

AgroP-11



**AN ASSESSMENT OF VARIOUS IMAGE  
TRANSFORMATION TECHNIQUES  
ON LANDSAT TM DATA  
FOR FOREST CLASSIFICATION**

*A case Study in Kobernausserwald Area, Austria*

by

**AGNELO FERNANDES**



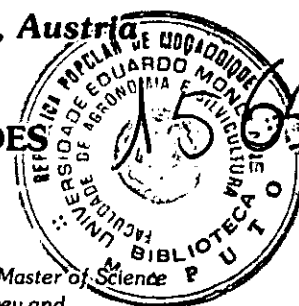
**INTERNATIONAL INSTITUTE FOR AEROSPACE SURVEY AND EARTH SCIENCES**

✓  
\* 634.0.6  
Feb  
Agrop-11

# AN ASSESSMENT OF VARIOUS IMAGE TRANSFORMATION TECHNIQUES ON LANDSAT TM DATA FOR FOREST CLASSIFICATION

*A case study in Kobernausserwald Area, Austria*

by  
**AGNELO DOS MILAGRES FERNANDES**



Submitted as a partial fulfilment of the requirements for the degree of Master of Science  
in Forest Survey at International Institute for Aerospace Survey and  
Earth Sciences (ITC), Enschede, The Netherlands

cidade Paris

## **Board of Examiners:**

**Dr. Ir. Alfred de Gier**

(Chairman of the Board - Head of Forest Science Division and Director of Studies)

**Prof. Ir. A. van Maaren**

(External Examiner - Wageningen Agriculture University)

**Dr. Yousif Ali Hussin**

(Supervisor - Forest Science Division, Department of Land Resources and Urban Sciences)

**Mr. Michael J. C. Weir, M.Sc.**

(Forest Science Division, Department of Land Resources and Urban Sciences)



**DEPARTMENT OF LAND RESOURCES AND URBAN SCIENCES  
FOREST SCIENCE DIVISION  
INTERNATIONAL INSTITUTE FOR AEROSPACE SURVEY AND EARTH SCIENCES (ITC)  
ENSCHDE, THE NETHERLANDS**

## ABSTRACT

---

LANDSAT Thematic Mapper data, collected over Kobernausserwald area in Austria, 13 May 1993, were digitally analysed using different techniques (linear transformations) such as Normalized Difference Vegetation Index, Vegetation Index, Transformed Vegetation Index, Principal Component Analysis, Tasselled Cap Transformations and untransformed bands 2, 3 and 4.

The above techniques were used for digital forest classification of the area with maximum likelihood classifier and then assessed to know their accuracies. The assessment were made through two statistical methods: confusion matrix and kappa statistics. Confusion matrix method gave good result in tasselled cap transformations and principal component analysis with overall accuracy more than 85%.

It was found conclusively that untransformed bands of LANDSAT TM data was quite useful and successful for forest classification. However it can be concluded the tasselled cap is the more recommendable one because of it's high overall accuracy (>90%).

## ACKNOWLEDGMENT

---

I would like to express my sincere gratitude to Food and Agriculture Organization of the United Nations who granted me a fellowship for two consecutive years for the completion of my studies at ITC.

My special thanks goes to Dr. Yousif Hussin who helped me to formulate the research topic, for his continuous supervision and precious comments during the research period without any reservation.

I wish to thank Dr. Ir. A. de Gier who advised me to reformulate the research problem and for his critical point of view, which helped to improve this research.

I would like to thank the staff of forestry division who guided me during my fieldwork period in Austria, in especially Mr. Michael Wier, M.Sc also for his good scottish english in correcting the research proposal and Ir. J. Remeijn also to provide the materials related to field area.

To Susana Barrera and Erasmo Rodríguez being a part of the remained community in my life, here in Holland when Michelle was absent, I thank them.

To Martinha and Moisés (the Mozambican community at ITC) for sharing the difficult moments of my life and cups of coffee during break hour.

Thanks are also due to my classmates, Adel, Acharya, Alma, Ali, Hargyono and Agung who remained together during the research period.

To the community (Jagdeesh, Juan, Montser, Monica, Moshe, Roxanna and Chris) for it's existence and for the beers shared during one year.

I would like to express my sincere gratitude to all those who helped me to pursue my studies at ITC.

To my mother, brothers, sisters in law and nephews to be so far and waiting patiently for my return.

Finally, my special warm thanks goes to my sweet wife, Michelle, for being patient during the first 14 months of my studies, when far away, and for the last 3 months in sharing her happiness with me in Holland.

## TABLE OF CONTENTS

---

|  |      |
|--|------|
| <b>ABSTRACT</b> .....                              | i    |
| <b>ACKNOWLEDGMENT</b> .....                        | ii   |
| <b>TABLE OF CONTENTS</b> .....                     | iii  |
| <b>LIST OF TABLES</b> .....                        | vi   |
| <b>LIST OF FIGURES</b> .....                       | viii |
| <b>CHAPTER 1: INTRODUCTION</b> .....               | 1    |
| 1.1. Justification of the Research .....           | 1    |
| 1.2. Objectives of the Research .....              | 2    |
| 1.3. Research Questions .....                      | 3    |
| 1.4. Description of the area of study .....        | 4    |
| 1.4.1. Physiography .....                          | 5    |
| 1.4.2. Climate, soil, vegetation .....             | 5    |
| 1.4.3. History and administration .....            | 7    |
| 1.4.4. Silviculture and Forest Management .....    | 8    |
| 1.4.4.1. Friedburg Forest District .....           | 9    |
| <b>CHAPTER 2: THEORY AND CONCEPTS</b> .....        | 11   |
| 2.1. General .....                                 | 11   |
| 2.2. Reflectance of the Main Ground Features ..... | 12   |
| 2.2.1. Vegetation .....                            | 12   |
| 2.2.3. Soils .....                                 | 14   |
| 2.2.4. Water .....                                 | 15   |
| 2.3. Satellite Imagery .....                       | 16   |

|   |           |
|---|-----------|
| 2.4. Linear Transformations . . . . .                     | 18        |
| 2.4.1. Rationing . . . . .                                | 18        |
| 2.4.1.1. Ratio Vegetation Index . . . . .                 | 19        |
| 2.4.1.2. Normalized Difference Vegetation Index . . . . . | 21        |
| 2.4.1.3. Transformed Vegetation Index . . . . .           | 22        |
| 2.4.2. Perpendicular Vegetation Index . . . . .           | 23        |
| 2.4.3. Tasselled Cap Transformation . . . . .             | 24        |
| 2.4.4. Principal Component Analysis . . . . .             | 27        |
| 2.5. Satellite Image Classification . . . . .             | 28        |
| 2.5.1. Unsupervised Classification . . . . .              | 28        |
| 2.5.2. Supervised Classification . . . . .                | 29        |
| 2.5.2.1. Box classification . . . . .                     | 31        |
| 2.5.2.2. Minimum Distance Classification . . . . .        | 32        |
| 2.5.2.3. Maximum Likelihood Classification . . . . .      | 33        |
| <b>CHAPTER 3: MATERIALS AND METHODOLOGY . . . . .</b>     | <b>35</b> |
| 3.1. Data . . . . .                                       | 35        |
| 3.2. Methodology . . . . .                                | 36        |
| 3.2.1. Image Pre-processing . . . . .                     | 37        |
| 3.2.1.1. Geometrical corrections . . . . .                | 37        |
| 3.2.1.2. Stretching . . . . .                             | 38        |
| 3.2.2. Image Processing . . . . .                         | 39        |
| 3.2.2.1. Band rationing . . . . .                         | 39        |
| 3.2.2.2. Tasselled Cap Transformation . . . . .           | 40        |
| 3.2.2.3. Principal Component Analysis . . . . .           | 41        |
| 3.2.2.4. Classification . . . . .                         | 41        |
| 3.2.3. Evaluation of the Classification Results . . . . . | 45        |
| 3.2.3.1. Quantitative Approach . . . . .                  | 45        |
| <b>CHAPTER 4: PRESENTATION OF THE RESULTS . . . . .</b>   | <b>49</b> |
| 4.1. General Remarks . . . . .                            | 49        |

|  |           |
|--|-----------|
| 4.2. Image Processing .....  | 49        |
| 4.3. Classified Maps .....   | 50        |
| 4.3.1. Untransformed Bands .....   | 51        |
| 4.3.2. NDVI, Bands 5 and 7 .....   | 52        |
| 4.3.3. VI, Bands 5 and 7 .....   | 53        |
| 4.3.4. TVI, Bands 5 and 7 .....  | 54        |
| 4.3.5. PC1, PC2 and Band 5 .....   | 55        |
| 4.3.6. Tasselled Cap Transformations .....   | 56        |
| 4.4. Statistic Results .....   | 57        |
| 4.4.1. Confusion Matrix for Classification of<br>Untransformed Bands .....           | 58        |
| 4.4.2. Confusion Matrix for Classification of<br>NDVI, Bands 5 and 7 .....           | 62        |
| 4.4.3. Confusion Matrix for Classification of<br>VI, Bands 5 and 7 .....             | 64        |
| 4.4.4. Confusion Matrix for Classification of<br>TVI, Bands 5 and 7 .....            | 66        |
| 4.4.5. Confusion Matrix for Classification of<br>PC1, PC2 and Band 5 .....           | 68        |
| 4.4.6. Confusion Matrix for Classification of<br>Tasselled Cap Transformations ..... | 70        |
| <b>CHAPTER 5: DISCUSSION AND CONCLUSIONS .....</b>                                   | <b>72</b> |
| 5.1. Introduction .....  | 72        |
| 5.2. Discussion of the Results .....   | 72        |
| 5.2.1. Accuracy Assessment .....   | 73        |
| 5.3. Conclusions .....   | 75        |
| <b>BIBLIOGRAPHIC REFERENCES .....</b>  | <b>78</b> |
| <b>APPENDICES .....</b>  | <b>83</b> |

