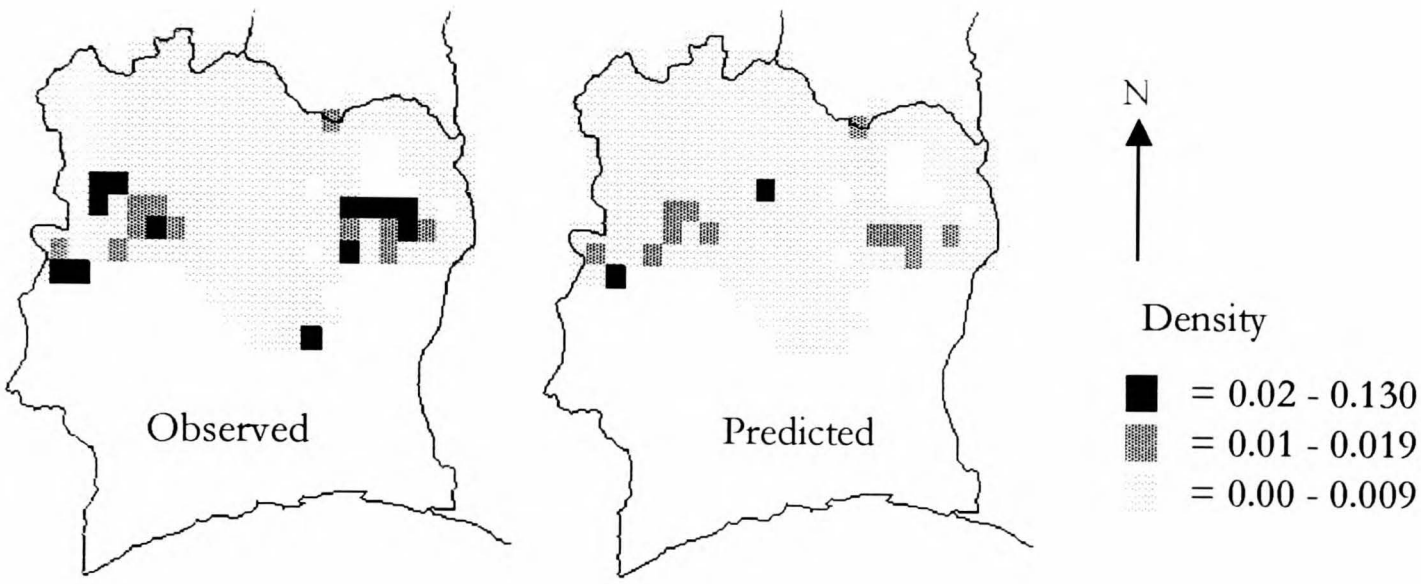


Figure 5.2.1 - 5.2.5. Discriminant analysis of predictions of the apparent density (flies per trap per day) of *Glossina morsitans* (5.2.1), *G. longipalpis* (5.2.2) *G. palpalis* (5.2.3) *G. tachinoides* (5.2.4) and *G. fusca* (5.2.5) in northern Côte d'Ivoire, based on the satellite predictor variables listed in Table 5.3. Fly apparent densities were divided into five or three (*G. fusca*) classes by the analysis to give approximately equal sample sizes.

TABLE 5.3.

The ten most important predictor variables used to describe the apparent density 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 except G. fusca*). See table 5.1 for the full names.*

Rank	*G. morsitans*	*G. longipalpis*	*G. palpalis*	*G. tachinoides*	*G. fusca*
1	Ch 3 Frange	Ch 3 phs 2	VPD phs 3	Price Fmax	Price Fmin
2	Price amp 1	Price amp 2	Ch 3 amp 2	Price amp 2	VPD amp 2
3	Ch3 amp 1	CCD amp 1	CCD phs 1	CCD amp 3	CCD phs 2
4	CCD phs 2	CCD amp 2	Price amp 1	Ch3 amp 3	ND Fmax
5	CCD amp 1	Price Frange	VPD amp 2	CCD Fmin	Ch3 phs 2
6	Ch 3 mean	Ch3 amp 1	Price amp 3	CCD phs 1	Ch 3 Fmin
7	CCD Fmax	CCD phs 3	CCD amp 3	Elev	CCD amp 3
8	VPD Fmin	CCD phs 2	Ch 3 Fmin	Ch 3 phs 1	CCD amp 2
9	Ch 3 Fmax	VPD Fmin	Ch3 amp 1	Ch 3 phs 3	ND amp 2
10	Ch3 phs 1	Price phs 3	CCD Fmax	VPD amp 1	ND Frange
	Density %	*Density* %	*Density* %	*Density* %	*Density* %
	0.00 0.00 60	0.00 - 0.00 79	0.00 - 0.38 68	0.00 0.00 83	0.00 0.00 100
	0.01 - 0.01 6	0.01 - 0.02 71	0.39 - 0.85 50	0.01 0.02 67	0.01 0.01 100
	0.02 - 0.07 70	0.03 - 0.08 71	0.86 1.49 41	0.03 0.09 65	0.02 - 0.13 8
	0.08 0.50 61	0.09 - 0.20 94	1.53 2.53 59	0.10 - 0.30 29	
	0.51 - 12.2 73	0.25 2.62 100	2.57 7.45 64	0.31 - 8.53 70	
Mean	41	83	56	63	69

TABLE 5.4.

Mean values of the predictor variables for each of the apparent density classes of G. morsitans. *See table 5.1. for the full names.*

Density	*Predictor variables*					
	VPD phs 3	*Ch 3 amp 2*	*CCD phs 1*	*Price amp 1*	*VPD amp 2*	*n*
0.00 - 0.00	1.42	1.10	5.46	7.69	0.52	41
0.01 0.01	1.45	1.19	5.44	7.63	0.51	42
0.02- 0.07	1.43	1.22	5.48	7.34	0.50	43
0.08 - 0.50	1.69	1.54	5.62	6.66	0.48	41
0.51 - 12.18	1.97	1.31	5.57	6.74	0.45	45
	Price amp 2	*CCD amp 3*	*Ch 3 Fmin*	*Ch 3 amp 2*	*CCD Fmax*	*n*
0.00 0.38	1.67	49.02	302.17	8.02	315.46	*41*
0.39 - 0.85	1.80	52.45	301.88	8.02	321.11	42
0.86 - 1.49	1.70	56.02	302.06	7.88	321.50	43
1.53 2.53	1.77	56.15	301.55	7.71	313.90	41
2.57 7.45	1.85	56.60	301.73	7.75	335.30	42

5.4 Discussion

5.4.1 Tsetse distribution

For the eight tsetse species investigated, the overall percentage of correct predictions was 85 % (with minimum of 70 % for *G. morsitans* and maximum of 100 % for *G. medicorum*). These results improved by an average of 2.4 percentage points on those found in Rogers *et al.* (1996), suggesting that the wider range of meteorological satellite sensor variables were of value in predicting tsetse fly distribution. Furthermore, the high levels of accuracy were achieved with ecologically diverse tsetse species with both widespread and restricted ranges. It was found that adding more predictor variables increased the accuracy and definition of predictions (see table 5.2), so that with ten predictor variables the analysis identified areas of suitability with a relatively high or low probability resulting in intermediate values within the range 0.35 to 0.65 being scarce (see figures 5.1.1 - 5.1.8).

The percentage false positives always considerably outnumbered the percentage of false negatives. The former indicate areas of apparent suitability for flies that were not occupied, or in which flies were not recorded by the surveys. These are common features of distribution maps since a species does not always occupy the entire area that is suitable for it and are not always found even when they are present. False negative predictions present a more serious problem and indicate where the discriminant analysis has failed to define the full range of conditions in which the tsetse population can survive on the basis of the predicator variables. The percentage of false negatives is highest (9 %) for *G. tachinoides*, with the errors due to the technique failing to identify the southern limits of its recorded distribution in Côte d'Ivoire (figure 5.1.4). The false negative predictions found in this chapter are in the majority of cases marginally higher than those in Rogers *et al.* (1996).

5.4.2 Tsetse abundance

The predictions of tsetse density were more variable with 63 % correct predictions across all the species and classes (minimum 41 % for *G. morsitans*, and maximum 83 % for *G. longipalpis*). *G. morsitans* was again the most poorly predicted of the species and this was probably related to it being inadequately sampling by the blue biconical traps used in the survey (see section 5.2.1). These results show a slight decrease in the accuracy reported by Rogers *et al.* (1996), but the difference is exaggerated by two very

low predictions in density classes with particularly small numbers of observations (see table 5.3).

5.4.3 General considerations.

There was an even spread of variables chosen in the rankings with MIR, CCD, TIR and VPD all important in describing both distributions and abundance datasets. The MIR AVHRR channel 3 data (see section 4.1.2) predominated in describing the tsetse distribution and the CCD was the most important variable in describing tsetse abundance. The reduced importance of the MIR imagery in determining abundance and its relatively greater importance in predicting distribution suggests that fly distribution limits in this part of Africa may be more sensitive to temperature, whilst abundance within the distributional limits is some function of rainfall and temperature.

Elevation was much less important in the prediction of tsetse than land-cover distributions, constituting only 7 % of the top-five rankings. In contrast, the Fourier amplitude variables occurred most often (43 %) with phase and range variables joint second (25 %). Among the Fourier variables chosen, the annual cycles were most important (35 %), with the bi-annual cycles and range values equal second (27 %) and with the tri-annual cycles (11 %) also contributing as predictors. The amplitudes were again the most important in the top five variables of the abundance data (at 52 %) with 24 % for the phase and range variables respectively. Slightly different trends were exhibited with the Fourier harmonics chosen, however, with bi-annual cycles (40 %) dominating, followed by both the annual cycle and range variables (24 %) and the tri-annual cycle (12 %). Interestingly therefore, the amplitude of a given variable throughout the year would seem to be the most important factor in determining species ranges and the amplitude of its bi-annual signal in determining population abundances within these limits.

The success of the approach is remarkable when the relatively poor quality of the distribution data and the long periods of time over which they were recorded are considered. The satellite sensor data were gathered during the 1980s, when particularly severe droughts affected much of the study area, resulting in rapid changes of the distributional limits of some of the species considered. *G. tachinoides*, for example, extended its range South, and apparently replaced *G. palpalis* over large areas of central Côte d'Ivoire (Clair 1987). Curiously, and perhaps significantly, when each dataset is

kept separate within the analysis, the predicted area of suitability for *G. tachinoides* shifts southwards into the same central and southern regions of Côte d'Ivoire that the species invaded in the 1980's. These fly advances may be facilitated by only slight changes in the average environmental conditions in the newly invaded areas. For example, the present analysis suggests that the difference between the mean NDVIs in sites of presence or absence of *G. tachinoides* varies between - 0.08 and + 0.04 (in Côte d'Ivoire and Burkina Faso respectively), each a fraction of the total range shown by this variable across these two countries (0.47, table 5.1). As some areas become more suitable for flies, however, other areas become less suitable, so that the distributional limits shift slowly with time. The expansion and subsequent contraction of tsetse fly distributions has been previously recorded, but it is evident that very small changes in global climate change will bring about significant range shifts in addition to these short-term variations (Rogers and Packer 1993).

Whilst the ability to describe the training set data is the first criterion for a successful statistical description of a species distribution, the technique is only of real use when it can be used to describe distributions in other places, and at other times. The maps presented in figures 5.1.1 5.1.8 show predicted distributions of the study species in Ghana and Togo. The predictions for Togo show both similarities and differences with the recently mapped tsetse distributions in this country (Rogers *et al.* 1994, Hendrickx *et al.* 1995). Predictions for *G. palpalis* and *G. tachinoides* are rather better than those for *G. morsitans* and *G. longipalpis*, although there has also been a southward extension of *G. tachinoides* in Togo that is not predicted by the present analysis, and does not appear on previous maps for tsetse in Togo (Ford and Katondo 1977). Clearly, therefore, the distribution of tsetse in Togo, and presumably elsewhere, has changed from the historical picture that forms the basis of much of the present analysis and, in the light of recent environmental changes in Africa, it is unlikely that non-contemporary satellite sensor data will give an entirely satisfactory fit. The ideal approach to tsetse mapping is obviously to use contemporary satellite sensor and fly distribution data to define the areas of suitability for each tsetse species and, from this, to make predictions for other places and times.

A further complication arises from the adaptation of each species to local conditions, about which very little is known at present. Species such as *G. morsitans*, and even its subspecies, occupy vast areas of Africa (*e.g.* greater than 40° of longitude for *G.*

m. submorsitans) and almost certainly show biological variation across such distances. Behavioural and ecological differences within *G. pallidipes*, a widespread species in East and southern Africa, have already been shown (Rogers 1990, Baylis and Nambiro 1993). These effects mean that the characterisation of a species' habitat in one area may not easily be extended to other, remote areas. Given contemporary satellite sensor and distributional data, however, the approach suggested may be used to estimate the degree of difference in habitat types across wide geographical areas. Whilst satellite sensor data may not immediately explain within-species differences across large geographical areas, they may, in the first instance, illuminate such differences and hence lead to a better biological understanding of them.

