# TOWARDS IMPROVED ESTIMATES IN WATER RESOURCES ASSESSMENTS: THE USE OF HYPERSPECTRAL AND MULTISPECTRAL IMAGERY TO CLASSIFY AND MAP VEGETATION IN SOUTHERN AFRICA

Report to the
**Water Research Commission**

by

**M. Govender[1], V Naiken[1], K Chetty[2], H Bulcock[2]**

[1]CSIR, Natural Resources of the Environment, Ecophysiology Research Group
[2]School of Bioresources Engineering and Environmental Hydrology,
University of KwaZulu-Natal

**WRC Report No 1684/1/08**
**ISBN 978-1-77005-735-7**

**July 2008**

**DISCLAIMER**

This report has been reviewed by the Water Research Commission (WRC) and approved for publication. Approval does not signify that the contents necessarily reflect the views and policies of the WRC, nor does mention of trade names or commercial products constitute endorsement or recommendation for use.

ACKNOWLEDGEMENTS

The project team wish to thank the Water Research Commission for funding this research study. We also wish to acknowledge contributions made by the following reference group members:

- Dr R Dube (Water Research Commission)
- Dr J van Aardt (CSIR)
- Dr O Mutanga (University of KwaZulu-Natal)
- Mr T Newby (Agricultural Research Council)
- Dr T Morgenthal (*ex* Agricultural Research Council)
- Mr J Bosch (CSIR)
- Dr C Jarmain (CSIR)
- Prof GGS Pegram (University of KwaZulu-Natal)
- Dr F Ahmed (University of KwaZulu-Natal)
- Dr M Roualt (University of Cape Town)
- Prof GPW Jewitt (University of KwaZulu-Natal)

This research is also supported by the European Space Agency (ESA ®) through the Tiger initiative, and acknowledgement is made for the use of Category 1Proba CHRIS data.

The project team would also like to thank Mondi for allowing access to the catchment and for data; and the Agricultural Research Council (ARC) for the use of ASD spectrometer during the field campaigns.

We acknowledge valuable technical support from Joshua Xaba and Lelethu Sinuka during field campaigns.

# EXECUTIVE SUMMARY

## 1. AIMS AND OBJECTIVES

The aims of this project were to:

- Review hyperspectral remote sensing technology and its application in vegetation and water resources studies.
- Measure and characterize the spectral signatures of selected vegetation classes for example indigenous, exotic and declared or candidate streamflow reduction activities for Southern Africa.
- Map selected and classified vegetation classes from the research site using hyperspectral imagery.
- Contribute the results towards an envisaged spectral library for vegetation in Southern Africa.

The aims and objectives of this research project were addressed as follows:

- A review comprising of an introduction to and brief history of remote sensing; a comparison between multispectral and hyperspectral remote sensing applications and their spatial and spectral differences; the fundamentals of image resolution, and details of contiguous spectral signatures; and the applications of hyperspectral remote sensing in water resources studies and more specifically on vegetation classification studies was undertaken. This review has been published in Water SA Volume 33 (2007). In addition, literature on data reduction methods, accuracy assessment and a concise summary of the national landcover database 2000 was reviewed.
- ASD hand-held leaf spectral reflectance signatures were acquired for exotic and declared/candidate streamflow reduction activities *viz*. *Acacia mearnsii, Pinus patula, Eucalyptus macarthurii, Eucalyptus grandis* and *Saccharum* within the Mistely-Canema estate.
- Supervised vegetation classification of the selected vegetation classes was investigated using both hyperspectral and multispectral imagery. Vegetation classes were classified using known regions of interest and vegetation indices

which were subjected to five different statistical classifying techniques. The results from this research study will be published in Water SA.

- A coordinated effort into the development of a vegetation spectral library is necessary in Southern Africa. Most studies have collected vegetation spectral signatures using the ARC ASD field spectrometer. Therefore, all vegetation spectral signatures and ancillary data acquired in this research project are being housed within ARC spectral database.

## 2. INTRODUCTION

Accurate monitoring and assessment of water resources is necessary for sustainable water resource management. Earth observation technologies have formed the basis for acquiring data remotely for many years (Landgrebe, 1999) and are viewed as time- and cost-effective techniques for undertaking large-scale monitoring (Okin *et al.*, 2001). Furthermore, the use of remote sensing technologies in water resources applications and more specifically in vegetation studies is well documented (Gitelson and Merzlyak, 1996; Zagolski *et al.,* 1996; Asner, 1998; Cochrane, 2000; Stone *et al.,* 2001; Coops *et al.,* 2002; Schmidt and Skidmore, 2003; Underwood *et al.,* 2003; Govender *et al.,* 2007).

Historical remote sensing applications have been confined to the use of low spectral resolution datasets such as multispectral airborne or satellite images and aerial photography. Low spectral resolution imagery covers only selected regions of the electromagnetic spectrum resulting in fewer spectral bands with coarse bandwidths, hence more generalized products. With the development of hyperspectral remote sensing technologies which extend across the electromagnetic spectrum producing high spectral resolution datasets, remote sensing has become a powerful analysis tool for applications in water resources management.

Hyperspectral imagery acquired using advanced hand-held spectroradiometers, airborne or satellite sensors have the advantage of high spatial and spectral resolution. High spectral resolution imagery covers a wide region of the electromagnetic spectrum (approximately 400 to 2500 nm) resulting in more spectral bands with narrower bandwidths (generally less than 10 nm) available for detailed analyses. The finer spectral resolution of a hyperspectral imager allows for detection of surface materials and their abundances, as well as inferences of biological and chemical processes. This capability plays an important role in the processing and analysis of remote sensing imagery resulting in more detailed products.

Remote sensing users within South Africa are only recently applying hyperspectral imaging technology to various scientific disciplines including water resources studies. The availability of hyperspectral data will provide researchers an opportunity to pursue more complex analyses and have a greater potential of producing more detailed end products. Within the scope of this research project improved vegetation classification and mapping in Southern Africa using hyperspectral imagery was investigated.

## 3. METHODOLOGY

The Mistley-Canema estate was selected as the study area for this research project. The estate is situated in the Seven Oaks district approximately 70 km from Pietermaritzburg in the Kwa-Zulu Natal midlands. In this study area, *Acacia mearnsii, Pinus patula, Eucalyptus macarthurii, Eucalyptus grandis* and *Saccharum* were considered based on their percentage area within the Mistely-Canema estate.

Remote sensing imagery, ground-based spectral measurements and ancillary data were collected during winter and summer thereby accounting for the significant differences in vegetation over the dry and wet seasons. Data sources selected for this study included the ASD hand-held spectrometer and Hyperion and Proba Chris satellite sensors. Proba Chris data were obtained from the European Space Agency (ESA) through the Tiger Initiative which is currently underway to support earth observation research in Africa. The acquired remote sensing images were geometrically and atmospherically corrected using the ENVI and Erdas remote sensing software. The Hyperion and Proba Chris images were georeferenced and orthorectified to UTM 36 S, datum WGS 84. Atmospheric corrections were performed using the empirical line method. Data reduction methods were used to manipulate high spectral resolution datasets, in order to reduce high data dimensionality and the complexity which arises when applying statistical classifiers to such datasets. The hyperspectral and multispectral datasets were reduced using an analysis of variance method and minimum noise fraction with an optimal band selection method based on known targets and background materials within the images.

Regions of interest were selected for the training and validation phases performed during supervised classification. These regions were selected to represent 8 year old *E. grandis,* 10 and 12 year old *E. macarthurri*, 12 and 16 year old *P. patula*, 4 year old *Saccharum* and 6 year old *A. mearnsii*. Two statistical measures were used to assess the spectral separability of the selected regions of interest: the Jeffries-Matusita and Transformed Divergence separability measures (Richards and Jia, 1999). All regions of interest when analysed produced a high spectral seperability. In this study five statistical classifying techniques were used in the training phase of the classification process. These techniques included maximum likelihood, minimum distance, mahalanobis distance, spectral angular mapper and parallelepiped. An iterative process was followed to determine the optimized parameters required for each statistical classifier. Thomlinson *et al*. (1999) suggested an overall accuracy of greater than 85% as indicator of superior classification. However, in this study only those statistical classifying techniques with an overall accuracy greater than 80% during the training phase were selected for validation. Accuracy assessments were used to determine which statistical classifier produced best classification results. The kappa coefficient of agreement and the overall accuracy (> 80%) were the two statistics most prominently used for decision making purposes (Foody, 2006). The user's accuracy was used to evaluate how well an individual vegetation class was classified for a specific statistical classifier method.

The use of vegetation indices as a technique for improved vegetation classification was also investigated in this study. Twelve vegetation indices were used in this study. Each vegetation index layer was classified using the parallelepiped and minimum distance statistical classifiers. Only those vegetation indices that attained an overall accuracy of greater than 80% were selected for validation.

## 4. RESULTS AND DISCUSSION

The mahalanobis distance and maximum likelihood statistical classifiers were selected from the training phase for validation. Validation was performed using the entire study area. Both methods performed well, with overall accuracies above 80%. However, the maximum likelihood classifications achieved the highest overall accuracies above 98%. Furthermore, classifications of each vegetation class at their respective age groups were evaluated to determine how accurately the statistical classifiers, mahalanobis distance and maximum likelihood performed at a species level classification. This evaluation was undertaken by comparing the user's accuracy for each vegetation class together with a visual assessment of their spatial area mapped within the study area.

During winter 2006 and summer 2007, the mahalanobis distance method best classified 4 year old *Saccharum* and 6 year old *A. mearnsii* using the Proba Chris multispectral dataset. In contrast, the maximum likelihood method best classified the 6 year old *A. mearnsii* and 10 and 12 year old *E. macarthurri* during winter and, the 6 year old *A. mearnsii* and 12 year old *E. macarthurri* during summer.

In general, both statistical classifiers best classified 4 year old *Saccharum* and 8 year old *E. grandis* when using the Hyperion reduced dataset A. Furthermore, when applying the maximum likelihood method to the Hyperion reduced dataset B, it was possible to classify and map 12 year old *P. patula* species.

Four vegetation indices were selected from the training phase and were best classified using the minimum distance method. The minimum distance method best classified the 6 year old *A. mearnsii* and the 10 and 12 year old *E. macarthurii* for all four vegetation indices.

## 5. CONCLUSIONS

The accuracy of classifying and mapping different vegetation classes is strongly dependant upon a unique combination of spectral bands. Furthermore spectral variability is evident across different vegetation classes within a season and between dry and wet seasons.  In general the Proba Chris multispectral imagery performed well for genus level classification whereas the Hyperion hyperspectral bands accurately classified both genera and species. The results of the classification and mapping of vegetation classes using selected vegetation indices were satisfactory. When applying a vegetation index only *A. mearnsii* and *E. macarthurii* species were accurately classified using the minimum distance statistical classifier method.

# TABLE OF CONTENTS

# LIST OF FIGURES

# LIST OF TABLES

# 1. INTRODUCTION

Water is one of the most valuable and essential resources that form the basis of all life. With the ever-increasing human population, there is constant stress exerted on water resources (McGwire *et al*., 2000). Accurate monitoring and assessment of our water resources is necessary for sustained water resource management. Earth observation technologies have formed the basis for acquiring data remotely for many years (Landgrebe, 1999) and are viewed as time- and cost-effective techniques for undertaking large-scale monitoring, which can be used to determine the quality, quantity and geographic distribution of water resources (Okin *et al*., 2001). The use of remote sensing imagery in water resources applications and more specifically in vegetation studies is well documented (Gitelson and Merzlyak, 1996; Zagolski *et al.,* 1996; Asner, 1998; Cochrane, 2000; Stone *et al*., 2001; Coops *et al*., 2002; Schmidt and Skidmore, 2003; Underwood *et al*., 2003; Govender *et al*., 2007).

The currently used national landcover 2000 database is a popular and highly utilized dataset based on remote sensing imagery. The remote sensing imagery used in these applications and in hydrological research in the past have been confined to the use of low spectral resolution datasets such as multispectral airborne or satellite images and aerial photography. Low spectral resolution imagery covers only selected regions of the electromagnetic spectrum resulting in fewer spectral bands with coarse bandwidths, hence more generalized products. On the other hand, hyperspectral remote sensing, which spans across the electromagnetic spectrum, has become a powerful analysis tool for applications in water resources management.

Hyperspectral imagery acquired using advanced hand-held spectroradiometers, airborne or satellite sensors have the advantage of high spatial and spectral resolution. High spectral resolution imagery covers a wide region of the electromagnetic spectrum (approximately 400 to 2500 nm) resulting in more spectral bands with narrower bandwidths (generally less than 10 nm) available for detailed analyses. The finer spectral resolution of a hyperspectral imager allows for detection of surface materials and their abundances, as well as inferences of biological and chemical processes. This capability plays an important role in the processing and analysis of remote sensing imagery resulting in more detailed end products.

Remote sensing users within South Africa are only recently applying hyperspectral imaging technology to various scientific disciplines including water resources studies. The availability of hyperspectral data will provide researchers an opportunity to pursue more complex analyses and have a greater potential to improve accuracy of end products. Within the scope of this research project improved vegetation classification and mapping in Southern Africa using hyperspectral imagery was investigated.

The specific aims of this project are:

- To review hyperspectral remote sensing technology and its application in vegetation and water resources studies.
- To measure and characterize the spectral signatures of selected vegetation classes for example indigenous, exotic and declared or candidate streamflow reduction activities for Southern Africa.
- To map selected and classified vegetation classes from the research site using hyperspectral imagery.
- To contribute the results towards an envisaged spectral library for vegetation in Southern Africa.

The report is structured as follows:

- A brief overview of the literature review paper is reported in Section 2 and is attached as Appendix A. In addition, a review on data reduction methods, accuracy assessments and a concise summary of the national landcover database are reported in Section 2. The methodology Section 3 includes a description of the Mistley-Canema study area, data collection during the winter and summer field campaigns, pre-processing of earth observation data which includes geometric and atmospheric corrections and data reduction techniques. Procedures for classification and assessment of the remote sensing imagery are also reported in Section 3.

- Results of the classification and mapping of vegetation are presented and discussed in Section 4.
- Conclusions and recommendations for further research are discussed in Section 5.

# 2. LITERATURE REVIEW

A review on the use of remote sensing technologies in water resources studies and more importantly vegetation classification has been undertaken in this project. The literature review comprised of an introduction to and brief history of remote sensing; a comparison between multispectral and hyperspectral remote sensing applications and their spatial and spectral differences. Also included are sections on the fundamentals of image resolution, and details of contiguous spectral signatures. Emphasis is placed on the applications of hyperspectral remote sensing in water resources studies and more specifically on vegetation classification studies. This review has been published and is included as Appendix A.

In addition, it was deemed necessary to include a review of literature on data reduction methods, accuracy assessment and a concise summary of the national landcover database 2000.

## 2.1 Data Reduction

Data dimensionality refers to the number of spectral bands of a remote sensing image (Hsu, 2007). As the dimensionality increases with the number of spectral bands, the number of training samples required for training a specific classification method increases (Hsu, 2007).  Bellman (1961) termed this phenomenon as the ‘curse of dimensionality”, which led to the “peaking phenomenon” or “Hughes phenomenon” characterized by an insufficient training sample which is not representative of the research area (Hughes, 1968). According to Hsu (2007), the amount of training data required for non-parametric classifiers increases exponentially with respect to an increase in dimensionality. The “curse of dimensionality” is enhanced with the advent of hyperspectral instruments that produce images with hundreds of spectral bands simultaneously. Increase in dimensionality affects the complexity of the classifiers and therefore the overall performance of hyperspectral image classification (Hsu, 2007).

Intuitively the "curse of dimensionality" can be subdued by selecting a sufficiently large sample size. For more complex classifiers, a larger ratio between sample size and dimensionality is expected. However, in practice, the size of the training samples is limited for most hyperspectral applications (Hsu, 2007). An effective method of dealing with high-dimensional data characteristic of hyperspectral data is to reduce the number of spectral bands. Therefore the reduction of hyperspectral data by reducing the number of spectral bands to be used in the application is seen as an effective control measure when using high dimensional data.

The ultimate goal of reducing the dimensionality of hyperspectral data is to substantially reduce the number of spectral features without jeopardising the integrity of the output information (Hsu, 2007). Four data reduction techniques *viz.* principle component analysis, canonical variate analysis, minimum noise fraction (MNF) and analysis of variance (ANOVA) with a Tukey's HSD post hoc test are briefly discussed in the subsequent sections.

Principle component analysis is a commonly used technique for reducing data dimensionality and highlighting variation in data (Kalacska, Bohlman, Sanchez-Azofeifa, Castro-Esau, Caelli, 2007). This method is based on the orthogonalization of the components of the input data such that the outputs from the principle component analysis will be uncorrelated with each other. By ordering the principal components from high to low variation, the principal components with the least contribution to variation in the dataset can be eliminated (Kalacska *et al.*, 2007).

Canonical variate analysis also known as linear or canonical discriminant analysis is a multivariate statistical tool that has also been successfully used to optimally separate different groups based on their spectral features (Mutanga, 2005). The objective of canonical variate analysis is to discriminate between predetermined, well defined groups of sampling entities based on a suite of characteristics (Mutanga, 2005).

Minimum noise fraction is another commonly used data reduction transformation technique. It is a linear transformation that involves two separate principal components analysis rotations. MNF is used to isolate "noise" in the data and by removing these spectral bands computational requirements are reduced when processing hyperspectral imagery (Boardman and Kruse, 1994).

An alternate method which uses analysis of variance with a Tukey's HSD post hoc test has been used to reduce dimensionality of hyperspectral data (Ismail *et al.*, 2007). When performing an analysis of variance, the Tukey's HSD post hoc test is used to determine significant differences between group means. A frequency distribution is used to indicate the occurrence of significant bands and highlight bands which could be more important for discriminating between different vegetation classes (Ismail *et al.*, 2007).

## 2.2 Accuracy Assessment

In a statistical context, accuracy defines the bias and precision in a dataset. However, in remote sensing applications the term classification accuracy is used to express the degree to which a derived remote sensing image classification represents reality (Foody, 2004). Accuracy assessments are widely accepted as a fundamental component when investigating the classification of remotely sensed data (Congalton and Green, 1993; Congalton, 1994; Foody, 2004). However, the design and implementation of accuracy assessment procedures and representative field validation studies are a challenging and time-consuming component of remote sensing studies which rarely receive adequate attention (Mundt *et al*., 2005). Accuracy assessments are important not only in evaluating the suitability between different classification techniques but also provide a guide to the quality of a map, its appropriateness for use in a particular study, as well as understanding possible errors and the implications thereof (Foody, 2004).

According to Congalton (1994) accuracy assessments have evolved with increasing rigor and detail attached to the analysis process. Four major stages of accuracy assessments have evolved overtime. The first stage involved a basic visual appraisal of the derived map which was deemed a very subjective approach. The second stage involved a more objective way of quantifying the geographical accuracy between derived images and their location on the earth or reference data. A major limitation of this approach was its non-site specific nature as; a derived map could display correct classes in correct proportions but at incorrect geographical locations (Congalton, 1994). The third stage involved the derivation of accuracy metrics based on comparisons between class labels in the classified maps and reference data for a set of specific locations. This included measures such as the number of cases correctly allocated often referred to as the overall accuracy (Congalton, 1994). The last stage identified was a refinement of the third stage in which greater use on the correspondence of the predicted classes to those made on the ground was made. This stage used a confusion or error matrix (Congalton, 1994).

The confusion or error matrix forms the building blocks of an accuracy assessment (Congalton and Green, 1993; Congalton, 1994; Foody, 2004). The confusion or error matrix can be described as a square array of numbers set out in rows and columns which express the number of sampling units (pixels) assigned to a particular class, relative to the actual class verified by ground-truth data (Congalton and Green, 1993). According to Foody (2004) the matrix provides a basis on which to describe classification accuracy and characterizes errors which can be used to improve the classification outputs.

Several measures or parameters of classification accuracy may be derived from a confusion matrix (Foody, 2004). Overall accuracy, which is the number of correctly allocated classes, is most commonly used in vegetation studies. The diagonal elements of the matrix represent the correctly classified pixels. Therefore, an absolutely correct classification will generate a diagonal matrix (Boschetti *et al.*, 2004). The overall accuracy is computed by dividing the total correctly classified pixels (i.e. the sum of the major diagonal), by the total number of pixels in the error matrix (Congalton, 1991). The error matrix also allows for the calculation of specific accuracy measures such as the user and producer accuracy and, omission and commission errors (Liu *et al.*, 2007). Producer's accuracy indicates the probability of a reference pixel being correctly classified. The user's accuracy indicates the probability with which a pixel classified on the map represents the same category or class on the ground (Liu *et al.*, 2007). The converse of user and producer accuracy is the error of commission and omission. According to Boschetti *et al.,* (2004) the commission error is the percentage of pixels classified into a class to which they do not belong (according to the reference data). The omission error is that percentage of pixels which belong to certain class in the reference data and which have been omitted (Boschetti *et al*., 2004).

Another statistical measure used to describe the accuracy of spatial data sets is the Kappa coefficient of agreement (Foody, 2004; Congalton and Mead, 1983). The strength of the Kappa analysis lies in the fact that it provides two tests of significance (Congalton, 1994). The first test of significance allows for assessing whether an individual map generated from remotely sensed data is significantly better compared to a map which has been generated by randomly assigning labels to areas. The second test of significance allows for the comparison of any two matrices to investigate whether they are significantly different from each other. The KHAT statistic is the result of performing a Kappa analysis. Landis and Koch (1977) categorised the possible ranges of the KHAT statistic into three groups where a value greater than 0.8 represents strong agreement; a value between 0.8 and 0.4 represents moderate agreement and a value below 0.4 represents poor agreement.

Foody (2004) advocates the use of accuracy assessments as a fundamental step in the classification process. This process should include at least one quantitative metric of classification accuracy together with appropriate confidence limits and, should be compared in a statistically rigorous manner to provide an objective basis for comment and interpretation.

## 2.3 A Concise Summary of the National Landcover Database

The primary objective of the national landcover 2000 database was to generate an up-to-date land-cover map of South Africa using remote sensing data to provide a long-term assessment of landscape changes resulting from natural and man-impacted environments. According to Majeke *et al*. (2006) a set of guidelines based on three key issues were produced to accomplish this objective. These issues were confirming minimum information requirements of land-cover database, determining most appropriate data formats and mapping technologies, and developing a strategic plan for long-term continuity of land-cover monitoring.

Several end-user requirements were considered in the process of developing the national landcover 2000 database, to ensure that a strategic, national level land-cover information and long-term monitoring database was generated. The database needed to complement existing and future detailed local mapping activities and should be compatible with national and international standards (Majeke *et al*., 2006). Furthermore, the process needed to be flexible, adaptable and compatible for future technologies, suitable for national-level mapping and, implementation was entirely conditional to the availability of funds (Majeke *et al*., 2006). In addition, end products needed to meet the direct information needs of key end-users, with future datasets matching or exceeding previous datasets with regard to issues of scale and detail (Majeke *et al*., 2006).

A key requirement when developing the landcover database was to ensure that the final information content was suitable for a wide range of uses with the quality of landcover information being dependent on the design and format of the classification scheme used. It was recommended that an a-prior, hierarchically structured approach be used, whereby landcover and landuse classes can be pre-determined and contained within separate hierarchical levels. Thirty classes were created using photo-interpretation techniques for the national landcover 1994 project based on 1:250 000 scale maps and Landsat TM imagery which were acquired on a specific date. The same reference classification system was modified into forty nine classes for the national landcover 2000 project based on improved mapping capabilities with multi-temporal digital Landsat ETM imagery. The additional classes used in national landcover 2000 database represented sub-divisions of broader classes previously defined in the national landcover 1994 database. For example forest plantations were further sub-divided into clear-felled, pine, eucalypt, wattle and mixed subclasses.

The national landcover 2000 database adopted an international land cover classification system to facilitate comparison and integration across the world (Majeke *et al*., 2006).

## 3. METHODOLOGY

A description of the study area, the collection of data during the winter and summer field campaigns, pre-processing of the satellite imagery, and the procedures of classification and accuracy assessment used in this study are discussed in subsequent sections.

### 3.1 Description of the Mistley-Canema Study Area

The Mistley-Canema estate is situated in the Sevenoaks district approximately 70 km from Pietermaritzburg in the Kwa-Zulu Natal midlands, South Africa (Figure 1). An initial scoping field visit of the Mistley-Canema estate was undertaken by the project team in June 2006. At least five or six appropriate vegetation classes ranging from exotic, natural and candidate streamflow reduction activities (e.g. pine, eucalyptus, wattle, grass, and sugarcane) are located within this estate.

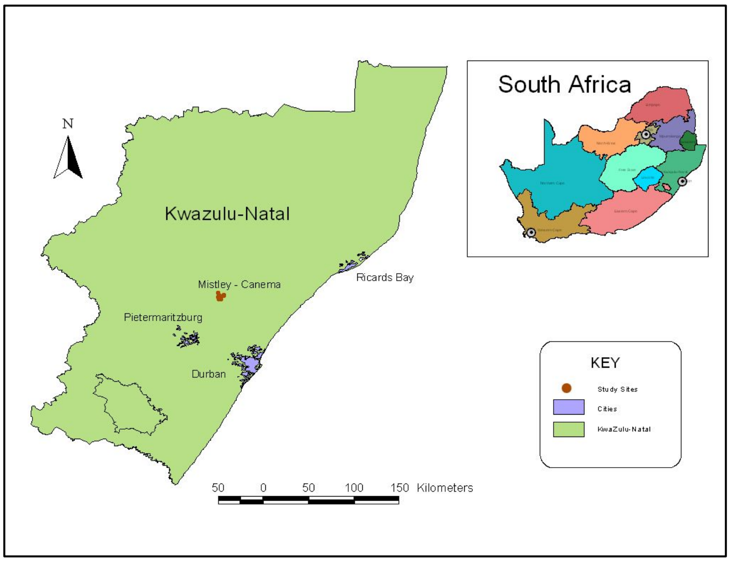



Figure 1 Location of the Mistley-Canema estate in KwaZulu-Natal

The estate is considered to be a beneficial study site as the Two Streams CSIR experimental research catchment is located within this estate (Everson *et al*., 2007). Measurements being carried out for the past six years at this experimental research site could be used to complement further hydrological and plant physiology studies using remote sensing data.

The natural vegetation of the area was previously *Themeda triandra* grassland. Only a few relic patches remain. The area is generally hilly with a high percentage (46%) of arable land. The soils are highly leached and forestry and sugarcane is considered ecologically suitable and is the most widespread vegetation classes on this estate (Everson *et al*., 2007). Common vegetation classes include *Saccharum* (sugarcane), *Acacia* (wattle), *Eucalyptus* (gum) and *Pinus* (pine) species as well as grass. For this project, *Acacia mearnsii, Pinus patula, Eucalyptus macarthurii, Eucalyptus grandis* and *Saccharum* were selected based on their percentage area within the Mistely-Canema estate.

The climate in this area is humid, with an annual rainfall ranging from 800 to 1280 mm, and a mean annual temperature of 17ºC. The bio-resource group of the study area is the moist midlands mistbelt (Camp, 1997). Heavy mists are common and are important feature providing additional moisture particularly to the forests (Everson *et al*., 2007). Climate hazards include occasional droughts, usually of short duration, occasional hail and frost which vary from slight to severe conditions. Hot "berg winds" followed by cold fronts make for unpredictable conditions, particularly in spring and early summer (Camp, 1997).