## LIST OF TABLES

---

|  |    |
|--|----|
| Table 1 - The seven spectral regions of LANDSAT TM . . . . .   | 17 |
| Table 2 - Coefficients for LANDSAT TM (Crist and Cicone, 1984) . . . . .   | 25 |
| Table 3 - Forest codes for classification . . . . .  | 43 |
| Table 4 - Number of true pixels per class collected to be verified over all<br>area. . . . .   | 46 |
| Table 5 - Variability among the new generated bands. . . . .   | 50 |
| Table 6a - Accuracy calculated through confusion matrix for classification of<br>untransformed bands of LANDSAT TM (2, 3 & 4). . . . . | 58 |
| Table 6b - The proportion for each class obtained from table 6a, in order to<br>calculate $\text{var}(K)$ . . . . .                    | 58 |
| Table 7 - Accuracy per class for classification of NDVI, bands 5 and 7 of<br>LANDSAT TM. . . . .                                       | 62 |
| Table 8 - Accuracy per class for classification of VI, bands 5 and 7 of<br>LANDSAT TM. . . . .   | 64 |
| Table 9 - Accuracy per class for classification of TVI, bands 5 and 7 of<br>LANDSAT TM. . . . .  | 66 |
| Table 10 - Accuracy per class for classification of PC1, PC2 and Band 5 of<br>LANDSAT TM. . . . .                                      | 68 |



|  |    |
|--|----|
| Table 11 - Accuracy per class for classification of Tasseled Cap<br>Transformations of LANDSAT TM. ....  | 70 |
| Table 12 - Performance of the accuracy in different classes for different<br>classifications .....       | 74 |
| Table 13 - Resume of overall accuracy for different maps using different<br>linear transformations. .... | 75 |
| Table 14 - Resume of the results through Kappa statistics. ....  | 75 |

## LIST OF FIGURES

---

|   |    |
|---|----|
| Figure 1. Location of the study area . . . . .  | 4  |
| Figure 2. Precipitation over 25 years . . . . .   | 6  |
| Figure 3. Temperature over 25 years . . . . .   | 6  |
| Figure 4. An idealized reflectance curve from a healthy vegetation . . . . .  | 13 |
| Figure 5. Major influences on spectral properties of the living leaf . . . . .  | 14 |
| Figure 6. Soil reflectance from a fine silty soil . . . . .   | 15 |
| Figure 7. Spectral reflectance curves of clear and turbid water . . . . .   | 16 |
| Figure 8. Relationship between VI and bare soil NIR reflectance for various<br>vegetation density levels . . . . .                                    | 20 |
| Figure 9. The concept of Perpendicular Vegetation Index . . . . .   | 23 |
| Figure 10. TM tasselled cap brightness, greenness and wetness functions<br>define two planes - the plane of vegetation and the plane of soils . . . . | 26 |
| Figure 11. Basic steps in supervised classification . . . . .   | 31 |
| Figure 12. Box classification. Ranges values within training data define<br>decision boundaries . . . . .   | 31 |
| Figure 13. Minimum distance classifier . . . . .  | 32 |

|   |    |
|---|----|
| Figure 14. Probability density functions defined by a maximum likelihood classifier . . . . .   | 34 |
| Figure 15. Schematic flowchart representing the main steps followed in this research . . . . .  | 36 |
| Figure 16. False colour composite using bands 2, 3 and 4 of LANDSAT TM representing the study area . . . . .                                | 42 |
| Figure 17. Map A representing forest cover types using untransformed bands of LANDSAT TM . . . . .  | 51 |
| Figure 18. Map B representing forest cover types using NDVI, bands 5 and 7 of LANDSAT TM . . . . .  | 52 |
| Figure 19. Map C representing forest cover types using VI, bands 5 and 7 of LANDSAT TM . . . . .  | 53 |
| Figure 20. Map D representing forest cover types using TVI, bands 5 and 7 of LANDSAT TM . . . . .   | 54 |
| Figure 21. Map E representing forest cover types using PC1, PC2 and band 5 of LANDSAT TM . . . . .  | 55 |
| Figure 22. Map F representing forest cover types using derived bands of Tasselled Cap Transformation of LANDSAT TM . . . . .                | 56 |
| Figure 23. Distribution of the reliability and accuracy percentages among the classes using bands 2, 3 and 4 of LANDSAT TM . . .            | 61 |
| Figure 24. Distribution of the reliability and accuracy percentages among the classes of classification using NDVI, bands 5 and 7 . . . . . | 63 |

|   |    |
|---|----|
| Figure 25. Distribution of reliability and accuracy percentages among the<br>classes of the classification using VI, bands 5 and 7 from<br>LANDSAT TM . . . . .     | 65 |
| Figure 26. Distribution of reliability and accuracy percentages of the<br>classification for TVI, bands 5 and 7 from LANDSAT TM . . . . .                           | 67 |
| Figure 27. Distribution of the reliability and accuracy percentages among the<br>classes of the classification for PC1, PC2 and band 5 from<br>LANDSAT TM . . . . . | 69 |
| Figure 28. Distribution of reliability and accuracy percentages among the<br>classes of classification using Tassel'd Cap Functions from<br>LANDSAT TM . . . . .    | 71 |

## **CHAPTER 1: INTRODUCTION**

---

In recent years, a limited number of investigations have indicated that sensors on board of remote sensing satellites may have potential use for land cover mapping of large areas. The outstanding characteristics of data from such satellites relates to their high temporal resolution, imagery being available for the whole globe on a daily basis.

Several studies have established the utility of digital image data such as those from LANDSAT Thematic Mapper (TM), Multispectral Scanner (MSS), SPOT MSS and NOAA AVHRR for automated classification of forest, although considerable studies have been conducted, in detail, with TM data.

In this research, LANDSAT TM data, as one of the best type image data due to it's high spectral and spatial resolutions, is carefully analyzed for certain type of forest land cover classification.

The present chapter has all the necessary background information and a general overview of the area of study. Apart of that, this chapter also explains the objectives of the thesis research.

### **1.1. Justification of the Research**

Forests cover large areas of the global land surface. For many developing countries, it represents an important income source for their economies. Due to over exploitation, forests are currently under constant risk.

The protection of forests from disasters (e.g. fire, disease, erosion) over extensive area is difficult without having any information such as condition, area, species, age classes and volume. With these informations, it is possible to make a proper management

of the forest by identifying and selecting the appropriate area for different management purposes, such as, harvesting, protection, etc.

Having all these informations collected, there is a need to store the referred information properly, for better and comprehensive use. For this reason, forest management maps play an important role in organizing gathered information for further strategies and policies determination in order to make the best use of forest.

Different approaches for mapping, like land survey, aerial photographs and satellite imagery can be used depending on the level of detail required and the extension of the area under study. For large areas, satellite imagery has been shown effective for forest classification and consequently mapping.

It is recognized that different satellite imagery can give different results in terms of information extraction. These different results relate to differences in spatial and spectral resolution. LANDSAT TM provides at present the best combination of spatial (30 m x 30 m) and spectral (seven bands, from visible to thermal infrared spectrum) resolutions. For this reason, LANDSAT TM is widely used by forest management agencies. However, most of these agencies apply only one standard image processing technique for forest mapping or, in another words, for forest classification. The referred technique sometimes provides unexpected results.

## **1.2. Objectives of the Research**

Knowing that the multispectral nature of most of remote sensing image data tends to be responsible for spectral transformations (linear transformations) that generate new sets of image components, image transformation can be one of the methods used to increase information by combining correlated bands and diminishing repeated data in selected bands of each satellite image.