5.5 Conclusion and summary

It is clear that remotely sensed satellite sensor imagery can be a powerful tool in our ability to investigate large area phenomena such as the distribution and abundance of insect vectors of disease. The application of temporal Fourier processing to multi-temporal meteorological satellite sensor imagery allows characterisation of habitat "fingerprints" in the form of means, the seasonal timing and seasonal extremes of values of temperature, atmospheric moisture, rainfall and vegetation surrogates. Insect vector distributions clearly depend on habitat types and so should be amenable to statistical descriptions based on such satellite habitat fingerprinting. In future, higher spatial and spectral resolution satellite sensors will provide a much more "fine-grained" view of natural habitats (Hay *et al.* 1996b) and it is timely to prepare now for the wealth of new data these satellite sensors will provide. Furthermore, the rapid advances in computing technology and development of novel image processing and analytical techniques will contribute to an increasing ability to predict the distribution and abundance of natural resources and disease vectors using remotely sensed satellite imagery.

A statistical description of a tsetse fly habitat or distribution is, however, no substitute for a full biologically based understanding of the same phenomenon. Such an understanding comes from a study of the underlying demographic processes but, as explained elsewhere (Rogers and Randolph 1993, Randolph 1994), the information on such processes is often lacking, leaving the statistical approach as the only one available at present. The long-term aim of the present work is to produce risk maps for tsetse-borne diseases based on a sound biological understanding of epidemiological processes. Satellite imagery provides a way of revealing the patterns in epidemiological processes from which such an understanding will eventually arise.

Chapter 6

" ... Science is built up of facts, as a house is built up of stones; but an accumulation of facts is no more science than a heap of stones a house ... "

Henri Poincaré.

CONCLUSIONS AND GENERAL DISCUSSION

In this last chapter the work presented in this thesis is discussed in the wider epidemiological context. The results are first summarised and possibilities for future research highlighted. The potential obstacles to the use of this work by epidemiologists in developing countries are considered with some perspectives on how these technologies might evolve with the satellite borne sensors of the future.

6.1 Remote sensing of tsetse fly distribution and abundance

Data from the AVHRRs on-board the NOAA polar-orbiting meteorological satellites and from the HRRs on-board the Meteosat geostationary meteorological satellites have been shown to be of use in predicting the distribution and abundance of the tsetse fly (Diptera: *Glossinidae*). This was achieved by first investigating which methods of deriving monthly land surface temperature, atmospheric moisture and rainfall indices (for the period 1988 to 1990) were most accurate, using a test dataset of meteorological records collected for the same years across continental Africa. These indices were found to provide relatively accurate surrogates for point meteorological records and explained on average 54 % (range 19 - 74 %) of the variance in monthly land ground temperature records, 63 % (range 40 - 81 %) of the variance in monthly VPD records and 53 % (range 34 84 %) in monthly rainfall totals. The indices that described the meteorological station data most accurately formed a refined dataset of surrogate meteorological information at 8 x 8 km spatial resolution that was used for the subsequent analysis. This proxy meteorological data and a range of SVIs were then subject to temporal Fourier analysis to parameterise the seasonal variation in these

variables as the average annual mean, range and timing of maximum value for the duration of the analysis. The advantages of temporal Fourier analysis over canonical principle components analysis approach were outlined. These Fourier data, in combination with elevation information from a DEM, were then used in a discriminant analysis to predict the land-cover of Nigeria. The mean level of correct prediction across eight land-cover types was 73 % (range 42 100 %). The NDVI was found to be the best predictor of the SVIs and was hence used in combination with the proxy meteorological and DEM data to predict the distribution of eight and abundance of five tsetse fly species in Côte d'Ivoire and Burkina Faso. This was achieved to an overall accuracy of 85 % correct predictions for species distributions (range 70 - 100 %) and 63 % correct predictions for species abundances (range 41 83 %). These results were considered particularly good when the inherent problems of remotely sensed (see section 1.3) and the species distribution (see section 5.2.1) data were considered along with their non-contemporary acquisition.

There is obvious potential for the wider application of these techniques and datasets to other vector borne diseases, as all terrestrial invertebrates are dependent to some degree on climate. The application of these techniques to large homeothermic animals is less straightforward however, but identical techniques and similar datasets have been used to predict, with accuracies comparable to the tsetse distribution work, large ungulate distributions in East Africa (Anderson 1996).

6.2 Remote sensing and epidemiology in developing regions

The perception that remote sensing is not appropriate in technologically developing regions is powerful and manifest in the form of frequent objections to the cost of image processing equipment, expertise and ground truth (Bos 1990, Arambulo and Astudillo 1991, Barinaga 1993, Kleiner 1995). These problems are real concerns but are diminishing as, (i) computer processing and data storage facilities become relatively cheaper, (ii) changes in the philosophy of satellite sensor image dissemination have resulted in meteorological satellite sensor data becoming more widely and freely available (Mulcahy and Clarke 1994, Justice *et al.* 1995) and (iii) the history of successful application of remote sensing techniques in epidemiology expands. The application of remote sensing has also been shown to be cost effective in mosquito control campaigns (Fleetwood *et al.* 1981, Welch *et al.* 1989a) and to provide baseline datasets which can be subsequently expanded and find application far beyond the motivation for their original

assembly (Clarke *et al.* 1991). There is no reason to assume this should not be the case in the developing tropics and is of critical importance in the evaluation of vector control strategies where resources are often acutely limited (Mills 1994). Furthermore, these predictions have already proved useful in Ghana where no recent tsetse distribution data are available (Doku 1995).

For remotely sensed data to become routinely used in disease vector monitoring and as an integral part of subsequent management procedures, it is important to know for how long we can expect to receive such imagery, to what extent the user community will have access and what improvements might reasonably be expected in the near future. These issues are addressed in the next section. A full exploration of the future potential in epidemiology of a range of new satellite borne sensors scheduled for orbit by the Millennium is given in Hay (1997).

6.3 Future satellite sensors

Continuity of data from the NOAA AVHRR polar orbiting satellites is guaranteed in the short term by the successful launch of NOAA -14 in December 1994 and in the longer term, as plans have been published for a further six NOAA satellites in collaboration with the U.S. Department of Defence (Power 1995). These new satellites will have slightly modified specifications with the AVHRR channel 3 being split into a 1.6 μm channel in the daylight phase of the orbit and remain in its original 3.55 - 3.93 μm configuration during the night time phase (Mandt 1995). This continuity in data collection is particularly important for long-term monitoring. Furthermore, as the archive of meteorological satellite data grows the possibility to model the responses of disease vector populations to changing ecological conditions (*i.e.* anthropogenic climate change) increases.

The range of satellite sensors for Earth observation have a variety of resolutions (spectral, spatial and temporal) that have been variously utilised with respect to disease control and management. The future will see a further increase in the range of data available, as well as its speed of dissemination and thus present new challenges and opportunities for synergism between the epidemiological and remote sensing communities.

Superseding and complimenting the NOAA instruments are the next generation of NASA Earth Observing System (EOS) satellites constructed with the explicit objective of producing "*a regular global dataset of well calibrated data of high radiometric resolution for a wide array of earth system sciences*" (Running *et al.* 1994). They will have an onboard Moderate Resolution Imaging Spectroradiometer (MODIS) with 36 channels spanning the spectral range 0.415 - 14.235 μm, with spatial resolutions of between 250 x 250 m and 1 x 1 km, a repeat time of only two days (Asrar and Greenstone 1995). The EOS satellite carrying MODIS is scheduled for launch in 1998.

It is proposed that various outputs, including more precise vegetation indices and land surface temperature products, will be provided together with a range of biophysical variables. Furthermore, the aim is that the information will be collected and disseminated rapidly via the Internet, giving previously unparalleled access to contemporary data on large area ecosystem processes (Price *et al.* 1994). This represents a massive leap in both the spectral quality (*i.e.* narrower channels can exploit parts of the electromagnetic spectrum in which signal attenuation from atmospheric constituents is minimal) and in the amount of data that will be available in near real-time to research workers. In preparation for these changes it would be prudent for epidemiologists to start considering the collection of georeferenced field data.

Similar advances in the fidelity and volume of satellite sensor data available to researchers will arise with the launch in 1998 of the European Organisation for the Exploitation of Meteorological Satellites' (EUMETSAT) Meteosat Second Generation (MSG). This satellite will be geostationary and maintain a view of the same hemisphere including Europe, Africa, and the Middle East as the present Meteosat systems (see section 1.2.3.2). The main payload will be the Spinning Enhanced Visible and Infra-Red Imager (SEVIRI) which will provide 1 x 1 km to 3 x 3 km spatial resolution images, every 15 minutes in 12 spectral channels ranging from 0.5 μm in the visible region to 13.4 μm in the infrared domain (Schmetz *et al.* 1995). The obvious advantage of this is the massive increase in temporal resolution enabling the possibility to look at developmental changes in populations over large areas. The increasing frequency of data capture will also allow for the more efficient screening of clouds and atmospheric attenuants.

6.4. Future research needs

Our understanding of the spatial and temporal distribution of arthropod vectors and the diseases they transmit has been enhanced by understanding the statistical associations between the ecological variables and processes observed remotely and vector biology. This understanding will only develop, however, when processed-based models using remote sensing data as real-time inputs to equations describing the survivorship of intermediate host species are generated. Epidemiologists are often fortunate in that a rudimentary understanding of the underlying vector biology is often available because the diseases transmitted have long been a threat to man and hence a focus of research. Further close investigation of specific test-sites will be required if the information to drive these models is to be generated. It is a priority to understand these interactions however, since they will enable "best guesses" to be made of the impact of global climate change on disease distributions (Rogers and Packer 1993, Rogers 1994, Patz *et al.* 1996).

The most detailed epidemiological understanding will come about through relating dynamic demographic processes (such as mortality rates) to satellite sensor data, rather than using satellite sensor imagery for purely static mapping exercises. The desire for such a detailed mechanistic or biological understanding must be tempered however, by the pragmatism of funding agencies and those in the developing tropics who are more concerned with a statistically accurate, operational control system than knowing precisely the causative relationships between a disease vector and its habitat.

Remote sensing of ecological processes is not a complete replacement for fieldwork. It provides a complementary approach which can extend our understanding of epidemiological processes in space and time in a way which has hitherto not been logistically feasible (Rogers and Williams 1993). In addition, remote sensing highlights the need for even more comprehensive surveying to obtain ground data that enables a more rigorous basis for its spatial extrapolation and to provide the inputs to processed-based models. This is especially true when the ecology of the diseases and their intermediate hosts, as well as human behaviour patterns that affect transmission vary substantially in space and time, significantly complicating our understanding.

6.5 Overall conclusion

This work should be viewed as a beginning. The utility of remotely sensed meteorological satellite sensor data in predicting tsetse fly and thus disease vector distribution and abundances has been demonstrated. So far this has been an essentially academic exercise. What remains is to tailor the techniques, and make available the datasets developed, so that those involved in disease vector control will have access to important and timely information on the distribution and abundance of their target organisms.

ACKNOWLEDGMENTS

" ... *In most of mankind gratitude is merely a secret hope for greater favours* ... "

Duc de la Rochefoucauld.

First and foremost I should like to thank my supervisors David Rogers, Willy Wint and Sarah Randolph. They have provided unflagging support, advice and assistance during the last 3 years and to them I am indebted.

Thanks are also due to the many people who have allowed access to satellite data without which this work would not have been possible. Compton Tucker and Chris Justice are thanked for access to the Global Inventory Monitoring and Modelling Systems (GIMMS) GAC AVHRR data, with which many of the current ideas were developed prior to the public release of the PAL dataset. Thanks are also due to Fred Snidjers of the FAO - ARTEMIS project for allowing access to Meteosat CCD data and to Tony Piccolo for arranging for its shipment. They both are to be further commended for their patience with my many queries about the data.

A formal acknowledgement is requested of those who use PAL data: " data used in this study include data produced through funding from the Earth Observing System Pathfinder Program of NASA's Mission to Planet Earth in cooperation with National Oceanic and Atmospheric Administration. The data were provided by the Earth Observing System Data and Information System (EOSDIS), Distributed Active Archive Center at Goddard Space Flight Center which archives, manages, and distributes this dataset". Compton Tucker, Chris Justice and Steve Prince are also thanked for providing encouragement and advice during work towards a Planetary Biology Internship at the NASA GSFC during the early stages of this work. I am also grateful to Scott Goetz for help and advice on retrieving vapour pressure deficit from NOAA - AVHRR towards the end of this work. Doreen Boyd is thanked for information on the elusive channel 3 and permission to plagiarise a figure 1.3.

I valued the comments of David Bourn, Mike Packer, Tim Robinson, Robert Green, Andrew Harrison and Sasha Norris who all read and promptly commented upon early drafts of the thesis. I would also like to thank Sarah Tomlin for the preparation of figure 1.1. and for reminding me that there is life beyond the thesis.