Thirteen different soil forms were identified within the study area with the Inanda soil formation being dominant. Inanda soils are dark brown to brown, of a sandy clay loam texture. These soils are characterized by a humic A horizon (30-50 cm) on a red apedal clay loam B horizon (90-120 cm), with a 10% gravel content on a saprolite. A majority of the vegetation classes of interest are located on the Inanda soil form as shown in Figure 2. Some of the *Eucalyptus Grandis* is located on the Hutton soil form, while *Pinus patula* and *Eucalyptus Macarthurri* are found on Glenrosa and Nomanci soils respectively. Hutton soil forms are reddish brown in colour and have a medium sandy clay loam texture. Huttons are characterized by a topsoil (40-50 cm) on a red, apedal, clay loam subsoil (>150 cm), with 10% surface rocks or boulders present. Glenrosa is pale brown to brown in colour with a sandy loam texture. Glenrosa soil forms are characterized by topsoils (20-40 cm) on hard lithocutanic subsoils. Nomanci soil forms have a sandy clay loam texture and are brown in colour, topsoil depth 30-60 cm on a hard saprolite or rocky subsoil.

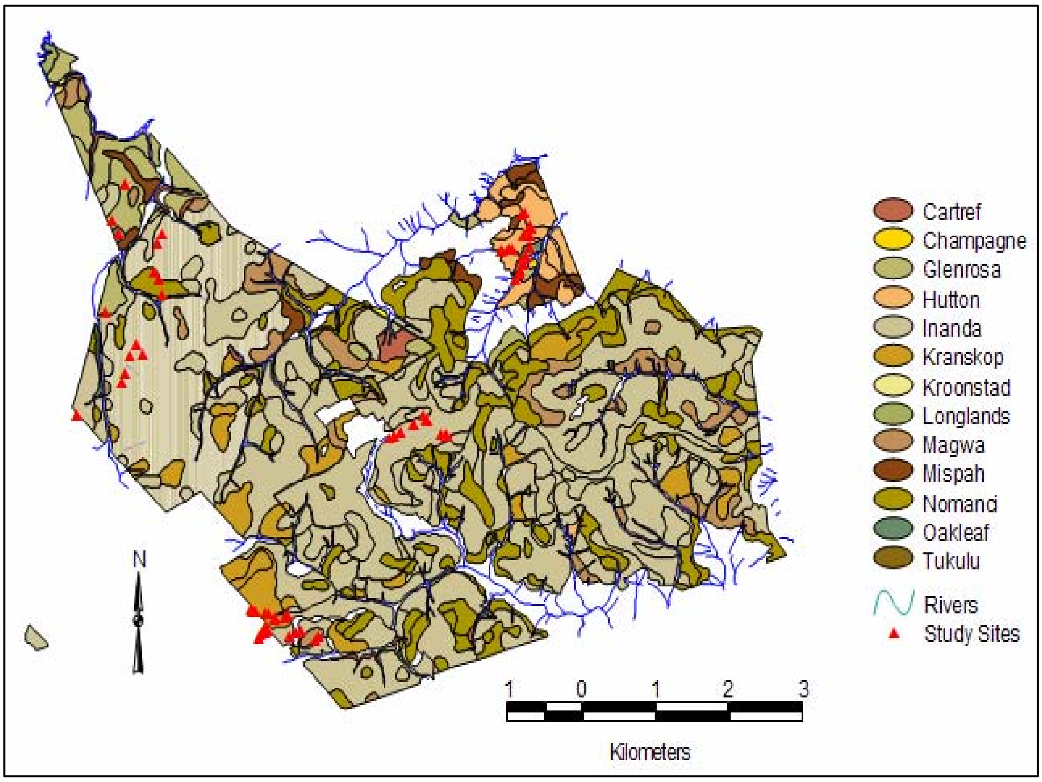



Figure 2 Soils map for the Mistley-Canema estate

### 3.2 Data Collection

Remote sensing data were acquired over the Mistley-Canema site, with concurrent ground measurements of leaf spectral reflectance required for the calibration and classification of vegetation. Ancillary data collected included geographic coordinates and ground control points, aspect, soils information and biophysical descriptions of each forestry compartment. All data were collected during the winter (July 2006) and summer (February 2007) to account for plant physiological differences over the dry and wet seasons.

#### 3.2.1 Selection of the hyperspectral remote sensing data sources

Hyperspectral remote sensing data can be collected using hand-held, airborne or satellite imagers and sensors. A review of commonly available remote sensing data sources was conducted for this study (Table 1). Even though this project is focussed on the application of hyperspectral data, the review included multispectral data sets which have a higher spectral resolution and are sometimes referred to as pseudo-hyperspectral data. The following criteria were considered when selecting the most appropriate data source(s) for this research: spectral, spatial and temporal resolutions; availability, accessibility and cost of data and, handheld, satellite or airborne data sources.

Airborne hyperspectral campaigns are costly, as there are currently no hyperspectral imagers located in South Africa. Imagers such as Casi or Hypmap are exported from international organizations, therefore the costs of airborne campaigns are easily escalated by equipment, processing and travel costs.

Table 1 A review of commonly available remote sensing data sources

| Sensor | Type | Airborne or Satellite | Spatial Resolution | Spectral Resolution | Cost |
|---|---|---|---|---|---|
| Modis | Multispectral | Satellite | 250 To 1000 m | 36 Bands | Free |
| Aster | Multispectral | Satellite | 15-90 m | Total of 14 bands VIS, SWIR,TIR | Free |
| Chris Proba | Multispectral | Satellite | 17 m | 36 bands | Free |
| Meris | Multispectral | Satellite | 1150 km, global coverage every 3 days | 15 bands 400 nm–1050 nm | Free |
| Hyperion | Hyperspectal | Satellite | 30 m | 220 Bands 400-2500 nm | $750 Plus $250 For Corrected Data |
| Hymap | Hyperspectal | Airborne | 3-5 m | 128-200 Bands 400-2500 nm | Based on flight, imager and processing costs |
| Aviris | Hyperspectal | Airborne | 17 m | 224 Bands 400-2500 nm @ 10 nm Intervals | $500 For Archived Data |
| Casi | Hyperspectal | Airborne | Depending On Aircraft Altitude | 288 Bands 400-1000 nm @1.8 nm Intervals | Based on flight, imager and processing costs e.g. approx. 0.73$ per hectare for imager costs only |
| Hydice | Hyperspectal | Airborne | 1-4 m Depending On Aircraft Altitude | 210 Bands 400-2500 nm | Based on flight, imager and processing costs |
| Analytical Spectral Devices Spectrometer | Hyperspectral | Hand-held | N/A | 2150 bands 350-2500 nm | Approx. R5000 p/d (for ASD with operator) |

It was concluded from this review, that the most appropriate data sources for this project is hyperspectral hand-held and satellite data, resulting in the selection of the

Hyperion sensor (United States Geological Survey Earth Resources Observation Systems) and the ASD spectrometer. In addition, a special motivation for earth observation data was made to the European Space Agency (ESA) through the Tiger Initiative which is currently underway, to support earth observation research in Africa. The project team was fortunate to receive support from ESA for free pseudo-hyperspectral Proba Chris data in this research project.

### 3.2.2 Satellite remote sensing data acquisition

During winter 2006, both Hyperion and Proba Chris imagery were acquired. The Hyperion image was acquired on the 12 July 2006, with less than 25% cloud cover. The Hyperion image was a 42 by 7.7 km scene displaying a 30 m spatial resolution, comprising 242 spectral bands ranging from 357 to 2576 nm. Two Proba Chris images were acquired on the 20 and 22 July 2006 using pre-selected spectral bands, which are commonly used when vegetation studies at high spatial resolutions are intended. A Chris image is a 13 by 13 km scene, at a spatial resolution of 17 m, comprising 18 spectral bands ranging from 485.6 to 796.1 nm. To complement the spatial data collected during the winter season, two earth observation datasets, Hyperion and Proba Chris were planned for the summer field campaign. However, all three of the Hyperion images acquired in January and February 2007 contained a high percentage of cloud cover (greater than 20%) and therefore, could not be utilised in the project. Proba CHRIS images were obtained on the 13 February 2007 for the summer campaign.

### 3.2.3 Ground-truthing surveys

The winter and summer field campaigns took place on the 12$^{th}$, 25$^{th}$ and 26$^{th}$ July 2006 and, the 6$^{th}$, 7$^{th}$ and 8$^{th}$ of February 2007 respectively. The sampling method undertaken involved identifying each tree species, their respective age classes and location within the forestry compartments. On the 1st day of each campaign, sample trees were demarcated to allow for efficient ASD spectral reflectance sampling on the subsequent days (Analytical Spectral Devices Inc, 5335 Sterling Drive, Suite A, Boulder, Colorado 80301, USA). On the 2$^{nd}$ and 3$^{rd}$ day of each campaign the project team split into three groups to undertake the following tasks: locate sample trees within demarcated compartments, fell trees where necessary and collect samples for ASD spectral reflectance measurements; record geographic coordinates of sample trees using a global positioning system (Trimble Navigator GPS); collect ASD spectral reflectance signatures from calibration sites; record geographic coordinates of calibration sites; record ancillary data which included the collection of soil samples at each sample tree from the top 10 cm of the A horizon using a bucket auger; record geographic coordinates of ground control points (GCPs) required for geometric corrections e.g. tar roads, bare fields, soccer field, grassland.

The compartments selected for sampling during the winter field campaign included two four year old *Saccharum* sites, two six year old *A. mearnsii* sites, a twelve, fifteen, and sixteen year old *P. patula* site, a ten and twelve year old *E. macarthurii* site and, an eight year old *E. grandis* site (Table 2). The same sampling sites were used in the summer field campaign with the exception of the sixteen year old *P. patula* which was felled prior to the summer field campaign. A random sampling method was used across the compartments, with 10 samples selected in each compartment. In addition calibration sites were used for atmospheric correction.

Table 2 Sample sites selected for ASD leaf spectral reflectance measurements and calibration sites during winter 2006

| **Site Name** | **Scientific name** | **Block Number** | **Area (ha)** | **Latitude (S)** | **Longitude (E)** | **Elevation (m)** | **Aspect** | **Age** | **Open/Closed Canopy** |
|---|---|---|---|---|---|---|---|---|---|
| Sugar cane | *Saccharum* | N/A | N/A | 29°12' 55.0" | 28 ° 39’ 3.1’’ | 1099 | NE | N/A | closed |
| Sugar cane | *Saccharum* | S015 | 8.3 | 29°11' 45.1" | 29 ° 37’ 52.2’’ | 1114 | E | 4 | closed |
| Wattle | *Acacia mearnsii* | C009 | 29.3 | 29°12' 49.4" | 30 ° 38’ 57’’ | 1102 | NW | 6 | closed |
| Wattle | *Acacia mearnsii* | C015 | 38 | 29°12' 52.24" | 30 ° 39’ 7.3’’ | 1073 | NW | 6 | closed |
| Pine | *Pinus patula* | B035 | 8 | 29 ° 11’ 21.7’’ | 30 ° 37’ 45.5’’ | 1066 | W | 15 | closed |
| Pine | *Pinus patula* | B019 | 9.2 | 29 ° 10’ 59.3’’ | 30 ° 37’ 52.7’’ | 1005 | W | 16 | closed |
| Pine | *Pinus patula* | B008 | 13.6 | 29 ° 10’ 12.3’’ | 30 ° 38’ 9.6’’ | 1024 | NW | 12 | closed |
| Gum/Euc | *Eucalyptus macarthurii* | B011 | 31 | 29 ° 10’ 13.2’’ | 30 ° 38’ 18.3’’ | 1000 | W | 12 | closed |
| Gum/Euc | *Eucalyptus macarthurii* | B030 | 7.7 | 29 ° 11’ 24.9’’ | 30 ° 38’ 35’’ | 1045 | NW | 10 | closed |
| Gum/Euc | *Eucalyptus grandis* | F008 | 33.5 | 29 ° 10’ 22.6’’ | 30 ° 40’ 57.1’’ | 901.5 | N | 8 | closed |

| **Calibration Sites** | **Latitude (S)** | **Longitude (E)** | **Elevation (m)** | **Additional Information** |
|---|---|---|---|---|
| Green Field | 29°09' 44" | 30°37' 50.2" | 992 | Irrigated field |
| Tar Road | 30°12' 56.2" | 30°37' 54.8" | 1116 | At the entrance to the Mistley-Canema Estate |
| Bare Soil | 29°11' 45.1" | 30°40' 21.2" | 1114 | Near the tower, near the sugarcane |
| Grass at Block F008 | 29°10' 23.4" | 30°40' 55.7" | 922 | *Themeda triandra* grass near the *E. Grandis* in block F008 |
| Bare Soil | 29°12' 55.0" | 30°39' 2.7" | 1103 | Near the wattle and sugarcane at block C009 |
| Bare Soil | 29°12' 48.6" | 30°38' 56.9" | 1104 | Opposite A. mearnsii near block C015 |
| Dark Soil | 29°11' 36.8" | 30°37'40.4" | 1069 | Felled area at the corner of block C009 |
| Tower | 29°11' 37.5" | 30°40'12.9" | 1133.5 | Near the Sugar cane at block S015 |

The ASD FieldSpec Pro spectrometer being a portable field instrument with a 25º field-of-view was used to obtain quantitative measurements of radiant energy easily and efficiently. Detailed model specifications are reported in FieldSpec Pro user guide (Analytical Spectral Devices Inc., 2002). Illustrations of ASD hand-held spectral reflectance measurements of wattle leaves, pine needles and sugarcane are shown in Figure 3 respectively.

Figure 3 Hand-held ASD spectral reflectance measurements from wattle leaves, pine needles and sugarcane.

The ASD spectrometer measures contiguous spectral signatures from 350 to 2500 nm. Due to atmospheric effects, the instrument was calibrated with a white reference at each site. All spectral signatures were measured between 10 am and 2 pm, under clear sunlit conditions. To minimise differences in background reflectance, leaf-bearing branches were bunched together to permit a solid target of overlapping leaves to be presented to the field spectrometer. Under conditions of poor sunlight resulting from shade or within closed canopy areas, branches were cut and immediately overlaid and sampled under bright sunlight. When leaves where too high, trees were felled at the time using a chainsaw, and thereafter branches were cut and immediately sampled. To obtain a representative leaf spectral measurement from each tree, five measurements each consisting of 5 replicate spectrometer readings were taken.

### 3.3 Pre-processing of remote sensing data

Pre-processing of the acquired remote sensing data involved geometric and atmospheric correction and, data reduction involving two methods as detailed in the subsequent sections.

#### 3.3.1 Geometric correction

The remote sensing data were geometrically corrected using GCPs collected during the field campaigns and a geo-rectified spatial coverage of the forestry compartments obtained from Mondi. A Hyperion product which was georeferenced and orthorectified to UTM 36 S, was obtained from the USGS during July 2006. However, upon analysing the data it was discovered that this image still contained a 400 m geometric shift and therefore required further geometric corrections. All images were projected to UTM 36S, datum WGS 84 using Erdas remote sensing software.

#### 3.3.2 Atmospheric correction

The conversion of radiance into reflectance data is of importance when comparisons are to be made between datasets. The conversion involved the removal of atmospheric absorptions and scattering as well as removal of the shape of the solar irradiance spectrum. Atmospheric and radiometric corrections were performed using the ENVI remote sensing software. Different methods which were investigated included internal average reflectance, flat field, empirical line and the FLAASH atmospheric correction model.

Interference effects of the atmosphere have resulted in significant noise levels in both datasets. The empirical line method was utilized to atmospherically correct both the Hyperion and Proba Chris datasets. This method compares radiance values coming back from the surface to reflectance values measured on the ground with a calibrated hand-held spectrometer (Research Systems Inc, 2005). Using several ground truth data targets the relationship between radiance at sensor and reflectance on the ground can be extracted. Since the effect of the atmosphere is multiplicative (by gasses) and additive (by aerosols) linearity is assumed in each wavelength (i.e. image layer) and a gain and offset are used as estimates of these atmospheric effects on radiance. After calculating these for all wavelengths the gains and offsets could be applied on the image as a whole and the reflectance in all pixels could be estimated (Research Systems Inc, 2005). Calibration sites which were collected using the ASD spectrometer during the winter and summer field campaigns were used to convert radiance into reflectance data.

### 3.3.3 Data reduction

Data reduction methods are used to manipulate high spectral resolution datasets, to reduce high data dimensionality and the complexity which arise when applying statistical classifiers to such datasets. The hyperspectral and multispectral datasets were reduced using analysis of variance. Furthermore, a second data reduction method was investigated to assess the affect of using different hyperspectral bands for the classification.

#### 3.3.3.1 Analysis of variance

The data reduction procedures as outlined by Ismail *et al.*, 2007 were adopted as the first method to obtain an optimal set of spectral bands for both remote sensing datasets. In this method it was assumed that a distinct collection of spectral bands can be used to discriminate between genus/species. Thirty spectral sample points per genus/species were used to determine the significant differences between spectral wavelengths using a one-way analysis of variance (95% confidence limit) with Tukey's Post hoc test for twenty one class pairs (i.e. Euc-Euc8, Euc-Euc10, Euc-Pine12, Euc-Pine15, Euc-Wattle, Euc-Sugar, Euc8-Euc10, Euc8-Pine12, Euc8-Pine15, Euc8-wattle, Euc8-Sugar, Euc10-Pine12, Euc10-Pine15, Euc10-Wattle, Euc10-Sugar, Pine12-Pine15, Pine12-wattle, Pine12-sugar, Pine15-Wattle, Pine15-Sugar, Wattle-Sugar). The number of significant bands at each wavelength for all classes were summed and plotted on a histogram to indicate the frequency of significant bands for all class pairs (Figure 4) as well as individual classes (Figures 5 to 8) and, to discriminate spectrally between classes. The frequency analysis highlighted significant regions of the electromagnetic spectrum for both the Proba Chris and Hyperion imagery *viz*. NIR region (700 nm to 900 nm) and IR and SWIR region (900 nm to 1400 nm) respectively.

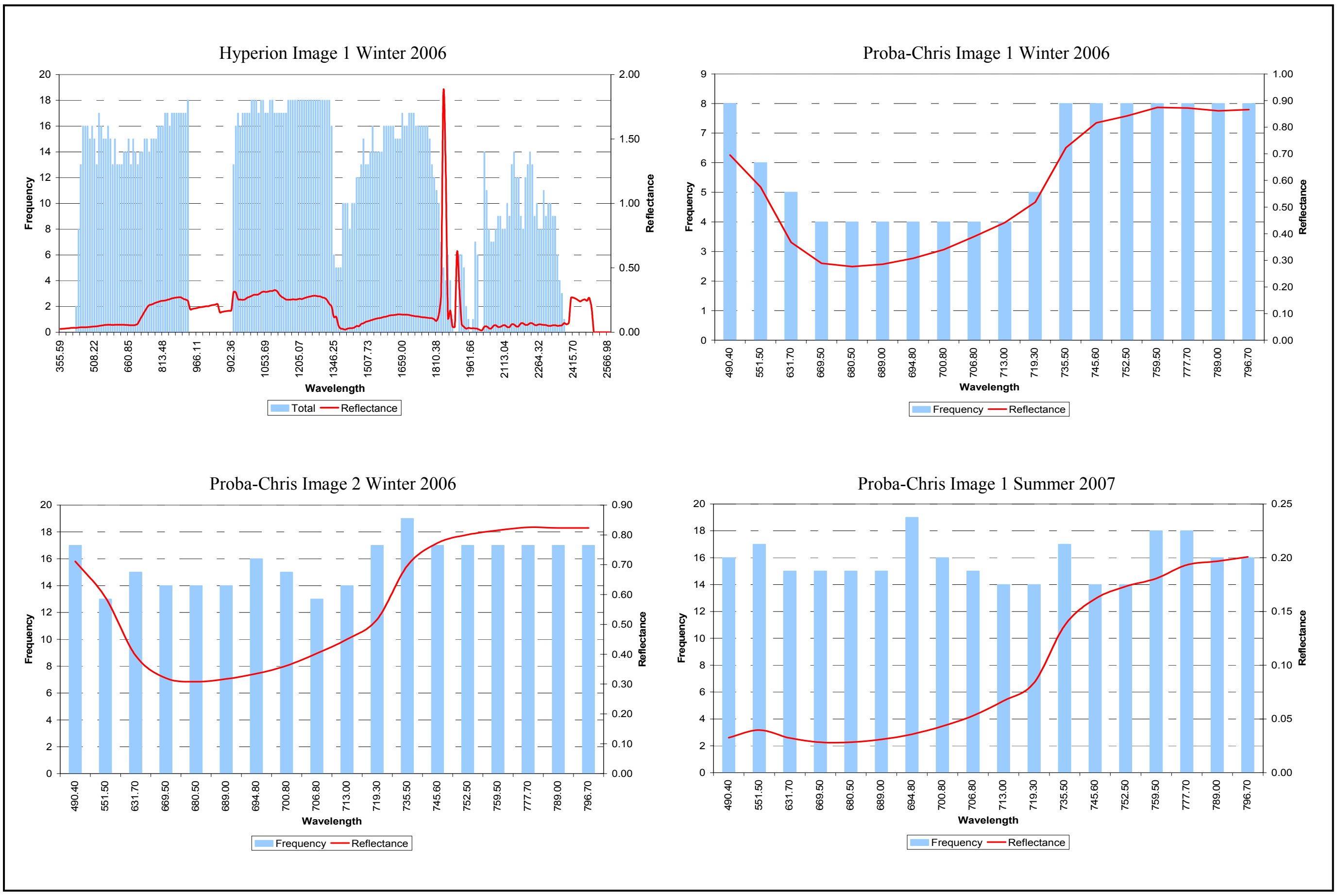



Figure 4 Frequency distribution of the significant spectral bands and the average spectral reflectance signature for all class pairs extracted from Hyperion and Proba Chris imagery.

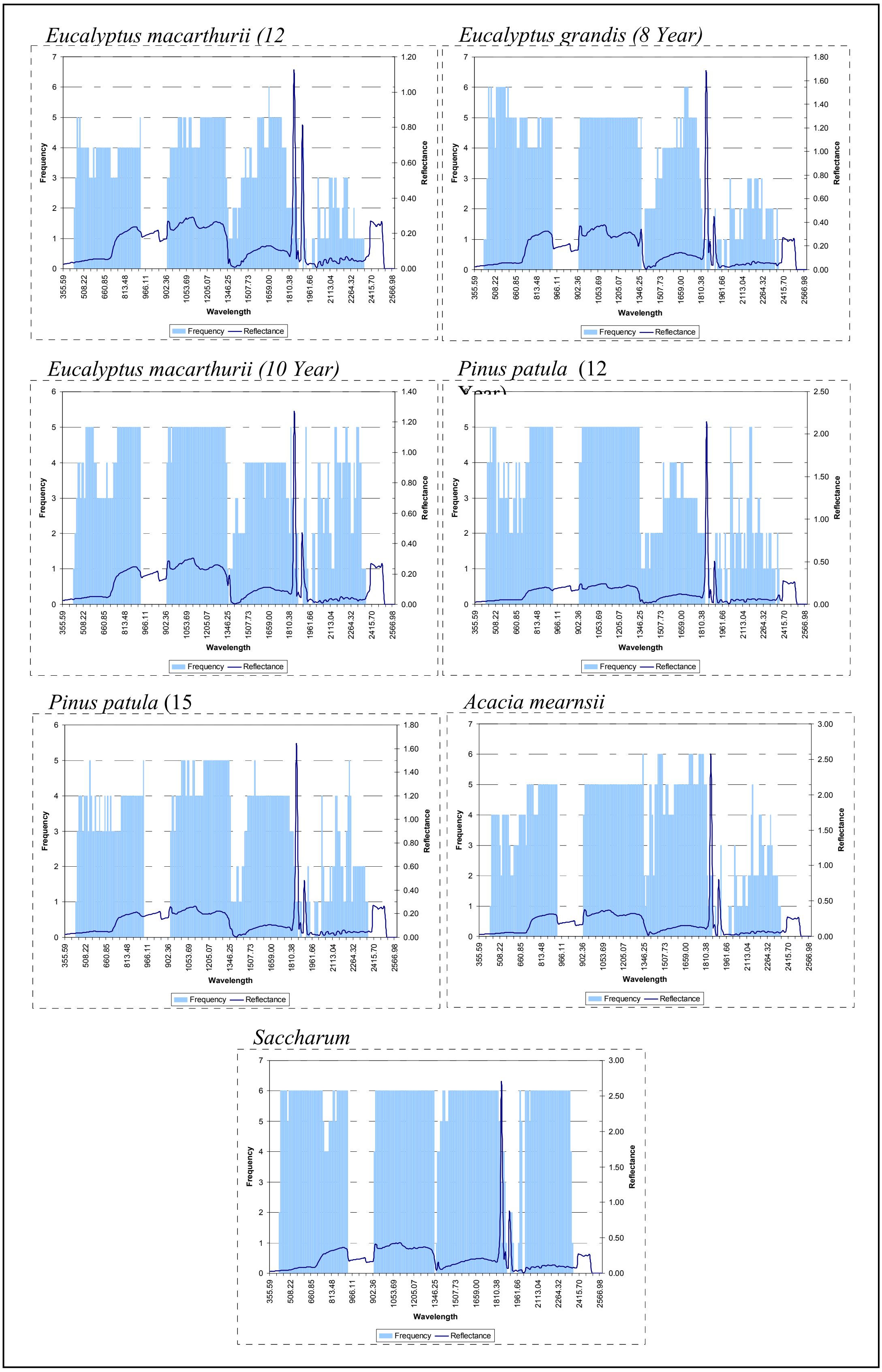



Figure 5 Frequency distribution of the significant spectral bands and the average spectral reflectance signature for individual classes extracted from Hyperion image acquired in winter 2006.

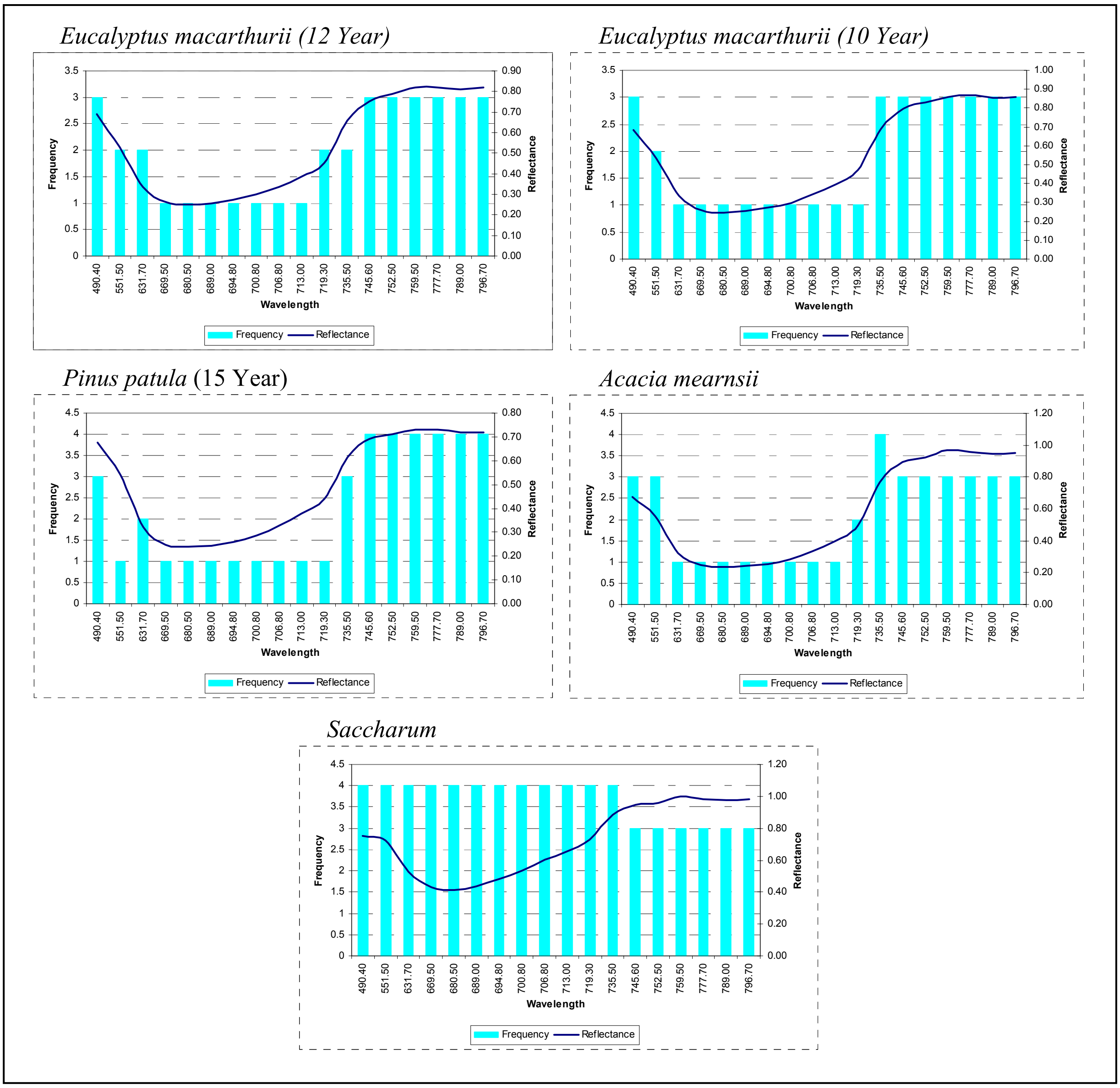



Figure 6 Frequency distribution of the significant spectral bands and the average spectral reflectance signature for individual classes extracted from Proba Chris image 1 acquired in winter 2006.