Hence, some different image processing techniques will be evaluated in order to enhance the needed information especially the one concerning to forest cover types differentiation.

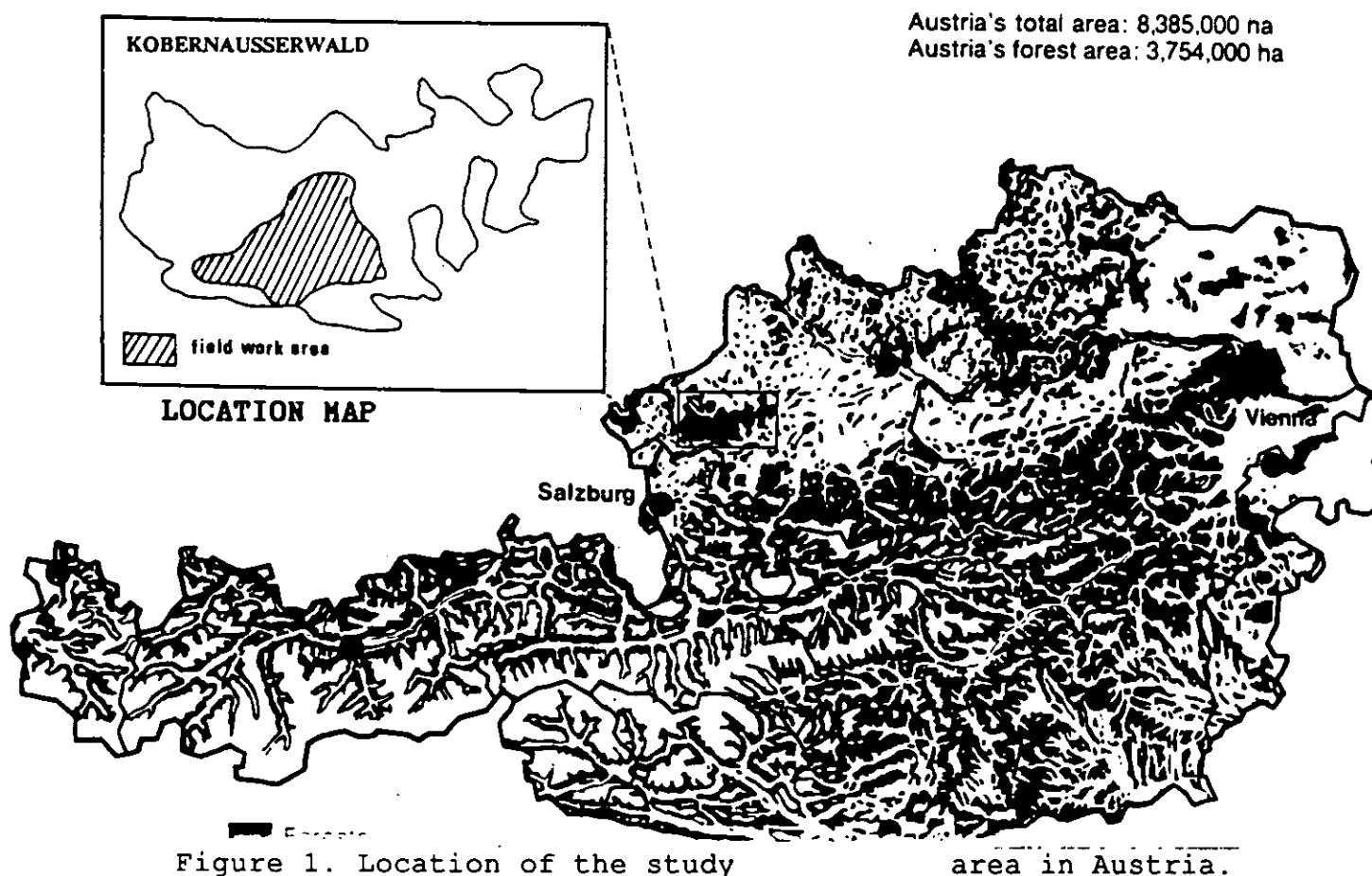
So, the objective of the present research by using two statistical methods, to assess which image processing techniques yields the best result for visual interpretation and multispectral land cover classification by:

- 1) Comparing results between various transformed and untransformed LANDSAT TM images for forest tree species and cover types differentiation purposes.
- 2) Reaching a ranking of goodness for the above comparison according to the statistical analysis.
- 3) Combining all transformed images in one image data set for image classification purpose.

### **1.3. Research Questions**

- 1 - How suitable is LANDSAT TM for forest cover differentiation ?
- 2 - Which image enhancement and transformation are the most suitable for forest classification ?
- 3 - Which statistical method is the most suitable for determination of the accuracy ?

### 1.4. Description of the area of study



The test site for this study, located in Kobernausserwald which involves two forest district (Matighofen and Friedburg), has an area of approximately 10,500 hectares which is located 40 Km north - east of Salzburg, Austria. Geographically the area is situated between  $48^{\circ} 1'$  and  $48^{\circ} 10'$  northern latitude and  $13^{\circ} 10'$  and  $13^{\circ} 27'$  eastern longitude (Figure 1).



The Kobernausserwald area covers about 15,000 hectares of forest, of which 10,500 are the property of the Austrian Federal Forest Service (Bundesforste). The remainder is under private ownership or it is community forest land.

#### **1.4.1. Physiography**

The area of study lies on a hilly and undulating terrain at a range of elevation 400 - 750 meters above sea level, with the lowest point at 430 meters in the northwest and highest point at 767 meters in the northeast. The forest has main drainage oriented westward in the western part and southward in the central and eastern side of the forest. Slopes are generally not steeper than 25%, although occasionally 80% has been measured, especially along the west facing slopes.

Geologically, the substratum belongs to the tertiary period with carbon - bearing lacustrine deposits of sand and clay. During the Pliocene period deposition of gravel and silicates took place by rivers from the central alpine region. These deposits reach thickness of 200 meters in the west part of the area, tapering off to 30 meters in the east. The gravel of the study area is composed of grey quartz and crystalline rocks mixed with sand and occasionally bound with calcite into conglomerate. The combination of relief and geologic conditions influence the drainage system.

#### **1.4.2. Climate, soil, vegetation**

The test site is situated in the oceanic climatic zone, rich in precipitation varying from 1000 to 1500 mm, with an average annual temperature of 7°C. Readings over 25 years (1926 - 1950) shows a mean annual precipitation of 1150 millimetres, which

672 are concentrated in the months of May to September with a peak in July (Figure 2). For the same period, the mean annual temperature is  $7.6^{\circ}\text{C}$ , where January is the coolest month (mean daily temperature is  $-2.7^{\circ}\text{C}$ ) and the hottest month is July with an average of  $16.6^{\circ}\text{C}$  daily temperature (Figure 3).

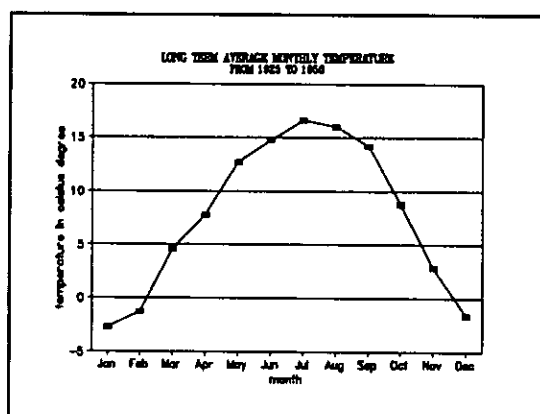


Figure 3. Temperature over 25 years (Roswita, 1975).

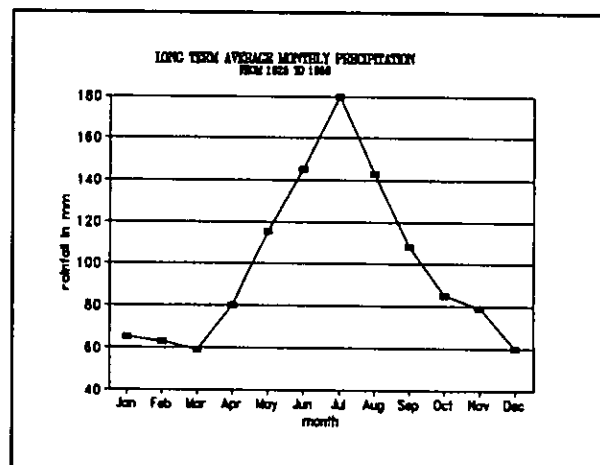


Figure 2. Precipitation over 25 years (Roswita, 1975)

Night frost normally occurs in September to mid June. In the winter, snow cover is usually high with layers between 40 to 100 cm thick.