This work was funded by the Overseas Development Administration (ODA) under scheme R5794 and administered through the Natural Resources Institute (NRI). Green College is also thanked for providing a Graduate Award during each of my three years of study, a Texas Instruments Computer Scholarship in 1995 and 1996 and an Academic Grant towards conference attendance in the final year.

BIBLIOGRAPHY

" ... Manuscripts containing innumerable references are more likely a sign of insecurity than a mark of scholarship ... "

W. C. Roberts.

Achard, F. and 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. and 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.

Anyamba, A. and Eastman, J. R. (1996). Interannual variability of NDVI over Africa and its relation to El - Nino Southern Oscillation. *International Journal of Remote Sensing*, **17**: 2533-2548.

Anderson, J. R., Hardy, E. E., Roach, J. T. and Witmer, R. E. (1976). *A land use and land cover classification system for use with remote sensing data.* United States Geological survey Professional Paper 964. US Government Printing Office: Washington DC., USA.

Anderson, R. (1996). *The modelling of East African ungulate distributions, past and present, with satellite, climate and vegetation data.* Final Honours School Biological sciences undergraduate project thesis. University of Oxford: Oxford, UK.

Anderson, R. M. and May, R. M. (1991). Indirectly transmitted microparasites. In *Infectious diseases of humans: dynamics and control*, pp. 374-429. Oxford University Press: Oxford, UK.

Andres, L., Salas, W. A. and Skole, D. (1994). Fourier-analysis of multi-temporal AVHRR data applied to land cover classification. *International Journal of Remote Sensing*, **15**: 1115-1121.

Arambulo, P. V. and Astudillo, V. (1991). Perspectives on the application of remote sensing and geographic information system to disease control and health management. *Preventive Veterinary Medicine*, **11**: 345-352.

Asrar, G. (1989). *Theory and applications of optical remote sensing.* John Wiley and Sons: New York, USA.

Asrar, G. and Greenstone, R. (1995). *The Mission to Planet Earth (MTPE) Earth Observing System (EOS) reference handbook.* NASA, Goddard Space Flight Centre: Greenbelt, Maryland, USA.

Baldwin, D. and Emery, W. J. (1993). Systematized approach to AVHRR navigation. *Annals of Glaciology*, **17**: 414-420.

Bailey, C. L. and Linthicum, K. J. (1989). Satellite remote sensing, a future technology for monitoring vector populations and predicting arbovirus outbreaks. In *Arbovirus Research in Australia, Proceedings of the Fifth Symposium*, pp. 111-116. Commonwealth Science and Industrial Research Organisation: Canberra, Australia.

Barinaga, M. (1993). Satellite data rocket disease control efforts into orbit. *Science*, **261**: 31-32.

Barnes, C. M. and Cibula, W. G. (1979). Some implications of remote sensing technology in insect control programs including mosquitoes. *Mosquito News*, **39**: 271-282.

Barrett, E. C. and Curtis, L. F. (1992). *Introduction to environmental remote sensing*, 3rd Edition. Chapman and Hall: London, UK.

Baylis, M. and Nambiro, C. O. (1993). The responses of *Glossina pallidipes* and *G. longipennis* (Diptera: Glossinidae) to odour-baited traps and targets at Galana Ranch, south-eastern Kenya. *Bulletin of Entomological Research*, **83**: 145-151.

Beck, L. R., Rodriguez, M. H., Dister, S. W., Rodriguez, A. D., Rejmankova, E., Ulloa, A., Meza, R. A., Roberts, D. R., Paris, J.F., Spanner, M. A., Washino, R. K., Hacker, C. and Legters L. J. (1994). Remote sensing as a landscape epidemiologic tool to identify villages at high risk for malaria transmission. *American Journal of Tropical Medicine and Hygiene*, **51**: 271-280.

Beck, L. R., Wood, B. L. and Dister, S. W. (1995). Remote sensing and GIS: new tools for mapping human health. *Geo Info Systems*, **5**: 32-37.

Becker, F. (1987). The impact of land surface emissivity on the measurement of land surface temperature from a satellite. *International Journal of Remote Sensing*, **10**: 1509-1522.

Becker, F. and Z.-L. Li. (1990a). Temperature independent spectral indeces in the thermal infrared bands. *Remote Sensing Environment*, **32**: 17-33.

Becker, F. and Z.-L. Li. (1990b). Towards a local split window method over land surfaces. *International Journal of Remote Sensing*, **11**: 369-393.

Becker, F. and Z.-L. Li. (1995). Surface temperature and emissivity at various scales: definition, measurement and related problems. *Remote Sensing Reviews*, **12**: 225-253.

Belward, A. S. (1992). Spatial attributes of AVHRR imagery for environmental monitoring. *International Journal of Remote Sensing*, **13**: 193-208.

Belward, A. S. and Lambin, E. (1990). Limitations to the identification of spatial structures from AVHRR data. *International Journal of Remote Sensing*, **11**: 921-927.

Bernstein, R. L. (1982). Sea surface temperature estimation using the NOAA-6 satellite Advanced Very High Resolution Radiometer. *Journal of Geophysical Research Oceans and Atmospheres*, 87: 9455-9465.

Bidlingmayer, W. L. and Klock, J. W. (1955). Notes on the influence of salt-marsh topography on tidal action. *Mosquito News*, **15**: 231-235.

Bird, R. E. (1984). A simple solar spectral model for direct-normal and diffuse horizontal irradiance. *Solar Energy*, **32**: 461-171.

Bos, R. (1990). Application of remote sensing. *Parasitology Today*, **6**: 39.

Bourn, D and Wint, W. (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.

Boyd, D. S. and Curran, P. J. (1996). Using remote sensing to reduce uncertainties in the global carbon budget: the potential of radiation acquired in the middle infrared wavelengths. *Remote Sensing Reviews*, in press.

Boyd, D. S., Foody, G. M., Curran, P. J., Lucas, R. M. and 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.

Brown, O. W., Brown, J. W. and Evans R. H. (1985). Calibration of advanced very high resolution radiometer observations. *Journal of Geophysical Research*, **90**: 11667-11677.

Brown, S. A., Folk, M., Goucher G. and Rew, R. (1993). Software for portable scientific data management. *Computers in Physics*, **7**: 304-308.

Bruce-Chwatt, L. J. (1986). *Essential Malariology*, 2nd Edition. William Heinemann: London, UK.

Brush, R. J. H. (1988). The navigation of AVHRR imagery. *International Journal of Remote Sensing*, **9**: 1491-1502.

Bryan, J. H., Foley, D. H. and Sutherst, R. W. (1996). Malaria transmission and climate change in Australia. *Medical Journal of Australia*, **164**: 345-347.

Burgdorfer, W., Barbour, A. G., Hayes, S. F., Benach, J. L., Grunwaldt, E. and Davis, J. P. (1982). Lyme disease a tick-borne spirochetosis? *Science*, **216**: 1317.

Bursell, E. (1957). The effect of humidity on the activity of tsetse flies. *Journal of Experimental Biology*, **34**: 42-51.

Bursell, E. (1958). The water balance of tsetse pupae. *Philosophical Transactions of the Royal Society London B*, **241**: 179-210.

Bursell, E. (1959). The water balance of tsetse flies. *Transactions of the Royal Entomological Society London*, **111**: 205-235.

Burt, P. J. A., Colvin, J. and Smith, S. M. (1995). Remote sensing of rainfall by satellite as an aid to *Oedaleus Senegalensis* (Orthoptera, Acrididae) control in the Sahel. *Bulletin of Entomological Research*, **85**: 455-462

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

Buxton, P. A. (1955). Tsetse and climate. In *The natural history of tsetse flies. An Account of the biology of the Genus* Glossina *(Diptera)*, pp. 215-259. London School of Hygiene and Tropical Medicine Memoir. No 10. H. K. Lewis and Co: London, UK.

Byers, H. R. and Braham, R. R. Jn. (1948). Thunderstorm structure and circulation. *Journal of Meteorology*, **5**: 71-86.

Challier, A. and Laveissiere, C. (1973). Un nouveau piège pour la capture des glossines (Glossina: Diptera, Muscidae): description et essais sur la terrain. *Cahiers ORSTOM Series Entomologie, Médecine et Parasitologie*, **11**: 315-317.

Chatfield, C. (1980). *The analysis of time-series, an introduction.* Chapman and Hall: London, UK.

Che, N. and Price J. C. (1992). Survey of radiometric calibration results and methods for visible and near-infrared channels of NOAA-7, NOAA-9, and NOAA-11 AVHRRs. *Remote Sensing of Environment*, **41**: 19-27.

Chuvieco, E. and Martin M. P. (1994). A simple method for fire growth mapping using AVHRR channel 3 data. *International Journal of Remote Sensing*, **15**: 3141-3146.

Cibula, W. G. (1976). *Applications of remotely sensed multispectral data to automated analysis of marshland vegetation.* NASA Technical note D-8139. Health Applications Office

(HAO), National Aeronautics and Space Administration (NASA): Washington DC., USA.

Clarke, K. C., Osleeb, J. P., Sherry, J. M., Meert, J-P. and Larsson, R. W. (1991). The use of remote sensing and geographic information systems in UNICEF's dracunculiasis (Guinea worm) eradication effort. *Preventive Veterinary Medicine*, **11**: 229-235.

Clair, M. (1987). Données récentes sur la repartition des Glossines au Niger, au Burkina Faso et en Côte d'Ivoire.. In *International Scientific Council for Trypanosomiasis Research and Control (ISCTRC), Nineteenth Meeting, Lomé, Togo, 1987*, pp. 345-350. Organization of African Unity - Scientific and Technical Research Commission (OAU - STRC): Nairobi, Kenya.

Cline, B. L. (1970). New eyes for epidemiologists, aerial photography and other remote sensing techniques. *American Journal of Epidemiology*, **92**: 85-89.

Colvin, J and Gibson, G. (1992). Host-seeking behaviour and management of tsetse. *Annual Review of Entomology*, **37**: 21-40.

Colwell, J. E. (1974). Vegetation canopy reflectance. *Remote Sensing of Environment*, **3**: 175-183.

Colwell, R. N., Brewer, W., Landis, G., Langley, P., Morgan, J., Rinker, J., Robinson, J. M. and Sorem, A. L. (1963). Basic matter and energy relationships involved in remote reconnaissance. *Photogrammetric Engineering*, **29**: 761-799.

Colwell, R. N. (1983a). *Manual of remote sensing. I. Theory, instruments and techniques, Volume I.* 2nd Edition. American Society of Photogrammetry: Falls Church, Virginia, USA.

Colwell, R. N. (1983b). *Manual of remote sensing. II. Interpretation and applications, Volume II.* 2nd Edition. American Society of Photogrammetry: Falls Church, Virginia, USA.

Cooper, D. I. and 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.

Cracknell, A. P. (1996). *The Advanced Very High Resolution Radiometer.* Taylor and Francis Ltd: London, UK.

Cracknell, A. P. and Hayes, L. W. B. (1991). *Introduction to remote sensing.* Taylor and Francis Ltd: London, UK.

Curran, P. J. (1985). *Principles of remote sensing.* Longman Group Ltd: London, UK.

Curran, P. J. and Williamson, H. D. (1987). GLAI estimation using measurements red, near infrared and middle infrared radiance. *Photogrammetric Engineering and Remote Sensing*, **53**: 181-186.

Dadhwal, V. K., Parihar, J. S. Medhavy, T. T., Ruhal, D. S., Jarwal, S. D. and Khera, A. P. (1996). Comparative performance of Thematic Mapper middle-infrared bands in crop discrimination. *International Journal of Remote Sensing*, **17**: 1727-1734.

Dalu, G. (1986). Satellite remote sensing of atmospheric water vapor. International Journal of Remote Sensing, 7: 1089-1097.

D'Amico, F., Moussa, A., Sarda, J. and Gouteux, J. P. (1992). Distribution et importance des gîtes à *Glossina fuscipes* Newstead 1910 dans l'agglomération de Bangui (République Centrafricaine). *Bulletin de la Société de Pathologie Exotique Filiales*, **85**: 64-68.

Daniel, M. and Dusbábek, F. (1994). Micrometeorological and microhabitat factors affecting maintenance and dissemination of tick-borne diseases in the environment.. In *Ecological dynamics of tick borne zoonoses*, pp. 91-138, editors Sonenshine, D. E. and Mather, T. N. Oxford University Press: Oxford, UK.

Daniel, M. and Kolar, J. (1990). Using satellite data to forecast the occurrence of the common tick *Ixodes ricinus* (L.). *Journal of Hygiene, Epidemiology, Microbiology and Immunology*, **34**: 243-252.

De Muynck, A. and Rogers, D. J. (1989). Workshop on modelling sleeping sickness epidemiology and control. *Annales de la Société Belge de Médecine Tropicale*, **69**: 1-256.

Desowitz, R. S. (1993). The first trial and tribulations of the malaria vaccine. In *The malaria capers: more tales of parasites, people, research and reality*, pp. 221-233. W. W. Norton and Company: New York, USA.

Defence Mapping Agency. (1992). Development of the Digital Chart of the World. Government Printing Office: Washington, DC., USA.