Figure 7 Frequency distribution of the significant spectral bands and the average spectral reflectance signature for individual classes extracted from Proba Chris image 2 acquired in winter 2006.


Figure 8 Frequency distribution of the significant spectral bands and the average spectral reflectance signature for individual classes extracted from Proba Chris image acquired in summer 2007.

The Jeffries-Matusita separability index was not utilized to test for band seperability as the analysis of variance methodology sufficiently reduced the data dimensionality of both the hyperspectral and multispectral datasets (Richards and Jia, 1999; Vaiphasa *et al*., 2005; Ismail *et al*., 2007). The most significant spectral bands with the highest frequencies were selected for each individual class. The Hyperion hyperspectral dataset reduced significantly from 242 to 27 spectral bands (referred to as Hyperion reduced set A), ranging from 900 to 1400 nm, while the Proba Chris multispectral data reduced from 18 to 9 spectral bands, ranging from 450 to 550 nm and 700 to 800 nm. All geometrically and atmospherically corrected imagery were spectrally subset to the significant spectral bands obtained from the frequency analysis.

#### 3.3.3.2 Minimum noise fraction and band max

The second method applied only to the hyperspectral data, used minimum noise fraction to segregate noise in the datasets, identify and remove these "noisy bands" and to reduce the computational requirements for analyses as adopted by (Green *et al*., 1988). In addition to removing the "noisy bands" from the Hyperion dataset, an optimal subset of bands was determined using known regions of interest for vegetation classes (targets) and background materials in the image. The process is used to isolate the spectral bands which best distinguish selected vegetation classes (targets) from background materials which were verified during field surveys. The optimal spectral subset therefore includes a subset of bands in which "noise" had been removed and which best characterized the vegetation classes selected for this study. The Hyperion hyperspectral dataset reduced significantly from 242 to 23 spectral bands (referred to as Hyperion reduced set B), ranging from 793 nm to 2375 nm, as listed in Table 3.

Table 3 Reduced spectral subset for the Hyperion hyperspectral image

| Hyperion Spectral Band Number | Wavelength (nm) |
|---|---|
| 44 | 793.13 |
| 45 | 803.30 |
| 46 | 813.48 |
| 54 | 894.88 |
| 56 | 915.23 |
| 83 | 972.99 |
| 84 | 983.08 |
| 87 | 1013.30 |
| 88 | 1023.40 |
| 94 | 1083.99 |
| 135 | 1497.63 |
| 136 | 1507.73 |
| 137 | 1517.83 |
| 167 | 1820.48 |
| 188 | 2032.35 |
| 189 | 2042.45 |
| 190 | 2052.45 |
| 203 | 2183.63 |
| 204 | 2193.73 |
| 207 | 2224.03 |
| 219 | 2345.11 |
| 220 | 2355.21 |
| 222 | 2375.30 |

## 3.4 Procedures for Classification and Assessment

Two methods of classification were investigated in this study *viz*. vegetation classification using vegetation spectral signatures measured in the field and known regions of interest (ROIs) verified during field surveys. Vegetation classification using hand-held spectral signatures collected during the field campaigns were unsuccessful.

### 3.4.1 Selection of regions of interest

Regions of interest were selected for the training and validation phases performed during supervised classification. Regions of interest may be defined within the context of remote sensing, as areas of an image which contain pixels of the same spectral characteristics which represent the same vegetation type.

Regions of interest were selected using GCPs and compartment boundary overlays of the Mistley-Canema estate, which were verified during field surveys. These regions were selected to represent 8 year old *E. grandis,* 10 and 12 year old *E. macarthurri*, 12 and 16 year old *P. patula*, 4 year old *Saccharum* and 6 year old *A. mearnsii*.

Two statistical measures were used to assess the spectral separability of the selected regions of interest: the Jeffries-Matusita and Transformed Divergence separability measures (Richards and Jia, 1999). The Jeffries-Matusita and Transformed Divergence separability measures range from 0 to 2.0 and give an indication of how well the selected region of interest pairs are statistically separate from each other (Richards, 1999). Values greater than 1.9 indicate that the region of interest pairs have good separability.

All regions of interest compared in Table 4 are seen to have high spectral seperability (values > 1.9). When comparing the regions of interest of the 10 and 12 year old *Eucalyptus* macarthurri, their spectral seperability was slightly reduced to between 1.5 and 1.85. However, a reduced spectral seperability was expected since this investigation involved differentiation between two age classes of the same species.

Table 4 An assessment of the spectral separability between different pairs of regions of interest

| Region of interest pairs | Proba Chris July 2006 | | Proba Chris January 2007 | | Hyperion July 2006 | |
|---|---|---|---|---|---|---|
| | Jeffries-Matusita | Transformed Divergence | Jeffries-Matusita | Transformed Divergence | Jeffries-Matusita | Transformed Divergence |
| 10 year old *Eucalyptus macarthurri* and *8 year old | - | - | 1.99999300 | 2.00000000 | 1.99136500 | 1.99979119 |
| 10 and 12 year old *Eucalyptus macarthurri* | 1.82955160 | 1.95724725 | 1.56849083 | 1.84130076 | 1.97273246 | 1.99999205 |
| 10 year old *Eucalyptus macarthurri* and 6 year old | 1.99995248 | 1.99999712 | 2.00000000 | 2.00000000 | 1.99886583 | 1.99999472 |
| 10 year old *Eucalyptus macarthurri* and 4 year old | 2.00000000 | 2.00000000 | 2.00000000 | 2.00000000 | 1.99977471 | 2.00000000 |
| 10 year old *Eucalyptus macarthurri* and 12 year old | 1.98065921 | 2.00000000 | 1.99999873 | 2.00000000 | 1.99850587 | 2.00000000 |
| 8 year old *Eucalyptus grandis* and 12 year old *Eucalyptus* | - | - | 1.99984939 | 2.00000000 | 1.99997146 | 2.00000000 |
| 8 year old *Eucalyptus grandis* and 12 year old *Pinus patula* | - | - | 1.99997390 | 1.99999925 | 1.99997197 | 2.00000000 |
| 8 year old *Eucalyptus grandis* and 4 year old *Saccharum* | - | - | 2.00000000 | 2.00000000 | 1.99999844 | 2.00000000 |
| 8 year old *Eucalyptus grandis* and 6 year old *Acacia mearnsii* | - | - | 1.99997722 | 2.00000000 | 1.99999195 | 1.99999999 |
| 12 year old *Eucalyptus macarthurri* and 6 year old | 1.99999993 | 2.00000000 | 2.00000000 | 2.00000000 | 1.99978638 | 1.99999976 |
| 12 year old *Eucalyptus macarthurri* and 4 year old | 2.00000000 | 2.00000000 | 2.00000000 | 2.00000000 | 1.99997202 | 1.99999999 |
| 12 year old *Eucalyptus macarthurri* and 12 year old | 2.00000000 | 2.00000000 | 1.99996885 | 1.99999971 | 1.99284234 | 1.99999996 |
| 6 year old *Acacia mearnsii* and 4 year old *Saccharum* | 2.00000000 | 2.00000000 | 2.00000000 | 2.00000000 | 1.97101221 | 1.99728235 |
| 6 year old *Acacia mearnsii* and 12 year old *Pinus patula* | 2.00000000 | 2.00000000 | 1.97382987 | 1.99998401 | 2.00000000 | 2.00000000 |
| 4 year old *Saccharum* and 12 year old *Pinus patula* | 2.00000000 | 2.00000000 | 2.00000000 | 2.00000000 | 1.99999962 | 2.00000000 |

* 8 year old *Eucalyptus grandis were unclassified in the winter 2006 Proba Chris imagery due to intense cloud cover over*

### 3.4.2 Training and validation phase of classification

Training regions were spatially subset from the Hyperion and Proba Chris images.
In this study five statistical classifying techniques were used in the training phase of the classification process. These techniques included maximum likelihood, minimum distance, mahalanobis distance, spectral angular mapper and parallelepiped. The ENVI 4.3 software was used to perform each statistical classifying method. An iterative process was followed to determine the optimized parameters required for

each statistical classifier. Parameter optimization was only required for the parallelepiped and minimum distance classifiers.

Accuracy assessments were performed for each of the classification techniques to determine which statistical classifier produced best classification results. According to Thomlinson *et al*., (1999) an overall accuracy of greater than 85% is an indicator of superior classification. However, in this study only those statistical classifying techniques with an overall accuracy greater than 80% during the training phase were selected for validation.

For validation the best statistical classifying methods were used to classify the entire study area. The kappa coefficient of agreement and the overall accuracy (> 80%) were the two statistics most prominently used for decision making purposes (Foody, 2006) and to assess the validated areas. The user's accuracy was used to evaluate how well an individual vegetation class was classified for a specific statistical classifier method.

### 3.4.3 Classification using vegetation indices

Vegetation indices or band ratios have been derived to emphasise different plant physiological features and have been used to distinguish between different vegetation within a mosaic of other landuses (McGwire *et al.,* 2000; Underwood *et al.,* 2003; Yamano *et al.,* 2003; Galvao *et al.,* 2005). Vegetation indices are also used to mask out vegetated areas which are then used in the classification process (Underwood *et al.,* 2003; Yamano *et al.,* 2003; Kotch *et al.,* 2005). In this study the use of vegetation indices to classify and map *Acacia mearnsii, Pinus patula, Eucalyptus macarthurii, Eucalyptus grandis* and *Saccharum* using hyperspectral data was investigated. The following commonly used vegetation indices which have been linked to different plant physiological characteristics were investigated: carter index, inverse carter, normalised difference infrared index (NDII), normalised difference vegetation index (NDVI), normalised difference water index (NDWI), pigment index, red-green ratio, red edge normalised difference vegetation index (RENDVI), Vogelmann index (VOGa), Vogelmann index (VOGc) and the waterband index (WBI) listed in Table 5.

Vegetation indices were derived for the Hyperion and Proba Chris datasets. Vegetation index layers were classified using the parallelepiped and minimum distance classifiers. Accuracy assessments were performed for each of the classification techniques to determine which statistical classifier produced best classification results. An overall accuracy greater than 80% during the training phase was selected for validation.

Table 5 Vegetation indices

| Physiological/ Biophysical Variables | Vegetation Spectral Index | Specific Wavelengths (nm) | Reference | Chris band selection (nm) | Hyperion 2006 band selection (nm) |
|---|---|---|---|---|---|
| Plant pigments | Pigment Index | (750/710) | Gamon and Surfus (1999) | (752.5 / 713.0) | (752.43 / 711.72) |
| Plant pigments | Carter Index | (695/760) | Carter (1994) | (694.8 / 759.5) | (691.37 / 762.60) |
| Plant pigments | Inverse Carter Index | (760/695) | Carter (1994) | (759.5 / 694.8) | (762.60 / 691.37) |
| Robust vegetation index derived from the highest absorption and reflectance regions of chlorophyll | Normalised Difference Vegetation Index (NDVI) | (750-676)/ (750+676) | Tucker (1979) | (572.5-680.5)/ (572.5+680.5) | (752.53-671.02)/ (752.53+671.02) |
| Modified NDVI for high spectral resolution data | Red Edge Normalised Difference Vegetation Index (RENDVI) | (750-705)/ (750+705) | Gitelson and Merzlyak (1994); Sims and Gamon (2002) | (752.5-706.8)/ (752.5+706.8) | (752.60-701.55)/ (752.60+701.55) |
| Plant water status | Water Band Index (WBI) | (900/970) | Penuelas *et al.*, (1995) | N/A | (905.05 / 966.11) |

| | | | | | |
|---|---|---|---|---|---|
| Plant water status | Normalised Difference Water Index (NDWI) | (857-1241)/ (857+1241) | Gao (1995) | N/A | (854.18-1245.36)/ (854.18+1245.36) |
| Plant water status | Normalised Difference Infra-red Index (NDII) | (819-1649)/ (819+1649) | Hardisky (1983); Jackson *et al*., (2004) | N/A | (823.65-1648.90)/ (823.65+1648.90) |
| Chlorophyll | Red/Green ratio | (683/510) | Gamon and Surfus (1999) | (680.5 / 490.4) | (681.20 / 508.22) |
| Foliage chlorophyll concentration, canopy leaf area, and water content | Vogelmann Index (VOGa) | (740/720) | Vogelmann *et al*., 1993 | (745.60 / 719.30) | (742.25 / 721.90) |
| Foliage chlorophyll concentration, canopy leaf area, and water content | Vogelmann Index (VOGc) | (734-747)/ (715+720) | Vogelmann *et al*., 1993 | (735.5-745.6)/ (713.0+719.3) | (732.07-742.45)/ (711.72+721.90) |
| LAI | Soil adjusted vegetation index (SAVI) (1+L)(NIR-R)/(NIR+R+L) L=0.5 | (1+0.5)(750-676)/ (750+676+0.5) | Schultz and Engman (2000) | (1.5)(752.5-680.5)/ (752.5+680.5+0.5) | (1.5)(752.43-671.02)/ (752.43+671.02+0.5) |

# 4. RESULTS AND DISCUSSION

## 4.1 Spectral Differences in Vegetation Classes Identified from ASD hand-held measurements

It is important to understand the basic spectral principles of vegetation before proceeding to canopy and landscape scale applications. Leaf-level reflectance measurements having many narrow spectral bands extending across the visible (VIS), near-infrared (NIR) and shortwave infra-red (SWIR) regions of the electromagnetic spectrum are characteristic of hand-held hyperspectral ASD measurements. The collection of leaf-level hyperspectral reflectance signatures can also be categorised as "best case scenario" hyperspectral data A typical vegetation spectral reflectance signature is illustrated in Figure 9. Plant physiological and structural properties such as plant pigments, cell structure and water content are typically related to the visible, near infrared and short wave infrared regions respectively, which certainly affect vegetation classification using earth observation data.

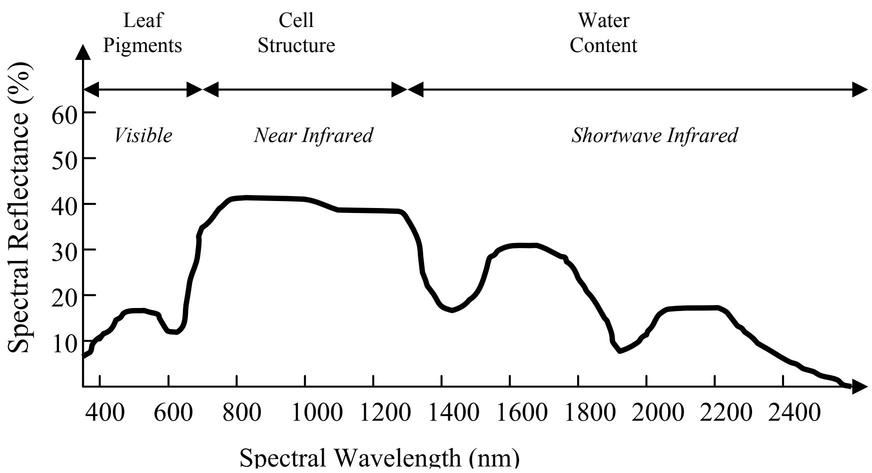



Figure 9 Typical reflectance sensitivities as controlled by leaf pigments, cell structure and water content

The mean reflectance spectra obtained for all tree species sampled in the Mistley-Canema estate during July 2006 are shown in Figure 10. Differences in amplitude between reflectance spectral of the differences tree species sampled using ASD spectrometer are illustrated in Figure 10. Furthermore, the 'noise' peaks represent the atmospheric water absorption bands found between 1820 to 1940 nm in the short wave infrared region of the spectrum.

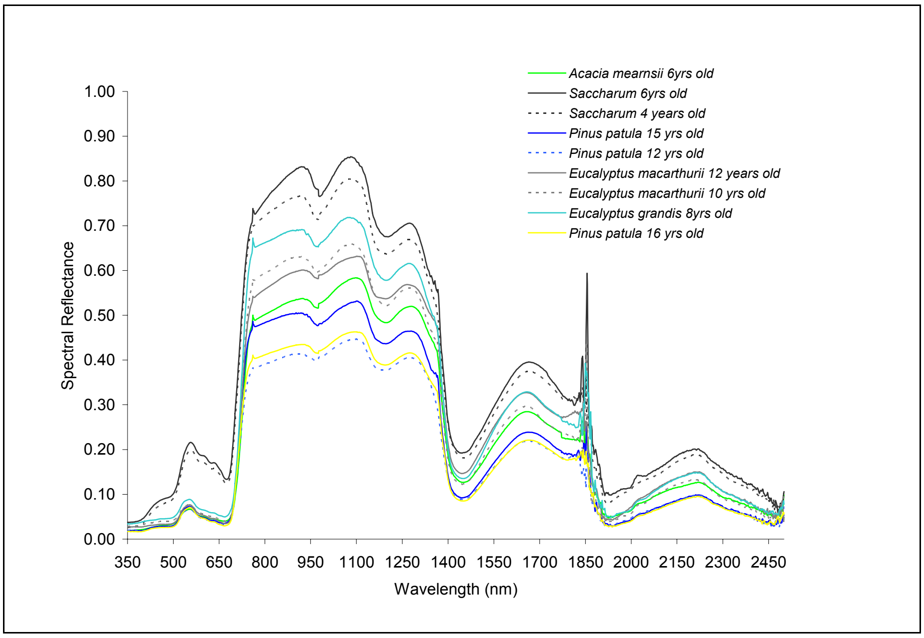



Figure 10 The average reflectance spectra sampled from 350 to 2450 nm for the different tree species sampled at Mistley-Canema

Spectral differences which manifest in the visible and near infrared regions of the spectrum are often used to describe different vegetation classes. Reflectance and first derivative spectra extending from 350 to 700 nm are shown in Figure 11. First derivative spectra were used to enhance the shape differences between the spectral signatures for each tree species sampled. Figure 10 and 11 demonstrate the spectral reflectance differences between tree species sampled within the Mistley-Canema estate. By accurately capturing these spectral differences, various vegetation indices or band ratios could be investigated to improve classification of the selected vegetation classes.

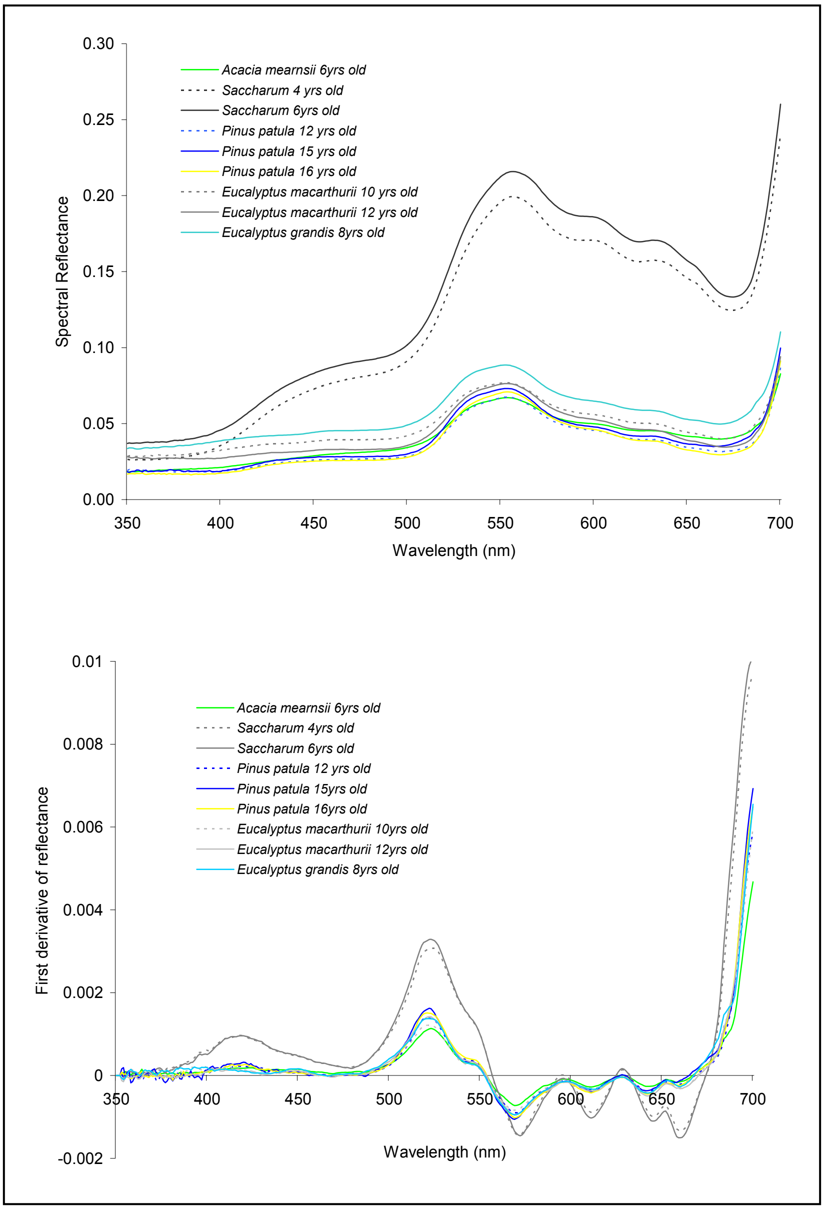



Figure 11 Mean reflectance and first derivative spectra extending from 350 to 700 nm for different species sampled at Mistley-Canema

Vegetation spectral reflectance indices or band ratios are often used to emphasize different physiological, biochemical or structural properties of vegetation. Vegetation indices are used to accentuate characteristics of vegetation and are now commonly being applied in different environmental and hydrological studies to predict parameters such as plant water content, leaf area, pigment concentrations, biomass and plant water stress. Numerous vegetation spectral indices have been reported in scientific literature, but only a selected number of indices are commonly used and have been tested for different vegetation types.

Eleven vegetation spectral reflectance indices which are commonly used to detect physiological, structural and biochemical variability amongst vegetation were derived from the ASD hand-held mean vegetation spectral reflectance signatures (Figure 12). These vegetation indices were used to enhance the spectral differences of *Acacia mearnsii, Saccharum, Pinus patula* and the *Eucalyptus* species sampled in July 2006. Spatial differences in the spectral reflectance properties of these vegetation types were investigated after the winter and summer field campaigns were completed.

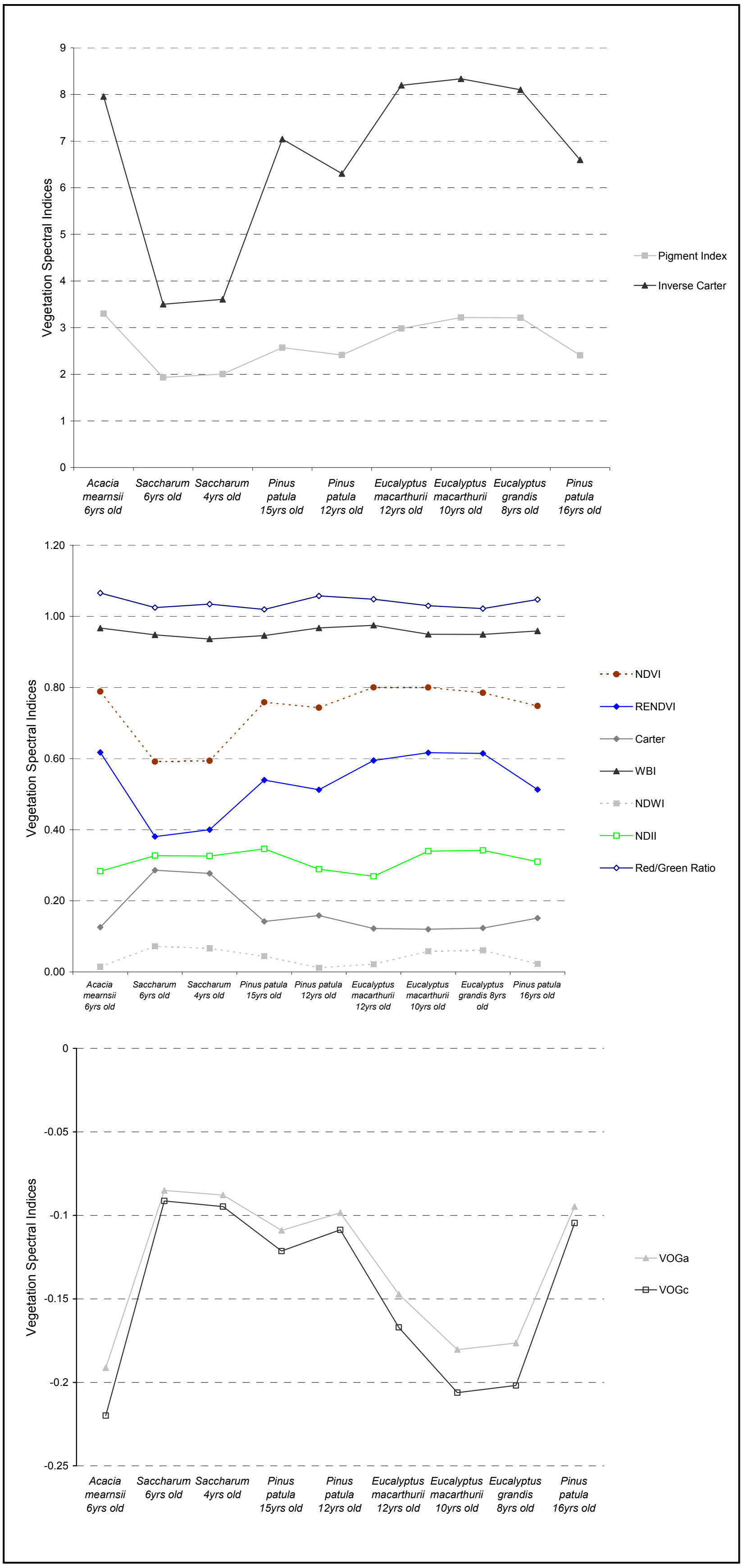



Figure 12 Vegetation indices derived from leaf-level hyperspectral signatures

## 4.2 Comparison of Spectral Signatures extracted from ASD Hand-held and Satellite Data

Vegetation spectral reflectance signatures collected using the ASD hand-held spectrometer and those extracted from the satellite datasets were compared to assess spectral differences captured using different remote sensing technologies. Average vegetation spectral reflectance signatures extracted from each data source for each vegetation class and age group were compared (Figure 13). A common axis was used to group the spectral signatures, as each data source has its own unique collection of pre-selected spectral bands which were sampled at different band widths. Discontinuous spectral signatures were extracted from the Proba Chris datasets, which span from 490 to 800 nm. However, continuous spectral signatures were extracted from the Hyperion hyperspectral satellite data ranging from 355 to 2500 nm at 10 nm bandwidth. Detailed spectral signatures were collected using the ASD spectrometer, with continuous spectral bands ranging from 355 to 2500 nm at a 1 nm bandwidth.

