Cranfield
UNIVERSITY
Silsoe

PhD Thesis
Academic year 1999-2000

Christophe Sannier

Strategic monitoring of crop yields and rangeland conditions in Southern Africa with Remote Sensing

Supervisor:
J.C. Taylor

January 2000

This thesis is submitted in fulfilment of the requirements of the Degree of Doctor of Philosophy

ABSTRACT

The monitoring of vegetation resources is of vital importance for Southern African countries because of the dominance of agriculture in the economy. The use of remote sensing techniques in a national or local planning context is particularly adapted to the Southern African conditions because large areas can be covered regularly with minimal requirements for field based infrastructure. Furthermore, relatively low-cost receiving stations have been installed in the meteorological departments and other local institutions, which makes satellite data available in real-time. Real-time acquisition is essential for the operational monitoring of vegetation development and remote sensing plays a significant role in three main areas:

- Inventory or mapping of cover types
- Monitoring of vegetation conditions relative to the norm
- Estimates of biomass

In this research operational techniques were developed in each of these areas with the participation and involvement of users.

Remote sensing and field survey techniques for inventory and mapping of cover types were adapted and developed from existing experience in the European context to match the requirements in Southern Africa. The need for an unbiased sample of field observations, for the calibration of digital classification of satellite imagery was identified and methodology for its collection demonstrated. Methods developed for the inventory of crop types in Europe were successfully adapted to the African rangeland. The levels of classification accuracy achieved were similar to that obtained in the European context for a classification scheme of equivalent complexity.

A Vegetation Productivity Indicator (VPI) was developed for monitoring vegetation conditions based on real-time acquisition of NOAA HRPT imagery from a local receiving station and a historical Normalised Difference Vegetation Index (NDVI) archive. The VPI maps show departure from normal vegetation response using methodology similar to the analysis of extreme events in hydrology, in near real-time. The method was successfully implemented in Zambia to monitor maize production and in Namibia to monitor rangeland. The VPI was significantly correlated with rainfall. The technique was successfully transferred to the Department of Meteorological Services in Botswana where VPI maps are produced routinely and presented to the inter-ministerial drought committee for assessing rangeland conditions.

Methodology for rapid biomass assessment was developed using simple physiognomic plant parameters. Field Measurements were taken in four different cover types (grassland, steppe and shrub and tree savanna) and correlated with the NDVI derived from the satellite observations in Etosha National Park, Namibia in near real-time. The pooled regression relationship which was obtained was highly statistically significant. However, the regression model excluding the two savanna types exhibited a higher correlation suggesting that there might be a separate relationship between savanna biomass and NDVI. Biomass maps were produced using the pooled relationship and their potential for operational targeting of areas suitable for prescribed burning was illustrated.

Although the methods and techniques in this work were developed using time series of NOAA-AVHRR and the NDVI, they can all be adapted to include data from new sensors systems and other vegetation indices as they become available. Methods demonstrated in this work can be integrated to form a suitable framework for a national vegetation resources monitoring system. This would assist Southern African governments in making decisions related to vegetation resources by providing sound and timely technical advice.

ACKNOWLEDGEMENTS

The work presented in this thesis could not have been undertaken without the sponsorship of the Overseas Development Administration, which has since become the Department for International Development. The majority of the work was undertaken as part of the Africa Regional Remote Sensing Project and Dr Chris Sear who co-ordinated the project at the Natural Resources Institute should be acknowledged for allowing the work carried out for the project to be part of my research. Part of the work, particularly the study presented in the appendix was also sponsored by the European Commission.

Colleagues at Cranfield University at Silsoe should also be thanked for their support and I would like especially to express gratitude to Dr Graham Thomas and Professor Roy Morgan for their help and advice. I also wish to thank my supervisor Professor John Taylor for his support during the past few years.

The contribution of the institutions collaborating to this work in southern Africa should also be recognised. These were the Ministry of Environment and Tourism in Namibia through the Etosha Ecological Institute, The Zambian Meteorological Department, the Department of Meteorological Services in Botswana and the Ministry of Agriculture of Botswana. I would especially like to thank staff at the Etosha Ecological Institute who took a special interest in my work and the assistance they provided in undertaking my fieldwork was well appreciated.

I would like to express my thanks to Ursula and my relatives for their patience and their moral support. Finally, my gratitude goes to my father, Bernard, for his encouragement and useful advice throughout the years, but who regrettably could not see this work reaching its completion.

TABLE OF CONTENTS

CHAPTER 1 INTRODUCTION

1 The importance of vegetation resources in Southern Africa

The monitoring of vegetation resources is crucial for southern African countries. Most of their economies are heavily dependent on the primary sector as a source of income. Although mining is the largest source of revenue in some Southern African countries, crop and livestock production often comes in second place. In addition, agriculture employs by far the largest proportion of the population (Moyo et al. 1993).

In general, the climate in Southern Africa is arid to semi-arid. This means that often, only a small part of the country is suitable for agriculture. For instance, only 6% of Botswana is classified as land with arable potential (Moyo et al. 1993), which means that in most of the country, the only possible farming activity is extensive livestock production. Furthermore, rainfall variations are very high and crop (in the absence of irrigation) and rangeland vegetation growth is limited by rainfall availability. This has an effect on the agricultural production, which is illustrated in Figure 1, showing the maize production of Zambia. From one year to the next, maize production can vary by almost a factor of 2, which poses a serious problem in terms of food supply.

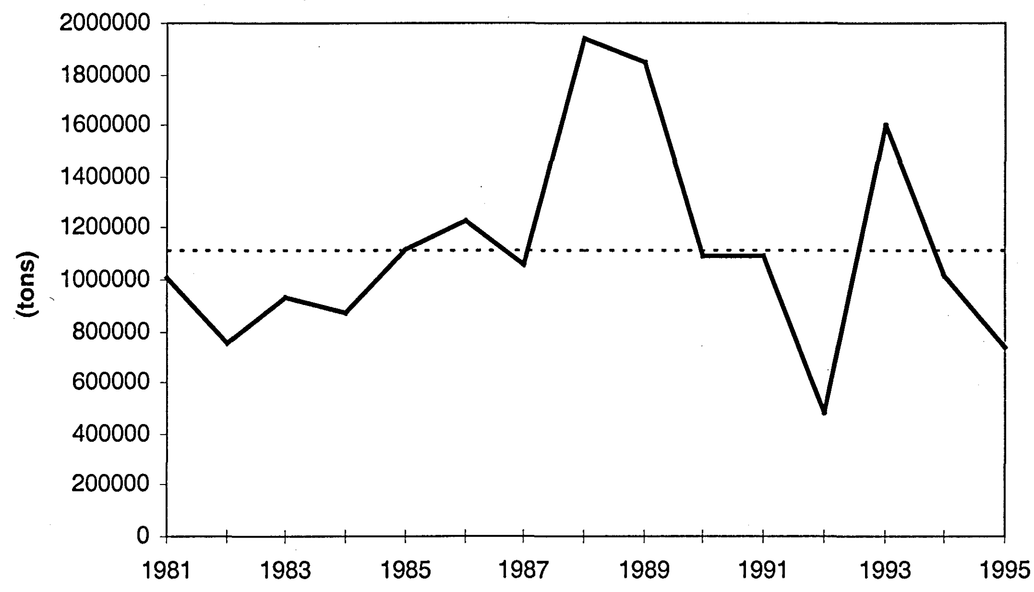


Figure 1. Annual variation in maize production in Zambia (———) and average production between the 1981-92 period (-------).

These variations are also observed for rangeland vegetation in terms of its annual net primary production. Because the animal population is not immediately adjusted to reflect these variations, a major concern in Southern Africa has been land degradation caused by overgrazing. According to the FAO report on the state of food and agriculture (FAO 1997), the main causes of soil degradation in Africa are overgrazing followed by arable farming and deforestation. Scherr and Yadav (1997) add that nutrient depletion, salinisation and soil erosion are some factors induced from the above causes.

Furthermore, Southern African countries face difficulties in keeping up with the rise in the needs for food supply generated by the continuous and rapid growth of the population. As a consequence, many countries have struggled to be self sufficient (FAO 1996). This has certainly been a major factor in the increased pressure on land which was witnessed during the past two decades. This has resulted in the degradation of land which poses serious threat to the sustainability of agriculture as it is implemented in most parts of Africa.

Nevertheless, Southern African countries still have some of the largest unspoilt natural areas in the world. With the increased concern about the environment, it is important to preserve these rich natural ecosystems. But it is also an opportunity for Southern African countries to diversify their economies through the development of tourism.

Therefore, it is vital for Southern African governments to try to create a balance between the sustainable development of agriculture and the preservation of natural ecosystems.

2 The need for vegetation monitoring: role of satellite remote sensing

The implementation of a suitable system for monitoring progress in achieving the balance between sustainable development and nature preservation is crucial if the impact of management strategies is to be measured. This can then contribute to the formulation of new management strategies. The conventional approach to the monitoring of natural resources is one that has been implemented over a number of years in the developed world and which is based on a network of governmental field offices located within each district. The lack of infrastructure and resources in most African countries has meant that the development of conventional monitoring systems is not sustainable. For example, in Botswana, the country is divided into agricultural districts with a field office for each district. Because of the country's low population density, each agricultural district covers an area too large to be monitored regularly on the ground from the field office. Instead, an approach based on a limited amount of field observations combined with remotely sensed data is likely to provide an appropriate solution. Furthermore, today's changes are very rapid and existing conventional systems in the developed world are now often questioned as to their ability to predict adverse situations.

The institutional context is also particularly important to ensure the sustainability of the vegetation monitoring system put in place. National meteorological departments usually have the combination of technical capabilities and institutional linkages at national levels. However, more detailed analysis is possible at sub-national level. This is normally the case in Southern African national parks, which have the requirement and a good level of technical abilities and resources to implement such a system.

Lambin et al. (1993) reviewed the various levels at which remotely sensed data can be used for monitoring agricultural production in the Sahel. However, some of their findings can be extended to environmental monitoring and to most parts of Africa. Their conclusions showed that remotely sensed data could be effectively used but that the following factors need to be addressed:

(i) the network of ground calibration sites needed to be increased

(ii) the telecommunication infrastructure should be improved

(iii) projects for the development of national or regional soil and topographic mapping in digital format should be increased

(iv) greater co-operation between researchers should be encouraged

These observations are also valid for the development of a famine early warning system. Hutchinson (1991) outlined the basic requirements for a famine early warning system and identified the role of satellite-derived data.

Apart from traditional meteorological applications such as rainfall estimates, the estimation of cropped areas, the monitoring of vegetation conditions and the assessment of forage and crop yields were identified as possible areas of applications for remote sensing. This thesis presents the results of research aimed at the development of robust methods to fulfil some of these tasks with the use of remote sensing technology. The sections below will look in turn at the background and requirements for some of the areas of application of satellite-derived data identified by Hutchinson (1991), which form the main researchable objectives of this thesis.

2.1. *Inventory of natural resources: Land cover mapping*

Before being able to monitor vegetation dynamics, it is important to know the initial conditions by carrying out a detailed inventory of land resources. This type of work requires many human and technical resources and consequently tends to be carried out every decade at best. High-resolution satellite imagery such as SPOT, Landsat or IRS normally provides the best compromise between spatial resolution and coverage. Aerial photography is usually too costly and lower resolution satellite imagery does not allow the accurate separation of cover types. Traditionally, the analysis of such imagery in Africa has been undertaken with very little input from field observations. Most of the time, the reasons put forward are the lack of access and the difficulty to locate features on the ground. As outlined in Taylor et al. (1997) and presented in the appendix, the value of a well designed survey is crucial as it will permit the accurate and unbiased measurement of the performance of satellite derived thematic maps and make reliable area estimates of cover types. The technique most widely employed is that of area frame sampling which has been used extensively in the USA (Cotter and Nealon 1987, Allen and Hanuschak 1988) and Europe (Taylor et al. 1997). Lambin et al. (1993) have expressed some concerns about the success of implementing area frame samples in Africa. However, they were specifically referring to the identification of crop types in farming systems where the small size of fields and mixed cropping practices does not allow the application of traditional area frame surveys. In cases where land cover mapping refers to the mapping of agricultural systems and natural vegetation, area frame sample surveys are still highly suited for the calibration and assessment of satellite derived classification. Taylor et al. (1997) present a synthesis of the results of the application of area frame sampling in the European Union under the MARS (Monitoring Agriculture with Remote Sensing) programme. The main objective of this part of the research was to test whether the MARS methods could be used in an African context. Chapter 2 of this thesis explores the application of the MARS methodology in two regions of Africa, Northern Nigeria and Etosha National Park in Namibia.

2.2. *Early warning systems: measurement of vegetation conditions*

Once the various cover types over the area under scrutiny are known, it is possible to monitor their evolution through time. One major advance over the past decade in terms of the availability of remotely sensed data has been the development of global and continental scale satellite image archives. The main driving forces behind the increased use of such archives has been the study of global change and the implementation of satellite based early warning systems. The purpose of the archive in the case of the early warning systems is to have a reference data set which is compared with current imagery. This then allows an evaluation of the vegetation conditions. One of the first early warning systems was developed by FAO with the Global Information and Early Warning System (GIEWS) which started in 1973 (Hutchinson 1991). Remotely sensed data is incorporated in GIEWS through the support of the Africa Real Time Environmental Monitoring Information System (ARTEMIS) which is in operation since 1988 (Hielkema and Snijders 1994). Most African countries now have such a system at their disposal, which is supported at various degrees by international aid organisations. The complexity of Remote Sensing derived products range from a purely qualitative use of the imagery to satellite based estimates of vegetation activities and/or rainfall.

Overall, vegetation based products are fairly rudimentary. Usually, they consist of an interpretation of the level of vegetation activity based on satellite derived vegetation index images. These products can be the result of the interpretation of a single image or the interpretation of an image resulting from the difference between a current vegetation index image and a previous one or a mean image for the time of year considered. To strengthen the usefulness of remote sensing in this field, there is a need to improve the robustness and reliability of these methods and also to make the satellite derived products more accessible to the users. Decision-makers will make better use of these outputs if they can be easily interpreted without requiring expertise in the field of remote sensing. Chapter 3 investigates a novel approach for the development of such products and assesses their reliability in an operational context.

2.3. Determining vegetation quantity

The assessment of crop yields or biomass quantity in general is a primary application of remote sensing for food security and environmental monitoring. Knowing crop yields combined with the knowledge of crop areas allows estimates of production to be made. This is of particular interest to decision-makers as the estimate is available several weeks before harvest and therefore it allows some time to take appropriate actions. However, most attempts to relate vegetation quantity or crop yields to satellite imagery have been of an experimental nature.

Most applications for biomass monitoring in Africa require the use of NOAA-AVHRR (National Oceanographic and Atmospheric Administration - Advanced Very High-Resolution Radiometer) imagery for reasons discussed in the next section. The calibration of the imagery in terms of vegetation quantity is most often achieved by making field measurements, which are then related to satellite observations using statistical methods. Because of the low resolution of AVHRR, 1.2km^2 at nadir, site selection and sampling strategy within the site must be carefully defined to ensure that measurements are a true reflection of the whole site. The lack of planning in the calibration survey is one of the main reasons why many calibration exercises have not been fully successful. Despite these problems, many accounts of biomass calibration with NOAA-AVHRR have been reported in the literature (Prince and Astle 1986, Prince and Tucker 1986, Kennedy 1989, Diallo et al. 1991, Franklin and Hiernaux 1991 and Prince 1991a). This is because of the wide range of potential operational applications it can provide. Measurements of biomass levels over wide areas can be used to produce fire risk maps. Areas with high risk can be targeted for controlled burning, thus preventing biomass to build up and larger fires to occur. Satellite derived biomass maps can also be used as an input to models for estimating carrying capacity for animals on ranges or natural reserves. These investigations demonstrated that AVHRR could be effectively calibrated with levels of green biomass. A step further towards operational application was made by Diallo et al. (1991) where NOAA-AVHRR imagery was combined with ground observations for around 30 sites throughout Senegal to produce biomass maps. The resultant maps are based on integrated NDVI, which means that they can only be produced when the growing season is completed. There is certainly a need to monitor biomass changes in real time and Chapter 6 describes how methods were developed to assess biomass rapidly in the field over suitable sites and how these observations were successfully related to contemporary AVHRR images.

In estimating crop yields, the situation is somewhat different. Again, many studies undertaken in Africa (Azzali 1991, Rasmussen 1992, Groten 1993 and Maselli et al. 1993) have shown that there is a relationship between crop yields and AVHRR, providing that the crops covered a significant area and that crop yields were strongly correlated with natural vegetation activity. This is because, in Africa, crops are mostly cultivated in small parcels amongst large areas of semi-natural vegetation. However, yield is only part of the equation and what decision-makers really want is an early estimate of the total production. If yield is known, this requires the estimation of crop areas which is difficult to achieve in real time because crop statistics are normally available only several months after harvest. Therefore, another option is to assess production directly. One such approach is investigated in Chapter 3.

3 Data requirements

As described in the previous sections, the monitoring of vegetation using remotely sensed data can be performed at various levels from a purely qualitative assessment (e.g. greenness) to more quantitative methods (yield, area, biomass). This will have an impact on the type of imagery required for a specific application. A relevant way to classify satellite imagery is according to its resolution. Spatial resolution is the type of resolution which is most often used. High-resolution satellites such as SPOT, IRS or Landsat typically have spatial resolution between 6 and 30m and are particularly adapted for the mapping of land cover patterns.

However, these sensors are not suited to map vegetation dynamics over large areas. This is because there is a trade-off between spatial resolution and swath or area covered by the satellite. The greater the spatial resolution the smaller the swath covered by the sensor. For example, the SPOT satellite with one of the highest spatial resolution of the remote sensing satellites currently in operation, can only cover a 60km swath. The implication of the smaller swath is also of a lower temporal resolution. Temporal resolution is the interval between two visits over the same satellite orbit path. Again for SPOT, it takes 22 days to revisit the same area of the Earth's surface whereas this only takes 16 days for Landsat TM which has a higher swath. In the case of SPOT, this temporal resolution can be reduced thanks to the pointing capability of its sensors. This is further reduced if, as in the case of SPOT, more than one satellite is in orbit. However, the coverage will still be limited to a small area. Even if larger swaths were possible, the amount of data to be transmitted and processed from the satellite would become very large reaching the limits of the available transmission bandwidth. Furthermore, the monitoring of vegetation growth does not really require high spatial resolution which is only necessary to identify individual cover types. Therefore, to monitor vegetation dynamics, a different type of sensor with a lower spatial resolution is required and this type of sensor already exists in the form of AVHRR. The use of AVHRR data for land applications and more specifically for vegetation monitoring began with the launch of the NOAA-6 satellite in 1979 (Tucker 1996). For the first time, it became possible to calculate vegetation indices from AVHRR data thanks to channel 1 being restricted to the upper portion of the visible electro-magnetic spectrum (0.55-0.70μm) as opposed to a wider waveband previously (0.55-0.90μm). However, the equatorial crossing time of NOAA-6 (in the morning, i.e. low solar energy) was inappropriate for vegetation monitoring and the first NOAA satellite that could effectively be used for land applications was NOAA-7, launched in 1981, which had an afternoon pass. Since then, various NOAA satellites have been launched maintaining the continuity of AVHRR data. The AVHRR sensor consists of 5 wavebands with two in the visible/near infrared portion of the spectrum, one in the middle infrared and two in the thermal infrared. Each band has a spatial resolution of 1.1 x 1.1km at nadir. Another of its main characteristics is its wide field of view, ±55°, which allows a swath of about 2,500km (Cracknell 1997). However, this wide field of view leads to serious distortions on the edge of the swath restricting the use of AVHRR imagery for quantitative analysis to the central portion of the swath. Despite this limitation, it is possible to get daily coverage of a particular region, especially when two satellites are available.

Full resolution AVHRR data are only available in two forms; firstly, through direct reception of AVHRR, known as HRPT (High Resolution Picture Transmission). Secondly, a small portion of full resolution data is recorded. This mode is known as Local Area Coverage (LAC). The remainder of the data from a given orbit is recorded on board the satellite at a lower spatial resolution, this mode is known as Global Area Coverage (GAC). GAC results from the combination of averaging and sampling of the LAC pixels resulting in a 5.5x3.3km spatial resolution at nadir (Belward 1992). GAC data resolution is too low to allow the selection of suitable calibration sites for measuring vegetation quantity because sites would have to be at least the size of one pixel in a reasonably homogeneous area and even more in heterogeneous vegetation. It would be totally impractical to survey such a site in the field. Therefore, the use of GAC is restricted to global monitoring studies or the assessment of vegetation conditions. LAC data are required for more detailed studies such as the measurement of vegetation quantity.

Currently, there is a fairly wide range of ground stations available throughout the world allowing complete coverage with HRPT data, although this has not always been the case. Previously, only GAC data were systematically archived daily at NASA. It is now fairly straightforward to order AVHRR data from one of these receiving stations, especially with the introduction of the Satellite Active Archive of NASA, where AVHRR imagery can be ordered and downloaded via the Internet (http://www.saa.noaa.gov). However, as said earlier only a small portion of LAC data are recorded and the only way for the user to make sure that the data are acquired is to receive them locally. This is particularly important for real time vegetation monitoring such as required by an early warning system, because it is crucial to ensure that these data are acquired and available for analysis immediately. Therefore, the only solution is the direct reception of AVHRR data. Fortunately, access to AVHRR data is free and a number of low cost PC based systems are developed and installed in various parts of the world (Gower 1992, Williams and Rosenberg 1993). For the work described in this thesis, HRPT data originated from LARST (Local Applications of Remote Sensing Technology) receivers originally developed by BURS (Bradford University Remote Sensing) Ltd and now installed in over 50 countries throughout the world. The processing of the AVHRR data received includes radiometric corrections taking account of sensor degradation, cloud screening, geometric corrections based on orbital parameters and manual shift, NDVI calculation and maximum value compositing. These can effectively be used in real-time for various applications such as fire, sea surface temperature or vegetation monitoring.

It is somewhat difficult to develop operational products from the analysis of NDVI images from a single season. The comparison of a current image with historical imagery is much more meaningful. Therefore, there is a requirement for a long-term archive of AVHRR data. Unfortunately, the archiving of AVHRR LAC data has not been consistent. The only global AVHRR LAC data set available at present is the International Geosphere Biosphere Programme (IGBP) 1km global data set (Belward 1996), but the data are only available for a couple of years from April 1992 and between 1995 and 1996. This is obviously not sufficient for long term temporal analysis. As a result, current AVHRR LAC scenes need to be compared with existing GAC archives. There are a number of GAC archives that have been produced over the past two decades (Belward 1996). The first one was the GVI (Global Vegetation Index) data set. Unfortunately, it suffers from a number of shortcomings, one of which is a very low resolution (at best 15km) which makes it impracticable to compare with LAC data. A number of other data sets were derived from GAC data and produced by the GIMMS (Global Inventory Monitoring and Modelling Study) group at NASA. One of the first data set they produced was for Africa with a 10 day Maximum Value Composite data set covering the period from 1981 to recent. A similar data set was compiled for the FAO's ARTEMIS project (Hielkema and Snijders 1994). The FAO-ARTEMIS data set is of particular interest as it can easily be obtained on a CD-ROM. The data are 10-day maximum value composites at 7.6-km resolution covering a 10-year period from 1981 to 1991. Thirdly, the Monitoring Tropical Vegetation (MTV) Unit of the Joint Research Centre (JRC) of the Commission of European Communities, produced their own data set (Malingreau and Belward 1994). The data are available from 1981 to mid-1992 as daily mosaics at 5-km resolution. Finally, a fourth data set was developed for the entire globe by the NASA AVHRR Pathfinder project. The so-called Pathfinder AVHRR Land (PAL) data set covers the entire globe from 1981 to 1994 with 10-day maximum valued composites at 8-km resolution. PAL is different to the other data sets cited so far in that it incorporates atmospheric corrections. Although this may improve the final product, most processing chains of HRPT data do not include any atmospheric corrections other than the calculation of maximum value composites. The main limitation in the use of these different archives is that when a new one becomes available covering a longer time period, little information is available about the compatibility with the previous archive used. Chapter 4 looks at the issue of the compatibility of PAL with traditional data sets such as the FAO-ARTEMIS.

4 Thesis Structure

The research presented here has been published, accepted or submitted in the *International Journal of Remote Sensing* or conference proceedings. The thesis is made up of following papers:

Chapter 2. TAYLOR J.C., BIRD A.C., SANNIER C.A.D., PRATT N., 1996, Calibration and validation of thematic maps from remote sensing in Developing Countries: Need and Method. In *Remote Sensing and GIS for Natural Resource Management,* edited by L.J. Rosenberg, C.H. Power and I. Downey (Chatham, UK: Natural Resources Institute. ISBN 0 85954 454-0), pp. 39-46.

Chapter 3. SANNIER C.A.D., TAYLOR J.C., du PLESSIS W. and CAMPBELL K., 1998, Real-Time Vegetation Monitoring with NOAA-AVHRR in Southern Africa for Wildlife Management and Food Security Assessment. *International Journal of Remote Sensing*, **19**, 621-639.

Chapter 4. SANNIER C.A.D, TAYLOR J.C. and CAMPBELL K., 1998, Compatibility of FAO-ARTEMIS and NASA Pathfinder AVHRR Land NDVI data archives for the African continent. *International Journal of Remote Sensing*, **19**, 3441-3450.

Chapter 5. SANNIER C.A.D, TAYLOR J.C, SLADE G., NTABENI T., NGAKANE S. and CAMPBELL K, 1998. Real-Time Rangeland Monitoring in Botswana with NOAA-AVHRR. In *RSS98 Developing International Connections.* Remote Sensing Society annual conference.

Chapter 6. SANNIER C.A.D., TAYLOR J.C., du PLESSIS W. and CAMPBELL K., 1998. Rapid biomass assessment of natural vegetation for real-time monitoring with NOAA-AVHRR in semi-arid Africa. *International Journal of Remote Sensing.* Submitted.

Appendix. TAYLOR J.C., SANNIER C.A.D., DELINCÉ J. and GALLEGO F.J. 1997, *Regional Crop Inventories in Europe Assisted by Remote Sensing: 1988-1993*, European Commission, EUR 17319, 80pp.

A CD-ROM containing relevant data to the thesis is also included. It is organised according to the thesis structure, with a separate directory for chapter 2 to 6. An index file describes the files contained in each directory.

References. The list of references includes all the references already listed for chapter 2 to 5 together with the references specific to chapter 1 and 2

5 Disclosure

The publications included in this thesis are from the work carried out as a research officer at Cranfield University at Silsoe and later developed as a lecturer in the same university.

The work presented in this thesis was mostly carried out under the Africa Regional Remote Sensing Project funded by the Department for International Development. Dr Ken Campbell and Wynand du Plessis who are part of the authorship of several of the papers presented played a role in terms of early discussions on the methodology for the first and participation to fieldwork activities for the second. Professor John Taylor proposed and managed the DfID project and was my supervisor.

Where I am the first author, I took the major role in all aspects of the research and write up. Where I am the second author, I had a leading role in data collection and analysis, methodology and some aspects of the write up but the final structure of the paper resides with the first author. Where I am third author, I made a significant contribution to one aspect of the work but the overall organisation and structure of the paper is not directly mine.

Chapter 2 is from the proceedings of a one-day Remote Sensing Society conference that took place at the University of Greenwich, Chatham Maritime on Remote Sensing and GIS for Natural Resource Management. My involvement concerned the part describing the work done in Etosha National Park which consisted of field data collection and analysis, image processing and some written material, tables and illustrations. Chapter 3 is the paper by Sannier *et al.* (1998a) concerned with the initial work carried out to develop and apply the VPI (Vegetation Productivity Indicator) methodology for wildlife management in Etosha National Park, Namibia and for crop monitoring in Zambia. The bulk of the writing up, the structure of the paper and the data analysis were my responsibility with the help and advice of Prof. John Taylor. In the paper by Sannier *et al.* (1998b) which forms chapter 4, I was responsible for the data analysis and write up with revisions from the other two authors. Chapter 5 by Sannier *et al.* (1998c) was presented at the 24th Remote Sensing Society annual conference and relates the application of the VPI for rangeland monitoring in Botswana. The roles of the first two authors were similar as in the previous paper and the other authors were involved in the production of the outputs in Botswana. In Chapter 6, the authors' inputs are identical to that of Chapter 3. In the Appendix, I was responsible for data analysis and for some early drafts but the first author had the most significant input in the final form of the report.

CHAPTER 2

Calibration and Validation of Thematic Maps from Remote Sensing in Developing Countries: Need and Method

J. C. Taylor[1], A. C. Bird[1], C. Sannier[1], N. Pratt[1] and W. Du Plessis[2]
[1]Silsoe College, Cranfield University, Silsoe, Bedford, MK45 4DT, UK.
[2]Etosha Ecological Institute, P.O. Okaukuejo via Outjo, Namibia.

ABSTRACT

Thematic maps derived from satellite image classification and aerial photographic interpretation usually contain significant errors. These result from spectral confusion, mixed pixels, the classification procedure and the scheme of land classification being applied. Experience in the use of thematic classifications for monitoring agriculture in the European Union (EU) has shown the need for statistically designed sample ground surveys to measure classification errors and to compensate for them when using the thematic data. Methodology for calibration and validation of image classifications has been successfully developed and is now widely used for work in Europe. It involves the design of an area-frame sample to select fixed-size areas at random locations in the study area. Land cover maps of the sample areas are then produced by field survey. An unbiased confusion matrix is produced by comparing the field survey with the classification, to show the pattern of error. This, or regression, is applied to obtain area estimates as class pixel counts alone are highly inaccurate. Difficulty of access to randomly selected sites is perceived as a problem when applying the methods in Africa. This work reports surveys of irrigated land area in northern Nigeria and of natural vegetation in Etosha National Park in Namibia where European methods of ground calibration have been tested and successfully adapted for use in the developing world.

INTRODUCTION

The objective of this paper is to draw attention to the need for calibration and validation of thematic maps produced from remote sensing, and to show how methods developed in Europe have been successfully adapted for applications in development projects in Nigeria and Namibia.

Thematic maps from classification of remotely sensed imagery, including aerial photographic interpretation (API), usually contain errors. These can be measured by comparing the thematic map with ground data using a confusion matrix (Story and Congalton, 1986). This is generated by cross-tabulating the frequencies of occurrence of class combinations, obtained from a random double-sample of ground reference data and the classification. As an example, Table 1 presents the confusion matrix generated for a regional crop inventory in the UK in 1992. The overall agreement is the sum of the diagonal elements divided by the total number of observations in the matrix. The off-diagonal row elements represent the mis-classification of ground classes which are included in the image classification. The diagonal element expressed as a percentage of the row total gives the so-called user or mapping accuracy of the classification for that class. The off-diagonal column elements represent the mis-classification of a ground class into other image classes. The diagonal element expressed as a percentage of the column total gives the accuracy that the producer of the classification has achieved for the class. The agreement between a digital classification and ground survey, given by a confusion matrix, is frequently used to estimate its accuracy, assuming the reference data are accurate.

In recent years, considerable experience in the application of remote sensing technology has been gained in the EU's Monitoring Agriculture with Remote Sensing (MARS) project. Digital classifications of Landsat TM and SPOT imagery were carried out over many different parts of the EU and the following observations were made from this collective experience.

Table 1. Confusion matrix showing the relationship between ground survey data and the digital classification of SPOT imagery in the 1992 regional crop inventory of Beds, Cambs and Northants, UK

		Reference Data										
		Woods	Inland Water	Urban	Wheat	Barley	Summer Crops	Grasses	OSR	Other	TOTAL	User Accuracy
	Woods	15	3	1	2	1		2			24	63%
	Inland Water		4		1						5	80%
Image Data	Urban			11			1			1	13	85%
	Wheat	2	1	2	155	8	1	3		1	173	90%
	Barley	8		4	18	16	3	17		3	69	23%
	Summer Crops		1	13		3	43	3	1	5	69	63%
	Grasses			8	6	7	10	37		1	69	54%
	OSR				1				30		31	97%
	Other	2		12	1	3	11	25		16	70	23%
	TOTAL	27	9	51	184	38	69	87	31	27	523	
	Producer Accuracy	56%	44%	22%	84%	42%	63%	42%	97%	59%	Overall Accuracy 63% Kappa 54% Var (kappa) 0.000599	

- The results presented in Table 1 are typical for a classification based on a single image. The overall accuracy was usually between 60 and 75% for classifications having between 8 and 12 classes.
- The overall accuracy tends to be lower as the number of classes increases.
- Accuracy of individual classes is very variable and can be anything between 0 and nearly 100%, depending on spectral separability.
- The accuracy of classifications based on the combination of two or more images is usually improved to about 80–85% for the same numbers of classes mentioned above.

There are several explanations for the occurrence of mis-classification errors.

- Many cover types are botanically and morphologically similar and have very similar reflectance properties. Mis-classification then results from spectral confusion. This is exacerbated when more sub-categories have been included.
- A frequently ignored source of mis-classification is the presence of mixed pixels. These occur along the borders between land parcels of different cover types. The effect of mixed pixels depends on the size of the pixel relative to the parcel sizes. As the size of the pixel approaches the size of land parcels, a very high proportion of the pixels may represent mixtures of all sorts of class combinations. On-going research at Silsoe has shown that in SPOT images of agricultural areas in England, the proportion of mixed pixels was between 13 and 35% (M. J. Dufour, personal communication).
- The digital classification procedure employed also influences the accuracy of classification. The main factors are: the selection process for training data, i.e., how representative the sample is of the spectral

properties of each class; and the classification algorithm used as each one defines the spectral boundaries for each class in a different way.

- The land classification scheme being applied also affects the accuracy of classification which will depend on whether class definitions are based on land cover or land use. In the latter case, variations in land use may be unrelated to spectral differences. Classifications which require separation of classes based on plant species may also be difficult because of the inherent similarity of spectral properties.

The overall accuracy using API is generally higher than for digital classification but there is wide variation in the accuracy of individual classes (Taylor *et al.*, 1991).

The above experience confirms that significant mis-classification errors usually exist in thematic maps produced from remote sensing. This poses important problems for the use of the data:

- as classes are not accurately identified at all locations, and errors vary from map to map, accurate assessment of change by simple cross tabulation of classifications from two dates is not possible;
- estimates of class areas, which are important for agricultural inventory, cannot be accurately measured from pixel counts.

In Table 2, a summary of results from the regional inventory reported by Taylor and Eva (1992) illustrates the latter point and shows the effect of using two different classification procedures. The columns PC-W and PC-UW are the class areas obtained by pixel counts in two separate digital classifications. The columns REG-W and REG-UW are the corresponding areas, and their 95% confidence intervals, estimated by combining the ground survey data with the respective digital classifications using the regression method described by Taylor and Eva (1993). The MAFF column gives, where available, the areas estimated by a census carried out independently by the Ministry of Agriculture, Fisheries and Food (MAFF). Comparison of the area estimates for each cover type shows that the pixel counts are widely different and that they are influenced by varying the digital classification algorithm, hence the degree of mis-classification. On the other hand, the regression estimates are similar to each other and to the MAFF census figures. This shows that the differences caused by varying the classification algorithms have been corrected by the regression technique.

Experience in the use of thematic classifications for monitoring agriculture and natural land cover in the EU has shown the general need to measure classification errors and to compensate for them. The acquisition of an unbiased sample of ground observations of sufficient size is crucial for this so that confusion matrices can be produced or regression estimates made.

The effect of bias can easily be illustrated by examining the effect of over-sampling one class. For example, if we increased the sampling of wheat in Table 1 by a factor of two, we would expect the mis-classifications of the reference data to be in the same proportion. Thus, the wheat column of Table 1 would have every element multiplied by two. This would increase the estimate of the user accuracy of wheat to 95%. However, the user accuracies for the other classes would be under-estimated because of the increased number of wheat commission errors. The user accuracy of barley for example would be reduced to 18%.

The confusion matrix will be biased if the samples for each class are not proportional to the class areas. As these are not known before the classification, a random sample design is used to obtain the necessary data. The MARS project

Table 2. Areas (ha) of cover types in the same region of England estimated by different techniques using digital classifications and by agricultural census

CLASS	PC-W[1]	PC-UW[2]	REG-W[3]	REG-UW[4]	MAFF[5]
Woodland	35153	42834	29636 ±19%	29409 ±25%	na
Inland water	5510	4123	5744 ±58%	6279 ±44%	na
Urban	111775	22518	82594 ±14%	70668 ±20%	na
Wheat	210459	159938	238003 ±6%	236736 ±6%	227637
Barley	22839	93969	53900 ±27%	55502 ±24%	50585
Summer crops	66990	105593	89888 ±13%	96494 ±15%	82587
Grass and forage	192582	107509	130339 ±12%	124691 ±13%	114491
Rape	29946	29047	44244 ±10%	44095 ±11%	46643

[1]pixel count, area-weighted discriminant functions; [2]pixel count, un-weighted discriminant functions; [3]regression estimate, area-weighted discriminant functions; [4]regression estimate, un-weighted discriminant functions; [5]MAFF agricultural census

employed an area-frame sample, developing methods used by USDA for crop area estimation in the 1970s (Hanuschak *et al.*, 1979). Between 1988 and 1993, there were study sites in 10 countries, generally covering areas of around 20 000 km^2. The study regions were divided into fixed-size areas by a regular grid to produce the sampling frame. Random samples of the fixed-size areas (referred to as segments) were selected. Many of the MARS study sites used a stratified sample design. All the sample segments were visited in the field to identify and map each land parcel within the segment. This required that the enumerators were suitably trained in the identification of crops and other land cover classes, and were equipped with suitable documents to enable them to locate the segments and draw maps of sufficient accuracy in the field. For example, the survey documents for each segment in the UK study consisted of: a false colour composite satellite imagette of the 1 km square segment and the surrounding 0.5 km border at 1:10 000 scale; a transparent 1:10 000 OS overlay of the segment and surrounding area; a 1:25 000 map of the segment and surroundings; a transparent overlay on which to draw parcel boundaries; a proforma on which to record parcel numbers and crop types; and a 1:50 000 OS map of the area for road navigation.

Area-frame sampling was successfully used in Libya by Latham *et al.* (1983) to measure irrigated areas, but its general application in Africa is perceived as being limited. The requirement for a random sample of locations generally means that access will be very difficult for some sites. The temptation is to ignore these even though doing so would invalidate the error assessment. The following are two examples of applications of area-frame sampling in Africa where feasible methods have been developed.

VEGETATION MAPPING IN ETOSHA NATIONAL PARK

Classification of vegetation in Etosha National Park was carried out as part of a project to measure vegetation status in near real time using NOAA-AVHRR images (Sannier *et al.*, 1995). The classification scheme used in Etosha was adapted from the Yangambe classification (Boughey, 1957) and aimed to separate vegetation classes according to the height, density and main species of woody vegetation.

A randomly aligned, systematic area-frame sample was chosen to facilitate aerial as well as ground survey. The sampling frame was the Universal Transverse Mercator (UTM) grid and the sample segments were 1 km square. The location of one sample segment within a 10 x 10 km block was determined randomly to give a 1% sampling rate, and the same location was used within each adjacent block to give the systematic sample of 220 segments inside the Park as shown in Figure 1. Three types of field survey documents were produced, based on geo-referenced Landsat TM imagery map products. These were: 1:125 000 scale image maps of the whole Park, in sheets covering 20 x 20 km, to assist navigation; 1:30 000 scale extracts, covering a 4.5 x 4.5 km area centred in each segment, to locate them; and 1:10 000 scale extracts of the segments for mapping. In addition, surveyors were supplied with booking forms, which included the class definitions for ready reference in the field, to help maintain consistency of class identification.

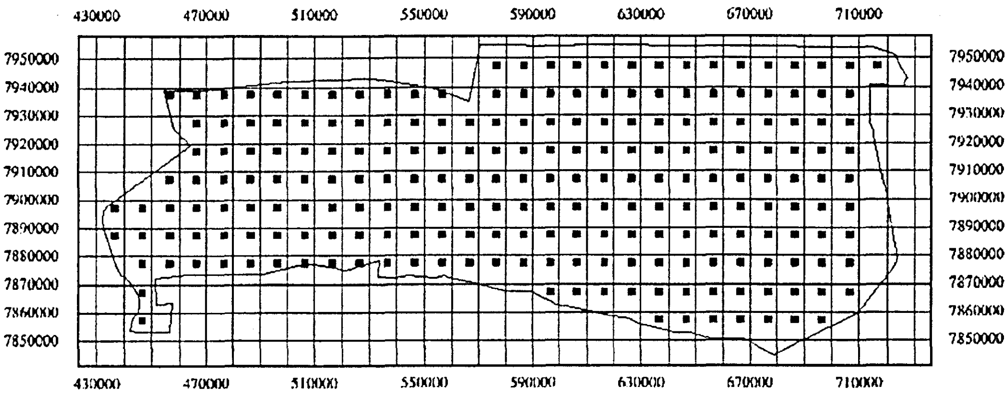


Figure 1. Area-frame sample of Etosha National Park

A total of 82 sites were surveyed on the ground and all sites were surveyed from the air. The ground survey was implemented in order to check the air survey. Prior to the field/aerial visit, the 1:10 000 TM extracts were photo-interpreted and parcel boundaries were drawn on to acetate overlays. On the ground, the position within the segment was determined by a GPS (Global Positioning System). This instrument was indispensable because there were no features to assist navigation most of the time. Parcel boundaries were checked and modified when necessary. The vegetation type of each parcel was determined and recorded on the booking form. From the air, the plane was navigated by the GPS to each site at a local altitude of 200–250 m and a speed of 120 kn. One observer recorded the vegetation with the help of the field survey documents, while another filmed the site through the window using a standard camcorder. After the flight, the video was used to verify the interpretation made in the plane. All segments were surveyed in about 30 h of flying; ground survey took about 10 times as long.

A random sub-sample of the survey data was used to derive spectral signatures for 33 land cover classes. The number of training pixels for each class was made approximately proportional to the class areas. Some of the classes had multi-modal statistical distributions and were sub-divided using a clustering routine. The classification was performed using the maximum likelihood algorithm. Two TM images covering the east and west of the Park were classified separately and the classifications were mosaiced later.

A confusion matrix was produced from the random sub-sample of survey locations. This enabled the agreement of the classification with survey data to be calculated for different levels of class groupings. Table 3 is the confusion matrix for seven main classes and shows that the overall agreement was 68%. Savanna cover types can be difficult to classify because boundaries between classes are often diffuse. The confusion matrix reflects this. Comparison of the ground survey and air survey results, also in a confusion matrix, indicated the classes which could not be reliably discriminated by visual assessment. Some types of mis-classifications were judged to be unimportant for the production of vegetation status maps and when these cases were allowed for, the agreement rose to over 80%. The confusion matrix will allow the digital classification to be used for various purposes as quantitative assessment of the errors can be made.

Table 3. Matrix for Etosha vegetation survey

Classification	Reference Data 1	2	3	4	5	6	7	Total	User Acc
1 Bare Ground	**2454**	21	1		1	11		2488	99%
2 Grassland	54	**140**	16	5	62	15	12	304	46%
3 Steppe	46	75	**407**	48	46	35		657	62%
4 Grass Savanna		2	13	**50**	74	46	4	189	26%
5 Shrub Savanna	17	23	34	21	**1824**	793	82	2794	65%
6 Low Tree Savanna	1	16	6	2	932	**1539**	151	2647	58%
7 High Tree Savanna		6			284	363	**687**	1340	51%
Total	2572	283	477	126	3223	2802	936	**10419**	
Producer Acc	95%	49%	85%	40%	57%	55%	73%	Overall Acc	68%

ESTIMATING THE AREA OF SMALL-SCALE IRRIGATION IN NORTHERN NIGERIA

An agricultural monitoring project was carried out in the region of northeast Nigeria adjacent to the border with Niger as part of the programme reported by Bird *et al.* (1995). The study area presented a complex agricultural setting with dryland agriculture on better drained sands, and various forms of flood-related or irrigated cultivation in the floodplain areas (Kimmage and Adams, 1990). In this study, the area of land under irrigation was the main focus. The aim of the project was to test the methodology which might be appropriate to the area in order to produce same-season inventory figures. For this reason, the study was based on those areas of irrigation which could be covered by one SPOT XS satellite image. The actual region studied was restricted to zones where small-scale irrigation was known to be feasible. Two zones were defined around the river system which was the main water source. These were delineated from a grid of 500 x 500 m squares overlaid on photomosaics created from aerial photographs independently of this work. Squares which contained a potential surface water resource, such as the river, a channel or a pond, were allocated to zone 1, while the squares adjacent to these were allocated to zone 2. This led to a study area of 252.5 km^2 within the SPOT scene. A field survey was designed, based on a sample size of 500 x 500 m, using an unaligned systematic random sample with a sampling fraction of 5%. This amounted to 49 ground segments, each of which was mapped during the field survey. The field survey was carried out during 10–24 January 1994, which coincided with the peak of the irrigation season. A SPOT satellite image was acquired on 11 January 1994.

Within the study area defined by the image, the first zone contained 32 segments, of which 11 included irrigation sites; the second zone had 17 segments of which only three contained irrigation. A supervised classification was applied to the imagery with training pixels derived from the ground survey data. The training was based on a simple irrigation/non-irrigation scheme. The classification was carried out using two of the available algorithms, 'maximum likelihood' and 'k nearest neighbour', on the ILWIS software package (ILWIS, 1993). Low mapping or user accuracies were found for the irrigated area in the order of 40% for 'maximum likelihood' and 30% for 'k nearest neighbour'. Confusion occurred between irrigation, woodland and aquatic grassland, i.e., between those classes made up of green vegetation. The 'k nearest neighbour' method classified twice as many pixels as the 'maximum likelihood' method. The predicted locations of irrigation were the same, but the number of pixels classified in those locations differed. The resulting maps could be used as general indicators of where irrigation was likely to occur but not as accurate maps. Quantitative data were better derived from area estimates for the study area as a whole.

The area of irrigated land could be calculated in a number of ways from the data collected. The options were: direct expansion of the field data; pixel counting from the classified image; or the application of the regression estimator using both the field data and the image classification. Table 4 summarizes the results obtained from these techniques. The direct expansion of field data gave a wide confidence interval while the pixel count showed major differences between the two classification methods. By applying the regression estimator technique, the two classifications yielded much more comparable results. For both zones, it could be said that between 469 and 1107 ha of irrigation exist in a study area of 25 250 ha, or between 1.9% and 4.4%.

The results of this work indicated that first estimates of the area under small-scale irrigation could be made in a study area where previously, no information was available. The estimates were far from exact numbers, but were upper and lower limits of a range. They provided a starting point for land use planning with regard to irrigation and water supply in the area.

Table 4. Estimated area of small-scale irrigation

Total study area	25 250 ha
Irrigated area– by direct expansion of field data	840 ± 524 ha
Irrigated area– by classified image pixel count	
'maximum likelihood' classifier	1009 ha
'k nearest neighbour' classifier	2312 ha
Irrigated area – by regression estimator	
'maximum likelihood' classifier	788 ± 319 ha
'k nearest neighbour' classifier	749 ± 369 ha

CONCLUSIONS

The work in Africa has confirmed that field surveys which enable the collection of statistically unbiased data can be implemented in Africa. The use of GPS to locate sample sites has facilitated the collection of land cover data from aircraft in otherwise inaccessible locations. Data collection by ground survey was feasible but more time consuming.

The digital classifications carried out in Namibia and Nigeria contained similar problems to those in the MARS project, namely, high levels of mis-classification. Area estimates made from pixel counts alone depended on the classification algorithm used and were unreliable. The integration of an area-frame sample of ground survey data with results of digital classification led to more consistent estimates.

Proper calibration of digital classifications with statistically unbiased ground data is both necessary and feasible in Africa. The methods developed in the EU have been successfully adapted for this.

REFERENCES

Bird, A.C., Pratt, N.D. and Lawan, A.I. (1995) The development of GIS and remote sensing techniques in the Centre for Arid Zone Studies, North East Nigeria. Presented at *Remote Sensing and GIS for Natural Resource Management, RSS Workshop, 19 December 1995, Chatham, UK.*

Boughey, A.S. (1957) The physiognomic delimitation of West African vegetation types. *Journal of the West African Science Association,* 3(2): 148-165.

Hanuschak, G., Sigman, R., Craig, M., Ozga, M., Luebbe, R., Cook, P., Kleweno, D. and Miller, C. (1979) *Obtaining Timely Crop Area Estimates Using Ground-Gathered and Landsat Data. USDA Technical Bulletin,* no. 1609. Statistical Research Division, Economics Statistics and Co-operative Services, USDA.

ILWIS (1993) *ILWIS User's Manual.* Enschede, Netherlands: International Institute for Aerospace Survey and Earth Sciences.

Kimmage, K. and Adams, W.M. (1990) Small-scale farmer-managed irrigation in northern Nigeria. *Geoforum,* **20**: 435-443.

Latham, J.S., Ferns, D.C., Colwell, J.E., Reinhold, R. and Jebe, E.H. (1983) Monitoring the changing areal extent of irrigated lands of the Gefara Plain, Libya. *Advances in Space Research,* **2**(8): 57-68.

Sannier, C., Taylor, J.C., du Plessis, W. and Campbell, K. (1995) Application of remote sensing and GIS for monitoring vegetation in Etosha National Park. Presented at *Remote Sensing and GIS for Natural Resource Management, RSS Workshop, 19 December 1995, Chatham, UK.*

Story, M. and Congalton, R.G. (1986) Accuracy assessment: a user's perspective. *Photogrammetric Engineering and Remote Sensing,* **52**: 397-399.