DeFries, R. S. and Townshend, J. R. G. (1994). NDVI-derived land cover classifications at a global scale. *International Journal of Remote Sensing*, **15**: 3567-3586.

Dietz, K. (1988). Mathematical models for transmission and control of malaria. In *Malaria: principles and practice of malariology*, Volume 2, pp. 1091-1133. Churchill Livingstone: New York, USA.

Doku, C. R.. (1995). *Personal communication.* Unit Head, Tsetse and trypanosomiasis Control Unit: Tamale, Ghana.

Dougill, A. and Cox, J. (1995). Land degradation and grazing in the Kalahari: new analysis and alternative perspectives. *Pastoral Development Network*, **38c**: 1-22.

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

Dudhia, A. (1989). Noise characteristics of the AVHRR infrared channels. *International Journal of Remote Sensing*, **10**: 637-644.

Dugdale, G., Mcdougall, V. and Thorne, V. (1995). The TAMSAT method for estimating rainfall over Africa and its implications. In *RSS95 Remote Sensing in Action. Proceedings of the 21st Annual Conference of the Remote Sensing Society*, pp. 773-780. Remote Sensing Society: Nottingham, UK.

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

Eck, T. F. and Holben B. N. (1994). AVHRR split window temperature differences and total precipitable water over land surfaces. *International Journal of Remote Sensing*, **15**: 567-582.

Ehrlich, D., Estes, J. E. and Singh, A. (1994). Applications of NOAA-AVHRR 1 km data for environmental monitoring. *International Journal of Remote Sensing*, **15**: 145-161.

Ehrlich, D. and Lambin, E. F. (1996). Broad scale land-cover classification and interannual climatic variability. *International Journal of Remote Sensing*, **17**: 845-862.

Eklundh, L. and Singh, A. (1993). A comparative-analysis of standardized and unstandardized principal components analysis in remote sensing. *International Journal of Remote Sensing*, **14**: 1359-1370

Emanuel, K. A. (1994). Observed characteristics of precipitating convection. In *Atmospheric convection*, pp. 230-279. Oxford University Press: Oxford, UK.

Emery, W. J., Brown, J. and Nowak, Z. P. (1989). AVHRR image navigation: summary and review. *Photogrammetric Engineering and Remote Sensing*, **55**: 1175-1183.

Epstein, P. R., and Chikwenhere, G. P. (1994). Environmental factors in disease surveillance. *The Lancet*, **343**: 1440-1441.

Epstein, P. R. Rogers, D. J. and Slooff, R. (1993). Satellite imaging and vector-borne disease. *The Lancet*, **341**: 1404-1406.

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.

European Organisation for the Exploitation of Meteorological Satellites. (1995). *Meteosat high resolution image dissemination. Technical description.* EUM TD 02. EUMETSAT: Darmstadt, Germany.

Ewusie, J. Y. (1968). Preliminary studies on the phenology of some woody species of Ghana. *Ghana Journal of Science*, **8**: 126-133

Ewusie, J. Y. (1969). Some observations on the annual pattern of flowering of some tropical woody plants. *Ghana Journal of Science*, **9**: 74-79

Ewusie, J. Y. (1992). *Phenology in tropical ecology.* Ghana Universities Press: Accra, Ghana.

Fialho, R. F. and Schall, J. J. (1995). Thermal ecology of a malarial parasite and its insect vector: consequence for the parasite's transmission success. *Journal of Animal Ecology*, **64**: 553-562.

Flasse, S. and 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.

Fleetwood, S. C., Chambers, M. D. and Terracina, C. (1981). An effective and economical mapping system for the monitoring of *Psorophora columbiae* in rice and fallow fields in south-western Louisiana. *Mosquito News*, **41**: 174-177.

Fleig, A. J., Heath, D. F., Klenk, K. F., Oslik, N., Lee, K. D., Park, H., Bartia, P. K. and Gordon. D. (1983). *User's Guide for the Solar Backscattered Ultraviolet (SBUV) and the Total Ozone Mapping Spectrometer (TOMS) RUT-S and RUT-T Data Sets: October 31, 1978 to November 1980.* NASA Reference Publication 1112.

Flitcroft, I. D., Milford, J. R. and Dugdale, G. (1989). Relating point to average rainfall in semi-arid West Africa and implications for rainfall estimates derived from satellite data. *Journal of Applied Meteorology*, **28**: 252-266.

Fontés, J., Gastellu-Etchegorry, J. P., Amram, O. and Flouzat, G. (1995). A phenological model of the African continent. *Ambio*, **24**: 297-303.

Foody, G. M. (1996). Fuzzy modeling of vegetation from remotely-sensed imagery. *Ecological Modelling*, **85**: 3-12.

Foody, G. M., Boyd, D. S. and 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.

Foody, G. M. and Cox, D. P. (1994). Sub-pixel land-cover composition estimation using a linear mixture model and fuzzy membership functions. *International Journal of Remote Sensing*, **15**: 619-631.

Foody, G. M. and Curran P. J. (1994). Estimation of tropical forest extent and regenerative stage using remotely sensed data. *Journal of Biogeography*, **21**: 223-244.

Ford, J. (1968). The control of populations through limitation of habitat distribution as exemplified by tsetse flies. In *Insect Abundance*, pp. 109-118. Symposia of the Royal Entomological Society of London: number four. Blackwell Scientific Publications: Oxford, UK.

Ford, J. (1970). The geographical distribution of the *Glossina*. In *The African Trypanosomiases*, pp. 274-297. George Allen and Unwin Ltd: London, UK.

Ford, J. (1971). *The role of the trypanosomiases in African ecology, a study of the tsetse fly problem.* Clarendon Press Ltd: Oxford, UK.

Ford, J. and Katondo, K. M. (1977). *The distribution of tsetse flies (Glossina) in Africa in 1973.* Organization of African Unity - Scientific and Technical Research Commission (OAU STRC). Cook, Hammond and Kell: London, UK.

Ford, J. and Leggate, B. M. (1961). The geographical and climatic distribution of trypanosome infection rates in *G. morsitans* group of tsetse-flies. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, **55**: 383-397.

Freeman, T. and Bradley, M. (1996). Temperature is predictive of severe malaria years in Zimbabwe. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, **90**: 232.

Frezil, J. -L., Eouzan, J. -P., Alary, J. -C., Malonga, J. -R. and Ginoux, P. Y. (1980). Epidémiologie de la trypanosomiase humaine en République populaire du Congo, *Cahiers ORSTOM Series Entomologie, Médecine et Parasitologie*, **18**: 329-346.

Fuller, D. O. and Prince, S. D. (1996). Rainfall and foliar dynamics in tropical Southern Africa - potential impacts of global climatic change on savanna vegetation. *Climatic Change*, **33**: 69-96.

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

Giddings, L. E. (1976). *Remote sensing for the control of tsetse flies.* Technical memorandum. Health Applications Office, National Aeronautics and Space Administration (HAO - NASA): Washington DC., USA.

Global Land Information System. (1994). The 30 arc-sec DCW digital elevation models. http://edcwww.cr.usgs.gov/Webglis/glisbin/guide.pl?guide=/glis/hyper/guide/30asdcwdem.

Godfrey, D. G. (1978). Identification of economically important parasites. *Nature*, **273**: 600-604.

Goel, N. S., (1988). Models of vegetation canopy reflectance and their use in the estimation of biophysical parameters from reflectance data. *Remote Sensing Reviews*, **1**: 1-222.

Goodwin, R. F. W. (1973). The cost effectiveness of animal disease control. *Span*, **16**: 63-64.

Gopal, S and Woodcock, C. (1994). Theory and methods for accuracy assessment of thematic maps using fuzzy sets. *Photogrammetric Engineering and Remote Sensing*, **60**: 181-188.

Gordon, H. R., Brown, J. W. and Evans, R. H. (1988). Exact Rayleigh scattering calculations for use with the Nimbus-7 coastal zone color scanner. *Applied Optics*, **27**: 862-871.

Gorman, A. J. and Mcgregor, J. (1994). Some considerations for using AVHRR data in climatological studies: II. Instrument performance. *International Journal of Remote Sensing*, **15**: 549-565.

Goward, S. N., Markham, B., Dye, D. G., Dulaney, W. and Yang, J. (1991). Normalized difference vegetation index measurements from the Advanced Very High Resolution Radiometer. *Remote Sensing of Environment*, **35**: 257-277.

Goward, S. N. and Prince, S. D. (1995). Transient effects of climate on vegetation dynamics satellite observations. *Journal of Biogeography*, **22**: 549-564.

Goward, S. N., Waring, R. H., Dye, D. G. and Yang, J. (1994). Ecological remote sensing at OTTER: satellite macroscale observations. *Ecological Applications*, **4**: 322-343.

Green, C. H. (1994). Bait methods for tsetse control. *Advances in Parasitology*, **34**: 229-291.

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

Gregoire, J. M. (1990). Effects of the dry season on the vegetation canopy of some river basins of West Africa as deduced from NOAA-AVHRR data. *Hydrological Sciences Journal*, **35**: 323-338.

Gutman, G. G. (1993). Multi-annual time series of AVHRR-derived land surface temperature. *Advances in Space Research*, **14**: 27-30.

Gutman, G. and Ignatov, A. (1996). The relative merit of cloud/clear identification in the NOAA/NASA Pathfinder AVHRR Land 10-day composites. *International Journal of Remote Sensing*, **17**: 3295-3304.

Guyenne, T. D. (1987). *Introduction to the Meteosat operational system.* ESA BR 32. European Space Agency (ESA): Paris, France.

Hacker, C. S. and Roberts, D. R. (1994). Public health applications of satellite remote sensing. *Sistema Terra*, **3**: 34-36.

Hannan, N. P., Prince, S. D. and Holben, B. N. (1995). Atmospheric correction of AVHRR data for biophysical remote sensing of the Sahel. *Remote Sensing of Environment*, **51**: 306-316.

Harley, J. M. B. (1954). The breeding sites of the tsetse fly *Glossina morsitans*. *Acta Tropica*, **11**: 379-402.

Hastings, D. A. and Emery, W. J. (1992). The Advanced Very High Resolution Radiometer (AVHRR): a brief reference guide. *Photogrammetric Engineering and Remote Sensing*, **58**: 1183-1188.

Hay, S. I. and Lennon, J. J. (1996). A comparison of the accuracy of AVHRR retrieved radiometric brightness temperatures and interpolated thermodynamic temperatures for Africa. In preparation.

Hay, S. I., Packer, M. J. and Rogers, D. J. (1996a). A review of the impact of remote sensing on the study and control of invertebrate intermediate hosts of disease. *International Journal of Remote Sensing*, submitted.

Hay, S. I., Tucker, C. J., Rogers, D. J. and Packer, M. J. (1996b). Remotely sensed surrogates of meteorological data for the study of the distribution and abundance of arthropod vectors of disease. *Annals of Tropical Medicine and Parasitology*, **90**: 1-19.

Hay, S. I. (1997). Remote sensing and disease control: past, present and future. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, in press.

Hayes, R. O., Maxwell, E. L., Mitchell, C. J. and Woodzick, T. L. (1985). Detection, identification, and classification of mosquito larval habitats using remote sensing

scanners in earth-orbiting satellites. *Bulletin of the World Health Organization*, **63**: 361-374.

Hendrickx, G., Rogers, D. J., Napala, A. and Slingenbergh, J. H. W. (1995). Predicting the distribution of riverine tsetse and the prevalence of bovine trypanosomiasis in Togo using ground-based and satellite data. In *International Scientific Council for Trypanosomiasis Research and Control (ISCTRC), Twenty Second Meeting, Kampala, Uganda, 1993*, pp. 218-227. Organization of African Unity Scientific and Technical Research Commission (OAU - STRC): Nairobi, Kenya.

Herring, T. A. (1996). The global positioning system. *Scientific American*, **274**: 44-50.

Hide, G. (1994). The elusive trypanosome. *Parasitology Today*, **10**: 85-86.

Hiernaux, P. (1993). Droughts, not animals, degrade Sahelian rangelands. *International Livestock Centre for Africa (ILCA) Newsletter*, **12**: 6.

Hill, M. O. (1991). Patterns of species distribution in Britain elucidated by canonical correspondence analysis. *Journal of Biogeography*, **18**: 247-255.

Hoare, C. A. (1972). *The trypanosomes of mammals*. Tyndall and Cox Ltd: London, UK.

Holben, B. N. (1986). Characteristics of maximum-value composite images from temporal AVHRR data. *International Journal of Remote Sensing*, **7**: 1417-1434.

Holben, B. N., Eck, T. F. and Fraser, R. S. (1991). Temporal and spatial variability of aerosol optical depth in the Sahel region in relation to vegetation remote sensing. *International Journal of Remote Sensing*, **12**: 1147-1163.

Holben, B. N. and Shimabukuro, Y. E. (1993). Linear mixing model applied to coarse spatial resolution data from multispectral satellite sensors. *International Journal of Remote Sensing*, **14**: 2231-2240.