In general, the mean spectral signatures extracted for the different vegetation classes followed the same trend. However, amplitudinal differences exist between the various data sources and, these differences are consistent for each vegetation class. These differences are more pronounced for the spectral bands ranging from 700 nm to 1800 nm extracted from the satellite data are compared with ASD spectrometer signatures. Furthermore, the amplitudinal differences are similar between the satellite datasets when compared to the ASD data. This trend is often associated with atmospheric interference which is highly variable between earth observations and ground level measurements.

Variation in spectral signatures extracted using satellite compared to ASD data is most significant for *Saccharum* (Figure 13g) and least significant for 12 year old *E. macarthurii*, and 15 and 12 year old *P. patula* (Figure 13a, d and e). *Saccharum* leaf blades were sampled non-destructively directly above the canopy whereas *E. macarthurii* and *P. patula* leaves were sampled destructively using the stacking methodology. Therefore, the amplitudinal range between satellite and ASD spectral signatures can be attributed to these different sampling techniques.

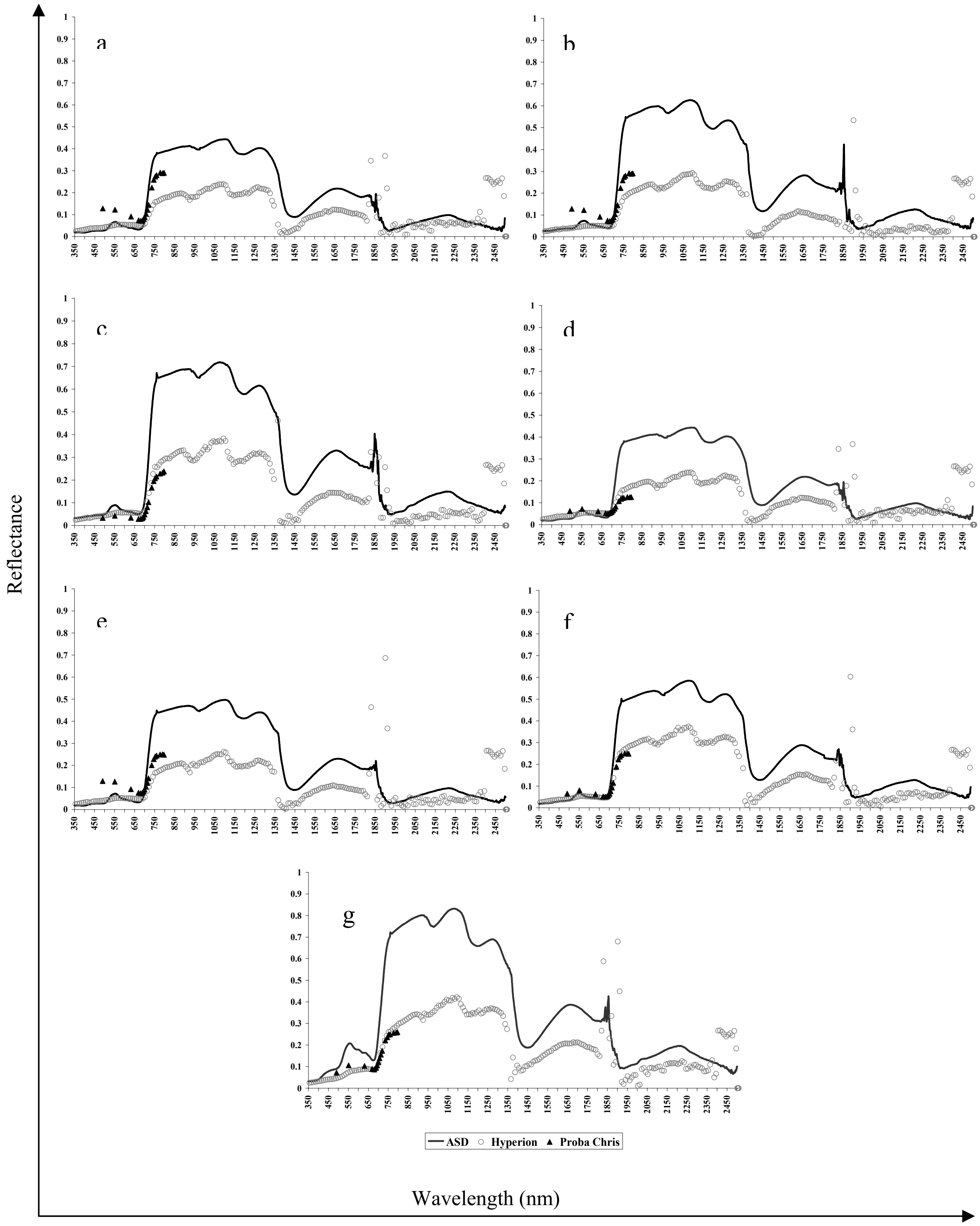



Figure 13 Average vegetation spectral reflectance signatures extracted for a) 12 year old *E. macarthurii*, b) 10 year old *E. macarthurii* c) 8 year old *E. macarthurii*, d) 15 year old *P. patula*, e) 12 year old *P. patula* f) *A. mearnsii* and g) *Saccharum*.

## 4.3 Classification and Mapping of Vegetation Using Known Regions of Interest

The classification and mapping of vegetation using known regions of interest was carried out in two phases *viz*. the training and validation phases. Results of the training and validation phases are detailed in sections 4.3.1 and section 4.3.2 respectively.

### 4.3.1 Training phase

Training regions were spatially subset from the Hyperion and Proba Chris images. Regions of interest were selected to represent 8 year old *E. grandis*, 10 and 12 year old *E. macarthurri*, 12 and 16 year old *P. patula*, 4 year old *Saccharum* and 6 year old *A. mearnsii*. An iterative process was followed to determine the optimized parameters required for each of the five statistical classifier methods. The performance of each statistical classifying method was evaluated using accuracy assessments (>80% overall accuracy) as shown in Table 6. The training results of two Hyperion reduced datasets A (27 spectral bands) and B (23 spectral bands) are shown independently in Table 6. From the training phase, the mahalanobis distance and the maximum likelihood methods most accurately classified the vegetation classes used in this study, achieving overall accuracies above 80%. For all four datasets, the parallelepiped method did not perform well, achieving overall accuracies ranging from 3% to 64%.

Table 6 Accuracy assessment: training phase

| Supervised Classification Methods | Proba Chris (winter 2006 survey) | | Proba Chris (summer 2007 survey) | | Hyperion reduced set A (winter 2006 survey) | | Hyperion reduced set B (winter 2006 survey) | |
|---|---|---|---|---|---|---|---|---|
| | **Kappa Coefficient** | **Overall Accuracy (%)** | **Kappa Coefficient** | **Overall Accuracy (%)** | **Kappa Coefficient** | **Overall Accuracy (%)** | **Kappa Coefficient** | **Overall Accuracy (%)** |
| Parallelepiped | 0.55 | 63.92 | 0.56 | 62.82 | 0.14 | 28.48 | 0.03 | 3.98 |
| Minimum Distance | 0.88 | 90.36 | 0.58 | 65.13 | 0.38 | 46.86 | 0.53 | 59.42 |
| Mahalanobis Distance | 0.90 | 91.91 | 0.91 | 92.75 | 0.83 | 85.45 | 0.78 | 80.86 |
| Maximum Likelihood | 0.99 | 99.22 | 0.97 | 97.48 | 0.99 | 98.93 | 0.99 | 98.93 |
| Spectral Angle Mapper | 0.75 | 75.21 | 0.76 | 77.36 | 0.59 | 64.93 | 0.39 | 48.09 |

The classification of the vegetation classes using the different classifying techniques were further verified by mapping individual vegetation classes across the training area and comparing their spatial distribution to true colour images and spatial maps of forestry compartments obtained from Mondi. Figure 14 illustrates the classification and mapping of the best discriminated vegetation classes using Proba Chris and Hyperion remote sensing imagery. Each classification can be compared to the true colour image shown in Figure 14a. Figures 14b, d, f, h represent the mahalanobis distance classifications using Proba Chris images obtained during winter and summer, and the Hyperion reduced datasets A and B acquired in winter respectively. Figures 14c, e, g, j represent the maximum likelihood classifications using Proba Chris images obtained during winter and summer, and the Hyperion reduced datasets A and B acquired in winter respectively.

For the Proba Chris multispectral images, the mahalanobis classifier consistently classified the *A. mearnsii* and *Saccharum* satisfactorily during winter (Figure 14b) and summer (Figure 14d). Conversely, the maximum likelihood method only classified *A. mearnsii* adequately, for the same set of images (Figure 14c and e). Furthermore, the maximum likelihood method satisfactorily classified the 10 and 12 year old *E. macarthurii* and 12 year old *P. patula* species during winter (Figure 14c); and 12 year old *E. macarthurii* during summer (Figure 14d). Due to *E. macarthurii* compartments being felled prior to the summer campaign, the proportions of this species classified and mapped during winter and summer is variable.

The Hyperion hyperspectral reduced dataset A (27 spectral bands ranging from 900 to 1400 nm) best classified *A. mearnsii* and *Saccharum* using the mahalanobis distance method (Figure 14f); and *A. mearnsii* and the 12 year old *P. patula* using the maximum likelihood method (Figure 14g). Similarly, *A. mearnsii* was satisfactorily classified when applying both statistical classifiers to the reduced Hyperion data set B which comprised of 23 spectral bands ranging from 793 nm to 2375 nm (Figures 14h and j).

The training results imply that classification of different genera and species are affected by optimal band selection and the statistical classifier method adopted. Furthermore, discriminating between different genera and species is influenced by seasonal physiological changes in vegetation which are highlighted during the dry winter and wet summer seasons.

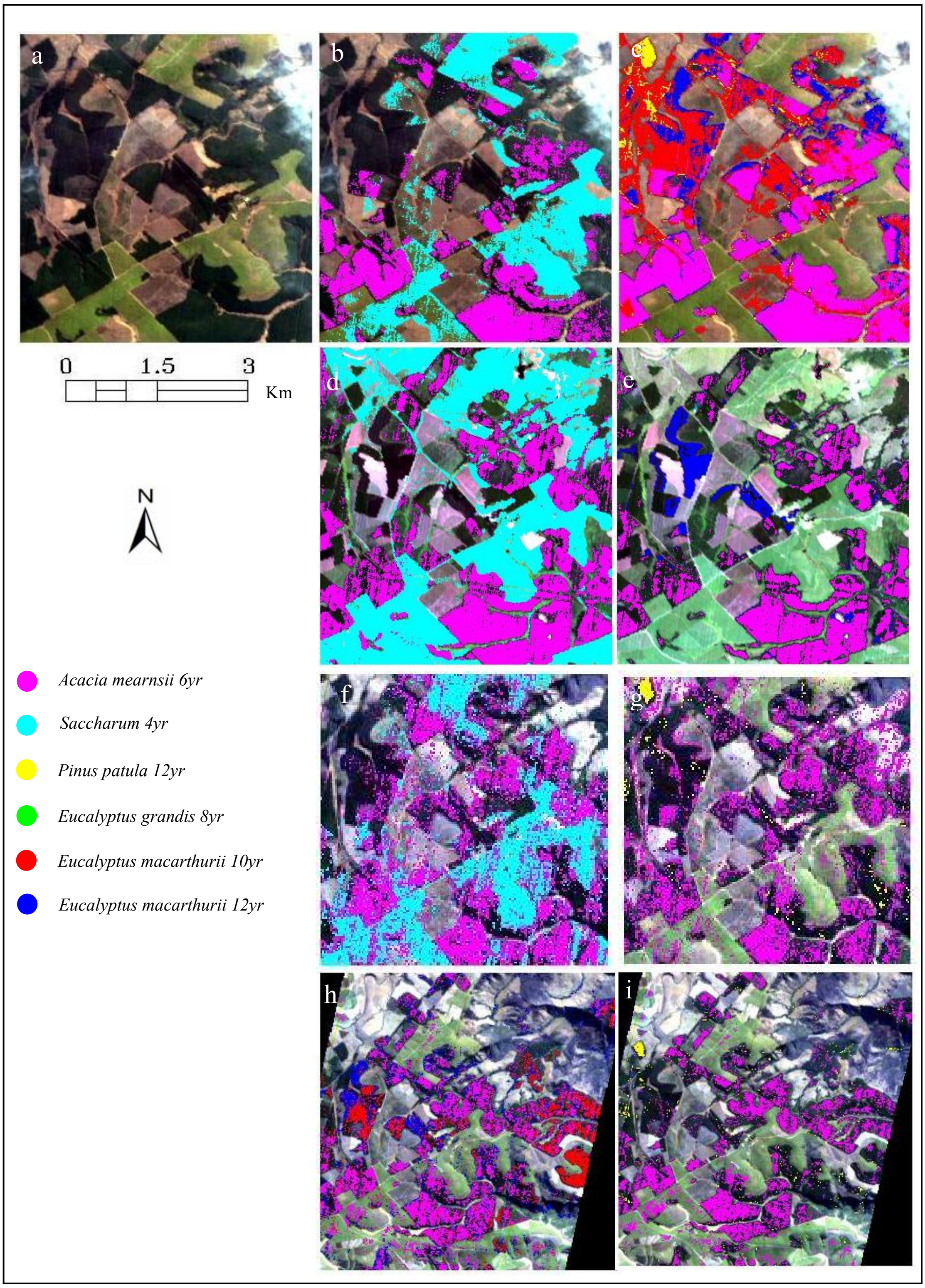



Figure 14 The classification and mapping of the best discriminated vegetation classes using Proba Chris and Hyperion remote sensing imagery during the training phase: a) true colour image, Proba Chris winter classification using the b) mahalanobis distance and c) maximum likelihood classification, Proba Chris summer classifications using d) mahalanobis distance and e) maximum likelihood, Hyperion reduced

set A winter classification using f) mahalanobis distance and g) maximum likehood and, Hyperion reduced set B winter classification using h) mahalanobis distance and i) maximum likelihood methods.

### 4.3.2 Validation phase

The mahalanobis distance and maximum likelihood statistical classifiers were selected from results of the training phase for validation. Validation was performed using the entire study area. Accuracy assessments were used to evaluate each classification as shown in Table 7. Both methods performed well, with overall accuracies above 80%. However, for all four datasets the maximum likelihood classifications achieved the highest overall accuracies above 98%.

Table 7 Accuracy assessments: validation phase

| **Supervised Classification Methods** | **Proba Chris (winter 2006 survey)** | | **Proba Chris (summer 2007 survey)** | | **Hyperion reduced set A (winter 2006 survey)** | | **Hyperion reduced set B (winter 2006 survey)** | |
|---|---|---|---|---|---|---|---|---|
| | **Kappa Coefficient** | **Overall Accuracy (%)** | **Kappa Coefficient** | **Overall Accuracy (%)** | **Kappa Coefficient** | **Overall Accuracy (%)** | **Kappa Coefficient** | **Overall Accuracy (%)** |
| Mahalanobis Distance | 0.83 | 85.89 | 0.90 | 91.60 | 0.88 | 89.13 | 0.78 | 80.86 |
| Maximum Likelihood | 0.97 | 98.80 | 1.00 | 99.86 | 0.98 | 98.68 | 0.99 | 98.93 |

The classifications of each vegetation class at their respective age groups were evaluated to determine how accurately the statistical classifiers, mahalanobis distance and maximum likelihood performed at a species level. This evaluation was undertaken by comparing the user's accuracy for each for each vegetation class (Table 8), with a visual assessment of their spatial distribution.

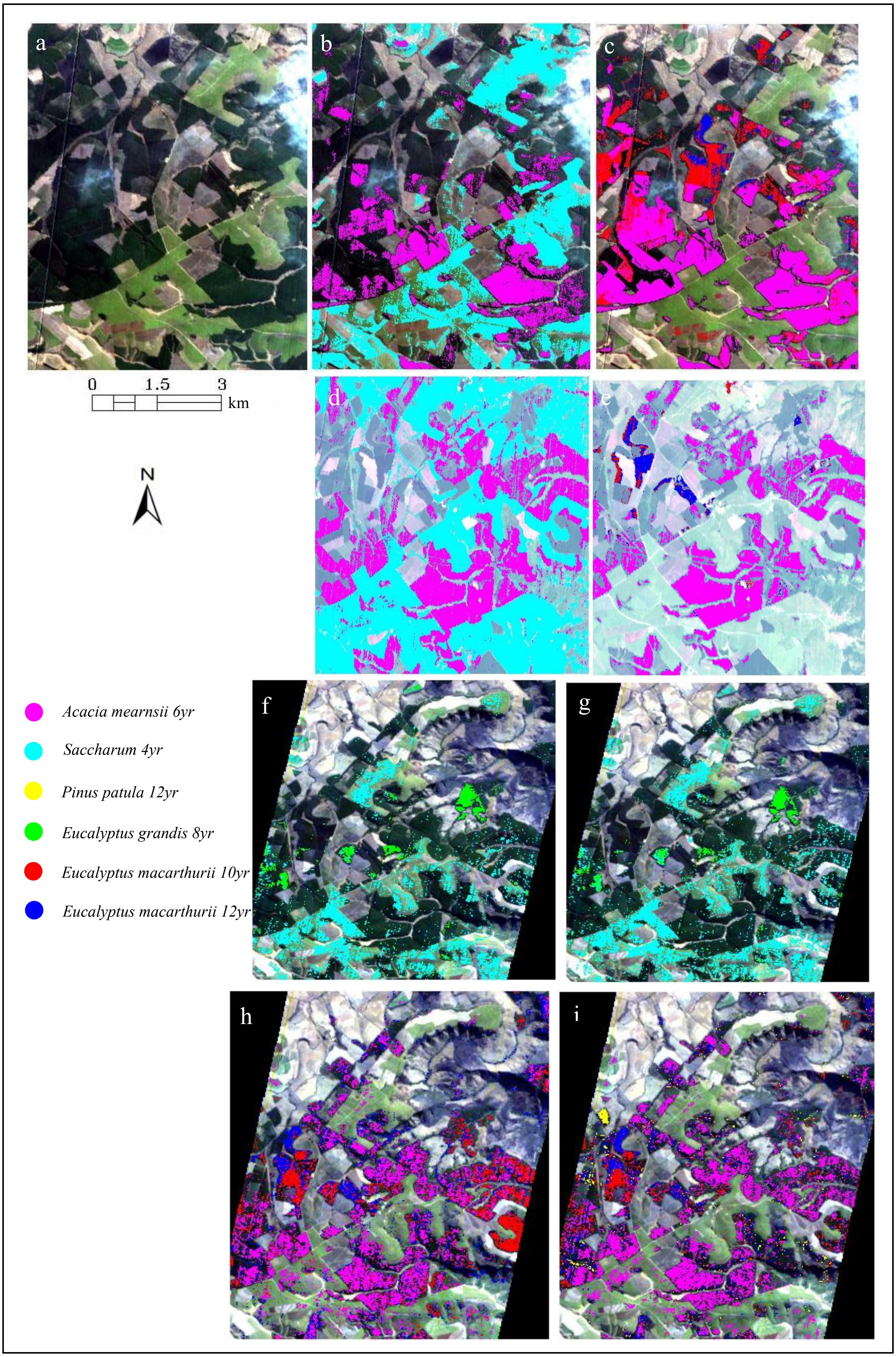
Table 8 User Accuracy: validation phase

| Vegetation Classes | User Accuracy (%) Proba Chris (winter 2006 survey) | | User Accuracy (%) Proba Chris (summer 2007 survey) | | User Accuracy (%) Hyperion reduced set A (winter 2006 survey) | | User Accuracy (%) Hyperion reduced set B (winter 2006 survey) | |
|---|---|---|---|---|---|---|---|---|
| | **Mahalanobis Distance** | **Maximum Likelihood** | **Mahalanobis Distance** | **Maximum Likelihood** | **Mahalanobis Distance** | **Maximum Likelihood** | **Mahalanobis Distance** | **Maximum Likelihood** |
| 8 year old *Eucalyptus grandis* | Cloud cover over site | Cloud cover over site | 100.00 | 100 | 96.39 | 100 | 77.65 | 100 |
| 10 year old *Eucalyptus macarthurri* | 78.40 | 100 | 76.14 | 99.04 | 76.67 | 99.07 | 80.70 | 99.08 |
| 12 year old *Eucalyptus macarthurri* | 76.56 | 96.24 | 68.03 | 100 | 88.64 | 97.73 | 80.46 | 96.63 |
| 12 year old *Pinus patula* | 78.50 | 100 | 100 | 100 | 81.05 | 100 | 75.00 | 98.81 |
| 16 year old *Pinus patula* | 75.00 | 100 | 100 | 100 | 76.14 | 96.91 | 75.64 | 98.95 |
| 4 year old *Saccharum* | 100.00 | 100 | 100 | 100 | 97.67 | 100 | 97.80 | 100 |
| 6 year old *Acacia mearnsii* | 100.00 | 100 | 100 | 100 | 86.02 | 98.92 | 78.57 | 98.82 |

The classifications from the validation phase were mapped across the Mistley-Canema estate. Figure 15 illustrates the best discriminated vegetation classes when applying the mahalanobis distance and maximum likelihood classifiers for the Proba Chris and Hyperion remote sensing imagery. These classifications can be compared to true colour images in Figure 15a. Figures 15b, d, f, h represent the mahalanobis distance classifications using Proba Chris images obtained during winter and summer, and the Hyperion reduced sets A and B acquired in winter respectively. Figures 15c, e, g, i represent the maximum likelihood classifications using Proba Chris images obtained during winter and summer, and the Hyperion reduced sets A and B acquired in winter respectively. The trends identified for each dataset are discussed in the subsequent paragraphs.

During winter 2006 and summer 2007, the mahalanobis distance method best classified 4 year old *Saccharum* and 6 year old *A. mearnsii* using the Proba Chris multispectral dataset (Figure 15b and d) .In contrast, the maximum likelihood method best classified the 6 year old *A. mearnsii* and 10 and 12 year old *E. macarthurri* during winter (Figure 15c); and 6 year old *A. mearnsii* and 12 year old *E. macarthurri* during summer (Figure 15e). However, the proportion of *E. macarthurri* classified and mapped during winter compared to summer is different (Figures 15c and 15e), as a result of selected compartments being felled prior to the summer field campaign. In general, when applying both statistical classifiers to the Proba Chris multispectral imagery, 6 year old *A. mearnsii* is classified and mapped satisfactorily during the dry and wet season.

The results of validation phase for the Hyperion reduced dataset A (27 spectral bands ranging from 900 to 1400 nm) and B (23 spectral bands ranging from 793 nm to 2375 nm) acquired during winter clearly demonstrate the importance of utilizing an optimal set of spectral bands to best discriminate between vegetation classes. In general both statistical classifiers best classify 4 year old *Saccharum* and 8 year old *E. grandis* when using the Hyperion reduced dataset A (Figure 15f and g). In contrast, to 6 year old *A. mearnsii* and 12 year old *E. macarthurri* best classified using the mahalanobis distance and maximum likelihood methods with the Hyperion reduced dataset B (Figures 15h and i). Furthermore, when applying the maximum likelihood method to the Hyperion reduced dataset B, it was possible to classify and map 12 year old *P. patula* species as shown in Figure 15i.



a
b
c
0 1.5 3
km
N
d
e
Acacia mearnsii 6yr
Saccharum 4yr
Pinus patula 12yr
Eucalyptus grandis 8yr
Eucalyptus macarthurii 10yr
Eucalyptus macarthurii 12yr
f
g
h
i

Figure 15 The classification and mapping of each vegetation class using Proba Chris and Hyperion remote sensing imagery during the validation phase: a) true colour image, Proba chris winter classification using b) mahalanobis distance and c) maximum likelihood classification, proba chris summer classifications using d) mahalanobis distance and e) maximum likelihood, Hyperion reduced set A winter classification f) mahalanobis distance and g) maximum likehood and, Hyperion reduced set B winter classification h) mahalanobis distance and i) maximum likehood methods.

The validation results indicate that each remote sensing dataset with its unique combination of spectral bands have classified and mapped the individual vegetation classes to different degrees of accuracy. Variability exists within a season and between statistical classifiers across the dry and wet seasons. The validation results indicate an overall trend of genus level classification when using the Proba Chris multispectral bands, compared to genus and species level classification when using a unique set of Hyperion hyperspectral bands. These results will be published and are reported in Appendix B.

## 4.4 Classification and Mapping of Vegetation Using Vegetation Indices

Classification and mapping of vegetation in the Mistley-Canema estate using vegetation indices was also investigated in this study. Details of this investigation are reported in the subsequent sections of this document.

### 4.4.1 Training phase

Training regions and regions of interest previously selected in this study were used in this part of the investigation. Vegetation indices were derived for each of the datasets. The parallepiped and minimum distances classifiers were used for the classification of the vegetation classes. Other statistical classifiers could not be utilised since they required a minimum of two spectral bands or layers. The performance of each statistical classifying method was evaluated using accuracy assessments (>80% overall accuracy) as shown in Table 9. In general, the classification and mapping of the selected vegetation classes during the training phase were unsatisfactory for the Proba Chris summer dataset and the Hyperion reduced dataset A. Only selected indices classified using the minimum distance method performed well with the Proba Chris winter dataset. Therefore no seasonal variations were identified in this investigation. The pigment index, red edge normalized difference vegetation index, soil adjusted vegetation index and vogelmann a index, for the Proba Chris winter dataset achieved an overall accuracy > 80%, and a kappa coefficient of good agreement > 0.8. These were therefore selected for the validation phase.