Taylor, J.C., Bird, A.C., Keech, M.A. and Stuttard, M.J. (1991) *Landscape Change in the National Parks of England and Wales - Final Report. Vol. 1. Main Report.* Silsoe College, Bedford.

Taylor, J.C. and Eva, M.D. (1992) *Regional Inventories on Beds, Cambs and Northants (UK).* Final Report. Contract No. 4817-92-06 ED ISP GB Joint Research Centre, Commission of the European Communities. Silsoe College, Silsoe.

Taylor, J.C. and Eva, M.D. (1993) Operational use of remote sensing for estimating crop areas in England. In: *Towards Operational Applications. Proceedings of the 19th Annual Conference of the Remote Sensing Society, Chester, UK, 1993.* Nottingham: Remote Sensing Society.

CHAPTER 3

INT. J. REMOTE SENSING, 1998, VOL. 19, NO. 4, 621–639

Real-time vegetation monitoring with NOAA-AVHRR in Southern Africa for wildlife management and food security assessment

C. A. D. SANNIER and J. C. TAYLOR
School of Agriculture Food and Environment, Cranfield University, Silsoe, Bedford, MK45 4DT, England, UK

W. DU PLESSIS
Etosha Ecological Institute, P.O. Okaukuejo, via Outjo, Namibia

and K. CAMPBELL
Natural Resources Institute, Central avenue, Chatham Maritime, Chatham, Kent ME4 4TB, England, UK

(*Received 15 November 1996; in final form 26 May 1997*)

Abstract. The monitoring of vegetation in Southern Africa with satellite data has become increasingly important over the past decade because it is linked to variation in agricultural production and climate change with implications for wildlife management and tourism. This study shows how maps of vegetation status were produced in near real time from NOAA images acquired from the local receiving stations in Etosha National Park, Namibia and in Zambia. Map products based on the NDVI were put into historical context and stratified to remove effects of the main vegetation types in order to assess vegetation status. The historical data were extracted from the FAO ARTEMIS NDVI archive and processed to obtain a statistical distribution of the NDVI for each 10-day period of the year and vegetation type by applying techniques commonly used in hydrology for the prediction of extreme events. The quintile probability ranges were used to define five classes of a Vegetation Productivity Indicator (VPI). LAC NDVI images obtained in real-time from the receiving station were processed to derive a VPI map for each 10-day period. In Etosha National Park and in Zambia, the VPI was strongly related to the rainfall and the VPI maps provided improved information on the spatial variations. The weighted average VPI for the main agricultural region of Zambia was significantly correlated with maize production.

1. Introduction

Droughts are one of the main causes of food shortages in Southern Africa (Hutchinson 1991). The time available for strategic planning to prevent famine is increased if they can be identified early. Also, the mapping of affected areas allows improved targeting of aid measures. If the impact of the drought on crop production can be quantified early there are considerable benefits for planning finance and the logistics of supply. The ability of a crop to produce is related to plant development. Drought-stunted plants yield poorly and well developed plants yield better. Variation in the annual development of vegetation in agricultural areas is therefore linked in a general way to variation in production.

Wildlife is also a major resource for Southern African economies through tourism and big-game hunting and better management strategies are needed because competing land use is restricting the traditional range for wildlife to national parks and conservation areas. Also, climate change may have a large negative impact on habitats. Therefore, the interaction between the wildlife population, the possibility of climate change, and the vegetation must be understood to sustain the wildlife resource and monitoring the vegetation condition is vital in this process.

The study area for application of NOAA-AVHRR for food security assessment shown in figure 1, was Zambia, an area of 753 000 km^2, with a tropical climate which is tempered by altitude. Although, rainfall is much higher than in neighbouring countries, varying between 700 mm in the south-west and 1400 mm in the north-west, it is also very variable from year to year. The annual population growth rate is 3·7 per cent which is one of the highest recorded in Africa but the growth rate of the country's agricultural output has been much lower causing an increase in Zambia's food imports (Moyo *et al.* 1993). Also, the agricultural production has been very erratic. For example, the annual maize production between 1981 and 1993 varied from 484 Mt in 1992 to 1845 Mt in 1988. The very low levels of agricultural

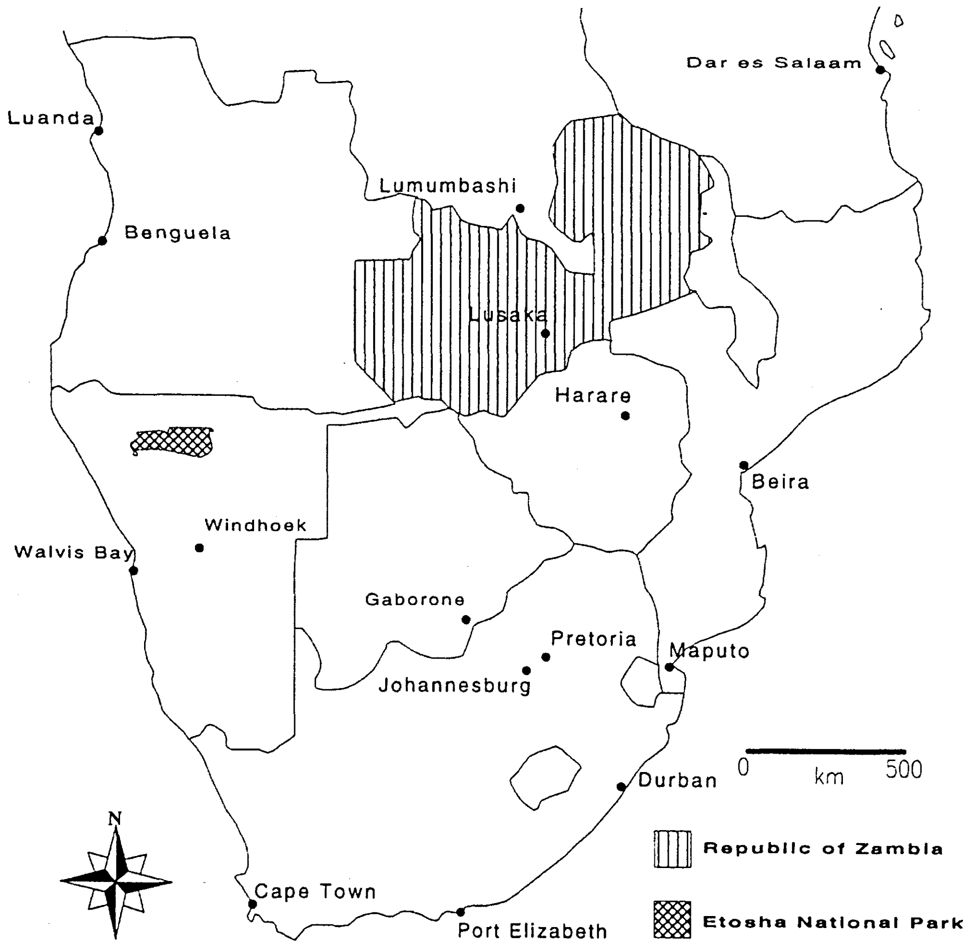


Figure 1. Map of Southern Africa and location of study areas.

production are mainly caused by droughts as 97·5 per cent of the cultivated area does not benefit from irrigation practice (Azzali 1991). Therefore, the early identification of the extent of droughts is crucial and has been a priority of the Zambian Meteorological Department.

The study area for application of NOAA-AVHRR in wildlife management was Etosha National Park (22 300 km^2), one of the biggest wildlife reserves in the world, situated in Northern Namibia, as seen in figure 1. The main feature of the Park is a saline desert called the Etosha pan which covers about a fifth of the area. The Pan usually remains dry except during exceptional rainy seasons when it can contain shallow water for a short period. The climate is semi-arid with a rainfall gradient increasing from around 300 mm in the west to 450 mm in the east (le Roux 1988). Vegetation types vary from grassland to fairly dense savanna in the south-east.

Previous studies (Hutchinson 1991, Lambin *et al.* 1993), have shown the advantage of satellite remote sensing and particularly NOAA-AVHRR, for the monitoring of vegetation conditions compared to more traditional methods such as interpretation of rainfall measurements. The available network of rain gauges is not usually dense enough to allow a reliable interpolation of rainfall over a whole country. Flitcroft *et al.* (1989) showed that spatial interpolation of short-term rainfall is unreliable in similar climates in West Africa when the distance between gauges is more than 15 km.

The most widely used satellite derived indicator of vegetation activity is the NDVI (Normalised Difference Vegetation Index) and several studies have used it to assess crop yields or savanna primary production but these have not suggested an operational methodology (Prince and Astle 1986, Kennedy 1989, Diallo *et al.* 1991, Prince 1991, Rasmussen 1992). For the non-technical user, NDVI image maps are difficult to interpret. Firstly, because an explicit relation between NDVI and vegetation condition is not available and secondly, there may be different relations for each vegetation type. Previous studies (Kogan 1990, Maselli *et al.* 1993) have shown the influence of geographical variations on the interpretation of the NDVI. Groten (1993) developed a statistical method for forecasting crop yield per unit area for the whole of Burkina Faso from NDVI profiles. However, the total crop production is the main interest for government and aid agencies and forecasting yield per unit area is only part of the equation.

There is a need to compare the current NDVI with historical data in order to assess the extremity of the event and hence infer the actual vegetation status. Major aid organisations (USAID-FEWS, FAO) have set up operational early warning systems which compare current NDVI images with the previous dekad (10-day period) or with the mean image for the dekad (Le Compte 1989, Hutchinson 1991, Lambin *et al.* 1993). The latter method is very simple but relies on the temporal variation of the NDVI for a location and a given dekad being normally distributed. This assumption maybe unreasonable because the lower limit of the NDVI is bounded by the response for bare soil. Kogan (1990) took a different approach and defined a Vegetation Condition Index (*VCI*) as:

$$\mathrm{VCI} = 100\left(\frac{\mathrm{NDVI} - \mathrm{NDVI}_{\mathrm{min}}}{\mathrm{NDVI}_{\mathrm{max}} - \mathrm{NDVI}_{\mathrm{min}}}\right) \tag{1}$$

where $\mathrm{NDVI}_{\mathrm{max}}$ and $\mathrm{NDVI}_{\mathrm{min}}$ are the maximum and minimum NDVI values in the time series, for the dekad. This assumes that the current range represents the maximum possible variation and that all values of the NDVI within the range occur with the same frequency and therefore have the same probability. This is also maybe an unrealistic assumption.

This work proposes a Vegetation Productivity Indicator (VPI), an alternative method to compare the current NDVI with the historical NDVI archive to assess vegetation condition. The method estimates the statistical distribution of the NDVI empirically from the available data without limiting assumptions and is sensitive to the background vegetation type. We investigate the real-time application for wildlife management and food security assessment using NOAA AVHRR HRPT from local receiving stations described by Williams and Rosenberg (1993).

2. Methodology

The interpretation of vegetation status could be applied pixel by pixel if the statistical distribution of the NDVI for every pixel location is calculated. However, the amount of processing required makes this approach less suitable for the operational application. Also, the pixel by pixel approach cannot be applied to LAC (Local Area Coverage) images because a time series at that spatial resolution does not yet exist. Therefore, the approach was simplified by stratifying the study areas into zones with homogeneous NDVI response so that the 7·6 km resolution ARTEMIS (Africa Real Time Environment Monitoring Information System) image archive, produced by FAO, could be used to estimate the statistical distribution of the NDVI.

In Etosha, this was done using a vegetation map of the Park, developed from a supervised classification of Landsat-TM imagery and an extensive sample survey of the Park's vegetation using methodology adapted from Taylor *et al.* (1996). The vegetation types were field-mapped in a randomly aligned systematic sample of 1 km square areas equivalent to 1 per cent of the Park area, excluding the pan. Landsat-TM imagery was geometrically corrected with a RMS error less than 30 m using the UTM (Universal Transverse Mercator) projection. A subset of the field data was used to train the maximum likelihood classifier. The supervised classification was performed using 33 classes and these were grouped, post classification, into four categories: bare ground (including salt pan), grassland (and steppe), shrub savanna and tree savanna. The overall agreement of the vegetation map with ground observations for these broad classes was estimated to be 89 per cent using a confusion matrix.

In Zambia, a different approach was used. There were no suitable maps available for stratifying the NDVI images. Existing land cover and vegetation maps of the country were compiled in the seventies (Edmonds 1976, Schultz 1975) and changes have occurred since then. Also, the high number of classes and the difficulty to aggregate them make these maps unsuitable for use with remotely-sensed data.

The ARTEMIS image archive was used to stratify Zambia. The images are derived from NOAA AVHRR and are composites of the maximum NDVI value recorded during dekads (10-day intervals) as described by Holben (1986) at 7·6 km spatial resolution covering the whole of the African continent from August 1981 to June 1991. Images consisting of the temporal mean NDVI were produced for each dekad. The 36 images thus produced were used to perform an unsupervised classification with the ISODATA algorithm (ERDAS 1993). Several classifications having different preselected numbers of clusters were assessed by visual comparison with the existing land cover maps and by inspection of the NDVI profiles formed from the signatures means such as shown in figure 2. The statistical distributions of the clusters were also checked to ensure that they were unimodal. A 9-cluster classification was found to be the most

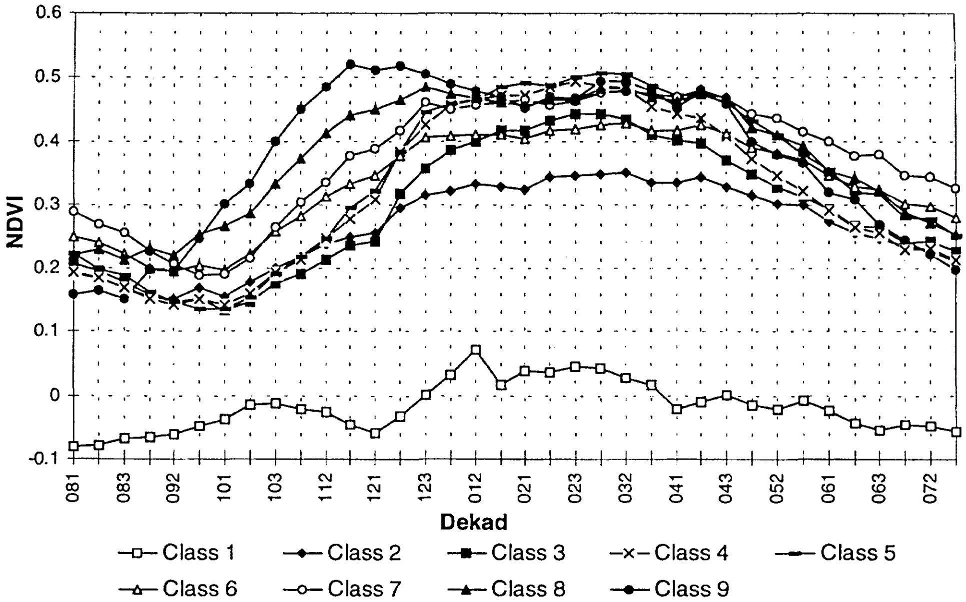


Figure 2. Average profiles over a ten year period for the nine classes in Zambia, derived from the unsupervised classification of the FAO-ARTEMIS NDVI archive.

suitable having a good overall resemblance to previous vegetation maps, unimodal cluster distributions and good separation between the profiles. The differences between each class in figure 2 can be easily seen especially during the green-up period. Once the class signatures were determined the image was classified using the maximum likelihood algorithm. The resulting classification was re-projected from the Hammer-Haitoff projection to Plate-Carrée to fit the local receiving station output.

The stratification of each study area allows representative time-series NDVI profiles for each stratum to be extracted. In the case of Etosha, homogeneous areas of each of the main cover types, equivalent in size to ARTEMIS pixels, were identified using the TM vegetation classification and the field survey data. Several sites were chosen to represent each of the main cover classes except for the grassland and steppe class which only covers a small area of the Park (10 per cent) around the pan where only one pure site could be found.

Time series profiles, already corrected for sensor degradation by the method of Loss (1993), were extracted for the 10-year period covered by the ARTEMIS data set (August 1981 to June 1991). The method suggested by Groten (1993) was used to estimate NDVI values in parts of the time series affected by cloud cover. The average NDVI value on the Pan for the corresponding dekad was also subtracted to maintain the compatibility with the process applied to the LAC data, described later. However, this correction was small, because the ARTEMIS data set was already calibrated. Once the time series were corrected for cloud cover, the data were averaged for each dekad resulting in 10-year averaged seasonal profiles in figure 3. The profiles for each site in Shrub and Tree Savanna in figure 4(a) show the within class variation which is greater for Tree Savanna but the general pattern is always distinct from that of Shrub Savanna. The start of the season is always

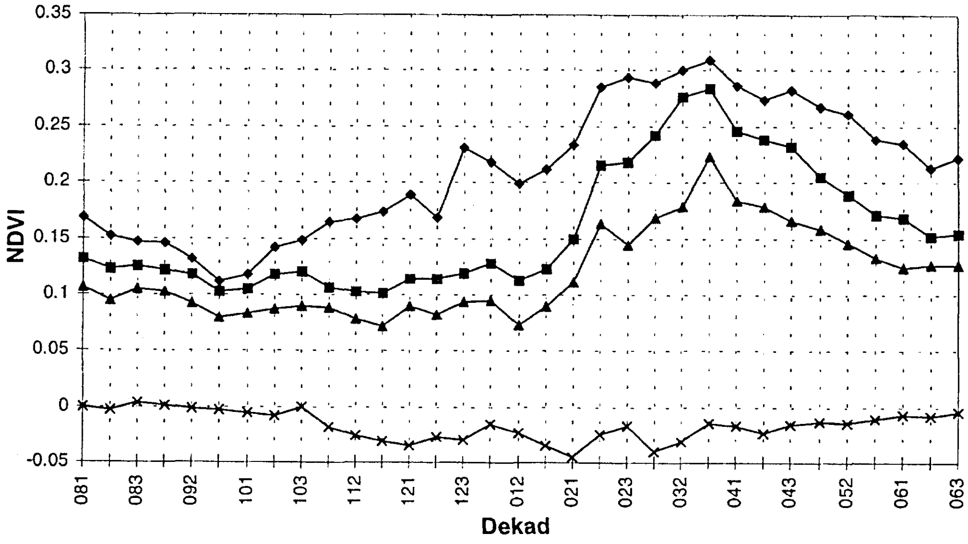


Figure 3. Average profiles over a 10-year period for four land cover types in Etosha, from the ARTEMIS NDVI archive for (– × –) Bare Ground, (–▲–) Grassland/Steppe, (–■–) Shrub Savanna and (–◆–) Tree Savanna.

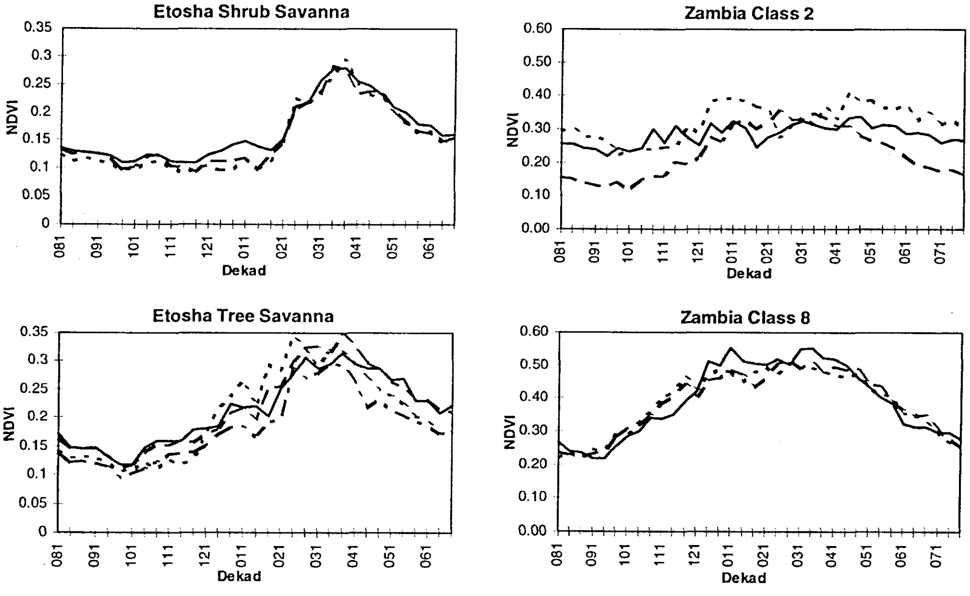


Figure 4. Comparison of 10-year averaged NDVI time response for different locations in selected classes in (*a*) Etosha, and (*b*) Zambia.

about three months earlier for Tree Savanna and the maximum NDVI is also always higher.

In Zambia, the maximum likelihood classification was thresholded so as to include only the pixels which had at least a 95 per cent probability to belong to a class, leaving 20 per cent of the pixels classified compared with the original classification. In the absence of any suitable surrogate information, this was to identify the purest pixels to be used to extract the historical NDVI data representative of each class. Candidate sites were selected in groups of pixels of the same class. Three well spaced sites were selected per class. Class 1 was not included in the analysis because it represented water bodies. Cloudy pixels were removed using the same method as in Etosha. An average profile for each class is shown in figure 2 and individual profiles for each site of selected classes in figure 4(b). The variation between classes is generally greater than within each class although class 2 exhibits considerable variation because it occurred mostly on the edge of water bodies and the variation is attributed to mixed pixels and changes of water levels.

The probability distribution of the NDVI for each dekad during the growing season and each stratum was calculated with the methodology used for assessing the probability of extreme hydrological events (Linsley *et al.* 1975). The NDVI values extracted from the historical data were ranked from lowest to highest for each dekad and stratum. This enabled us to compute the probability (p) of having an NDVI less or equal to a given value by applying the formula defined by Weibull (1939): $p = m/n + 1$, where m is the rank and n is the number of years. This can also be expressed as a return period, $T_r = 1/p$, which is the average number of years between occurrences of the event (Linsley *et al.* 1975). The probability was plotted against the corresponding NDVI values as for hydrological events.

In hydrology, the aim is often to extrapolate the distribution in order to predict the size of events which have very low return periods and special statistical distributions are fitted to the data to do this. In the present case, the time series of 10 years is too short to fit any distribution. Therefore, a simple least square fit polynomial was used to interpolate estimates of the NDVI for specified probabilities. This enabled us to calculate quintile ranges of NDVI for each vegetation class and each dekad and to define the five classes in table 1.

Example probability plots for the first dekad of March are shown in figure 5 for each of the Etosha main vegetation communities. One can clearly see in figure 5 that a given NDVI will be associated with a different probability level depending on the vegetation type. For instance, an NDVI of 0·3 corresponds to a very high probability for Grassland/Steppe, average for Shrub Savanna and very low for Tree Savanna.

NOAA-AVHRR HRPT from local receiving stations provided images daily with sufficient coverage to enable vegetation monitoring in near real-time in both study areas. A receiving station was installed at the Etosha Ecological Institute, Okaukuejo in 1993 to monitor the vegetation dynamics and fire risks. NDVI images are produced daily but are found to be difficult to interpret by the Park management who are not specialists in remote sensing. We aimed to produce VPI maps with the highest possible spatial detail and with the local reception facility we had the advantage of

Table 1. Description of VPI classes in relation to probability and return periods.

VPI class	Probability level	Return period (years)
Very low	$p<0·2$	$T_r>5$
Low	$0·2<p<0·4$	$5<T_r<2·5$
Average	$0·4<p<0·6$	$T_r<2·5$
High	$0·6<p<0·8$	$5<T_r<2·5$
Very high	$p>0·8$	$T_r>5$

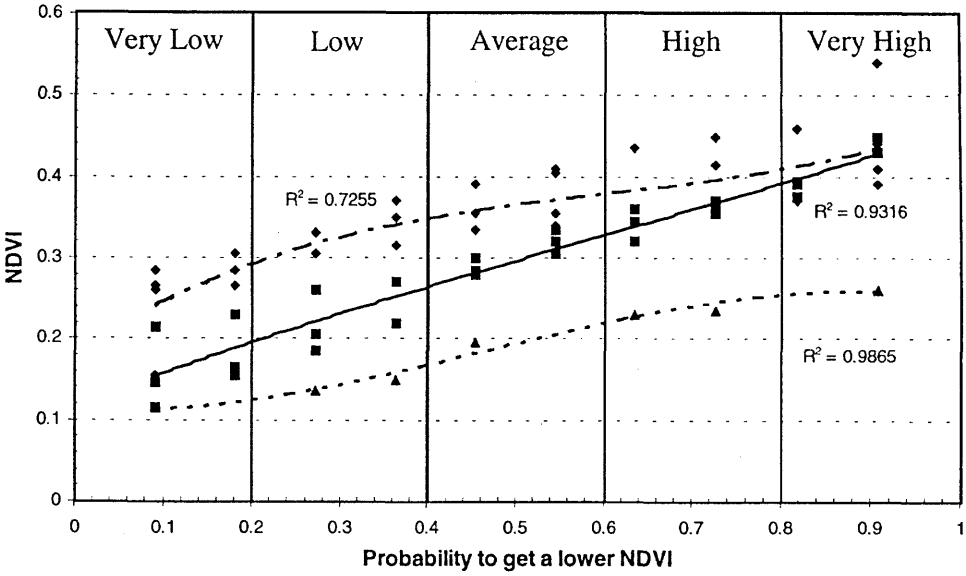


Figure 5. Determination of NDVI probability distribution for (–▲–) Grassland/Steppe, (–■–) Shrub Savanna and (–◆–) Tree Savanna in Etosha using a polynomial fit for the first dekad of March and limits of the Vegetation Productivity Indicator (VPI) classes, Very Low to Very High.

being able to select the highest quality LAC image to represent the reporting period. We used single images which were near nadir to get the highest spatial resolution. Visual inspection was used to select images with the least cloud cover and atmospheric interference. LAC images were geo-referenced using the TM based vegetation map of the Park. Final RMS (root mean square) errors were between 0·5 and 0·9 pixel and the LAC images were re-sampled using cubic convolution and a 500 m output pixel size.

In Zambia, the receiving station was installed in 1993 in the Meteorological Department and our objective was to cover the whole country so the ARTEMIS spatial resolution was used. Therefore, retaining maximum spatial detail was not required and images were geo-referenced using the standard processing chain of the receiving station software. This consisted of correcting the image using orbital parameters and manual shifting to fit the vector map of the area. This method leads to a geometric accuracy of about 1–2 LAC pixels. ARTEMIS pixels are produced by a combination of sampling and averaging LAC pixels. Since it was not possible to replicate this procedure, the arithmetic average of the LAC pixels was used in the change of pixel size.

The AVHRR channels 1 and 2 were radiometrically corrected using the method described by Kaufman and Holben (1993). NDVI images were computed and a further correction was applied in Etosha National Park to minimise atmospheric noise by using methodology similar to Taylor *et al.* (1992) and part of the Etosha pan as a standard target. The NDVI of selected sites in the Pan were extracted and averaged and this value was subtracted from all the pixels in the image. It can be argued that the pan is a feature big enough to affect the atmosphere. However, the study area surrounds the Pan and therefore, is likely to be affected by the same atmospheric interferences. The equivalent correction was also applied to the ARTEMIS data so that NDVI values computed for each source were compatible.

The empirical correction against a standard target is useful in real-time applications because the accurate satellite calibration parameters are not always available and it helps to take into account sensor degradation which is unknown at the time. However, raw data should be kept to allow post-processing when new calibration parameters become available. In Zambia, the maximum value composite images were used to remove atmospheric interference. This was not applied in Etosha because compositing reduces the spatial detail.

The following procedure was used to produce the VPI maps. The appropriate probability distribution for the NDVI was determined from the date and by reference to its position in the stratification map. The probability of having an NDVI equal to or smaller than the current value was calculated from the polynomial equation and this was used to assign it to the appropriate VPI class defined by table 1. This process was repeated for the whole image, producing the VPI map for the period considered.

3. Results and discussion

The graphs in figure 6 show how the NDVI thresholds for selected classes vary through the season. The status of specific sites can be followed by plotting the NDVI values on these 'templates' as shown for two different seasons for a maize growing area in Kalomo District, Zambia in figure 7. The average NDVI of several contiguous pixels was used to minimise local spatial variation and the resulting time series was smoothed using a three-term moving average to reduce temporal noise. In the 1987–88 season, maize production was 15 per cent of the average between 1982 and 1991 for Kalomo District whereas in 1981–82 it was only 55 per cent of the average. The area that was selected is not meant to be representative of the whole country but around 8 per cent of Zambia's maize is produced in Kalomo and some very useful conclusions can be drawn from the scrutiny of the graph. The 1981–82 season started with high VPI probably related to an early start of the rains which did not continue and consequently, the VPI became very low throughout most of the season. This scenario is particularly damaging because farmers start sowing as soon as rains begin and if rain stops shortly after, most of the seeds are lost and many farmers cannot afford to sow a second time, hence the much reduced production. In 1987–88, the start of the wet season seemed late but once it began, the rains were probably well distributed through the season, inferred by the late rise in the NDVI which continued until the end of January when a very high status was reached and there was no drop during the main part of the growing season. Although the start of the growing season was one or two dekads late, the crop was well watered once planting was done and yielded much higher than normal.

In Etosha VPI maps were produced from images acquired during the 1995 season. The comparison of the VPI map and the NDVI image in figure 8 illustrates the difficulty to interpret NDVI images and how VPI maps solve the problem. At point A of figure 8 (a) the NDVI is fairly low compared to the rest of the image but the VPI is high. This is because the vegetation type for that area is Grassland/Steppe which normally has a lower NDVI than any other vegetation type. Point B is situated in a Tree Savanna region and has a higher NDVI than point A but the VPI in the vicinity is low to very low. Tree Savanna is normally expected to have the

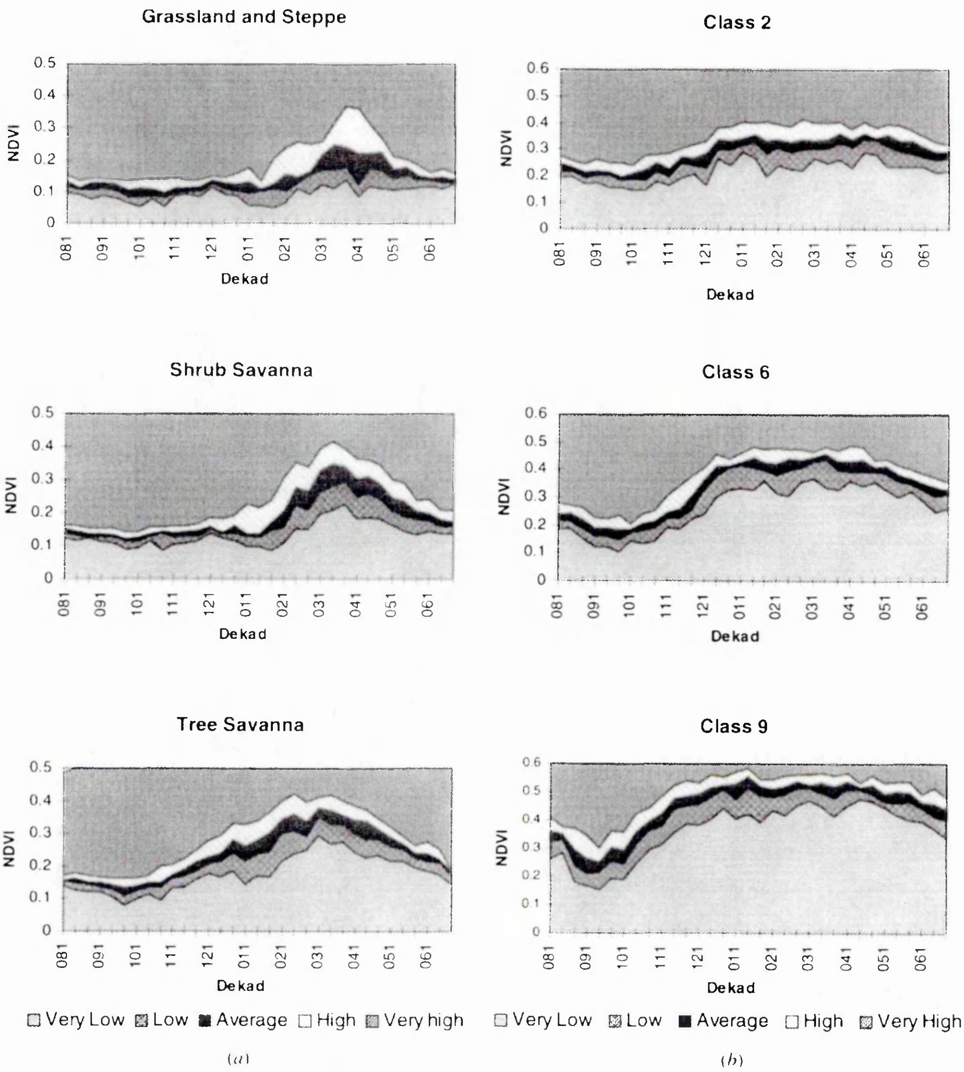


Figure 6. Seasonal variations of Vegetation Productivity Indicator classes, for selected classes in (*a*) Etosha, and (*b*) Zambia.

highest NDVI values in Etosha but the VPI map shows current values are much lower than normal.

Various studies have shown there is a strong relation between the NDVI and rainfall in semi-arid environments (Hess *et al.* 1996, Di *et al.* 1994, Bonifacio *et al.* 1993, Davenport and Nicholson 1993, Justice *et al.* 1991). In Etosha, daily rainfall is measured at five stations in the Park and we explored the correspondence there between rainfall and VPI to evaluate the performance of VPI maps produced at or near the end of January, February and March, 1995.

The rainfall data were cumulated from the beginning of January because although it often rains in November and December, these are isolated storms and contribute marginally to the net primary biomass production for a whole season. The main rains usually occur from January through March. For comparison, the cumulative

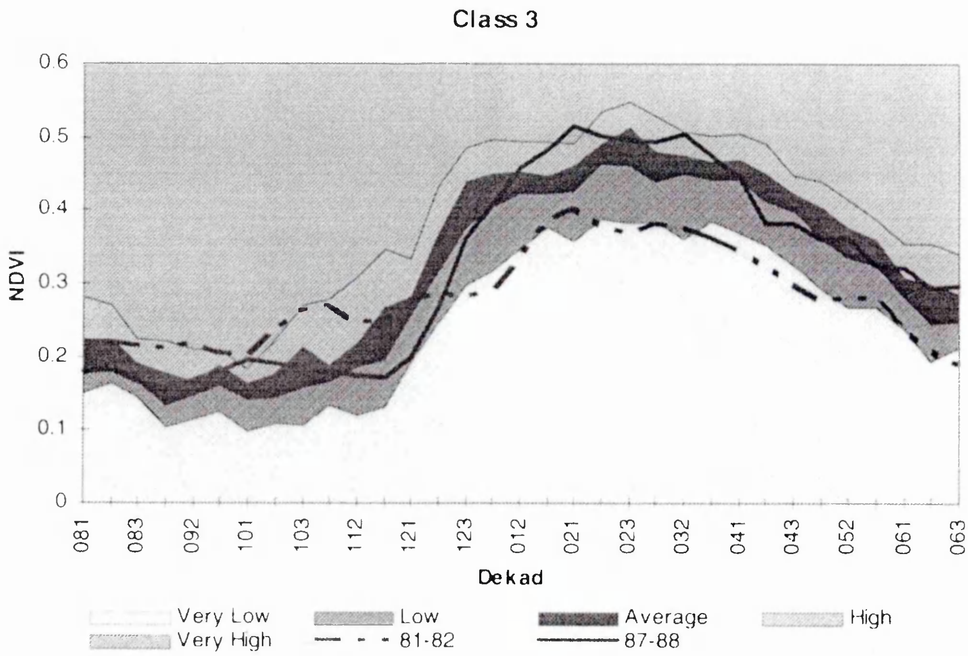


Figure 7. Comparison of the Vegetation Productivity Indicator in two different seasons (1981–82 and 1987–88) in Zambia for the same location in the Kalomo district.

rainfall up to each date of VPI map was ranked for the 1982–91 period enabling us to determine quintile ranges in a similar analysis to that of the NDVI. The rainfall status thus determined is compared to the VPI at the five rainfall stations in figure 9 for the 1995 season. The rainfall and VPI classes follow the same trend across the Park spatially and temporally and there was never more than a one class difference between them.

In Zambia, we compared rainfall status at 30 rainfall stations distributed throughout the territory with the VPI of the nearest pixel in the ARTEMIS NDVI images for the period 1982–91. NDVI time series were extracted for the end of March which marks the end of the rainy season. We did not use an average of several pixels as advised in some studies (Kogan 1990) because of the 7·6 km pixel size and the localised nature of African rainfall events. The cumulative seasonal rainfall up to two dekads before the NDVI image was used. The Spearman's rank correlation coefficient, $\mathbf{r}_s$ (Steel and Torrie 1960, Burt and Barber 1996) was used to compare the vegetation and rainfall status classes at each station. Ordinal measures of association, or rank correlation coefficients, are appropriate whenever the relation between two variables is monotonic increasing or decreasing. In our case, a monotonic increasing function is expected, i.e., the higher the VPI the higher the rainfall status. The results in table 2 generally show a very high correlation between vegetation and rainfall status over time. Most values of $\mathbf{r}_s$ are above 0·8 and are highly significant. Only Mbala station shows a $\mathbf{r}_s$ which is not significant at the 95 per cent confidence level. This station is situated next to Lake Tanganyika, and at the scale of ARTEMIS the pixel may cover a mixture of land and water from time-to-time, due to image mis-registration which influences the NDVI in a way uncorrelated to the rainfall. The next lowest values of $\mathbf{r}_s$ are for two rainfall stations located in the north-west of the country, close to the

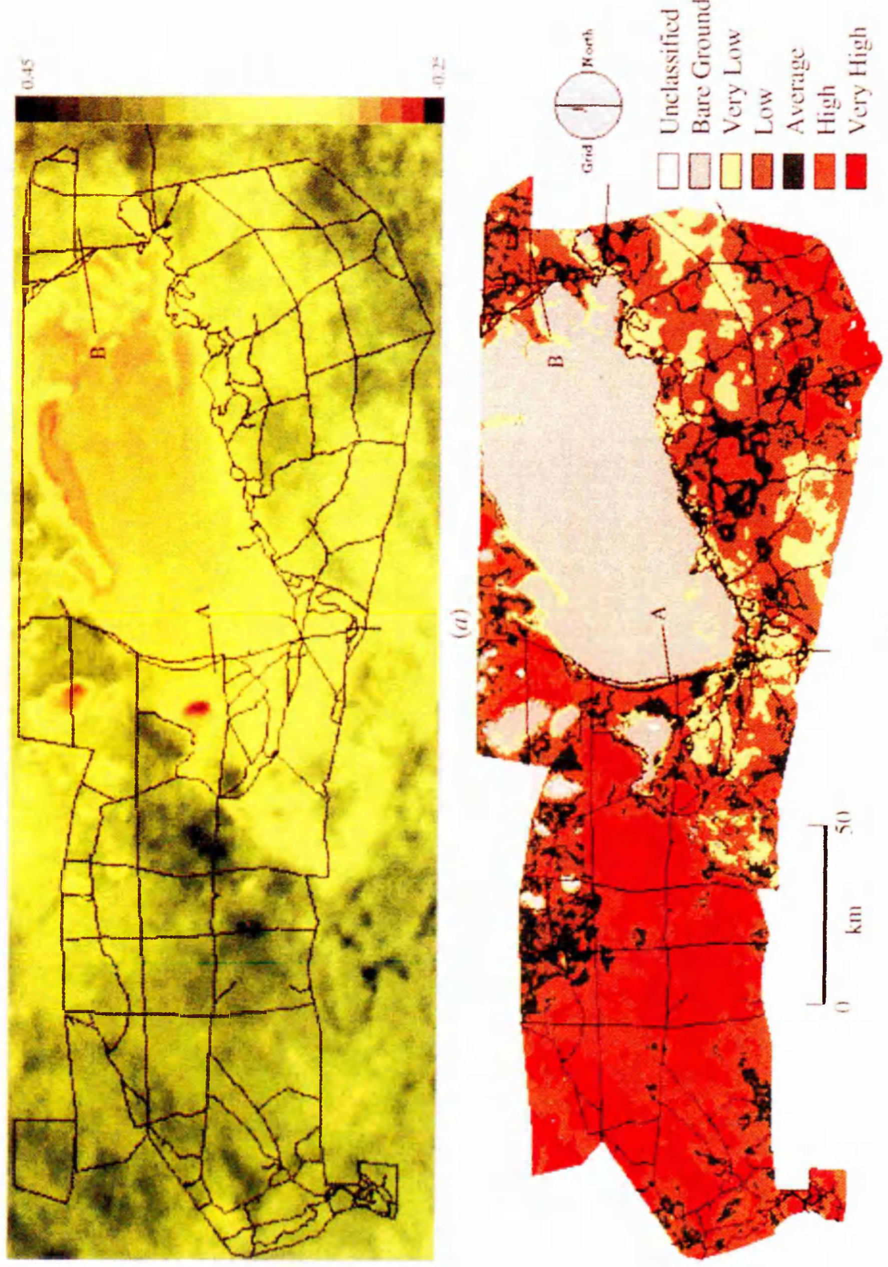


Figure 8. NDVI image (a) and Vegetation Productivity Indicator (VPI) map, (b) of Etosha National Park, Namibia, 27 March 1995. Points A and B have similar NDVI values but A is in grassland and has a high VPI and B is in tree savanna with low VPI.

Table 2. Results of Spearman's rank correlation between rainfall and VPI for 30 rainfall stations in Zambia.

Stations	r_s	t	Significance level (%)	Latitude	Longitude
SENANG01	0·97	11·226	99·9	16° 7′S	23° 16′E
LUSAKA02	0·96	10·200	99·9	15° 19′S	28° 27′E
MFUWE001	0·96	10·200	99·9	13° 16′S	31° 56′E
KABOMP01	0·93	7·005	99·9	13° 36′S	24° 12′E
MUMBWA01	0·91	6·172	99·9	14° 59′S	27° 4′E
CHOMA001	0·90	5·946	99·9	16° 51′S	27° 4′E
KAFUE001	0·90	5·946	99·9	15° 46′S	27° 55′E
KABWE001	0·90	5·739	99·9	14° 25′S	28° 29′E
KASAMA01	0·90	5·739	99·9	10° 13′S	31° 8′E
LUSAKA01	0·89	5·548	99·9	15° 25′S	28° 19′E
PETAUK01	0·89	5·548	99·9	14° 15′S	31° 17′E
MTMAKU01	0·88	5·209	99·9	15° 33′S	28° 15′E
SERENJ01	0·88	5·209	99·9	13° 14′S	30° 13′E
KASEMP01	0·87	4·914	99·0	13° 32′S	25° 51′E
KAWAMB01	0·87	4·914	99·0	9° 48′S	29° 5′E
NDOLA001	0·87	4·914	99·0	13° 0′S	28° 39′E
CHIPAT01	0·85	4·654	99·0	13° 33′S	32° 35′E
ZAMBEZ01	0·84	4·422	99·0	13° 32′S	23° 7′E
MAGOYE01	0·82	4·117	99·0	16° 8′S	27° 38′E
MANSA001	0·82	4·025	99·0	11° 6′S	28° 51′E
LIVING01	0·81	3·937	99·0	17° 49′S	25° 49′E
MONGU002	0·81	3·937	99·0	15° 15′S	23° 9′E
KAOMA001	0·81	3·852	99·0	14° 48′S	24° 48′E
SOLWEZ01	0·79	3·693	99·0	12° 11′S	26° 23′E
MISAMF01	0·78	3·477	99·0	10° 11′S	31° 13′E
MPIKA001	0·76	3·283	98·0	11° 45′S	31° 26′E
ISOKA001	0·74	3·106	98·0	10° 10′S	32° 38′E
KAFIRO01	0·72	2·894	95·0	12° 36′S	28° 7′E
MWINIL01	0·66	2·489	95·0	11° 45′S	24° 26′E
MBALA001	0·59	2·088	90·0	8° 51′S	31° 20′E

border with Zaire where the average rainfall is above 1300 mm for both stations and during the 1982–91 period, varied between 900 and 1500 mm. Here, the lower correlation may be because the rainfall is less of a limiting factor for vegetation growth.

We have shown that the VPI reflects temporal variations in the rainfall at a wide range of locations. The VPI maps show the spatial pattern and thus allow the mapping of drought affected areas. To demonstrate this further, we have investigated the relation between VPI and maize production in the group of Districts in southern Zambia, shown in figure 10 which are responsible for about 70 per cent national production, maize being by far the most important food crop. In 1988, maize production was 163 per cent of the norm and the VPI, in figure 10(*a*) is high or very high, contrasting with figure 10(*b*) in January 1995 when the maize production was only 62 per cent of the norm and the VPI is very low.

The frequency distribution of the VPI in maize producing districts, in the second dekad of January, were determined for the period 1982–91 and for 1995, corresponding to the peak time of maize vegetative development (Azzali 1991). In 1988, cloud cover was a problem so the third dekad of January was used instead. The results presented in figure 11 are ranked according to the level of production expressed as

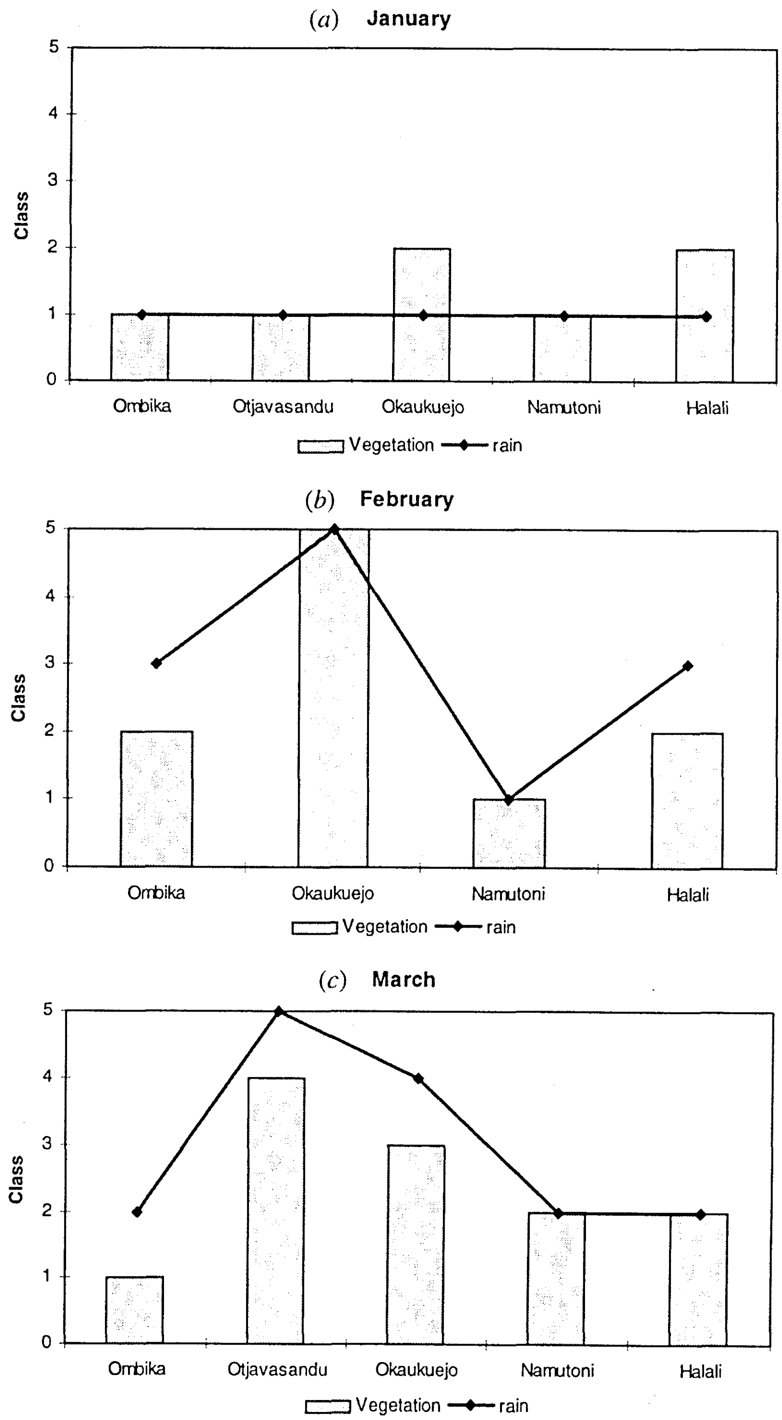


Figure 9. Comparison of the VPI and rainfall status classes (1 = very low, 2 = low, 3 = average, 4 = high, 5 = very high) between January and March 1995 for the five rainfall stations of Etosha National Park. Otjavasandu was covered by clouds in February therefore no comparison was available.