The low temperature, together with high precipitation, causes podzolization to take place in most soils in the forest. The soils are highly acidic (PH 3 - 4) with an unsaturated humus complex because of the low humus content which has been filtered down from the upper horizons to the lower depth. The vegetation characteristics is a function of soils. In some parts the trees develop shallow and laterally spreading root systems due the shallow subsoil water table. In flat terrain where the gley - podzol are covered with humus, the main indicator is the blue berry (*Vaccinium spp.*), with norway spruce (*Picea abies*) as predominant tree species. On sand podzol, the grass species *Deschampsia spp.* is found together with *Vaccinium spp.* and *Lycopodium spp.*. In wet areas *Sphagnum spp.* also occurs under the spruce, mixed with silver fir (*Abies alba*) and beech (*Fagus sylvatica*).

### **1.4.3. History and administration**

According to historical documents dated 1007, the entire forest of the Kobernausserwald was donated by the Emperor Heinrich II to Monastery of Bamberg. In 1377 it was transferred to the Kuchler family i.e. the Ministry of the Archbishop of Salzburg. The area was sold to the Duke of Bavaria in 1439. It was under the ownership of the Duke of Bavaria when the first detailed descriptions of the Kobernausserwald were documented and a forest act (Forstordnung Von Ludwig des Reichen) formulated in 1468 to control careless exploitation.

In 1779, after the Treaty of Teschen, the Kobernausserwald was transferred to Austrian ownership. Proper scientific forest management started during this period in Kobernausserwald forest with a rotation of 130 years for norway spruce and beech trees.

The Imperial Forest law of 1852 further improved the forest act and regulations of the middle ages. Added guiding rules were made to protect and conserve environmental effects of forests by prohibiting the clearing of forests, making reforestation obligatory and imposing regulations on "Protection Forest" and "Protection Forest by Decree".

After World War I, between 1939 and 1945, Kobernausserwald was under German control, followed by the Trust Management of the Austrian State Forest Service. Through the Austrian State Treaty of 1955 the whole forest became the property of the Republic of Austria.

Due to the pressing needs and growing awareness on sustainable management of forest and environmental problems, Forest Law 1975 was made and came into force in January 1976. This Forest Law regulates the entire field of forestry defining utilization, protection, environmental and recreational aspects.

Kobernausserwald was originally divided into 3 forest districts of Schneegattern, Friedburg and Mattighofen, but later combined in 1980 with headquarters in Friedburg.

In 1988, reorganization of administrative area was again made and the property was redivided into two forest districts, Friedburg and Mattighofen.

These two districts, until now are headed by two professionally trained District Forest Officers, each performing a separate function.

#### **1.4.4. Silviculture and Forest Management**

Kobernausserwald is a commercial forest, although originally the main function was for the maintenance of game. The rotation period is 100 to 120 years, with a revision of the management plan every 10 years. Thinning operations are done in a five years cutting cycle and often as strip felling to allow access for timber extraction, and also to break the stand compositions into smaller units for easy and proper management.

During good seed years, matured stands are partially cleared to allow natural regeneration with a few mother tress left standing to provide seeds and to protect young regeneration trees. In areas where mother trees are not available, enrichment planting is introduced. Natural regeneration is found to be better in the valleys than at high elevations and largely influenced by the intensity of thinning. The existing standing volume of the forest is approximately 3.3 million m<sup>3</sup>, mainly norway spruce.

#### 1.4.4.1. Friedburg Forest District

Friedburg Forest District, which is a part of the Kobernausserwald and also the main area, has a total area of 5781.2 hectares of forest. This area has the following tree composition:

|   |       |
|---|-------|
| . Norway spruce ( <i>Picea abies</i> )  | 76.9% |
| . Beech ( <i>Fagus sylvatica</i> )      | 14.7% |
| . Silver fir ( <i>Abies alba</i> )      | 5.4%  |
| . Scot pine ( <i>Pinus silvestris</i> ) | 1.1%  |
| . Other species (mainly coniferous)     | 1.9%  |

The dominance of the spruce is due to the heavy cutting of the beech in the past and also due the present commercial reason.

The mean volume of the forest is about 320 m<sup>3</sup> per hectare with an annual growth increment of 8 to 9 m<sup>3</sup>. The Federal Forest Service cut about 45,000 m<sup>3</sup> annually out of which 11,000 m<sup>3</sup> harvested from thinning. Timber felling are carried out all year round. Nowadays most of the timber extraction is done by tractors to the near forest road, although horses are occasionally used to reduce damage on standing trees. Timber extraction are served by trucks using the well maintained road network within the forest.

This forest district has a high road density of about 32 meters per hectare. All weather roads criss-crossing the entire forest, making it possible to use highly mechanised logging, even in adverse weather conditions. Roads were constructed to make the area more accessible for proper silvicultural and forest management, but sometime has caused some damage to the forest.

## CHAPTER 2: THEORY AND CONCEPTS

---

This chapter describes an overview of the theory of satellite remote sensing applied to forestry.

### 2.1. General

Renewable and non-renewable natural resources such as forests, soils, water, minerals, etc., are very valuable environmentally and economically. These resources usually occupy very large area.

To study land and earth resources over large areas, it is desirable to identify and choose the most efficient tool in terms of time consuming, cost and accuracy for information gathering.

The primary interest in study forest, in particular, made remote sensing an important and useful tool for mapping forests, including studies of timber volume, insect infestation, and site quality, etc. In addition, it can also be attributed for detecting and monitoring changes in major ecological units, known as important source of genetic diversity. Human activities are rapidly destroying large areas of tropical forest by shifting cultivation or hunting, and only by means of remote sensing it is possible to understand the nature and the locations of these changes.

To consistently and repeatedly monitor forest cover types over large areas, it is desirable to use remote sensing data and automated image analysis techniques.

Several types of remote sensing data, including aerial photography, Multispectral Scanner (MSS), Simulated Thematic Mapper (TMS), Thematic Mapper (TM), Laser,

and Radar data have been used by forest agencies to detect, identify, classify, and evaluate various forest cover types and their changes over the area under study.

In this paper, only LANDSAT TM will be under consideration because of its magnitude and acceptance by the foresters till recently as the best satellite data available in the world because of certain characteristics described below.

## **2.2. Reflectance of the Main Ground Features**

Since remote sensing involves measurements of reflected, emitted or backscattered electromagnetic energy from the earth's surface, it is indispensable to understand the reflection energy along the wavelengths of the spectrum of distinct features on the ground caused by different and particular properties of each of them.

Experience, has shown that many earth surface features of interest can be identified, mapped and studied on the basis of their spectral characteristics. Thus, to utilize remote sensing data effectively, one must know and the spectral characteristics of the particular features under investigation in any given application. Likewise, one must know which factors influence these characteristics.

### **2.2.1. Vegetation**

The reflectance characteristics of vegetation depends on leaf properties including the orientation and the structure of the canopy leaves.



Sometimes vegetation in visible portion of the electromagnetic spectrum (EM) of may induce to mix up with water. But going further in EM spectrum, the accentuation is more clear specially in near infrared as it can be seen in Figure 4.