Table 9 Accuracy assessment: training phase of classification using vegetation indices

| Vegetation Index | Statistical Classifier | Proba Chris (winter 2006 survey) | | Proba Chris (summer 2007 survey) | | Hyperion reduced set A (winter 2006 survey) | |
|---|---|---|---|---|---|---|---|
| | | Kappa Coefficient | Overall Accuracy (%) | Kappa Coefficient | Overall Accuracy (%) | Kappa Coefficient | Overall Accuracy (%) |
| Carter | Parallelepiped | 0.39 | 51.63 | 0.22 | 33.19 | 0.43 | 50.54 |
| | Minimum Distance | 0.53 | 62.83 | 0.37 | 47.48 | 0.46 | 53.45 |
| Inverse Carter | Parallelepiped | 0.35 | 48.21 | 0.18 | 29.20 | 0.42 | 49.77 |
| | Minimum Distance | 0.55 | 64.23 | 0.35 | 45.69 | 0.50 | 56.66 |
| Normalized difference vegetation index | Parallelepiped | 0.32 | 46.50 | 0.23 | 34.24 | 0.38 | 46.55 |
| | Minimum Distance | 0.61 | 68.74 | 0.38 | 48.00 | 0.45 | 52.68 |
| Pigment | Parallelepiped | 0.54 | 63.92 | 0.16 | 28.05 | 0.49 | 55.90 |
| | Minimum Distance | 0.82 | 85.54 | 0.45 | 54.73 | 0.54 | 60.80 |
| Red-Green ratio | Parallelepiped | 0.46 | 56.92 | 0.25 | 36.34 | 0.22 | 32.16 |
| | Minimum Distance | 0.75 | 79.94 | 0.56 | 64.08 | 0.30 | 40.28 |
| Red edge normalized difference vegetation index | Parallelepiped | 0.54 | 63.92 | 0.21 | 32.25 | 0.39 | 47.32 |
| | Minimum Distance | 0.81 | 84.91 | 0.45 | 54.20 | 0.52 | 58.96 |
| Soil adjusted vegetation index | Parallelepiped | 0.31 | 45.26 | 0.18 | 29.73 | 0.23 | 33.08 |
| | Minimum Distance | 0.80 | 84.14 | 0.35 | 45.48 | 0.27 | 36.75 |
| Vogelmann a index | Parallelepiped | 0.49 | 59.72 | 0.00 | 12.40 | 0.46 | 52.99 |
| | Minimum Distance | 0.75 | 80.09 | 0.42 | 52.31 | 0.49 | 56.66 |
| Vogelmann c index | Parallelepiped | 0.23 | 70.61 | 0.00 | 12.29 | 0.21 | 32.47 |
| | Minimum Distance | 0.63 | 37.79 | 0.25 | 37.29 | 0.27 | 37.21 |
| Normalized infrared index | Parallelepiped | | | | | 0.25 | 35.68 |
| | Minimum Distance | | | | | 0.30 | 39.66 |
| Normalized difference water index | Parallelepiped | | | | | 0.13 | 25.27 |
| | Minimum Distance | | | | | 0.19 | 30.47 |
| Water band index | Parallelepiped | | | | | 0.16 | 27.57 |
| | Minimum Distance | | | | | 0.24 | 34.92 |

The classification of the vegetation classes using the different classifying techniques were further verified by mapping individual vegetation classes across the training area and comparing their spatial distribution to true colour images and spatial maps of forestry compartments obtained from Mondi. Figure 16 illustrates the classification of the various vegetation indices that had an overall accuracy of greater than 80%.

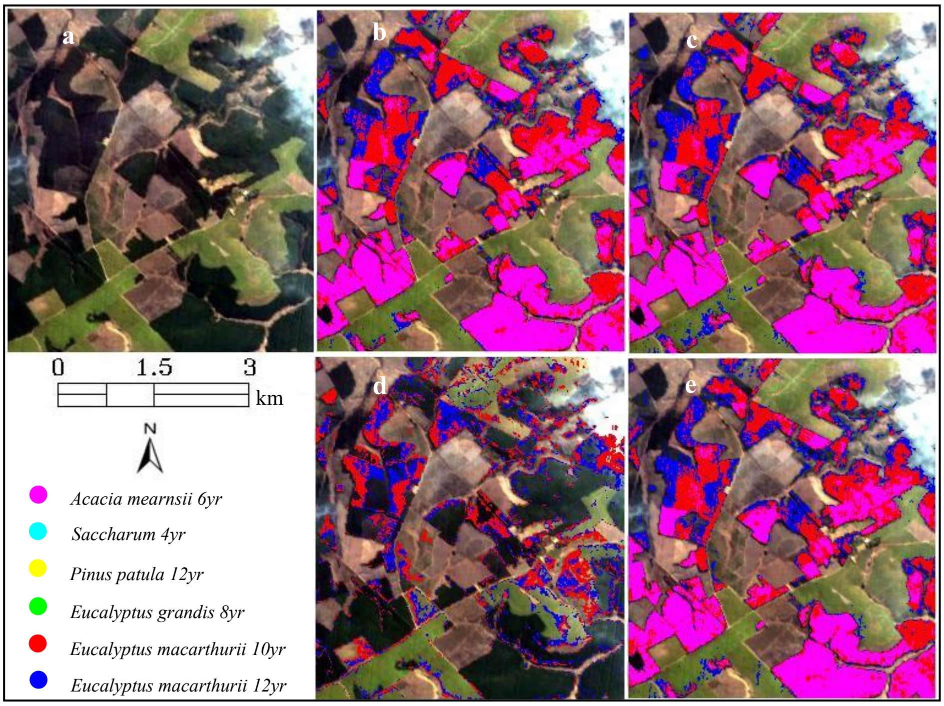



Figure 16 Classification of a) true colour image, b) pigment index, c) red edge normalized difference vegetation index, d) soil adjusted vegetation index layer and e) vogelmann a index using the minimum distance statistical classifier method.

Three species were mapped predominantly well using the pigment index, red edge normalized difference vegetation index, soil adjusted vegetation index layer and the vogelmann a index. These included the 6 year old *A. mearnsii*, and the 10 and 12 year old *E. macarthurrii*. Only a few vegetation classes were accurately classified and mapped using the vegetation indices. However, a good distinction is made between the *E. macarthurrii* age classes as shown in Figure 16b to e.

### 4.4.2 Validation phase

Four vegetation indices were selected for validation using the minimum distance classifier. Validation was performed using the entire study area. An accuracy assessment was used to evaluate each classification as shown in Table 10. Accuracy assessments indicate a decrease in classification with lower overall accuracies ranging from 68 to 70%.

Table 10 Accuracy assessment: validation phase of classification using vegetation indices

| Vegetation Index | Statistical Classifier | Proba Chris (winter 2006 survey) | |
|---|---|---|---|
| | | Kappa Coefficient | Overall Accuracy (%) |
| Pigment | Minimum Distance | 0.62 | 68.31 |
| Red edge normalized difference vegetation index | Minimum Distance | 0.63 | 69.37 |
| Soil adjusted vegetation index | Minimum Distance | 0.64 | 70.44 |
| Vogelmann a index | Minimum Distance | 0.63 | 69.11 |

Vegetation classes best discriminated during the validation phase were mapped across the study area. Figure 17 illustrates the classification of the various vegetation indices when applying the minimum distance classifier to the vegetation indices. These classifications can be compared to the true colour image in Figure 17a. The minimum distance method best classified the 6 year old *A. mearnsii* and the 10 and 12 year old *E. macarthurii* for all four vegetation indices. From a visual inspection of the classification results, the soil adjusted vegetation index performed poorly, even though it had achieved the highest overall accuracy (70.44%) during the validation phase. Furthermore, certain regions were incorrectly classified as *A. mearnsii* instead of *Saccharum* when using this index. *E. macarthurii* age classes were discriminated as shown in Figure 17 b to e.

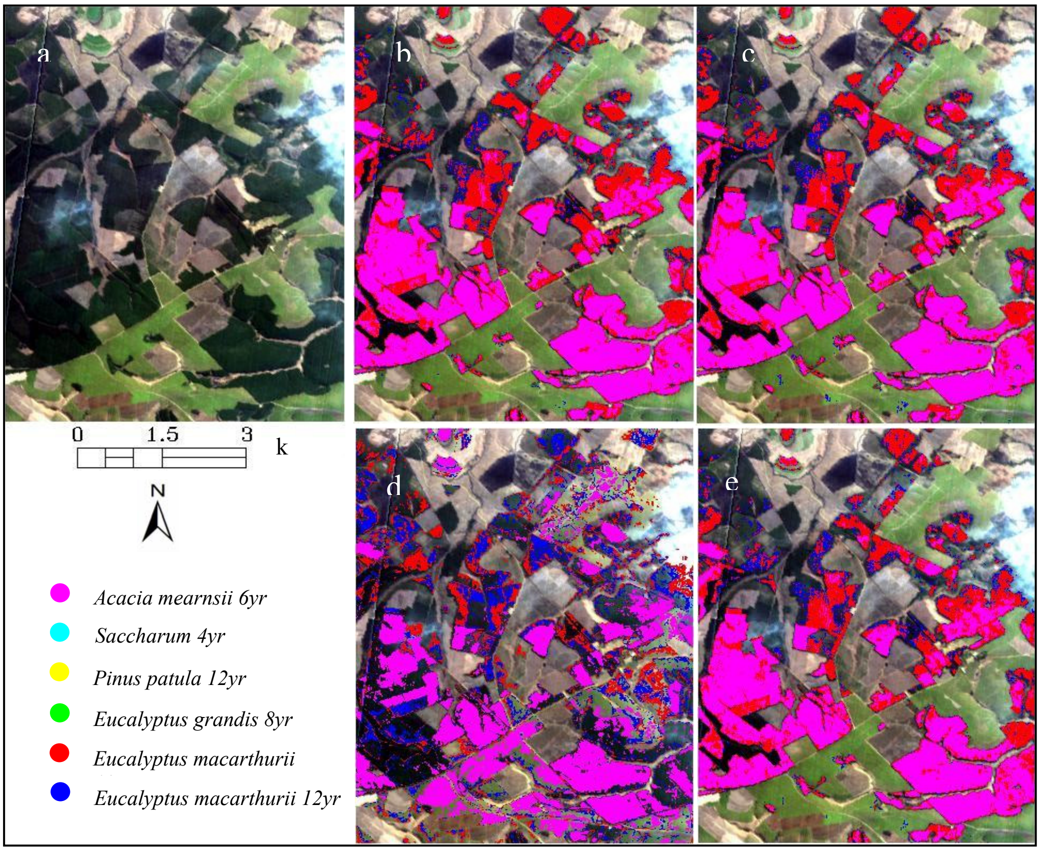



Figure 17 Validation classification of a) true colour image, b) pigment index, c) red edge normalized difference vegetation index, d) soil adjusted vegetation index layer and e) vogelmann a index using the minimum distance statistical classifier method.

# 5. CONCLUSIONS AND RECOMMENDATIONS

Recent advances in remote sensing have led the way for the use of hyperspectral remote sensing technologies in vegetation and water resources studies. Hyperspectral data has overcome the constraints of low spectral resolution, characteristic of existing multispectral data. However, according to the review of existing hyperspectral data sources, very few satellite technologies are available for operational or research applications. This study utilized the most accessible and cost effective hyperspectral technologies *viz*. the ASD hand-held spectrometer and the satellite Hyperion hyperspectral sensor. Futhermore, in this study the classification of different vegetation classes using both hyperspectral and multispectral imagery were investigated.

The classification and mapping of different genera and species selected in this study were strongly affected by seasonal physiological changes in vegetation, the range of spectral bands used in the classification and, the use of different statistical classifier methods. Seasonal variations in the spectral characteristics of vegetation may occur throughout the year and through different stages of development in their growth cycle. The discrimination between vegetation classes surveyed during winter and summer in this study indicate that the spectral characteristics of vegetation are significantly influenced by seasonal physiological changes in vegetation which are most prominent during the dry and wet seasons. The results from the two data reduction methods illustrate that different combinations of spectral bands can classify and map genera and species to different degrees of accuracy. In this study the key spectral wavelengths which best discriminate *Acacia mearnsii, Pinus patula, Eucalyptus macarthurii, Eucalyptus grandis* and *Saccharum* using the hyperion hyperspectral data were identified. In general, the results from this study suggest that multispectral remote sensing data generally produced genus level classification whereas hyperspectral remote sensing data facilitated genus and species level classification. This implies that classification at a finer resolution using an optimal and unique set of spectral bands is possible and can improve vegetation classification. Different statistical classifier methods and vegetation spectral reflectance indices were investigated in this project. The statistical classifiers which best classified *Acacia* *mearnsii, Pinus patula, Eucalyptus macarthurii, Eucalyptus grandis* and *Saccharum* during winter and summer included the mahalanobis distance and maximum likehood methods. Only selected vegetation spectral reflectance indices, the pigment index, red edge normalized difference vegetation index, soil adjusted vegetation index layer and vogelmann a index, performed well with the Proba Chris winter dataset.

Several remote sensing datasets were collated during this study. The ASD hand-held vegetation spectral signatures and ancillary data collected for *Acacia mearnsii, Pinus patula, Eucalyptus macarthurii, Eucalyptus grandis* and *Saccharum* during winter and summer will contribute towards a vegetation spectral library for Southern Africa. A database of spectral reflectance signatures is being managed by the Agricultural Research Council.

Data limitations and instrument problems addressed in this study are reported below for the benefit of future applications. These limitations range from technical problems and time-based obstacles to climatic issues, all related to the remote sensing aspects of this project and, which could not be predicted or controlled by the project team. These can be summarized as follows: instrumentation problems with the Proba Chris sensor during winter; erratic climatic conditions during field surveys; delays in data requests for Hyperion imagery from USGS and, “noisy” or poor quality Hyperion imagery.

## 5.1 Transferability

The methodology adopted in this classification study can be applied to other study areas and be used as the foundation for improving vegetation classification using an optimal set of spectral bands which best characterize a particular vegetation class. Key procedures which were followed in this research study are outlined in Figure 18.

Several statistical classifiers were used in this study as shown in Figure 18, because it is clear from the literature and evident in the results obtained from this project that their performance is highly variable for different vegetation classes. It is recommended that more than one method be used in vegetation classification studies. Furthermore, the project team re-iterates the statements made by Congalton (1994) and strongly recommends that statistical accuracy assessments and visual inspections of the results are performed to access the accuracy of classification methods.

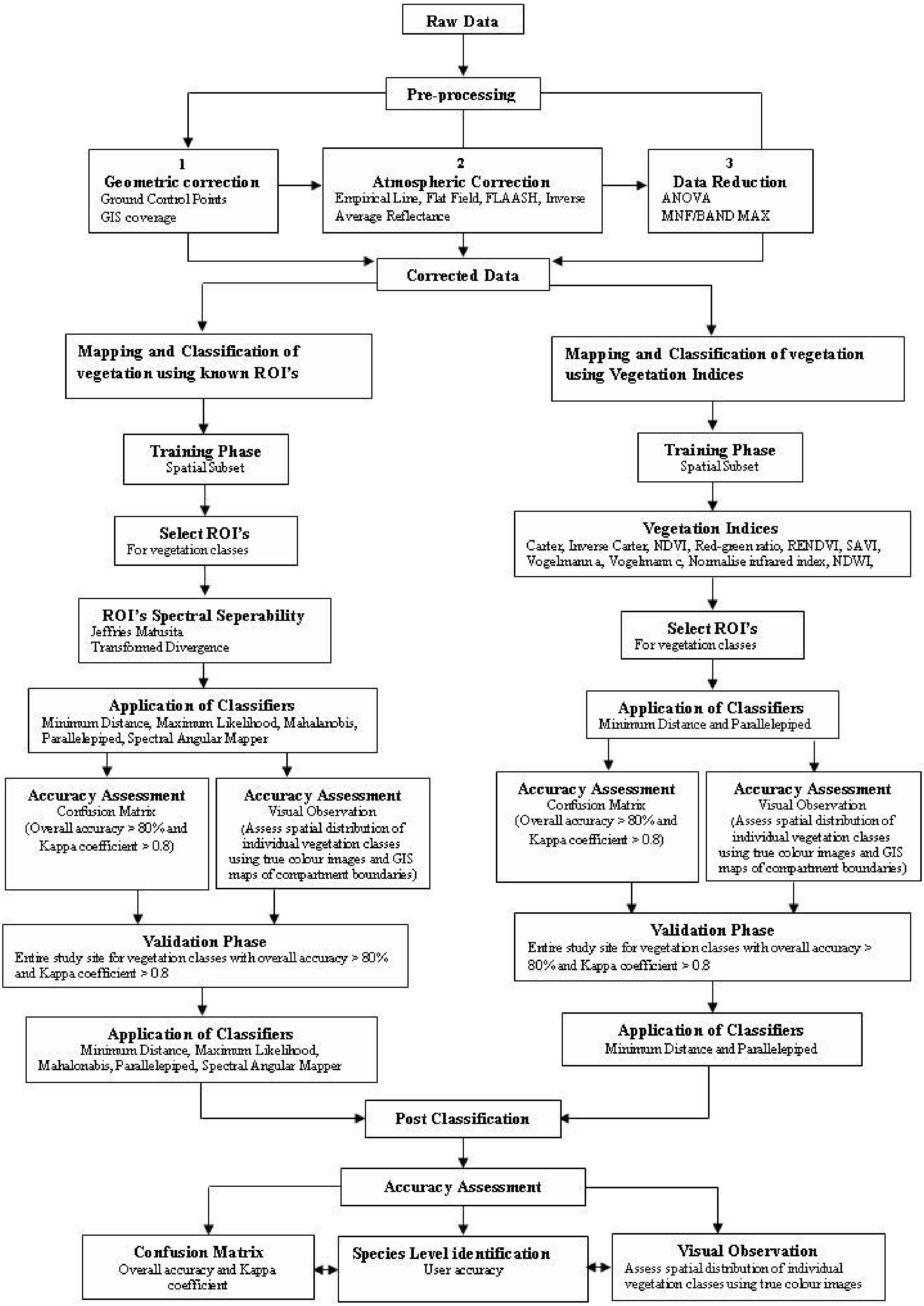



Figure 18 Procedures of classification.

Other aspects of this study which may be transferred and used in future applications include:

- Unique Proba Chris multispectral and Hyperion hyperspectral bands can be applied in other vegetation studies. Key spectral wavelengths identified which enhance the spectral differences amongst the selected vegetation classes could be applied to other localities and verified using different remote sensing data sources.
- The optimal sets of spectral bands which best characterize the vegetation classes selected for this study could initially be applied to other vegetation types. However, due to high spectral variability between different vegetation classes or mixed vegetation types it is recommended that the data reduction methods reported in this study be used obtain the optimal set of spectral bands which would best characterize the vegetation type under investigation.
- The spectral and spatial resolution chosen for a particular study is dependent upon the objective of the study and the scale of investigation. In this study a high spectral resolution was necessary to capture spectral differences amongst the different vegetation classes under investigation. However, a medium spatial resolution of 30 m for the Hyperion hyperspectral data and 17 m for the Proba Chris multispectral data was deemed an adequate spatial resolution required for the Mistley-Canema estate. It is important when undertaking vegetation or water resources type studies to utilize remote sensing data which meet both the spectral and spatial requirements needed for a particular application. In general, as the scale of the investigation increases, lower spatial resolutions which are more freely available and of lower cost can be used. However, generalizations with regards to the spatial and spectral resolution of datasets required for a particular study are discouraged.
- The mahalanobis distance and maximum likehood statistical classifiers best classified *Acacia mearnsii, Pinus patula, Eucalyptus macarthurii, Eucalyptus grandis* and *Saccharum* during winter and summer. Their performance in classifying other vegetation classes should be further investigated.
- The vegetation indices, pigment, red edge normalized difference vegetation index, soil adjusted vegetation index and vogelmann a index which

performed well in this study should be evaluated in different localities and for different vegetation classes.

## 5.2 Recommendations for Future Research

It is recommended, that future remote sensing research studies utilize hyperspectral data in conjunction with multispectral data. Such datasets can be utilized in vegetation and in hydrological studies for:

- Improved spatial estimates of plant water use, plant water content and plant biomass;
- More detailed classifications of alien invasive species; and
- Classifying regions/zones of land degradation and biodiversity loss;

which will assist in improved decision making in water resource management.

Future hyperspectral research should also focus on means of improving the accuracy of predictions by incorporating hyperspectral data into currently used systems. Developers of remote sensing technologies should strive toward optimizing sensor systems. Optimal hyperspectral bands which best characterize different vegetation classes should be incorporated into multispectral systems through sensor modifications or spectral filters. Using this approach researchers will no longer be restricted to existing sensor specifications, but can tailor remote sensing model inputs to a particular application. The ultimate goal is to move from impractical large data volumes and expensive sensor development towards practical operational low data volumes, high temporal resolution and cheaper sensor development. Such recommendations should feed into broader space observation programmes such as the Multi Sensor Micro Satellite Imager (MSMI) programme which is a joint venture between a consortium of partners including universities, private companies and research councils in South Africa and Belgium.

# 6. REFERENCES


Analytical Spectral Devices Inc. 2002. FieldSpec Pro User Guide. 5335 Sterling Drive, Suite A, Boulder, Colorado 80301, USA.

Asner, G.P. 1998. Biophysical and biochemical sources of variability in canopy reflectance. Remote Sensing of Environment, 64, 234-253.

Bellman, R. 1961. Adaptive control processes: A guided tour. Princeton University Press, Princeton, New Jersey.

Boardman, J. W. and Kruse, F. A. 1994. Automated spectral analysis: a geological example using AVIRIS data, north Grapevine Mountains, Nevada: in Proceedings, ERIM Tenth Thematic Conference on Geologic Remote Sensing. Environmental Research Institute of Michigan, Ann Arbor, MI, I-407 - I-418.

Boschetti, L., Eva, H.D., Brivio, P.A. and Gregoire J.M. 2004. Lessons to be learned from the comparison of three satellite-derived biomass burning products. Geophysical Research Letters, VOL. 31, L21501.

Camp, K.G.T. 1997. The Bioresources Groups of KwaZulu-Natal. Cedara Report N/A/97/6. KwaZulu-Natal department of Agriculture, Pietermaritzburg.

Carter, G.A. 1994. Ratios of leaf reflectance in narrow wavebands as indicator of plant stress. International Journal of Remote Sensing, 15, 697-704.

Congalton , R.G. and Mead, R.A. 1983. A quantitative method to test for consistency and correctness in photointerpretation. Photogrammatic Engineering & Remote Sensing, 49, 69 - 74.

Congalton, R.G. 1991. A review of assessing the accuracy of classifications of remotely sensed data. Remote Sensing of Environment, 37, 35-46.

Congalton, R.G. and Green, K. 1993. A practical look at the sources of confusion in error matrix generation. Photogrammetric Engineering & and Remote Sensing, 59, 641-644.

Congalton, R.G. 1994. Accuracy assessment of remotely sensed data: future needs and directions. In: Proc. 12th Pecora Symp, American Society for Photogrammetry and Remote Sensing, Bethesda, MD (1994), 385–388.

Congalton, R.G. 2001. Accuracy assessment and validation of remotely sensed and other spatial data. International Journal of Wildland Fire, 10, 321-328.

Cochrane, M.A. 2000. Using vegetation reflectance variability for species level classification of hyperspectral data. International Journal of Remote Sensing, 21 (10), 2075-2087.

Coops N., Dury, S., Smith, M.L., Martin, M. and Ollinger, S. 2002 Comparison of green leaf eucalypt spectra using spectral decomposition. Australian Journal of Botany, 50, 567-576.

Everson, C.E., Gush M.B., Moodley M., Jarmain C., Govender M. and Dye P. 2007. Effective management of the riparian zone vegetation to significantly reduce the cost of catchment management and enable greater productivity of land resources. WRC Report No. 1284/1/07, ISBN 978-1-77005-613-8.

Everson. C.S., Moodley, M., Gush, M.B., Jarmain, C.J., Govender, M. and Dye, P.J. 2006. Can effective management of riparian zone vegetation significantly reduce the cost of catchment management and enable greater productivity of land resources. WRC Report K5/1284. Water Research Commission, Pretoria.

Foody, G.M. 2004. Thematic map comparison: Evaluating the statistical difference in classification accuracy. Photogrammetric Engineering & Remote Sensing, 70(5), 627-633.

Foody, G.M. and Mathur, A. 2006. The use of training sets containing mixed pixels for accurate hard image classification: Training on mixed spectral responses for classification by SVM. Remote Sensing of Environment, 103, 179-189.

Galvao L.S., Formaggio, A.R. and Tisot D.A. 2005. Discrimination of sugarcane varieties in South Eastern Brazil with EO-1 Hyperion data. Remote Sensing of Environment, 94, 523-534.

Gamon, J.A. and Surfus, J.S. 1999. Assessing leaf pigment content and activity with a reflectometer. New Phytologist, 143, 105-117.

Gao, B.C. 1995. Normalized Difference Water Index for Remote Sensing of Vegetation Liquid Water from Space. Proceedings of SPIE 2480, 225-236.

Gitelson, A.A. and Merzlyak, M.N. 1994. Spectral Reflectance Changes Associated with Autumn Senescence of Aesculus Hippocastanum L and Acer Platanoides L Leaves. Spectral Features and Relation to Chlorophyll Estimation. Journal of Plant Physiology, 143, 286-292.

Gitelson A.A. and Merzlyak M.N. 1996. Signature analysis of leaf reflectance spectra: algorithm development for remote sensing of chlorophyll. Journal of Plant Physiology, 148, 494-500.

Govender, M., Chetty, K. and Bulcock, H. 2007. A review of hyperspectral remote sensing and its application in vegetation and water resource studies. Water SA, 33, April edition, 8p.

Green, A. A., Berman, M., Switzer, P. and Craig, M. D. 1988. A transformation for ordering multispectral data in terms of image quality with implications for noise removal. IEEE Transactions on Geoscience and Remote Sensing, 26(1), 65-74.

Hardisky, M.A., Klemas, V. and Smart, R.M. 1983. The Influences of Soil Salinity, Growth Form, and Leaf Moisture on the Spectral Reflectance of Spartina Alterniflora Canopies. Photogrammetric Engineering and Remote Sensing , 49, 77-83.

Hsu, P.H. 2007. Feature extraction of hyperspectral images using wavelet and matching pursuit. *ISPRS Journal of Photogrammetry & Remote Sensing (In Press).*

Hughes, G.F. 1968. On the mean accuracy of statistical pattern recognizers. IEEE Transactions on Information Theory IT-14(1), 55-63.

Ismail, R., Mutanga, O. and Ahmed, F. 2007. In review. Discriminating Sirex noctilio attack in pine forest plantations in South Africa using high spectral resolution data. In: Kalacska, M., Sanchez-Azofeifa, A. (Eds.), Hyperspectral Remote Sensing of Tropical and Sub-Tropical Forests. Taylor and Francis: CRC Press, 350.

Jackson, T.L., Chen, D., Cosh, M., Li, F., Anderson, M., Walthall, C., Doriaswamy, P. and Hunt, E.R. 2004. Vegetation Water Content Mapping Using Landsat Data Derived Normalized Difference Water Index for Corn and Soybeans. Remote Sensing of Environment, 92, 475-482.

Kalacska, M., Bohlman, B., Sanchez-Azofeifa, G.A., Castro-Esau, K. and Caelli, T. 2007. Hyperspectral discrimination of tropical dry forest lianas and trees: Comparative data reduction approaches at the leaf and canopy levels. Remote Sensing of Environment (In Press).

Koch M., Inzana J. and El Baz F. 2005. Applications of hyperion hyperspectral and aster multispectral data in characterizing vegetation for water resource studies in arid lands. Geology and Remote Sensing, Salt Lake City Annual Meeting, 44-5.

Landis, J. and Koch, G. 1977. The measurement of observer agreement for categorical data. Biometrics, 33, 159-174.

Landgrebe D. 1999. On information extraction principles for hyperspectral data. Cybernetics, 28, part c, 1, 1-7.

Landgrebe D. 1999. Some fundamentals and methods for hyperspectral image data analysis. Systems and Technologies for Clinical Diagnostics and Drug Discovery II, 3603, 6.

Mc Gwire K., Minor T. and Fenstermaker L. 2000. Hyperspectral mixture modeling for quantifying sparse vegetation cover in arid environments. Remote Sensing of Environment, 72, 360-374.

Majeke, B., Madau, H., Poti, L. and Van Aart, J. 2006. Updating the National Landcover Database of Southern Africa. The International Archives of the Photogrammetry, Remote Sensing and Spatial Information Sciences, 34, 10.

Mundt, J.T., Glenn, N.F., Weber, K.T., Prather, T.S., Lass, L.W. and Pettingill, J. 2005. Discrimination of hoary cress and determination of its detection limits via hyperspectral image processing and accuracy assessment techniques. Remote Sensing of Environment, 96, 509 – 517.

Mutanga, O. 2005. Discriminating Tropical Grass Grown under Different Nitrogen Treatments using Hyperspectral Data Resampled to HYMAP. International Journal of Geoinformatics, 1(2), 21-32.

Okin, G.S., Roberts, D.A., Murray, B. and Okin, W.J. 2001. Practical limits on hyperspectral vegetation discrimination in arid and semiarid environments. Remote Sensing of Environment, 77, 212-225.