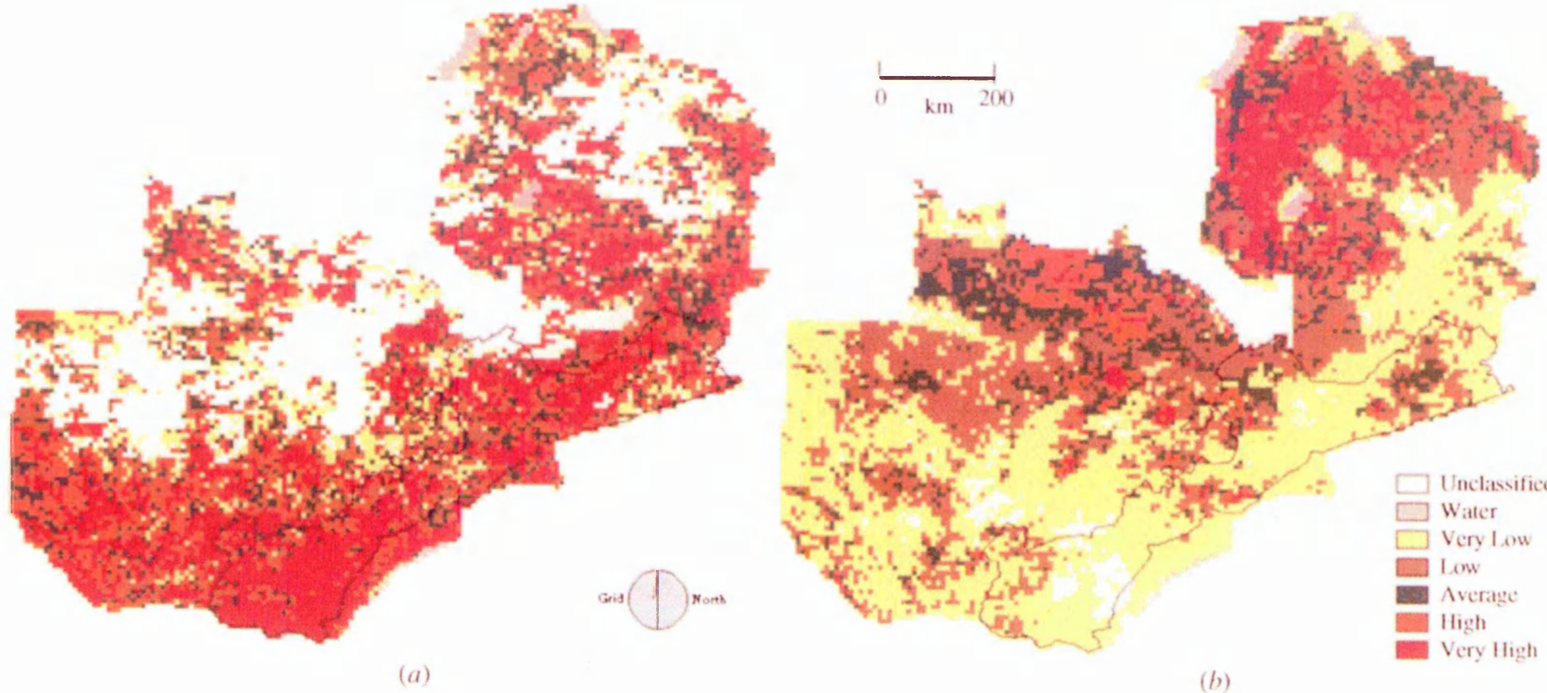


Figure 10. Vegetation Productivity Indicator maps of Zambia (*a*) third dekad of January, 1988, and (*b*) second dekad of January 1995. Boundary of main maize growing region is outlined in black.

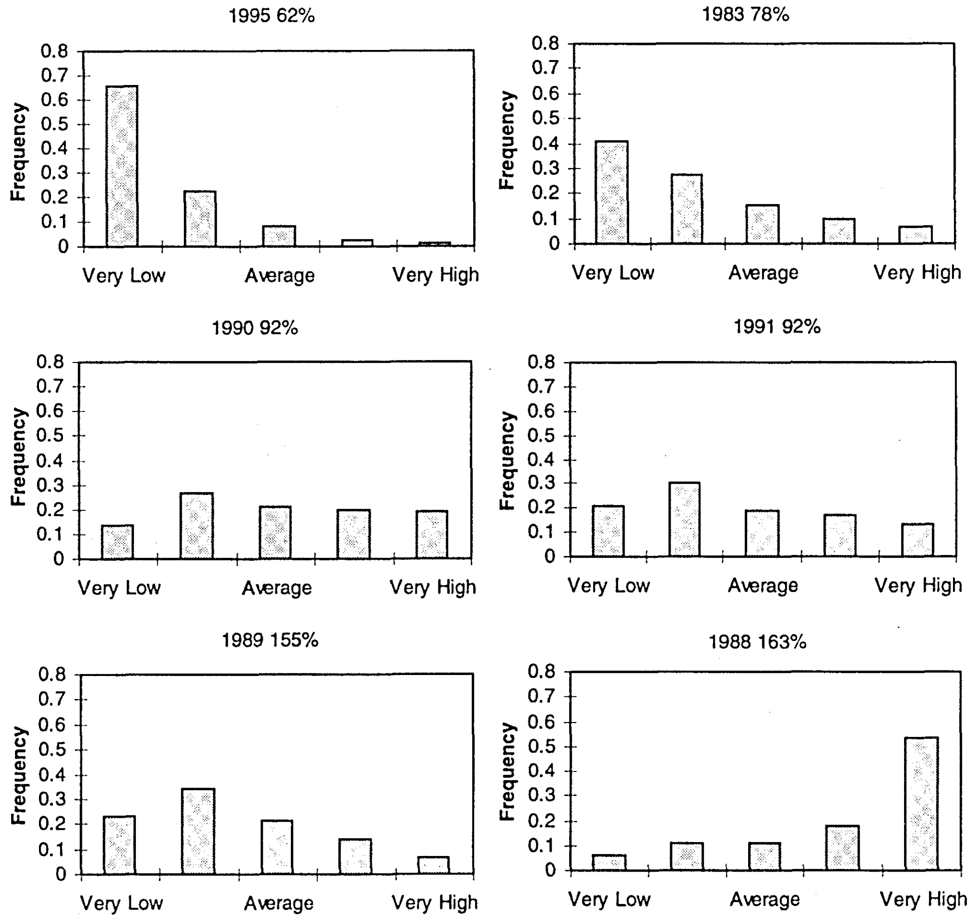


Figure 11. Distribution of the Vegetation Productivity Indicator for selected years with different maize production, expressed as a percentage of the average production between 1982 through 1991.

a percentage of the average for the period 1982–91. One would expect to have a skewed distribution toward the very low class for very low production years and a skewed distribution toward the very high class for very high production years and in general this is demonstrated, particularly in years with low production. However, this is not the case in 1989 with 155 per cent of the average production, where the low VPI is the most frequent class. A likely explanation is a reduction in the NDVI caused by cloud contamination of images during the growing season. In years when production is higher than average, rainfall is higher which means more cloud cover. In 1989, the second dekad of January image did not seem particularly cloudy over the study area but both the dekad before and after were badly affected. It is a well-known fact that when there are many clouds, the maximum value composite might comprise of a single cloud free image which can then contain adverse atmospheric effects. However, in years with low production, because of lower rainfall, cloud is less a problem which then makes the method more reliable. It is crucial to identify cloud cover problems when interpreting the VPI map and this can be easily done in the context of local data reception when each image is routinely assessed visually.

Figure 12 shows the relation between national annual maize production and the weighted average VPI in the study Districts, except for 1989 for the reasons mentioned above. The coefficient of determination of 0·60 is highly significant ($p = 0{\cdot}009$ for the F significance test) and demonstrates the potential for developing quantitative estimates of drought impact.

4. Conclusion and further development

We developed and demonstrated methodology to monitor the condition of vegetation, independent of type, using statistical techniques for estimating extreme events, borrowed from the field of hydrology and applied to NDVI images acquired by local receiving stations using archived NDVI imagery for the time-series analysis.

The VPI thus defined and mapped was highly correlated to rainfall and thus enabled the identification of drought affected areas. In Zambia, the weighted average VPI was highly correlated to maize production and this indicates a possibility for quantitative assessment of drought impact.

The current methodology has been installed for operational use in Etosha National Park where the information is routinely available to scientists studying wildlife movements and population dynamics and to managers monitoring fire risk and managing burning programs. Shortly, we expect to install the method for operational use in the Zambian Meteorological Department where it will contribute information to the Crop Weather Bulletin which is used by governmental departments and motivated farmers.

The methodology is generally applicable and with the increasing availability of global data sets such as the NOAA Pathfinder programme and the possibility to acquire current AVHRR images on the internet through the NOAA/SAA (Satellite Active Archive) it should be usable anywhere on the globe.

The methodology was better at identifying conditions with lower than normal NDVI because the occurrence of higher than usual NDVI is accompanied by higher

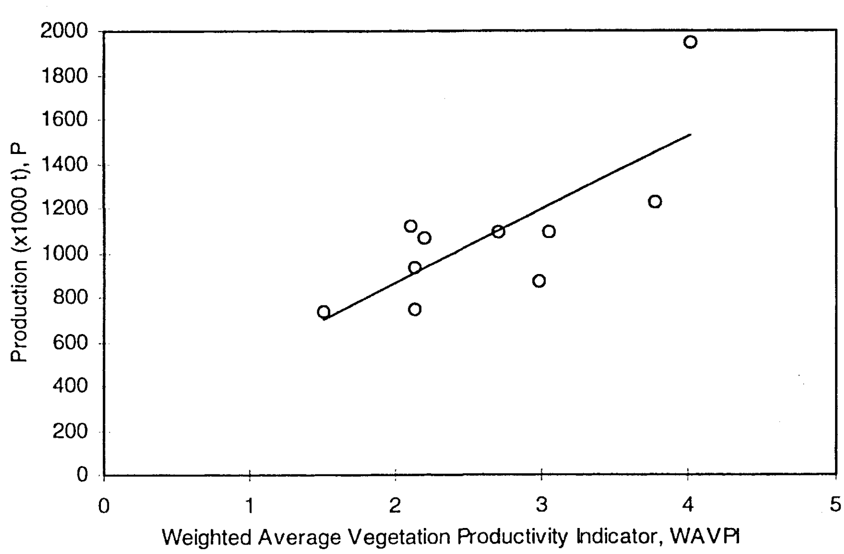


Figure 12. Relation between total Zambian maize production and the weighted average Vegetation Productivity Indicator over the main maize growing region of Zambia for 1982–88, 1990–91 and 1995, $P = 331\,WAVPI + 201$, $r^2 = 0{\cdot}6$.

rainfall and atmospheric interference during the growing season, thus making the use of maximum value compositing less effective for removing low, cloud-contaminated NDVI values.

The development of a VPI map based on an integrated NDVI is a future objective, to assess end of season conditions. It is also necessary to investigate further the sensitivity of the VPI to variations in rainfall and primary production. Archived AVHRR NDVI data is now available for up to 15 years, and the extra five years of data should be added to improve the capability to identify extreme events. The addition of extra historical data should also improve the robustness of the method and a more quantitative use of the VPI maps should also be investigated for crop yield forecasting.

Acknowledgments

This work is part of an on-going collaborative effort with the Etosha Ecological Institute of the Ministry of Environment and Tourism, Namibia and the Zambian Meteorological Department, funded by the U.K. Overseas Development Administration and coordinated by the Natural Resources Institute, Chatham, U.K. Dr Graham Thomas should also be acknowledged for providing the C programs.

References

AZZALI, S., 1991, Interpretation of crop growth patterns by means of NDVI- time series in Zambia. *Geocarto International*, **3**, 15–26.

BONIFACIO, R., DUGDALE, G., AND MILFORD, J. R., 1993, Sahelian rangeland production in relation to rainfall estimates from Meteosat. *International Journal of Remote Sensing*, **14**, 2695–2711.

BURT, J. E., and BARBER, G. M., 1996, *Elementary Statistics for Geographers* (New York, London: The Guilford Press).

DAVENPORT, M. L., and NICHOLSON, S. E., 1993, On the relation between rainfall and Normalized Difference Vegetation Index for diverse vegetation types in East Africa. *International Journal of Remote Sensing*, **12**, 2369–2389.

DI, L., RUNDQUIST, D. C., and HAM, L., 1994, Modelling relationships between NDVI and precipitation during vegetative growth cycles. *International Journal of Remote Sensing*, **15**, 2121–2136.

DIALLO, O., DIOUF, A., HANAN, N. P., NDIAYE, A., and PRÉVOST, Y., 1991, AVHRR monitoring of savanna primary production in Senegal, West Africa: 1987–1988. *International Journal of Remote Sensing*, **12**, 1259–1279.

EDMONDS, A. C. R., 1976, *The Republic of Zambia, Vegetation Map 1:500 000* (Forest Department, the Government of Zambia).

ERDAS, 1993, *ERDAS Field Guide* (Atlanta: ERDAS Inc.).

FLITCROFT, I. D., MILFORD, J. R., and DUGDALE, G., 1989, Relating point to area rainfall in semi-arid West Africa and the implications for rainfall estimates derived from satellite data. *Journal of Applied Meteorology*, **58**, 193–207.

GROTEN, S. M. E., 1993, NDVI-crop monitoring and early yield assessment of Burkina Faso. *International Journal of Remote Sensing*, **14**, 1495–1515.

HESS, T., STEPHENS, W., and THOMAS, G., 1996, Modelling NDVI from decadal rainfall data in the North East Arid Zone of Nigeria. *Journal of Environmental Management*, **48**, 249–261.

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

HUTCHINSON, C. F., 1991, Uses of satellite data for famine early warning in sub-Saharan Africa. *International Journal of Remote Sensing*, **12**, 1405–1421.

JUSTICE, C. O., DUGDALE, G., TOWNSHEND, J. R. G., NARRACOTT, A. S., and KUMAR, M., 1991, Synergism between NOAA-AVHRR and Meteosat data for studying vegetation development in semi-arid West Africa. *International Journal of Remote Sensing*, **12,** 1349–1368.

KAUFMAN, Y. J., and HOLBEN, B. N., 1993, Calibration of the AVHRR visible and near- IR bands by atmospheric scattering, ocean glint and desert reflection. *International Journal of Remote Sensing*, **14,** 21–52.

KENNEDY, P., 1989, Monitoring the vegetation of Tunisian grazing lands using the Normalized Difference Vegetation Index. *Ambio*, **18,** 119–123.

KOGAN, F. N., 1990, Remote sensing of weather impacts on vegetation in non-homogeneous areas. *International Journal of Remote Sensing*, **11,** 1405–1411.

LAMBIN, E. F., CASHMAN, P., MOODY, A., PARKHURST, B. H., PAX, M. H., and SCHAAF, C. B., 1993, Agricultural production monitoring in the Sahel using remote sensing: present possibilities and research needs. *Journal of Environmental Management*, **38,** 301–322.

LE COMTE, D. M., 1989, Using AVHRR for early warning of famine in Africa. *Photogrammetric Engineering and Remote Sensing*, **55,** 168–169.

LE ROUX, C. J. G., GRUNOW, J. O., MORRIS, J. W., BREDENKAMP, G. J., and SCHEEPERS, J. C., 1988, A classification of the vegetation of the Etosha National Park. *South African Journal of Botany*, **54,** 1–10.

LINSLEY, R. K., KOHLER, M. A., and PAULHUS, J. L. H., 1975, *Hydrology for Engineers* (New York, Toronto, London: McGraw-Hill).

LOSS, S. O., 1993, Calibration adjustment of the NOAA AVHRR normalised difference vegetation index without recourse to channel 1 and 2 data. *International Journal of Remote Sensing*, **14,** 1907–1917.

MASELLI, F., CONESE, C., PETKOV, L., and GILABERT, M. A., 1993, Environmental monitoring and crop forecasting in the Sahel through the use of NOAA-NDVI data. A case study: Niger 1986–89. *International Journal of Remote Sensing*, **14,** 3471–3487.

MOYO, S., O'KEEFE, P., and SILL, M., 1993, *The Southern African Continent*, Profiles of the SADC Countries (London: Earthscan Publications Ltd).

PRINCE, S. D., 1991, Satellite remote sensing of primary production: comparison of results for Sahelian grasslands 1981–1988. *International Journal of Remote Sensing*, **12,** 1301–1311.

PRINCE, S. D., and ASTLE, W. L., 1986, Satellite Remote Sensing of Rangelands in Botswana. I. Landsat MSS and Herbaceous Vegetation. *International Journal of Remote Sensing*, **7,** 1533–1553.

RASMUSSEN, M. S., 1992, Assessment of millet yields and production in northern Burkina Faso using integrated NDVI from the AVHRR. *International Journal of Remote Sensing*, **13,** 3431–3442.

SCHULTZ, J., 1975, Republic of Zambia, Land Use 1:750 000 (University of Zambia, Ministry of Rural Development, the Government of Zambia).

STEEL, R. G. D., and TORRIE, J. H., 1960, *Principles and Procedures of Statistics* (New York, Toronto, London: McGraw-Hill).

TAYLOR, J. C., D'SOUZA, G., and COLEMAN, V. R., 1992, Vegetation Conditions and Yield Indicators in England and Wales using NOAA-AVHRR data (Monitoring of Vegetation and Yield Indicators). *Proceedings of the conference on the Applications of Remote Sensing to Agricultural Statistics held in Villa Carlotta, Belgirate, Lake Maggiore, Italy, on 26–27 November 1991* (Brussels, Luxembourg: Commission of the European Communities), p. 249–256.

TAYLOR, J. C., SANNIER, C. A. D., DELINCÉ, J., and GALLEGO, F. J., 1996, Regional Crop Inventories in Europe Assisted by Remote Sensing: 1988–1993. Synthesis Report of the MARS Project—Action 1. Final report, Commission of the European Communities, Joint Research Centre, Ispra Establishment.

WEIBULL, W., 1939, A statistical Theory of the Strength of Materials, *Ing. Vetenskapsakad. Handl. (Stockh.)*, **151,** 15.

WILLIAMS, J. B., and ROSENBERG, L. J., 1993, Operational reception, processing and application of satellite data in developing countries: Theory and practise. In: *Towards Operational Applications, Proceedings of 19th Annual Conference of the Remote Sensing Society, Chester, U.K.*, on 16–17 September 1993 edited by K. Hilton (Chester: The Remote Sensing Society), pp. 76–81.

CHAPTER 4

INT. J. REMOTE SENSING, 1998, VOL. 19, NO. 17, 3441–3450

Compatibility of FAO-ARTEMIS and NASA Pathfinder AVHRR Land NDVI data archives for the African continent

C. A. D. SANNIER, J. C. TAYLOR

School of Agriculture, Food and Environment, Cranfield University, Silsoe, Bedford, MK45 4DT, England, UK

and K. CAMPBELL

Natural Resources Institute, Central Avenue, Chatham Maritime, Chatham, Kent ME4 4TB, England, UK

(Received 22 December 1997; in final form 1 July 1998)

Abstract. There is a need for a consistent time-series of NDVI data for vegetation monitoring in Africa. In this work, we explore the compatibility of the FAO-ARTEMIS and NASA Pathfinder AVHRR Land NDVI archives. Values from the two archives were found to be significantly and systematically different even for the recently corrected form of the Pathfinder archive. Differences are attributable to the different processing chains used on essentially the same raw data. The two archives should not be mixed together for time-series analysis of NDVI, without correction. Although differences varied for periods of the record covered by different satellites, these were practically small and we present a single linear relation to correct one series values to the same basis as the other.

1. Introduction

NOAA AVHRR is a recognised source of information for monitoring vegetation resources (Lambin *et al.* 1993). NDVI images representing current conditions are compared to historical data and departures from normal conditions are determined using techniques such as the ones described by Kogan (1990) or by Sannier *et al.* (1998). These techniques depend on the existence of consistent archives. This work investigates the compatibility of the FAO-Africa Real Time Environmental Monitoring Information System (ARTEMIS) archive with the NASA Pathfinder AVHRR Land (PAL) archive for the African continent.

The ARTEMIS data, has been widely used (Groten 1993, Hielkema and Snijders 1994, Sannier *et al.* 1998) and comprises 356 images covering a period from the first dekad (10-day period) of August 1981 to the third dekad of June 1991. The image for the first dekad of January 1983 is missing. Currently, no update is planned and there is a need to expand the period covered in order to improve the reliability of time-series analyses. The data are derived from the FAO-ARTEMIS/NASA-GSFC (National Aeronautic and Space Administration-Goddard Space Flight Centre) archive and are based on NOAA-AVHRR GAC (Global Area Coverage) data and consist of 10-day maximum value composite (MVC) NDVI images. Radiometric corrections are applied directly to the NDVI without recourse to the individual channel 1 and 2 data whilst also accounting for sensor degradations (Los 1993). The spatial resolution is 7.6 km and the Hammer-Aitoff projection was used. Values are stored in 8-bit form with the following relation between the NDVI and Digital Numbers (DN).

$$\mathrm{NDVI} = \frac{\mathrm{DN} - 82}{256} \tag{1}$$

Thus the maximum NDVI accommodated is 0.68. Clouds are coded as 0, equivalent to a NDVI of −0.32. A complete description of ARTEMIS data processing is given by Hielkema and Snijders (1994).

The PAL archive also originated from the NASA-GSFC data but the processing is different to ARTEMIS. As of June 1998, the archive comprised 474 10-day MVC NDVI images covering a period between the second dekad of July 1981 and the first dekad of September 1994. The PAL archive covers the entire globe but continental subsets are available. The archive includes 10-day MVC NDVI images, monthly MVCs and individual channel data. The data processing, described by James and Kalluri (1994) consists of: navigating GAC images; radiometric calibration based on the work published by Rao (1993); atmospheric corrections for Rayleigh scattering; cloud testing; binning to 8-km pixel size using the Goode Homolosine projection described by Steinwand (1994); NDVI calculation and maximum value compositing. NDVI images are stored in 8-bit form with the following relation between NDVI and DN:

$$\mathrm{NDVI} = \frac{\mathrm{DN} - 128}{128} \tag{2}$$

This accommodates NDVI values between −1 and +1. Thus the radiometric resolution of ARTEMIS (0.004 NDVI per DN) is higher than that of the PAL (0.008 NDVI per DN).

Recently, errors in the PAL processing were discovered and a complete post-processing of the 10-day composite data was carried out. The errors were in the calculation of the solar zenith angle and in the atmospheric correction. Although the first did not affect the NDVI, the second had a considerable impact. The post-processed data have been made available to the public on the Goddard DAAC (Distributed Active Archive Centre) Internet site at the beginning of 1998 and a full description is made by Smith (1998). Because, some users might already have the old PAL data set, both PAL data sets were analysed for this letter. Table 1 shows variations in satellite sensors used in each archive.

2. Comparison of individual images

The NDVI values at equivalent pixel locations were compared in randomly chosen pairs of images. These were first brought to the same projection using third order transform equations to account for non-linear differences in the projections at continental scale. Transform coefficients were determined by the method of least

Table 1. Satellite data sources for the ARTEMIS and PAL archives.

Satellite	ARTEMIS dekads	PAL dekads
NOAA-7	81081–85011	81072–85021
NOAA-9	85012–88103	85022–88111
NOAA-11	88111–91063	88112–94091

Table 2. Reduced Major Axis (RMA) linear fit results for a selection of dekadal images from the ARTEMIS and Pathfinder data set, *ARTEMIS = f (PAL)*.

Dekad	Satellite	Original PAL archive					Post-processed PAL archive				
		Slope	Intercept	n	R^2	p	Slope	Intercept	n	R^2	p
84031	NOAA-7	0.873	0.053	1279	0.94	<0.001	0.756	0.025	1287	0.91	<0.001
88111	NOAA-9/11*	0.858	0.051	1202	0.83	<0.001	0.713	0.012	1257	0.85	<0.001
91051	NOAA-11	0.898	0.051	1320	0.91	<0.001	0.818	0.019	1296	0.89	<0.001

* NOAA-9 for PAL, NOAA-11 for ARTEMIS.

squares using 25 ground control points scattered over the whole of the African continent. Coastlines and inland features such as water bodies and pans were chosen. The RMS error was 0.79 pixels, which was acceptable considering the coarse resolution of the data. The comparison was done every 20 pixels. Visual inspection of the ARTEMIS images indicated that NDVI values less than −0.04 were water bodies or unmasked clouds. Therefore, pixels with values less than this were not considered because water bodies were masked out completely in the PAL data. Also, differences in the cloud screening procedures meant that they could be properly identified in one data set but not in the other, again leading to inconsistent comparisons. NDVI values higher than 0.68 in the PAL archives were also ignored because this was the maximum value allowed by the digital scaling in the ARTEMIS data set. The number of data points affected by both thresholds varied on each image but never exceeded more than 0.3% of the total.

Both data sets derive from the same raw data therefore it is not reasonable to attribute errors wholly to one of the archives. Instead, we assumed them to be the same order of magnitude in each archive and used Reduced Major Axis (RMA) regression described by Till (1973) to compare NDVI values. This also has the advantage that a single equation can be used to describe the transform relation (i.e. the regression of Y on X gives the same line as the regression of X on Y). The coefficients of determination (r^2) and the RMA equation coefficients for three image pairs are shown in table 2. There was generally a very high correlation between the two data sets and similar relations. The lower correlation for the first dekad of November 1988 (88111, table 2) could be because ARTEMIS used an image from NOAA-11 whereas PAL used one from NOAA-9. Also, visual inspection of the NOAA-9 image showed it was of bad quality with a lot of cloud cover and atmospheric interference. All the equations were significantly different from the one to one relationship implying that the two data sets are not directly inter-changeable.

3. Temporal analysis of spectrally uniform sites

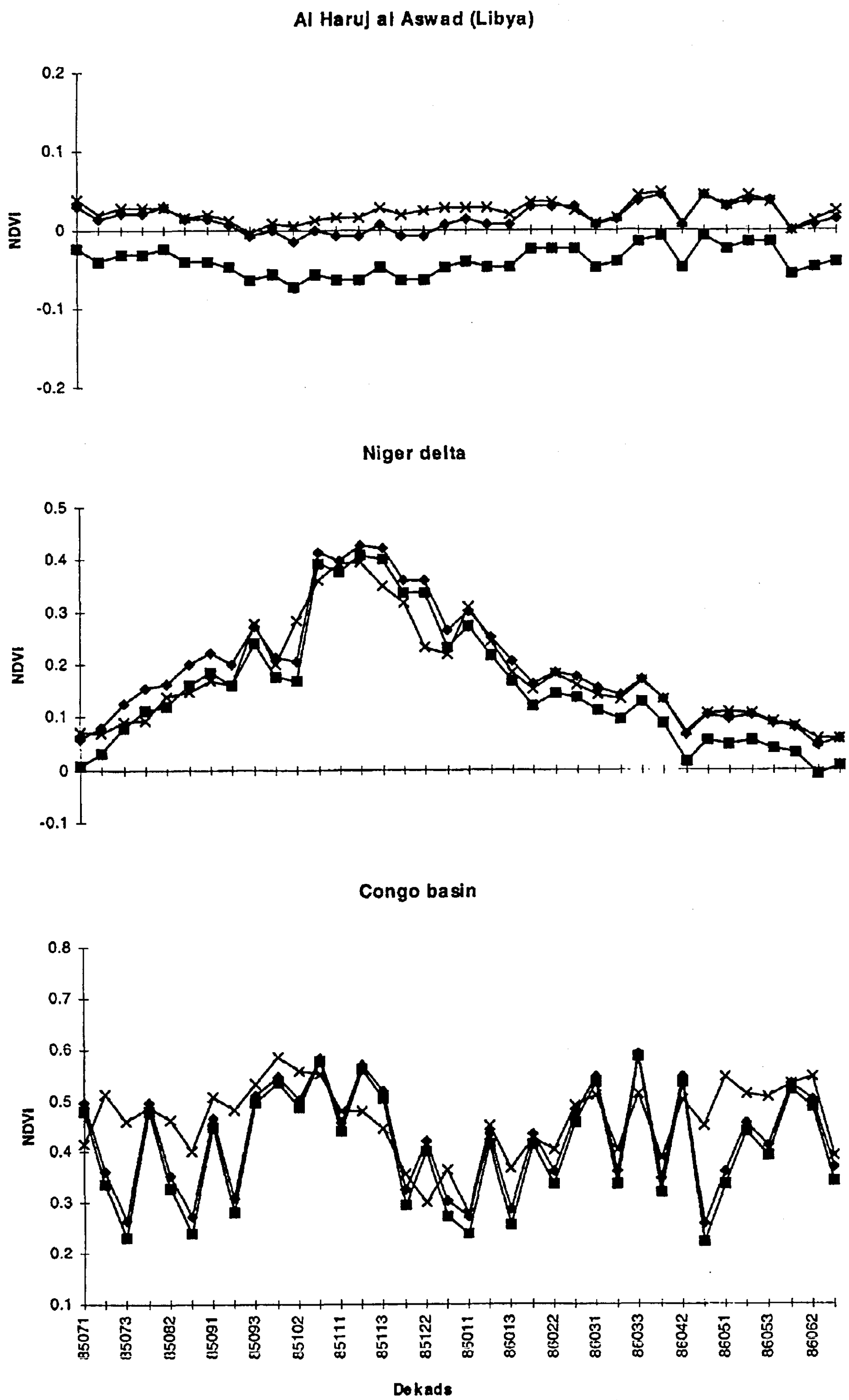
At such a coarse scale, pixel to pixel comparisons may be inaccurate because of re-sampling errors. There is also a need to assess the variation through time as well as spatially. Seven sites, identified in table 3, were chosen across Africa, based on their homogeneity, their visual definition in the imagery and their representation of the range of NDVI occurring in Africa. The median values of the 3 by 3 array of pixels surrounding the centre of the site were compared. This compensates for small errors in geometric registration of the two archives and minimises the influence of noisy NDVI values such as undetected cloudy pixels that remained after removing cloudy and missing data from the time-series by applying a simple linear interpolation

Table 3. List and location of sites for the comparison of ARTEMIS and Pathfinder archives.

Geographical name	Position
Nile Delta	30°55′N/30°58′E
Al Haruj al Aswad (Libya)	27°23′N/17°19′E
Rub al Khali (Saudi Arabia)	19°11′N/53°01′E
Niger Delta	14°25′N/4°32′W
Congo Basin	0°46′N/23°16′E
Okavango Delta	19°10′S/22°29′E
Etosha Pan	18°53′S/16°14′E

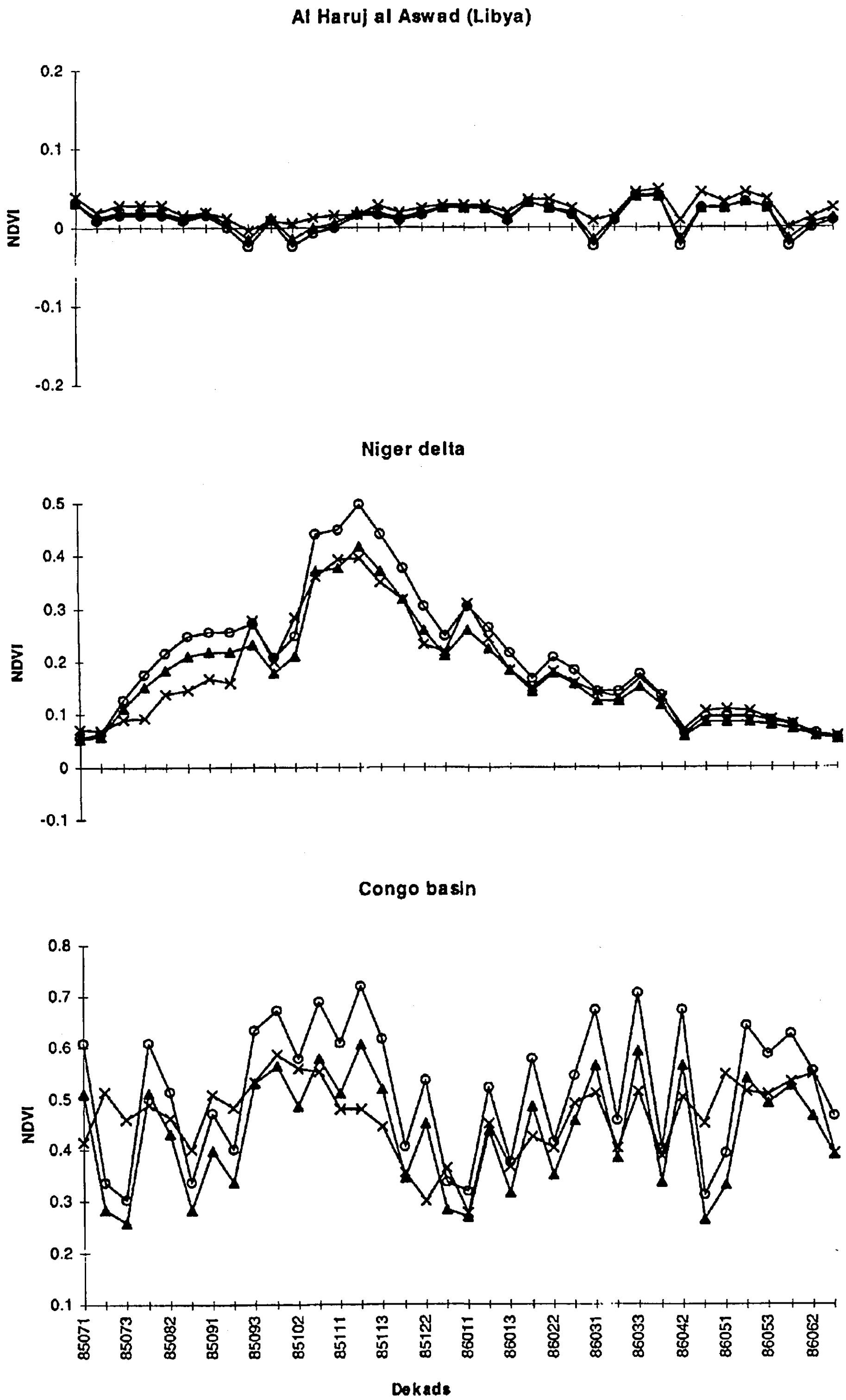
Table 4. Reduced Major Axis (RMA) linear fit results between the ARTEMIS and PAL archives for seven sites over the entire time series and each satellite, *ARTEMIS* = *f* (*PAL*).

Data	Original PAL archive					Post-processed PAL archive				
	Slope	Intercept	n	R^2	p	Slope	Intercept	n	R^2	p
All	0.925	0.051	2497	0.94	<0.001	0.832	0.005	2437	0.92	<0.001
NOAA-7	0.910	0.046	867	0.93	<0.001	0.825	−0.007	844	0.90	<0.001
NOAA-9	0.923	0.060	938	0.94	<0.001	0.818	0.016	920	0.93	<0.001
NOAA-11	0.946	0.044	665	0.95	<0.001	0.857	0.005	646	0.95	<0.001



(*a*)

Figure 1. NDVI time responses for three sites and a selected season for (*a*) ARTEMIS (– x –) and original PAL archives, before correction (–■–) and after correction (–◆–) and for (*b*) ARTEMIS (– x –) and post-processed PAL archive, before correction (–O–) and after correction –▲–).


Al Haruj al Aswad (Libya)
0.2
0.1
NDVI
0
-0.1
-0.2
Niger delta
0.5
0.4
0.3
NDVI
0.2
0.1
0
-0.1
Congo basin
0.8
0.7
0.6
NDVI
0.5
0.4
0.3
0.2
0.1
85071
85073
85082
85091
85093
85102
85111
85113
85122
86011
86013
86022
86031
86033
86042
86051
86053
86062
Dekads

(*b*)

between the previous and following dekad. The time-series of median values for each site in figure 1 shows differences between the two data sets. In the Libyan and Saudi sites where the NDVI remains relatively constant, the original PAL series was always

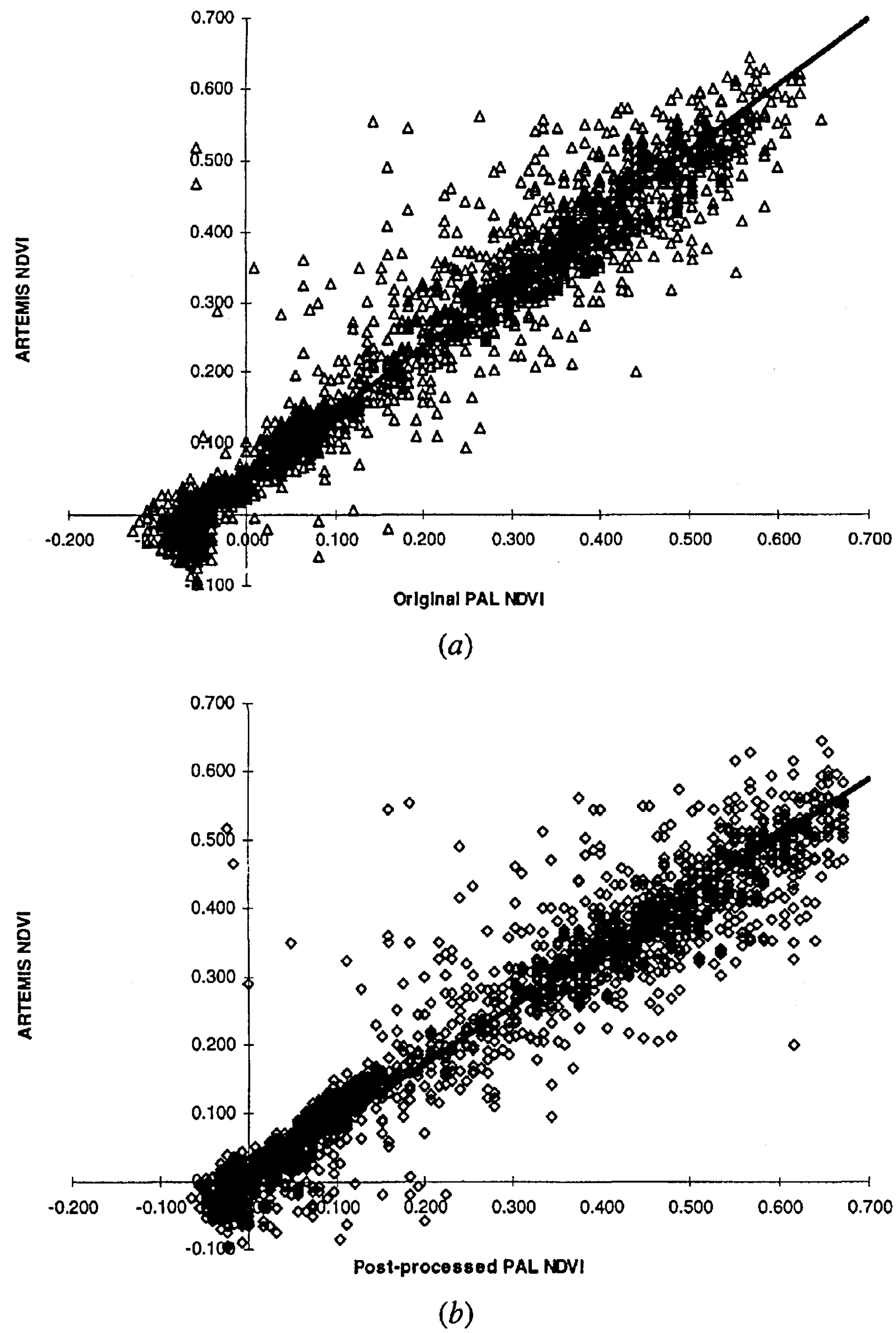


Figure 2. Reduced Major Axis (RMA) linear fit for seven sites over the entire time series between (*a*) the ARTEMIS and the original PAL archives, $n=2497$, $r^2=0.94$, ARTEMIS=0.925. (PAL)+0.051 and (*b*) ARTEMIS and the post-processed PAL archives, $n=2437$, $r^2=0.92$, ARTEMIS=0.832 (PAL)+0.005.

below that of the ARTEMIS. The average NDVI offset being 0.06 in Libya and 0.05 in the Saudi desert. The other curves show that the difference is greater at low NDVI values and smaller for high NDVI values.

The original PAL NDVI values are consistently lower than ARTEMIS. The main difference between the two archives is the correction of atmospheric effects in PAL and this should lead to higher NDVIs than in ARTEMIS. This seems to verify that there was a problem in the PAL processing chain. However, this appears to have been corrected for the post-processed PAL where the two data sets are fairly close to each other at low NDVI values but post-processed PAL has much higher values than ARTEMIS at high NDVI levels which is consistent with the expected effect of atmospheric correction.

The data for the seven sites were pooled together as the range of NDVI values at individual sites could be very small (e.g. the Libyan, Saudi, Congo and Etosha Pan sites). The analysis was also split according to the three different NOAA satellites. Results are shown in table 4. The slope of the relation for NOAA-11 is significantly different from the others ($Z = 2.14$ for original PAL between NOAA-7 and 11 and $Z = 2.68$ for post-processed PAL between NOAA-9 and 11). The intercepts were not significantly different (Z always less than 1). Thus the relation between the two archives remains relatively constant over time for NOAA-7 and NOAA-9 but a separate relation could be used for NOAA-11. However, further inspection reveals minimal differences in the values of NDVI compared to using the single pooled relations shown in figure 2 which represent the whole data set. The benefit of applying the relation can be seen in a representative section of the time series in figure 1 where the systematic differences between ARTEMIS and PAL profiles are much reduced.

4. Conclusions

The NDVI values from the PAL and ARTEMIS data sets are significantly and systematically different. Relations were derived and used to adjust one data set to the other thus allowing time-series analyses developed for ARTEMIS data to be used with the PAL archive. The difference between the two archives is attributed to the more complex processing applied to PAL. Further work is needed to assess whether the extra level of processing applied to PAL would improve relations between satellite imagery and biophysical variables such as crop yield or biomass quantity.

Acknowledgments

The authors wish to thank the Distributed Active Archive Center at the Goddard Space Flight Center for producing the data in their present form and distributing them. The original data products were produced under the NOAA/NASA Pathfinder program, by a processing team headed by Ms Mary James of the Goddard Global Change Data Center; and the science algorithms were established by the AVHRR Land Science Working Group, chaired by Dr John Townshend of the University of Maryland. Goddard's contributions to these activities were sponsored by NASA's Mission to Planet Earth program.

References

Groten, S. M. E., 1993, NDVI-crop monitoring and early yield assessment of Burkina Faso. *International Journal of Remote Sensing*, **14**, 1495–1515.

HIELKEMA, J. U., and SNIJDERS, F. L., 1994, Operational use of environmental satellite remote sensing and satellite communications technology for global food security and locust control by FAO: The ARTEMIS and DIANA systems. *Acta Astronautica*, **32**, 603–616.

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.

KOGAN, F. N., 1990, Remote sensing of weather impacts on vegetation in non-homogeneous areas. *International Journal of Remote Sensing*, **11**, 1405–141.

LAMBIN, E. F., CASHMAN, P., MOODY, A., PARKHURST, B. H., PAX, M. H., and SCHAAF, C. B., 1993, Agricultural production monitoring in the Sahel using remote sensing: present possibilities and research needs. *Journal of Environmental Management*, **38**, 301–322.

LOS, S. O., 1993, Calibration adjustment of the NOAA AVHRR normalised difference vegetation index without recourse to channel 1 and 2 data. *International Journal of Remote Sensing*, **14**, 1907–1917.

RAO, C. R. N., 1993, 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, NOAA/NESDIS, Washington, DC.

SANNIER, C. A. D., TAYLOR, J. C., DU PLESSIS, W., and CAMPBELL, K., 1998, Real-time vegetation monitoring with NOAA-AVHRR in Southern Africa for wildlife management and food security assessment. *International Journal of Remote Sensing*, **19**, 621–639.

SMITH, P., 1998, Coding errors in the Pathfinder AVHRR Land (PAL) data set's atmospheric correction algorithm. Goddard DAAC home page. URL: http://daac.gsfc.nasa.gov/CAMPAIGN – DOCS/LAND – BIO/PAL – coding – errors.pdf.

STEINWAND, D. R., 1994, Mapping raster imagery to the Interrupted Goode Homolosine projection. *International Journal of Remote Sensing*, **15**, 3463–3471.

TILL, R., 1973, The use of linear regression in geomorphology. *Area*, **5**, 303–308.

CHAPTER 5

REAL-TIME RANGELAND MONITORING IN BOTSWANA WITH NOAA-AVHRR

C.A.D. Sannier[1], J.C. Taylor[1], G. Slade[2], T. Ntabeni[2], S. Ngakane[3], K. Campbell[4]
contact e-mail: c.sannier@cranfield.ac.uk

[1] Cranfield University, Silsoe College, Silsoe, Bedford MK45 4DT.
[2] Department of Meteorological Services, Gaborone, Botswana.
[3] Ministry of Agriculture, Botswana Range Inventory and Monitoring Project, Gaborone, Botswana
[4] NRI, Chatham Maritime, Chatham, Kent ME4 4TB.

ABSTRACT

Rangeland is a major resource in Botswana where cattle ranching is one of the main sources of national income. The climate of Botswana is predominantly arid to semi-arid which means that the impact of highly variable rainfall on vegetation growth is a primary factor. The use of satellite derived vegetation index maps is a very efficient way to characterise these variations but they are difficult to interpret for the non-specialist user. A VPI (Vegetation Productivity Indicator) was used to assess the severity of current conditions compared to the historical record. VPI maps are now produced operationally at the Department of Meteorological Services and disseminated to users. Field checking showed very good correspondence between local field conditions and the VPI maps. These maps and a number of derived products are regularly presented to the inter-ministerial drought committee which decides on the applications for drought relief and other drought related issues.

SURVEILLANCE DES PÂTURAGES EN TEMPS RÉEL AU BOTSWANA AVEC NOAA-AVIIRR

Mots clefs: NDVI, Surveillance en temps réel, Botswana.

Résumé: Les pâturages sont une ressource importante pour un pays comme le Botswana où l'élevage du bétail est une des sources principales de revenu pour l'économie. Le climat du Botswana est à prédominance aride à semi-aride, ce qui signifie que l'impact des chutes de pluie sur la croissance de la végétation est un facteur primordial. De plus, la distribution des chutes de pluie est très variable aussi bien dans le temps que dans l'espace. L'utilisation de cartes d'indice de végétation provenant de satellite est un moyen très efficace de caractériser ces variations, mais elles sont difficiles à interpréter pour l'utilisateur non spécialiste. Un des problèmes principaux dans l'interprétation des images NDVI (Indice de Différence de Végétation Normalisé) est l'effet des types de végétation. Une stratification du pays a été trouvée, basée sur des données NDVI historiques définissant des zones de réponse NDVI homogène. Un VPI (Indicateur de Productivité de Végétation) a été développé : il compare les valeurs NDVI actuelles à la distribution statistique des NDVI pour la couche correspondante, tirés des archives NDVI historiques. Cette étude décrit l'application de la méthodologie VPI au contexte du Botswana. C'est le fruit d'un effort de collaboration de l'Université Cranfield de G.B., responsable du développement de la méthodologie VPI, du NRI (Institut des Ressources Naturelles, G.B.), qui a coordonné le projet, du Département des Services Météorologiques du Botswana, qui avait la charge de recueillir et traiter les données satellite, et du Projet d'Inventaire des Pâturages et de Surveillance du Botswana (BRIMP) du Ministère de l'Agriculture du Botswana, pour la dissémination et le contrôle sur le terrain des produits. Des cartes VPI sont maintenant produites de façon opérationnelle au Département des Services Météorologiques. Le contrôle sur le terrain par l'équipe BRIMP a montré une très bonne correspondance entre les conditions locales sur le terrain et les cartes VPI. Ces cartes et un certain nombre de produits dérivés sont régulièrement présentés au comité interministériel de la sécheresse, qui décide des applications pour des aides de sécheresse et autres problèmes liés à la sécheresse.

1. INTRODUCTION

Botswana is a land-locked country located in southern Africa with a total area of 576,000 km^2. The population is just over 1 million but is increasing rapidly (Moyo et al. 1993). Most of the country's climate is semi-arid with a mean annual rainfall varying between 250 mm in the south up to 650 mm in the north. The rainfall is highly variable both spatially and temporally. Only 5% of the country's land is fit for agriculture (Powell and Pulles 1996), although population pressure is forcing this to be increased into less suitable areas. The rest of the land is mostly occupied by rangeland. Indeed, the country's largest source of income after diamond extraction is beef production.

In the past two decades the country was hit by a number of droughts and according to Ringrose and Matheson (1987) the combination of low rainfall years and an increase in the animal population have led to overgrazing and ultimately land degradation in the most affected areas. However, it is difficult to say whether the situation has become irreversible or whether it will improve as soon as the rains return to normal. This situation has led government and donor agencies to seek better ways to monitor the current rangeland vegetation conditions in order to make better management decisions.

Previous studies (Hutchinson 1991, Lambin et al. 1993), have shown the advantage of satellite remote sensing, particularly NOAA-AVHRR (National Oceanic and Atmospheric Administration - Advanced Very High Resolution Radiometer), for the monitoring of vegetation conditions compared to methods using interpretation of rainfall measurements. The available network of rain gauges, especially in the case of Botswana, is not sufficient to allow a reliable interpolation of spatial variation in annual rainfall over the whole country. The most widely used satellite derived indicator of vegetation activity is the NDVI (Normalised Difference Vegetation Index).

There is a need to compare the current NDVI with historical data in order to assess whether vegetation conditions are better or poorer than usual and in the latter case to know if conditions are sufficiently extreme to adversely affect livestock and crops. Major aid organisations (USAID-FEWS, FAO) have set up operational early warning systems which compare current NDVI images with the previous dekad (10-day period) or with the mean image for the dekad (Hutchinson 1991, Lambin et al. 1993). The latter method assumes the annual variation of the NDVI for a location and a given dekad to follow a gaussian distribution. This assumption is unreasonable because the lower limit of the NDVI is bounded by the response for bare soil. An improved method, developed and tested by Sannier et al. (1998a) in Namibia and Zambia estimates the statistical distribution from the NDVI time-series by applying techniques commonly used in hydrology for the prediction of extreme events and defines a Vegetation Productivity Indicator (VPI). This paper describes how this methodology was developed and adapted for the Botswana Early Warning System (EWS) at the Department of Meteorological Services (DMS) and the Ministry of Agriculture of Botswana through the Botswana Range Inventory and Monitoring Project (BRIMP).