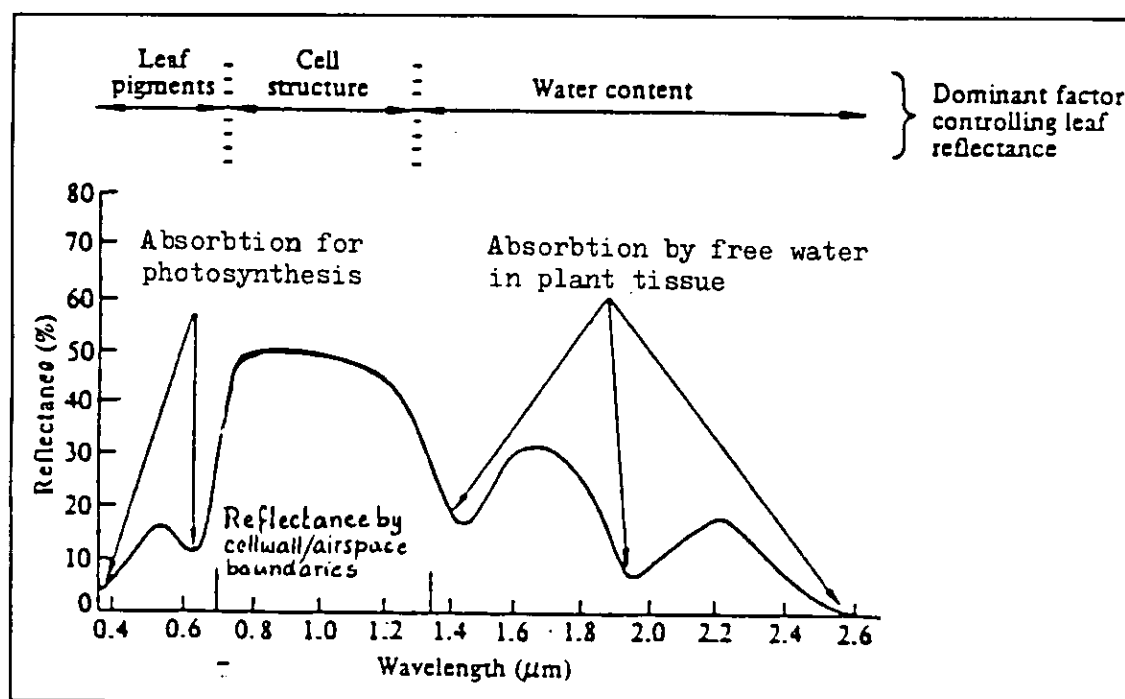
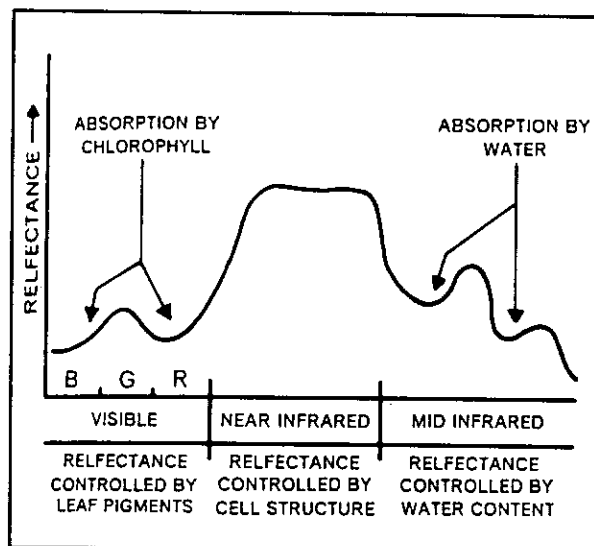


Figure 4. An idealized reflectance curve from a healthy vegetation (Kannegieter, 1987)

In figure 4 the ideal reflectance curve of the healthy vegetation is shown along the wavelength of the spectrum. In the visible part ( $0.45 \mu\text{m} - 0.69 \mu\text{m}$ ), the reflectance of the blue and red light is slightly less than green light since they are absorbed by the plant, mainly because of the presence of chlorophyll, for photosynthesis process. Vegetation reflects more in the green part of the electromagnetic spectrum. In the infrared region, the highest reflectance is concentrated in the near infrared, because of the cell structure which is proportional to the leaf development. In the middle infrared, the amount of free water in the leaf tissue determines the reflectance. More free water implies less reflectance.

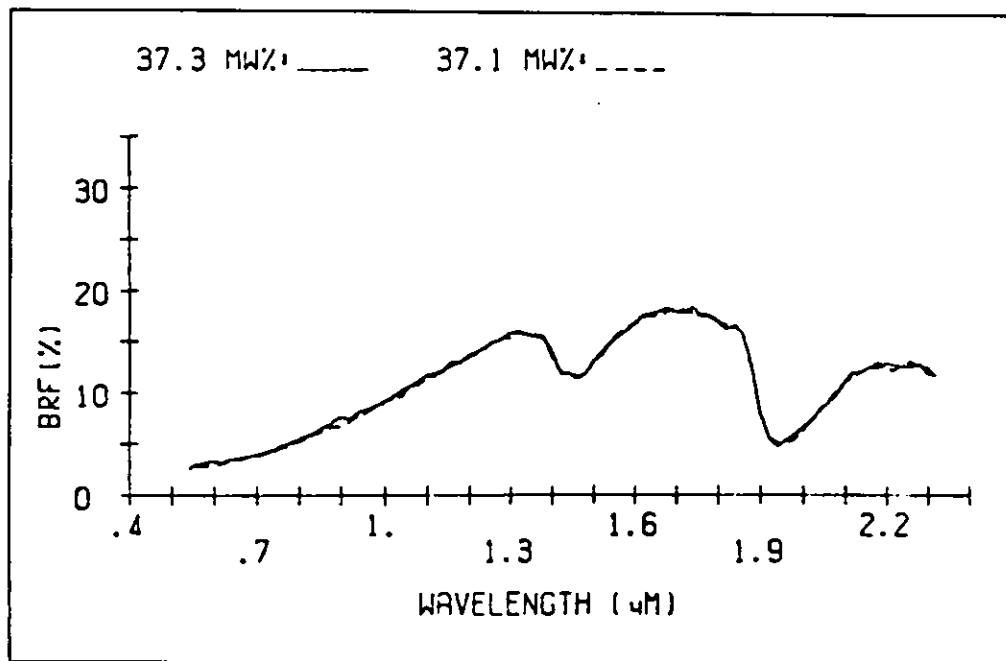
The blue and red portions of the spectrum are chosen to coincide with chlorophyll-absorption bands and the green portion of the spectrum with the green reflection peak. Whereas spectral response in the visible portions is controlled primarily by plant pigments, the response in the near-IR (infrared) is controlled mainly by the physical structure of the mesophyll layer of the leaves and in middle infrared as there are two strong water absorption bands, the reflectance is controlled by water content in that region of the spectrum (Figure 5).



**Figure 5.** Major influences on spectral properties of the living leaf (after Campbell, 1987).

### 2.2.3. Soils

Different soils, have different reflectance curve. The reflectance from the soil depends on several factors. The factors which contribute to variation of soil reflectance among others are: soil colour, particle size distribution, organic matter content, soil moisture content, presence of carbonate, iron oxide content, surface crust and soil salinity.



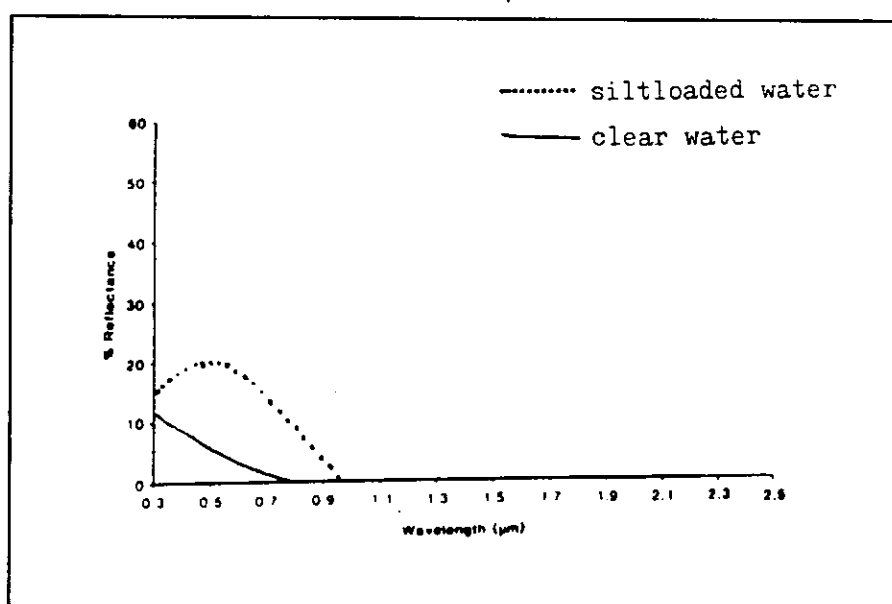
**Figure 6.** Soil reflectance from a fine silty soil (after Stoner and Baumgardner, 1980).

In Figure 6, the reflectance is increasing linearly from green band (0.5 - 0.6  $\mu\text{m}$ ) through red band (0.6 - 0.7  $\mu\text{m}$ ). In the middle infrared part of the spectrum, there are two strong water absorption bands.

#### 2.2.4. Water

Generally speaking, clear and deep water looks like blue to our eyes. So, if the water is too clear, the blue portion of the spectrum gives high reflectance, decreasing subsequently in the green band followed by red band. In the near infrared and middle infrared regions, the radiation is completely absorbed by the water, making the reflectance curve concave (Figure 7).

When the water is turbid, the green band has the high reflectance, and in blue and red bands, the reflectance decrease, making the curve convex (Figure 7).



**Figure 7.** Spectral reflectance curves of clear and turbid water.

### 2.3. Satellite Imagery

The reflected or emitted electromagnetic radiation can be detected electronically. Electronic sensors generate an electrical signal that corresponds to the reflected energy variations of the ground scene and records it as a digital numbers.