Hopkins, C. C., Hollinger, F. B., Johnson, R. F., Dewlett, H. J., Newhouse, V. F. and Chamberlain, R. W. (1975). The epidemiology of St. Louis encephalitis in Dallas, Texas, 1966. *American Journal of Epidemiolgy*, **102**: 1-15.

Horler, D. N. H. and Ahern, F. J. (1986). Forestry information content of Thematic Mapper data. *International Journal of Remote Sensing*, **7**: 405-428.

Huete, A. R. (1988). A soil adjusted vegetation index (SAVI). *Remote Sensing of Environment*, **25**: 295-309.

Huete, A. R. and Escadafal, R. (1991). Assessment of biophysical soil properties through spectral decomposition techniques. *Remote Sensing of Environment*, **35**: 149-159.

Huete, A. R., Jackson, R. D. and Post, D. F. (1985). Spectral response of a plant canopy with different soil backgrounds. *Remote sensing of Environment*, **17**: 37-53.

Hugh-Jones, M. (1989). Applications of remote sensing to the identification of the habitats of parasites and disease vectors. *Parasitology Today*, **5**: 244-251.

Hugh-Jones, M. (1991a). Satellite imaging as a technique for obtaining disease related data. *Revue Scientifique et Technique de l'Office International Epizooties*, **10**: 197-204.

Hugh-Jones, M. (1991b). The remote recognition of tick habitats. *Journal of Agricultural Entomology*, **8**: 309-315.

Hugh-Jones, M., Barre, N., Nelson, G., Wehnes, K., Warner, J., Garvin, J. and Garris, G. (1992). Landsat-TM identification of *Amblyomma variegatum* (Acari: Ixodidae) habitats in Guadeloupe. *Remote Sensing of Environment*, **40**: 43-55.

Hugh-Jones, M., Barre, N., Nelson, G., Wheyness, C., Warner, J., Garris, G. and Hubbet, W. (1988). Remote recognition of *Amblyomma variegatum* habitats in Guadeloupe using Landsat-TM imagery. *Acta Veterinaria Scandinavica*, **84**: 259-261.

Hugh-Jones, M. and O'Neil, P. (1986). The epidemiological uses of remote sensing and satellites. In *Proceedings of the Fourth International Symposium on Veterinary Epidemiology and Economics*, pp. 113-118. Singapore Veterinary Association: Singapore.

Huh, O. K. (1991). Limitations and capabilities of the NOAA satellite advanced very high resolution radiometer (AVHRR) for remote sensing of the Earth's surface. *Preventive Veterinary Medicine*, **11**: 167-183.

Hutchinson, M. F. (1988). Calculation of hydrologically sound digital elevation models. In Third International Symposium on Spatial Data Handling, pp. 117-113. International Geographic Union: Columbus, USA.

Hutchinson, M. F. (1989). A new procedure for gridding elevation and stream line data with automatic removal of spurious pits. *Journal of Hydrology*, **106**: 211-232.

Institut d'Elevage et de Médecine Vétérinaire des Pays Tropicaux. (1982). *Cartographie de la repartition des glossines*. IEMVT: Maison-Valfort, France.

Jackson, R. D. and Huete, A. R. (1991). Interpreting vegetation indices. *Preventive Veterinary Medicine*, **11**: 185-200.

James, M. E. and 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.

Jeyaseelan, A. T. and Thiruvengadachari, S. (1993). Suspected Mt. Pinatubo aerosol impact on the NOAA AVHRR NDVI over India. *International Journal of Remote Sensing*, **14**: 603-608.

Jordan, A. M. (1986). *Trypanosomiasis control and African rural development.* Longman: London, UK.

Jordan, A. M. (1992). Degradation of the environment: an inevitable consequence of trypanosomiasis control? *World Animal Review*, **70/71**: 2-7.

Jovanovic, P. (1987a). Satellite remote sensing imagery in public health. *Acta Astronautica*, **15**: 951-953.

Jovanovic, P. (1987b). Remote sensing of environmental factors affecting health. *Advances in Space Research*, **7**: 11-18.

Justice, C. O., Bailey, G. B., Maiden, M. E., Rasool, S. I., Strebel, D. E. and Tarpley, J. D. (1995). Recent data and information system initiatives for remotely sensed measurements of the land surface. *Remote Sensing of Environment*, **51**: 235-244.

Justice, C. O., Eck, T. F., Tanré, D. and Holben, B. N. (1991). The effect of water vapour on the normalized difference vegetation index derived for the Sahelian region from NOAA AVHRR data. *International Journal of Remote Sensing*, **12**: 1165-1187.

Justice, C. O., Markham, B. L., Townshend, J. R. G. and Kennard, R. L. (1989). Spatial degradation of satellite data. *International Journal of Remote Sensing*, **10**: 1539-1561.

Justice, C. O., Townshend, J. R. G., Holben, B. N. and Tucker, C. J. (1985). Analysis of the phenology of global vegetation using meteorological satellite data. *International Journal of Remote Sensing*, **10**: 1539-1561.

Kalluri, S. (1996). Errors in the computation of the solar zenith angles in the PAL data. http://daac.gsfc.nasa.gov/CAMPAIGN_DOCS/LAND_BIO/zenith_angle_memo.html.

Kaufman, Y. L. and Holben, B. N. (1993). Calibration of the AVHRR visible and near-IR bands by atmospheric scattering, ocean glint and desert refection. *International Journal of Remote Sensing*, **14**: 21-52.

Kaufman, Y. L. and Remer, L. A. (1994). Detection of forests using MID-IR reflectance: an application for aerosol studies. *IEEE Transactions on Geosciences and Remote Sensing*, **32**: 672-683.

Kaufman, Y. L. and Tanré, D. (1992). Atmospherically resistant vegetation index (ARVI) for EOS-MODIS. *IEEE Transactions on Geoscience and Remote Sensing*, **30**: 261-270.

Kerber, A. G. and Schutt, J. B. (1986). Utility of AVHRR channel 3 and 4 in land-cover mapping. *Photogrammetric Engineering and Remote Sensing*, **52**: 1877-1883.

Khan, B., Hayes, L. and Cracknell, A. P. (1992). The optimisation of higher-order resampling methods in the multiprocessor environment. *Parallel Computing*, **18**: 1335-1347.

Khan, B., Hayes, L. W. B. and Cracknell, A. P. (1995). The effects of higher-order resampling on AVHRR data. *International Journal of Remote Sensing*, **16**: 147-163.

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.

Kineman, J., Boyle, S. C., Mealy, A., Ohrenschall, M., Colby, J. D., Della Mana, S. and Mellon, D. (1990). *Global change database project: pilot project for Africa. Data set documentation, version 1.0 (Dakar Workshop).* World Data Centre A for Solid Earth Geophysics, United States Department of Commerce: Bolder, Colorado, USA.

Kitron, U., Otieno, L. H., Hungerford, L. L., Odulaja, A., Brigham, W. U., Okello, O. O., Joselyn, M., Mohamed-Ahmed, M. M. and Cook, E. (1996). Spatial analysis of the distribution of tsetse flies in the Lambwe Valley, Kenya, using Landsat TM satellite imagery and GIS. *Journal of Animal Ecology*, **65**: 371-380.

Kleiner, K. (1995). Satellites wage war on disease. *New Scientist*, **148**: 9.

Knox, L. (1995). The Earth science data and information system project. http://spsosun.gsfc.nasa.gov/ESDIShome.html.

Krasnopolsky, V. M. and Breaker, L. C. (1994). The problem of AVHRR image navigation revisited. *International Journal of Remote Sensing*, **15**: 979-1008.

Kuzoe, F. A. S. (1993). Current situation of African trypanosomiasis. *Acta Tropica*, **54**: 153-162.

Lambin, E. F. and 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.

Lambin, E. F. and Ehrlich, D. (1996). The surface temperature-vegetation index space for land cover and land-cover changes analysis. *International Journal of Remote Sensing*, **17**: 463-487.

Laveissiere, C. and 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. and 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.

Leake, J. P., Musson, E. K. and Chope, H. D. (1934). Epidemiology of epidemic encephalitis, St. Louis type. *Journal of the American Medical Association*, **103**: 728-731.

Lennon, J. J. and Turner, J. R. G. (1995). Predicting the spatial distribution of climate: temperature in Great-Britain. *Journal of Animal Ecology*, **64**: 370-392.

Leprieur, C., Kerr, Y. H. and Pichon, J. M. (1996). Critical assessment of vegetation indexes from AVHRR in a semiarid environment. *International Journal of Remote Sensing*, **17**: 2549-2563.

Lessard, P., L'Eplattenier, R., Norval, R. A. I., Kundert, K., Dolan, T. T., Croze, H., Walker, J. B., Irvin, A. D. and Perry, B D. (1990). Geographical information systems for studying the epidemiology of cattle diseases caused by *Theileria parva*. *The Veterinary Record*, **126**: 255-262.

Li, Z. -L. and Becker, F. (1993). Feasibility of land surface temperature and emissivity determination from AVHRR data. *Remote Sensing of Environment*, **43**: 67-85.

Lillesand, T. M. and Kiefer, R. W. (1994). *Remote sensing and image interpretation.* 3rd Edition. John Wiley and Sons, Inc.: New York, USA.

Linthicum, K. J., Bailey, C. L., Glyn Davies, F. and Tucker, C. J. (1987). Detection of Rift Valley fever viral activity in Kenya by satellite remote sensing imagery. *Science*, **235**: 1656-1659.

Linthicum, K. J., Bailey, C. L., Tucker, C. J., Angleberger, D. R., Cannon, T., Logan, T. M., Gibbs, P. H. and Nickeson, J. (1991). Towards real-time prediction of Rift Valley fever epidemics in Africa. *Preventive Veterinary Medicine*, **11**: 325-334.

Linthicum, K. J., Bailey, C. L., Tucker, C. J., Gordon, S. W., Logan, T. M., Peters, C. J. and Digoutte, J. P. (1994). Man-made ecological alterations of Senegal River basin on Rift Valley Fever transmission. *Sistema Terra*, **3**: 45-47.

Linthicum, K. J., Bailey, C. L., Tucker, C. J., Mitchell, K. D., Logan, T. M., Davies, F. G., Kamau, C. W., Thande, P. C. and Wagateh, J. N. (1990). Application of polar-orbiting, meteorological satellite data to detect flooding in Rift Valley Fever virus vector mosquito habitats in Kenya. *Medical and Veterinary Entomology*, **4**: 433-438.

Lord, C. C., Woolhouse, M. E. J., Heesterbeek, J. A. P. and Mellor, P. S. (1996). Vector-borne diseases and the basic reproduction number: a case study of African horse sickness. *Medical and Veterinary Entomology*, **10**: 19-28.

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. and Redmond, R. L. (1995). Tau coefficients for accuracy assessment of classification of remote sensing data. *Photogrammetric Engineering and Remote Sensing*, **61**: 435-439.

Macdonald, G. (1953). The analysis of malaria epidemics. *Tropical Diseases Bulletin*, **50**: 227-235.

Malingreau, J. P. and 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.

Malingreau, J. P., Stephens, G. and Fellows, C. (1985). 1982- 83 forest fire of Kalimantan and North Borneo: satellite observations for detection and monitoring. *Ambio*, **14**: 314-321.

Malingreau, J. P., Tucker, C. J. and Laporte, N. (1989). AVHRR for monitoring global tropical deforestation. *International Journal of Remote Sensing*, **10**: 855-867.

Mandt, G. (1995). *NOAA polar-orbiting operational environmental satellite program: status and plans*. In Proceedings of the Meteorological Satellite Data Users' Conference, pp. 29-35. EUMETSAT: Darmstadt, FDR.

Marriott, F. H. C. (1974). *The interpretation of multiple observations*. Academic Press: London, UK.

Marsh, S. E., Walsh, J. L., Lee, C. T., Beck, L. R. and 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.

Martyn, D. (1992). The climates of Africa. In *Developments in Atmospheric Science, 18. Climates of the World*, pp. 199-261. Elsevier: Amsterdam, The Netherlands.

Matson, M. and Holben, B. (1987). Satellite detection of tropical burning in Brazil. *International Journal of Remote Sensing*, 8: 509-516.

Mattingley, P. F. (1969). *The biology of mosquito-borne diseases*. George Allen and Unwin Ltd: London, UK.

Maudlin, I. (1991). Transmission of African trypanosomiasis: interactions among tsetse immune system, symbionts and parasites. *Advances in Disease and Vector Research*, **7**: 117-148.

McGregor, J. and Gorman, A. J. (1994). Some considerations for using AVHRR data in climatological studies, I. Orbital characteristics of NOAA satellites. *International Journal of Remote Sensing*, **15**: 537-548.

McIntyre, J., Bourzat, D and Pingali, P. (1992). *Crop livestock interaction in sub-Saharan Africa.* Regional and sectoral study series. World Bank: Washington DC., USA.

Menenti, M., Azzali, S., Verhoef, W. and van Swol, R. (1993). Mapping agroecological zones and time-lag in vegetation growth by means of Fourier analysis of time-series of NDVI images. *Advances in Space Research*, **13**: (5)233-(5)237.