2. IMPLEMENTATION OF REAL-TIME VEGETATION MONITORING IN BOTSWANA

Real-time vegetation monitoring with satellite imagery is only possible if reliable local reception of satellite data is available. Over the last decade, the LARST (Local Application of Remote Sensing Technology) consortium led by NRI (Natural Resources Institute) has developed low cost satellite receivers capable of acquiring NOAA-HRPT (High Resolution Picture Transmission). The system comprises of an antenna and a receiver connected to a PC with the appropriate capture card and software. One such system was installed at the Department of Meteorological Services in Gabarone where NOAA-AVHRR data are acquired on a daily basis and used for production of NDVI images.

The ARTEMIS NDVI archive used previously covers the 1981-91 period. There was a need to extend this period to improve the implementation of the VPI method. The NASA-GSFC (National Aeronautics and space Administration - Goddard Space Flight Centre) Pathfinder AVHRR Land (PAL) data set was available up to September 1994. Additional data were also available on the ADDS (Africa Data Dissemination Services) Internet site of the USGS (United States Geological Survey) which were processed the same way as the ARTEMIS for the period starting August 1995 up to present (the data are being continuously updated). The combined data set covers a 15-year period.

The ARTEMIS and PAL archives were compared by Sannier et al. (1998b) in the overlap period and showed that although there was a strong relationship between ARTEMIS and PAL NDVI values, they were significantly different. However, the PAL archive NDVI values for the 1991-94 period were matched to the ARTEMIS values using a regression relationship. Comparison of the ADDS data with ARTEMIS values over stable targets such as deserts showed that the values were directly compatible with ARTEMIS. A 15-year time series of dekadal MVC NDVI images for Botswana and its surroundings was extracted from the various data sources identified in Table 1.

Table 1. Data sources for the 15-year Botswana data set

Time Period	Data Source
1st dekad August 1981 to 3rd dekad June 1991	FAO-ARTEMIS
1st dekad July 1991 to third dekad July 1994	NASA/GSFC PAL
1st dekad July 1995 to third dekad June 1997	USGS/FEWS/ADDS

The ISODATA algorithm (ERDAS 1995) was used to perform several unsupervised maximum likelihood classifications on the thirty-six 15-year mean dekadal images, varying clustering parameters and the number of classes. Cloudy pixels were eliminated in the calculation of the 15-year averages. Finally, a 16-class classification was selected to stratify the study area for the VPI method. The NDVI statistical distribution for each dekad and each stratum was determined using the method described by Sannier et al (1998a) . This was to determine the NDVI values corresponding to quintile probability thresholds used to define five vegetation status classes (very low, low, average, high and very high) in each stratum. Figure 2 shows these thresholds plotted against time for two locations.

3. DEVELOPMENT OF OUTPUTS

The methodology was commissioned in the Ministry of Agriculture and DMS in February 1998 and a workshop was held in Gabarone to discuss the format and dissemination pathways for outputs including maps, graphs and tabulated statistics. A map for the monthly report to the inter-ministerial drought committee during the rainy season was identified. This consisted of the maximum VPI value obtained from 3 dekads, was simple to implement and was thought to further remove cloudy pixels. An example of a monthly VPI map is shown in Figure 1.

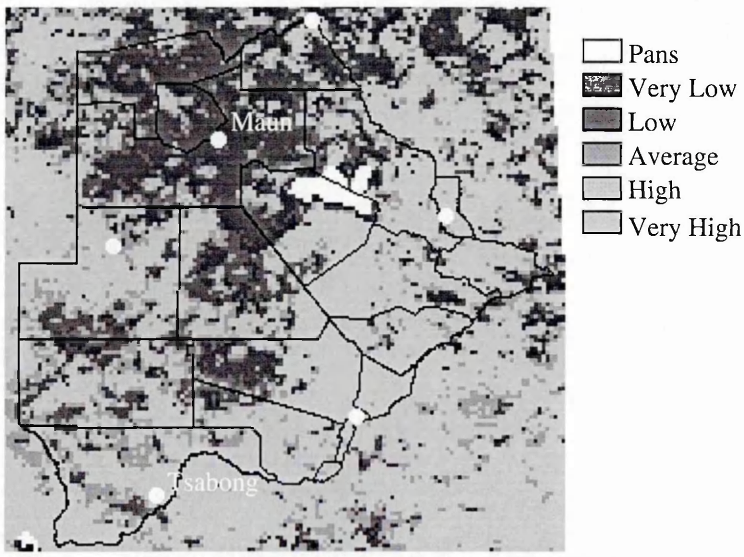


Figure 1. VPI map of Botswana for December 1997

The VPI maps show the spatial distribution of rangeland vegetation conditions for the whole of Botswana at a particular moment in time. It was also considered important to monitor vegetation condition at a single location over time by plotting the current NDVI on the vegetation status profiles such as the one shown in Figure 2. This supplements the information from the maps and assists the production of seasonal forecasts. Macros were developed to enable timely production of these plots for six main locations in the country (Gaborone, Kasane, Francistown, Maun, Guanzi and Tsabong). In Tsabong (Figure 2a) the vegetation response was both exceptionally early and exceptionally high whereas in Maun (Figure 2b) the growing season was two months later than usual and was still low at the time of the last image plotted. DMS and Ministry of Agriculture staff were shown these methods and were trained to produce similar graphs .

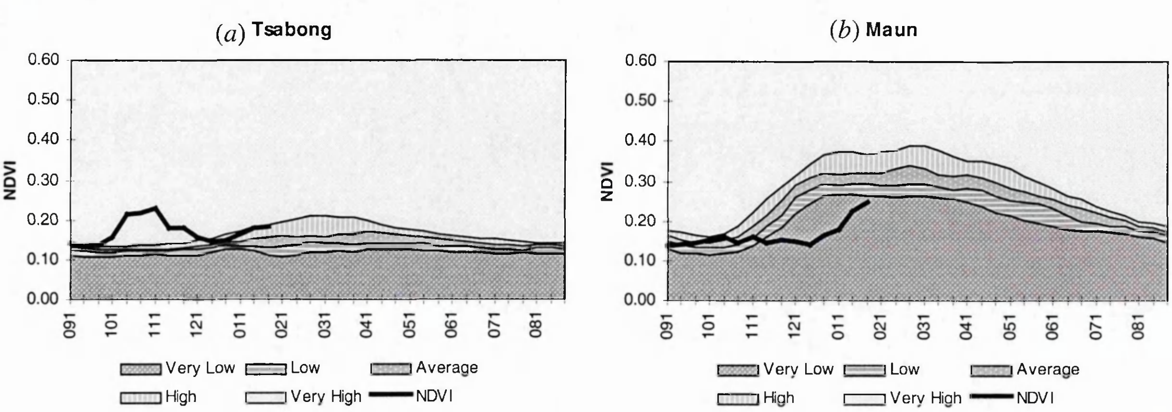


Figure 2. Evolution of vegetation conditions (*a*)in Tsabong and (*b*) in Maun, Botswana during the 1997-98 season

There was also a need for a simple output to indicate conditions at the level of agricultural districts. The number of pixels below average conditions for each district was imported into a pre-formatted spreadsheet and to produce a graph (Figure 3) of the percentage of each district suffering from potential drought conditions. This information contributes to the rapid identification of districts with problems.

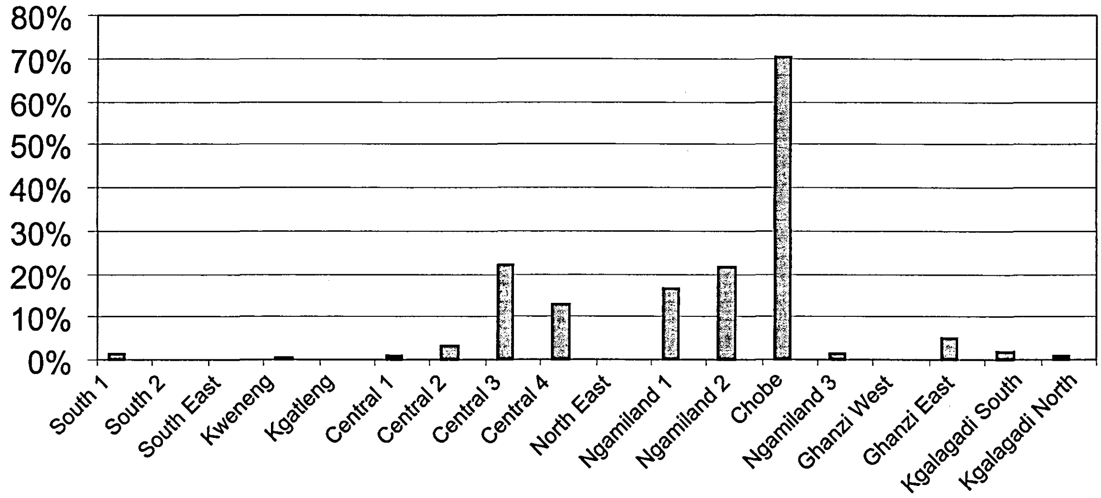


Figure 3. Proportion of agricultural districts of Botswana below average in September 1997.

4. DISSEMINATION AND FIELD CHECKING OF OUTPUTS

DMS and BRIMP have held workshops and seminars, given presentations and provided interpretations of the VPI maps for the members of the Botswana EWS to raise awareness of the new information described above and provide training in its use. This will be an on-going process which will be expanded to include users at a district level to increase the use and impact of the products. The VPI maps and associated products have also been distributed to a wide range of general users in Botswana through 10-day and monthly agro-meteorology bulletins produced by the Department of Meteorological services. The VPI maps are presented to the Botswana Early Warning Technical Committee during monthly meetings; are used to monitor potential drought conditions and as a visual summary of the current status of Botswana's vegetation. The technical data from the EWTC is then used by the Inter Ministerial Drought Committee to make specific recommendations and policy advice.

National Drought Assessment Tours (DAT) are conducted by multi-disciplinary teams at least twice a year to assess drought conditions on the ground. The VPI maps have been used by the DAT teams to target visits to worst hit areas and to assess the spatial extent of conditions that are identified in the field. The DAT teams visit each district and consult with the District Drought Committees. BRIMP and DMS have also supplied enlarged district level VPI maps to the District Drought Committees through the district range ecologists who use the VPI data for monitoring range conditions in the district.

5. CONCLUSIONS

The implementation of the VPI methodology at the DMS and its use by BRIMP seems to have worked extremely well and VPI maps are now being produced operationally and can be distributed to the relevant authorities. Initial field checking by BRIMP staff also showed that VPI products were effectively picking up variations in vegetation development compared to the norm. VPI maps are now being used by range managers as a monitoring tool.

The 1997/98 season has been declared a drought year by the President of Botswana. The use of the VPI maps for drought monitoring has made a significant contribution to the identification of the severity of drought in each area, the spatial extent of the area affected and the drought relief measures to be introduced.

ACKNOWLEDGEMENT

This study, funded by the UK Deparment for International Development (DfID) is a collaborative effort between Cranfield University from the UK, responsible for the development of the VPI methodology, the NRI (Natural Resources Institute, UK), which co-ordinated the project, the Department of Meteorological Services of Botswana, in charge of acquiring and processing the satellite data, and the Botswana Rangeland Inventory and Monitoring Project (BRIMP) from the Ministry of Agriculture of Botswana, for the dissemination and field checking of the products.

REFERENCES

ERDAS, 1995, *ERDAS Field Guide* (Atlanta : ERDAS Inc.).

HUTCHINSON, C.F., 1991, Uses Of Satellite Data For Famine Early Warning In Sub- Saharan Africa. *International Journal of Remote Sensing,* **12**, 1405-1421.

LAMBIN, E.F., CASHMAN, P., MOODY, A., PARKHURST, B.H., PAX, M.H., SCHAAF, C.B., 1993, Agricultural Production Monitoring In The Sahel Using Remote Sensing: Present Possibilities And Research Needs. *Journal of Environmental Management,* **38**, 301-322.

MOYO, S., O'KEEFE, P. and SILL, M., 1993, *The Southern African Continent, Profiles of the SADC Countries* (London: Earthscan Publications Ltd)

POWELL, M.J. and PULLES J.H.M., 1996, *Animal Production Simulation and Range Assessment Model for Botswana (APSRAMB), Volume I: Theory, User Manual, and Vegetation Inventory.* FAO/UNDP/Goverment of Botswana, Land Use Planning for Sustainable Agriculture Development, Project BOT/94/001, Field Document 17, 152pp.

RINGROSE S. and MATHESON W., 1987, Spectral Assessment of Indicators of Range degradation in the Botswana Hardveld Environment. *Remote Sensing of Environment,* **23**, 379-396.

SANNIER C.A.D., TAYLOR J.C., du PLESSIS W. and CAMPBELL K., 1998a, Real-Time Vegetation Monitoring with NOAA-AVHRR in Southern Africa for Wildlife Management and Food Security Assessment. *International Journal of Remote Sensing*, **19**, 621-639.

SANNIER C.A.D., TAYLOR J.C. and CAMPBELL K., 1998b, Compatibility of FAO-ARTEMIS and NASA Pathfinder AVHRR Land NDVI data archives for the African continent. *International Journal of Remote Sensing*, submitted.

CHAPTER 6

Real-Time Monitoring of Vegetation Biomass with NOAA-AVHRR in Etosha National Park, Namibia, for fire risk assessment

C.A.D. SANNIER and J.C. TAYLOR

School of Agriculture Food and Environment

Cranfield University, Silsoe, Bedford, MK45 4DT, UK

and W. DU PLESSIS

Etosha Ecological Institute

P.O. Okaukuejo, via Outjo, Namibia

Abstract. Estimates of biomass production are important in a wildlife reserve such as Etosha National Park, Namibia, for assessment of fire risk and subsequent selection of sites for controlled burning, for estimating forage supply and for improved understanding of nutrient cycles. We present methodology for using locally acquired NOAA-AVHRR images to make estimates of biomass in near-to-real-time. To this end, techniques for rapid measurement of the biomass of herbaceous and woody vegetation are developed using a rising disc pasture meter and individual plant dimensions. A field sampling methodology is presented to make biomass estimates compatible with the scale of AVHRR spatial resolution sufficiently close to the time of satellite overpasses to enable correlation with the NDVI from single images. Initial results show high correlations of biomass with NDVI for individual vegetation cover classes which appear to be temporally stable. There seem to be different regression equations for the different savanna vegetation types. These results were exploited to produce Biomass maps that could be used directly for the management of prescribed burning.

1. Introduction and operational context

Semi-arid environments such as those in Africa are characterised by highly variable rainfall patterns from localised storm systems which result in highly variable vegetation development, both spatially and temporally. In previous work, we presented methodology to assess vegetation conditions in near-to-real-time relative to historical variation (Sannier et al. 1998). The general intensification of land use in semi-arid Africa points to a need for better quantification and monitoring of vegetation resources. In this work we develop methodology for near-to-real-time biomass monitoring in the context of wildlife management, a major source of revenue for many African countries through tourism. The study area is Etosha National Park, (22,300 km^2), one of the biggest wildlife reserves in the world, situated in Northern Namibia (Figure 1). The most obvious feature of the Park is a saline desert called the Etosha Pan which covers about a fifth of the area. The Pan usually remains dry although during exceptional rainy seasons it may contain shallow water for short periods. The climate is semi-arid with a rainfall gradient increasing from around 250 mm in the West to 400 mm in the East (le Roux 1988). Vegetation consists of grassland, steppe and savanna with differing amounts of shrub and tree cover. Grass and steppe occur in relatively small plains mainly surrounding the pan and are very important grazing areas for game. The woody cover in the savanna broadly follows the rainfall gradient, with the highest density in the north and east, decreasing towards the west although there can be considerable local variation. Enclosure of game, even in such a large area as Etosha is known to have severely restricted the range of certain species of plains ungulates and animal populations have changed considerably since the Park boundary was fenced. The Park Management has a program of controlled burning to revitalise park vegetation and to limit fire hazard. In recent years, in spite of this, there have been uncontrolled fires doing serious damage to large areas in both eastern and western areas of the Park. As outlined by Holechek et al. (1995), biomass is an essential vegetation attribute in the monitoring of rangland conditions. Near-to-real-time monitoring of biomass is needed in wildlife management for estimating fuel load for fire risk assessment; for assessing the food supply and for understanding biological processes such as nutrient cycles.

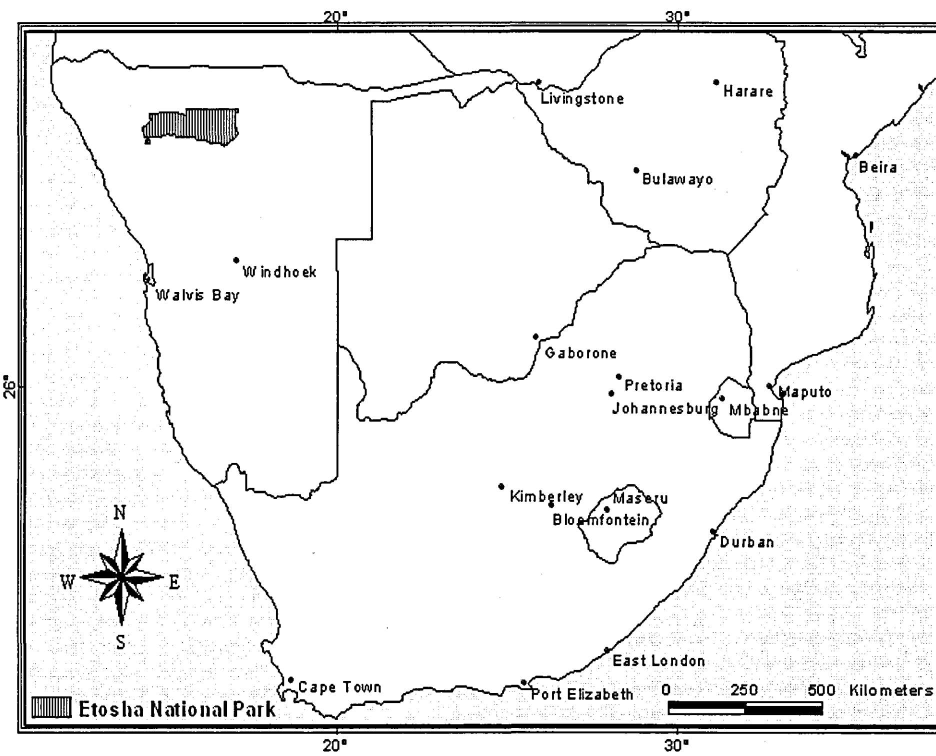


Figure 1. Location of Study area

The use of remote sensing techniques is particularly relevant as it provides the capability to cover large areas in near-to-real-time, which would be impossible to achieve with any other methods. Early work on biomass estimation with remote sensing concentrated on the investigation of a relationship between ground biomass measurements and vegetation indices derived from ground or airborne radiometer data (Pearson et al. 1976, Wagenaar and de Ridder 1986, Lamprey and de Leeuw 1986, Franklin et al. 1991, Hanan et al. 1993, McCloy et al. 1993). The Normalised Difference Vegetation Index (NDVI) described by Tucker (1979) was the most commonly used vegetation index although the PVI (Perpendicular Vegetation Index) first described by Richardson and Wiegand (1977), was also used by Taylor et al. (1986) in Niger. The results from these studies all identified strong relationships between green biomass and Vegetation indices and improved the understanding of the influence of various components in the vegetation. Some also attempted to relate airborne NDVI to satellite NDVI derived from the NOAA-AVHRR sensor. Higher resolution satellites were discarded because of the high cost of the imagery and their low temporal resolution.

Prince (1991b) and Franklin and Hiernaux (1991) used physical models of vegetation productivity. Although this approach is useful to understand vegetation reflectance characteristics, it has been rarely applied because its operational use is hampered by the necessity to create an extensive network of field stations to calibrate the model. This is unrealistic in semi-arid Africa where resources are limited. As a consequence, empirical approaches based on regression relationships are often preferred. Most studies (Prince and Tucker 1986, Prince 1991a, Wylie et al. 1991, Diallo et al. 1991) estimated the end-of-season biomass based on the integrated NDVI. However, Kennedy (1989) and Garcia-Daguerre (1993) obtained good relationships between green biomass and NDVI for single date AVHRR images in Tunisia and Mexico respectively.

Although encouraging results were obtained from early studies, they did not lead to operational applications mainly due to difficulties in image acquisition at that time and to the high cost of field calibration. In this work we demonstrate near-to-real-time biomass monitoring in Etosha National Park using an on-site NOAA-AVHRR receiving station to acquire imagery and rapid field methods of biomass estimation at selected sites to calibrate them by regression.

2. Rapid measurement of plant biomass

Rapidity is needed for economy and for timeliness of observations. Vegetation develops rapidly and is consumed rapidly in the Park so our aim was to compare image observations with ground estimates made within a few days of image acquisition time. To estimate the total green biomass per unit area, it was necessary to calculate contributions from herbaceous and woody components at the scale of NOAA-AVHRR pixels. This was done by a statistical estimator using sample observations.

2.1. Herbaceous Biomass

Several techniques for the estimation of herbaceous biomass have been developed. The absolute gravimetric or direct harvest method applied to small sub-plots is generally too time-consuming for operational work. Visual estimation is rapid and reasonably accurate when carried out by local experts (Pieper 1988). Wagenaar and de Ridder (1986) used a double sample of clipped and visually estimated samples in order to calibrate the visual estimation. The survey was carried out by two separate teams to check the consistency of the estimates. A combination of gravimetric and visual estimation using a regression relationship, was also applied by Garcia-Daguerre (1993). The method is fairly easy to implement but the calibration stage can be time consuming because each surveyor needs to be calibrated separately to maintain a certain level of objectivity.

Mitchell (1982) described the use of a rising plate meter for estimating grass biomass in Australian pastures. The rising plate or disc meter needs to be calibrated against clipped quadrats. The disc meter is dropped from a standard height onto the grass and there is a strong relationship between the height above ground of the disc meter and the amount of dry biomass. Scrivner et al. (1986) used a rising plate meter to estimate forage production and utilisation on improved annual pastures in California and also found a linear relationship. Trollope and Potgieter (1986) successfully used a Disc Pasture Meter (DPM) in the Kruger National Park in South Africa to estimate grass fuel loads. It has the advantage of being objective, rapid and easy to operate and was adopted for herbaceous biomass assessment in this work. However, judgement is required when making measurements in stony ground to avoid false readings.

Previous work in Etosha by Kannenberg (1995) produced the DPM calibration curve in Figure 2a using the same calibration procedure as Trollope and Potgieter (1986). The curve includes points from a wide range of locations and the single curve seems to be generally applicable for all Etosha grasses. The linear model for the regression produced a high coefficient of determination (r^2) but the scatter of points for biomass below 2000 kg/ha seems to deviate systematically below the regression line. So the regression model will over-estimate the biomass in this range. Secondly, the scatter of the data increases with the increase in biomass. Possibly because the ability of the grass to support the pasture meter weight is more variable in larger plants at the higher biomass levels. This means that the assumption of uniform variance is violated for the linear regression. The alternative logarithmic model in Figure 2b, proposed by us for this work, is an improvement and has a higher r^2. The scattergram in Figure 2a suggests that the pasture meter gives a positive reading in conditions of zero biomass. This is probably because of the roughness of the soil surfaces in the Park or the tendency for the meter measuring staff to penetrate the surface slightly when the meter is operated. Both models predict a positive biomass for zero DPM reading. In practice, we noted the presence of biomass visually when the DPM readings were zero and in the absence of alternatives adopted the model prediction for these base levels.

2.2. Woody plant biomass

Green biomass estimation of woody plants is usually by a regression relationship between dry matter weight obtained from direct harvest and some plant parameter, usually specific dimensions (Pieper 1988). Bille (1980) reviewed the relationship between stem diameter and leaf biomass derived in various parts of Africa. Cissé (1980) investigated the relationship between leaf biomass and various plant physical parameters for the Sahel. Tietema (1993) derived a relationship between total tree fresh weight and plant dimensions for 14 tree species in Botswana. The dominant woody species in Etosha, accounting for about 85% of the shrubs and trees in the savanna, is Mopane

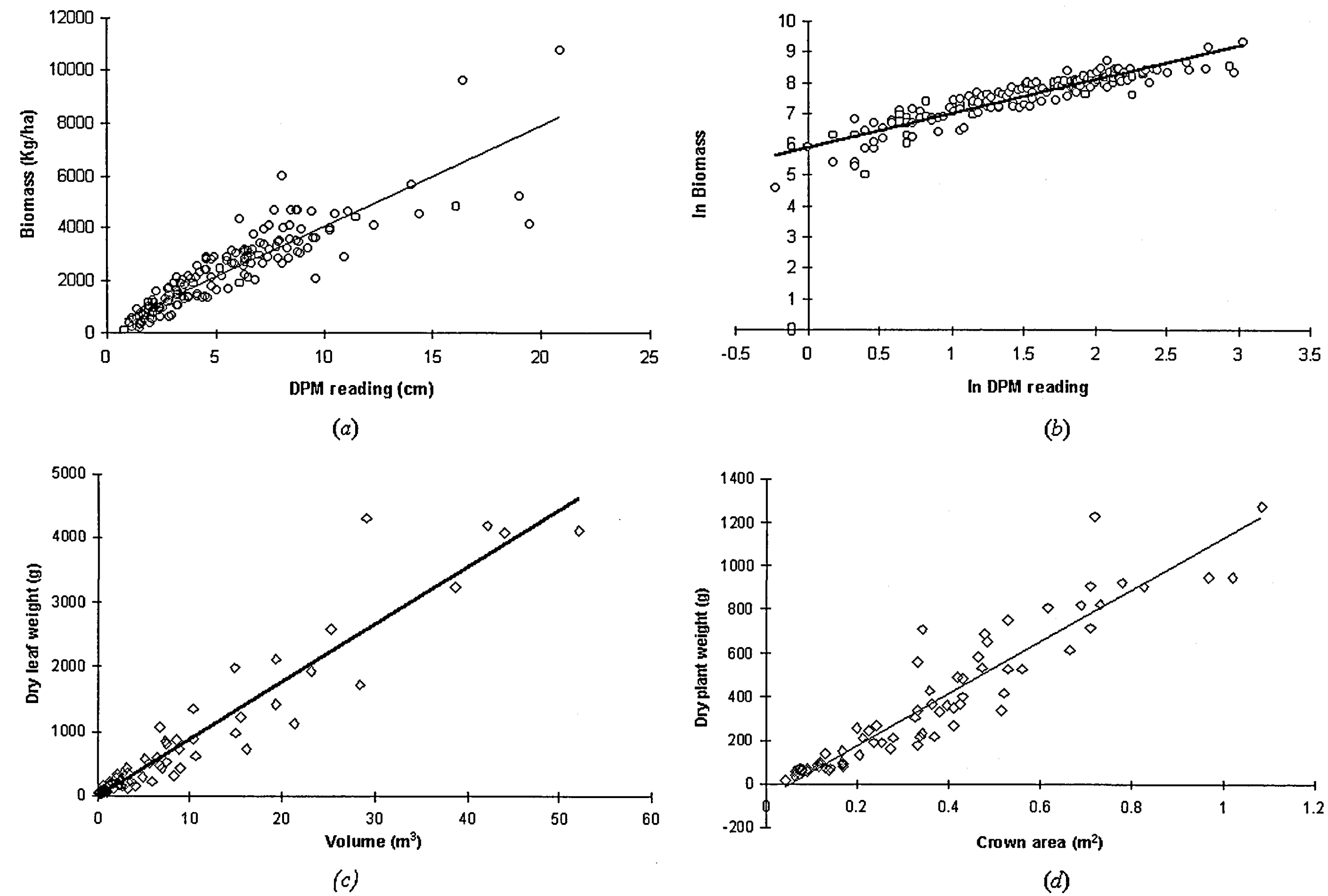


Figure 2. Calibration of (*a*) the Disc Pasture Meter for herbaceous biomass using a linear regression model, $n = 160$, $y = 380x + 248$, $r^2 = 0.78$ ($p < 0.01$), (*b*) the Disc Pasture Meter for herbaceous biomass using a natural logarithmic model, $n = 160$, $y = x^{1.09} \cdot 352.6$, $r^2 = 0.83$ ($p < 0.01$), (*c*) *Colosphospermum mopane* using the linear relationship between plant volume and dry leaf weight, $n = 494$, $y = 88.6x$, $r^2 = 0.90$ ($p < 0.01$) and, (*d*) *Leucosphaera bainesii* using dry plant weight and crown area, $n = 72$, $y = 1185x - 58$, $r^2 = 0.88$ ($p < 0.01$)

(*Colophospermum mopane*) which occurs in both forms. Data relating the branch diameter of Mopane to leaf biomass was available from previous work at the Etosha Ecological Institute (EEI, Du Plessis 1995). This was used to create a rapid field technique to estimate the biomass of Mopane trees and shrubs. Table 1 shows the average leaf biomass of Mopane for stems and branches in different size categories. The leaf biomass of a randomly selected sample of 80 Mopane trees and shrubs was estimated in the field by counting the number of primary stems in each of the size categories and using the average leaf weights from the table. The height of the plant and its crown diameter in two perpendicular directions was also recorded. At the time of the sampling, the Mopane was at or near to being in full leaf. The estimated dry leaf weight was best correlated to the volume of the plant calculated as a cylinder with diameter equal to the average crown size and height equal to the estimated tree height as shown in Figure 2c. All correlations we tried were less reliable for large trees. This is because the natural pattern of growth is more likely to have been altered in older trees by damage caused by animal utilisation, disease or other natural causes.

Table 1. Average leaf weight of Mopane stems in full leaf, in different size ranges.

Stem diameter class (cm)	Average leaf dry weight (g)	SEStandard Error (g)
0 to 0.5	1.6	0.1
0.5 to 1	6.9	0.4
1 to 2	28.3	1.8
2 to 3	84.8	8.2
3 to 4	171.2	11.3
4 to 5	239.7	26.4
5 to 6	387.2	81.3
6 to 7		
7 to 8	785.2	83.6
8 to 10	1240.7	209.6
10 to 12	1595.0	223.9
12 to 14	1714.8	190.3
14 to 16	2683.0	216.6
21 to 28	2883.2	774.0

The dominant shrub in Etosha steppe areas is Leucosphaera (*Leucosphera bainesii*), accounting for around 80% of steppe shrubs. Unlike Mopane, no previous work on the assessment of plant biomass had been done on Leucosphaera. Seventy-two plants were randomly selected in the field; height and perpendicular crown diameters measured then harvested. Then, the total dry plant weight was determined by standard oven drying. The total dry plant weight was considered more appropriate to use because the plants almost completely disappear during the dry season, therefore any plant material above the ground was considered new material. The best relationship between dry plant biomass and plant dimensions was found with crown area and is shown in Figure 2d. Unlike Mopane, plant volume was not better related to biomass because Leucosphaera is a smaller plant which develops itself horizontally rather than vertically.

3. Biomass estimation for calibration of NOAA-AVHRR observations

3.1. Sampling strategy

The criteria for selecting calibration sites were that they: be of sufficient size and internally homogeneous, to reduce the effects of errors in co-location of the ground observations with NDVI values; be accessible and, reflect the range of biomass levels in the Park. Since the aim was to develop regression relationships of NDVI versus biomass and not to make direct expansion estimates from the sites themselves, site selection was not random. Sites were chosen to reflect the variation of main grass, steppe and savanna types in the Park.

The formula derived by Justice and Townshend (1981) gives a guideline for the minimum size, a, of a sampling unit in relation to the geometric accuracy: $a = p(1+2l)$ where p is the pixel dimensions in distance units and l the geometric accuracy in number of pixels. For example, a 1.1km pixel size for NOAA-AVHRR and a geometric accuracy of 0.5 pixel should result in a sampling unit of 2.2km on each side. Generally, it is not feasible to have calibration sites large enough to meet this ideal and in past studies smaller sites have been used. We initially selected candidate homogeneous locations, several square km in size by photo-interpretation of geometrically corrected false colour composites from Landsat TM imagery. By selecting a 1 km^2 site in the middle of a larger homogeneous area, we expected to minimise the effects of geometric correction errors as variation of the biomass in the immediate surrounding area was unlikely to be great and also, surrounding pixels would not be mixed responses including other vegetation types. Field checking with geo-coded enlargements of the TM imagery verified the homogeneity of the selected sites.

Biomass estimates are based on sample observations within the sites. Various sampling schemes have been used which usually attempt to represent the range of vegetation types in the study area.

The biomass sampling was carried out along a 1 km transect through the centre of the 1 km square sample site. Because of our previous selection procedure, it is assumed that the site is isotropic and homogeneous so that average biomass along the transect represents the average for the whole of the 1km square area. The sampling scheme for herbaceous vegetation is shown in figure3. DPM measurements were made at ten equally spaced locations along the transect. Navigation was assisted by a Landsat TM enhanced geo-coded image hardcopy and a handheld GPS. Five clustered DPM readings were taken on each side of the transect. This resulted in a total of 100 DPM measurements per site which is the value suggested by Trollope and Potgieter (1986).

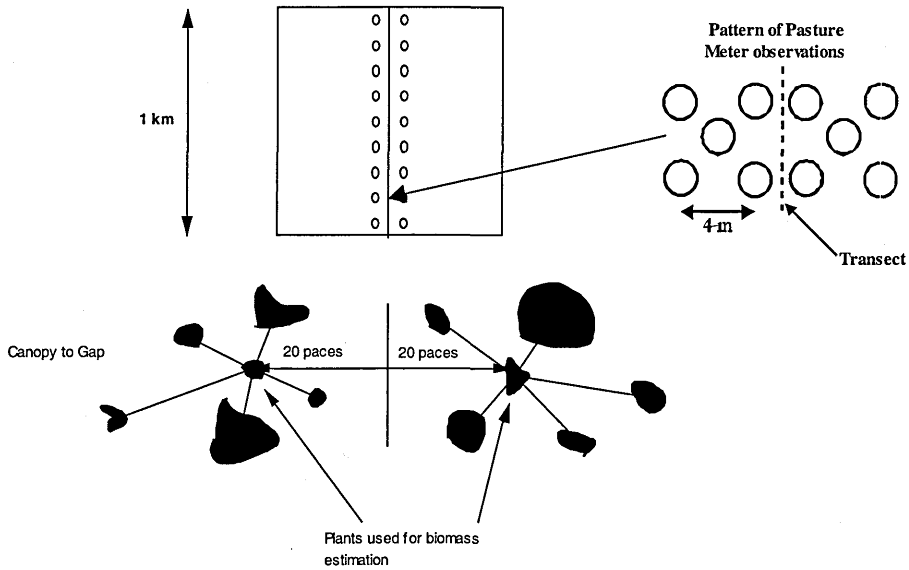


Figure 3. Sampling strategy for biomass measurements

The woody plant biomass per unit area is the product of the biomass per plant and the number of plants per unit area. The number of plants per unit area is estimated by dividing proportion of plant cover (canopy area/unit ground area) by the canopy area per plant. The crown to gap method (Westfall and Panagos 1984, Walker et al. 1988) is a very fast and unbiased method for estimating the percentage canopy cover of woody vegetation (trees, shrubs, dwarf shrubs etc.) over relatively large areas from ground observations or surrogates such as photographs and was selected for this study. Two adjacent plants approximately the same distance from the observer are selected. The ratio ϕ, the distance between their crowns (G) to the diameter of the crown of one of them (K) in Equation 1

$$\phi = \frac{G}{K} \tag{1}$$

can be: estimated by eye or by a transparency gauge (Westfall and Panagos 1984) or measured on photographs. The average ratio $\bar{\phi}$, for at least 25 pairs of plants is used in Equation 2, described by Walker et al. (1988) to estimate the percentage covered by the crowns, C:

$$C = \frac{100\pi}{2\sqrt{3}}\left[\frac{1}{\left(\bar{\phi}+1\right)^2}\right] \tag{2}$$

The sampling of woody vegetation was done on the same 10 locations along the transect as shown in figure 3. A plant was randomly selected on each side of the transect by walking 20 paces perpendicular to the transect using the vehicle as a reference. Then the closest plant was selected. In the case of Leucosphaera, plant dimensions were recorded before being harvested and in the case of Mopane crown dimensions were recorded together with stem dimensions. The reference plant was also used to perform the canopy to gap procedure as shown in figure 3 by choosing five other plants closest to it and evaluating visually the canopy to gap ratio between the reference plant and each of the five other plants measurement. Care was taken to select only the closest plants as this would otherwise have biased the results.

The whole procedure was repeated on the other side of the transect. In total, for each site, 100 DPM measurements, 100 canopy to gap ratio estimations and 20 plant dimensions were taken. The 20 plants selected were also used to develop the calibration procedure described in 2.2. This was reduced to 12 plants in one Leucosphaera site due to lack of time before sunset.

It was decided that, for the first year (1995), the survey would be limited to grassland. The three grassland sites shown in figure 4 were selected on the high resolution imagery. The grassland sites were surveyed four times during the 1995 growing season. During the 1996 season, steppe and savanna sites were included in the survey. Four steppe sites, four savanna sites and three grassland sites were surveyed between mid-March and early April.

3.2. Determination of site total green biomass

The estimation of herbaceous biomass per site is the most simple because DPM measurements are directly related to biomass per unit area. DPM readings were recorded in a spreadsheet. Each cluster of five DPM readings (figure 3) was averaged. This was to create observations comparable with the original calibration procedure. The calibration equation was applied to the averaged value resulting in 20 grass biomass value per site. The overall biomass of the site was calculated by taking the mean of the 20 biomass values. The same method was used in woody vegetation sites without taking into account of the woody cover because grass also often grew under the trees and shrubs and a reading of zero was recorded when the DPM fell on a shrub.

When assessing grass biomass in the field with the DPM, it was felt that great care needed to be taken when individual measurements were below 2.5 cm, which corresponds to about 1,000 kg/ha of biomass. Below 2 cm, small irregularities of the terrain were likely to have a greater influence on the measurement than the amount of biomass and a reading was visually estimated using the DPM calibration curve to derive the corresponding biomass. Between 2 and 2.5 cm, the DPM reading was recorded providing that the terrain was flat and without stones.

The estimation of woody biomass includes several parameters and the process of averaging was carefully considered because of non-linearity. Each location along and on each side of the transect was treated individually. This was to take into account of all the variations within the site and the non-linearity of the canopy to gap ratio method. A

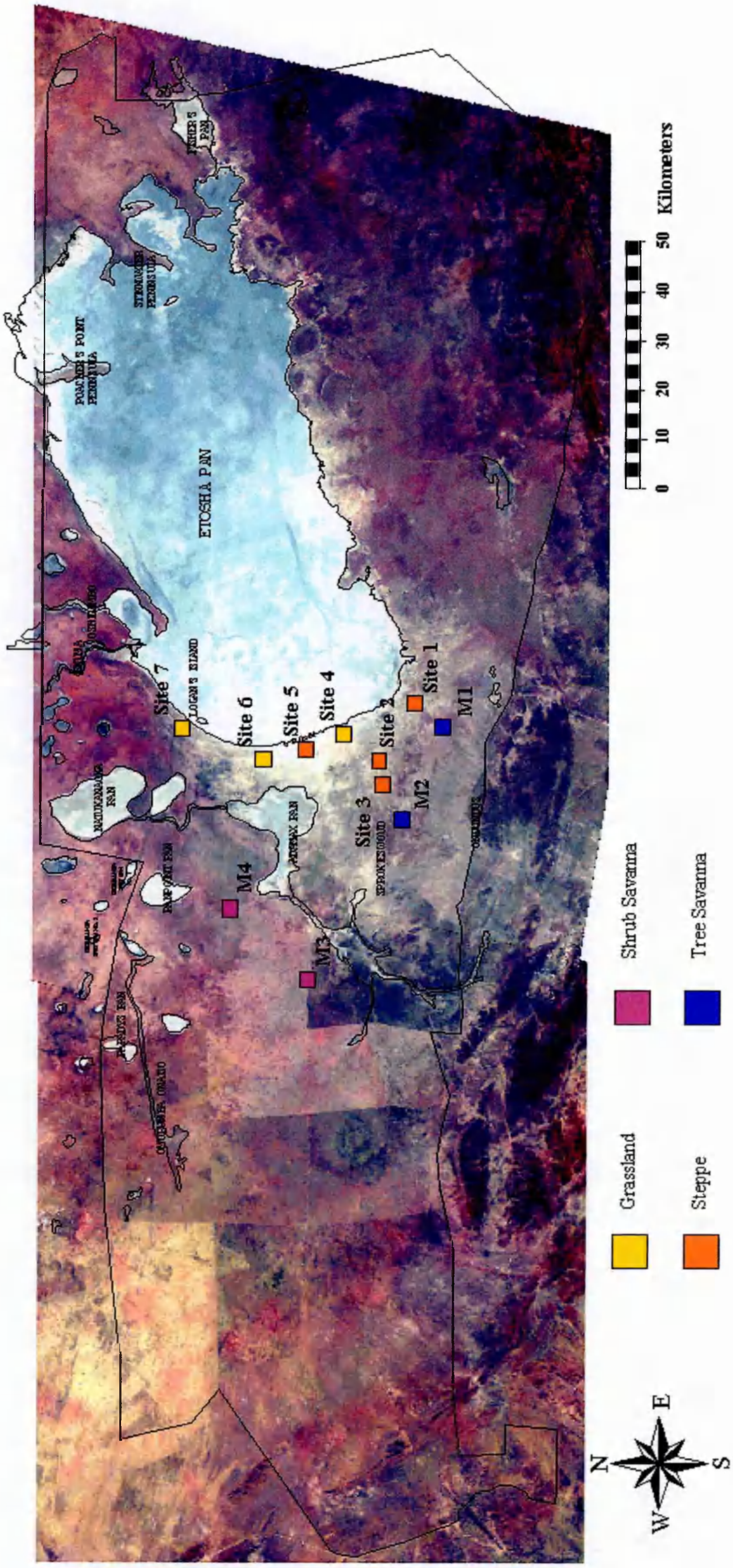


Figure 4. Landsat TM false colour composite mosaic (red: near infrared, green: red and, blue: green), February/March 1992, of Etosha National Park showing the location of biomass survey sites (vector overlay courtesy of EEI)

biomass value was calculated for each of the 20 sampled plants associated with 5 measurements of density (canopy to gap ratio). These five estimates of density were averaged to establish the density at each plant location. The density combined with the plant area gives the number of plants per unit area. The number of plants per unit area multiplied by the plant weight gives the biomass corresponding to the plant and density considered. 20 plants and five measurements of density per plant resulted in 100 observations of biomass per site. The estimate of biomass for the whole site was the arithmetic mean of these 100 biomass observations plus the grass biomass.

Weight biomass per plant, *P*, in kilogrammes is determined by the relationship between plant dimensions and biomass defined in 2.2. The density cover, *C*, is determined according to equation (2). The number of plants per hectare, *N*, is:

$$N = \frac{10000}{A} \cdot C \quad (3)$$

Finally, the total woody biomass per hectare and plant, W_i, is equal to:

$$W = P \cdot N \quad (4)$$

In the case of Leucosphaera, the calculation of biomass is simplified because the crown area is used in both the calculation of *P* and *N* which are cancelled out. Therefore, for Leucosphaera, woody biomass is only related to *C* and the slope of the calibration between plant biomass and area.

For Mopane, where the relationship with biomass is based on volume, woody biomass is a function of the slope of the calibration, height and *C*. It means that once the relationship between biomass and plant dimensions has been established, it is not required to measure plant area in order to derive estimates of woody biomass. The total woody biomass per site is equal to the average of the 20 estimates derived per plant. The combined grass and woody biomass is obtained by adding the two estimates.

4. Relationship between estimated biomass and NOAA-AVHRR observations

4.1. NOAA data processing

NOAA-AVHRR imagery was acquired in real-time from the LARST receiving station installed at the EEI since 1993 (Williams and Rosenberg, 1993). This allowed us to produce NDVI images in near-to-real-time, to select only the best images and to make sure that the images would coincide with the fieldwork. A total of seven images were selected (4 for the 1995 season and 3 for the 1996 season). Data processing consisted of geometric and radiometric correction of the imagery and NDVI calculation. No atmospheric correction was performed because it was thought that careful selection of imagery, free of clouds and nearest as possible to nadir would be more efficient in keeping atmospheric effects to a minimum rather than applying an approximate atmospheric correction. Existing methods for removing atmospheric effects often assume constant effects over the entire scene and require ground meteorological data that is not realistic to obtain for near-to-real-time application. However, the use of normalised radiance values is important because data from different satellites and at different time periods are being used.

Although, the receiving station software can also output a geo-referenced product, this was discarded because the geometric correction performed is based on the satellite orbital parameters and on the manual shifting of a warped map on the image. This procedure yields an accuracy of about 1-2 pixels in best cases. This is fine for the monitoring over a whole country but in our case, the area covered is much smaller (22,300km^2) and a better geometric correction will reduce the area needed for each monitoring site. Moreover, the output from the receiving station is re-projected in Latitude/Longitude, again, this is fine for large area but when ground observations are collected, a map grid is more appropriate than geographic co-ordinates. Therefore, an interactive geometric correction using ground control points was performed on channel 1 and 2 for each image. The high-resolution Landsat TM geo-referenced imagery was very useful for this purpose as a template allowing an accurate positioning of ground control points and fast image to image registration. A minimum of well-scattered 12 ground control points were selected and a second order transformation model was applied to take account of the deformation of AVHRR imagery. The total RMS error obtained was never more than 0.6 of the input pixel size and most of the time less than 0.5 pixel. All seven NOAA-AVHRR images used were re-sampled using cubic convolution with a 500m pixel size and geo-referenced to the UTM map projection. This was the best method to estimate the NDVI at our sampling locations. Nearest neighbour preserves NDVI values exactly but pixels can be offset by about 1 pixel size because of the geometric correction error and re-sampling approximation. A non-linear interpolation procedure, such as cubic convolution, is therefore, appropriate to represent the local variations of the NDVI around a sample point and to estimate a representative value.

DN values were extracted from the imagery for each channel and for each site. The data were input into and spreadsheet where radiometric corrections and NDVI calculations were carried out.

Radiometric corrections were based on the work published by Kaufman and Holben (1993) and Los (1993) for NOAA9 data and Rao and Chen (1996) for NOAA14, which took sensor degradations into account.

4.2. Relationship between green biomass and NDVI

A total of 15 observations were available for grassland for three sites taken at five time intervals. The data were analysed together because it was important to investigate whether the relationship was stable for near-to-real-time monitoring. Figure 5a shows the scatter of points for all the grassland sites. There is a very obvious outlier which corresponds to site 4 (figure 4). When examining the plot more closely, points from site 4 seem to cover a fairly wide range of NDVI but always have a similar biomass value. Firstly, although grassland is crucial for wildlife, it covers a small proportion of the Park and it is difficult to find sites which cover a sufficiently large area. When site 4 was first selected from visual interpretation of the TM imagery, it only covered a slightly smaller than 2 by 2 km area of what appeared to be grassland. On the first year (1995), an initial field visit showed that the vegetation was mostly composed of annual short grass species. This was confirmed during subsequent visits during the 1995 growing season although very low biomass levels were observed. However, in 1996, there was a much higher level of biomass due to higher rainfall and although there was some of the same short grass species present, most of the vegetation was dominated by broad leave herbaceous plants. They did not give a measurable DPM reading, which explains the low biomass values recorded, but were likely to produce a significant response in NDVI. Site 4 effectively belonged to a different vegetation type and a method other than the DPM, e.g. based on basal cover, would have been necessary to assess that vegetation type. However, this is a very unusual vegetation type in Etosha and is not representative of the Park's vegetation. Therefore, it was decided to discard site 4 from the analysis. Figure 5b shows the relationship without site 4, showing a much higher coefficient of determination.

Table 2. Biomass components of savanna sites

	Grass biomass		Woody biomass		Total
	(kg/ha)	(%)	(kg/ha)	(%)	(kg/ha)
M1	353	34	694	66	1047
M2	196	17	979	83	1175
M3	1016	74	359	26	1375
M4	923	68	429	32	1352

When incorporating all sites surveyed together, the relationship seems much weaker (figure 5c). However, one can also clearly identify that the points that are falling the furthest from the regression curve are the savanna sites. This would tend to show that there is a different calibration curve for savanna. The savanna points tend to fall below the calibration curve which means that the NDVI is much higher than that for the other sites in relation to the amount of biomass measured. It should be noted that Mopane leaves are very broad and bright green leaves, which are likely to reflect more near-infrared light than grass. This can explain why savanna sites exhibit a higher NDVI than grassland. Furthermore, the proportion of woody biomass in the total site biomass was calculated as shown in table 2, and it appeared that savanna site M3 (figure 4), which is the closest to the regression curve, has only 26% of its total biomass from woody vegetation. M1 and M2 (figure 4), which are the furthest from the regression curve have respectively 64 and 83% of woody biomass. M4, which is slightly closer to the regression curve, has 32% of woody biomass. This tends to corroborate the influence of Mopane leaves on the NDVI, although more data would need to be collected to confirm this hypothesis.