Penuelas, J., Filella, I., Biel, C., Serrano, L. and Save, R. 1995. The Reflectance at the 950-970 Region as an Indicator of Plant Water Status. International Journal of Remote Sensing , 14, 1887-1905.

Research Systems Inc. 2005. RSI Training Series: Spectral Analysis with ENVI. 4990 Pearl East Circle, Boulder, CO 80301, 106.

Richards, J.A. and Jia, X. 1999. Remote sensing digital image analysis. Springer, Berlin, 363.

Schmidt, K.S. and Skidmore, A.K. 2003. Spectral discrimination of vegetation types in a coastal wetland. Remote Sensing of Environment, 85, 92– 108.

Schultz, G.A. and Engman, E.T. 2000. Editors, *Remote Sensing in Hydrology and Water Management*. Springer, Berlin, 483.

Sims, D.A. and Gamon, J.A. 2002. Relationships Between Leaf Pigment Content and Spectral Reflectance Across a Wide Range of Species, Leaf Structures and Developmental Stages. Remote Sensing of Environment, 81, 337-354.

Stone, C., Chisholm, L. and Coops, N. 2001. Spectral reflectance characteristics of eucalypt foliage damaged by insects. Austrian Journal of Botany, 49, 687-698.

Thomlinson, J.R., Bolstad, P.V. and Cohen, W.B. 1999. Coordinating methodologies for scaling landcover classifications from site-specific to global: steps toward validating global map products. Remote Sensing of the Environment, 70, 16-28.

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

Underwood, E., Ustin, S. and Dipietro, D. 2003. Mapping nonnative plants using hyperspectral imagery. Remote Sensing of the Environment, 86, 150-161.

Vaiphasa, C., Ongsomwang, S., Vaiphasa, T. and Skidmore, A.K. 2005. Tropical mangrove species discrimination using hyperspectral data: a laboratory study. Estuarine, Coastal and Shelf Science, 65, 371-379.

Vogelmann, J.E., Rock, B.N. and Moss, D.M. 1993. Red Edge Spectral Measurements from Sugar Maple Leaves. International Journal of Remote Sensing, 14, 1563-1575.

Yamano, H., Chen, J. and Tamura, M. 2003. Hyperspectral identification of grassland vegetation in Xilinahot, Inner Mongolia, China. International Journal of Remote Sensing, 24, 3171-3178.

Zagolski, F., Pinel, V., Romier, J., Alcayde, D., Gastellu-Etchegorry, J.P., Giordano, G., Marty, G., Mougin, E. and Joffre, R. 1996. Forest canopy chemistry with high spectral resolution remote sensing. International Journal of Remote Sensing, 17, 1107-1128.

**Appendix A**

# A REVIEW OF HYPERSPECTRAL REMOTE SENSING AND ITS APPLICATION IN VEGETATION AND WATER RESOURCE STUDIES

*M Govender[1], K Chetty[2] and H Bulcock[2]*

[1]*CSIR Natural Resources and the Environment, c/o School of Applied and Environmental Sciences, University of KwaZulu-Natal, Private Bag X01, Scottsville , 3209, South Africa*

[2] *School of Bioresources Engineering and Environmental Hydrology, University of KwaZulu- Natal, Private Bag X01, Scottsville , 3209, South Africa*

**KEYWORDS:** hyperspectral, multispectral, spectral resolution, spatial resolution, vegetation classification, water resources

**ABSTRACT**

Multispectral imagery has been used as the data source for water and land observational remote sensing from airborne and satellite systems since the early 1960s. Over the past two decades, advances in sensor technology have made it possible for the collection of several hundred spectral bands. This is commonly referred to as hyperspectral imagery. This review details the differences between multispectral and hyperspectral data; spatial and spectral resolutions and, focuses on the application of hyperspectral imagery in water resource studies and, in particular the classification and mapping of land use and vegetation.

**INTRODUCTION**

Water is one of the most valuable and essential resources that form the basis of all life. With the ever-increasing human population, there is constant stress exerted on water resources (McGwire *et al.*, 2000). Accurate monitoring and assessment of our water resources is necessary for sustained water resource management. Earth observation data have formed the basis for acquiring data remotely for many years (Landgrebe, 1999) and are viewed as a time-and cost-effective way to undertake large-scale monitoring (Okin *et al.*, 2001), which can be used to determine the quality, quantity and geographic distribution of this resource.

Multispectral imagery has been used as the data source for water and land observational remote sensing from airborne and satellite systems since the 1960s (Landgrebe, 1999). Multispectral systems commonly collect data in three to six spectral bands in a single observation from the visible and near-infrared region of the electromagnetic spectrum. This crude spectral categorization of the reflected and emitted energy from the earth is the primary limiting factor of multispectral sensor systems. Over the past two decades, advances in sensor technology have overcome this limitation of earth observation systems, with the development of hyperspectral sensor technologies. Hyperspectral systems have made it possible for the collection of several hundred spectral bands in a single acquisition, thus producing many more detailed spectral data. However, with the advances in hyperspectral technologies practical issues related to increased sensor or imager costs, data volumes and data processing costs and times would need to be considered especially for operational modes.

This review details the differences between multispectral and hyperspectral data; highlights commonly used remote sensing terminology, and focuses on the use of hyperspectral imagery in water resource studies and, in particular vegetation applications.

## DIFFERENCES BETWEEN MULTISPECTRAL AND HYPERSPECTRAL DATA

Multispectral airborne and satellite systems have been employed for gathering data in the fields of agriculture and food production, geology, oil and mineral exploration, geography and urban to non-urban localities (Landgrebe, 1999). The advantage of using satellite remote sensing systems was to provide both the synoptic view space provides and the economies of scale, since data over large areas could be gathered quickly and economically from such platforms (Landgrebe, 1999).

Multispectral remote sensing systems use parallel sensor arrays that detect radiation in a small number of broad wavelength bands. According to Smith (2001a), most multispectral satellite systems measure between three and six spectral bands within the visible to middle infrared region of the electromagnetic spectrum. There are, however, some systems that use one or more thermal infrared bands. Multispectral remote sensing allows for the discrimination of different types of vegetation, rocks and soils, clear and turbid water, and selected man-made materials (Smith, 2001a). To obtain data of a higher spectral resolution compared to multispectral data, hyperspectral sensors on board satellites or airborne hyperspectral imagers are used (Smith, 2001b).

Hyperspectral remote sensing imagers acquire many, very narrow, contiguous spectral bands throughout the visible, near-infrared, mid-infrared, and thermal infrared portions of the electromagnetic spectrum. Hyperspectral sensors typically collect 200 or more bands enabling the construction of an almost continuous reflectance spectrum for every pixel in the scene. Contiguous, narrow bandwidths characteristic of hyperspectral data allow for in-depth examination of earth surface features which would otherwise be 'lost' within the relatively coarse bandwidths acquired with multispectral scanners.

Over the past decade, extensive research and development has been carried out in the field of hyperspectral remote sensing. Now with commercial airborne hyperspectral imagers such as CASI and Hymap and the launch of satellite-based sensors such as Hyperion, hyperspectral imaging is fast moving into the mainstream of remote sensing and applied remote sensing research studies. Hyperspectral images have found many applications in water resource management, agriculture and environmental monitoring (Smith, 2001a). It is important to remember that there is not necessarily a difference in spatial resolution between hyperspectral and multispectral data but rather in their spectral resolutions.

## FUNDAMENTALS OF IMAGE RESOLUTION

Remotely sensed images are characterised by their spatial and spectral resolutions. The terms spatial and spectral resolution represent pixels of an image displayed in a geometric relationship to one another; and variations within pixels as a function of wavelength respectively. These fundamentals of image resolution are explained below.

**Spectral resolution**

Spectral resolution refers to the number and width of the portions of the electromagnetic spectrum measured by the sensor. A sensor may be sensitive to a large portion of the electromagnetic spectrum but have poor spectral resolution if it captures a small number of wide bands. A sensor that is sensitive to the same portion of the electromagnetic spectrum but captures many small bands within that portion would have greater spectral resolution. The objective of finer spectral sampling is to enable the analyst, human or computer, to distinguish between scene elements. Detailed information about how individual elements in a scene reflect or emit electromagnetic energy increases the probability of finding unique characteristics for a given element which allows for better distinction from other elements in the scene (Jensen, 1996). Multispectral remote sensing systems record energy over several separate wavelength ranges at various spectral resolutions. Hyperspectral sensors, detect hundreds of very narrow spectral bands throughout the visible, near-infrared, and mid-infrared portions of the electromagnetic spectrum. A distinct advantage of their very high spectral resolution facilitates fine discrimination between different targets based on their spectral response in each of the narrow bands (Landgrebe, 1999).

**Spatial Resolution**

Spatial resolution defines the level of spatial detail depicted in an image. This may be described as a measure of the smallness of objects on the ground that may be distinguished as separate entities in the image, with the smallest object necessarily being larger than a single pixel. In this sense, spatial resolution is directly related to image pixel size. The spatial property of an image is a function of the design of the sensor in terms of its field of view and the altitude at which it operates above the surface (Smith, 2001b). Each of the detectors in a remote sensor measures the energy received from a finite patch on the ground surface. The smaller the individual patches are, the higher the spatial resolution and the more detail one can spatially interpret from the image (Smith, 2001b). Currently hyperspectral imagery acquired with satellite systems are usually in the order of 30 m or finer whereas airborne systems generally acquire higher spatial resolution data usually in the order of 5 m or finer.

## CONTIGUOUS SPECTRAL SIGNATURES

An understanding of spectral signatures is essential in the understanding and interpretation of a remotely sensed image. Different materials are discriminated by wavelength-dependent absorptions, and these images of reflected solar energy are known as spectral signatures. The property that is used to quantify these spectral signatures is called spectral reflectance. This is a ratio of the reflected energy to incident energy as a function of wavelength (Smith, 2001b). The graph of the spectral reflectance of an object as a function of wavelength is termed the spectral reflectance curve (Lillesand and Kiefer, 1999).

The configuration of the spectral reflectance curves is important in the determination of the wavelength region(s) in which remote sensing data is acquired as the spectral reflectance curves give insight into the spectral characteristics of an object (Lillesand and Kiefer, 1999). Spectral signatures obtained from multispectral images are discrete compared to the contiguous signatures obtained from hyperspectral images. Contiguous spectral signatures allow for detailed analysis through the detection of surface materials and their abundances, as well as inferences of biological and chemical processes.

The three basic contiguous spectral reflectance signatures of earth features used are those for green vegetation, dry bare soil and clear water. The Fig. 1 below shows the average reflectance curves for each feature.

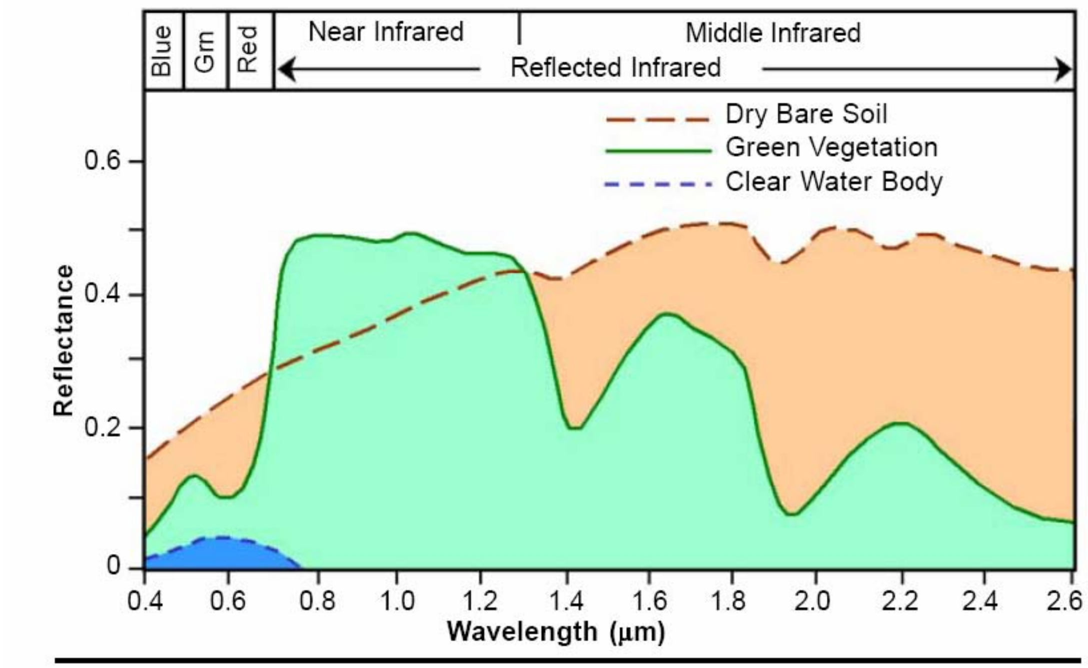



Figure 1 Spectral signatures for dry bare soil, green vegetation and clear water. (Smith, 2000b)

In green vegetation, the valleys in the visible portion of the spectrum are determined by the pigmentation of the plant. For example, chlorophyll absorbs strongly in the blue (450 nm) and red (670 nm) regions, also known as the chlorophyll absorption bands. Chlorophyll is the primary photosynthetic pigment in green plants (Smith, 2000b). This is the reason for the human eye perceiving healthy vegetation as green, due to the strong absorption of the red and blue wavelengths and the reflection of the green wavelengths. When the plant is subjected to stress that hinders normal growth and chlorophyll production, there is less adsorption in the red and blue regions and the amount of reflection in the red waveband increases (Smith, 2000b).

The spectral reflectance signature illustrates a dramatic increase in the reflection for healthy vegetation at around 700 nm. In the NIR between 700 and 1 300 nm, a plant leaf will typically reflect between 40 to 50%, the rest is transmitted, with only about 5% being adsorbed. Structural variability in leaves in this range allows one to differentiate between species, even though they might look the same in the visible region (Lillesand and Kiefer, 1999). Beyond 1 300 nm the incident energy upon the vegetation is largely absorbed or reflected with very little transmittance of energy. Three strong water absorption bands are noted at 1 400, 1 900 and 2 700 nm.

The spectral curve for bare soil shows far less variation in reflectance compared to that of green vegetation. This is due to the factors that affect soil reflectance acting over less specific spectral bands. Factors that affect soil reflectance include moisture content, soil texture, surface roughness, presence of iron oxide, and organic matter content. The factors that influence soil reflectance are complex, variable and interrelated (Lillesand and Kiefer, 1999).

The most distinctive characteristic of water is the absorption in the NIR and beyond. There are various conditions of water bodies that manifest themselves in the visible wavelengths. There are complex energy-matter interactions at these wavelengths that depend on a number of factors. These factors include the interaction with the water surface and material suspended in the water. Clear water reflects the greatest at 600 nm. The presence of organic and inorganic materials greatly influences the transmittance of the water and therefore the reflectance is dramatically affected. For example, a water body with high amounts of suspended sediments will reflect better than ‘clear’ water. An increase in chlorophyll will decrease the reflectance in the blue wavelengths and increase in the green. This can be useful in the detection of algae in water using remote sensing (Lillesand and Kiefer, 1999).

## APPLICATIONS OF HYPERSPECTRAL REMOTE SENSING IN WATER RESOURCES

Remote sensing technology has been widely used in water resource applications (Gitelson and Merzlyak, 1996; Zagolski *et al*., 1996; Asner, 1998; McGwire *et al*., 2000; Stone *et al*., 2001; Coops *et al*., 2002; Underwood *et al*., 2003) and in particular hyperspectral remote sensing is emerging as the more in-depth means of investigating spatial, spectral and temporal variations in order to derive more accurate estimates of information required for water resource applications. This section briefly highlights applications of hyperspectral remote sensing in water resources, and is followed by a detailed review of the methods and applications of land use and vegetation classification.

**Flood detection and monitoring** is constrained by the inability to obtain timely information of water conditions from ground measurements and airborne observations at sufficient temporal and spatial resolutions. Satellite remote sensing allows for timely investigation of areas of large regional extent and provides frequent imaging of the region of interest (Felipe *et al*., 2006). Until recently, near real-time flood detection was not possible, but with sensors such as Hyperion on board the EO-1 satellite this has been vastly improved (Felipe *et al*., 2006). According to research conducted by Felipe *et al*. (2006) automated spacecraft technology reduced the time to detect and react to flood events to a few hours. Advances in remote sensing, have resulted in the investigation of early warning systems with potential global applications. Most recent studies from NASA and the US Geological Survey are utilizing satellite observations of rainfall, rivers and surface topography into early warning systems (Brakenridge *et al*., 2006). Specifically, scientists are now employing satellite microwave sensors to gauge discharge from rivers by measuring changes in river widths and satellite based estimates of rainfall to improve warning systems (Brakenridge *et al*., 2006). Procedure for the detection of flooded areas with satellite data were also investigated by Glaber and Reinartz (2002). Moisture classes in flood plain areas in relation to high water changes, the accumulation of sediments and silts for different land use classes and erosive impacts of floods were investigated (Glaber and Reinartz, 2002). The estimation of discharge and flood hydrographs from hydraulic information obtained from remotely sensed data was assessed by Roux and Dartus (2006). Remote sensed images as used to estimate the hydraulic characteristics which are then applied in routing modules to generate a flood wave in a synthetic river channel. Optimization methods are used to minimise discrepancies between simulations and observations of flood extent fields to estimate river discharge (Roux and Dartus, 2006).

**Detection of water quality** conditions and parameters is one of the major advantages of hyperspectral remote sensing technologies. Hyperspectral reflectance has been widely used to assess water quality conditions of many open water aquatic ecosystems. This includes classifying the trophic status of lakes (Koponen *et al*., 2002; Thiemann and Kaufmann, 2002) and estuaries (Froidefond *et al*., 2002) characterising algal blooms (Stumpf, 2001) and assessment of ammonia dynamics for wetland treatments (Tilley *et al*., 2003). Predictors of total ammonia concentrations using remotely sensed hyperspectral signatures of macrophytes in order to monitor changes in wetland water quality were developed by Tilley *et al*. (2003). Hyperspectral spectrometers have also proved useful in determining the total suspended matter, chlorophyll content (Hakvoort *et al*., 2002; Vos *et al*., 2003) and total phosphorus (Koponen *et al*., 2002). Much research has been undertaken in the estimation of chlorophyll content from remotely sensed images which is then used as an estimate for monitoring algal content and hence water quality. Since wavelengths corresponding with the peak reflectance of blue-green and green algae are close together it is more difficult to differentiate between them. However, hyperspectral imagers allow for improved detection of chlorophyll and hence algae, due to the narrow spectral bands which are acquired between 450 nm and 600 nm. (Hakvoort *et al*., 2002). Estimation and mapping of water quality constituents such as concentrations of dissolved organic matter, chlorophyll or total suspended matter from optical remote sensing technologies have proved to be useful and successful and, is being investigated for operational use (Hakvoort *et al*., 2002).

**Wetland mapping** has gained increased recognition for the ability to improve quality of ecosystems (Schmidt and Skidmore, 2003). Sustainable management of any ecosystem requires, among other information, a thorough understanding of vegetation species distribution. Hyperspectral imagery has been used to remotely delineate wetland areas and classify hydrophytic vegetation characteristics of these ecosystems (Schmidt and Skidmore, 2003; Becker *et al*., 2005). Research undertaken by Schmidt and Skidmore (2003) promoted the use of high spatial and spectral resolution data for improved mapping of salt marsh vegetation of similar structure. The hyperspectral analysis identified key regions of the electromagnetic spectrum which provided detailed information for discriminating between and identifying different wetland species (Schmidt and Skidmore, 2003). Becker *et al*. (2005) performed a similar study based on coastal wetland plant communities which are spatially complex and heterogeneous. This study also emphasized the importance of hyperspectral imagery for identifying and differentiating vegetation spectral properties from narrow spectral bands focussing on the visible and near-infrared regions (Becker *et al*., 2005). A number of studies have investigated the potential of providing timely data for mapping and monitoring submerged aquatic vegetation which has been identified as one of the most important aspects of ecosystem restoration and reconstruction (Lin and Liquan, 2006). Such species have been termed ecological engineering species and the quantification of their coverage and spectral reflectance properties is currently being researched (Lin and Liquan, 2006).

**Measures of plant physiology and structure** such as leaf area index, water content, plant pigment content, canopy architecture and density have been investigated extensively over the past decade (Asner, 1998; Datt, 1998; Ceccato *et al*., 2001; Gitelson *et al*., 2002; Champagne *et al*., 2003; Gupta *et al*., 2003; Merzlyak *et al*., 2003; Pu *et al*., 2003; D'Urso *et al*., 2004; Schlerf *et al*., 2005; Stimson *et al*., 2005; Chun-Jiang *et al*., 2006) using hyperspectral imagery. These applications investigate the spectral reflectance properties of plants, identifying key spectral wavebands related to specific plant physiological and structural characteristics, hence deriving sensitive vegetation spectral indices for their non-destructive estimation. Remote sensing data have been exploited to estimate canopy characteristics by using empirical approaches based on spectral indices (D'Urso *et al*., 2004; Schlerf *et al*., 2005). Analysis of hyperspectral remote sensing data has been carried out to estimate LAI for agricultural crops and forests. Schlerf *et al*. (2005) investigated several narrow band and broad band vegetation indices in order to explore whether hyperspectral data may improve the estimation of biophysical variables such as LAI, canopy crown and crown volume when compared to multispectral analyses. The spectral and spatial information content of the satellite data was exploited to validate canopy reflectance models such as PROSPECT and SAILH (D'Urso *et al*., 2004). Results obtained for the crops under investigation encourage the use of canopy reflectance models in the inverse mode in order to retrieve other vegetation parameters such as chlorophyll content, dry matter and canopy geometrical characteristics like mean leaf inclination angle (D'Urso *et al*., 2004). The accurate estimation of plant water status and plant water stress is essential to the integration of remote sensing into precision agricultural and forestry management. The potential to spectrally estimate plant physiological properties over relatively large areas, and to predict plant water status and plant water stress was demonstrated by Champagne *et al*., 2003 for agricultural crops; and Stimson *et al*., 2005 and Eit el al., 2006 for forestry species. Their results indicate the potential use of vegetation spectral indices derived from various scales of remote sensing data for determining plant physiological properties and characteristics. These studies amongst others clearly indicate the improved estimates of plant physiological and structural characteristics from hyperspectral data, allowing for much more detailed spectral analyses and hence more accurate estimates.

**Evapotranspiration** (ET) estimates are essential in a wide range of water resource applications such as water and energy balance computations, in irrigation schemes, reservoir water losses, runoff prediction, meteorology and climatology (Medina *et al*., 1998). ET cannot be estimated directly from satellite observations; however, hyperspectral remote sensing can provide a good estimate of components of energy balance algorithms used to derive spatial estimates of ET. Earlier studies have focused on the estimation of evaporative fraction (EF) rather than on ET as the estimation of available radiant energy was difficult to obtain. According to Batra *et al*. (2006) an estimation of EF is defined as a ratio of ET and available radiant energy and has been estimated successfully using AVHRR and MODIS data. Several recent studies (Medina *et al.,* 1998; Kite and Droogers, 2000; Loukas *et al*., 2005; Batra *et al*., 2006; Eichinger, *et al*., 2006) have been conducted using more detailed hyperspectral data, ancillary surface data and atmospheric data for improved spatial estimates of ET. The availability of water, radiant energy and the removal of water vapour away from the surface are the major factors that control ET. However these factors in turn depend on other variables such as soil moisture, land surface temperature, air temperature, and vegetation cover, vapour pressure, and wind speed which may vary between regions, seasons, and time of day. Generally these factors are accounted for by using a combination of remote sensing data, ancillary surface data and atmospheric data for the estimation of ET values (Batra *et al*., 2006), and has lead to extensive measurements of surface fluxes, meteorological and soil variables (Wang *et al*., 2006). Batra *et al*. (2006) successfully estimated ET based on the extension of the Priestly-Taylor equation and the relationship between remotely sensed surface temperature and vegetation spectral indices.

**Land use and vegetation classification**

Land use and vegetation classification is generally performed using supervised and unsupervised classification methods which are commonly available in most data processing systems. The key difference in the two methods lies in the training stage of supervised classification which involves identifying areas of specific spectral attributes for each landcover or landuse type of interest to the analyst. In comparison unsupervised image classification into spectral classes as based solely on the natural groupings from the image values. Furthermore, remote sensing applications and specifically landuse or vegetation classification is seldom done without some form of ground truthing or collection of reference data. Ground-based spectral measurements are commonly done using portable, field spectrometers. Field spectrometers are utilized in forestry, agriculture and other environmental studies; and the spectral signatures obtained can be used in classification and mapping of vegetation, mapping of ecosystem productivity, crop type or yield mapping, and in the detection of plant stress for water resource operations and management.

**Applications**

Traditional methods for landscape-scale vegetation mapping require expensive, time-intensive field surveys. Remotely sensed data for the classification and mapping of vegetation provides a detailed accurate product in a time and cost effective manner. The availability of satellite and airborne hyperspectral data with its increased spatial and more critically fine spectral resolution offers an enhanced potential for the classification and mapping of land use and vegetation. Due to the large number of wavebands, image processing is able to capitalize on both the biochemical and structural properties of vegetation (Underwood *et al*., 2003). The need for exploring these spectral properties is particularly important when we consider the limitations of using traditionally available wavebands, where majority of the land cover is grouped and identification of individual species is difficult.

Vegetation has unique spectral signatures. Vegetation spectra are often used for the training stage of image classification. The spectral reflectance signatures of healthy vegetation have characteristic shapes that are dictated by various plant attributes. In the visible portion of the spectrum, the curve is governed by absorption effects of chlorophyll and other leaf pigments. Leaf structure varies significantly between plant species, and can also change as a result of stress. Thus species type, plant stress and canopy state can all affect near-infrared reflectance and making it difficult to distinguish between different vegetation types (Smith, 2001a), as illustrated in Fig. 2.

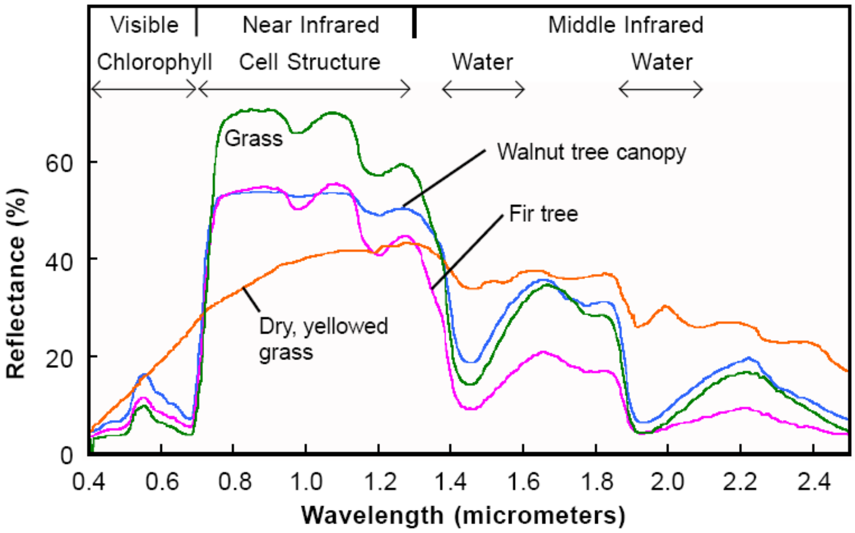



Figure 2 Reflectance spectra of different types of green vegetation compared to a spectral signature for senescent leaves (Smith, 2001a).