Meyer, D. J. (1996). Estimating the effective spatial resolution of an AVHRR time series. *International Journal of Remote Sensing*, **17**: 2971-2980.

Meyer, W. B. and Turner II, B. L. (1992). Human population growth an global land-use/cover change. *Annual Review of Ecology and Systematics*, **23**: 39-61.

Meyer, W. B. and Turner II, B. L. (1994). *Changes in land use and land cover: a global perspective.* Cambridge University Press: Cambridge, UK.

Milford, J. R. and Dugdale, G. (1990). Monitoring of rainfall in relation to the control of migrant pests. *Philosophical Transactions of the Royal Society London B*, **328**: 689-704.

Mills, A. (1994). The economics of vector control strategies for controlling tropical diseases. *American Journal of Tropical Medicine and Hygiene*, **50**: 151-159.

Moloo, S. K. (1993). The distribution of *Glossina* species in Africa and their natural hosts. *Insect Science and its Application*, **14**: 511-527.

Molyneux, D. H. (1993). Control. In *Modern parasitology: a textbook of parasitology*, 2nd edition. Editor F. E .G. Cox, pp. 243-263. Blackwell Scientific Publications: Oxford, UK.

Monbiot, G. (1994). *No man's land. An investigative journey through Kenya and Tanzania.* Macmillan: London, UK.

Monteith, J. L and Unsworth, M. H. (1990). *Principles of Environmental Physics.* 2nd Edition. Edward Arnold: London, UK.

Mouchet, J. and Carnevale, P. (1981). Malaria endemicity in the various phytogeographic and climatic areas of Africa, south of the Sahara. *South East Asian Journal of Tropical Medicine and Public Health*, **12**: 439-440.

Mulchay, K. A. and Clarke, K. C. (1994). Government digital cartographic data policy and environmental research needs. *Computers, Environment and Urban Systems*, **18**: 95-101.

Mukhebi, A. W. (1992). Economic impact of theileriosis and its control in Africa. In *The epidemiology of theileriosis in Africa.* Edited by R. A. I. Norval., B. D., Perry and A. S. Young, pp. 379-245. Academic Press: London, UK.

Murray, M. and Gray, A. R. (1984). The current situation on animal trypanosomiasis in Africa. *Preventive Veterinary Medicine*, **2**: 28-30.

Myneni, R. B., Hall, F. G., Sellers, P. J. and Marshak, A. L. (1995a). The interpretation of spectral vegetation indexes. *IEEE Transactions on Geosciences and Remote Sensing*, **33**: 481-486.

Myneni, R. B., Maggion, S., Iaquinta, J., Privette, J. L., Gobron, M., Pinty, B., Kimes, D. S., Verstrate, M. M. and Williams, D. L. (1995b). Optical remote sensing of vegetation: modelling caveats and algorithms. *Remote Sensing of Environment*, **51**: 169-168.

Nash, T. A. M. (1933). A statistical analysis of the climatic factors influencing the density of tsetse flies, *Glossina morsitans* Westw. *Journal of Animal Ecology*, **2**: 197-203.

Nash, T. A. M. (1969). *Africa's bane. The tsetse fly.* Collins: London, UK.

Nationaal Lucht- en Ruimtevaartlaboratorium. (1988). *Computer programme description of the ARTEMIS thematic processing.* NLR document TR88179L. NRL: Amsterdam, Netherlands.

National Aeronautics and Space Administration. (1973). *The use of remote sensing in mosquito control.* Health Applications Office, NASA (HAO - NASA): Washington DC., USA.

National Center for Supercomputing Applications. (1990). *Hierarchical Data Format (HDF) Users Guide.* NCSA, University of Illinois: Urbana-Champaign, USA.

National Geophysical Data Center. (1993). *ETOPO5 5-minute gridded elevations/bathymetry for the world.* Data announcement 93-MGG-01. NGDC: Boulder, Colorado, USA.

National Oceanic and Atmospheric Administration. (1972). *Substation observations. National Weather Service (NWS) observing handbook.* No. 2. Government Printing Office: Washington DC., USA.

National Oceanic and Atmospheric Administration. (1990). *Monthly climate data for the word. Prepared in cooperation with the World Meteorological Organisation.* Volume 41, 42 and 43; numbers 1-12. National Climate Data Centre (NCDC), NOAA: North Carolina, USA.

Nemani, R. R. and Running, S. W. (1988). Estimation of regional surface resistance to evapotranspiration from NDVI and thermal-IR AVHRR data. *Journal of Applied Meteorology*, **28**: 276-284.

Nemani, R., Pierce, L., Running, S. and Goward, S. (1993). Developing satellite-derived estimates of surface moisture status. *Journal of Applied Meteorology*, **32**: 548-557.

Nerry, F., Labed, J. and Stoll, M. P. (1990). Spectral properties of land surfaces in the thermal infrared. Part I: laboratory measurements of absolute spectral emissivity and reflectivity signatures. *Journal of Geophysical Research*, **95**: 7027-7044.

Njoku, E. (1963). Seasonal periodicity in the growth and development of some forest in Nigeria. I. Observation on mature trees. *Journal of Ecology*, **51**: 617-624.

Njoku, E. (1964). Seasonal periodicity in the growth and development of some forest in Nigeria. II. Observation on seedlings. *Journal of Ecology*, **52**: 19-26.

Norman, J. M., Divakarla, M. and Goel, N. S. (1995). Algorithms for extracting information from remote thermal-IR observations of the Earth's surface. *Remote Sensing Environment*, **51**: 157-168.

Norval, R. A. I., Perry, B. D., Gebreab, F. and Lessard, P. (1991). East Coast fever, a problem of the future for the horn of Africa? *Preventive Veterinary Medicine*, **10**: 163-172.

Olsson, L. (1993). On the causes of famine - drought, desertification and market failure in the Sudan. *Ambio*, **22**: 395-403.

Olsson, L. and Eklundh, L. (1994). Fourier series for analysis of temporal sequences of satellite sensor imagery. *International Journal of Remote Sensing*, **15**: 3735-3741.

Openshaw, S. (1996). Geographical information systems and tropical diseases. *Transactions of the Royal Society of Tropical medicine and Hygiene*, **90**: 337-339.

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

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

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

Patt, F. S. and Gregg, W. W. (1994). Exact closed-form geolocation algorithm for earth survey sensors. *International Journal of Remote Sensing*, **15**: 3719-3734.

Patz, J. A., Epstein, P. R., Burke, T. A. and Balbus, J. M. (1996). Global climate change and emerging infectious diseases. *Journal of the American Medical Association*, **275**: 217-223.

Pender, J. and Rosenberg, J. (1995). *Impact of tsetse control on land use in the semi-arid zone of Zimbabwe. Phase I: Classification of land use by remote sensing imagery.* Natural Resources Institute (NRI) Bulletin 66. NRI: Chatham, UK.

Pereira, A. C. and Setzer, A. W. (1996). Comparison of fire detection in savannas using AVHRRs channel 3 and TM images. *International Journal of Remote Sensing*, **17**: 1925-1937

Perry, B. D., Kruska, R., Lessard, P., Norval, R. A. I. and Kundert, K. (1991). Estimating the distribution and abundance of *Rhipicephalus appendiculatus* in Africa. *Preventive Veterinary Medicine*, **11**: 261-268.

Perry, B. D., Lessard, P., Norval, R. A. I., Kundert, K. and Kruska, R. (1990). Climate, vegetation and the distribution of *Rhipicephalus appendiculatus* in Africa. *Parasitology Today*, **6**: 100-104.

Petty, G. W. (1995). The status of satellite based rainfall estimation over land. *Remote Sensing of Environment*, **51**: 125-137.

Pfirman, E. S. and Hogue, J. (1995). *IDA, image display and analysis, User's guide.* Version 4.1. United States Agency for International Development (USAID) Famine Early Warning System (FEWS) project: Arlington, USA.

Pilson, R. D. and Leggate, B. M. (1962). A diurnal and seasonal study of the resting behaviour of *Glossina pallidipes* Aust. *Bulletin of Entomological Resources*, **53**: 551-562.

Pinty, B. and Verstraete, M. M. (1992). GEMI, a non-linear index to monitor global vegetation from satellites. *Vegetatio*, **101**: 15-20.

Pitcairn, M. J., Rejmankova, E., Wood, B. L. and Washino, R. K. (1988). Progress report on the use of remote sensing data to survey mosquito larval abundance in California rice fields. *Proceedings of the Californian Mosquito and Vector Control Association*, **56**: 158-159.

Platt, R. B., Love, G. J. and Williams, E. L. (1958). A positive correlation between relative humidity and the distribution and abundance of *Aedes vexans*. *Ecology*, **39**: 167-169.

Pope, K. O., Rejmankova, E., Savage, H. M., Arredondo-Jimenez, J. I., Rodriguez, M. H. and Roberts, D. R. (1994). Remote-sensing of tropical wetlands for malaria control in Chiapas, Mexico. *Ecological Applications*, **4**: 81-90.

Pope, K. O., Sheffner, E. J., Linthicum, K. L., Bailey, C. L., Logan, T. M., Kasischke, E. S., Birney, K., Njogu, A. R. and Roberts, C. R. (1992). Identification of central Kenyan Rift Valley fever virus vector habitats with Landsat TM and evaluation of their flooding status with airborne imaging radar. *Remote Sensing of Environment*, **40**: 185-196.

Potts, W. H. (1970). Systematics and identification of the *Glossina*. In *The African Trypanosomiases*, pp. 243-273. George Allen and Unwin Ltd: London, UK.

Power, C. (1995). NOAA 14 is successfully launched. *The Remote Sensing Society Newsletter*, **81**: 2.

Prabhakara, C. and Dalu, G. (1976). Remote sensing of surface emissivity at 9 μm over the globe. *Journal of Geophysical Research*, **81**: 3719-3724.

Prabhakara, C., Dalu, G. and Kunde, V. G. (1974). Estimation of sea surface temperature from remote sensing in the 11- to 13- μm window region. *Journal of Geophysical Research*, **79**: 5039-5044.

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. and Platt, C. M. R. (1991). *Land surface temperature measurement from the AVHRR*. In Proceedings of the 5th AVHRR data users' meeting, pp. 433-438. EUM P 09. European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT): Darmstadt, Germany.

Price, J. C. (1982). On the use of satellite data to infer surface fluxes at meteorological scales. *Journal of Applied Meteorology*, **21**: 1111-1122.

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.

Price, J. C. (1991). Timing of the NOAA afternoon passes. *International Journal of Remote Sensing*, **12**: 193-198.

Price, R. D., King, M. D., Dalton, J. T., Pedelty, K. S., Ardanuy, P. E. and Hobish, M.K. (1994). Earth science data for all: EOS and the EOS data and information system. *Photogrammetric Engineering and Remote Sensing*, **60**: 277-285.

Prince, S. D. and Goward S. N. (1995). Global primary production - a remote-sensing approach. *Journal of Biogeography*, **22**: 815-835.

Prince, S. D. and Goward S. N. (1996). Evaluation of the NOAA/NASA Pathfinder AVHRR Land data set for primary production modelling. *International Journal of remote Sensing*, **17**: 217-221.

Rao, C. R. N. (1993a). *Degradation of the visible and near-infrared channels of the Advanced Very High Resolution Radiometer on the NOAA-9 spacecraft: assessment and recommendations for corrections.* NOAA Technical Report NESDIS-70. National Oceanic and Atmospheric Administration (NOAA) - National Environmental Satellite, Data, and Information Service (NESDIS): Washington DC., USA.

Rao, C. R. N. (1993b). *Non-linearity corrections for the thermal infrared channels of the advanced very high resolution radiometer: assessment and recommendations.* NOAA Technical Report NESDIS-70. National Oceanic and Atmospheric Administration (NOAA) - National Environmental Satellite, Data, and Information Service (NESDIS): Washington DC., USA.

Randolph, S. E. (1993). Climate, satellite imagery and the seasonal abundance of the tick *Rhipicephalus appendiculatus* in southern Africa, a new perspective. *Medical and Veterinary Entomology*, **7**: 243-258.

Randolph, S. E. (1994). Population dynamics and density dependent seasonal mortality indices of the tick *Rhipicephalus appendiculatus* in eastern and southern Africa. *Medical and Veterinary Entomology*, **8**: 351-368.

Randolph, S. E. (1996). Abiotic and biotic determinants of the seasonal dynamics of the tick *Rhipicephalus appendiculatus* in South Africa. *Medical and Veterinary Entomology*, in press.

Rejmankova, E., Rejmanek, M,. Pitcairn, M. J. and Washino, R. K. (1988). Aquatic vegetation in rice fields as a habitat for *Culex tarsalis* and *Anopheles freeborni. Proceedings of the Californian Mosquito Vector Control Association*, **15**: 160-163.

Rejmankova, E., Roberts, D. R., Pawley, A., Manguin, S. and Polanco, J. (1995). Predictions of adult *Anopheles albimanus* densities in villages based on distances to remotely-sensed larval habitats. *American Journal of Tropical Medicine and Hygiene*, **53**: 482-488.