Different sensors are available for different purposes like, LANDSAT Thematic Mapper (TM) and Multispectral Scanner (MSS), SPOT MSS, NOAA AVHRR among the others. The latest generation of LANDSAT 5, is carrying two sensors on board, MSS and TM with different spectral and spatial resolutions.

The TM sensor carried on LANDSAT 4 and 5, is an improvement of the MSS sensor in terms of spatial (30 m x 30 m for all TM bands except band 6 which has 120 m x 120 m and 79 m x 79 m for MSS) and spectral (7 bands - blue, green, red, NIR, MI, TIR, FIR - for TM and 4 bands - green, red, NIR, NIR - for MSS) resolutions.

Although the design is based upon the same principles as the MSS system, the improvements are more complex in a way to give better geometric fidelity, greater radiometric detail and more spectral information.

The seven spectral regions performed by TM sensor are given on table 1.

Table 1 - The seven spectral bands of LANDSAT TM

| <b>BAND</b> | <b>REGION</b> | <b>WAVELENGTH<br/>(<math>\mu\text{m}</math>)</b> | <b>MAJOR ADVANTAGES,<br/>APPLICATIONS</b> |
|-------------|---------------|--|---|
| 1           | blue          | 0.45-0.52  | separation of soil and vegetation         |
| 2           | green         | 0.52-0.60  | reflection from vegetation                |
| 3           | red           | 0.63-0.69  | chlorophyll absorption                    |
| 4           | NIR           | 0.76-0.90  | delineation of water bodies               |
| 5           | MIR           | 1.55-1.75  | vegetative moisture                       |
| 6           | TIR           | 10.4-12.5  | hydrothermal mapping                      |
| 7           | MIR           | 2.08-2.35  | plant heat stress                         |

LANDSAT has a sun synchronous orbit with an inclination of  $99.09^\circ$  and positioned in an altitude of 919 Km for LANDSAT 1, 2 and 3 and LANDSAT 4 and 5 are in an altitude of 750 Km. One complete circle of the globe is done in 103 minutes and passes through the equator at 9.30 hrs. (local time) with a repeat cycle of 18 days.

Digital data for satellite images has contributed to the growth of image processing, pattern recognition and image analysis for land cover purposes.

## **2.4. Linear Transformations**

Most of the remote sensing data have repeated data over the contiguous bands. This is due to the existence of positive correlation among the visible bands and negative correlation between the near infrared and visible red bands. For this reason the satellite imagery data is responsible for spectral transformations that generate new sets of image components or bands.

Image transformation is used to gain more information by combining correlated bands and reduce duplicated data over all bands in each satellite image.

Image processing techniques are evaluated in order to extract the needed information especially the one concerning to vegetation differentiation information for foresters.

There are three major types of linear transformation such as band ratios, principal components analysis, tasselled cap transformation.

### **2.4.1. Rationing**

The infrared spectrum (from near to middle) is the important part of the electromagnetic spectrum because of the high response in reflectance due to the characteristics and the condition of the vegetation. It is used mainly for monitoring biomass vegetation. The ratios involve the red and infrared bands of a satellite image data set.

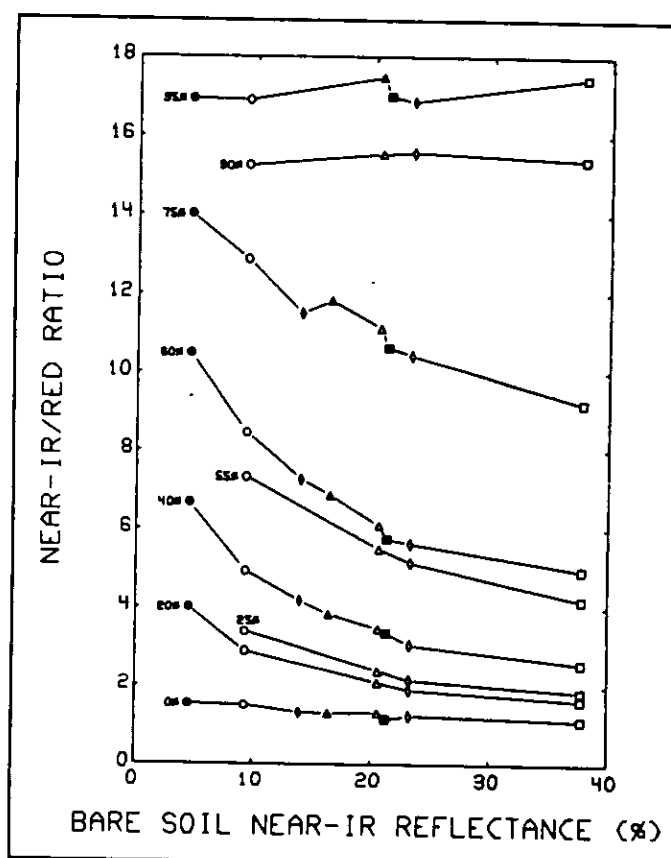
The use of ratios for estimating biomass or leaf area index was first reported by Jordan (1969) who used a radiance ratio  $0.8/0.675$  (infrared over the red part of the spectrum) to derive the leaf area index for forest canopies in tropical rain forest. The

followed research was conducted by Pearson and Miller (1972), who found a hand held spectral radiometer for estimating grass canopy biomass. The vegetation indices are also performed to minimize the variability due the external factors. They depend on canopy geometry, leaf and soil optical properties to provide good estimation of leaf area index (LAI).

Most of the ratios, started to be developed on LANDSAT MSS and on NOAA AVHRR data, because of the availability of data on that time.

#### **2.4.1.1. Ratio Vegetation Index**

The Ratio Vegetation Index, was initially settled for quantify and enhance the green vegetation on NOAA AVHRR data to avoid to work with all data set.



**Figure 8.** Relationship between RVI and bare soil NIR reflectance for various vegetation density levels (Huete et al., 1985).

Pearson and

Miller (1972)

defined Ratio Vegetation Index (RVI) as simple division of the red band over near infrared band ( $RVI = R/NIR$ ) to enhance the contrast between soil and vegetation by minimizing the effects of illumination conditions. However, it is sensitive to optical properties of soil background (Baret and Guyot, 1990). For a certain amount of vegetation, darker soil substrate result in higher vegetation index values (Elvidge and Lyon, 1985; Huete et al., 1985). A plot of the RVI ratio as a function of soil background reflectance for various vegetation level is shown in Figure 8.

The differences in NIR/R ratios, caused by various soil backgrounds, increased in magnitude with increasing vegetation amount to 60% green cover. Thus, the



influence of soil background in this greenness measure was greater at higher vegetation densities (50 - 60% cover).

Carnegie et al. (1974) used a ratio of LANDSAT MSS (band 7/band 5) and found that the ratio curves, plotted as function of time, peaked during the period of greatest forage production.

#### **2.4.1.2. Normalized Difference Vegetation Index**

The Normalized Difference Vegetation Index (NDVI) is similar to RVI, because NDVI was derived from it. When Rouse et al. (1973, 1974) analyzed the LANDSAT MSS data, they found that although the RVI could be used for a measurement of relative greenness, location and cycle deviations over the area would introduce a large error component. To compensate the errors, they reached to a new formula so called NDVI (Normalized Difference Vegetation Index), which is:

$$\text{NDVI} = \frac{\text{IR} - \text{R}}{\text{IR} + \text{R}}$$

where,

IR = infrared band

R = red band

Rouse et al. (1974) in their research, found that NDVI varies from a minimum at bare soil reflectance to a maximum for a fully developed canopy with a value somewhat less than one. According to Campbell (1987), NDVI is tailored to produce desirable statistical properties in the resulting values.

Also NDVI as RVI, is affected by external factors like atmospheric hazards resulting false values.

The utility of this ratio, relates to the strong absorption by chlorophyll in the red part of the spectrum and strong reflection by the mesophyll layer of healthy vegetation in the near infrared. Hence, the NDVI is frequently found to be strongly related with vegetation parameters, such as leaf biomass and leaf area index (Tucker 1979, Curran 1980).

#### 2.4.1.3. Transformed Vegetation Index

To avoid working with negative ratio values and the possibility that the variances of the ratio would be proportional to the mean values, Rouse et al. (1973, 1974) added a constant of 0.5 and a square root was applied to NDVI.