Several studies have derived band ratios or spectral indices which are useful for emphasizing certain physiological features, and can be used to distinguish between different vegetation within a mosaic of other land uses (McGwire *et al*., 2000; Underwood *et al*., 2003; Yamano *et al*., 2003; Galvao *et al*., 2005). Commonly used indices include normalized difference vegetation index, soil-adjusted vegetation index, modified soil-adjusted vegetation index (McGwire *et al*., 2000). Vegetation indices are also used to mask out vegetated areas from remote sensed imagery which are then used in the classification process (Underwood *et al*., 2003; Yamano *et al*., 2003; Koch *et al*., 2005). Similarly, principal component analysis has been applied as a data enhancement method when analysing remote sensing images to distinguish between vegetated and non-vegetation areas (Castro-Esau *et al*., 2004). These techniques facilitate the classification of vegetated areas, especially over large spatial scales where the landscape is unknown.

Classification algorithms or statistical classifiers such as spectral angle mapper, gaussian maximum likelihood and parallelepiped classifier use reflectance spectra as reference data for identifying classes from both multispectral and hyperspectral images. Reference spectra are measured or collected from pure and single image pixels or larger training areas. The quality or ‘pureness’ of the reference spectra is an important factor that defines the classification results (Castro-Esau *et al*., 2004).

Improved landcover and landuse classification is expected from hyperspectral data due to the improved quality in the reference spectra. Examples of these classification algorithms or statistical classifiers are discussed below.

**Gaussian maximum likelihood** estimates of the parameters are computed, and individual pixels are assigned to the class which maximises the likelihood function of the dataset. As a pixel by pixel method this approach does not take contextual information about the neighbouring classes into account in labelling a pixel. However, increased information provided by the spatial extent of the classes of the neighbours tends to mitigate the effects of noise, isolated pixels, and individual pixels (Castro-Esau, 2004).

**Spectral angle mapper** classifies by comparing the spectral angles between the reflectance spectrum of the classified pixel and the reference spectrum obtained from training data or a spectral library. Each pixel is assigned to a class according to the lowest spectral angle value (Kruce *et al*., 1993).

**Parallelepiped** classification is a decision rule method based on the standard deviation from the mean of each defined and trained class. A threshold of each class signature is used to determine if a given pixel falls within a class. Pixels which fall inside the parallelepiped are assigned to the class; however, those within more than one class are grouped into an overlap class. Pixels ungrouped are considered as unclassified (Lillesand and Kiefer, 1999).

## DISCUSSION AND CONCLUSION

Recent advances in remote sensing have led the way for the development of hyperspectral sensors and the application of hyperspectral data. Hyperspectral remote sensing is a relatively new technology to South Africa that is being used in the detection and identification of minerals, terrestrial vegetation, and man-made materials as well as in the field of water resources and environmental applications (Nagy and Jung, 2005).

The availability of hyperspectral data, has overcome the constraints and limitations of low spectral and spatial resolution imagery, and discreet spectral signatures.

Hyperspectral images provide high spectral resolution data, with many narrow contiguous spectral bands allowing for detailed applications.

This review highlights the vast extent to which hyperspectral technologies have been applied to the water resources and environmental sectors. In particular, the remotely sensed classification and mapping of land use and vegetation has been adopted internationally, since traditional methods of classifying and mapping land use and vegetation require expensive, time-intensive field surveys. Remote sensed data allow for the classification and mapping of vegetation in a time and cost-effective manner.

Different processing and statistical techniques such as spectral vegetation indices and principal component analysis are often used as data enhancement methods to mask vegetated areas within remotely sensed images. Vegetated areas are then classified using one or more recognized classification algorithms such as spectral angle mapper, gaussian maximum likelihood and parallelepiped, resulting in more accurate products borne from hyperspectral data sources. As hyperspectral imaging techniques evolve and are introduced into new fields their primary contribution will be in exploring and developing new applications primarily on the selection of optimal spectral band parameters i.e. band position and widths. Practical issues of sensor costs, data volumes and data processing mechanisms will need to addressed for operational use, and thus still favour multispectral systems.

In general this review illustrates the enhanced capability of hyperspectral technologies in vegetation and water resource studies and, allows water resource managers to make informed management decisions with the relevant detail in an efficient timeframe.

## REFERENCES

ASNER GP (1998) Biophysical and biochemical sources of variability in canopy reflectance. *Remote Sens*. *Environ.* **64** 234-253.

BATRA N, ISLAM S, VENTURINI V, BISHT G and JIANG L (2006) Estimation of comparison of evapotranspiration from MODIS and AVHRR sensors for clear sky days over the Southern Great Plains. *Remote Sens*. *Environ.* **103** 1-15.

BECKER BL, LUSCH DP and QI J (2005) Identifying optimal spectral bands from in-situ measurements of Great Lakes coastal wetlands using second derivative analysis. *Remote Sens. Environ*. **97** 238-248.

BRACKENRIDGE R, ANDERSON E and NGHIEM SV (2006) Satellite microwave detection and measurement of river floods. *NASA Spring Annual General Conference 2006.* www.nasa.gov/vision/earth/lookingatearth/springagu_2006.html (Accessed 4 October 2006).

CASTRO-ESAU KL, SANCHEZ-AZOFEIFA GA and CAELLI T (2004) Discrimination of lianas and trees with leaf level hyperspectral data. *Remote Sens. Environ.* **90** 353-372.

CECCATO P, FLASSE S, TARANTOLA S, JACQUEMOUD S and GREGOIRE JM (2001) Detecting vegetation leaf water content using reflectance in the optical domain. *Remote Sens. Environ*. **77** 22-33.

CHAMPAGNE CM, STAENZ K, BANNARI A, MCNAIRN H and DEGUISE JC (2003) Validation of a hyperspectral curve fitting model for the estimation of water content of agricultural canopies. *Remote Sens. Environ*. **87** 148-160.

CHUN-JIANG Z, JI-HUA W, LIANG-YUN L, WEN-JIANG H and QI-FA Z (2006) Relationship of 2100-2300nm spectral characteristics of wheat canopy to leaf area index and leaf N as affected by leaf water content. *Pedosphere* **16** 333-338.

COOPS N, DURY S, SMITH ML, MARTIN M and OLLINGER S (2002) Comparison of green leaf eucalypt spectra using spectral decomposition. *Austral. J. Bot*. **50** 567-576.

DATT B (1998) Remote sensing of chlorophyll a, chlorophyll b, chlorophyll a+b, and total carotenoid content in eucalyptus leaves. *Remote Sens*. *Environ* **66** 111-121.

D'URSO GD, DINI L, VUOLO F, ALONSO L and GUANTER L (2004) Retrieval of leaf area index by inverting hyperspectral multiangular CHRIS PROBA data from SPARC 2003. Proceedings of the second CHRIS Proba workshop, ESA/ESRIN, Frascati Italy, 28-30 April.

FELIPE IP, DOHM JM, BAKER VR, DOGGETT T, DAVIES AG, CASTANO R, CHIEN S, CICHY B, GREELEY R, SHERWOOD R, TRAN D, RABIDEAU G (2006) Flood detection and monitoring with the autonomous sciencecraft experiment onboard EO-1. *Remote Sens. Environ.* **101** 463-481.

EICHINGER WE, COOPER DI, HIPPS LE, KUSTAS WP, NEALE CMU and PRUEGER JH (2006) Spatial and temporal variation in evapotranspiration using Raman Lidar. *Adv. Water Resour*. **29** 369-381.

FROIDEFOND J, GARDEL L, GUIRAL D, PARRA M and TERNON J (2002) Spectral remote sensing reflectances of coastal waters in French Guiana under the Amazon influence. *Remote sens. Environ.***80** 225-232.

GALVAO LS, FORMAGGIO, AR and TISOT DA (2005) Discrimination of sugarcane varieties in South Eastern Brazil with EO-1 Hyperion data. *Remote Sens. Environ*. **94** 523-534.

GITELSON AA and MERZLYAK MN (1996) Signature analysis of leaf reflectance spectra: algorithm development for remote sensing of chlorophyll. *J. Plant Physiol*. **148** 494-500.

GITELSON AA, ZUR Y, CHIVKUNOVA, OB and MERZLYAK MN (2002) Assesing carotenoid content in plant leaves with reflectance spectroscopy. *Photochem. Photobiol.* **75** 272-281.

GLABER C and REINARTZ, P (2004) Multitemporal and multispectral remote sensing approach for flood detection in the Elbe-Mulde region 2002. *Acta hydrochmica et hydrobiologica* **33** Issue 5.

GUPTA RK, VIJAYAN D and PRASAD TS (2003) Comparative analysis of red-edge hyperspectral indices. *Adv. Space Res*. **32** 2217-2222.

HAKVOORT H, de HAAN J, JORDANS R, VOS R, PETERS S and RIJKEBOER M (2002) Towards airborne remote sensing of water quality in the Netherlands-validation and error analysis. *J. Photogr. Remote Sens*. **57** 171-183.

JENSEN, JR (1996) *Introductory Digital Image Processing: A Remote Sensing Perspective.* Prentice-Hall, New Jersey.

KITE G and DROOGERS P (2000) Comparing evapotranspiration estimates from satellites, hydrological models and field data. *J. Hydrol*. **229** 3-18.

KOCH M, INZANA J and EL BAZ F (2005) Applications of hyperion hyperspectral and aster multispectral data in characterizing vegetation for water resource studies in arid lands. *Geol. Remote Sens*. Salt Lake City Annual Meeting (16-19 October, 2005), Paper No. 44-5.

KOPONEN S, PULLIAINEN J, KALLIO K and HALLIKAINEN M (2002) Lake water quality classification with airborne hyperspectral spectrometer and simulated MERIS data. *Remote Sens. Environ.* **79** 51-59.

KRUCE F, LEFKOFF A, BOARDMAN J, HEIDEBRECHT K, SHAPIRO A, BARLOON P and GOETZ A (1993) The spectral image processing system (SIPS) interactive visualization and analysis of imaging spectrometer data. *Remote Sens. Environ*. **44** 145-163.

LANDGREBE D (1999) On information extraction principles for hyperspectral data. *Cybernetics* **28**, part c, 1, 1-7.

LANDGREBE D (1999) Some fundamentals and methods for hyperspectral image data analysis. *Systems and Technologies for Clinical Diagnostics and Drug Discovery* II, 3603. 6p.

LILLESAND TM and KIEFER RW (1999) *Remote Sensing and Image Interpretation.* John Wiley & Sons, Inc. New Jersey.

LIN Y and LIQUAN Z (2006) Identification of the spectral characteristics of submerged plant *Vallisneria spiralis*. Acta Ecologica Sinica. **26** 1005-1011.

LOUKAS A, VASILIADES L, DOMENIKIOTIS C and DALEZIOS NR (2005) Basin wide actual evapotranspiration using NOAA/AVHRR satellite data. *Phys. Chem. Earth* **30** 69-79.

MCGWIRE K, MINOR T and FENSTERMAKER L (2000) Hyperspectral mixture modeling for quantifying sparse vegetation cover in arid environments. *Remote Sens. Environ*. **72** 360-374.

MEDINA JL, CAMACHO E, RECA J, LOPEZ R and ROLDAN J (1998) Determination and analysis of regional evapotranspiration in Spain based on remote sensing and GIS. *Phys. Chem. Earth* **23** 427-432.

MERZLYAK MN, SOLOVCHENKO AE and GITELSON, AA (2003) Reflectance spectral features and non-destructive estimation of chlorophyll, carotenoid and anthocyanin content in apple fruit. *Postharvest Biol. Technol.* **27** 197-211.

NAGY, Z and JUNG, A (2005) A case study of the anthropogenic impact on the catchment of Mogyorod-brook, Hungary. *Phys. Chem. Earth* **30** 588-597.

OKIN GS, ROBERTS DA, MURRAY B, OKIN WJ (2001) Practical limits on hyperspectral vegetation discrimination in arid and semiarid environments. *Remote Sens. Environ*. **77** 212-225.

PU R, GE S, KELLY NM and GONG P (2003) Spectral absorption features as indicators of water status in coast live oak leaves. *Int. J. Remote Sens*. **24** 1799-1810.

ROUX H and DARTUS D (2006) Use of parameter optimization to estimate a flood wave: potential applications to remote sensing of rivers. *J. Hydrol*. **328** 258-266.

SCHLERF M, ATZBERGER C and HILL J (2005) Remote sensing of forest biophysical variables using HyMap imaging spectrometer data. *Remote Sens. Environ.* **95** 177-194.

SCHMIDT KS and SKIDMORE AK (2003) Spectral discrimination of vegetation types in a coastal wetland. *Remote Sens. Environ.* **85** 92– 108.

SMITH RB (2001a) Introduction to hyperspectral imaging. www.microimages.com (Accessed: 11/03/2006)

SMITH RB (2001b) Introduction to Remote sensing of the environment. www.microimages.com (Accessed 24/03/2006)

STIMSON HC, BRESHEARS DD, USTIN SL and KEFAUVER SC (2005) Spectral sensing of foliar water conditions in two co-occurring conifer species: *Pinus edulis* and *Juniperus monosperma*. *Remote Sens. Environ.* **96** 108-118.

STONE C, CHISHOLM L and COOPS N (2001) Spectral reflectance characteristics of eucalypt foliage damaged by insects. *Aust. J. Bot.* **49** 687-698.

STUMPF RP (2001) Applications of satellite ocean color sensors for monitoring and predicting harmful algal blooms. *Hum. Ecol. Risk Assess*. **7** 1363-1368.

THIEMANN S and KAUFMANN H (2002) Lake water quality monitoring using hyperspectral airborn data – a semiempirical multisensor and multitemporal approach for the Mecklenburg Lake District, Germany. *Remote Sens. Environ*. **81** 228-237.

TILLEY DR, AHMED M, SON, JH and BADRINARAYANAN H (2003) Hyperspectral reflectance of emergent macrophytes as an indicator of water column ammonia in an oligohaline, subtropical marsh. *Eco. Eng*. **21** 153-163.

UNDERWOOD E, USTIN S and DIPIETRO D (2003) Mapping nonnative plants using hyperspectral imagery. *Remote Sens. Environ*. **86** 150-161.

VOS RJ, HAKVOORT JHM, JORDANS, RWJ and IBELINGS BW (2003) Multiplatform optical monitoring of eutrophication in temporally and spatially variable lakes. *Sci. Total Environ*. **312** 221-243.

WANG K, LI Z and CRIBB M (2006) Estimation of evaporative fraction from a combination of day and night land surface temperatures and NDVI: A new method to determine the Priestly – Taylor parameter. *Remote Sens. Environ*. **102** 293-305.

YAMANO H, CHEN J and TAMURA M (2003) Hyperspectral identification of grassland vegetation in Xilinahot, Inner Mongolia, China. *Int. J. Remote Sens*. **24** 3171-3178.

ZAGOLSKI F, PINEL V, ROMIER J, ALCAYDE D, GASTELLU-ETCHEGORRY JP, GIORDANO G, MARTY G, MOUGIN E and JOFFRE R (1996) Forest canopy chemistry with high spectral resolution remote sensing. *Int. J. Remote Sens*. **17** 1107-1128.

**Appendix B**

# A comparison of satellite hyperspectral and multispectral remote sensing imagery for improved classification and mapping of vegetation

*M Govender[1], K Chetty[2], V Naiken[1] and H Bulcock[2]*

*[1]CSIR Natural Resources and the Environment, c/o School of Environmental Sciences, University of KwaZulu-Natal, Private Bag X01, Scottsville , 3209, South Africa*
*[2] School of Bioresources Engineering and Environmental Hydrology, University of KwaZulu- Natal, Private Bag X01, Scottsville , 3209, South Africa*

**Keywords:** hyperspectral, multispectral, satellite data, statistical classifiers, vegetation classification

## Abstract
In recent years the use of remote sensing imagery to classify and map vegetation over different spatial scales has gained wide acceptance in the research community. Many national and regional datasets have been derived using remote sensing data. However, much of this research was undertaken using multispectral remote sensing datasets. With advances in remote sensing technologies, the use of hyperspectral sensors which produce data at a higher spectral resolution is being investigated. The aim of this study was to compare the classification of selected vegetation types using both hyperspectral and multispectral satellite remote sensing data. Several statistical classifiers including maximum likelihood, minimum distance, mahalanobis distance, spectral angular mapper and parallelepiped methods of classification were used. Classification using mahalanobis distance and maximum likelihood methods with an optimal set of hyperspectral and multispectral bands produced overall accuracies greater than 80%.

## Introduction
Remotely sensed data are commonly used for the classification and mapping of vegetation over large spatial scales, replacing traditional classification methods, which require expensive and time-intensive field surveys. Since the early 1960s,

multispectral airborne and satellite remote sensing technologies have been used as a common source for the remote classification of vegetation (Landgrebe, 1999). Multispectral remote sensing technologies, in a single observation, collect data from three to six spectral bands from in the visible and near-infrared region of the electromagnetic spectrum. This crude spectral categorization of the reflected and emitted energy from the earth is the primary limiting factor of multispectral sensors. Over the past two decades, the development of airborne and satellite hyperspectral sensor technologies have overcome the limitations of multispectral sensors. Govender *et al*., 2007 have reviewed the application of hyperspectral imagery in the classification and mapping of land use and vegetation and, in particular in water resources studies.

Hyperspectral sensors collect several, narrow spectral bands from the visible, near-infrared, mid-infrared, and short-wave infrared portions of the electromagnetic spectrum. These sensors typically collect two hundred or more spectral bands, enabling the construction of an almost continuous spectral reflectance signature. Furthermore, narrow bandwidths characteristic of hyperspectral data permit an in-depth examination of earth surface features which would otherwise be 'lost' within the relatively coarse bandwidths acquired with multispectral data.

Hyperspectral data at a finer spectral resolution can be used to improve vegetation classification, by detecting biochemical and structural differences in vegetation (Yamano *et al*., 2003; Underwood *et al*., 2003; Galvao *et al*., 2005; Koch *et al*., 2005). Research is underway to determine if data of a higher spectral resolution could be used to discriminate vegetation at individual genus and species level.

The aim of this study was to compare the classification of selected vegetation classes using both hyperspectral and multispectral satellite remote sensing data. The two data sources selected included the Proba CHRIS (Compact High Resolution Imaging Spectrometer) multispectral sensor and Hyperion hyperspectral sensor. Supervised classification was undertaken using the statistical classifiers, maximum likelihood, minimum distance, mahalanobis distance, spectral angular mapper and parallelepiped methods.

**Methods**

**Description of study area**

The Mistley-Canema Estate is situated in the Sevenoaks district approximately 70 km from Pietermaritzburg in the KwaZulu-Natal midlands (Fig. 1).

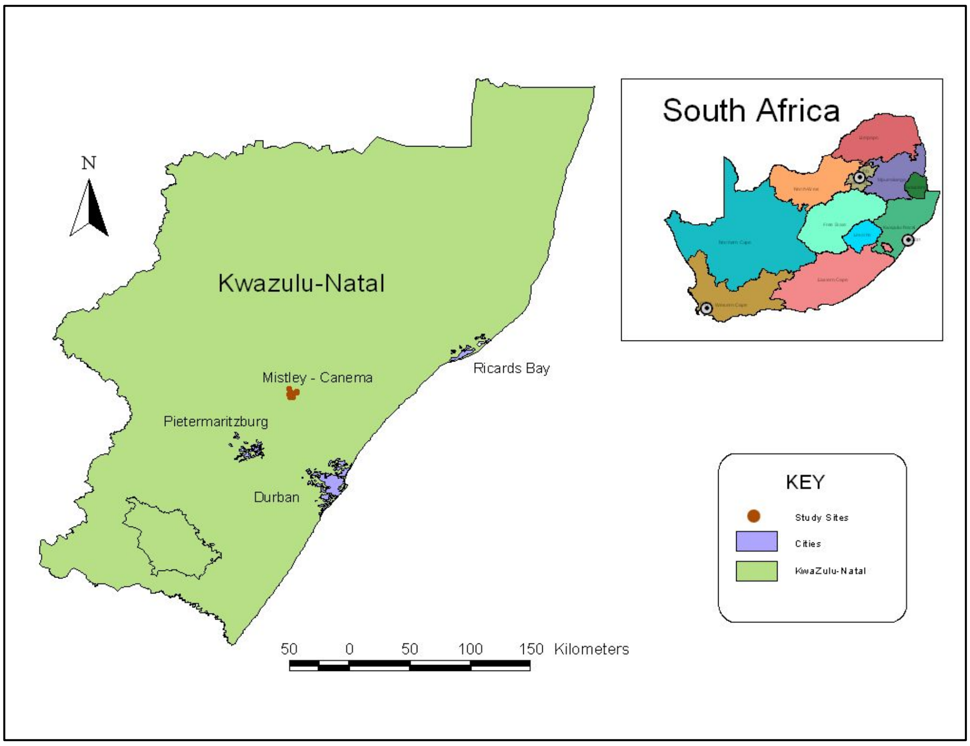



Figure 1 Location of the Mistley-Canema study site

The bio-resource group of the study area is the moist midlands mistbelt (Camp, 1997). The climate is humid, with annual rainfall ranging from 800 to 1 280 mm and a mean annual temperature of 17ºC. The natural vegetation of the area was previously *Themeda triandra* grassland. Only a few relic patches remain. The soils are highly leached and forestry and sugarcane is considered ecologically suitable and are the most widespread vegetation classes on this estate (Everson *et al*., 2006). Vegetation classes selected for this study were *Acacia mearnsii, Pinus patula, Eucalyptus macarthurii, Eucalyptus grandis* and *Saccharum*.

Thirteen different soil types were identified within the study area with the Inanda soil formation being dominant and on which a majority of the vegetation classes of interest are located. Some of the *E. grandis* sites are located on the Hutton soil form and *P. patula* and *E. macarthurri* are found on Glenrosa and Nomanci soils respectively.

**Selection of satellite hyperspectral data sources**

The spectral, spatial and temporal resolution, as well as the cost and availability of data, were considered when reviewing the most appropriate data source(s) for this research. The Hyperion hyperspectral sensor (United States Geological Survey Earth Resources Observation Systems) and the multispectral Proba CHRIS sensor (European Space Agency) were then selected for use in this study. A Hyperion image is a 42 by 7.7 km scene displaying a 30 m spatial resolution, comprising 242 spectral bands ranging from 357 to 2 576 nm. A CHRIS image is a 13 by 13 km scene, at a spatial resolution of 17 m, comprising 18 spectral bands ranging from 485.6 to 796.1 nm.

**Data collection**

Remote sensing data were acquired over the Mistley-Canema site, with concurrent ground measurements of leaf spectral reflectance required for the calibration and classification of vegetation. Ancillary data collected included geographic coordinates and ground control points, aspect, soils information and biophysical descriptions of each forestry compartment. The spectral characteristics of vegetation differ from season to season. Therefore data were collected during the winter (July 2006) and summer (February 2007) to account for plant physiological differences over the dry and wet seasons and enhanced spectral differences.

**Satellite remote sensing data**

During winter 2006, both Hyperion and Proba CHRIS imagery were acquired. The Hyperion image was acquired on 12 July 2006, with less than 25% cloud cover. Two Proba CHRIS images were acquired on 20 and 22 July 2006 using pre-selected spectral bands, which are commonly used when vegetation studies at high spatial resolutions are intended. To complement the spatial data collected during the winter season, two earth observation datasets, Hyperion and Proba CHRIS were planned for the summer field campaign. However, all three of the Hyperion images acquired in January and February 2007 contained a high percentage of cloud cover (greater than

20%) and therefore, could not be utilised in the project. Proba CHRIS images were obtained on 13 February 2007 for the summer campaign allowing for seasonal comparisons using the Proba CHRIS datasets only.

**Ground-truthing surveys**

Ground-truthing surveys should be undertaken within fifteen days of acquiring satellite remote sensing imagery (Ahmed, 2006). The winter and summer field campaigns took place on the 12$^{th}$, 25$^{th}$ and 26$^{th}$ July 2006 and, the 6$^{th}$, 7$^{th}$ and 8$^{th}$ of February 2007 respectively. The sampling method undertaken involved identifying each tree species, their respective age classes and location within the forestry compartments. The compartments selected for sampling during the winter field campaign included two *Saccharum* sites (4 years old); two *A. mearnsii* sites (6 years old); *,* three *P. patula* sites of 12, 15, and 16 years old; ; two *E. macarthurii* sites (10 and 12 years old) and an 8- year-old *E. grandis* site. The same sampling sites were used in the summer field campaign with the exception of the 16- year-old *P. patula* which was felled prior to the summer field campaign. A random sampling method was used across the compartments, with 10 samples selected in each compartment. In addition calibration sites were selected within the study area for use in atmospheric correction. The ASD handheld spectrometer was used to obtain quantitative measurements of radiant energy easily and efficiently.

**Pre-processing of remote sensing data**

**Geometric and atmospheric correction**

Images were geometrically corrected using ground control points collected during the winter field campaign and a geo-rectified spatial coverage of the forestry compartments located in the catchment. All images were projected to UTM 36S, datum WGS 84.

Interference effects of the atmosphere had resulted in significant noise levels in both datasets. The empirical line method was utilized to atmospherically correct both the Hyperion and Proba CHRIS datasets. This method compares radiance values coming back from the surface to reflectance values measured on the ground with a calibrated hand-held spectrometer (Research Systems Inc, 2005). The calibration sites which were collected using the ASD spectrometer during the winter field campaign were used to convert radiance into reflectance data.

**Data reduction**

Data reduction methods are used to manipulate high spectral resolution datasets, to reduce high data dimensionality and the complexity which arise when applying statistical classifiers to such datasets. Furthermore, by reducing data dimensionality an optimal set of spectral bands which best characterise a specific vegetation class can be identified. Two data reduction methods were applied to the hyperspectral and multispectral datasets. These included the analysis of variance and minimum noise fraction with maximum band selection.

The data-reduction procedures as outlined by Ismail *et al*., 2007 were adopted as the first method to obtain an optimal set of spectral bands for both remote sensing datasets. In this method it was assumed that a distinct collection of spectral bands can be used to discriminate between genus/species. Thirty spectral sample points per genus/species were used to determine the significant differences between spectral wavelengths using a one-way analysis of variance (95% confidence limit) with Tukey's *post hoc* test for 21 class pairs. Frequency distributions were then plotted, using the total number of significant bands at each spectral wavelength for all the classes and was used to spectrally discriminate between classes. The frequency analysis highlighted significant regions of the electromagnetic spectrum for both the Proba CHRIS and Hyperion imagery *viz*. NIR region (700 nm to 900 nm) and IR and SWIR region (900 nm to 1400 nm) respectively. The Jeffries-Matusita separability index was not utilized to test for band separability as the analysis of variance methodology sufficiently reduced the data dimensionality of both the hyperspectral and multispectral datasets (Richards and Jia, 1999; Vaiphasa *et al*., 2005; Ismail *et al*., 2007). The most significant spectral bands with the highest frequencies were selected for each individual class. The Hyperion hyperspectral dataset reduced significantly from 242 to 27 spectral bands (referred to as Hyperion reduced set A), ranging from 900 to 1400 nm, while the Proba CHRIS multispectral data reduced from 18 to 9 spectral bands, ranging from 450 to 550 nm and 700 to 800 nm.

The second method applied only to the hyperspectral data, used minimum noise fraction to segregate noise in the datasets, identify and remove these 'noisy bands' and to reduce the computational requirements for analyses as adopted by (Green *et al*., 1988; Lee *et al*., 1990). In addition to removing the 'noisy bands' from the Hyperion dataset, an optimal subset of bands was determined using known regions of interest for vegetation classes (targets) and background materials in the image. The optimal spectral subset included a reduced dataset in which 'noise' had been removed and which best characterized the vegetation classes selected for this study. This approach increased the accuracy of the classification process. The Hyperion hyperspectral dataset reduced significantly from 242 to 23 spectral bands (referred to as Hyperion reduced Set B), ranging from 793 nm to 2 375 nm.