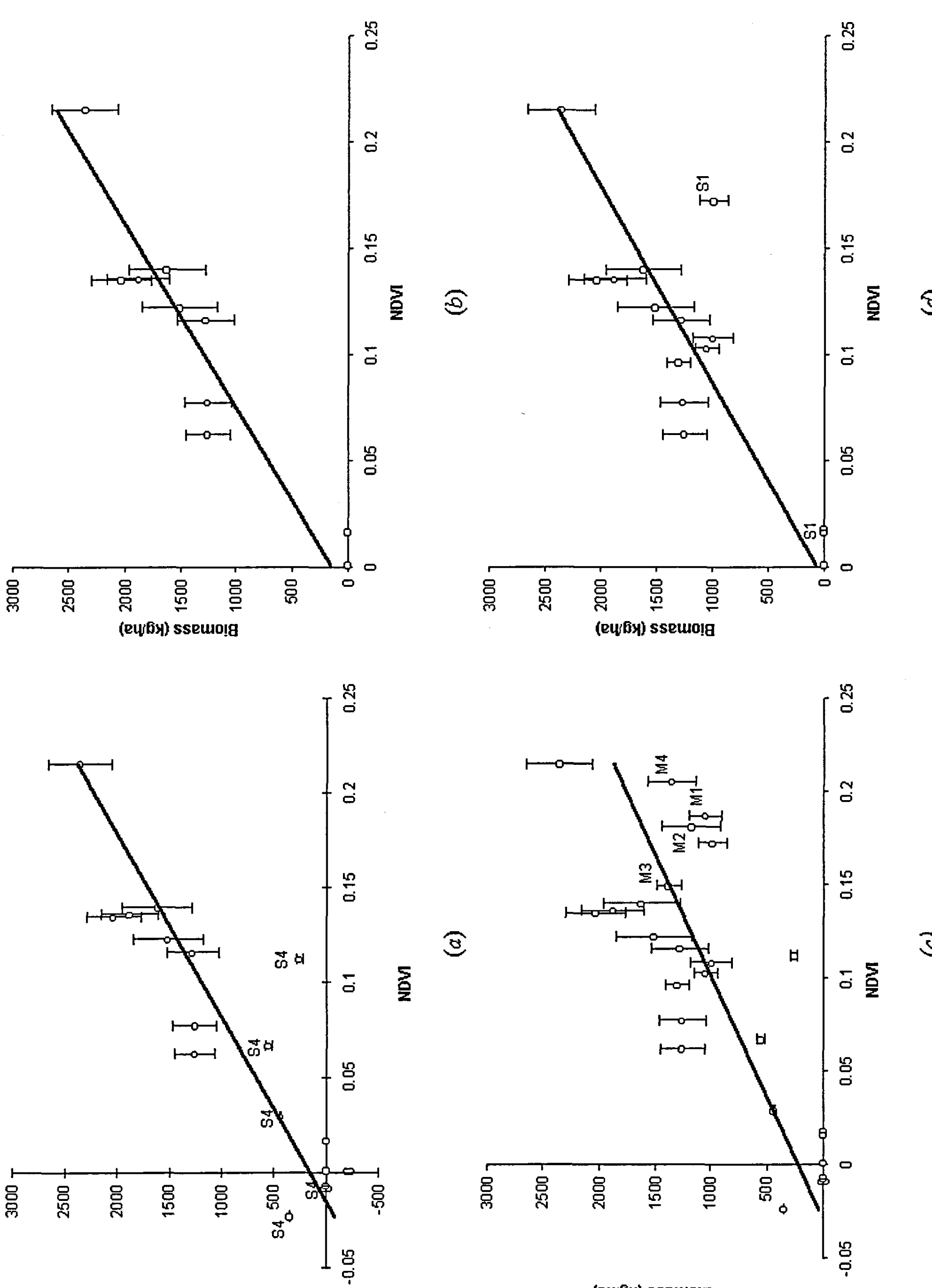


Figure 5. Relationship between NDVI and biomass for (*a*) all grassland sites, $n = 15$, $y = 10354x + 156$, $r^2 = 0.76$ ($p < 0.01$), (*b*) excluding site 4, $n = 10$, $y = 11551x + 138$, $r^2 = 0.89$ ($p < 0.01$), (*c*) all grassland, steppe and savanna sites, $n = 25$, $y = 7735x + 208$, $r^2 = 0.61$ ($p < 0.01$) and, (*d*) Grassland and Steppe sites, excluding site 4, $n = 16$, $y = 10806x + 57.7$, $r^2 = 0.77$ ($p < 0.01$); error bars represent the standard deviation of the field biomass estimate

Figure 5d shows the combined relationship for grass and steppe sites. The coefficient of determination obtained is much higher than when savanna sites were included. However, it is not as high as when only grassland sites are considered. This is mainly due to site 1 which falls below the regression curve. The expected biomass for that site should be much higher. A close look at the field data (see table 3) shows that the average plant diameter is generally much smaller in site 1 than in the other steppe sites, 0.52m for site 1 compared to at least 0.63m in other sites. It is possible that with such small plants the canopy to gap ratio method to estimate the density is not as reliable. However, it was not possible to test this at this stage. The other three steppe sites fit on the regression line. Furthermore, the regression curves for grassland alone and grassland/steppe combined are not statistically different.

Table 3. Average plant diameter for *Leucosphaera bainesii* sites

	Average plant diameter (m)
Site 1	0.54
Site 2	0.69
Site 3	0.68
Site 5	0.63

5. Production and applications of biomass maps

The pooled relationship developed in the previous section (figure 5c) can be used to transform NDVI images acquired at the NOAA HRPT receiving station into biomass maps. There are a number of ways in which these biomass maps can be used. This could include the monitoring of animal movement in relation to fodder availability during a growing season or the monitoring of the Park's animal carrying capacity from year to year and throughout the study area. Furthermore, the comparison of biomass maps with animal census could provide a better understanding of the factors influencing animal population dynamics.

However, a more direct application of biomass maps is the correlation between fuel loads and fire risks. Fires occur naturally in the Park and are normally triggered by lightning (Heady 1975). Under favourable conditions, wildfires can spread over large areas and can cause major damage to wildlife and vegetation. However, controlled or prescribed burning is often used to prevent the occurrence of wildfires by reducing fuel loads. Furthermore, controlled fires may benefit wildlife through positive effects on vegetation regeneration and habitat diversity (Heady 1975, Holechek et al. 1995). Controlled fires have been used in Etosha National Park for the above reasons and the Park was divided in a number of fire blocks as shown in figure 6. Trollope and Potgieter have shown that in the Kruger National Park, biomass fuel loads needed to reach at least 1500kg/ha to propagate. Therefore, it is possible to use biomass maps reclassified according to a series of thresholds indicating the levels of fire risk. This is illustrated in figure 6, where biomass maps were produced at the end of the rainy season for 1995, 1996 & 1997 using the pooled NDVI / Biomass regression relationship shown in figure 5c.

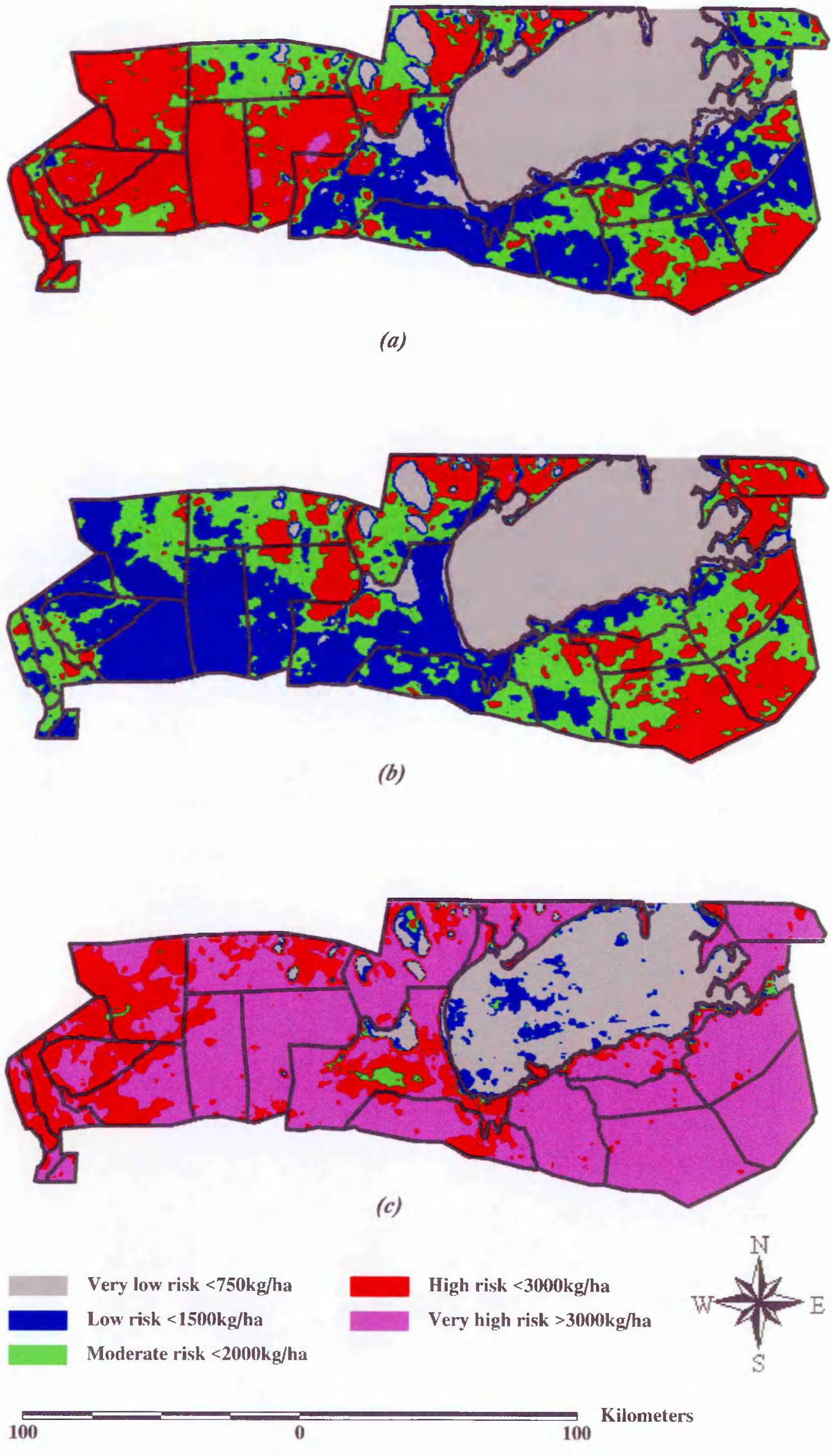


Figure 6. Biomass maps of Etosha National Park for (*a*) 27 March 1995, (*b*) 18 march 1996 and, (*c*) 22 March 1997

Looking especially at 1995 and 1996, it becomes apparent how such maps could be used to target, for controlled burning, fire blocks corresponding to the high to very high risk classes. 1997 was an exceptional year, with rainfall in excess of 40 to 60% of the median between 1981 and 1996, which explains the extremly high levels of biomass reached throughout the Park. As a result, all the conditions for an extensive fire were met and a wildfire started at the end of the dry season, which burnt nearly 21% of the Park's area outside the pans, crossing several fire breaks. The conditions of 1997 were very unusual, but it is possible that if some controlled burning had taken place in fire blocks where biomass was already relatively high in previous seasons, the wildfire that took place in 1997 would not have spread so extensively.

6. Conclusion and further development

Firstly, this study has demonstrated methods for near-to-real-time monitoring of biomass quantity with NOAA-AVHRR. The value and reliability of the DPM has been demonstrated and was shown to be suitable to measure biomass of grassland in the Park. However, its use should really be limited to the grass types for which it has been calibrated. It is possible, in certain circumstances, that several calibration curves might be required depending on the grass types present. The DPM is also not suitable to reliably measure biomass below 1000kg/ha and, other techniques, such as visual estimation, need to be used.

Concerning woody plant biomass, the set of techniques that were derived seem reasonably reliable. The calibration of green plant biomass based on dimensions worked particularly well. The canopy to gap method was also very rapid to implement and seem to be a fairly reliable way to assess canopy cover, although it is necessary to investigate further the precision of the method in relation to plant size. Once the calibration of plant biomass with dimensions has been determined, the measurement of woody biomass becomes extremely rapid especially for plant species for which the calibration is based on plant area. In this case, the only parameter required is the canopy cover. For plant species for which the calibration is based on volume, the parameters required are the canopy cover and plant height. This makes the assessment of biomass monitoring sites much faster and it is possible to increase the sample size from 20 to 40 plants within a site allowing a better characterisation of the site's variations.

It would also be desirable to extend the work to other Etosha plant species, although the existing calibrations that were developed are valid for about 85% of the Park

The need to develop a suitable sampling scheme was also identified. The assistance of high resolution imagery for site selection was proved to be very useful and allowed the identification of homogeneous sites for selected cover types at the scale of NOAA. It is also crucial to develop a suitable sampling scheme for the measurement of biomass within the site. The random selection of plants along the transect is particularly important and the measurement of canopy cover needs to be based on the selected plants. This allows the measurement of variations within the site and the assessment of the precision of the estimate.

It was also demonstrated that single AVHRR images, received locally, could be calibrated against biomass allowing near-to-real-time monitoring of biomass quantity. Nevertheless, the number of points available for the calibration are still limited and more observations would be needed to confirm that the calibration remains stable through space and time. It already appears that the same calibration could be used for grassland and steppe. This is particularly encouraging because grassland and steppe are present in the same areas and are difficult to differentiate at the scale of NOAA. However, it seems that savanna sites need a different calibration especially when the proportion of woody biomass reaches a certain level. More observations need to be collected to be able to confirm this theory and to determine this threshold.

It was shown that biomass maps could be used for the planning of prescribed burning. If local reception of NOAA data is possible, biomass maps can be produced in near-to-real-time and a direct application is to target areas suitable for controlled burning of fire blocks mainly to prevent large scale wildfires. However, more work is required on refining the relationship between biomass and the NDVI for different vegetation communities, but also to investigate whether the effective burning threshold varies according to the vegetation type. The main factor influencing the development of a fire is the amount of grass biomass and the proportion of grass to total green biomass will vary according to the vegetation type, which suggests that fire thresholds will be lower for pure grassland than for savanna. However, this remains to be investigated and stresses the importance of a suitable stratification of the Park's vegetation. Finally, biomass maps could potentially be used for several other purposes linked to rangeland and wildlife management such as monitoring of animal movement and assessment of carrying capacity.

Acknowledgements

This work is part of a collaborative effort with the Etosha Ecological Institute of the Ministry of Environment and Tourism, Namibia, funded by the UK Department for International Development and co-ordinated by the Natural Resources Institute, Chatham, UK. Dr. Ken Campbell from NRI should be thanked especially for his support in the realisation of this study. Finally, staff at EEI should be acknowledged for their participation in some of the field survey activities.

References

BILLE, J.C., 1980 Measuring the primary palatable production of browse plants. *Browse in Africa, the current state of knowledge*. ed. H.N. Le Houérou (International Livestock Centre for Africa, P.O. Box 5689, Addis Ababa, Ethiopia)

CISSÉ, M.I., 1980 The browse production of some trees in the sahel: Relationships between maximum foliage biomass and various physical parameters. *Browse in Africa, the current state of knowledge*. ed. H.N. Le Houérou (International Livestock Centre for Africa, P.O. Box 5689, Addis Ababa, Ethiopia)

DIALLO, O., DIOUF, A., HANAN, N.P., NDIAYE, A., PRÉVOST, Y., 1991. AVHRR Monitoring of Savanna Primary Production in Senegal, West Africa: 1987-1988. *International Journal of Remote Sensing,* **12**, 1259-1279.

DU PLESSIS, W. 1995 Biomass calibration of *Colophospermum mopane* in Etosha National Park. Personal communication

FRANKLIN, J., HIERNAUX, P.H.Y., 1991. Estimating Foliate and Woody Biomass in Sahelian and Sudanian Woodlands using a Remote Sensing Model. *International Journal of Remote Sensing*, **12**, 1387-1404.

FRANKLIN, J., PRINCE, S.D., STRAHLER, A.H., HANAN, N.P., SIMONETT, D.S. 1991 Reflectance and Transmission Properties of West African Savanna Trees From Ground Radiometer Measurements. *International Journal of Remote Sensing*, **12**, 1369-1385

GARCIA-DAGUER, R R , 1993. *Coarse Resolution Remote Sensing for Rangeland Monitoring in North-Central Mexico*. PhD thesis, Silsoe College.

HANAN, N.P., PRINCE, S.D., FRANKLIN, J. 1993 Reflectance Properties of West African Savanna Trees From Ground Radiometer Measurements. II. Classification of Components. *International Journal of Remote Sensing*, **14**, 1081-1097

HEADY, H.F., 1975, *Rangeland Management* (New York: McGraw Hill)

HOLECHEK, J.L., PIEPER, R.D. and HERBEL, C.H., 1995, *Range Management: Principles and Practices* - Second Edition (New Jersey: Prentice Hall)

JUSTICE, C.O. and TOWNSHEND, J.R.G., 1981, Integrating ground data with remote sensing. *Terrain Analysis and Remote Sensing*. ed. J.R.G. Townshend (London: George Allen & Unwin).

KANNENBERG, N. 1995 Grass-biomasse in Savannenbiomen des Etosha National Park: Anwendung des Disc Pasture Meter (DPM).

KAUFMAN, Y.J., HOLBEN, B.N., 1993. Calibration of The AVHRR Visible and Near-IR Bands by Atmospheric Scattering, Ocean Glint and Desert Reflection. *International Journal of Remote Sensing*, **14**, 21-52.

KENNEDY, P., 1989. Monitoring the Vegetation of Tunisian Grazing Lands using the Normalized Difference Vegetation Index. *Ambio,* **18**, 119-123.

LAMPREY, R.H. , DE LEEUW, P.N., 1986 ILCA/UNEP-AVHRR Calibration Project, Methodology And Results For June 1985 (Nairobi: ILCA).

LE ROUX, C.J.G., GRUNOW, J.O., MORRIS, J.W. BREDENKAMP, G.J., SCHEEPERS, J.C., 1988, A Classification of the Vegetation of yhe Etosha National Park. *South African Journal of Botany,* **54**, 1-10.

LOS, S.O., 1993, Calibration adjustment of the NOAA AVHRR normalised difference vegetation index without recourse to channel 1 and 2 data. *International Journal of Remote Sensing,* **14**, 1907-1917.

MCCLOY, K.R., SCHONEVELD, R., Kemp, D., 1993. Measurement of Pasture Parameters from Reflectance Data. *International Journal of Remote Sensing* **14**, 1107-1118.

MICHELL, P., 1982, Value of a Rising-plate Meter for Estimating Herbage Mass of Grazed Perennial Ryegrass - White Clover Swards. *Grass andForage Science* **37**, 81-87.

PEARSON, R.L., TUCKER, C.J., MILLER, L.D., 1976. Spectral Mapping of Shortgrass Prairie Biomass. *Photogrammetric Engineering and Remote Sensing,* **42**, 317-323.

PIEPER, R.D., 1988. Rangeland Vegetation Productivity And Biomass. *Handbook of vegetation science vol. 14. Vegetation science applications for rangeland analysis and management.* ed. P.T. Tueller. (Dordrecht: Kluwer Academic Pub).

PRINCE, S.D., TUCKER, C.J., 1986. Satellite Remote Sensing of Rangelands in Botswana. II. NOAA-AVHRR And Herbaceous Vegetation. *International Journal of Remote Sensing*, **7**, 1555-1570.

PRINCE, S.D., 1991a Satellite Remote Sensing Of Primary Production: Comparison Of Results For Sahelian Grasslands 1981-1988. *International Journal of Remote Sensing,* **12**, 1301-1311

PRINCE, S.D., 1991b A Model of Regional Primary Production for Use with Coarse Resolution Satellite Data. *International Journal of Remote Sensing*, **12**, 1313-1330

RAO, C.R.N. and CHEN, J., 1996, Post-launch Calibration of The Visible and Near-infrared Channels of The Advanced Very High Resolution Radiometer on The NOAA-14 Spacecraft. *International Journal of Remote Sensing,* **17**, 2743-2747

RICHARDSON, A.J., WIEGAND, C.L. 1977 Distinguishing Vegetation from Soil Background Information. *Photogrammetric Engineering and Remote Sensing*, **43**, 1541-1552

SANNIER C.A.D., TAYLOR J.C., DU PLESSIS W. and CAMPBELL K, 1998 Real-Time Vegetation Monitoring with NOAA-AVHRR in Southern Africa for Wildlife Management and Food Security Assessment. *International Journal of Remote Sensing,* **19**, 621-639.

SCRIVNER, J.H., CENTER, M., JONES, M.B., 1986, A Rising Plate Meter for Estimating Production and Utilization. *Journal of Range Managemen,t* **39**:475-477.

TAYLOR, J.C., BELWARD, A.S., HEWETT, D.G., and WYATT, B.K., 1986, *Mapping of vegetation and soils and estimation of biomass in theSahel.* Final report to the Commission of the European Communities on: Monitoring of pastureland ecosystems in the Sahel and mapping of cloud cover and rainfall.

TIETEMA, T., 1993. Biomass Determination of Fuelwood Trees and Bushes of Botswana, Southern Africa. *Forest Ecology and Management* **60**, 257-269.

TROLLOPE, W.S.W., POTGIETER, A.L.F.,1986 Estimating Grass Fuel Loads with a Disc Pasture Meter in The Kruger National Park. *Journal of the Grassland Society of South Africa*, **3**, 148-152

TUCKER, C.J., 1979. Red and Photographic Infrared Linear Combinations for Monitoring Vegetation. *Remote Sensing of Environment* **8**, 127-150.

WAGENAAR, KT , DE RIDDER, N, 1986. *Estimates of Biomass Production and Distribution in the I L P Project Zone In 1985, Based on Satellite NDVI Values.* (ILCA: Addis Ababa).

WALKER, J., CRAPPER, P.F., PENDRIDGE, L.K. 1988 The Crown-gap Ratio (c) And Crown Cover: The Field Study. *Australian Journal of Ecology,* **13**, 101-108

WESTFALL, R.H., PANAGOS, M.D., 1984 A Cover Meter for Canopy and Basal Cover Estimations. *Bothalia* **15**, 241-244

WILLIAMS, J.B. and ROSENBERG, L.J., 1993, Operational reception, processing and application of satellite data in developing countries : Theory and practise. In, K. Hilton (editor) *Towards Operational Applications,* Proceedings of 19th Annual Conference of the Remote Sensing Society, Chester, U.K., on 16-17 September 1993 (Chester: The Remote Sensing Society), 76-81.

WYLIE, B.K., HARRINGTON, J.A. JR., PRINCE, S.D., DENDA, I., 1991. Satellite and Ground-based Pasture Production Assessment in Niger: 1986-1988. *International Journal of Remote Sensing*, **12**, 1281-1300.

CHAPTER 7 CONCLUSIONS AND RECOMMENDATIONS

This thesis was presented as a series of papers with their own conclusions that have been published or have been submitted for publication in refereed journals or in conference proceedings. This chapter is not aimed at repeating the conclusions of each of the individual papers. Rather, it attempts to summarise the main findings of the research and to make recommendations on further research required to improve techniques using remote sensing for a better understanding of vegetation dynamics in Southern Africa.

In previous chapters of the thesis, the role of remote sensing for collecting strategic information about vegetation in Southern Africa was outlined. A particular emphasis was given to demonstrating that it is possible to apply robust methods to the operational assessment of vegetation in Southern Africa. This was done through the combination of widely and regularly available remote sensing data sources (NOAA-AVHRR), survey design, field measurements and time series analysis.

As stressed in previous chapters, products derived from the use of remotely sensed data for strategic monitoring of vegetation currently offer a low level of sophistication which only makes qualitative analyses possible. The main reason for this is the difficulty to extract vegetation related information from existing satellite-derived products such as the NDVI or a thematic classification. As a result, environmental satellite and especially meteorological satellite data are under-utilised. Therefore, the main aim of the research was to attempt to develop techniques that would facilitate information extraction from satellite data in the context of strategic monitoring of vegetation in Africa. In the first chapter, three main areas were identified where remote sensing could play a significant role. These were:

- Land cover / use mapping
- Monitoring of vegetation conditions
- Crop yield and vegetation biomass measurements

Each of these areas will be reviewed in turn to identify the main findings of the research and some recommendations for further work. The issue of data availability is particularly crucial especially for real-time operation and will also be looked at in relation to these areas of applications.

1 Land cover mapping

The implementation of area frame sampling methodology initially developed for the European context and here applied in Africa proved to be a success. In this work, it was demonstrated that the same levels of accuracy are achievable in an African context compared to Europe for an equivalent complexity of the classification scheme. However, the importance of the choice of an appropriate classification scheme has to be recognised. It appears that the use of a structural scheme which allows classes to grouped at various hierarchical levels is best adapted for use with satellite imagery when it is difficult to know a priori which cover types are distinguishable. This approach was particularly successful in Etosha National Park where it was applied in natural vegetation.

Such a classification scheme was implemented within an area frame sampling strategy and proved perfectly adapted for this type of survey. The practicality of this method for operational assessment has sometimes been questioned by the remote sensing community because of problematical access to field sites and the difficulty to identify distinguishable features on the image that are also identifiable in the field. However, the use of a GPS and surrogate information when necessary, obtained for example from a low altitude aerial survey, illustrated in Chapter 2, showed that it was perfectly possible. Another reason often put forward against the implementation of such a survey is the higher cost it implies. However, the survey described in chapter 2 was realised at a fairly low cost: 2 man months with a vehicle and about £400 worth of aviation fuel to survey a 22,000 km^2 area at a 1% sampling rate. The cost of flying was especially cheap because the Namibian Ministry of Environment and Tourism provided the aircraft and the pilot. This may not always be possible but the cost would still remain fairly low considering only a standard light aircraft is required. It should also be added that an extensive ground survey was carried out which took most of the two man months. The ground component took about 10 times the time required by the aerial component. This was to check the accuracy of the aerial survey. In an operational context, the ground component could be reduced to a calibration level only aimed at becoming familiar with the land cover types present in the area.

Furthermore, chapter 2 showed that extracting quantitative information from thematic maps requires an unbiased ground calibration sample. However, although the value of the combination of digital classification with unbiased field survey has been demonstrated, thematic mapping, undertaken in African countries, is often based upon a more traditional combination of photo-interpretation and limited ground survey whenever classes are too difficult to distinguish on the image. Several reasons can be put forward. This can be because such a method is too sophisticated or too costly to be implemented when dealing with a whole country. But equally, it could be because commercial organisations or commissioning organisations do not have the technical knowledge to implement this method or judge its benefits. Rogers (1995) suggested that the rate of adoption of an innovation was affected by five factors:

- *Relative advantage* – degree to which an innovation is perceived as better than the idea it supersedes
- *Compatibility* – degree to which an innovation is perceived as consistent with the existing values, past experiences and needs of potential adopters
- *Complexity* – degree to which an innovation is perceived as relatively difficult to understand and use
- *Trialability* – degree to which an innovation may be experimented with on a limited basis
- *Observability* – degree to which the results of an innovation are visible to others

From the results shown in chapter 2 and in the appendix, it is clear that there are many *relative advantages* to perform land cover mapping in Africa using a combination of unbiased ground observations and digital classification rather than a traditional approach of photo-interpretation. Because the methodology is more robust, results are more consistent and quantitative estimates can be derived with a measurable degree of confidence. However, the implementation of these new techniques requires a greater move toward information technology which up until recently has been lagging behind the rest of the world in African countries. This may have had an effect on the degree of *compatibility* and provide a partial explanation to the relatively low uptake. The level of *complexity* may also be another factor hampering the adoption of the technique. It is true that traditional methods still require expert knowledge but there is a change of emphasis in the nature of the expertise required. Although it is possible to conduct pilot studies to demonstrate the value of digital classification versus traditional methods, this would already require a large investment in the new technology which denotes a low *trialability* level. Finally, *observability* has also been a problem due to the lack of visibility of the benefits to the user: both digital classification and traditional methods produce maps and class areas can be calculated. Users do not always realise that the map produced using traditional methods can be inaccurate because measurement of map accuracy has only recently become a serious issue with the advent of digital handling of spatial databases.

From assessing the five attributes controlling the rate of adoption of an innovation according to Rogers (1995), it becomes clear why digital classification has not been more widely adopted. A sign of hope is that recently prices of IT equipment in some African countries have been falling dramatically both in absolute and relative terms. This is due to a combination of the general trend in the IT market but more importantly to a reduction of import duties on IT equipment. For example, in Namibia, five years ago a standard PC used to cost at least twice as much as in the UK, now prices are more or less equivalent. As a result, many African countries are now entering the IT revolution and this will inevitably have a positive impact on digital analysis of satellite imagery. However, this would only improve the situation for only one of Rogers attributes, *Compatibility*. The adoption of the innovation would most likely require more than just one attribute to be fulfilled. As said earlier, at the moment, the *relative advantage* of a mapping exercise based on the combination of an area frame sample and digital classification of satellite imagery is not very well perceived by users. This is probably due to a low degree of *observability* of the differences with more traditional techniques. The main reason is that up until now, decisions were mostly based on political ground and little emphasis was put on robust unbiased methods of gathering information. However, there is a chance that the drive towards better governance in African countries will increase the demand for such methods. Techniques such as the combination of area frame sampling with digital classification put great emphasis on an objective and quantitative approach. In order to produce maps based on unbiased observations, the use of a properly designed field survey is the only alternative dismissing any other more traditional techniques.

The availability of the data is obviously another important issue to consider in using remote sensing for thematic mapping. It was shown in the appendix that results could be considerably affected when the imagery could not be obtained on time. However, the available sources of high-resolution imagery have improved in recent years and delivery time from image supplier is now of the order of one to two weeks. Furthermore, environmental satellite imagery can now be delivered on CD-ROM which means that no specialised hardware is required to read the data in the system. Suppliers have also created a series of pre-processed products (radiometric and geometric corrections and image enhancement techniques have been applied) implying that it is no longer necessary to be an image-processing specialist to start making use of satellite imagery. Although image suppliers try in this way to make their products available to a wider public, the type of analysis described in chapter 2 and in the appendix would still require expert users. Nevertheless, the cost of high-resolution satellite imagery is still a major factor when deciding whether to use satellite imagery or not (Harris 1997) and prices are still too high especially for African users. This means that it is unlikely that the imagery will be purchased outside of well-funded projects. Although the price policy of SPOTimage does not seem likely to change, it appears that the US government will radically change its policy for Landsat 7, only charging US$600 per scene.

2 Monitoring of vegetation conditions

In this thesis it was shown that the monitoring and assessment of vegetation conditions requires at least two elements. Firstly, meteorological satellite data such as NOAA-AVHRR must be acquired in real-time to ensure that data cover the country or study area considered and also to ensure that the information derived is not obsolete when it becomes available to users. This normally implies the use of a local receiving station. Fortunately, low-cost NOAA receiving stations have now been available for a number of years and installed in several African countries. These receivers allow the acquisition of AVHRR data in those countries on a routine basis with a system that is relatively simple to operate.

Other benefits include a better sense of ownership of the outputs produced and the capacity to tailor products to users' requirements. Furthermore, the operation of a receiving station is perceived as a "high tech" employment opportunity in countries where these jobs are not widely available. It should also be stressed that for the above reasons and because there are no other methods available to monitor factors affecting food security in real time, the uptake of these techniques is much faster than for land cover mapping. This was demonstrated in chapter 5 which shows how the vegetation monitoring techniques described in chapter 3 were successfully transferred to Botswana in an operational context.

The receivers output standard calibrated products such as individual channels with reflectance values or the NDVI. These products, notably the NDVI in the case of vegetation monitoring, are difficult to interpret on their own and especially by non-specialists. They need to be compared to a historical archive which is the second element required for monitoring vegetation conditions with remote sensing. The comparison of the current image acquired by the receiver with the archive allows users to determine whether conditions for a specific area are higher or lower than normal. Such a technique was developed for the purpose of this research and named the VPI. The development and implementation of the VPI was described in Chapter 3. The main comparative advantage of the VPI is a greater reliability than other NDVI derived products and the ability to be understood by non-specialists users. This should facilitate the dissemination of the information through the hierarchy, thus improving the way decisions are taken by providing objective and reliable information to decision makers.

The initial satellite archive that was used (FAO-ARTEMIS NDVI) covered a 10-year period which was the maximum available at the time. The archive was developed by FAO for the purpose of monitoring and its main advantage was that it had the same level of processing as the receiving station's outputs. This is an essential requirement of the analysis otherwise the comparison of the current with archive data will not be possible.

Subsequently, a second archive, PAL, became available which covered a longer time period and was planned to be continuously updated. This archive is particularly applicable for time series analysis allowing extension of the period of record to a total of 15 years. The extension of the time series is crucial to identify extreme events more precisely. However, the PAL archive had a different level of processing than that of ARTEMIS or receiving station's outputs, notably incorporating atmospheric corrections. Therefore, it was necessary to check the level of compatibility between the two archives before starting to use the extended time series. This study was described in Chapter 4 and showed that there were significant differences between the two data sets but also that it was possible to apply an empirical correction based on linear regression to make the two archives compatible. Furthermore, the study showed that the atmospheric corrections applied to PAL did not appear to bring any improvements to the data although this remains to be confirmed by a more detailed study. It should also be noted that atmospheric corrections are difficult to apply and more particularly in the context of real-time applications. Most of the time, input parameters for a semi-deterministic atmospheric model are missing or incomplete and do not take into account spatial variations. Instead, approximations or averaged values have to be used. As a result, there is little evidence in the literature that such atmospheric corrections improve the relationship between remotely sensed data and biophysical parameters. Therefore, it seems that atmospheric corrections should be discarded from processing chains unless their degree of reliability improves significantly.

The extended time series was used to implement the VPI in a third study area, Botswana, where it was produced operationally. VPI maps were generated in near real-time for the 1996-97 and 1997-98 rainy seasons. More importantly the maps were presented and well-received by decision-makers.

Currently the resolution of the VPI maps is limited by the pixel resolution of the imagery used to generate the stratification. In the case of Zambia and Botswana this is the resolution of the historical NOAA archive which is about $64km^2$. Clearly this degree of precision is insufficient for some regional planning for which the full resolution of AVHRR would be more adapted. The application of the VPI in Etosha showed that it was possible to apply the technique with LAC data. This was made possible thanks to the stratification derived from the digital classification of Landsat TM. The use of TM for a whole country for such an exercise can not realistically be envisaged but TM could be replaced efficiently with a time series of 1km AVHRR. Unfortunately a long-term archive of AVHRR at that resolution does not exist at present.

One of the main outputs from this work was the development of the VPI which was applied successfully in three different sites. Chapter 3 stressed a number of theoretical advantages of the VPI technique compared to existing ones but it would be necessary to carry out more research to compare the VPI with traditional techniques to see whether these advantages are observed in practise. The main advantages of the VPI are on the one hand the capability to undertake return period analysis and secondly the VPI does not make any assumption about the nature of the statistical distribution of the NDVI. However, for some cases, the assumption made by other methods may be close enough to reality to be acceptable. Furthermore, the way in which the VPI has been used is still very much for a qualitative approach, which is perfectly adapted for the purpose of an early warning system. However, this does not make use of the technique's full potential and further work is required to develop more quantitative analysis based on the VPI.

3 Vegetation quantity measurements

It was also demonstrated in Chapter 6 that AVHRR could be used for real-time monitoring of biomass providing that the quality of the imagery was carefully controlled (close to nadir and as free as possible of atmospheric contamination). This is easily achievable in the context of local reception where the operator can assess the imagery to be used. However, success in the use of AVHRR for vegetation monitoring also requires the field data to be representative of a calibration site suitable for AVHRR. Special care was taken to make sure that this was the case and a number of existing techniques for biomass measurements were successfully adapted and integrated to use for calibration of satellite imagery. This stresses the importance of the expert's assessment of the imagery compared to a fully automated processing chain that cannot provide the same level of precision. A good example is geometric correction. Most automated processing chains of AVHRR rely on orbital parameters and manual shifts of the image to match a vector file of the country or region considered. This cannot achieve the level of precision reached when using properly selected ground control points.

The correlation between pooled field observations of biomass and satellite observations of NDVI, was strong and consistent, thus making the derivation of a unique relation possible. However, the scrutiny of the results seems to suggest that different relationships exist for separate cover types, although this would have to be confirmed by further data collection and analysis. Again this stresses the importance of stratification and this is where the role of a high resolution satellite derived vegetation map could prove extremely useful to improve separation between cover types. This would allow the application of a different relation for each vegetation cover thus improving the reliability of the resulting biomass map. Another approach to solve this problem would be linked to the derivation of quantitative products from the VPI mentioned at the end of the previous section. This should prevent the necessity of applying a different relation for each cover type because the variation in cover type response is removed when the VPI is used. An example is briefly outlined at the end of Chapter 3, where it seems that the VPI could be used for the prediction of maize production in Zambia.

As mentioned above, the quality of the AVHRR imagery is another important factor to consider. It is well known that some of the problems associated with AVHRR are the lack of on-board calibration that prevents reliable and constant monitoring of sensor degradation, and the geometric deformation and other effects due to the large scan angle. For many years AVHRR was the only operational sensor of its category but new sensors that are specifically adapted to vegetation monitoring are now or will soon become operational. SPOT-4 Vegetation has recently been launched and provides a much improved level of calibration and geometry thanks to the push-broom configuration of its sensor. With its blue waveband, it also provides a way of implementing more reliable atmospheric corrections. A relevant area of research would be to evaluate the significance of the improved level of reliability of products derived from SPOT4 Vegetation compared to AVHRR. The use of WiFS (Wide Field Sensor) on board IRS 1C and 1D is also worth considering and would provide enhanced information. It has a higher spatial resolution (188m pixel size) while maintaining a fairly high temporal resolution (4 to 5 days), which should make possible the monitoring of vegetation at a much larger map scale and could be used for applications such as catchment management. More relevant at national level, the use of Meteosat Second Generation data opens up new possibilities for land applications (Cihlar et al. 1999). Increased spatial and spectral resolution (compared to the first generation of Meteosat) makes the data comparable to AVHRR with a much higher temporal resolution (one image every 15 minutes) and a better geometric fidelity. For instance, the high repeat coverage over the same area offer interesting opportunities for the monitoring of bidirectional reflectances (Cihlar et al. 1999). However, a major challenge with the new sources of data will be to maintain the continuity with archived data which is essential for the application of a methodology such as the VPI.

4 Integrated vegetation monitoring at national level

Each of the separate elements of this research, land cover mapping, vegetation monitoring and biomass assessment, only has limited interest on its own in an operational context but it is the integration of the elements that make them particularly powerful as a natural resources monitoring tool. One could imagine a multi-scale survey where detailed land cover mapping would be carried out regularly on a few selected areas. This could be linked to fairly detailed fieldwork to estimate vegetation quantity and linked to AVHRR for allowing a complete coverage of the country. This type of system would provide real-time and reliable information and give a more quantitative aspect to early warning and monitoring systems. As a result the information produced would complement or sometimes could even replace information produced by governmental statistical services. Traditionally, statistical services in African countries were inherited from colonial times and are not necessarily adapted to the country's requirements or data gathering capacity. Estimates often come too late to be really useful and are derived using inappropriate methods for which no precision estimates or data quality measures are available.

Unfortunately, current management decisions in Africa are rarely taken on the basis of objective information. Much work still remains to be done in raising the awareness of decision-makers to the value of information derived from remote sensing techniques. Many good and useful products can be derived from the integration of satellite imagery and field observation but too often these products do not leave the offices of the experts that have been producing them. Remote sensing will certainly gain more recognition as a science when the information it produces will become an integral part of the decision making process in the management of vegetation resources. Remote sensing techniques such as the ones developed and tested in this work can provide an up to date, reliable and unbiased source of information for vegetation resources. In this work methods were developed and tested for land cover mapping, vegetation monitoring and biomass assessment. These methods were developed with an operational approach in mind and it is now possible to apply and combine these techniques to contribute to the development of an integrated vegetation monitoring system at national level, thus enabling countries heavily dependant on agriculture and rangeland to better manage their vegetation resources.

REFERENCES

ALLEN D.A. and HANUSCHAK G.A. 1988, *The Remote Sensing Applications Program of the National Agricultural Statistics Service: 1980-1987.* United States Department of Agriculture, National Agricultural Statistics Service, Research and Applications Division, SRB Staff Report Number SRB-88-08, 43 pp.

AZZALI, S., 1991, Interpretation Of Crop Growth Patterns By Means Of NDVI- Time Series In Zambia. *Geocarto International*, **3,** 15-26.

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

BELWARD, A.S., 1996, AVHRR data sets for global terrestrial ecosystem monitoring. In, G. D'Souza, A.S. Belward and J.P. Malingreau (Editors) *Advances in the Use of NOAA AVHRR Data for Land Applications, Euro Courses, Remote Sensing, Volume 5* (Dordrecht, Boston, London: Kluwer Academic Publishers), pp.1-19.

BILLE, J.C., 1980 Measuring the primary palatable production of browse plants. *Browse in Africa, the current state of knowledge.* ed. H.N. Le Houérou (International Livestock Centre for Africa, P.O. Box 5689, Addis Ababa, Ethiopia)

BIRD, A. C., PRATT, N. D. and LAWAN, A. I., 1995, The development of GIS and Remote sensing techniques in the Centre for Arid Zone Studies, North East Nigeria, RSS Workshop, Chatham Dec 19 1995.

BONIFACIO, R., DUGDALE, G., MILFORD, J.R., 1993, Sahelian Rangeland Production In Relation To Rainfall Estimates From Meteosat. *International Journal of Remote Sensing,* **14**, 2695-2711.

BOUGHEY, A. S., 1957. The physiognomic delimitation of west African vegetation types. *J. W. Afr. Sci. Ass.* Vol. 3 (2): 148-165.

BURT, J.E. and BARBER, G.M., 1996, *Elementary Statistics for Geographers.* The Guilford Press: New York, London.

CIHLAR, J., BELWARD, A. and GOVAERTS, Y., 1999, Meteosat Second Generation Opportunities for Land Surface Research and Applications. EUMETSAT Scientific Publications, ISBN 92-9110-031-5.

CISSÉ, M.I., 1980 The browse production of some trees in the sahel: Relationships between maximum foliage biomass and various physical parameters. *Browse in Africa, the current state of knowledge.* ed. H.N. Le Houérou (International Livestock Centre for Africa, P.O. Box 5689, Addis Ababa, Ethiopia)

COTTER J. and NEALON J. 1987. *Area Frame Design for Agricultural Surveys.* United States Department of Agriculture, National Agricultural Statistics Service, Research and Applications Division, 67 pp.

CRACKNELL, A.P., 1997. *The Advanced Very High Resolution Radiometer.* Taylor and Francis: London, ISBN 0-7484-0209-8, 534 pp.

DAVENPORT, M.L., and NICHOLSON, S.E., 1993, On the relation between rainfall and Normalized Difference Vegetation Index for diverse vegetation types in East Africa. *International Journal of Remote Sensing*, **12**, 2369-2389.

DI, L., RUNDQUIST, D.C. and HAM, L., 1994, Modelling relationships between NDVI and precipitation during vegetative growth cycles. *International Journal of Remote Sensing,* **15**, 2121-2136.

DIALLO, O., DIOUF, A., HANAN, N.P., NDIAYE, A., PRÉVOST, Y., 1991, AVHRR Monitoring Of Savanna Primary Production In Senegal, West Africa: 1987-1988. *International Journal of Remote Sensing,* **12**, 1259-1279.

DU PLESSIS, W. 1995 *Biomass calibration of Colophospermum Mopane in Etosha National Park.* Personal communication

EDMONDS, A.C.R., 1976, *The Republic of Zambia, Vegetation Map 1:500000* (Forest Department, the Government of Zambia).

ERDAS, 1995, *ERDAS Field Guide* ERDAS Inc: Atlanta.

FAO, 1996, *The state of food and agriculture 1996* ISBN 92-5-103858-9 (Rome: FAO).

FAO, 1997, *The state of food and agriculture 1997* ISBN 92-5-104005-2 (Rome: FAO).

FLITCROFT, I.D., MILFORD, J.R. and DUGDALE, G., 1989, Relating point to area rainfall in semi-arid West Africa and the implications for rainfall estimates derived from satellite data. *Journal of Applied Meteorology*, **58**, 193-207.

FRANKLIN, J., HIERNAUX, P.H.Y., 1991. Estimating Foliate And Woody Biomass In Sahelian And Sudanian Woodlands Using A Remote Sensing Model. *International Journal of Remote Sensing* **12**(6):1387-1404.

FRANKLIN, J., PRINCE, S.D., STRAHLER, A.H., HANAN, N.P., SIMONETT, D.S. 1991 Reflectance And Transmission Properties Of West African Savanna Trees From Ground Radiometer Measurements. *International Journal of Remote Sensing* **12**(6):1369-1385

GARCIA-DAGUER, R R , 1993. *Coarse Resolution Remote Sensing For Rangeland Monitoring In North-Central Mexico.* PhD thesis, Silsoe College, 1993.

GOWER, J.F.R., 1992, Low cost satellite sensor image reception for NOAA HRPT and other compatible data. *International Journal of Remote Sensing,* **14**, 177-181.

GROTEN, S.M.E., 1993, N D V I-crop Monitoring And Early Yield Assessment Of Burkina Faso. *International Journal of Remote Sensing,* **14**, 1495-1515.

HANAN, N.P., PRINCE, S.D., FRANKLIN, J. 1993 Reflectance Properties Of West African Savanna Trees From Ground Radiometer Measurements. I I. Classification Of Components. *International Journal of Remote Sensing* **14**(6):1081-1097

HANUSCHAK, G., SIGMAN, R., CRAIG, M., OZGA, M., LUEBBE, R., COOK, P., KLEWENO, D. and MILLER, C., 1979, Obtaining Timely Crop Area Estimates Using Ground-Gathered and Landsat Data, Statistical Research Division, Economics Statistics and Co-operative Services, USDA, Technical Bulletin No. 1609.

HESS, T., STEPHENS, W. and THOMAS, G., 1996, Modelling NDVI from decadal rainfall data in the North East Arid Zone of Nigeria. *Journal of Environmental Management.*

HIELKEMA, J.U. and SNIJDERS, F.L., 1994, Operational use of environmental satellite remote sensing and satellite communications technology for global food security and locust control by FAO: The ARTEMIS and DIANA systems. *Acta Astronautica*, **32**, 603-616.

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

HUTCHINSON, C.F., 1991, Uses Of Satellite Data For Famine Early Warning In Sub-Saharan Africa. *International Journal of Remote Sensing,* **12**, 1405-1421.

ILWIS 1993, *ILWIS User's Manual,* International Institute for Aerospace Survey and Earth Sciences, Enschede, Netherlands, pp 6.48-49.

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.

JUSTICE, C.O. and TOWNSHEND, J.R.G., 1981, Integrating ground data with remote sensing. *Terrain Analysis and Remote Sensing*. ed. J.R.G. Townshend (London: George Allen & Unwin).

JUSTICE, C.O., DUGDALE, G., TOWNSHEND, J.R.G., NARRACOTT, A.S., KUMAR, M., 1991, Synergism Between N O A A - A V H R R And Meteosat Data For Studying Vegetation Development In Semi-arid West Africa. *International Journal of Remote Sensing,* **12**, 1349-1368.

KANNENBERG, N. 1995 Grass-biomasse In Savannenbiomen Des Etosha National Park: Anwendung Des Disc Pasture Meter (DPM), unpublished.

KAUFMAN, Y.J., HOLBEN, B.N., 1993, Calibration Of The AVHRR Visible And Near- I R Bands By Atmospheric Scattering, Ocean Glint And Desert Reflection. *International Journal of Remote Sensing,* **14**, 21-52.

KENNEDY, P., 1989, Monitoring The Vegetation Of Tunisian Grazing Lands Using The Normalized Difference Vegetation Index. *Ambio,* **18**, 119-123.

KIMMAGE, K. and ADAMS, W.M., 1990, Small-scale farmer-managed irrigation in Northern Nigeria, *Geoforum,* 20:435-43.

KOGAN, F.N., 1990, Remote Sensing Of Weather Impacts On Vegetation In Non-homogeneous Areas. *International Journal of Remote Sensing,* **11**, 1405-141.

LAMBIN, E.F., CASHMAN, P., MOODY, A., PARKHURST, B.H., PAX, M.H. and SCHAAF, C.B., 1993, Agricultural Production Monitoring In The Sahel Using Remote Sensing: Present Possibilities And Research Needs. *Journal of Environmental Management,* **38**, 301-322.

LAMPREY, RH , DE LEEUW, PN, 1986 *ILCA/UNEP-AVHRR Calibration Project, Methodology And Results For June 1985*. ILCA: Nairobi, 1986

LATHAM, J. S., FERNS, D. C., Colwell, J. E., Reinhold, R. and Jebe, E. H., 1983, Monitoring the Changing Aereal Extent of Irrigated Lands of the Gefara Plain, Libya, Adv. Space Res., V2, No 8, pp57-68.

LE COMTE, D.M., 1989, Using AVHRR for early warning of famine in Africa. *Photogrammetric Engineering and Remote Sensing*, **55**, 168-169.

Le Roux, C.J.G., Grunow, J.O., Morris, J.W. Bredenkamp, G.J., Scheepers, J.C., 1988, A Classification Of The Vegetation Of The Etosha National Park. *South African Journal of Botany,* **54**, 1-10.

Le Roux, C.J.G., Grunow, J.O., Morris, J.W. Bredenkamp, G.J., Scheepers, J.C., 1988, A Classification Of The Vegetation Of The Etosha National Park. *South African Journal of Botany,* **54**, 1-10.

Linsley, R.K., Kohler, M.A. and Paulhus, J.L.H., 1975, *Hydrology for Engineers* (New York, Toronto, London: McGraw-Hill).

Los, S.O., 1993, Calibration adjustment of the NOAA AVHRR normalised difference vegetation index without recourse to channel 1 and 2 data. *International Journal of Remote Sensing,* **14**, 1907-1917.

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.