The formula for Transformed Vegetation Index (TVI) is:

$$TVI = \frac{\sqrt{IR - R}}{\sqrt{IR + R}} + 0.5$$

where, IR = infrared band

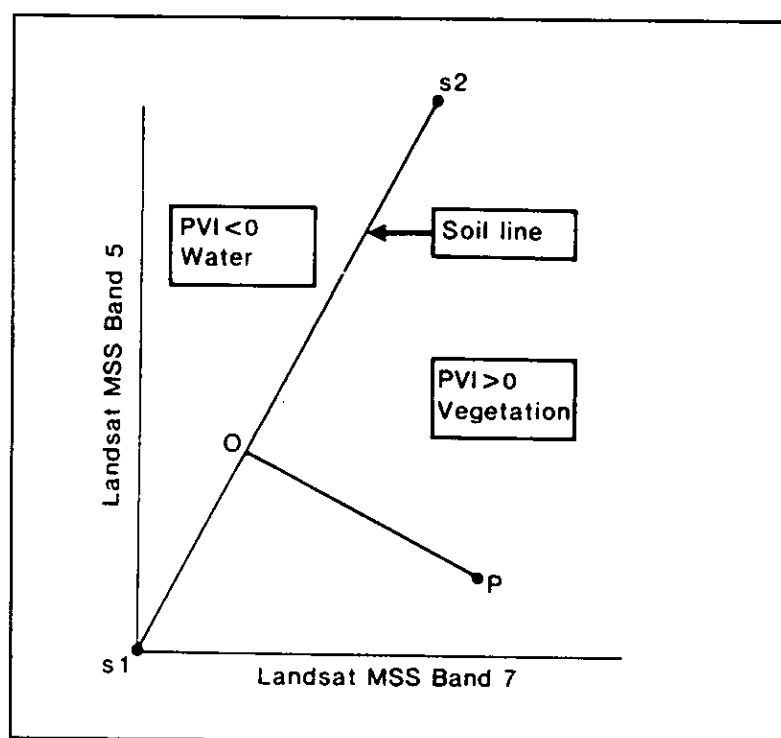
R = red band

In the same research, Rouse et al. (1974), found a close relationship between TVI and green biomass on LANDSAT MSS.

### 2.4.2. Perpendicular Vegetation Index

Richard and Weigand (1977) approached the problem of variable soil brightness by developing perpendicular vegetation index (PVI). It is defined as the perpendicular distance of a point represented by the red and near infrared reflectance from the bare soil line which lies along the line S1-S2 (Figure 9), and can be expressed as:

$$PVI = \sqrt{[(R_{soil} - R_{veg})^2 + (IR_{soil} - IR_{veg})^2]}$$



**Figure 9.** The concept of the Perpendicular Vegetation Index, which is the perpendicular distance from point P to the soil line S1-S2. (Campbell, 1987).

If the canopy cover is high, the PVI will underestimate Leaf Area Index (LAI) for wet soils.

The PVI error increases as LAI increases because at high LAI the soil brightness effects on reflectance are almost perpendicular to the PVI (Baret and Guyot, 1990).

### **2.4.3. Tasselled Cap Transformation**

Kauth and Thomas (1976) used the same idea of PVI to develop the called Tasselled Cap Transformation (TCT) for LANDSAT MSS, but in their model they used all four MSS bands. The TCT is intended to define a new coordinate system in terms of which the soil line and the region of vegetation are more clearly represented. The derived axes are called brightness, greenness, yellowness and nonesuch.

The brightness is correlated with variations in the soil reflectance. The greenness is related with variations in the vigour of green vegetation. The yellowness is associated with the variations in the yellowing of senescent vegetation. The fourth one, nonesuch, as it is being related with atmospheric conditions, it is not widely used in vegetation studies.

Crist (1983) and Crist and Cicone (1984 a,b) extend the study of TCT to data from LANDSAT TM (excluding the thermal band). The TM were found to contain significant information in a third dimension which was defined as wetness (yellowness in MSS). The third TM Tasselled Cap is contrast between the sum of the two longer infrared TM bands (5 & 7) and the sum of the visible and near infrared bands (1 & 4).

The coefficients of the three TM Tasselled Cap functions shown on Table 2 express the the amount of correlation for each band to identify one of the functions (brightness, greenness and wetness).

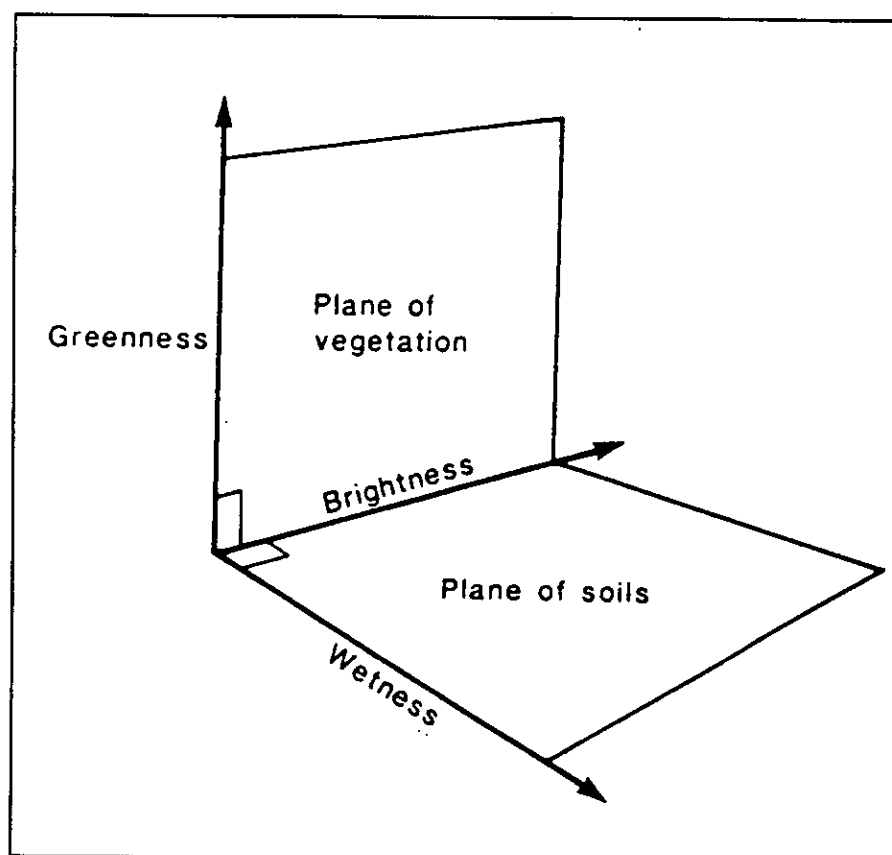
Table 2 - Coefficients for LANDSAT TM (Crist and Cicone, 1984)

| TM band: | 1       | 2       | 3       | 4      | 5       | 7       |
|----------|---------|---------|---------|--------|---------|---------|
| Bright.: | 0.3037  | 0.2793  | 0.4343  | 0.5585 | 0.5082  | 0.1863  |
| Green.:  | -0.2848 | -0.2435 | -0.5436 | 0.7243 | 0.0840  | -0.1800 |
| Wetn.:   | 0.1509  | 0.1793  | 0.3299  | 0.3406 | -0.7112 | -0.4572 |

The brightness function is a weighted average of the six bands, while greenness is a visible /near infrared contrast, with a little contributions from bands five and seven. Wetness is a contrast between the mid-infrared bands (5 and 7) and red/near infrared bands (3 and 4). These functions define a three dimensional space in which the positions of individual pixels can be computed using the above mentioned coefficients.

The plane defined by the greenness and brightness functions is termed by Crist and Cicone (1984c), the "plane of vegetation", while the functions brightness and wetness define the "plane of soils" (Figure 10).

Tasseled Cap Transformations deal with findings of the data structures inherent to a certain sensor and set of scene classes, and adjusting the viewing perspective such that these structures can be easily and completely observed (Crist & Kauth, 1986).



**Figure 10** – TM Tasseled Cap brightness, greenness and wetness functions define two planes – the plane of vegetation and the plane of soils (based on Crist and Cicone, 1984c).

The direct connection between the features and the actual scene characteristics improve and enhance the interpretation of the observed spectral variation and the prediction of the spectral effects of particular changes in the scene characteristics (Crist & Cicone, 1983).

Studies done by Kouth & Thomas (1976), Kauth et al. (1979), Crist & Cicone (1984c) and Crist et al. (1986), using the TCT technique, mentioned that to obtain reliable and better results, some corrections to the image data set is required, before running the referred technique. These corrections are necessary due to the effect of haze, clouds, cloud shadow and plant canopy shadow, which affect the reflectance

of the objects and confuse the analyst, especially when the feature under observation is mixed with one of the referred effects mentioned above (e.g. shadow of plants on a dark soil tone),

#### **2.4.4. Principal Component Analysis**

In general, adjacent bands in a multispectral remotely sensed image are correlated with each other. This correlation implies the existence of redundancy in the data, in another words, some information is repeated.