**Procedures for classification**

Regions of interest were selected for the training and validation phases performed during supervised classification. Regions of interest may be defined within the context of remote sensing as areas of an image which contain pixels of the same spectral characteristics which represent the same vegetation type.

Regions of interest were selected using ground control points and compartment boundary overlays of the Mistley-Canema estate, which were verified during field surveys. These regions were selected to represent 8-year-old *E. grandis,*10- and 12-year-old *E. macarthurri*, 12- and 16-year-old *P. patula*, 4-year- old *Saccharum* and 6-year-old *A. mearnsii*.

Two statistical measures were used to assess the spectral separability of the selected regions of interest: the Jeffries-Matusita and transformed divergence separability measures (Richards and Jia, 1999). All regions of interest when analysed produced a high spectral separability.

A review of several studies on vegetation classification indicates that different statistical classifiers perform well with different vegetation types and using different remote sensing data sources (Kruce *et al*., 1993; Lillesand and Kiefer, 1999; Castro-Esau, 2004; South *et al*., 2004; Belluco *et al*., 2006). Due to this variability, five commonly used statistical classifying techniques were used in the training phase of the classification process. These techniques included maximum likelihood, minimum distance, mahalanobis distance, spectral angular mapper and parallelepiped. An iterative process was followed to determine the optimized parameters required for each statistical classifier. Thomlinson *et al*. (1999) suggested an overall accuracy of greater than 85% as indicator of superior classification. However, in this study only those statistical classifying techniques with an overall accuracy greater than 80% during the training phase were selected for validation. Accuracy assessments were used to determine which statistical classifier produced best classification results. The kappa coefficient of agreement and the overall accuracy (> 80%) were the two statistics most prominently used for decision making purposes (Foody and Mathur, 2006). The user's accuracy was used to evaluate how well an individual vegetation class was classified for a specific statistical classifier method.

**Results and discussion**

**Training phase of classification**

Training regions were spatially subset from the Hyperion and Proba CHRIS images. Regions of interest were selected to represent 8-year-old *E. grandis*, 10- and 12- year-old *E. macarthurri*, 12- and 16-year-old *P. patula*, 4-year-old *Saccharum* and 6-year-old *A. mearnsii*. An iterative process was followed to determine the optimized parameters required for each of the five statistical classifier methods. The performance of each statistical classifying method was evaluated using accuracy assessments (>80% overall accuracy) as shown in Table 1. The training results of two Hyperion reduced Datasets A (27 spectral bands) and B (23 spectral bands) are shown independently in Table 1. From the training phase, the mahalanobis distance and the maximum likelihood methods most accurately classified the vegetation classes used in this study, achieving overall accuracies above 80%. For all four datasets, the parallelepiped method did not perform well, achieving overall accuracies ranging from 3% to 64%.

Table 1 Accuracy assessment: training phase

| Supervised classification methods | Proba CHRIS (winter 2006 survey) | | Proba CHRIS (summer 2007 survey) | | Hyperion reduced set A (winter 2006 survey) | | Hyperion reduced set B (winter 2006 survey) | |
|---|---|---|---|---|---|---|---|---|
| | Kappa coefficient | Overall accuracy (%) | Kappa coefficient | Overall accuracy (%) | Kappa coefficient | Overall accuracy (%) | Kappa coefficient | Overall accuracy (%) |
| Parallelepiped | 0.55 | 63.92 | 0.56 | 62.82 | 0.14 | 28.48 | 0.03 | 3.98 |
| Minimum distance | 0.88 | 90.36 | 0.58 | 65.13 | 0.38 | 46.86 | 0.53 | 59.42 |
| Mahalanobis distance | 0.90 | 91.91 | 0.91 | 92.75 | 0.83 | 85.45 | 0.78 | 80.86 |
| Maximum likelihood | 0.99 | 99.22 | 0.97 | 97.48 | 0.99 | 98.93 | 0.99 | 98.93 |
| Spectral angle mapper | 0.75 | 75.21 | 0.76 | 77.36 | 0.59 | 64.93 | 0.39 | 48.09 |

The classification of the vegetation classes using the different classifying techniques were further verified by mapping individual vegetation classes across the training area and comparing their spatial distribution to true colour images and spatial maps of forestry compartments obtained from Mondi. Figure 2 illustrates the classification and mapping of the best discriminated vegetation classes using Proba CHRIS and Hyperion remote sensing imagery. Each classification can be compared to the true colour image shown in Fig. 2a. Figures 2b, d, f, h represent the mahalanobis distance classifications using Proba CHRIS images obtained during winter and summer, and the Hyperion reduced Datasets A and B acquired in winter respectively. Figs. 2c, e, g, j represent the maximum likelihood classifications using Proba CHRIS images obtained during winter and summer, and the Hyperion reduced Datasets A and B acquired in winter respectively.

For the Proba CHRIS multispectral images, the mahalanobis classifier consistently classified the *A. mearnsii* and *Saccharum* satisfactorily during winter (Fig. 2b) and summer (Fig. 2d). Conversely, the maximum likelihood method only classified *A. mearnsii* adequately, for the same set of images (Figs. 2c and e). Furthermore, the maximum likelihood method satisfactorily classified the 10- and 12-year-old *E. macarthurii* and 12-year-old *P. patula* species during winter (Fig. 2c); and 12-year-old *E. macarthurii* during summer (Fig. 2d). Due to *E. macarthurii* compartments being felled prior to the summer campaign, the proportions of this species classified and mapped during winter and summer is variable.

The Hyperion hyperspectral reduced Dataset A (27 spectral bands ranging from 900 to 1 400 nm) best classified *A. mearnsii* and *Saccharum* using the mahalanobis distance method (Fig. 2f); and *A. mearnsii* and the 12-year-old *P. patula* using the maximum likelihood method (Fig. 2g). Similarly, *A. mearnsii* was satisfactorily classified when applying both statistical classifiers to the reduced Hyperion Dataset B which comprised of 23 spectral bands ranging from 793 nm to 2 375 nm (Figs. 2h and j).

The training results imply that classification of different genera and species are affected by optimal band selection and the statistical classifier method adopted. Furthermore, discriminating between different genera and species is influenced by seasonal physiological changes in vegetation which are highlighted during the dry winter and wet summer seasons.

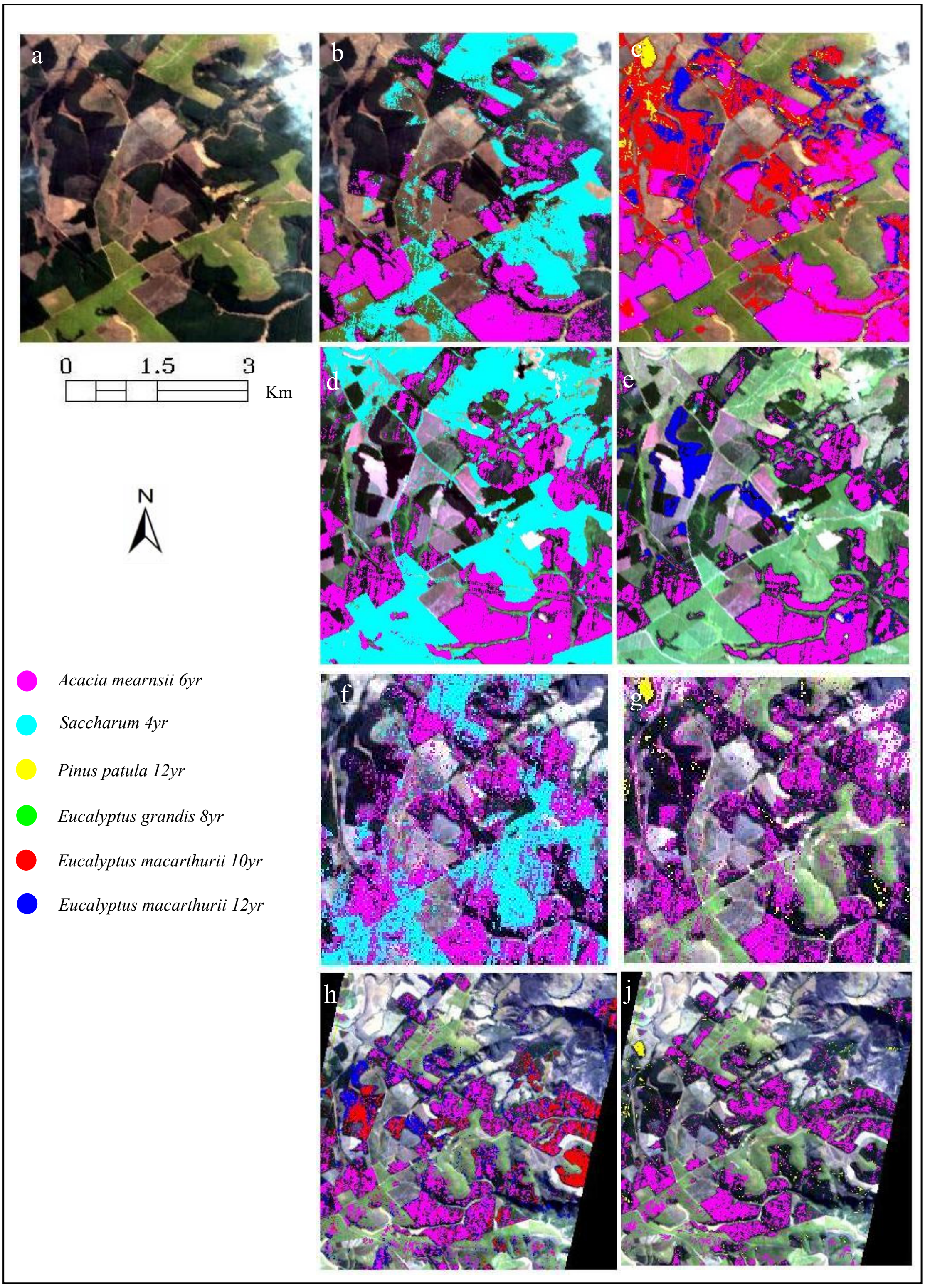



Figure 2 The classification and mapping of the best discriminated vegetation classes using Proba CHRIS and Hyperion remote sensing imagery during the training phase: a) true colour image, Proba CHRIS winter classification using the b) mahalanobis distance and c) maximum likelihood classification, Proba CHRIS summer classifications using d)

mahalanobis distance and e) maximum likelihood, Hyperion reduced Set A winter classification using f) mahalanobis distance and g) maximum likelihood and, Hyperion reduced set B winter classification using h) mahalanobis distance and i) maximum likelihood methods.

**Validation phase of classification**

The mahalanobis distance and maximum likelihood statistical classifiers were selected from results of the training phase for validation. Validation was performed using the entire study area. Accuracy assessments were used to evaluate each classification as shown in Table 2. Both methods performed well, with overall accuracies above 80%. However, for all 4 datasets the maximum likelihood classifications achieved the highest overall accuracies above 98%.

Table 2 Accuracy assessments: validation phase

| Supervised classification methods | Proba CHRIS (winter 2006 survey) | | Proba CHRIS (summer 2007 survey) | | Hyperion reduced set A (winter 2006 survey) | | Hyperion reduced set B (winter 2006 survey) | |
|---|---|---|---|---|---|---|---|---|
| | Kappa coefficient | Overall accuracy (%) | Kappa coefficient | Overall accuracy (%) | Kappa coefficient | Overall accuracy (%) | Kappa coefficient | Overall accuracy (%) |
| Mahalanobis distance | 0.83 | 85.89 | 0.90 | 91.60 | 0.88 | 89.13 | 0.78 | 80.86 |
| Maximum likelihood | 0.97 | 98.80 | 1.00 | 99.86 | 0.98 | 98.68 | 0.99 | 98.93 |

The classifications of each vegetation class at their respective age groups were evaluated to determine how accurately the statistical classifiers, mahalanobis distance and maximum likelihood performed at a species level. This evaluation was undertaken by comparing the user’s accuracy for each for each vegetation class (Table 3), with a visual assessment of their spatial distribution.

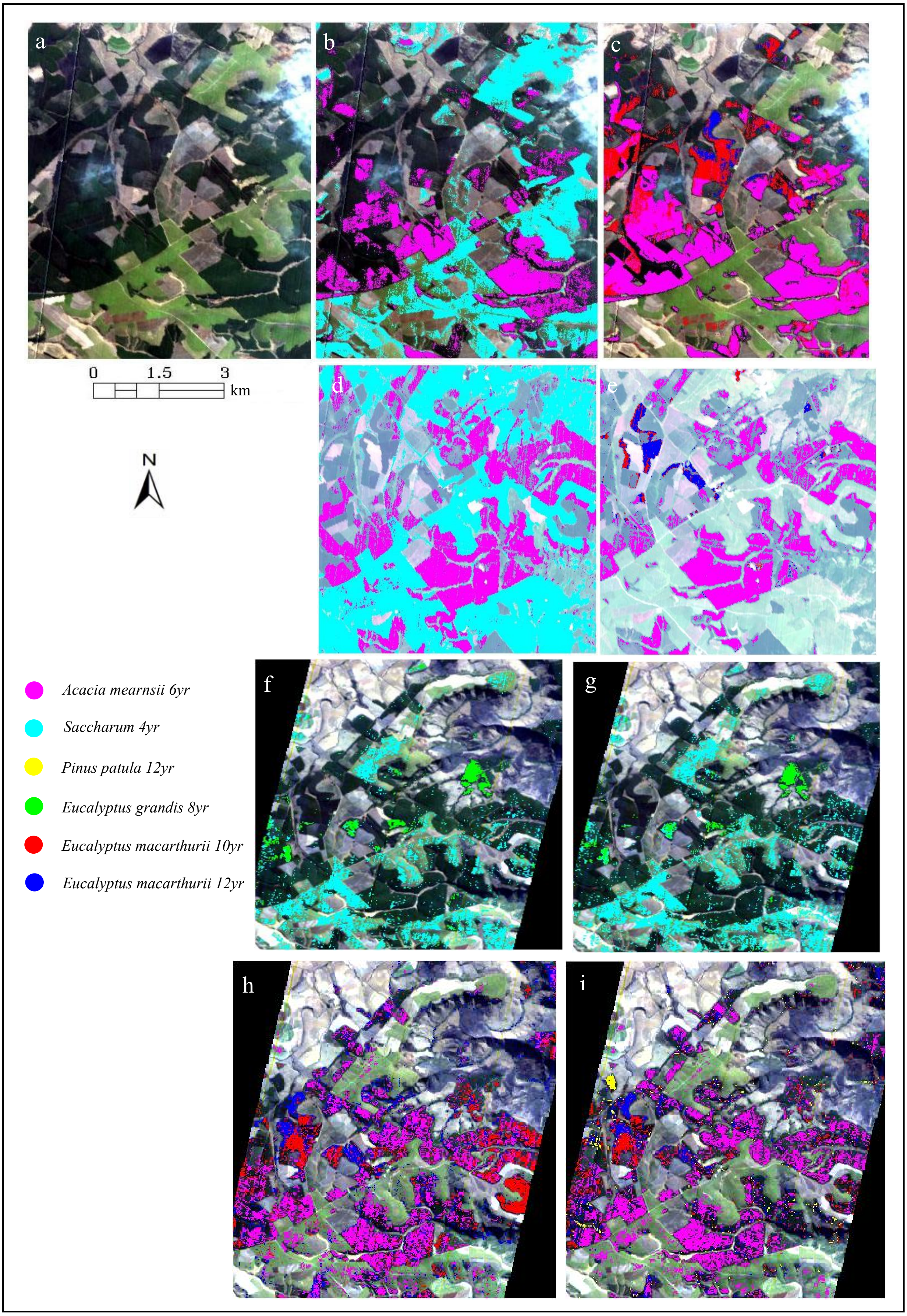
Table 3 User Accuracy: validation phase

| Vegetation classes | User accuracy (%) Proba CHRIS (winter 2006 survey) | | User accuracy (%) Proba CHRIS (summer 2007 survey) | | User accuracy (%) Hyperion reduced set A (winter 2006 survey) | | User accuracy (%) Hyperion reduced set B (winter 2006 survey) | |
|---|---|---|---|---|---|---|---|---|
| | Mahalanobis Distance | Maximum Likelihood | Mahalanobis Distance | Maximum Likelihood | Mahalanobis Distance | Maximum Likelihood | Mahalanobis Distance | Maximum Likelihood |
| 8-year-old *Eucalyptus grandis* | Cloud cover over site | Cloud cover over site | 100.00 | 100 | 96.39 | 100 | 77.65 | 100 |
| 10- year - old *Eucalyptus macarthurri* | 78.40 | 100 | 76.14 | 99.04 | 76.67 | 99.07 | 80.70 | 99.08 |
| 12- year- old *Eucalyptus macarthurri* | 76.56 | 96.24 | 68.03 | 100 | 88.64 | 97.73 | 80.46 | 96.63 |
| 12- year- old *Pinus patula* | 78.50 | 100 | 100 | 100 | 81.05 | 100 | 75.00 | 98.81 |
| 16- year - old *Pinus patula* | 75.00 | 100 | 100 | 100 | 76.14 | 96.91 | 75.64 | 98.95 |
| 4 –year- old *Saccharum* | 100.00 | 100 | 100 | 100 | 97.67 | 100 | 97.80 | 100 |
| 6 –year- old *Acacia mearnsii* | 100.00 | 100 | 100 | 100 | 86.02 | 98.92 | 78.57 | 98.82 |

The classifications from the validation phase were mapped across the Mistley Canema estate. Figure 3 illustrates the best discriminated vegetation classes when applying the mahalanobis distance and maximum likelihood classifiers for the Proba CHRIS and Hyperion remote sensing imagery. These classifications can be compared to true colour images in Fig. 3a. Figures 3b, d, f, h represent the mahalanobis distance classifications using Proba CHRIS images obtained during winter and summer, and the Hyperion reduced Datasets A and B acquired in winter respectively. Figures 3c, e, g, i represent the maximum likelihood classifications using Proba CHRIS images obtained during winter and summer, and the Hyperion reduced Datasets A and B acquired in winter respectively. The trends identified for each dataset are discussed in the subsequent paragraphs.

During winter 2006 and summer 2007, the mahalanobis distance method best classified 4-year-old *Saccharum* and 6-year-old *A. mearnsii* using the Proba CHRIS multispectral dataset (Figs. 3b and d) .In contrast, the maximum likelihood method best classified the 6-year-old *A. mearnsii* and 10- and 12-year- old *E. macarthurri* during winter (Fig. 3c); and 6-year-old *A. mearnsii* and 12-year-old *E. macarthurri* during summer (Fig. 3e). However, the proportion of *E. macarthurri* classified and mapped during winter compared to summer is different (Figs. 3c and 3e), as a result of selected compartments being felled prior to the summer field campaign. In general, when applying both statistical classifiers to the Proba CHRIS multispectral imagery, 6-year-old *A. mearnsii* is classified and mapped satisfactorily during the dry and wet season.

The results of validation phase for the Hyperion reduced Dataset A (27 spectral bands ranging from 900 to 1 400 nm) and B (23 spectral bands ranging from 793 nm to 2 375 nm) acquired during winter clearly demonstrate the important of utilizing an optimal set of spectral bands to best discriminate between vegetation classes. The general both statistical classifiers best classify 4-year-old *Saccharum* and 8-year-old *E. grandis* when using the Hyperion reduced Dataset A (Figs. 3f and g). In contrast to 6-year-old *A. mearnsii* and 12-year-old *E. macarthurri* best classified using the mahalanobis distance and maximum likelihood methods with the Hyperion reduced Dataset B (Figs. 3h and i). Furthermore, when applying the maximum likelihood method to the Hyperion reduced Dataset B, it was possible to classify and map 12-year-old *P. patula* species as shown in Fig. 3i.



a
b
c
0
1.5
3
km
N
d
e
Acacia mearnsii 6yr
Saccharum 4yr
Pinus patula 12yr
Eucalyptus grandis 8yr
Eucalyptus macarthurii 10yr
Eucalyptus macarthurii 12yr
f
g
h
i

Figure 3 The classification and mapping of each vegetation class using Proba CHRIS and Hyperion remote sensing imagery during the validation phase: a) true colour image, Proba CHRIS winter classification using b) mahalanobis distance and c) maximum likelihood classification, proba CHRIS summer classifications using d) mahalanobis distance and e) maximum likelihood, Hyperion reduced set A winter classification f) mahalanobis distance and g) maximum likehood and, Hyperion reduced set B winter classification h) mahalanobis distance and i) maximum likehood methods.

The validation results indicate that each remote sensing dataset with its unique combination of spectral bands have classified and mapped the individual vegetation classes to different degrees of accuracy. Variability exists within a season and between statistical classifiers across the dry and wet seasons. The validation results indicate an overall trend of genus level classification when using the Proba CHRIS multispectral bands, compared to genus and species level classification when using a unique set of Hyperion hyperspectral bands.

**Conclusions**

The classification and mapping of the different genera and species selected in this study were strongly affected by seasonal physiological changes in vegetation as illustrated using the multispectral Proba CHRIS remote sensing imagery, the range of spectral bands used in the classification and, the use of different statistical classifier methods. The discrimination between vegetation classes surveyed during winter and summer using the multispectral data indicate that the spectral characteristics of vegetation are significantly influenced by seasonal physiological changes in vegetation which are most prominent during the dry and wet seasons. The results from the two data reduction methods illustrate that different combinations of spectral bands which were extracted from narrow-band hyperspectral data and the broad-band multispectral data can classify and map genera and species to different degrees of accuracy. In this study the key spectral wavelengths which best discriminate *Acacia mearnsii, Pinus patula, Eucalyptus macarthurii, Eucalyptus grandis* and *Saccharum* using the Hyperion hyperspectral data and Proba CHRIS multispectral data were identified. In general, the results from this study suggest that multispectral remote sensing data generally produced genus level classification whereas hyperspectral remote sensing data facilitated genus and species level classification. This implies that classification at a finer spectral resolution using an optimal and unique set of spectral bands is possible and can improve vegetation classification. Different statistical classifier methods and vegetation spectral reflectance indices were investigated in this project. The statistical classifiers which best classified *Acacia mearnsii, Pinus patula, Eucalyptus macarthurii, Eucalyptus grandis* and *Saccharum* during winter and summer included the mahalanobis distance and maximum likelihood methods.

It is also recommended, that the developers of remote sensing technologies should utilize the findings of these type of studies in order to optimize sensor systems. Optimal hyperspectral bands which best characterize different vegetation classes should be incorporated into multispectral systems through sensor modifications or spectral filters. Using this approach researchers will no longer be restricted to existing sensor specifications, but can tailor remote sensing model inputs to a particular application. The ultimate goal is to move from impractical large data volumes and expensive sensor development towards practical operational low data volumes, high temporal resolution and cheaper sensor development.

**Acknowledgements**

The authors wish to thank Water Research Commission for funding this study. This work is also supported by the European Space Agency (ESA ®) through the Tiger initiative, and acknowledgement is made for the use of Category 1Proba CHRIS data. The authors would also like to acknowledge and thank Dr Jan van Aardt and Dr Onisimo Mutanga for their comments and support in this study and the Agricultural Research Council for the use of ASD spectrometer during the field campaigns.

## References


AHMED F (2006) Personal communication. University of KwaZulu-Natal, Pietermaritzburg, South Africa.

BELLUCO E, CAMUFFO M, FERRARI S, MODENESE L, SILVESTRI S, MARANI A and MARANI M (2006) Mapping salt-marsh vegetation by multispectral and hyperspectral remote sensing. *Remote Sens. Environ*. **105** 54-67.

CAMP KGT (1997) The Bioresources Groups of KwaZulu-Natal. Cedara Report N/A/97/6. KwaZulu-Natal Department of Agriculture, Pietermaritzburg, South Africa.

CASTRO-ESAU KL, SANCHEZ-AZOFEIFA GA and CAELLI T (2004) Discrimination of lianas and trees with leaf level hyperspectral data. *Remote Sens. Environ.* **90** 353-372.

EVERSON CS, MOODLEY M, GUSH MB, JARMAIN CJ, GOVENDER M and DYE PJ (2006) Can effective management of riparian zone vegetation significantly reduce the cost of catchment management and enable greater productivity of land resources. WRC Report No. K5/1284. Water Research Commission, Pretoria, South Africa.

FOODY GM and MATHUR A (2006) The use of training sets containing mixed pixels for accurate hard image classification: Training on mixed spectral responses for classification by SVM. *Remote Sens. Environ*. **103** 179-189.

GALVAO LS, FORMAGGIO AR and TISOT DA (2005) Discrimination of sugarcane varieties in South Eastern Brazil with EO-1 Hyperion data. *Remote Sens. Environ*. **94** 523-534.

GOVENDER M, CHETTY K and BULCOCK H (2007) A review of hyperspectral remote sensing and its application in vegetation and water resource studies. *Water SA* **33** 145-151.

GREEN AA, BERMAN M, SWITZER P and CRAIG MD (1988) A transformation for ordering multispectral data in terms of image quality with implications for noise removal. *IEEE Trans. on Geosci. and Remote Sens.* **26** 65-74.

ISMAIL R, MUTANGA O and AHMED F (2007) Discriminating Sirex noctilio attack in pine forest plantations in South Africa using high spectral resolution data. In: Kalacska M and Sanchez-Azofeifa A (eds.) *Hyperspectral Remote Sensing of Tropical and Sub-Tropical Forests.* Taylor and Francis: CRC Press, Routledge, USA. p. 350.

KOCH M, INZANA J and EL BAZ F (2005) Applications of hyperion hyperspectral and aster multispectral data in characterizing vegetation for water resource studies in arid lands. Paper No. 44-5. *Proc. Geol. Remote Sens*. Annual Meeting. 16-19 October 2005, Salt Lake City, USA.

KRUCE F, LEFKOFF A, BOARDMAN J, HEIDEBRECHT K, SHAPIRO A, BARLOON P and GOETZ A (1993) The spectral image processing system (SIPS) interactive visualization and analysis of imaging spectrometer data. *Remote Sens. Environ*. **44** 145-163.

LANDGREBE D (1999) On information extraction principles for hyperspectral data. *Cybernetics* **28** part c, 1, 1-7.

LEE JB, WOODYATT AS and BERMAN M (1990) Enhancement of High Spectral Resolution Remote Sensing Data by a Noise – Adjusted Principal Components Transform. *IEEE Trans. on Geosci. and Remote Sens*. **28** 295-304.

LILLESAND TM and KIEFER RW (1999) *Remote Sensing and Image Interpretation.* John Wiley & Sons, Inc. New Jersey, USA.

RESEARCH SYSTEMS INC. (2005) *RSI Training Series: Spectral Analysis with ENVI*. 4990 Pearl East Circle, Boulder, CO 80301, USA,106pp.

RICHARDS JA and JIA X (1999) *Remote Sensing Digital Image Analysis*. Springer, Berlin, Germany, 363pp.

SOUTH S, QI J and LUSCH DP (2004) Optimal classification methods for mapping agricultural tillage practices. *Remote Sens. Environ*. **91** 90-97.

THOMLINSON JR, BOLSTAD PV and COHEN WB (1999) Coordinating methodologies for scaling landcover classifications from site-specific to global: steps toward validating global map products *Remote Sens. Environ.* **70** 16-28.

UNDERWOOD E, USTIN S and DIPIETRO D (2003) Mapping nonnative plants using hyperspectral imagery. *Remote Sens. Environ*. **86** 150-161.

VAIPHASA C, ONGSOMWANG S, VAIPHASA T and SKIDMORE AK (2005) Tropical mangrove species discrimination using hyperspectral data: a laboratory study. *Estuarine, Coastal and Shelf Sci*. **65** 371-379.

YAMANO H, CHEN J and TAMURA M (2003) Hyperspectral identification of grassland vegetation in Xilinahot, Inner Mongolia, China. *Int. J. Remote Sens*. **24** 3171-3178.