Maselli, F., Conese, C., Petkov, L., Gilabert, M.A., 1993, Environmental Monitoring And Crop Forecasting In The Sahel Through The Use Of NOAA-NDVI Data. A Case Study: Niger 1986-89. *International Journal of Remote Sensing,* **14**, 3471-3487.

McCloy, K.R., Schoneveld, R., Kemp, D., 1993. Measurement Of Pasture Parameters From Reflectance Data. *International Journal of Remote Sensing* **14**(6):1107-1118.

Mitchell, P., 1982, Value Of A Rising-plate Meter For Estimating Herbage Mass Of Grazed Perennial Ryegrass - White Clover Swards. *Grass andForage Science* **37**:81-87.

Moyo, S., O'Keefe, P. and Sill, M., 1993, *The Southern African Continent, Profiles of the SADC Countries* (London: Earthscan Publications Ltd)

Pearson, R.L., Tucker, C.J., Miller, L.D., 1976. Spectral Mapping Of Shortgrass Prairie Biomass. *Photogrammetric Engineering and Remote Sensing* **42**(3):317-323.

Pieper, R.D., 1988. Rangeland Vegetation Productivity And Biomass. *Handbook of vegetation science 14. Vegetation science applications for rangeland analysis and management*. ed. P.T. Tueller. Kluwer Academic Pub..

Powell, M.J. and Pulles J.H.M., 1996, *Animal Production Simulation and Range Assessment Model for Botswana (APSRAMB), Volume I: Theory, User Manual, and Vegetation Inventory.* FAO/UNDP/Goverment of Botswana, Land Use Planning for Sustainable Agriculture Development, Project BOT/94/001, Field Document 17, 152pp.

Prince, S.D. and Astle, W.L. 1986, Satellite Remote Sensing Of Rangelands In Botswana. I. Landsat MSS And Herbaceous Vegetation. *International Journal of Remote Sensing,* **7**, 1533-1553.

Prince, S.D. and Tucker, C.J., 1986. Satellite Remote Sensing Of Rangelands In Botswana. II. NOAA-AVHRR And Herbaceous Vegetation. International Journal of Remote Sensing **7**(11):1555-1570.

PRINCE, S.D., 1991a, Satellite Remote Sensing Of Primary Production: Comparison Of Results For Sahelian Grasslands 1981-1988. *International Journal of Remote Sensing,* **12**, 1301-1311.

PRINCE, S.D., 1991b A Model Of Regional Primary Production For Use With Coarse Resolution Satellite Data. *International Journal of Remote Sensing* **12**(6):1313-1330

RAO, C.R., CHEN, J.1996, Post-launch Calibration Of The Visible And Near-infrared Channels Of The Advanced Very High Resolution Radiometer On The NOAA-14 Spacecraft. *International J. Remote Sensing* **17**(14):2743-2747

RAO, C.R.N., 1993, *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, NOAA/NESDIS, Washington, DC.

RASMUSSEN, M.S., 1992, Assessment of millet yields and production in northern Burkina Faso using integrated NDVI from the AVHRR. *International Journal of Remote Sensing*, **13**, 3431-3442.

RICHARDSON, A.J., WIEGAND, C.L. 1977 Distinguishing Vegetation From Soil Background Information. *Photogrammetric Engineering and Remote Sensing* **43**(12):1541-1552

RINGROSE S. and MATHESON W., 1987, Spectral Assessment of Indicators of Range degradation in the Botswana Hardveld Environment. *Remote Sensing of Environment*, **23**, 379-396.

ROGERS E.M., 1995, *Diffusion of innovations.* Fourth edition. The Free Press: New York, 204-251

SANNIER C.A.D., TAYLOR J.C. and CAMPBELL K., 1998b, Compatibility of FAO-ARTEMIS and NASA Pathfinder AVHRR Land NDVI data archives for the African continent. *International Journal of Remote Sensing*, submitted.

SANNIER C.A.D., TAYLOR J.C., du PLESSIS W. and CAMPBELL K., 1998a, Real-Time Vegetation Monitoring with NOAA-AVHRR in Southern Africa for Wildlife Management and Food Security Assessment. *International Journal of Remote Sensing*, **19**, 621-639.

SANNIER C.A.D., TAYLOR J.C., SLADE G., NTABENI T., NGAKANE S. and CAMPBELL K., 1998c. Real-Time Rangeland Monitoring in Botswana with NOAA-AVHRR. In *RSS98 Developing International Connections.* Remote Sensing Society annual conference.

SANNIER, C., TAYLOR, J. C., du PLESSIS, W. and CAMPBELL, K., 1995, Application of Remote Sensing and GIS for Monitoring Vegetation in Etosha National Park, RSS Workshop, Chatham Dec 19 1995.

SCHERR, S.J. and YADAV, S., 1997 Land Degradation in the Developing World: Issues and Policy Options for 2020. 2020 Brief 44 (IFPRI).

SCHULTZ, J., 1975, *Republic of Zambia, Land Use 1:750000* (University of Zambia, Ministry of Rural Development, the Government of Zambia).

SCRIVNER, J.H., CENTER, M., JONES, M.B., 1986, A Rising Plate Meter For Estimating Production And Utilization. *Journal of Range Management* **39**(5):475-477.

SMITH P., 1998, *Coding errors in the Pathfinder AVHRR Land (PAL) data set's atmospheric correction algorithm.* Goddard DAAC home page. URL: http://daac.gsfc.nasa.gov/CAMPAIGN_DOCS/LAND_BIO/PAL_coding_errors.pdf.

STEEL, R.G.D. and TORRIE, J.H., 1960, *Principles and Procedures of Statistics* (New York, Toronto, London: McGraw-Hill).

STEINWAND, D.R., 1994, Mapping raster imagery to the Interrupted Goode Homolosine projection. *International Journal of Remote Sensing,* **15**, 3463-3471.

STORY, M., and CONGALTON, R. G., 1986, Accuracy assessment: a user's perspective, Photogrammetric Engineering and Remote Sensing, 52:397-399.

TAYLOR, J.C., BELWARD, A.S., HEWETT, D.G., and WYATT, B.K., 1986, *Mapping of vegetation and soils and estimation of biomass in theSahel.* Final report to the Commission of the European Communities on: Monitoring of pastureland ecosystems in the Sahel and mapping of cloud cover and rainfall.

TAYLOR, J. C., BIRD, A. C., KEECH, M. A. and STUTTARD, M. J., 1991, Landscape change in the National Parks of England and Wales - Final Report Vol I Main Report, Silsoe College, ISBN 1 871564 15 8, 97p

TAYLOR, J.C., D'SOUZA, G. and COLEMAN, V.R., 1992, Vegetation Conditions and Yield Indicators in England and Wales using NOAA-AVHRR data (Monitoring of Vegetation and Yield Indicators). *Proceedings of the conference on the Applications of Remote Sensing to Agricultural Statistics held in Villa Carlotta, Belgirate, Lake Maggiore, Italy, on 26-27 November 1991* (Brussels, Luxembourg: Commission of the European Communities), pp. 249-256.

TAYLOR, J. C. and EVA, M.D. 1992, *Regional Inventories on Beds, Cambs and Northants (UK).* Final report. Contract No. 4817-92-06 ED ISP GB Joint Research Centre, Commission of the European Communities. Silsoe College, Silsoe.

TAYLOR, J. C. and EVA, M.D. 1993, Operational use of remote sensing for estimating crop areas in England. In *Towards operational applications.* Proceedings of the 19^{th} Annual Conference of the Remote Sensing Society, Chester, UK.

TAYLOR JC, SANNIER CAD, DELINCÉ J and GALLEGO FJ 1997, *Regional Crop Inventories in Europe Assisted by Remote Sensing:1988-1993*, European Commission, EUR 17319, 80pp.

TIETEMA, T., 1993. Biomass Determination Of Fuelwood Trees And Bushes Of Botswana, Southern Africa. *Forest Ecology and Management* **60**:257-269.

TILL, R., 1973, The use of linear regression in geomorphology. *Area*, **5**, 303-308.

TROLLOPE, W.S.W., POTGIETER, A.L.F.,1986 Estimating Grass Fuel Loads With A Disc Pasture Meter In The Kruger National Park. *Journal Grassland Society South Africa* **3**(4):148-152

TUCKER C.J. 1996, History of the use of AVHRR data for land applications. In, G. D'Souza, A.S. Belward and J.P. Malingreau (Editors) *Advances in the Use of NOAA AVHRR Data for Land Applications, Euro Courses, Remote Sensing, Volume 5* (Dordrecht: Kluwer Academic), pp.1-19.

TUCKER, C.J., 1979. Red And Photographic Infrared Linear Combinations For Monitoring Vegetation. *Remote Sensing of Environment* **8**:127-150.

WAGENAAR, KT , DE RIDDER, N, 1986. *Estimates Of Biomass Production And Distribution In The I L P Project Zone In 1985, Based On Satellite NDVI Values*. ILCA: Addis Ababa, 1986.

WALKER, J., CRAPPER, P.F., PENRIDGE, L.K. 1988 The Crown-gap Ratio (c) And Crown Cover: The Field Study. *Australian Journal of Ecology* **13**:101-108

WEIBULL, W., 1939, A statistical Theory of the Strength of Materials, *Ing. Vetenskapsakad. Handl. (Stockh.),* **151**, 15.

WESTFALL, R.H., PANAGOS, M.D., 1984 A Cover Meter For Canopy And Basal Cover Estimations. *Bothalia* **15**(1,2):241-244

WILLIAMS, J.B. and ROSENBERG, L.J., 1993, Operational reception, processing and application of satellite data in developing countries : Theory and practise. In, K. Hilton (editor) *Towards Operational Applications, Proceedings of 19th Annual Conference of the Remote Sensing Society, Chester, U.K., on 16-17 September 1993 (Chester: The Remote Sensing Society)*,pp. 76-81.

WYLIE, B.K., HARRINGTON, J.A. JR., PRINCE, S.D., DENDA, I., 1991. Satellite And Ground-based Pasture Production Assessment In Niger: 1986-1988. *International Journal of Remote Sensing* **12**(6):1281-1300.

SPACE APPLICATIONS INSTITUTE

Agriculture

REGIONAL CROP INVENTORIES IN EUROPE ASSISTED BY REMOTE SENSING: 1988-1993

Synthesis Report of the MARS Project - Action 1

by
C. Taylor, C. Sannier, J. Delincé, F.J. Gallego

JOINT RESEARCH CENTRE
EUROPEAN COMMISSION

1997

EUR 17319 EN

SPACE APPLICATIONS IN

Agriculture

REGIONAL CROP INVENTORIES IN EUROPE ASSISTED BY REMOTE SENSING: 1988-1993

Synthesis Report of the MARS Project - Action 1

by
C. Taylor, C. Sannier,
Silsoe College, Cranfield University

J. Delincé[1], F.J. Gallego
JRC, Ispra

[1] Now with DG VI G4

JOINT RESEARCH CENTRE
EUROPEAN COMMISSION

1997

EUR 17319 EN

Cataloguing data can be found at the end of this publication

Printed in Italy

Contents

Tables summary

Figures summary

Summary

Background

In 1986, the European Community (EC) carried out a feasibility study, through its Directorate General for Agriculture (DG VI), to see if remote sensing technology could be used to improve the collection of agricultural information to assist the management of the Common Agricultural Policy (CAP). The working party visited the national statistical offices in France, Italy, Germany, United Kingdom and Spain; and, DG VI and the European Statistical Office (EUROSTAT) in the EC, to assess the statistical methodology for providing agricultural information then currently used by those member states. The potential use of remote sensing was assessed in consultation with centres in France, Italy, Germany, the United Kingdom and Spain and at the Joint Research Centre (JRC-Ispra). On-going projects using remote sensing to assist the acquisition of agricultural information in France, Italy and Spain were studied. The managers of the European Agricultural Guidance and Guarantee Fund (FEOGA) were consulted to identify the EC's needs for agricultural information.

The feasibility study identified several aspects of the existing agricultural information system where improvements were needed among these were:

- improved timeliness of information. The requirement was for first estimates of winter and spring crops in the first half of June and improved estimates of these and for summer crops in the first half of September.
- the development of an objective methodology which could be applied in all member states and be uniformly credible

Remote sensing methods were identified as the way forward to achieve these objectives and by 1988, the EC started a programme entitled Monitoring Agriculture with Remote Sensing (MARS), managed by the Joint Research Centre (JRC) at Ispra. One of the first aims in the MARS programme was to investigate an operational methodology for carrying out inventories of crop areas with the aid of remote sensing, within selected statistical reporting regions. This became known as Action 1. The main aim was to estimate the area of all crops of interest to the EC within specified geographical regions, which were used by member states for statistical reporting. A second aim was to obtain estimates of crop yields per unit area in the same regions so that estimates of total production could be made.

This report is a synthesis of the work, the results and an assessment of the achievements of Action 1. The report consists of this stand-alone executive summary and five additional sections. The first three enlarge on the main elements of the methodology: ground survey, remote sensing and the combination of these using regression. These sections also present a sysnthesis of the results and comments on their accuracy. PART 4 is an assessment of technical factors influencing the results. PART 5 is the Bibliography, listing all the contractors reports and other reference material which have been used to provide the basis of this work.

Geographical evolution

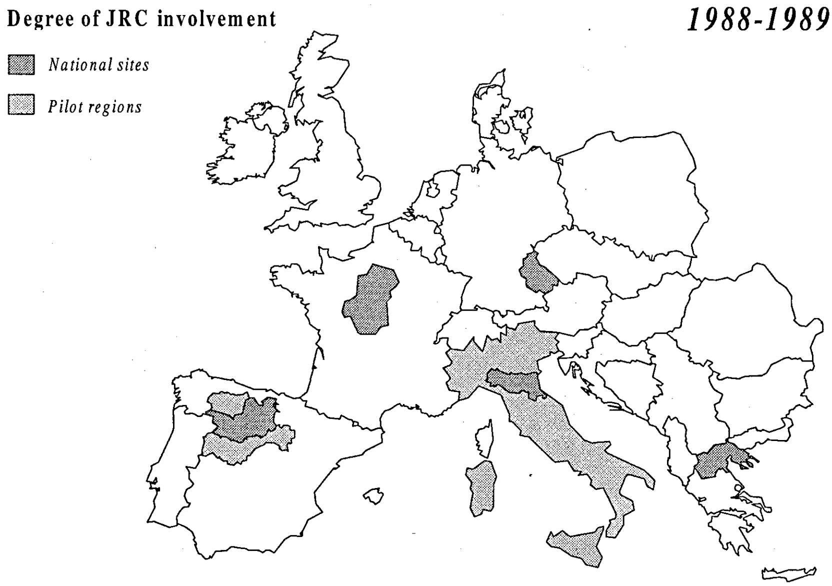


Figure 1.A

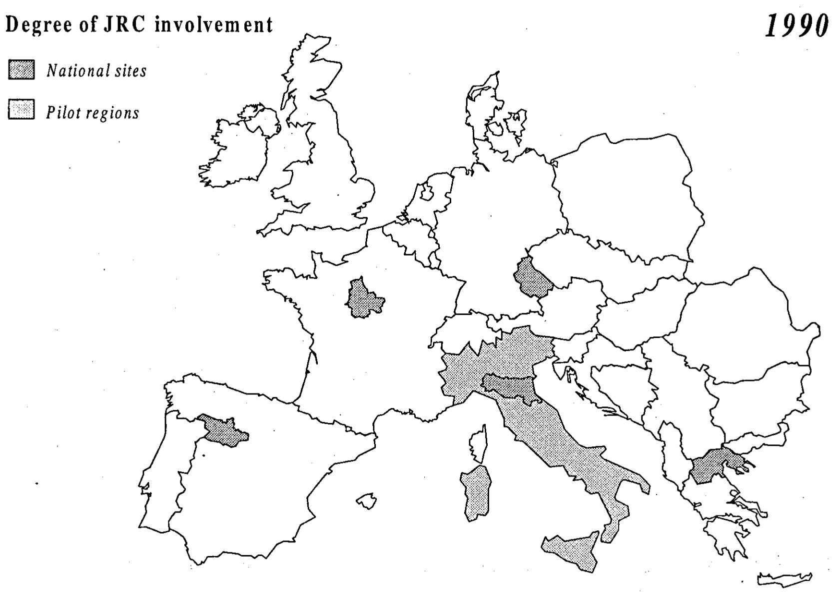


Figure 1.B

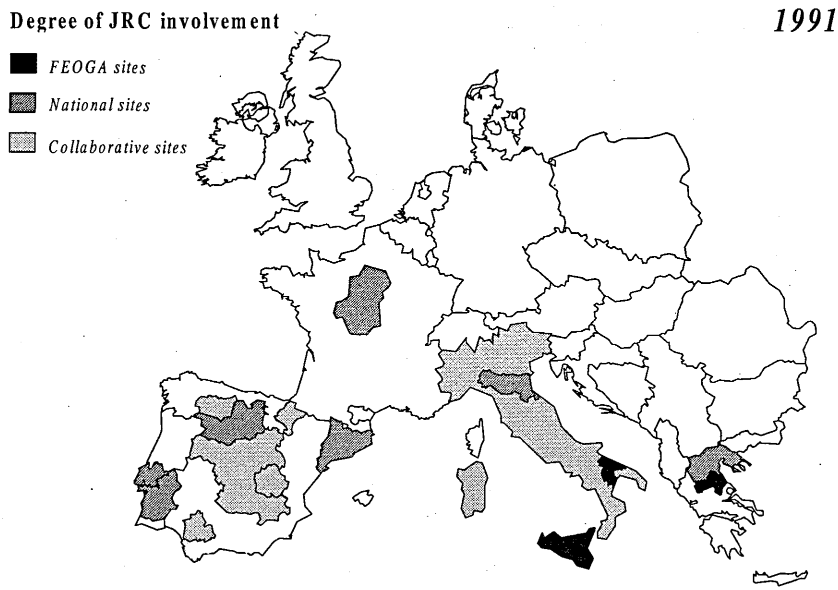


Figure 1.C

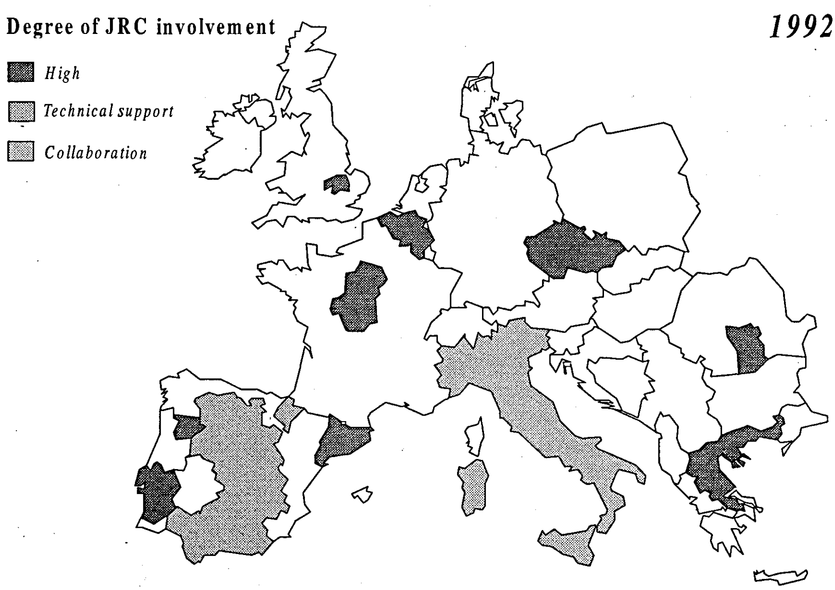


Figure 1.D

Figure 1 (A;B;C;D): Geographical evolution of test sites used for regional crop inventories in Europe: 1988 - 1992

The locations of the regional inventory programme study sites are shown in Figure 1. Five test sites were chosen in 1988, each covering approximately 20,000 sq. km. These were in France (Centre), Germany (Bayern), Greece (Makedonia), Italy (Emilia-Romagna) and Spain (Castilla-León). The sites were selected to be representative of cropping and agricultural systems in the EC and to be in areas less affected by cloud cover. This was to maximise the chances of obtaining clear satellite images. An investigation of existing archives of satellite imagery suggested that acquisition would be easier in continental or Mediterranean climates.
In 1990, some of the original test sites of Action 1 were increased in size to include neighbouring administrative districts. In 1991, the original test sites were included except in Germany. In Spain and Italy, the area was extended and a new study area was started in Portugal. In 1992, a large part of Spain and Greece were covered. Additional sites were the whole of Belgium and three counties in the southern United Kingdom. Two non-EC sites, the whole of the Czech Republic and a region of Romania were also included.two additional. The same method has been also applied in test sites surveyed for FEOGA in south of Italy and north of Greece. The objective of these was to obtain area estimates for durum wheat to check farmers declarations.

Administrative evolution

There were two underlying principles which guided the administrative evolution of Action 1: transfer of technology and transfer of supervision to the appropriate government services of the member states. The capacity for implementation of Action 1 methodology was extremely variable across the EC at the beginning. So initially, the work was carried out by private contractors, closely supervised by JRC and with varying inputs from government services. The technical capacity to carry out the work at local level was increased by collaborating with local government services, locally based private contractors and internationally renown contracting companies. By 1992, the surveys were mostly carried out under the supervision of the local Ministries of Agriculture and they were moving from a semi-operational to a fully operational approach. Thus by 1992, responsibility for the test sites was as shown in Figure 1 and can be grouped in three categories:

- EC collaborative sites
- non-EC collaborative sites
- national sites

The last categories includes, the AGRIT project of the Italian Ministry of Agriculture. In that case, there was no direct participation of the MARS project rather an exchange of experience and points of view. In other cases, projects were partially financed by the project

Technical Approach

The estimation of crop areas with remote sensing was first carried out in the USA by the National Agricultural Statistics Service (NASS) of the USDA, in the 1970s. The application of the USDA methods was under investigation in Italy in 1985 and some results were available for assessment by the MARS project. Independent studies were commissioned to consider the USDA method and alternatives and the USDA was closely consulted about their experiences with remote sensing and the gathering of agricultural information. The outcome was the method outlined below, and in Figure 2, which was based on the USDA technique modified for European conditions.

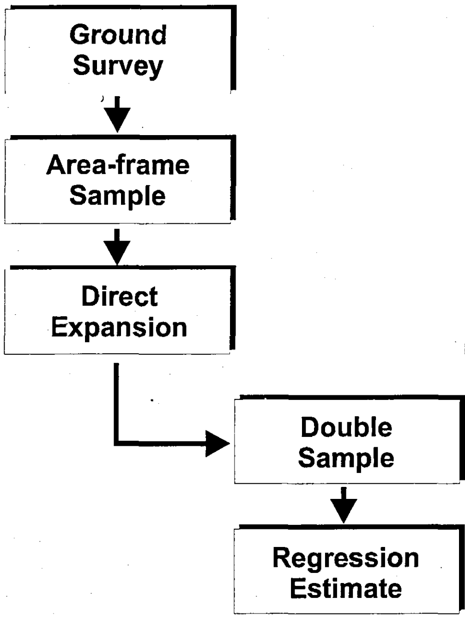


Figure 2: Technical approach for estimation of crop areas in Action 1

Initially, two separate area estimates are made. The first is by ground observations of crop areas in a random sample of fixed-size areas within the study region. The total area of each crop in the whole of the study region is then estimated by classical statistical methods. This methodology can be used alone as an agricultural information system. The second estimate is by digital classification of the satellite imagery. Satellite images comprise of measurements of surface reflectance properties made systematically over the landscape. Each measurement is an average value for a small square area (20 or 30 m square usually). With this technique, the whole of the region is classified by computer and the area can be calculated by counting the number of the small squares, known as pixels, in each class. The estimates by the first method suffer from high sampling errors. The classification method has no sampling error, however, area estimates are not accurate because significant numbers of pixels of the satellite image can be mis-classified. Finally, the third estimate uses the two sources of information together. The relationship between ground survey information and the satellite image classification is determined by regression methods. The so-called regression estimate which results is more accurate than either the ground survey or the image classification.

The improvement achieved by combining ground observations with satellite data can be calculated by a parameter known as the relative efficiency (RE). The RE can be used to estimate the additional size of ground survey sample needed to achieve an equivalent improvement in accuracy of area estimates. Thus, the value of applying remote sensing is the cost saving by reducing the ground sample for a desired level of accuracy in crop area estimation. The precision of area estimates from the ground survey and the regression can be measured by the Coefficient of Variation (CV). This can also be used to judge when there are serious discrepancies between Action 1 estimates and those from other sources.

Main findings

- **Crop area estimates from the ground survey.** The average precision of crop area estimates varied inversely with the crop area. For typical areas of main crops, the average precision's at regional, national and Community levels were 9.0, 3.3 and 1.2% respectively.

- **Improvement of crop areas estimates with remote sensing.** The improvement in precision of crop area estimates with remote sensing was variable but, on average, the improved precision's at regional national and Community levels were 5.7, 2.1 and 0.8%.

- **Cost-effectiveness of remote sensing.** The relative efficiency of remote sensing improved during the progress of Action 1. This was mainly because implementation improved with experience. By 1991, the overall average relative efficiency was 2.55. This was considerably better than thresholds commonly accepted for remote sensing to be cost-effective and those calculated in Action 1.

- **Satellite Imagery.** Cloud cover was not generally a problem at any of the study sites. There were a few organisational problems which were overcome. By the end of Action 1, coverage with high quality imagery in the specified time-windows was usually >90%. Delivery times were acceptable once a few organisational problems were overcome.

- **Timeliness.** There were no technical restrictions to cause delays unless images acquisition was a problem.

- **Yield survey:** The yield survey was a relatively small part of Action 1. There were problems with collection of yield data. The methodology to estimate crop yields with remote sensing requires more development.

- **Harmonisation of European agricultural information.** Action 1 methodology was readily adapted over a wide range of European conditions. Crop area stimates from Action 1 methods can be used to make comparisons with equivalent figures obtained with different methods used by member states and to detect serious discrepancies.

- **Technical guidelines.** The experience of Action 1 enabled a series of guidelines for quality control of future work to be identified. These are provided in PART 4.

- **Geopolitical considerations.** Action 1 stimulated considerable collaboration between statistical services, private consultants and research organisations within and between member states. The progress of work and results were disseminated in bulletins produced by JRC and at several international conferences organised by them.

1. Crop inventory by ground survey

The design of ground surveys using an area-frame sample is described. The statistical formulae for area calculations are presented Variations in the methodology applied and the field work procedures used during Action 1 are explained. The performance of the methodology is assessed.

1.1 General Statistical Methodology

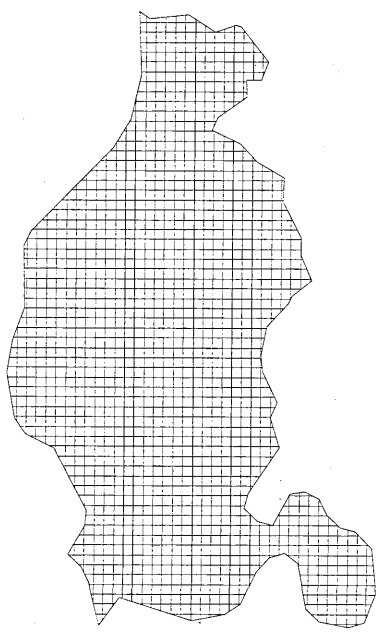


Figure 1.1: Squared grid overlaid on a small region.

1.1.1 Summary of approach at each study site

The site is completely sub-divided into a finite number of elementary areas referred to as segments. A random sample of the segments is drawn from the so-called area-frame and the area of each crop of interest is measured in each sample segment by field work. The approach for estimating crop areas for the site is by direct expansion of the field measurements. Crop yields are obtained for a sub-sample of segments and are used to estimate average yields. The site can be stratified to improve both area and yield estimates. The methodology is modified to account for geographical variation and to include preferences of individual Member States.

1.1.2 Sample selection

Figure 1.1 shows an area-frame of square segments overlaid onto the province of Varese, Italy which will be used to illustrate the recommended form of the methodology. To simplify the ground survey design, the boundary of the region is approximated to follow the grid, as shown in Figure 1.2, to avoid sub-division of segments. Thus, segments falling across boundaries are completely included if more than 50% of their area fell within the region and totally excluded otherwise.

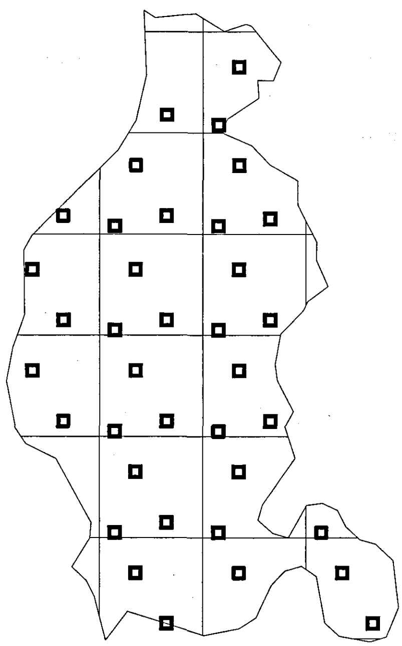


Figure 1.2: Systematic sample with a distance threshold

A systematic random sample with replication is generally used and designed as follows. The region is divided into blocks of segments (e.g. 10 x 10 segments), as shown in Figure 1.2, where the area-frame grid is not included for clarity. A sample segment is chosen randomly within a block. The systematic random sample is created by selecting segments at the equivalent locations in each of the blocks. The process is repeated to produce several replicates per block. Three replicates are shown in Figure 1.2. Some additional restrictions can be set on the distance between replicates to ensure a good distribution of the segments and to avoid any forbidden areas, e.g. edges of aerial photographs. Only the segments falling in the region are retained and surveyed.

N.B.: In the context of Action 1, the term random is always meant in the statistical sense. Sample locations are selected using unbiased techniques and not by ad hoc or haphazad selection.

1.1.3 Direct expansion estimates

The classical expansion formulae give an unbiased estimate for the area of land use, or for its proportion, y_C. It is better to compute the estimates as proportions rather than as absolute areas because this automatically takes account of errors resulting from small localised variations in the scale of segment maps and drawing or digitising errors. The estimate of the proportion of land area covered by class c is given by:

$$\bar{y}_c = \frac{1}{n}\sum_{i=1}^{n} y_i$$

with variance

$$Var(\bar{y}_c) = \left(1 - \frac{n}{N}\right)\frac{1}{n(n-1)}\sum_{i=1}^{n}(y_i - \bar{y}_c)^2$$

where: y_i is the proportion of sęgment i covered by class c, N is total number of segments in the region, n is number of segments in the sample. The proportion of the study region sampled $\left(\frac{n}{N}\right)$ is the sample fraction. When this is less than 5%, the correction factor for a finite population $\left(1 - \frac{n}{N}\right)$, can be omitted from the above formula (Cochran, 1977). The estimate of the class area is:

$$\hat{Z}_c = D\,\bar{y}_c$$

with variance $$Var(\hat{Z}_c) = D^2\,Var(\bar{y}_c)$$

where D is the area of the region.

1.1.4 Stratification

The aim of stratification is to sub-divide the region of interest into a number of non-overlapping sub-regions known as strata such that the variation of crop area per segment within each stratum is less than the variation between the strata. When stratification is carried out before sampling, the expansion formulae are applied to each stratum, *h*, as though it is a separate region. The area of a class and the variance in a stratum is thus:

$$\hat{Z}_h = D_h \bar{y}_h \qquad \text{and} \qquad Var(\hat{Z}_h) = D_h^2 Var(\bar{y}_h)$$

The total area of the class in the region is the sum of the area estimates in each stratum:

$$\hat{Z} = \sum_{h=1}^{H} \hat{Z}_h$$

with total variance equal to the sum of the variances of each stratum:

$$Var(\hat{Z}) = \sum_{h=1}^{H} Var(\hat{Z}_h) = \sum_{h=1}^{H} D_h^2 Var(\bar{y}_h)$$

1.1.5 Example of stratified sample design

Figure 1.3 shows the province of Vicenza in Italy divided into three strata based on landforms, which influenced the intensity of agriculture.

- plains: intensive agriculture
- hills: medium intensity agriculture
- mountains: low intensity agriculture

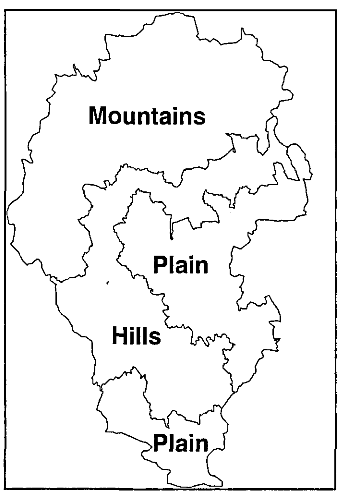


Figure 1.3: Stratification in Vicenza.

The efficiency of the survey is increased by varying the sampling rate according to the intensity of agriculture. A higher sampling rate is used for strata where agriculture is predominant and lower rates where it is marginal. The sample selection proceeds as before by overlaying the region with a square grid. The strata boundaries are approximated to the nearest segment applying the same rule as for the region boundary. The region is divided into blocks of segments and the number of replicates per block is chosen to provide the highest sampling rate needed. Lower sampling rates are achieved by eliminating replicates at random, as required. Blocks can cross several strata thus the replicates for that block can be in different strata. Replicates are kept in the sample if they correspond to the ones chosen for the strata covered by the block.

Figure 1.4 shows the final sample for the Province of Vicenza. In the example, the segment size was 500m x 500m and the block size was 10 km x 10 km giving 400 segments per block. The highest sampling rate was 1%, for intensive agriculture on the plains, requiring 4 replicates. The sampling rates for medium and low intensity agriculture were 0.5% and 0.25% for which 2 and 1 replicates respectively were chosen.

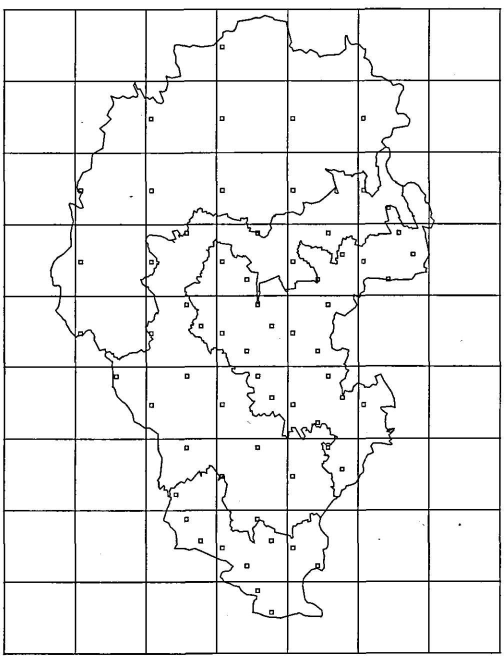


Figure 1.4: Final sample of segments in the Province of Vicenza.

1.1.6 Efficiency of stratification

The quality of a stratification is assessed by the relative efficiency which is the ratio between the variance of the estimates with and without stratification, given by:

$$\eta_{strat} = \frac{Var(\hat{Z}_{ran})}{Var(\hat{Z}_{strat})}$$

However, the same reduction in variance could also be achieved by increasing the size of the ground survey sample such that:

$$\eta_{strat} = \frac{\frac{1}{n}Var(y)}{\frac{1}{n_1}Var(y)} = \frac{n_1}{n}$$

where n_1 is the increased sample size. The stratification will be economical if its relative efficiency exceeds a threshold given by:

$$n(\eta_{strat} - 1)p > \tfrac{s}{t}$$

where n is the sample size, p is the cost of surveying additional segments, s is the cost of the stratification and t the number of years that the stratification is likely to be used.

1.2 Application and variations

1.2.1 Survey design

In Action 1, the area sampling frames were usually regular square grids. The exception was in Italy where the area frame was changed in 1990, to consist of segments which were irregular in shape and size. These were individually delineated using "natural" boundaries, such as roads, rivers and permanent field limits.

The aligned systematic random sample design with replication was normally used. A alternative approach, used in 1992 in the United Kingdom, was to select random locations separately for each block which produces an unaligned systematic random sample.

The recommended segment size was 49 ha. This required the production of a special grid (700 m x 700 m) for the sampling frame and the systematic sample was selected in a 11.2 km x 11.2 km block with four replicates at full sample rate. In many cases, the sampling frame was based on an existing map grid for convenience of application. In Spain, where the study area was greatly increased, the 49 ha segment size was retained, using part of the 1km map grid square, with three replicates selected within 10 km x 10 km blocks. A 100 ha (1 km square) segment was used in England and the south of Italy (FEOGA project). In Belgium, a 25 ha segment was used; selected within a 20 km x 20 km block with 11 to 22 replicates. In the Czech Republic, where very big fields were encountered, 400 ha segments (2km x 2km) were selected from a 40 km x 40 km block with 2 to 12 replicates. A 36 ha segment (600 m x 600 m), based on the existing national sampling system, was used in some areas in France.

There were various approaches employed to stratify the regions of interest. These may be summarised as follows:

- Identification and delineation of sub-areas with different agricultural systems by photo interpretation of geometrically corrected satellite imagery
- Use of existing statistical reporting sub-regions for which previous agricultural information was available. These were of variable size, generally less than 100 sq. km in area. Strata were formed by grouping the sub-regions according to the proportion of the area in agriculture or in the crops of interest.
- Use of an existing classification of agricultural regions
- Combination of existing thematic maps such as land-use, geology and soils within a GIS

After the first year, the stratifications were generally modified slightly and the number of segments sampled within each stratum was optimised for the main crops of interest by using the Neyman allocation. This assumes the unit cost of surveying segments is the same for all strata, and then the variance of the estimate, using a stratified sample, is minimised when the number of segments sampled in each stratum is proportional to $N_h Var_h(y)$.

The optimum allocation was generally different for each of the main crops of interest and a compromise allocation was usually made. When users believed it was desirable to keep the same sample of segments for several years and to be able to modify stratum boundaries adjacent strata were given the same sampling intensity.

1.2.2 Field work

The field work consisted of visiting each segment to produce an accurate map of crops and other land cover for all the field parcels contained within it. In some cases observations for crop yield estimation were also made. The enumerators for the ground surveys were usually students in agriculture (or a closely related field), employees from government agricultural and statistical services or, employees of the private consultancy companies contracted to do the work. The number of enumerators employed and the duration of the survey varied considerably. Ten to thirty segments were surveyed per enumerator during a two to four week period. In some cases, two ground surveys were needed in order to cover both winter/spring and summer crops. However, the second survey was usually achieved more rapidly because it was more an update rather than a full survey. Enumerators were usually given several days training prior to the field survey.

Several documents were prepared for the field work. These were:

- a photographic image of the segment and the immediate surroundings enlarged to 1:10,000 or 1:5,000 scale.
- a transparent overlay on which to draw the field boundaries
- the 1:25,000 or 1:50,000 map of the area surveyed by the enumerator with the segment locations delineated.
- the 1:10,000 map of the segment and its surroundings when available.
- a form to record the field codes and a legend.
- a list of the codes - based on the CRONOS codes used by EUROSTAT for crops and other land use classes
- a list of class definitions used by EUROSTAT
- a field manual describing correct procedures and guidelines for the field work

The photographic images of segments were produced from archived aerial photographs, orthophotos or satellite images. In extreme cases , as in Greece, cadastral maps of segments had to be specially produced for the field survey.

The field manual was translated into the national language of the enumerators and provided them with definitive step-by-step instructions for accurately surveying the segments. Details of the field manual and examples of field survey documents are given by Perdigão (1991). A summary of the instructions is as follows:

- *Finding the segment:* Use the 1:25,000 or 1:50,000 scale map to plan the most efficient route to the allocated segments. Use relevant landmarks shown on the map to assist the location of the segment.
- *Itinerary in the segment:* Examine the photographic images and/or 1:10,000/1:5,000 scale maps of the segment closely to determine the most efficient route to see all the field parcels, before starting. Paths and roads which are easily identified on the images and maps should be preferred.
- *Changes in field boundaries:* In cases when there is a disagreement between the parcel boundaries on the ground and on the field documents, the new boundary locations should be measured on the ground and drawn onto the documents.
- *Drawing field boundaries on the transparent overlay:* Areas less than 20m in width should not generally be mapped. Features such as roads, paths, hedges, railways, rivers and streams should be mapped as single lines unless they are more than 20m wide. Take care to ensure that lines representing the field boundaries join up properly.

As Action 1 progressed, responsibility for the ground survey was often transferred to the appropriate government service with the rest of the work done by the consultancy company. In principal, the dates of the survey were chosen to enable the maximum differentiation of the crops encountered consistent with time constraints. However, crop information was not always available to verify that this was achieved.

1.2.3 Data processing

The main task was to measure the area of the field parcels from the segment maps and to produce a computer database file containing the area and cover type for each field parcel. In the beginning of Action 1, the recommended method for measuring the areas of each parcel was by counting squares in 2 mm x 2 mm grid overlaid on the segment. This proved to be impractical and a special video digitising procedure was developed for the MARS project by JRC-Ispra. Application of this method required that enumerators used a specified format for the production of the segment maps, using black ink to draw parcel boundaries and red or green ink to label them. The segment maps were then scanned by a video camera with a red or green filter thus a digital file was produced, containing only the parcel boundary information. The areas of each parcel were then extracted using a series of digital processing techniques. The method is described fully by Stakenborg (1989) and allowed up to fifteen segments to be processed per hour. Other commercially available digitising hardware and software was used by some contractors. Summary results were produced for each segment, giving the total area of each crop. These were usually computed as a proportion of the segment area to reduce the effects of errors in digitising and map scale.

1.3 Results and Conclusions

1.3.1 Accuracy of area estimates

The Coefficient of Variation (CV) is a measure of the accuracy of unbiased area estimates from the ground survey and is given by:

$$CV\% = \frac{\sqrt{Var(\bar{y})}}{\bar{y}}.100$$

Results from the whole of Action 1 were pooled to produce Figure 1.5 which shows the relationship between the CV and the total estimated area of the crop in each study region. The CV decreased with increasing size of the crop area because the standard error, $SE = \sqrt{Var(\bar{y})}$ decreased. The main factors reducing the standard error, are an increase in the sample size and in particular, the number of non-zero observations.

An increase in the size of the sample fraction will also reduce the SE if the segment size is constant. However, no attempt was made to assess the effect of the different sampling fractions because these only varied between 1 and 1.5%.

Main crops typically cover areas of around 100,000, 1,000,000 and 10,000,000 ha at regional, national and European levels. Estimates of the corresponding average CVs, from Figure 1.5, are 9, 3.3 and 1.2% respectively and are therefore estimates of the typical accuracy of crop area estimates at those levels.

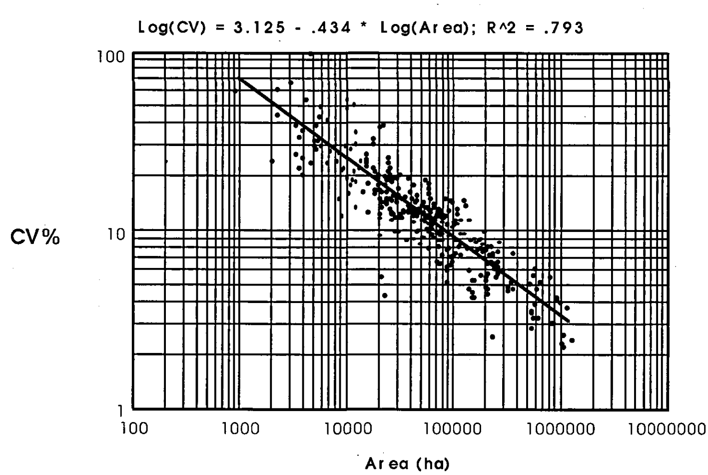


Figure 1.5: Relationship of Coefficient of Variation (CV %) to area of crop (ha) for ground surveys with sampling rates from 1 to 1.5 % in regional crop inventories in Europe

1.3.2 Timeliness of results

The general requirement in Action 1 was that area estimates for winter and spring crops from the ground survey should be available in early to mid June and in September for summer crops. The timing of ground surveys was governed by the crop cycles and was generally in the mid to late May period for spring and summer crops and in July for summer crops. The time for analysis of ground survey results for winter and spring crops was therefore quite short. However, most contractors were able to complete the analytical phase to meet the deadline. There were no real technical problems limiting the timely delivery of area estimates from the ground survey.

1.3.3 Efficiency of stratification

The relative efficiency of stratification η_{strat}, was calculated for each crop in each test site and for each year. The results were very variable and to simplify assessment of the benefits of stratification in Action 1, a weighted average value was computed with the weighting according to the proportion of the crop, *p* in each study site, as follows:

$$\eta_{weighted} = \frac{\sum p\eta_{strat}}{\sum p}$$

Cases with crops covering less than 1 % of the study area not included and the indicator was computed:

- by crop across all test-sites for each year
- by test-site across the crops for each year
- overall for each year across all crops and test-sites

The weighted average relative efficiency of stratification for main crops was generally around 1.5 as shown in Table 1.1. Only the figures for 1988-90 are presented because from 1991 not all the data was available. The efficiency of stratification was higher for the crops covering largest areas, such as wheat and barley. Relative efficiencies for crops such as maize oil seeds and dried pulses were below 1.30. The Neyman optimum allocation algorithm was applied in 1989 and did not seem to have brought any improvement in relative efficiency for any of the main crops.

Table 1.1 : Weighted efficiencies of stratification for main crops, across all study areas

	1988	1989	1990
Wheat	1.45	1.61	1.48
Barley	1.53	1.47	1.67
Maize	1.24	1.20	1.27
Oil seeds	1.06	1.15	1.08
Dried pulses	1.15	1.19	1.12

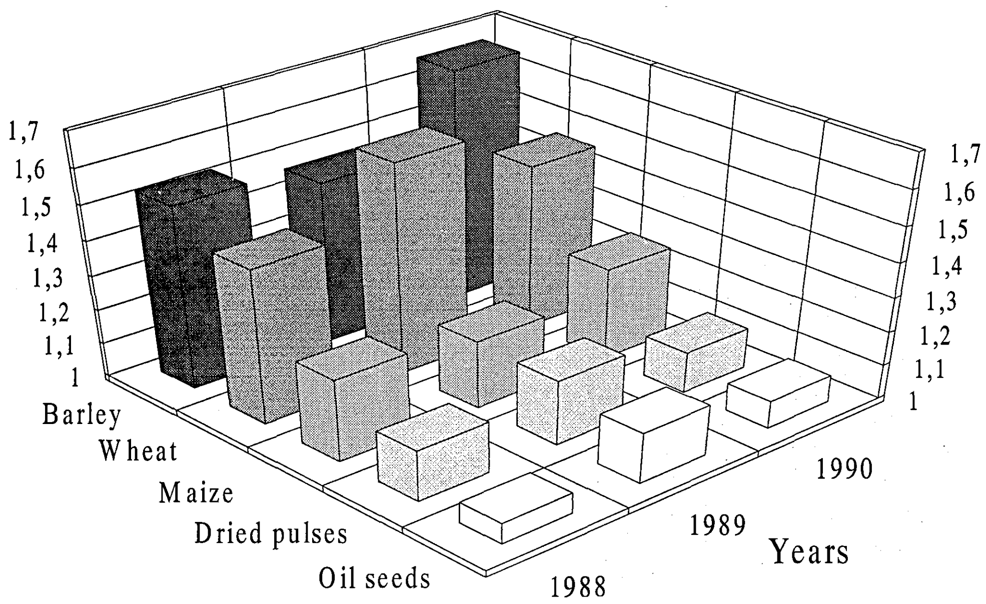
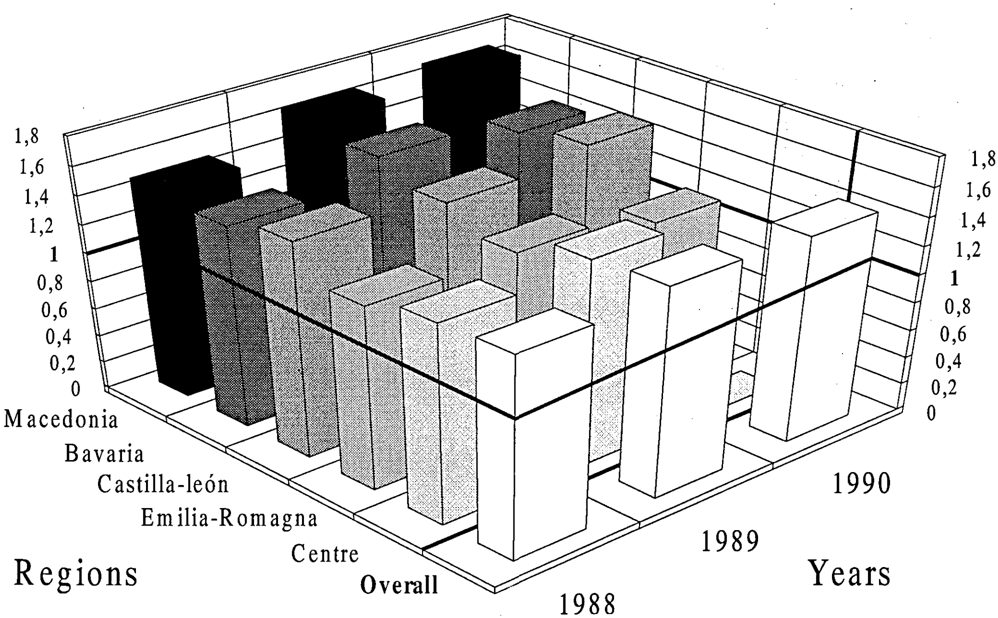


Table 1.2: Weighted efficiencies of stratification for study areas, across all crops

	1988	**1989**	**1990**
Centre	1.41	1.46	Abandoned
Bavaria	1.44	1.59	1.46
Emilia-Romagna	1.31	1.31	1.18
Castilla-león	1.53	1.46	1.54
Macedonia	1.49	1.67	1.70
Overall	1.43	1.49	1.47



The weighted averages for relative efficiency of stratification for each study area, shown in Table 1.2, are rather low and generally around 1.5. In Centre (France) stratification was abandoned in 1990. The effect of the Neyman optimum allocation algorithm from 1989 onwards is not apparent. The only site where there seems to be an improvement is Macedonia. The sub-division of Macedonia into strata was based on clear geographical distinction between mountains and areas of irrigated and non-irrigated agriculture which had different crop intensities and was carried out by photo interpretation of satellite imagery. In some other regions strata were defined by grouping existing small administrative sub-regions according to the relative proportions of the main crops of interest within them. Possibly, the crops were not homogeneously distributed within the administrative units. Thus these groupings were not a correct reflection of the geographical distribution of land with different cropping intensities and would account for the poor relative efficiencies.