Rejmankova, E., Savage, H. M., Rejmanek, M., Roberts, D. R. and Arredondo-Jimenez., J. I. (1991). Multivariate analysis of relationships between habitats, environmental

factors and occurrence of anophelene mosquito larvae (*Anopheles albimanus* and *A. pseudopuntipennis*) in southern Chiapas, Mexico. *Journal of Applied Ecology*, **28**: 827-841.

Rejmankova, E., Savage, H. M., Rodriguez, M. H., Roberts, D. R. and Rejmanek, M. (1992). Aquatic vegetation as a basis for classification of *Anopheles albimanus* Wiedemann (Diptera: Culicidae) larval habitats. *Environmental Entomology*, **21**: 598-603.

Richardson, A. J. and Weigand, C. L. (1977). Distinguishing vegetation from soil background information. *Photogrammetric Engineering and Remote Sensing*, **43**: 1541-1552.

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.

Riley, J. R. (1989). Remote sensing in entomology. *Annual Review of Entomology*, **34**: 247-271.

Ringrose, S and Matheson, W. (1995). Aspects of the darkening effect and its significance in the analysis of semi-arid vegetation cover. In *RSS95 Remote Sensing in Action. Proceedings of the 21st Annual Conference of the Remote Sensing Society*, pp. 968-975. Remote Sensing Society (RSS): Nottingham, UK.

Roberts, D. R. and Andre, R. G. (1994). Insecticide resistance issues in vector-borne disease control. *American Journal of Tropical Medicine and Hygiene*, **50**: 21-34.

Roberts, D. R. and Rodriguez, M. H. (1994). The environment, remote sensing, and malaria control. *Annals of the New York Academy of Sciences*, **740**: 396-402.

Roberts, D. R., Rodriguez, M., Rejmankova, E., Pope, K., Savage, H., Rodriguez-Ramirez, A., Wood, B., Salute, J. and Legters, L. (1991). Overview of field studies for the application of remote sensing to the study of malaria transmission in Tapachula, Mexico. *Preventive Veterinary Medicine*, **11**: 269-275.

Robinson, T. P. (1996). Spatial and temporal accuracy of coarse resolution products of NOAA - AVHRR NDVI data. *International Journal of Remote Sensing*, **17**: 2303-2321.

Rodriguez, A. D., Rodriguez, M. H., Hernandez, J. E., Dister, S. W., Beck, L. R., Rejmankova, E. and Roberts, D. R. (1996). Landscape surrounding human-settlements and *Anopheles albimanus* (Diptera, Culicidae) abundance in southern Chiapas, Mexico. *Journal of Medical Entomology*, **33**: 39-48.

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

Rogers, D. J. (1985). Trypanosomiasis "risk" or "challenge": a review. *Acta Tropica*, **45**: 5-23.

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

Rogers, D. J. (1988b). The dynamics of vector-transmitted diseases in human communities. *Philosophical Transactions of the Royal Society London B*, **321**: 513-539.

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

Rogers, D. J. (1991). Satellite imagery, tsetse and trypanosomiasis in Africa. *Preventive Veterinary Medicine*, **11**: 201-220.

Rogers, D. J. (1995). Remote sensing and the changing distribution of tsetse flies in Africa. In *Insects in a changing environment.* Edited by R. Harington and N. E. Stark, pp. 176-193. Harcourt Brace and Company Ltd: London, UK.

Rogers, D. J., Hay, S. I. and Packer, M. J. (1996). Predicting the distribution of tsetse flies in West Africa using temporal Fourier processed meteorological satellite data. *Annals of Tropical Medicine and Parasitology*, **90**: 225-241.

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

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

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

Rogers, D. J. and Randolph, S. E. (1988). Tsetse flies in Africa: bane or boon? *Conservation Biology*, **2**: 57-65.

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

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

Rogers, D. J. and Williams, B. G. (1993). Monitoring trypanosomiasis in space and time. *Parasitology (Supplement)*, **106**: 77-92.

Rogers, D. J. and 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.

Rosema, A. (1990). Comparison of Meteosat - based rainfall and evapotranspiration mapping in the Sahel region. *International Journal of Remote Sensing*, **11**: 2299-2309.

Roubaud, E. (1909). Recherches biologiques sur les conditions de viviparité et de larvaire de *Glossina palpalis* R. Desv. *Comptes Rendus de L'Academie des Sciences*, **148**: 195-197.

Running, S. W., Justice, C. O., Salmonson, V., Hall, D., Barker, J., Kaufman, Y. J., Strahler, A. H., Huete, A. R., Muller, J. P., Vanderbilt, V., Wan, Z. N., Teillet, P. and Carneggie, D. (1994). Terrestrial remote sensing science and algorithms planned for EOS/MODIS. *International Journal of Remote Sensing*, **15**: 3587-3620.

Rush, M. and Vernon, S. (1975). Remote sensing and urban public health. *Photogrammetric Engineering and Remote Sensing*, **41**: 1149-1155.

Sabins, F. F. (1987). *Remote sensing: principle and interpretation.* 2nd Edition. W. H. Freeman and Company: New York, USA.

Salisbury, J. W. and Daria, D. M. (1992). Emissivity of terrestrial materials in the 8 14 mm atmospheric window. *Remote Sensing of Environment*, **42**: 83-106.

Scofield, A. and Oliver, V. J. (1977). *A scheme for estimating convective rainfall from satellite imagery.* NOAA Technical memorandum NESS 86. National Oceanic and Atmospheric Administration (NOAA) - National Environmental Satellite, Data, and Information Service (NESDIS): Washington DC, USA.

Schultz, P. A. and Halpert, M. S. (1993). Global correlation of temperature, NDVI and precipitation. *Advances in Space Research*, **13**: 5277-5280.

Schmetz, J., Klaes, D. and Rattenborg, M. (1995). Meteorological products from current and future EUMETSAT satellite systems. In *Proceedings of the 1995 Meteorological Satellite Data Users' Conference*, pp. 233-240. EUM P 17. EUMETSAT: Darmstadt, Germany.

Sellers, P. J. (1985). Canopy reflectance, photosynthesis and transpiration. *International Journal of Remote Sensing*, **6**: 1335-1372.

Service, M. W. (1978). The effect of weather on mosquito biology. In *Weather and parasitic animal disease*, pp. 151-157. Technical note No. 159. World Meteorological Organisation: Geneva, Switzerland.

Service, M. W. (1989). Rice, a challenge to health. *Parasitology Today*, **5**: 11-12.

Shope, R. E. (1980). Arbovirus-related encephalitis. *Yale Journal of Biological Medicine*, **53**: 93-99.

Smith, G. R., Levin, R. H., Abel, P. and Jacobowitz, H. (1988). Calibration of the solar channels of NOAA-9 AVHRR using high altitude aircraft measurements. *Journal of Atmospheric and Oceanic Technology*, **5**: 631-639.

Smith, W. L. (1966). Note on the relationship between total precipitable water and the surface dew point. *Journal of Applied Meteorology*, **5**: 726-727.

Snedecor, G. M. and Cochran, W. G. (1967). *Statistical methods.* Sixth edition. The Iowa State University Press: Iowa, USA.

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.

Snijders, F. L. (1995). *Personal communication.* Remote Sensing Officer (Monitoring), Food and Agriculture Organisation, Rome.

Snyder, J. P. (1987). *Map projections a working manual.* U.S. Geological Survey professional paper 1395. United States Government Printing Office: Washington DC., USA.

Sonenshine, D. E. (1991). Tick Life cycles. In *Biology of Ticks*, Volume 1, pp. 51-66. Oxford University Press: Oxford, UK.

Sonenshine, D. E. (1993a). Ecology of non-nidocolous ticks. In *Biology of Ticks*, Volume 2, pp. 3-65. Oxford University Press: Oxford, UK.

Sonenshine, D. E. (1993b). Tick-borne and tick-caused diseases. Part 4. In *Biology of Ticks*, Volume 2, pp. 107-412. Oxford University Press: Oxford, UK.

Steelman, C. D. (1976). Effects of external and internal arthropod parasites on domestic livestock production. *Annual Review of Entomology*, **21**: 155-178.

Steinwand, D. R., Hutchinson, J. A., and Snyder, J. P. (1992). *Map projections from global and continental data sets, and an analysis of distortion caused by reprojection.* United States Geological Survey (USGS) Earth Resources Observation Systems (EROS) Data Centre contract report. EDC: Sioux Falls, USA.

Steinwand, D. R. (1994). Mapping raster imagery to the Interrupted Goode Homolosine Projection. *International Journal of Remote Sensing*, **15**: 3463-3471.

Stowe, L. L., McClain, E. P., Carey, R., Pellegrino, P., Gutman, G. G., Davis, P., Long, C. and Hart S. (1991). Global distribution of cloud cover derived from NOAA/AVHRR operational satellite data. *Advances in Space Research*, **3**: 51–54.

Stowe, L. L., Carey, R. M. and Pellegrino, P. P. (1992). Monitoring the Mt-Pinatubo aerosol layer with NOAA-11 AVHRR data. *Geophysical Research Letters*, **19**: 159-162

Sugita, M. and 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.

Sutherst, R. W. and Maywald, G. F. (1985). A computerised system for matching climates in ecology. *Agriculture, Ecosystems and Environment*, **13**: 281-299.

Swain, P. H. and Davis, S. M. (1978). *Remote sensing: the quantitative approach.* McGraw Hill International Book Company: New York, USA.

Swynnerton, C. F. M. (1936). The tsetse flies of East Africa. A first study on their ecology, with a view to their control. *Transactions of the Royal Entomological Society*, **84**: 1-579.

Takken, W. (1984). Studies on the biconical trap as a sampling device for tsetse (Diptera: *Glossinidae*) in Mozambique. *Insect Science and its Application*, **5**: 357-361.

Tanre, D., Holben, B. N. and Kaufman, Y. (1992). Atmospheric correction algorithm for NOAA-AVHRR products: theory and application. *IEEE Transactions on Geoscience and Remote Sensing*, **30**: 231-248.

Tatsuoka, M. M. (1971). *Multivariate Analysis: techniques for educational and psychological research.* John Wiley and Sons: New York, USA.

Ter Braak, C. J. F. (1986). Canonical correspondence-analysis a new eigenvector technique for multivariate direct gradient analysis. *Ecology*, 67: 1167-1179.

Theon, J. S. (1993). The Tropical Rainfall Measuring Mission (TRMM). *Advances in Space Research*, **14**: 159-165.

Thomson, M. C., Adiamah, J. H., Connor, S. J., Jawara, M., Bennett, S., D'Alessandro, U., Quinones, M., Langerock, P. and Greenwood, B. M. (1995a). Entomological evaluation of The Gambia's national impregnated bednet programme. *Annals of Tropical Medicine and Parasitology*, **89**: 229-241.

Thomson, M. C., Connor, S. J., Milligan, P., Flasse, S. and Service, M. W. (1995b). The potential in using environmental monitoring satellites to predict malaria epidemics in Africa. In *Proceedings of the 1995 Meteorological satellite data users' conference*, pp. 461-472. EUM P 17. European Organisation for the Exploitation of Meteorological Satellites (EUMETSAT): Darmstadt, Germany.

Thomson, M. C., Connor, S. J., Milligan, P. J. M. and Flasse, S. P. (1996). The ecology of malaria - as seen from Earth observation satellites. *Annals of Tropical Medicine and Parasitology*, **90**: 243-264.

Tom, C. H. and Miller, L. D. (1984). An automated land-use mapping comparison of the Bayesian maximum likelihood and linear discriminant analysis algorithms. *Photogrammetric Engineering and Remote Sensing*, **50**: 193-207.

Townhsend, J. R. G. (1994). Global data sets for land applications from the Advanced Very High Resolution Radiometer: an introduction. *International Journal of Remote Sensing*, **15**: 3319-3332.

Townshend, J. R. G., Goff, T. E. and Tucker, C. J. (1985). Multitemporal dimensionality of images of normalized difference vegetation index at continental scales. *IEEE Transactions on Geoscience and Remote Sensing*, **23**: 888-895

Townshend, J. R. G., Justice, C. O. and Kalb, V. (1987). Characterization and classification of South American land cover types using satellite data. *International Journal of Remote Sensing*, **8**: 1189-1207.

Townshend, J. G. R. and 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. and Tucker, C. J. (1984). Objective assessment of Advanced Very High Resolution Radiometer data for land cover mapping. *International Journal of Remote Sensing*, **5**: 497-504.

Trodd, N. M. (1995). Uncertainty in land cover mapping for monitoring land cover change. In *RSS95 Remote Sensing in Action. Proceedings of the 21st Annual Conference of the Remote Sensing Society*, pp. 1138-1145. Remote Sensing Society (RSS): Nottingham, UK.

Tucker, C. J. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment*, 8: 127-150.

Tucker, C. J., Holben, B. N. and Goff, T. E. (1984). Intensive forest clearing in Rondonia, Brazil, as detected by satellite remote sensing. *Remote sensing of Environment*, **15**: 255-261.

Tucker, C. J., Newcomb, W. W., Los, S. O. and Prince, S. D. (1991). Mean and inter-year variation of growing-season normalized difference vegetation index or the Sahel 1981-1989. *Remote Sensing of Environment*, 8: 127-150.