In general, stratification was most successful when different agricultural systems could be delineated as large units on satellite imagery by photo interpretation.

1.3.4 Yield estimates

Yield surveys were carried out after 1989. In 1989, fields for the yield survey were usually selected from a sub-sample of the segments (e.g. by choosing one repetition). Then, a square grid with a pre-defined number of points was used to select the fields. From 1990 the entire farm associated with a sample field was surveyed. In the farm survey, the sample was made with replacement that is if two sample points fell on the same farm, the farm was drawn twice. Generally, about 100 observations were sampled for each crop of interest. In the farm survey, the number of observations was around 400. There were several problems with the yield survey, which relied on the farmer to tell the yield. The first was that the farmer could not always be identified. Secondly, among the identified farmers, some were not willing to provide any information. Quite often farmers were rather reluctant to answer the questions they were asked. In Castilla-león for example, the enumerators were asked to comment about the reliability of the answers and believed that only 50 % of them were to be considered reliable. Thus, bias could be introduced by inaccurate farmer estimates of yield per unit area and the incomplete sample. Another problem for yield estimation in Europe is the variation in the stage of growth at the time estimates were required. In May, the cereals were near to harvest in Spain whereas in UK they were in still early growth stages. A relatively small proportion of Action 1 effort and resources was devoted to yield estimation and so more development of the methodology is needed.

CONCLUSIONS:

- **Precision of area estimates by ground survey increased with area of crop**
- **The typical accuracy of area estimates of main crops, with sampling fraction of about 1.5%, is 9.0, 3.3 and 1.2% at regional, national and European levels respectively**
- **The method was generally applicable**
- **Results were generally available at the specified time for management of the EU agricultural policy**
- **Overall, the relative efficiency of stratification was rather low, generally less than 1.5**
- **The relative efficiency of stratification was highest for crops covering the largest areas**
- **The efficiency was higher when stratification was based on large geographical units photo interpreted from satellite images**
- **The methodology for yield surveying needs more development**

2. Satellite Remote Sensing

Properties of SPOT and Landsat TM images, used in Action 1, are described. The strategy for image acquisition is reported. Fundamental concepts and techniques of remote sensing methods are explained. Methods of accuracy assessment for geometric correction and digital classification are explained. Variations in the implementation of remote sensing methods in Action 1 are discussed. The accuracy of geometric correction and of digital classification of crops is reported. Methodology for yield estimation is discussed.

2.1 Achievement of satellite image coverage

2.1.1 Satellite images used for regional inventories

Two types of satellite images were used in Action 1. The Landsat Thematic Mapper (TM) on board Landsats 4 and 5, has a spatial resolution of 30m in six waveband ranges within the visible, near infrared and middle infrared parts of the electromagnetic spectrum and 120m in a single waveband in the thermal infrared waveband. The satellite repeat cycle is 16 days. A full size TM scene covers an area 185 by 185 km. The SPOT HRV scanner is on board the SPOT 1 and SPOT 2 satellites (and now also SPOT 3). It produces two types of image: XS and PAN. SPOT XS has three bands in the visible and near infrared wavebands with 20m spatial resolution. SPOT PAN has a single band with 10m spatial resolution. The satellite repeat cycle is 26 days but it has the capability to point the sensors to either side of the orbit path and is thus able to get repeated coverage on selected sites every 4 or 5 days depending on the latitude. There are two identical sensors on the satellite. The full scene size is 60 by 60 km. When the sensors are operated in parallel this gives a full coverage of 110 km across the satellite track, there being about 10km overlap.

The predictable and repeated orbital paths followed by Landsat and SPOT make it possible to produce maps showing the location of the scene centres with a world referencing system of row and path numbers which can be used to specify the required area to be covered by the image. The actual image locations are generally accurate to within 10km of the nominal scene positions. More recently it has become possible for users to specify their own scene centre locations and the along-track dimensions of the image. It is also possible to buy parts of a scene, usually a quadrant. These developments make it possible to cover specific areas with less unusable imagery.

Experience has shown that the discriminability of crops on satellite images varies with time during the crop cycle. It may be necessary to use several images of the same area, taken at different times during the season, to achieve sufficient accuracy of crop identification. The timing of the overpasses for any specific site is known from the orbital parameters of the satellites. It is therefore possible to request images from those overpasses which take place during the desired time period. With SPOT, the frequency of coverage can be increased by requesting that the satellite be programmed to use its capability to point the sensors.

The main cause of failure to acquire images of suitable quality is atmospheric interference. This is because the satellites which are currently available cannot 'see' through cloud, haze or dust in the atmosphere. An assessment of the cloud cover can be made by viewing so-called quick-looks which are small scale pictures of the image often available in a catalogue or from a numerical code giving the cloud cover in 'tenths' over several sub-parts of the image (0 = cloud free, 9 = complete cloud cover). For example, when the image sub-parts are quadrants, 0000 means that the image is essentially cloud free in all quadrants, 9999 means it is essentially completely cloud covered. The order of the digits indicates which quadrant and is as follows: NE,NW,SE,SW. Satellites which can penetrate cloud cover use RADAR sensors such as that on board the ERS 1 satellite. The use of RADAR images for crop discrimination is still experimental and is therefore not considered here.

All satellite images undergo a certain amount of processing before they are sold to users. This is usually to carry out corrections to the data that are applicable to all images from a particular sensor and these are therefore designated as 'system' corrections. It is also possible to purchase images at a higher price, for which certain commonly required types of processing have been applied, such as geometric correction. It is important to determine if these additional levels of processing are really beneficial for regional inventory projects. A commonly held view is that they are not and that purchase of basic system corrected products is most cost-effective.

2.1.2 Acquisition of satellite images

Satellite images for all study areas were ordered and purchased centrally from the vendors. In the first year of the programme, both TM and SPOT imagery were purchased in order to evaluate which of the two high resolution sensors available on the market was the most suitable for Action 1. The perceived advantages of TM were the increased spectral information from the middle infrared wavebands and the larger area of coverage by a single image (34,000 km^2 for TM against 3,600 km^2 for SPOT). The advantages of SPOT were thought to be the improved spatial resolution (20m for SPOT against 30m for TM) and the increased frequency of coverage because of the ability to point the satellite. However, the priority was given to TM because it was thought to be easier to use as it required less scenes to cover the study area. Nevertheless, it was possible to substantially reduce the number of SPOT scenes (up to 30% less) by programming the acquisition of precisely located twin strips instead of standard scenes. Images from the same path, date and incident angle from both HRV sensors could be then easily be joined together. In fact, in 1990, in the Centre region of France, only six of these SPOT image groups, generated from 26 SPOT scenes, were needed to cover the whole study area (50,000km^2).

Later, the first suitable scene from either sensor was purchased on a competitive basis. Time windows for acquisition were specified to meet the requirements for timeliness of results and to coincide as far as possible with periods of optimum crop discrimination. The latter were defined for each study area by reference to the local cycle of crop development - the crop calendar. If no suitable images could be acquired in the optimum window, then images from a wider time period were considered. If no images were acquired within the extended time window, then image cover was abandoned.

2.2 Remote sensing methodology

2.2.1 Fundamental concepts

Satellite images are created by scanners which systematically measure electromagnetic energy (reflected solar or emitted terrestrial) from small areas of earth's surface, to produce an array of observations which make up the image. The small area from which individual observations are made is known as a picture element or "pixel" and its projected size on the ground determines the spatial resolution of the image. The satellite images used in Action 1 were mainly produced from measurements of reflected solar energy. Scanners make measurements in a number of waveband ranges simultaneously because the proportion of solar energy reflected from the earth's surface varies with wavelength depending on the surface properties. Satellite images are recorded in a digital format so that for every pixel there is a number representing the magnitude of the energy measured by the scanner, for each waveband. Atmospheric properties, sun angle and sensor calibration also influence the magnitude of the satellite observations. Since these and land cover properties change with time, the relationship between land cover types and the digital measurements has to be determined for each image. In practice, this is done empirically with reference to sample observations over areas of known cover type and map location. Geometric correction of the satellite images is necessary to facilitate this by bringing the images to a known scale and map projection, which is also needed for measuring areas and sub-dividing the image into administrative regions or strata. Digital classification of the satellite image is carried out by applying mathematical rules to the whole image which were determined from the double sample of satellite and ground observations. The classification results are used to produce a new image with all pixels of the same class given the same numerical value. These numerical values are then used to produce a thematic map of the area, each class being coded in a different colour. The computer system also counts the number of pixels assigned to each class and the totals are calculated. The pixel area is known for the sensor used to produce the image, therefore the area of each class can be calculated.

Unfortunately, the spectral information from satellite images is not totally reliable for distinguishing land cover categories and this complicates the operational use of remote sensing for regional inventories. There are two main causes. The first is because different cover types may have very similar reflectance patterns. This often happens when crops are closely related botanically. In such cases it may be difficult or even impossible to distinguish them by remote sensing. Other cover types may be completely different in terms of land use. e.g. bare soil surfaces and urban areas. But they may, as in this case, be spectrally similar because they are derived from similar base materials. In this case soils and construction materials frequently contain similar rock minerals. It is therefore fundamental to measure this so-called spectral confusion and to take account of it in estimating crop areas. The second cause is because the pixels are large enough to cover a mixture of underlying land cover classes. This frequently happens along field parcel boundaries when a pixel of 20 or 30 m dimension can include the cover types on both sides of the boundary plus the influence of the boundary itself. These so-called mixed pixels are usually mis-classified and can be a large proportion of the total if the land cover comprises of parcels near to the same size as the pixel.

2.2.2 Technical principles of geometric correction

Distortion in satellite images arises from several sources: the sensor; the changing attitude, height and speed of travel of the satellite; the angle of the orbit path in relation to north; and, the rotation of the earth under the satellite.

The first aim of geometric correction is to determine the mathematical relationship between the co-ordinates of each pixel, in the original image and the map co-ordinates. This is in the form of a pair of polynomial transform equations such as the following:

$$x = a_0 + a_1X + a_2Y + a_3X^2 + a_4XY + a_5Y^2 \ldots$$

$$y = b_0 + b_1X + b_2Y + b_3X^2 + b_4XY + b_5Y^2 \ldots$$

where x and y are the co-ordinates of the pixel in the original image and X and Y are the map co-ordinates. The order of the polynomials may be changed by adding or reducing the number of terms. The coefficients (a's and b's) are determined by the method of least squares. To do this, the map references for several clearly visible points in the image known as ground control points (GCP's) are measured.

The transform equations are then used to extract the correct pixel digital values from the original image for each location on a map grid, to form the corrected image. It is assumed that the pixel value determined at each grid position is the same for the whole pixel centred over the grid location and that the grid size is the pixel dimension for the corrected image. When the transform equations are used to locate the pixel value for a particular geo-referenced point, they usually indicate a location that falls between the pixel locations in the original image. It is therefore necessary to interpolate a value from the surrounding pixel values. There are several ways that this is done. The methods differ in complexity and this influences the processing time and hence the cost. They also influence the properties of the geometrically corrected image and thus are appropriate for different purposes. The simplest way of selecting the appropriate pixel value is to take the value from the one nearest to the point indicated by the transform equations. This is known as nearest neighbour re-sampling. Bi-linear and bi-cubic convolution are also commonly used and these methods involve interpolation with the nearest 4 and 16 pixels respectively.

2.2.3 Estimating the accuracy of geometric transformation

Once the coefficients of the transform equations have been determined, the accuracy of the transform is estimated by calculating the RMS error for the positioning of the ground control points. For each control point, we can calculate the position predicted from the transform equations:

$$\hat{x}_i = a_0 + a_1X_i + a_2Y_i + a_3X_i^2 + a_4X_iY_i + a_5Y_i^2 \ldots$$

$$\hat{y}_i = b_0 + b_1X_i + b_2Y_i + b_3X_i^2 + b_4X_iY_i + b_5Y_i^2 \ldots$$

The RMS errors in the x and y direction are:

$$RMS_x = \sqrt{\frac{\sum_{i=1}^{n}(x_i - \hat{x}_i)^2}{n-k}}; \quad RMS_y = \sqrt{\frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)^2}{n-k}}$$

where k is the number of terms in the transform equation. The absolute error, ε is

$$\varepsilon = \sqrt{RMS_x^2 + RMS_y^2}$$

When the errors for individual ground control points are greater than a chosen threshold most versions of geometric correction software will reject those outliers and then re-calculate the transform equation coefficients, thus applying some automated filtering out of erroneous control points. If the number of control points is large relative to the order of the transform equations and they are randomly and evenly distributed across the area being corrected, the RMS error is a good estimate of the actual error of geometric correction. In Action 1, the aim was to keep the error of geometric correction to less than the size of one satellite image pixel.

2.2.4 Technical principles of digital classification

A mathematical procedure known as discriminant analysis is used to classify the pixels by computer using the numerical values collected by the scanner. At each pixel location there is a numerical value for each waveband which are elements of a so-called measurement vector, $\mathbf{x}$ which is used to evaluate a mathematical discriminant function, $g_i(\mathbf{x})$ for each land cover class, ω_i. The discriminant function returning the highest numerical value identifies the class to which the pixel is allocated. In mathematical terms, the pixel at $\mathbf{x}$ belongs to class ω_i of M classes if

$$g_i(\mathbf{x}) > g_j(\mathbf{x}) \quad \textit{for all } j \neq i \textit{ and } i = 1, \ldots M$$

Various form of discriminant function have been used for classification of satellite images. In Action 1, the maximum likelihood method was used and the discriminant functions were of the form

$$g_i(\mathbf{x}) = \ln p(\mathbf{x}|\omega_i) + \ln p(\omega_i)$$

where ln is the natural logarithm, $p(\mathbf{x}|\omega_i)$ is the probability that the pixel at $\mathbf{x}$ belongs to class ω_i and $p(\omega_i)$ is the *a priori* probability of the class existence in the image.

The probability of finding a pixel belonging to class ω_i at position $\mathbf{x}$, is given by

$$p(\mathbf{x}|\omega_i) = (2\pi)^{-N/2}|\Sigma_i|^{-1/2}\exp\{-\tfrac{1}{2}(\mathbf{x}-\mathbf{m}_i)^T\Sigma_i^{-1}(\mathbf{x}-\mathbf{m}_i)\}$$

if all the satellite measurements of a given land cover class are Normally distributed. Each class can then be characterised as a multivariate distribution by estimates of the means of satellite observations in each of N wavebands - the so-called mean vector, $\mathbf{m}_i$, and the variance-covariance matrix, Σ_i.

The *a priori* probability, $p(\omega_i)$ of each class being in the image may be estimated as the proportion of the region of interest known to be covered by the class from independent previous sources of data. The alternative is to give each class equal *a priori* weightings equal to 1/M. The effect of unequal weightings compared to the case of equal weighting is to increase the number of pixels classified for the more common classes and reduce the number classified in the rare classes. Opinions differ as to the comparative benefits of unequal *a priori* weighting and both strategies were used in Action 1.

The discriminant functions require calibration prior to the classification process. Specifically, the mean vectors and variance-covariance matrices for each class must be estimated. Typical examples of each class are identified in a ground survey and pixel values are extracted from the equivalent locations on the geometrically corrected digital image by visual interpretation or map reference. This procedure is known as supervised training of the classifier when, as was most frequently done in Action 1, the operator controls the selection of the training data. Alternative procedures exist where the training is carried out automatically by determining the natural spectral groupings of pixels in the image. This process is known as clustering and in the context of digital classification-unsupervised training and was frequently used to sub-divide land cover classes which had more than one set of spectral properties.

The quality of the classification depends on the size of the training sample and how well it represents each of the classes. Textbooks frequently recommend that the minimum sample of pixels required to represent each class varies between 10 and 100 times the number of wavebands used in the classification. However, representativity of the sample is extremely important and the sample should be extracted so that the full spatial variation of each class is included. The application of the maximum likelihood method assumes that the class spectral properties are normally distributed but in fact, the method is robust to considerable departures from this assumption. However, inspection of the histograms of each class sample, showing the frequency distribution of pixel values, is necessary to check if distributions are multi-modal. This can occur when the same crop exists in the region at different growth stages or on different soil types. In such cases, the class must be sub-divided.

The maximum likelihood method will classify every pixel because it is based on continuous probability functions. There will be cases where the probability of a pixel belonging to any of the classes included in the discriminant analysis is so low that it is not reasonable to assign it to any class. A threshold value is chosen and below this, pixels are assigned to an "unclassified" category. Thresholding improves the precision of the classification but if it is set too high, the amount of unclassified data will become unacceptable. Thresholds equivalent to $p<0.01$ and $p<0.05$ were usually used in Action 1.

2.2.5 Assessing the accuracy of digital classification

A confusion matrix is a contingency table for comparing the results of one classification with another. In the present context, one classification is by the ground survey and the other is the digital classification based on reflectance properties. A typical example is shown in Table 2.1.

Table 2.1: Example confusion matrix comparing image classification with ground data in the UK, 1992

		Reference Data									
		Woods	Inland Water	Urban	Wheat	Barley	Summer Crops	Grasses	OSR	Other	TOTAL	User Accuracy
Image Data	Woods	15	3	1	2	1		2			24	63%
	Inland Water		4		1						5	80%
	Urban			11			1			1	13	85%
	Wheat	2	1	2	155	8	1	3		1	173	90%
	Barley	8		4	18	16	3	17		3	69	23%
	Summer Crops		1	13		3	43	3	1	5	69	63%
	Grasses			8	6	7	10	37		1	69	54%
	OSR*				1				30		31	97%
	Other	2		12	1	3	11	25		16	70	23%
	TOTAL	27	9	51	184	38	69	87	31	27	523	
	Producer Accuracy	56%	44%	22%	84%	42%	63%	42%	97%	59%	Overall Accuracy 63%	

* Oilseed Rape

The matrix is generated by sampling, at random, a series of equivalent spatial locations for which ground data and digital classification data are available. The elements of the confusion matrix are the number of occurrences a ground data class was assigned to a particular image class, within the sample. Thus in the example, 155 of the 523 sample points were classified as wheat by both ground observation and the image classification, and 18 points were wheat by the ground observation but were classified as barley on the image. Thus, the sum of the diagonal elements expressed as a percentage of the total number of observations represents the overall agreement between ground observations and the digital classification. The off-diagonal elements represent disagreements and the position of the element in the matrix identifies the type. Off-diagonal row elements are errors of commission. They represent the mis-classification of other ground classes which are included in the image class. The diagonal element expressed as a percentage of the row total gives the accuracy a user of the classification can expect for that class and hence is known as the user or mapping accuracy of the class. The off-diagonal column elements are errors of omission. They represent the mis-classification of a ground class into other image classes. The diagonal element expressed as a percentage of the column total gives the accuracy that the producer of the classification has achieved for the class. If the confusion matrix is not based on a random sample of points, then the user and producer acuracies will be biased estimates. In this context, the crucial property of the random sample is that each element in the confusion matrix is correctly weighted in proportion to the area it represents.

2.3 Application and results

2.3.1 Geometric correction

No particular method was imposed to the contractors. Usually the scenes from same date, sensor, path and viewing angle (for SPOT) were merged prior to any geometric corrections. In relatively flat terrain, a first order polynomial transformation was applied using at least 15 well-distributed GCP's giving an RMS error of less than one pixel. Sometimes second order transform equations were needed for SPOT scenes with a high of-nadir viewing angle. However, this high precision of geometric correction was not always possible in mountainous zones such as in Greece, especially in 1988 when only the 1:200,000 map was available to locate control points. In this case, the RMS error for digitising control points can, on its own be 50m to 100m. A digital elevation model DEM was used in Catalunya and the south of Italy, to improve accuracy but it is difficult to say if the additional cost was worthwhile because the improvement was in the mountainous areas where agriculture was known to be marginal.

Some contractors checked the geometric accuracy with an independent set of control points. Overall, the location error was generally less than two pixels in relatively flat areas and up to three pixels in mountainous areas.

The resampling procedure employed was usually nearest neighbour but bilinear interpolation and bicubic convolution were also used. It is usually believed that nearest neighbour should be preferred when doing a classification because the original DN's are kept. However, in Action 1, good final results were obtained with any of the resampling methods.

2.3.2 Classification of satellite imagery

After geometric correction, hardly any other pre-processing was applied to the imagery. Exceptions were in Germany, in 1990, where improved classification accuracy was obtained using an edge-preserving smoothing filter. However, this filter was also tested in Spain by another contractor with no significant improvement.

Two methods were used to extract training data mainly according to the type of software used. The first method consisted of manually digitising only the training fields on the screen. The operator selected fields of an appropriate size (usually > 3ha) with the aid of a database, where all the segment information was held, then digitised the training fields taking care of selecting only the purest pixels, away from field boundaries. In the second method, the whole segments were digitised, then parcels were selected from the database and pure pixels were selected using a process which automatically removed boundary pixels. This process was mainly implemented when using PEDITOR. In one case, in Greece, an additional survey was carried out to collect a separate dataset for the training. Otherwise, the training data was always a subset of the ground survey segments. Then, the training data for each class was extracted and clustering was either automatically or interactively performed to sub-divide the class if necessary so that their statistical distributions were approximately Normal.

Areas corresponding to the ground survey segments were usually extracted from the image and placed adjacent to one another into a single file to facilitate training and running of test classifications. A single file containing the ground survey data was also produced after segments were digitised and converted to raster format. The image segments were classified using a maximum likelihood algorithm usually with area-weighted a priori probabilities with the weights determined from the ground survey results.

The next step was to classify the image. The regions of interest in Action 1 were statistical reporting regions usually around 20,000 sq. km in size. Total coverage of these generally required several satellite images acquired on different dates under different conditions of illumination and sometimes by different sensors. Under these circumstances, the classification process was done separately for each stratum within each image.

The methodology described above was applied in most cases. However, in France, another method was applied in the first years, it consisted of a combination of supervised and unsupervised classification with a "box classifier". Each "box" was refined by applying different thresholds. This method had the advantage of being able to control each step of the process but was very cumbersome to implement and was limited to the processing of only three bands. It was replaced with PEDITOR of which the main asset was a perfect adaptation to the methodology and automation, although it was not user friendly.

2.3.3 Classification accuracy assessment

In Action 1, confusion matrices were used to estimate the 'accuracy' of a classification compared to ground data. The term 'accuracy ' implies that the ground observations are themselves 100% accurate and therefore represent truth. This was not always the case although the assumption was made that the ground observations were sufficiently accurate for the disagreements in the confusion matrix to be ascribed to errors in the computer classification.

If a confusion matrix is based on a random sample of points, the magnitudes of the matrix will give unbiased estimates of the actual distribution of errors. The matrix can then be used to estimate the actual amounts of confusion between classes. However, confusion matrices are frequently drawn up using the pixel values used for training the classifier. These tend to give biased estimates of the relationship between the ground data and the classification. This is firstly because the training data may not be a random sample and secondly, because they were used to develop the spectral signatures used in the classification and are not independent observations.

If several methods of classification are being tested, it is possible to use biased confusion matrices - derived for each classification and the training ground data - to compare the relative merits of the classifications but not to assess the actual performance. Such matrices tend to give erroneous estimates of classification accuracy. In the first years of Action 1, only the training pixels were used to produce confusion matrices. In most cases, confusion matrices were constructed using all the segment pixels. In other cases (e.g.: Castilla-León), random pixels were selected within the segments.

Confusion matrices produced with various methods in Action 1 made it difficult to compare classifications from different study areas or even classifications of the same area from year to year. Also, it was not possible in some cases to be sure that the confusion matrices were unbiased. This aside, the overall classification accuracies were between 60% and 70%. There was considerable variation of accuracy of individual classes in any one classification. The results presented in Table 2.1 were for the UK in 1992 and are typical in this respect. Average values, for Action 1, of user (mapping) and producer accuracy for each crop are presented in Table 2.2 and are generally rather low. Wheat obtained a higher accuracy, mainly because it was the most dominant crop. However, a lot of confusion occurred whenever differentiation between several types of wheat was tried. The same applied for barley. Crops covering a small area usually obtained mixed results depending on their characteristics. For example, rice, which is very localised with distinctive spectral characteristics, was accurately classified whereas tobacco which is usually much more scattered across the region was poorly classified.

One requirement of Action 1 was to be able to identify as many crops as possible. Depending on the location 5 to 10 agricultural themes, covering most of the crops of interest in the area, could be classified together with a few other non-agricultural ones.

Overall, although crops covering large area are more likely to be mapped best with classification of satellite imagery, it is really the spectral properties which will matter in the end. Thus, remote sensing was also particularly good for mapping crops with distinctive reflectance properties covering smaller areas, such as oil seed rape. However, the above accuracy figures indicate that area estimates made from pixel counts will not be satisfactory. Therefore remote sensing, on its own, is not a suitable technique for inventory work.

2.3.4 Timeliness

The time required by all the digital processing parts of Action 1 work depended mainly on computer power and availability of suitable software. Generally, there were no problems in accommodating this within the time frame required to deliver results.

2.3.5 Yield estimation by remote sensing

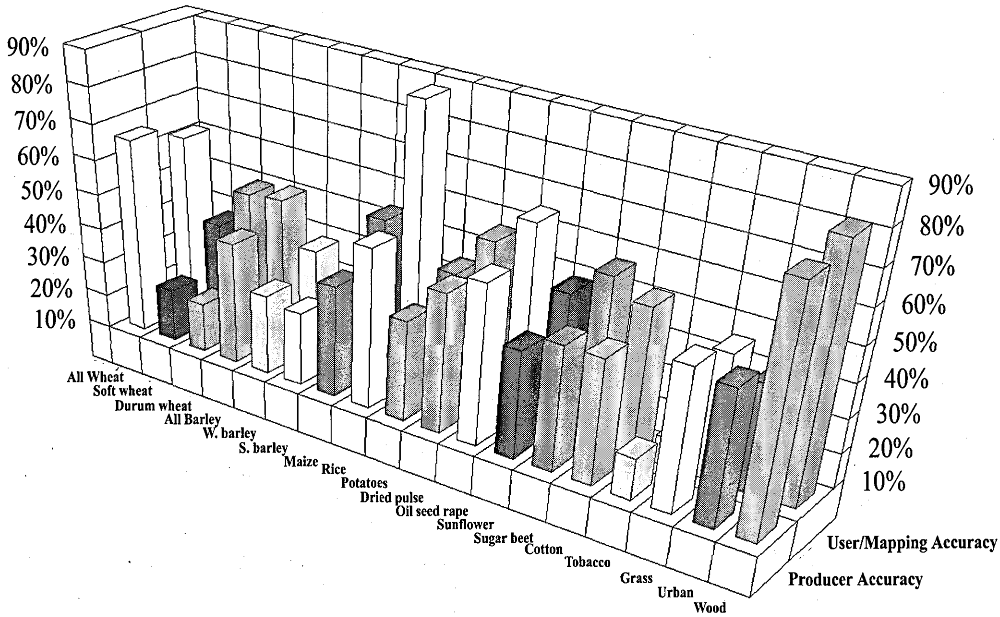
Estimation of crop yields by remote sensing was only a small component of Action 1 and really only progressed as far as determining correlation's between yield data obtained from farmers and vegetation indices, computed from the same satellite images used for the area estimates. Vegetation indices are calculated from linear combinations of spectral responses in the red (0.6-0.7 μm) and near infrared (0.7-1.1 μm) wavebands. The vegetation index mostly applied in Action 1 was the so-called normalised difference vegetation index (NDVI) which is defined as:

$$NDVI = (IR - R)/(IR+R)$$

where R and IR are the reflectance's in red and near infrared wavebands respectively which are calculated from the pixel values in the satellite images. The Index uses the fact that increasing amounts of photosynthetically active vegetation causes a reduction in R because of absorption by chlorophyll and an increase in IR because of increasing leaf layers on the plant canopy. Therefore NDVI increases with the amount of photosynthetically active vegetation (PAV) on the land surface.

Table 2.2: Average digital classification accuracies for crops in Action 1

Crop	Producer Accuracy	User/Mapping Accuracy
All Wheat	66%	61%
Soft wheat	25%	38%
Durum wheat	24%	50%
All Barley	45%	51%
W. barley	33%	39%
S. barley	31%	29%
Maize	42%	54%
Rice	56%	89%
Potatoes	39%	45%
Dried pulse	50%	57%
Oil seed rape	56%	65%
Sunflower	41%	49%
Sugar beet	46%	57%
Cotton	46%	52%
Tobacco	22%	18%
Grass	51%	46%
Urban	49%	41%
Wood	79%	82%



The relationship between the NDVI and crop yield is indirect. In cereal crops, grain yield is related to the above ground phytomass. NDVI is correlated to the Leaf Area Index and hence to the amount of phytomass. It has been demonstrated empirically that there is a positive correlation between the NDVI and cereal yield at single dates for field-plot experiments. These relationships vary with time during the season and with location. However, a regression estimator, similar to that described in PART 3, for crop yield using the NDVI could improve yield estimates.

Yield estimation with remote sensing was attempted in France, Germany, Greece and Spain in 1989 on a trial basis. Vegetation indices were calculated for the field parcels where crop yield data were available. The correlation's were not used for improving yield estimates because they were very variable and generally low. This was mainly attributed to unsuitable timing of the imagery as many of the crops were already beyond the growth stages when correlation's could be expected. Poor co-location of NDVI with yield data because of errors in geometric correction was also a factor. Poor data on crop yields in individual fields was also a problem. Resolving these difficulties was beyond the scope of the Action 1 therefore the methodology was not really developed and a more specific study is required to do this.

CONCLUSIONS:

- **The error of geometric correction was frequently greater than the ±1 pixel specified for Action 1**
- **This was mainly because of the effects of high relief in some areas or of low accuracy of available ground control**
- **Resampling method did not effect the final results**
- **Overall accuracies of digital classifications were 60 - 70%**
- **Mapping accuracy of digital classifications for individual crops varied considerably and was usually too low for area estimations to be based on pixel counts**
- **The time required for digital processing could be accommodated to enable timely delivery of results**
- **Yield estimation with remote sensing was not possible with the images acquired for Action 1 because the timing was wrong**
- **More development of a methodology for yield estimation with remote sensing is needed**

3. Crop inventory with remote sensing

The methodology is presented for improving crop estimates by combining ground survey data with digital classifications of satellite imagery using the regression estimator. The improvements of crop estimates are assessed. Cost-effectiveness of remote sensing is discussed. Results are compared with official statistics and with those of USDA-NASS.

3.1 Methodology

3.1.1 Introduction

The area estimates obtained from ground surveys contain unbiased sampling errors. Areas measured from digital classification have no sampling error because they are based on pixel counts covering the whole of the study region but these are biased because of mis-classification. Nevertheless, relationships between the areas of crops and the digital classification results can be expected and this presents the opportunity to combine the two sources of information to obtain improved area estimates using the so-called regression estimator described by Cochran (1977). The first step is to determine the relationship between area estimates by ground survey and digital classification.

3.1.2 Relationship between ground survey and digital classification

The relationship between ground survey and satellite image measurements of crop areas is determined using the sample segments. The proportion of each crop (y_i) in the sample segments was determined in the ground survey, described in PART 1. The equivalent proportions determined by digital classification (p_i) were calculated, to give pairs of observations as shown in Figure 3.1. The relationship between the two sets of information is determined by the linear regression of y on p given by:

$$y = \bar{y} + b(p - \bar{p})$$

where $\bar{y}$ and $\bar{p}$ are the sample mean values and b is the slope of the regression line.

3.1.3 The regression estimator

In the case of the satellite imagery, $\bar{p}$ can also be estimated from the digital classification. This so-called population estimate, $\bar{p}_{pop}$, is the proportion of pixels classified as the crop in the whole of the region of interest. This value is used in the regression equation to produce a correction for the sample estimate of the mean crop proportion per unit area, $\bar{y}$, and is known as the regression estimate, $\bar{y}_{reg}$, shown in Figure 3.2 and given by:

$$\bar{y}_{reg} = \bar{y} + b(\bar{p}_{pop} - \bar{p})$$

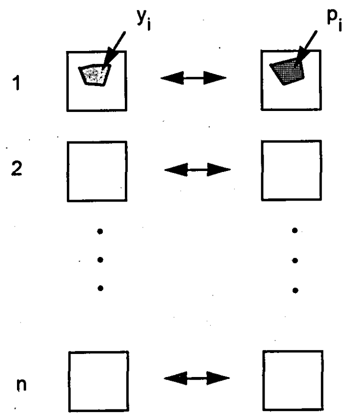


Figure 3.1: Sample of observations of crop areas by ground survey (y) and by digital classification (p), in the equivalent n segments

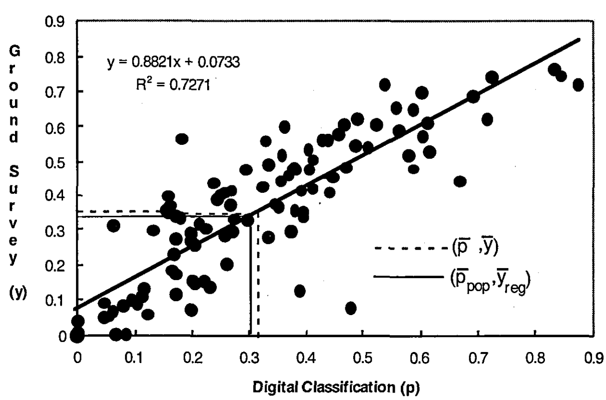


Figure 3.2: Relationship between digital classification and ground survey for wheat in the UK.

For large random samples (n>50) the variance is approximately:

$$Var(\bar{y}_{reg}) = \tfrac{1}{n} Var(y)(1 - r_{py}^2)$$

where r_{py}^2 is the coefficient of determination.

The estimate of crop area in the study region and its variance are then:

$$\hat{Z}_{reg} = D\bar{y}_{reg}$$

and

$$Var(\hat{Z}_{reg}) = D^2 Var(\bar{y}_{reg})$$

When the region of interest is stratified, regression estimates are made for each stratum separately and combined as shown for direct expansion in PART 1.

Thus, the regression estimator improves the accuracy of the area estimates by adjusting the estimate of the mean proportion of crop per sample unit from $\bar{y}$ to $\bar{y}_{reg}$ and reducing the variance.

The variance reduction is generally by far the main benefit, unless the ground sample is highly biased, and so is used as the measure of the improvement in area estimates using the regression method.

3.1.4 Neostratification imposed by satellite imagery

The application of remote sensing imposes additional requirements for sub-division of the region of interest because the digital classification has to be performed for each satellite image separately except when acquired by the same sensor and in the same orbit. This is because the images are affected by sensor characteristics and changes in illumination, atmosphere and ground surface properties as described in PART 2. Potentially, this means that separate classifications are required for each agronomically homogeneous stratum within each group of different satellite images. If a stratum is not covered by a single image, it must be divided. The position of the dividing line can be anywhere within the region of overlap of the images. The choice of position can be made to take account of the relative quality of the overlapping images, taking the line which uses more of the better quality image. It is also important to reduce the fragmentation of strata and this may also influence the positioning of the dividing line. This sub-division of the region of interest to accommodate variations in the satellite image coverage is a form of post-stratification because the boundaries of the satellite images are unknown at the time of sample selection. The process is referred here to as neostratification to distinguish it from the general concept of post stratification.

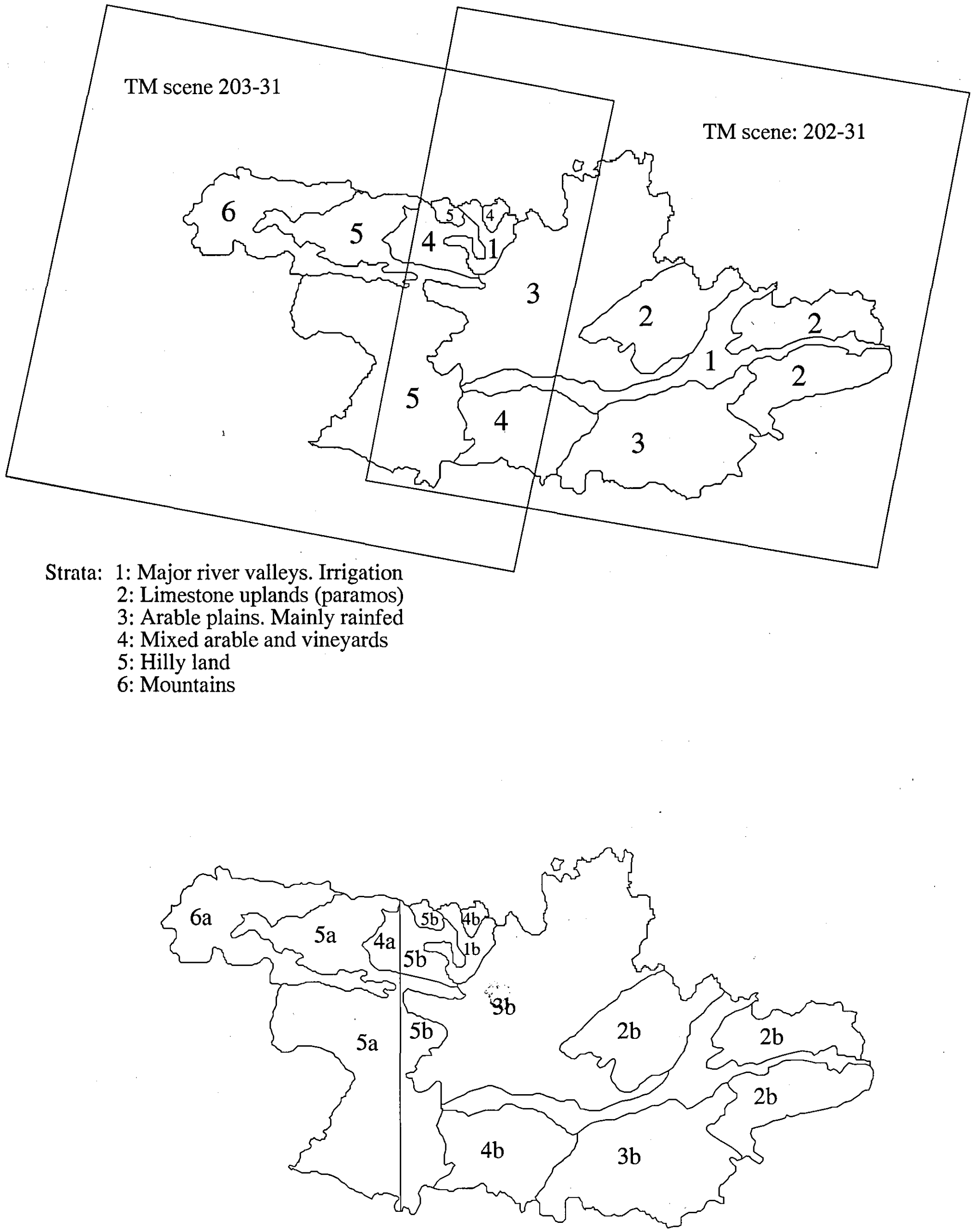


Figure 3.3: Stratification, Landsat TM image coverage, and neo-stratification in Castilla-León (Valladolid and Zamora) in 1988-89

Areas missed by image coverage can be treated as neostrata. If there are large areas covered by clouds, cloud shadows or haze, they can also be treated as neostrata. Note that it is also possible to deal with cloudy areas by creating a class "clouds" in the supervised classification, but this can induce a bias in the final estimates, unless it can be assumed that the distribution of clouds is uncorrelated with the land cover. Haze and cloud shadows can similarly be dealt with by sub-dividing the effected classes with a clustering algorithm prior to the supervised classification. Then a sub-set of classes will correspond to the land cover under haze or shade.

Figure 3.3 shows a relatively simple example of neostratification for the pilot region in Castilla-León (Valladolid and Zamora) in 1988-89. The region contained 6 strata and was covered by two Landsat TM scenes as shown in the top part of the figure. Any line in the overlap area can be considered for the neostratification. The line dividing the UTM zones 29 and 30 happened to fall in this area. To simplify operations this line was used to cut strata into neostrata. Neostrata marked with “b” were classified with image 202-31 and neostrata marked with “a” were classified with image 203-31. Neostrata 4b and 5b were merged for regression because they had too few segments, as well as neostrata 4a, 5a, and 6a.

For the purpose of making direct comparisons with regression estimates, it is necessary to rework the direct expansion of the ground survey data to conform with the neostratification. The new mean, $\bar{y}_{neo}$, will be slightly different from $\bar{y}$. The new variance is $Var(\bar{y}_{neo})$. The relative efficiency of the neostratification, $\eta_{neo} = Var(\bar{y})/Var(\bar{y}_{neo})$, can be quite different from the value obtained for the original ground survey stratification.

3.1.5 Cost benefit of remote sensing

The ratio between the variances of the ground survey area estimate and the regression estimate gives the relative efficiency of remote sensing which is:

$$\eta_{reg} = \frac{Var(\hat{Z})}{Var(\hat{Z}_{reg})} = \frac{1}{1 - r_{py}^2}$$

However, the same reduction in variance could also be achieved by increasing the size of the ground survey sample such that:

$$\eta_{reg} = \frac{\frac{1}{n}Var(y)}{\frac{1}{n_1}Var(y)} = \frac{n_1}{n}$$

where n_1 is the increased sample size.

Remote sensing will be economically efficient if the cost of ground surveying the additional n_1-n segments is greater than the cost of the remote sensing part of the project, given by:

$$n(\eta_{reg} - 1)p > R$$

where n is the original sample size, p is the cost of surveying each additional segment and R is the cost of the remote sensing part of the project. The break-even value of the remote sensing depends on the relative costs of remote sensing (images, image processing etc.) and ground survey which will vary according to circumstances. In developed countries, it is commonly accepted that the break-even occurs when the relative efficiency is close to 2.

However, remote sensing provides more than an improvement of the statistical precision. There is also the adjustment of the area estimate itself. Additional information on the location of the crops and other land uses is given plus the imagery may also be used to carry out stratification. Therefore, although relative efficiency is the main tool to measure whether or not remote sensing has been economical, the value of the additional information must also be taken into account when assessing overall economic efficiency.

3.1.6 Implementation of the methodology

There were two approaches to implementation of the regression calculations used in Action 1. The first was to use commercially available spreadsheet and database software. This worked well in the smaller and less complex studies but was not sufficiently automated when study areas increased in size and when many neostrata were involved. This pointed to the need for development and application of specialised software - the second approach. To this end, JRC produced a program to facilitate the extraction of ground survey statistics and carry out the calculations for the regression estimator. The program, called MARS-PED, was developed from the software package called PEDITOR which was produced for similar work in the USA by USDA-NASS in the 1970s and was also used by some contractors. MARS-PED was designed to link with commercially available software for image processing (ERDAS) and a Geographic Information System (ARC-INFO). However, MARS-PED, despite its advantages and the successful application of PEDITOR in France and in Germany, was not widely used in Action 1. This was because further development is needed to provide a user interface more in accordance with current commercial standards.

The main source of variation between contractors was the sophistication of their approach to deciding whether or not the application of regression estimators was valid. Some contractors made no attempt to assess the regression relationships, particularly in the early years, and applied them regardless of their quality. Others developed guidelines which are discussed in PART 4.

3.2 Results

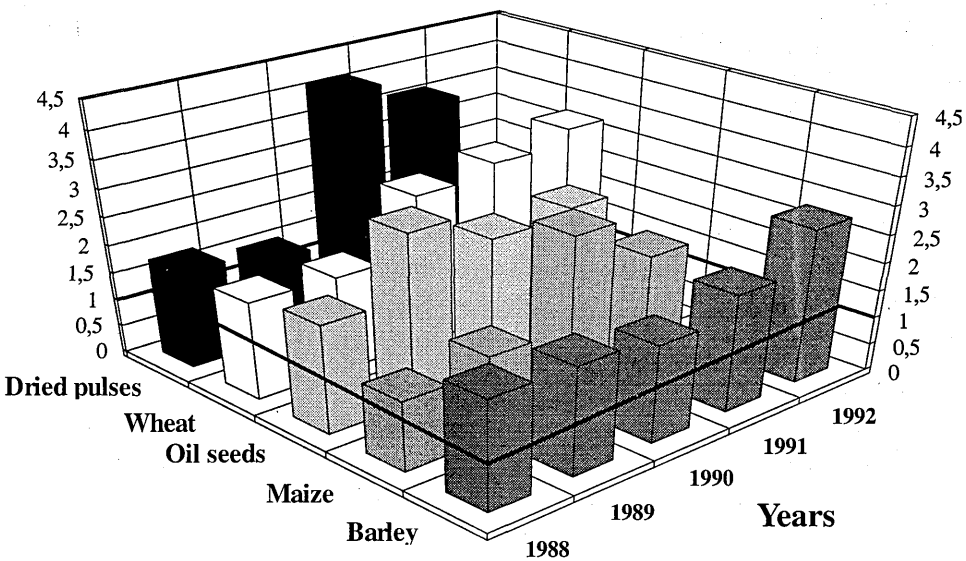
3.2.1 Improvement of estimates of main crop areas

There was considerable variation in the improvement of crop area estimates gained as a result of using the regression estimator from remote sensing. Area-weighted average relative efficiencies, as defined in PART 1, were calculated to summarise the results and to compare them for different crop types and geographical region, and to see if there were any consistent variations with time.

Table 3.1 below shows the area-weighted average of the relative efficiencies for each of the main crops of interest.

Table 3.1: Area-weighted relative efficiencies of regression estimates by crop

	1988	1989	1990	1991	1992
Wheat	1.78	1.75	2.78	2.98	3.24
Barley	1.95	1.93	1.75	2.12	2.78
Maize	1.25	1.48	3.05	2.26	n/a
Oil seeds	1.95	3.03	2.50	2.62	n/a
Dried pulses	1.77	1.53	4.14	3.61	n/a



In 1988 and 1989, the figures include all the data from the five original test sites. In 1990, data from Ile-de-France was added and the area of some test sites was extended. In 1991, the figures included data from: Centre, Ile-de-France, Castilla-León, Catalunya and Macedonia. And, in 1992, data from the United Kingdom, Belgium and Catalunya are included. The figures for the individual crops were calculated with data from at least four test-sites except for dried pulses (for which the coverage is only significant in Centre and Ile-de-France) and for 1992 (when only three sites applied the regression). In 1992, figures are only quoted for wheat and barley because the area covered by the other crops, where regression was applied, was considered too small to give representative results.

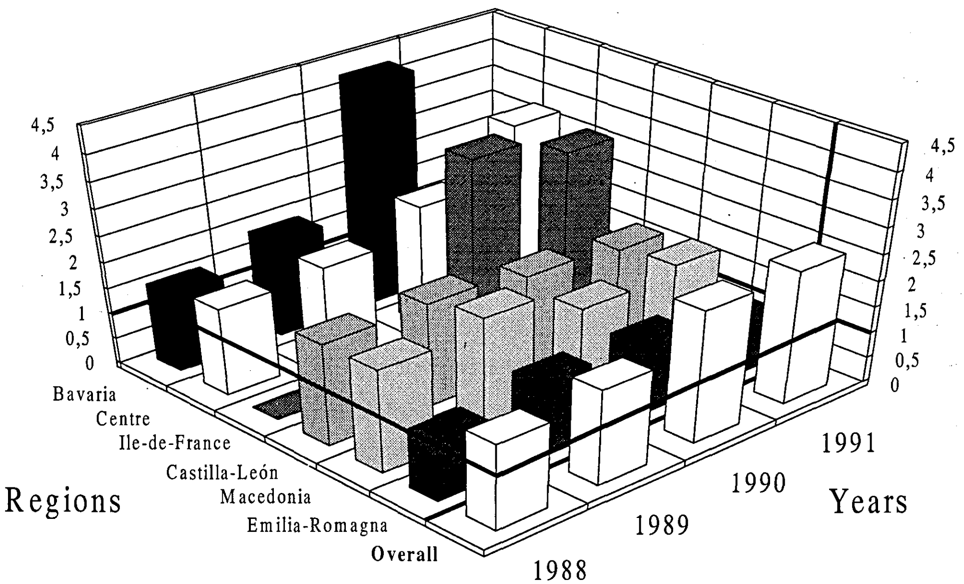
The results show a general improvement in the relative efficiency with time. By 1990, the relative efficiencies were above 2.5 for all main crops except barley. The relative efficiencies of regression estimates for wheat and barley at each study site tended to depend on the relative dominance of the crop. For example in Castilla-León barley was the dominant cereal and was estimated with a relative efficiency of 2.1 compared to 1.5 for wheat in 1991. This contrasted with the United Kingdom in 1992 where wheat was dominant and the relative efficiencies for wheat and barley were 3.7 and 1.0 respectively. This was caused by spectral confusion between wheat and barley causing mis-classification of the satellite image, the effect being worse for the minor crop.

3.2.2 Variation of results across test sites

Table 3.2 below shows the area-weighted average of the relative efficiencies for each of the main test sites. There was a general improvement as a function of time because of improvements in Bavaria, Centre and Ile-de-France. The overall relative efficiency rose to 2.55 in 1990 and 1991. Results were less satisfying in Castilla-León and in Macedonia where no improvement was found. The worst results were obtained in Emilia-Romagna where there was a slight improvement in the second year but a serious drop in 1991.

Table 3.2: Area weighted relative efficiencies of regression estimates by site and overall

	1988	1989	1990	1991
Bavaria	1.52	1.88	4.27	*
Centre	1.70	1.81	2.34	3.32
Ile-de-France	*	*	3.63	3.17
Castilla-León	1.97	2.01	1.84	1.70
Macedonia	1.98	2.20	1.66	1.82
Emilia-Romagna	1.15	1.44	1.43	1.11
Overall	1.65	1.83	2.55	2.58/2.36



There was consistent improvement in relative efficiency with time when the work was carried out by the same contractor indicating that there was a general need for contractors to gain experience of the methodology, particularly at the beginning of Action 1. Such was the case in Bayern and Centre. Ile-de-France was done by the same contractor used for Centre. However, in the case of Emilia-Romagna, there was a change in the methodology in 1990 (the square segments were replaced by cadastral segments as for the methodology applied in the USA), and in 1991, a different contractor undertook the work. Consequently , the learning process as described before cannot be expected in the case of Emilia-Romagna especially for 1991. That is why figures for this region were not included in the computation for 1991 in Table 3.1 and two figures for the overall weighted relative efficiency of regression estimates are shown in Table 3.2 for 1991. The lower value includes data from Emilia-Romagna and the other is without.

An important component of the regional inventory programme was the transfer of technology. In Castilla-León and Macedonia the work was gradually shifted from internationally renown consultant companies to local contractors as the same time that the study areas were extended. A few more years would probably be necessary in order to assess if a good grasp of the methodology was achieved. However, the implementation of remote sensing in Macedonia in 1992 showed an interesting improvement with a weighted average relative efficiency of 2.27. The MARS-PED software was used that year and was another possible factor for improved application of the methodology. Contractors in Bavaria and Centre used parts of PEDITOR in later years of Action 1.

Thus it seems that the use of specialised software to assist the statistical analysis was more generally an important factor. Also in 1992, the weighted relative efficiency for the new site in the UK was 3.2, indicating the continued trend of improvement although no specialised software for analysis was used.

There was considerable variation of relative efficiency of regression estimates for different neostrata within test sites. Factors such as the farming system and field size could have influenced the relative efficiency sufficiently to account for differences in the area-weighted values between the sites. Detailed analysis to quantify these effects was beyond the scope of this work but would be useful in future, to assist the development of more specific guidelines for the application of the regression method with remote sensing.

3.2.3 Improved accuracy of crop area estimates vs. class area

The typical accuracies of crop area estimates made with combined ground survey and remote sensing can be estimated with the aid of Figure 3.4, below and Tables 3.2 and 3.2.

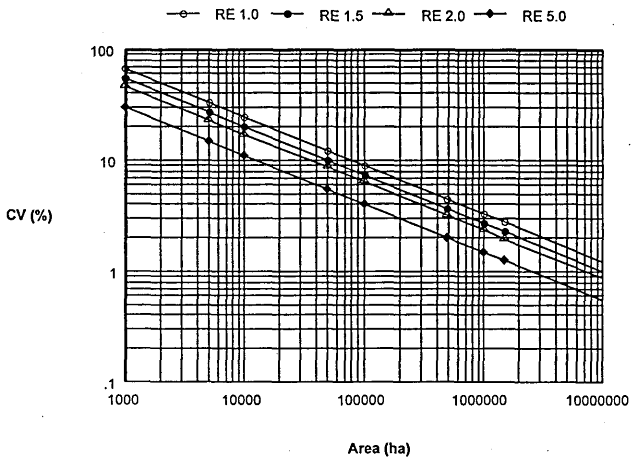


Figure 3.4: Relationship of Coefficient of Variation (CV %) to area of crop (ha) and Relative Efficiency (RE) of using remote sensing in regional crop inventories in Europe

The relative efficiency measures the improvement in precision. Figure 3.4 shows how this translates into reductions in the CV of area estimates as follows. A relative efficiency of 1.0 represents no improvement thus the corresponding line in Figure 3.4 is the relationship between CV and crop area for ground survey, shown in Figure 1.5. The CV for regression estimates is given by:

$$CV_{reg} = \frac{\sqrt{Var(\bar{y}_{reg})}}{\bar{y}_{reg}}.100$$

But on average, $\bar{y}_{reg} \approx \bar{y}$ therefore

$$CV_{ground} / CV_{reg} \approx \sqrt{RE}$$

thus

$$CV_{reg} \approx CV_{ground} / \sqrt{RE}$$

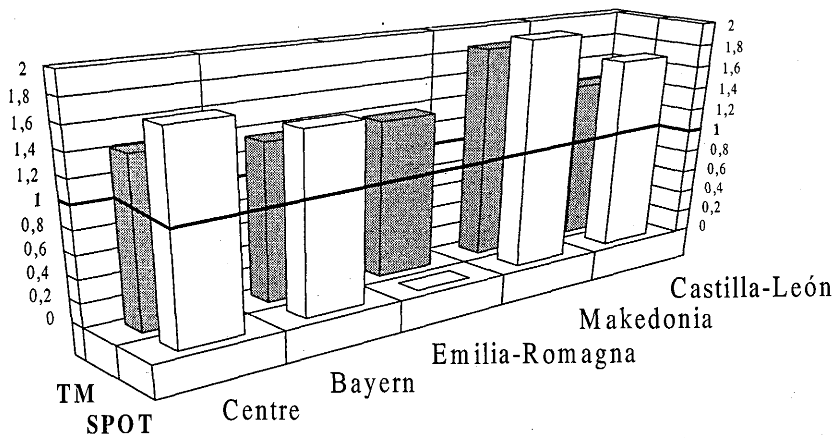
On the logarithmic scales in Figure 3.4 this leads to a series of lines parallel to and below the line for the ground survey data. Lines for REs of 1.5, 2.0 and 5.0 are shown as examples. Taking the same example as in PART 1, for crops typically covering 100,000, 1,000,000 and 10,000,000 ha respectively at regional, national and EU levels. The corresponding expected CVs for area estimates, using remote sensing with RE= 2.5, would be 5.7, 2.1 and 0.8%. A reduction of more than 30%.

3.2.4 Cost-effectiveness

In 1988, values of relative efficiencies, corresponding to the threshold for cost-effectiveness, were computed for each test site and satellite sensor and are presented in Table 3.3. The figures were based on cost estimates by the contractors. The limit is when the cost of remote sensing equals the cost of increasing the size of the ground survey as described in the methodology section. The figures obtained are always less than 2 which is usually quoted as the break even point from former studies such as the remote sensing applications programme carried out by the USDA-NASS. When the values in Table 3.3 are compared with Tables 3.1 and 3.2 it can be seen that the relative efficiencies were generally well in excess of the thresholds indicating that remote sensing had become cost-effective by the end of Action 1. The threshold values also varied with satellite sensor and with geographical location in the Community. This reflects the difference in price between images from SPOT and Landsat TM and variations in the cost of ground survey.

Table 3.3: Thresholds values of the Relative Efficiency of using remote sensing in 1988

	Centre	Bayern	Emilia-Romagna	Makedonia	Castilla-León
SPOT	1.71	1.54		2.00	1.71
TM	1.42	1.35	1.37	1.86	1.42



3.2.5 Comparison with national statistics

Some contractors compared the area estimates in Action 1 with equivalent figures obtained by the national statistical services. The comparisons, summarised in Table 3.4 were not exhaustive and do not provide a rigorous basis for evaluating either set of figures. They are only presented to illustrate how Action 1 results might be used in the context of harmonising data obtained with a variety of methodologies, as is the case with figures produces by national statistical services, or to detect gross discrepancies. For the latter, the 95% confidence interval for Action 1 estimates provides a statistical threshold, to indicate when further explanation of differences between national and Action 1 figures, is needed. Cases where the difference exceeded the threshold are indicated in the Table. The limits vary as a function of the total area of the crop as explained in Figures 1.5 and 3.4 and comparisons become less precise when the area of the crop is small. The regression estimates give narrower limits for the threshold, thus allowing more accurate detection of potential discrepancies. An example concerning harmonisation is the comparison between Action 1 figures in the UK with those obtained by the national statistical service in MAFF from an annual census by postal enquiry, direct to the farmers - a method not generally applicable throughout the Community. The results in Table 3.4 show that the MAFF Census figures for the test site are not significantly different from the Action 1 figures. Thus in this case confirming the consistency of the census.

3.2.6 Timeliness

It was not possible to carry out a rigorous analysis of the timeliness. This was because actual delivery times for results were not always apparent from the contractors reports. However, with a few exceptions which were related to occasional difficulties with satellite image acquisition, or matters of organisation, the regression estimates were available within acceptable time-frames. The technical implementation of the methodology per sé did not create significant time delays.

3.2.7 Comparison with USDA-NASS results

Crop area estimation with remote sensing, using the regression estimator, was carried out in the USA from 1980 to 1987. There were a number of differences in the methodology. Firstly NASS used Landsat MSS imagery which had a spatial resolution of approximately 80m compared to the higher resolution of SPOT (20m) and Landsat TM (30m) images available for Action 1. Secondly, the method of producing the sample frame was very different. The USDA used segments which were defined by natural boundaries as described by Cotter and Nealon (1987). This meant that their sizes varied and much more effort was required to define the sample frame than in Action 1. The method for defining the segment boundaries was by photo interpretation of satellite images. A similar approach was adopted in Italy after 1990. Thirdly, Action 1 was mainly carried out by sub-contracts subject to commercial competitive tenders and let mainly to private companies whereas the USDA project was implemented centrally by NASS. Also, Action 1 was carried out against a background of the non-uniform approaches of different nations within the Community compared to the uniform approach in the USA. Both of these factors meant there was potential for much greater variation in competence and approach in Action 1 compared to the USDA project.

Table 3.4: Action 1 estimates of crop areas expressed as a percentage of official estimates, by study site.

	Wheat	Barley	Maize	Oilseed rape	Sunflower	Sugar beet	Dried pulses
CENTRE 1988							
Official stats (x 1,000 ha)	*512*	*87.1*	*140.4*	*56.9*	*83*	*25.6*	*51.2*
Direct expansion (%)	106.62 •	**113.20**	**101.78**	103.34	97.23	**95.70**	**98.63**
Regression (%)	**95.22** •	114.58 •	105.77	**100.18**	**99.88**	85.94	107.42
MAKEDONIA 1988 (wheat only corresponds to soft wheat)							
Official stats (x 1,000 ha)	*195.5*		*30.4*		*16.3*	*10.3*	*5*
Direct expansion (%)	**93.71**		89.14		108.59	183.50 •	18.00 •
Regression (%)	81.07 •		**90.46**		**101.84**	**154.37**	12.00 •
BAYERN 1988 (wheat and barley are respectively winter wheat and winter barley)							
Official stats (x 1,000 ha)	*152.2*	*88.7*	*148.9*	*28*		*29.4*	*7*
Direct expansion (%)	105.39	**101.58**	104.70	**96.07**		98.64	64.29 •
Regression (%)	**100.72**	97.86		87.86			40.00 •
EMILIA ROMAGNA 1988 (wheat only corresponds to soft wheat)							
Official stats (x 1,000 ha)	*218.6*	*40.1*	*74.3*			*98.8*	*0.6*
Direct expansion (%)	93.69	106.73	**119.92**			**98.18**	866.67 •
Regression (%)	93.09	106.73	124.63 •			103.24	833.33 •
CASTILLA-LEON 1988							
Official stats (x 1,000 ha)	*127.2*	*527*	*7.3*		*30*	*34.2*	*18.7*
Direct expansion (%)	93.55	106.45	68.49		16.67 •	99.42	42.78 •
Regression (%)	**94.34**	**104.93** •					
CENTRE 1990							
Official stats (x 1,000 ha)	*928.5*	*169.5*	*215.5*	*95*	*234*	*28.7*	*102.9*
Direct expansion (%)	109.13 •	116.34 •	89.19	116.00	103.29	96.52	**99.42**
Regression (%)	**106.06** •	**115.10**	**104.69**	**102.84**	96.58	**100.00**	89.99
ILE-DE-FRANCE 1990							
Official stats (x 1,000 ha)	*270.5*	*49.3*	*56.3*	*29.4*	*24.8*	*44.0*	*64.9*
Direct expansion (%)	96.82	101.62	110.48	**91.50**	126.61	**84.59**	**92.50**
Regression (%)	96.56	**98.99**	**107.28**	90.48	**108.87**	83.41	73.04 •
BAYERN 1990 (wheat and barley are respectively winter wheat and winter barley)							
Official stats (x 1,000 ha)	*144.2*	*80.1*	*148.0*	*44.0*		*32.0*	*2.0*
Direct expansion (%)	**99.72**	110.74	107.03	103.41		**98.44**	100.00
Regression (%)	106.24 •		**103.99**			102.50	
UK 1992							
Official stats (x 1,000 ha)	*227.6*	*50.6*		*46.6*			
Direct expansion (%)	110.81	106.52		94.85			
Regression (%)	**104.57**			**94.85**			

Bold figures are those closest to the official figures

• indicates that the official estimate lies outside the 95% confidence interval of the Action 1 estimate

The results are reported by Allen and Hanuschak (1988). In 1980 estimates for main crops were made for the states of Iowa and Kansas (356755 sq. km). By 1985 the area had been increased to cover eight of the main agricultural states adding: Arkansas, Colorado, Illinois, Indiana, Missouri and Oklahoma, an area of 1,353,510 sq. km. The remote sensing estimates were closer to Agricultural Statistics Board final estimate than the June Enumerative Survey 21 out of 35 cases. The relative efficiencies are summarised in Table 3.5 and show a general trend of improvement with time although there was considerable variation. Generally, the relative efficiencies were remarkably similar to those in Action 1. Possibly the hoped for improvement in relative efficiency by using higher resolution satellite images in Action 1 was offset by smaller field sizes in Europe or as previously mentioned, the potential was not yet fully realised because more time was needed by contractors in general to build up experience with the methodology and to develop full technical competence. Another reason for similarity of the results is that the sampling frame in the USA was probably no so different in practice to that used in Action 1 because the majority of segment boundaries in the USA would coincide with the road system. In the mid-western states, the roads are mainly set out to follow a regular square grid of 1 mile dimension and farms were originally homesteads based on so-called quarter sections of 160 acres.

Table 3.5: Relative efficiencies of using remote sensing in eight states of the USA in the remote sensing applications program of USDA-NASS.

	1980	1981	1982	1983	1984	1985	1986	1987
Corn	1.6	1.7	1.2	1.5	2.2	1.6	1.7	2.0
Cotton				3.3	2.4	5.3	3.5	8.0
Rice		4.2		2.4	3.2	2.0	5.4	4.2
Sorghum		1.3			2.6	1.5	1.5	2.1
Soybeans	1.4	1.9	1.2	1.4	2.6	1.9	2.0	2.3
Wheat	1.4	1.8	2.0	1.6	1.9	2.3	2.2	2.6

CONCLUSIONS:

- **By 1990 the relative efficiency of the regression estimator was usually above 2.5 for main crops and rising towards 3 by 1992**
- **There was considerable variation in relative efficiency for each crop, tending to be higher when the crop was dominant in the site**
- **Relative efficiency of remote sensing improved with contractor experience**
- **Specialised software for statistical analysis assisted improvement of results**
- **Technology transfer was not completed and further improvements should be possible, raising relative efficiencies to above 3**
- **Relative efficiency was very variable within sites and could be influenced by factors such as farming system and field size**
- **On average remote sensing improved accuracy of area estimates by more than 30% by the end of Action 1**
- **Remote sensing was cost-effective by the end of Action 1**
- **Action 1 area estimates can be used to assist harmonisation of data produced by different methods**
- **Action 1 area estimates can be used as indicators of gross discrepancies in area estimates from different sources**
- **Technical implementation of remote sensing analysis per sé did not create unacceptable time delay**
- **Relative efficiencies of Action 1 results were similar to the USDA-NASS eight state survey**
- **Benefits of improved satellite imagery in Action 1 compared to USDA-NASS were probably obscured by variation of technical competence in Europe which improved during Action 1**

4. Technical Factors Influencing Results

Technical factors influencing the accuracy of Action 1 results are identified and discussed. Where possible guidelines for assessing future projects are developed.

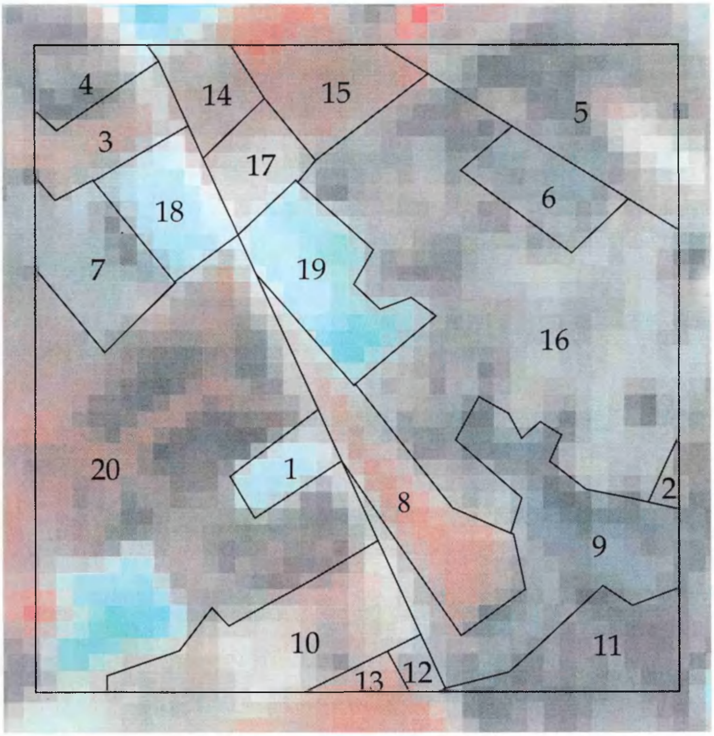
4.1 Accuracy of ground surveys

A high level of accuracy in the ground survey is fundamental for the success of all aspects of the Action 1 work. Errors affect the direct expansion estimates in PART 1; the training of digital classification algorithms in PART 2 and the coefficient of determination of regression estimates in PART 3. Errors in field work were of two types: mis-classification of the cover type and incorrect delineation of parcels.

Most of the problems concerning the mis-classification occurred when dealing with non-cropped areas (e.g. fallow and pasture land). There was some variation in the interpretation of the definitions for these classes across Europe and this caused the areas to vary more than expected from year to year.

In principle, 10 % of the segments were independently checked. However, no standard rules were systematically applied. For instance, the checked plots were not always randomly selected. In some cases segments with a lot of parcels or certain enumerators were preferably checked. Checking was done either by the supervisory staff or by a double survey using different enumerators in the survey team. When the government services undertook the ground survey, check results were not usually available and it was very difficult to assess the quality of the ground survey results. In some cases, photo interpretation of satellite images acquired for Action 1 and contemporary with the ground survey, were successfully used to identify field work errors. Examples are shown in Figure 4.1: field number 20, coded as wheat, contains a large patch of bare soil, that has escaped to the attention of the surveyor, may be because he looked from the road in the middle of the segment; the rest of the polygon has a rather heterogeneous response, and may correspond to several contiguous fields of wheat, which is not an error. The border between fields 14 and 17 is slightly mislocated. Field number 5 contains an area with scarce vegetation, but it is not clear whether it is a survey error.

In a few cases, in difficult terrain, ground survey sites were incorrectly located. Thus, there were various ways of measuring the ground survey errors, which appeared not to be negligible in several cases. It was not possible to quantify ground survey errors from reported information. However, the general precision of ground survey in Action 1 was considered acceptable. The methodology of the ground survey was readily adapted for local conditions which mainly related to accommodation of variation in crop types and parcel sizes, as discussed in PART 1.



Plot	Cover	Plot	Cover	Plot	Cover	Plot	Cover
1	fallow	6	waste land	11	high scrub	16	barley
2	barley	7	waste land	12	barley	17	oast
3	wheat	8	peas	13	peas	18	fallow
4	rye	9	waste land	14	barley	19	fallow
5	barley	10	sunflower	15	wheat	20	wheat

Figure 4.1: Examples of satellite imagery of segments showing discrepancy with ground survey data

The optimum size of the sample segments depends on the average field size and the homogeneity of the land use. In general, it is better to have a large number of small segments but this increases cost and operational difficulties. An intuitive rule arising from the experience of Action 1 is to make segments large enough so that they include 20 to 30 field parcels on average. When the average number of parcels exceeds an upper threshold, the fieldwork burden can be reduced by quartering the segment and sampling a quarter chosen randomly. In Spain, a threshold of 50 parcels was used.

CONCLUSIONS ON THE GROUND SURVEY:

- **Field survey checks can be assisted by the photo interpretation of satellite imagery in areas where field sizes are large (> several hectares)**
- **The precision of the ground survey methodology was generally sufficient for main crops with sample fractions between 0.5 and 1.5%**
- **The ground survey methods were flexible enough to adapt to local conditions**
- **Definition of some land cover classes needs to be improved to enable more consistent interpretation across Europe**

RECOMMENDATION FOR FUTURE GROUND SURVEYS:

- **Ground survey should be checked independently, preferably with double survey in a random sub-sample of segments**

4.2 Success of image coverage

At the beginning of the project, cloud cover was thought to be a serious limitation to the use of satellite imagery for regional inventories in central and northern parts of the EU. The number of images acquired for Action 1, within the specified time windows is summarised in Table 4.1.

Table 4.1: Number of images which were acquired within the time window specified for the identification of the crops of interest

	France	Germany	Greece	Italy	Spain	UK
1988	2/10,0/2*	10/10,0/2*	11/11,4/4*	5/6	13/13,1/2*	n/a
1989	2/2	1/3	4/4	3/3	3/4	n/a
1990	26/26	0/3	4/4	3/3	3/5	n/a
1991	3/4	n/a	5/5	2/2	5/5	n/a
1992	n/a	n/a	2/4	n/a	n\a	4/4

* The first ratio is about the SPOT imagery and the second TM; n/a not applicable.

However, experience with the MARS programme shows that this is rarely a problem with the current number of imaging satellites that are available and that cloud problems could be encountered in Mediterranean climates while suitable cloud free imagery could be obtained in Northern Europe under oceanic climatic conditions. Overall, 82% of the images acquired within the regional inventory programme were within the optimum window. This ratio was respectively 89% for SPOT and 75% for TM and suggests that the combination of programming and pointing capabilities of SPOT enabled more flexibility for acquisition of cloud free images at the desired date. Most of the remaining images were not very far outside of the acquisition window and were still used in the classification. Exceptional cases were in Germany in 1990 when no images could be obtained in June so the regression estimates could only be applied to summer crops, and in Spain in 1988.

Table 4.2 below shows the satellite coverage expressed as a percentage of the study area. The area not covered includes areas where no imagery was available and areas covered by clouds. Satellite image coverage was much better after 1988 and this is attributed to improvements in satellite programming. By 1989, satellite coverage's are all above or around 90% except for Italy in 1990. The number of operational satellites (1 Landsat and 2 SPOT) was sufficient to provide enough overpasses to acquire adequate cloud free imagery. The delivery time to contractors was reduced to about two weeks and was generally rapid enough not to inhibit the production of timely crop estimates. This contrasted with the earlier work by USDA-NASS where lack of timely imagery was a serious problem.

At the beginning of Action 1, both TM and Spot images were acquired to see if the type of imagery influenced the quality of the results. The area-weighted relative efficiencies of regression estimates from TM and SPOT images, over the same area in Table 4.3 show that there was no consistent advantage from using one type of imagery rather than the other. Accuracy of the results depended more on the timing of acquisition in the agricultural cycle than on the type of sensor.

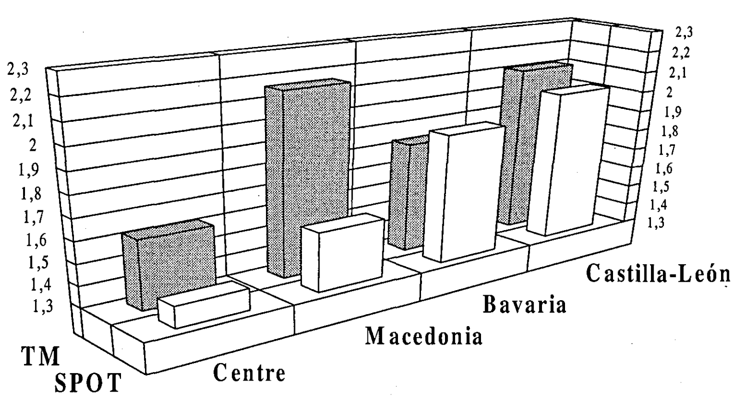
Table 4.2: Percentage of the study area covered by cloud-free satellite imagery

	France	Germany	Greece	Italy	Spain	UK
1988	88%,97%*	80%,90%*	90%,99%*	100%	95%,70%*	
1989	89%	100%	100%	90%	99%	
1990	100%	90%	100%	82%	90%	
1991	92%	n/a	100%	99%	90%	
1992	n/a	n/a	unknown	n/a	n/a	100%

* The first figure is about the SPOT imagery and the second TM; n/a Not Applicable.

Table 4.3: Area-weighted relative efficiency of regression estimates from TM and SPOT over equivalent areas

	SPOT	TM
Centre	1.40	1.62
Macedonia	1.58	2.16
Bavaria	1.92	1.83
Castilla-León	2.03	2.11



CONCLUSIONS ON SATELLITE IMAGE ACQUISITION:

- **Acquisition of improved to > 90 % coverage as vendors obtained experience and improved satellite programming**
- **Imagery was usually acquired within the specified time-window**
- **Cloud cover prevented image acquisition on a few occasions but the problem was not prevalent in any geographical location**
- **Delivery times became acceptably short**
- **There was no consistent advantage of SPOT over TM imagery**

RECOMMENDATION ON IMAGE ACQUISITION FOR FUTURE PROJECTS:

- **Centralised purchasing of images should be done if possible to assist co-ordination and programming of satellites**
- **Use all available satellites to increase chances of acquiring high quality images within the optîmum time-window**

4.3 Quality of geometric correction

Errors in geometric correction of satellite images causes a mis-match between ground survey observations and the satellite imagery. This affects the accuracy of the regression estimator in two ways. Firstly, by reducing the accuracy of the digital classification because the spectral properties of classes were not defined accurately in the training procedure. Secondly, the coefficient of determination of regressions can be reduced because the double sample used in the regression does not cover corresponding areas.

In general, contractors were not able to achieve enough precision in geometric correction to be able to ignore the mis-match. Also, there was evidence that some, particularly at the beginning of Action 1, did not appreciate that actual errors could be much larger than the RMS values described in PART 2. This occurred when high order polynomial transform equations were used and when the ground control points were poorly distributed or low in number. It was not possible to assess the size of errors introduced into regression estimates by poor geometric correction. In any case, most contractors circumvented the problem by introducing manual shifting or even sometimes warping the ground segments to achieve a best fit to the image by visual interpretation.

CONCLUSIONS ON GEOMETRIC CORRECTION:

- **There was evidence of technical incompetence by some contractors through inexperience which was later overcome**
- **Small manual corrections were necessary to accurately co-locate ground survey and image data**

RECOMMENDATION REGARDING GEOMETRIC CORRECTION:

- **Processing chain should allow independent check of the co-location of ground survey and satellite imagery**

4.4 Quality of regression relationships

The quality of the regression relationships between ground survey and digital classification was one of the main factors influencing the improvement in crop area estimates and was very variable throughout Action 1. There was a tendency at the beginning for regression estimates to be made without regard to the inherent quality of the relationships. This also occurred later with inexperienced contractors, and in cases where neostratification resulted in few observations to produce the regression. Some contractors were concerned about the validity of the regressions they were applying and arrived at the empirical rules to help decide whether to apply them.

The initial recommendation that each regression should have at least 30 observations was not sufficient on its own to ensure high quality. Regression relationships are not useful when the coefficient of determination r^2 becomes very low, as shown by the example in Figure 4.2 for barley in the UK. They become suspect when number of points is low and/or poorly distributed across a range of values. An example of this is shown in Figure 4.3. The r^2 is very high but is caused by a single point where there was a high proportion of the crop in a segment. All the other points are scattered around zero. If the single point happens to be an error, then the regression estimate and the supposedly high relative efficiency may be spurious because the regression relationship is badly affected if the point is removed.

Another guideline is to look for consistency of the regression slope with the confusion matrix when the r^2 is high. This is done by comparing the relative magnitudes of the row and column totals for a class with the slope of the regression. Take the example for wheat in the UK. In Table 2.1, the total number of observations classified as wheat was 173 and the total observations on the ground was 184. Therefore, wheat is slightly under classified by the image. This is consistent with the slope of the regression being near to one as is confirmed in Figure 3.2. When the image grossly over classifies, the slope would be much less than one and vice versa. In such cases, the regression estimate should be viewed with suspicion.

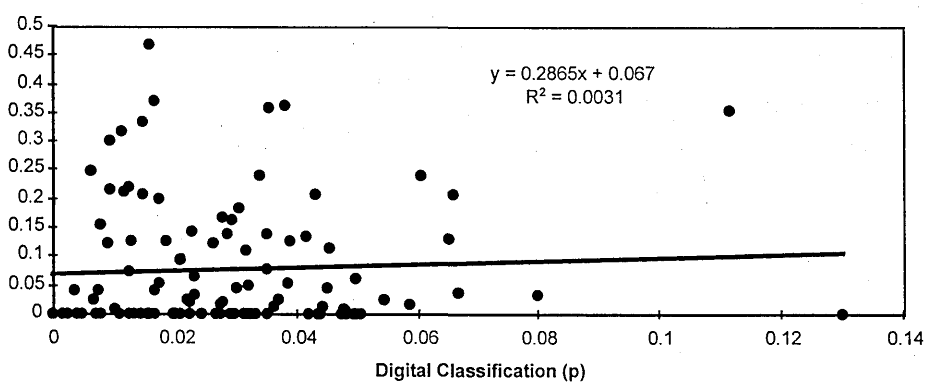


Figure 4.2: Example of poor regression relationship, because of low coefficient of determination, r^2.

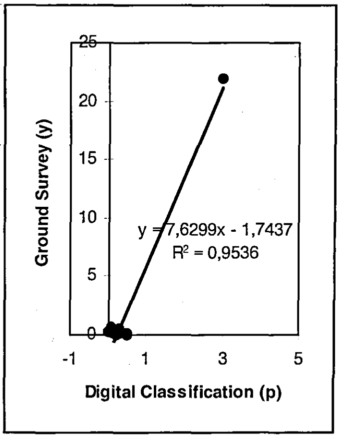


Figure 4.3: Example of an unreliable regression relationship, because of few and poorly distributed points

RECOMMENDED GUIDELINES FOR APPLICATION OF THE REGRESSION ESTIMATOR:

- **Coefficient of determination, r2 >0.3**
- **A producer accuracy >30% to ensure a sufficient number of well classified pixels**
- **A sufficient number of non zero points in the regression, well distributed across the range**
- **Slope of regression should be consistent with the ratio of row and column totals in the confusion matrix**

4.5 Comments on Neostratification

In Action 1, the same neostratification was used for all the tasks concerning the application of the regression estimator with remote sensing, i.e. image classification, estimation of regression parameters, and application of these parameters to compute the corrected estimates. The variance of the final regression estimators should be lower by using different ones for each, but the procedure becomes more difficult to manage and it is not clear if the decrease of the variance makes up for the higher complexity.

For supervised digital classification, the procedure is easier and the computation is faster with a few large strata and can give as good results as a larger number of small strata if pre-clustering is performed to sub-divide land use classes which may have different radiometric properties in different areas of the image. For the estimation of the regression parameter b, large strata are strongly recommended in order to ensure that there are sufficient observations to give good regression relationships. If this is not possible then a reasonable solution to avoid these risks is amalgamate several strata to estimate b and then compute the regression estimates separately in each of the individual strata. If the estimates of regression parameters are not very reliable, even after aggregation of several strata, the ground survey estimates and their corresponding variances are to be kept for the computation of the final results. This can happen for minor crops concentrated in small areas.

4.6 Relationship between classification accuracy and improvement in precision

Another reason why the regression estimates failed to improve the crop area estimates was when the accuracy of the digital classification was low. The RE and the corresponding mapping (user) accuracy was extracted for all of the regressions in the contractors' reports. This resulted in over 400 points; from all years of the project, except 1988 when no confusion matrices were produced; across all the study regions ; and, for all the crops. There were two problems with this data set. The first was that some of the regression relationships were of doubtful quality when judged by the criteria discussed in the previous section. It was not possible to eliminate all these by detailed examination of each one because the raw data were not available. However, the most likely spurious regressions are those with extremely high r^2. Therefore, all those with $r^2 > 0.96$ (RE > 25) were excluded from the analysis. The second problem is that the mapping (user) accuracies are determined from sample observations and are therefore subject to sampling errors and bias if the sample was not random. It was not possible to assess these errors for each point. Therefore, only an approximate analysis of the data was done. The data were grouped into ranges of the mapping (user accuracy) and a box plot showing the variation of the RE in each range is shown in Figure 4.4.

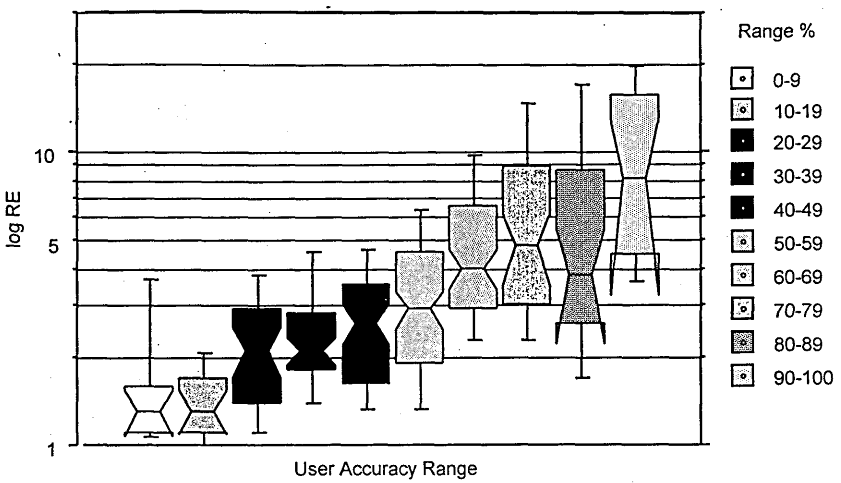


Figure 4.4: Box plot showing 10,25,50,75 and 90 percentiles for the relative efficiency (RE) of regression for different ranges of mapping (user) accuracy of the digital classification

The results show that relative efficiencies increase with increase in the mapping accuracy, as expected intuitively. Relative efficiencies >2 are expected when the mapping accuracy increases above 50% and >3 when the mapping accuracy rises above 60%. There is greater variation in the RE as mapping accuracy increases. Part of this may be attributable to the use of different methods by contractors, to produce the confusion matrices.

Thus, the relationship between the regression and the classification accuracy is an important factor to look at. If the relationship is inconsistent with Figure 4.4, the reason should be investigated and corrected otherwise the regression in question should not be used. For example a case when the RE>3 and the mapping accuracy was less than 10% - much higher than the normally expected value of around 1.

CONCLUSIONS ON RELATIONSHIP BETWEEN MAPPING ACCURACY AND REGRESSION ESTIMATES:

- **Relative efficiency increases with increased mapping accuracy of the digital classification**
- **Relative efficiencies >2 are expected when mapping accuracy >50%**
- **Relative efficiencies >3 are expected when the mapping accuracy >60%**

RECOMMENDATION FOR CHECKING FUTURE WORK:

- **Use Figure 4.4 to judge consistency between the digital classification and relative efficiency. Outliers indicate possibility of error in implementation of the methodology**
- **Check that the confusion matrixis based on a random sample. The column totals should be approximately proportional to the land cover areas and the row totals to the classification theme areas.**

4.7 Effect of different digital classification rules

Table 4.4: Areas (ha) of cover types in the same region of England estimated by different techniques using digital classifications and by agricultural census

CLASS	PC-W[1]	PC-UW[2]	REG-W[3]	REG-UW[4]	MAFF[5]
Woodland	35153	42834	29636 ±19%	29409 ±25%	na
Inland Water	5510	4123	5744 ±58%	6279 ±44%	na
Urban	111775	22518	82594 ±14%	70668 ±20%	na
Wheat	210459	159938	238003 ±6%	236736 ±6%	227637
Barley	22839	93969	53900 ±27%	55502 ±24%	50585
Summer Crops	66990	105593	89888 ±13%	96494 ±15%	82587
Grass and Forage	192582	107509	130339 ±12%	124691 ±13%	114491
Rape	29946	29047	44244 ±10%	44095 ±11%	46643

[1]pixel count, area-weighted discriminant functions; [2] pixel count, un-weighted discriminant functions [3]regression estimate, area-weighted discriminant functions; [4]regression estimate, un-weighted discriminant functions; [5]MAFF agricultural census

The mathematical rule used in the digital classification process influences the classification results. In Action 1, the maximum likelihood algorithm described in PART 2 was used but results will be different if unequal *a priori* weightings are used for the discriminant functions. Table 4.4 compares the results from two maximum likelihood classifications carried out for the UK test site. The column PC-W was for class areas obtained by pixel counts following classification using discriminant functions which were weighted according to class areas, estimated from the ground survey and PC-UW was for equally weighted discriminant functions. The columns REG-W and REG-UW are the corresponding areas and their 95% confidence intervals estimated by regression. The MAFF column gives, where available, the areas estimated by the census carried out independently by the Ministry of Agriculture, Fisheries and Food. Comparison of the area estimates for each cover type shows that the pixel counts are widely different, influenced by varying the digital classification algorithm and hence the degree of mis-classification. The regression estimates on the other hand, are similar to each other and to the MAFF census figures. This shows that the differences caused by varying the classification algorithms were corrected by the regression technique.

CONCLUSION ON INFLUENCE OF DIGITAL CLASSIFICATION RULE:

- **The regression estimate is relatively insensitive to variations in classification algorithm**
- **Area estimates using pixel counts should never be used unless the mapping and producer accuracies are simultaneously > 90%**

5. Bibliography

5.1.1 Preparatory studies for Action 1

- Aquater, SODETEG. 1987. Statistique d'occupation du sol. Rapport Final Préliminaire. CCR, Ispra.
- Logica 1987. Feasibility study of the EEC landuse statistics project. Draft final report. JRC, Ispra.
- Moulias J. 1986. Pré-évaluation de l'utilisation de la Télédétection dans le système communautaire d'Information Agricole. Rapport Final.

5.1.2 Contractors reports

- Aquater, SODETEG-TAI. 1989. Projet Agriculture. Inventaires Régionaux 1988 France-Italie. Volume I: Présentation Générale, région Emilia-Romagna et région Centre. Rapport final. CCR, Ispra.
- Aquater, SODETEG-TAI. 1989. Projet Agriculture. Inventaires Régionaux 1988 France-Italie. Volume II: Resultats segments et télédétection, région Emilia-Romagna. Rapport final. CCR, Ispra.
- Aquater, SODETEG-TAI. 1989. Project Agriculture. Inventaires Régionaux 1988 France-Italie. Volume III: Resultats segments et télédétection Région Centre. Rapport final. CCR, Ispra.
- Aquater, SODETEG-TAI. 1989. Projet Agriculture. Action 1: Inventaires Régionaux 1988 France-Italie. Volume I: Campagne 1989, Région Centre. Rapport final. CCR, Ispra.
- Aquater, SODETEG-TAI, SYSAME. 1989. Projet Agriculture. Action 1: Inventaires Régionaux. Region: France et Italie. Volume II: Rapport Final Campagne 1989. Région Emilie-Romagne. CCR, Ispra.
- Aquater, SYSAME. 1990. Projet Agriculture. Action 1: Inventaires Régionaux. Région: Italie. Volume I: Rapport Final Campagne 1990. Région Emilie-Romagne. CCR, Ispra.
- BDPA-SCETAGRI, BRGM, ADK. 1988. Projet Télédétection Agriculture. Action 1: Inventaires Régionaux. Région Makedonia. 1er rapport Intermédiaire. CCR, Ispra.
- BDPA-SCETAGRI, BRGM, ADK. 1988. Projet Télédétection Agriculture. Action 1: Inventaires Régionaux. Région Makedonia. Résultats de l'enquête-terrain. 2eme rapport Intermédiaire. CCR, Ispra.
- BDPA-SCETAGRI, BRGM, ADK. 1988. Projet Télédétection Agriculture. Action 1: Inventaires Régionaux. Région Makedonia. Volet Télédétection: Méthodologie et Résultats. 3eme rapport Intermédiaire. CCR, Ispra.
- BDPA-SCETAGRI, BRGM, ADK. 1989. Inventaires Régionaux en Makedonia (Grèce). Campagne 1989. Raport Final. CCR, Ispra.
- Belgian Ministry of Agriculture, University of Gent and University of Liège. 1992. MARS 1992 Regional Inventory, Belgium. Interim Report. CEC-JRC Ispra.
- BRGM, Organotecnica, SYSAME. 1990. Estimation des surface cultivées en blé dur, Grèce 1990. Thessalia-Makedonia Anatoliki. Résultats télédétection. FEOGA.
- GAF. 1988. Regional Inventories, Survey 1988. Final Report. JRC, Ispra.
- GAF. 1989. Regional Inventories Niederbayern-Oberpfalz, Survey 1989. Final Report. JRC, Ispra.

- GAF. 1989. Regional Inventories Niederbayern-Oberpfalz, survey 1990. Final Report. JRC, Ispra.
- GEOSYS, SGS-qualitest. 1990. Project FEOGA blé dur sud de l'Italie. Rapport Final. FEOGA.
- GEOSYS, SGS-qualitest. 1990. Project FEOGA blé dur sud de l'Italie. Rapport Final. Annexes. FEOGA.
- HTS Ltd., AURENSA. 1989. Spain: Castilla León. Regional Crop Inventory from satellite imagery 1988. Final Report. JRC, Ispra.
- HTS Ltd. 1989. Analysis of 1989 Crop Inventory data of Castilla-León. Final Report. JRC, Ispra.
- HTS Ltd., AURENSA. 1990. Spain: Castilla León. Regional Crop Inventory from satellite imagery 1990. Final Report Year 2. JRC, Ispra.
- HTS Ltd., AURENSA. 1990. Remote sensing programme. Agriculture Project. Regional crop inventory of Castilla-León, Spain: Summary Report. JRC, Ispra.
- HTS Ltd., AURENSA. 1990. Spain: Castilla León. Regional Crop Inventory from satellite imagery 1990. Final Report Year 3. JRC, Ispra.
- HTS Ltd, AURENSA. 1990. Estimation of summer crop areas in Valladolid and Zamora Provinces using satellite imagery. Final Report. JRC, Ispra.
- HTS Ltd, AURENSA. 1991. Regional crop inventory from satellite imagery 1990. Supplement to final report. Year 3. Summer crops production survey results. JRC, Ispra.
- HTS Ltd, AURENSA. 1991. Regional crop inventory from satellite imagery 1990. Final Results Burgos Province. JRC, Ispra.
- HTS Ltd, AURENSA. 1991. Spain: Castilla-León. Regional crop inventory from satellite imagery. 1991. Final Report Year 4. JRC, Ispra.
- Institut Cartogràfic de Catalunya 1991. MARS project. Catalunya 1991. JRC, Ispra.
- Kolá J. and Jelínková. 1992. Agriculture project. Crop area and yield inventories in the Czech Republic, 1992. Interim Report. GISAT, Prague, JRC, Ispra.
- Ministerio de Agriculture, Pesca y Alimentacion. 1992. Utilizacion de Imagenes de satelite paro la estimacion de superficies de cultivos en Segovia 1990. Boletin Mensual de Estadística 3:72-92.
- SYSAME, BRGM, ADK. 1990. Inventaires Régionaux en Makedonia (Grèce). Campagne 1990. Rapport Final. CCR, Ispra.
- SYSAME, BRGM, Organotecnica. 1990. Estimation des surface cultivées en blé dur, Grèce 1990. Résultats de l'enquête-terrain. Thessalia-Makedonia Anatoliki. FEOGA.
- SYSAME, BRGM, Organotecnica. 1990. Estimation des surface cultivées en blé dur, Grèce 1990. Thessalia-Makedonia Auatoliki. Annexes. Résultats détaillés de l'enquête terrain toutes cultures. FEOGA.
- SYSAME, ISPIF. 1992. Projet pilote, inventaires régionaux, Roumanie, 1992. Rapport final. CCR, Ispra.
- SYSAME, Organotechnica, BRGM. 1991. Action 1: Regional Inventories. Makedonia Region, 1991 campaign. Final Report. JRC, Ispra.
- SYSAME, SCEES. 1991. Projet Agriculture. Action 1: Inventaires Régionaux France. Résultats Terrain et Télédétection campagne 1990. Régions: Centre et Ile-de France. Rapport final. CCR, Ispra.
- SYSAME, SCEES. 1991. Projet Agriculture. Action 1: Inventaires Régionaux France. Résultats Terrain et Télédétection campagne 1991. Régions: Centre et Ile-de France. Rapport final. CCR, Ispra.
- SYSAME, SCEES. 1991. Projet Agriculture. Action 1: Inventaires Régionaux France. Résultats Terrain et Télédétection campagne 1992. Régions: Centre et Ile-de France. Rapport final. CCR, Ispra.
- SYSAME. 1992. Etude Pilote TER-UTI, Indre et Loire. Rapport Final. CCR, Ispra.

- Taylor J.C. and Eva H.D. 1992. Regional Inventories on Beds, Cambs and Northants (UK). Final report produced by Silsoe College for the JRC, Ispra, 70 pp.
- Taylor J.C. and Eva H.D. 1992. Regional Inventories on Beds, Cambs and Northants (UK). Summary report produced by Silsoe College for the JRC, Ispra, 9 pp.
- Tele-Expert. 1992. Technical Support to the Action 1 (Regional Inventories) of the MARS project in Greece. Final Report. JRC, Ispra.
- Tele-Expert. 1993. Technical support to the Action 1 (Regional Inventories) of the MARS project in Greece. Final Report. JRC, Ispra.
- Telespazio, EUROMED, GEIE. 1991. Action 1. Regional Inventories on selected administrative regions. Regional Inventories in Italy. Final Report. JRC, Ispra.
- Trabajos Catastrales SA. 1992. Estimation of area, yield and production of cereals in Navarra. Final report.

5.1.3 Associated JRC reports and publications

- Bernard A.C. and Meyer-Roux J. 1994. Proceeding of the Conference on MARS Project: Overview and Perspectives. Official Publication of the European Communities, EUR 15599, 168 pp.
- Gallego F.J.. 1995. Sampling Frames of Square Segments. Official Publication of the European Communities, EUR 16317, 72 pp.
- Perdigào V. 1991. Ground survey documents based on high resolution satellite imagery, 1990 campaign. Publication of the European Communities Joint Research Centre, Ispra, Italy, S.P.I. 91.29, 95 pp.
- Stakenborg J. 1989. Television Tracking Digitization of boundaries with a video camera. Publication of the European Communities Joint Research Centre, Ispra, Italy, S.P.I.89.09.
- Toselli F. and Meyer-Roux J. 1990. Proceedings of the Conference Application of Remote Sensing to Agricultural Statistics. Official Publication of the European Communities, EUR 12581, 386 pp.
- Toselli F. and Meyer-Roux J. 1992. Proceedings of the Conference Application of Remote Sensing to Agricultural Statistics. Official Publication of the European Communities, EUR 14262, 382 pp.

5.1.4 Other scientific sources

- Allen D.A. and Hanuschak G.A. 1988. The Remote Sensing Applications Program of the National Agricultural Statistics Service: 1980-1987. United States Department of Agriculture, National Agricultural Statistics Service, Research and Applications Division, SRB Staff Report Number SRB-88-08, 43 pp.
- Cochran W.G. 1977. Sampling Techniques, third edition. John Wiley and Sons, 413 pp.
- Cotter J. and Nealon J. 1987. Area Frame Design for Agricultural Surveys. United States Department of Agriculture, National Agricultural Statistics Service, Research and Applications Division, 67 pp.
- Richards J.A. 1993. Remote Sensing Digital Image Analysis. Springer-Verlag, 360 pp.
- Schowengerdt R.A. 1983. Techniques for Image Processing and Classification in Remote Sensing. Academic Press, Inc., 265 pp.
- Swain P.H. and Davis S.M. 1978. Remote Sensing: The Quantitative Approach. McGraw-Hill, Inc., 418 pp.

European Commission

EUR 17319 – Regional crop inventories in Europe assisted by remote sensing: 1988-1993

C. Taylor, C. Sannier, J. Delincé, F.J. Gallego

1997 – 80 pp. – 16.2 x 22.9 cm

Agriculture

This report is a synthesis of the work, the results and an assessment of the achievements of Action I of the MARS Project. The report consist of this stand-alone executive summary and five additional sections. The first three enlarge on the main elements of the methodology: ground survey, remote sensing and the combination of these using regression. These sections also present a synthesis of the results and comments on their accuracy. PART 4 is an assessment of technical factors influencing the results. PART 5 is the Bibliography, listing all the contractors reports and other reference material which have been used to provide the basis of this work.