Tucker, C. J. and Sellers, P. J. (1986). Satellite remote sensing of primary production. *International Journal of Remote Sensing*, **7**: 1395-1416.

Tucker, C. J., Townshend, J. R. G. and Goff, T. E. (1985a). African land-cover classification using satellite data. *Science*, **227**: 369-375.

Tucker, C. J., Vanpraet, C., Boerwinkel, E. and Gaston, A. (1983). Satellite remote sensing of total dry matter production in the Senegalese Sahel. *Remote Sensing of Environment*, **13**: 461-474.

Tucker, C. J., Vanpraet, C. L., Sharman, M. J. and Van Ittersum, G. (1985b). Satellite remote sensing of total herbaceous biomass production in the Senegalese Sahel: 1980-1984. *Remote Sensing of Environment*, **17**: 233-249.

Unwin, D. M. (1980). *Microclimate measurement for ecologists.* Biological techniques series 3. Academic Press: London, UK.

Van de Griend, A. A. and 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.

Verhoef, W., Menenti, M. and Azzali, S. (1996). A color composite of NOAA-AVHRR - NDVI based on time-series analysis (1981-1992). *International Journal of Remote Sensing*, **17**: 231-235.

Vermote, E., Tanre, D. and Herman, M. (1990). Atmospheric effects on satellite imagery, correction algorithms for ocean color or vegetation monitoring. *International Society for Photogrammetry and Remote Sensing*, 28: 46-55.

Verstraete, M. M. (1995). *Personal communication.* Team Leader, Monitoring of Tropical Vegetation Unit, Space Applications Institute, Joint Research Centre (European Commission): Ispra (VA), Italy.

Viovy, N., Arino, O. and Belward, A. S. (1992). The best index extraction slope (BISE): a method for reducing noise in NDVI time-series. *International Journal of Remote Sensing*, **13**: 1585-1590.

Vogt, J. (1992). *Characterising the spatio-temporal variability of surface parameters from NOAA AVHRR data: a case study for southern Mali.* Ph.D. Thesis. Institute for Remote Sensing Applications, Joint Research Centre: Ispra, Italy.

Wagner, V. E., Hill-Rowley, R., Narlock, S. A. and Newson, H. D. (1979). Remote sensing, a rapid and accurate method of data acquisition for a newly formed mosquito control district. *Mosquito News*, **39**: 283-287.

Washino, R. K. and Wood, B. L. (1994). Application of remote sensing to arthropod vector surveillance and control. *American Journal of Tropical Medicine and Hygiene (Supplement)*, **50**: 134-144.

Weinreb, M. P., Hamilton, G., Brown, S. and Koczor, R. J. (1990). Nonlinearity corrections in calibration of advanced very high-resolution radiometer infrared channels. *Journal of Geophysical Research (Oceans)*, **95**: 7381-7388.

Welch, J. B., Olson, J. K., Hart, W. D., Ingle, S. G. and Davis, M. R. (1989a). Use of aerial colour infrared photography as a survey technique for *Psorophora columbiae*

oviposition habitats in Texas ricelands. *Journal of the American Mosquito Control Association*, **5**: 147-160.

Welch, J. B., Olson, J. K., Yates, M. M, Benton, A. R. JR. and Baker, R. D. (1989b). Conceptual model for the use of aerial color infrared photography by mosquito control districts as a survey technique for *Psorophora columbiae* oviposition habitats in Texas ricelands. *Journal of the American Mosquito Control Association*, **5**: 369-373.

Whitfield, P. J. (1993). Parasitic Helminths. In *Modern Parasitology: a Textbook of Parasitology*, 2nd edition. Editor F. E. G. Cox, pp. 24-52. Blackwell Scientific Publications: Oxford, UK.

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.

Wilson, M. L. (1994). Rift Valley fever virus ecology and the epidemiology of disease emergence. *Annals of the New York Academy of Sciences*, **740**: 169-180.

Williams, B. (1993). Biostatistics - concepts and applications for biologists. Chapman and Hall: London, UK.

Williams, B., Rogers, D. J., Staton, G., Ripley, B. and Booth, T. (1992). Statistical modelling of georeferenced data: mapping tsetse distribution in Zimbabwe using climate and vegetation data. In *Modelling vector-borne and other parasitic diseases*, pp. 267-280. Proceedings of a workshop organised by International Laboratory for the Research into Animal Diseases (ILRAD) in collaboration with Food and Agriculture Organization (FAO) of the United Nations. ILRAD: Nairobi, Kenya.

Wooster, M. J., Sear, C. B., Patterson, G. and Haigh, J. (1994). Tropical lake surface temperatures from locally received NOAA-11 AVHRR data - comparison with *in-situ* measurements. *International Journal of Remote Sensing*, **15**: 183-189.

World Health Organisation. (1986). *Epidemiology and control of African trypanosomiasis. Report of a WHO Expert Committee, Technical Report Series 739.* WHO: Geneva, Switzerland.

World Health Organisation. (1994). World malaria situation in 1992. Part I. *Weekly Epidemiological Record*, **69**: 309-314

World Meteorological Organisation. (1971). *Guide to meteorological instruments and observing practices.* 4th Edition. WMO - No. 8.TP.3. WMO: Geneva, Switzerland.

World Meteorological Organisation. (1994). *Information on meteorological and other environmental satellites.* 3rd Edition. WMO - No. 411. WMO: Geneva, Switzerland.

Wood, B. L., Beck, L. R., Dister, S. W. and Spanner, M. A. (1994). The Global Monitoring and Disease Prediction Program. *Sistema Terra*, **3**: 38-39.

Wood, B. L., Beck, L. R., Lawless, J. G. and Vesecky, J. F. (1992a). Preliminary considerations for a small satellite to monitor environmental change associated with vector-borne disease. *Journal of Imaging Science and Technology*, **36**: 431-439.

Wood, B. L., Beck, L. R., Washino, R. K., Hibbard, K. A. and Salute, J. S. (1992b). Estimating high mosquito-producing rice fields using spectral and spatial data. *International Journal of Remote Sensing*, **13**: 2813-2826.

Wood, B. L., Beck, L. R., Washino, R. K., Palchick, S. M. and Sebesta, P. D. (1991a). Spectral and spatial characterization of rice field mosquito habitat. *International Journal of Remote Sensing*, **12**: 621-626.

Wood, B., Washino, R., Beck, L., Hibbard, K., Pitcairn, M., Roberts, D., Rejmankova, E., Paris, J., Hacker, C., Salute, J., Sebesta, P. and Legters, L. (1991b). Distinguishing high and low anopheline-producing rice fields using remote sensing and GIS technologies. *Preventive Veterinary Medicine*, **11**: 277-288.

Appendix I

NOTES ON AUTHORSHIP

At the time of submission material presented in this thesis has been published or is in press. Since all of this published information is in co-authored papers I should like to make clear the parts that are not exclusively my work. An accompanying statement from my supervisors confirms that the greater part of the work in the primary authored papers was initiated and performed by myself and that I had a significant intellectual and material input into the two papers on which I am second author. Those papers that have been published form appendix II.

1. Hay, S. I., Packer, M. J. and Rogers, D. J. (1996). A review of the impact of remote sensing on the study and control of invertebrate intermediate hosts and vectors for disease. *International Journal of Remote Sensing*, in press.
2. Hay, S. I., Tucker, C. J., Rogers, D. J. and Packer, M. J. (1996). Remotely sensed surrogates of meteorological data for the study of the distribution and abundance of arthropod vectors of disease. *Annals of Tropical Medicine and Parasitology*, **90**(1): 1-19.
3. Rogers, D. J., Hay, S. I., Wint, G. R. W. and Packer, M. J. (1996). Mapping land-cover of large areas using public-domain meteorological satellite data: a case study of Nigeria. *International Journal of Remote sensing*, Accepted.
4. Rogers, D. J., Hay, S. I. and Packer, M. J. (1996). Predicting the distribution of tsetse flies in West Africa using temporal Fourier-processed meteorological satellite data. *Annals of Tropical Medicine and Parasitology*, **90**(3): 225-241.
5. Hay, S. I. (1997). Remote sensing and disease control: past, present and future. *Transactions of the Royal Society of Tropical Medicine and Hygiene*, in press.

Appendix 2 originally presented here cannot currently be made freely available via ORA due to Copyright considerations.

LIST OF ACRONYMS

Acronym	*Meaning*
ABU	Atmospheric Boundary Layer
ANUDEM	Australian National University Digital Elevation Model
ARTEMIS	African Real Time Environmental Monitoring using Imaging Satellites
AVHRR	Advanced Very High Resolution Radiometer
BISE	Best Index Slope Extraction
CCD	Cold Cloud Duration
CLAVR	Clouds from AVHRR
CSIRO	Commonwealth Scientific and Industrial Research Organisation
DAAC	Distributed Active Archive Center
DDS	Digital Data Storage
DEM	Digital Elevation Model
DI-MOD	Disease Modelling
DJR	David John Rogers
DMA	Defence Mapping Agency
ECF	East Coast Fever
EDC	EROS Data Centre
EOS	Earth Observing Satellite
EOSDIS	Earth Observing System Data and Information Service
EMR	Electromagnetic Radiation
EMS	Electromagnetic Spectrum
ERDAS	Earth Resources Data Analysis Software
ERGO	Environmental Research Group Oxford Ltd
EROS	Earth Resources Observation Systems
ETOPO5	Earth Topograhic Five Minute Grid
EUMETSAT	European Organisation for the Exploitation of Meteorological Satellites
FAO	Food and Agriculture Organisation
FEWS	Famine Early Warning System
GAC	Global Area Coverage

GCP	Ground Control Point
GEMI	Global Environmental Monitoring Index
GIMMS	Global Inventory Monitoring and Modelling Systems
GIS	Geographic Information System
GLIS	Global Land Information System
GLO-PEM	Global Production Efficiency Model
GPS	Global Positioning System
GSFC	Goddard Space Flight Centre
GTZ	Deutsche Gesellschaft für Technische Zusammenarbeit
HAO	Health Applications Office
HDF	Hierarchical Data Format
HRPT	High Resolution Picture Transmission
HRR	High Resolution Radiometer
HRV	High Resolution Visible
ICIPE	International Centre for Insect Physiology and Ecology
IDA	Image Display and Analysis
IEMVT	Institut d'Élevage et de Médecine Vétérinaire des pays Tropicaux
IFOV	Instantaneous Field of View
ILCA	International Livestock Centre for Africa
ILRAD	International Laboratory for the Research into Animal Diseases
ISCTRC	International Scientific Council for Trypanosomiasis Research and Control
ITCZ	Inter Tropical Convergence Zone
LAC	Local Area Coverage
MI	Moisture Index
MIR	Middle Infrared
MODIS	Moderate Resolution Imaging Spectroradiometer
MSG	Meteosat Second Generation
MSS	Multi Spectral Scanner
MVC	Maximum Value Composit
NASA	National Aeronautics and Space Administration
NCDC	National Climate Data Centre
NDVI	Normalised Difference Vegetation Index
NESDIS	National Environmental Satellite, Data, and Information Service
NIR	Near Infrared

NOAA	National Oceanic and Atmospheric Administration
NOMCD	New Orleans Mosquito Control District
NLR	National Aerospace Laboratory (Nationaal Lucht- en Ruimtevaartlaboratorium)
NRI	Natural Resources Institute
OAU	Organisation of African Unity
ONC	Operational Navigation Chart
ORSTOM	Office de la Recherche Scientific et Technique Outre-Mer
PAL	Pathfinder AVHRR Land
PCA	Principal Components Analysis
PVI	Perpendicular Vegetation Index
RIM	Resource Inventory and Management Ltd
RSS	Remote Sensing Society
RVF	Rift Valley Fever
RVI	Ratio Vegetation Index
SAR	Synthetic Aperture Radar
SAVI	Soil Adjusted Vegetation Index
SCA	Scan Angle
SEVIRI	Spinning Enhanced Visible and Infra-Red Imager
SI	Systeme International d'Unites
SLE	Saint Louis Encephalitis
SPOT	Satellite Pour l'Observation de la Terre
SPSS	Statistical Package for the Social Sciences
SRI	Simple Ratio Index
STRC	Scientific and Technical Research Commission
SVI	Spectral Vegetation Index
SZA	Solar Zenith Angle
TAMSAT	Tropical Applications of Meteorological Satellite and Other Data
TIR	Thermal Infrared
TIROS	Television InfraRed Observing Satellite
TM	Thematic Mapper
TOMS	Total Ozone Mapping Spectrometer
TOVS	TIROS Operational Vertical Sounder
TRMM	Tropical rainfall monitoring mission
UN	United Nations

UNESCO	United Nations Educational, Scientific and Cultural Organisation
UNICEF	United Nations International Childrens Education Fund
USAID	United States Agency for International Development
USGS	United States Geological Survey
UV	Ultraviolet
WHO	World Health Organisation of the United Nations
WMO	World Meteorological Organisation of the United Nations