Multiband visible and near infrared images of vegetated areas show negative correlations between the near infrared and visible red bands and positive correlation among the visible bands. These correlations are because of the spectral characteristics of vegetation such as the vigour or greenness of the vegetation increases, the red reflectance diminishes and the near infrared reflectance increases. A principal component analysis transform of a multispectral image might therefore be expected to define the dimensionality of the data set and identify the principal axes of variability within the data.

A compression of the variability of the data made to reduce the number of bands in the image processing analysis for classification, and that will reduce the amount of time consuming needed for classification and reduce the space in the computer.

## **2.5. Satellite Image Classification**

Classification in general is grouping of the similar objects and separation of dissimilar ones.

Satellite image data have advantages over other data sources for large area land cover classification because of their frequent repeat cycles, large area sample, wide spectral range, and amenability to automated classification.

According to Skidmore and Turner, 1988, the major reason for forest land managers to adopt remotely sensed data for land cover classification, is the low accuracy obtained after automated classification (less than 80% of accuracy is unacceptable).

There are two approaches, supervised and unsupervised image classification which give different levels of accuracy depending on which one is chosen.

### **2.5.1. Unsupervised Classification**

The basic principle, of unsupervised classification is that values within a given cover type should be close together in the measurement space, whereas data in different classes should be well separated. The process is done by grouping homogeneous pixels (e.g. same reflectance) and previous information of the area is not needed. A general knowledge of the area is required in order to be able to interpret the meaning of the same classification.

The type of classification generated is spectral classes, and does not necessarily correspond to informational categories interested by the analyst. Thus, there is a need to link logically the spectral classes to informational categories.



In unsupervised classification the output might be numerous spectral classes which are not familiar to an analyst make it difficult to chain on all of them.

There are different clustering algorithms to determine the spectral classes in the data set. The two common ones are the so called "K-means" and the other is the use of algorithms that incorporate a sensitivity to image texture as a basis of establishing cluster centres.

The "K-means" locates the number of clusters in a multi- dimensional measurement space . The pixels are assigned to a cluster whose mean vector is the closest one. After all pixels have been classified, the revised mean vectors for each cluster are computed, and then used as a basis to reclassify the image data. In the other approach, texture is defined by the multidimensional variance observed in a moving window passed trough the image. The analyst sets a variance to the threshold below which a window is considered smooth (homogenous) and above which it is rough (heterogenous) along all the image to be analyzed. The mean of the first smooth becomes a centre of the first class, the mean of the first rough becomes the centre of the second class, and so forth. After defining the number of clusters, the classifier considers the distances between the defined cluster centres and merges the closest ones until the entire image is analyzed and separates the ones based on the defined analyst distance. Then, the final clustering is used to classify the image.

### **2.5.2. Supervised Classification**

In supervised classification, the interaction of the user is important. He bases on the knowledge of the area and other type of ancillary data (maps, aerial photographs, ground truth) to select training samples for different informational classes.

Three basic steps are involved in supervised classification. The first one, is training stage, where is done the identification of the representative training areas and development of a numerical description of the spectral classes for each land cover type. The second stage, classification stage, in which each pixel in the image data set is categorized for each informational class. If there is any pixel not similar to any training data set, the referred pixel is labelled as unknown. After the data set has been categorized, comes the third stage or the output stage. This can be in three ways, which are thematic maps, tables of full scene or subscene area statistics for the each different land cover (Figure 11).

The classification stage is the important one, because of the selection of the training fields. This selection has to be homogeneous to represent each category. Once, all training samples for different informational classes are defined, the data set is classified according to algorithm method. The most common ones are parallelepiped classifier, minimum distance to means classifier and gaussian maximum likelihood classifier.



Box classification, some times known as parallelepiped classification or level slice procedures, is based upon the ranges of values within the training data to define regions within multi- dimensional data space. The spectral values of unclassified pixels are projected into data space; those that fall within regions defined by the training data are assigned to the appropriate categories (Figure 12).

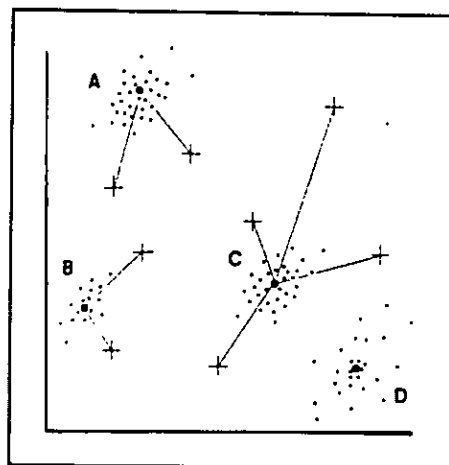
Difficulties can be find when there is an overlap between two informational classes. The pixels will be classified as unknown, or arbitrarily attributed to any of the categories.

### 2.5.2.2. Minimum Distance Classification

Minimum distance classification or the nearest neighbour classification is based on use of the central values of the spectral data that form the training data as a means of assigning pixels to informational classes (Figure 13).

The system computes the mean distance for each category. If it is closed to one mean category, the pixel will be assigned as belonging to this category. If the pixel is farther from any mean category, it would be classified as unknown.

This method is mathematically simple and computationally efficient, but some disadvantages are occurred in the process. The important one is the insensitivity to different degrees of variance in the spectral data.

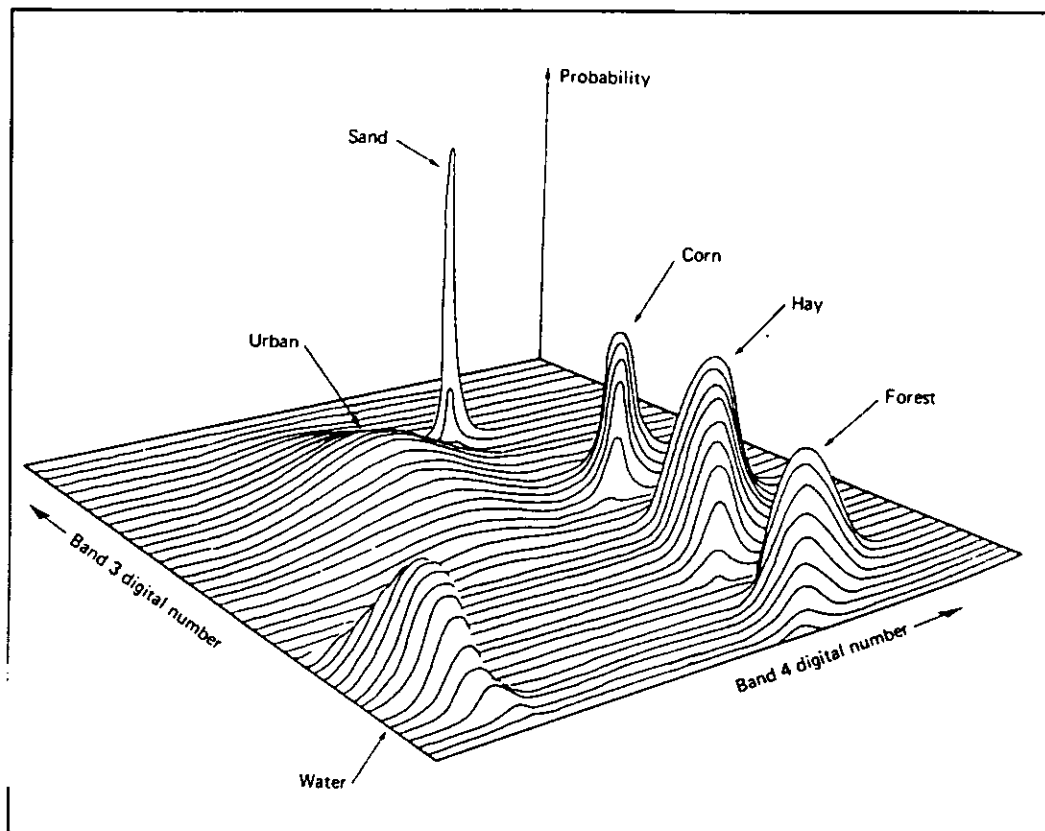


**Figure 13.** Minimum distance classifier. Here, the small dots represent pixels from training fields and the large ones the means for each class (Lillesand and Kiefer, 1987).

The classes D and B make a shape more ellipsoidal than the classes A and C. This is due the high variance in those classes (D and B). One pixel can be assigned as one class because it is close to that category, although the similarity with the neighbour category is more representative because of it's variance (Figure 13).

### **2.5.2.3. Maximum Likelihood Classification**

The maximum likelihood system computes the variance and covariance of the spectral class when is classifying an unknown pixel. There is an assumption of the pixels forming a determined category which is that they are normally distributed. Therefore, the distribution of a category response is described by the mean vector and the covariance matrix. With these parameters computed as statistical probability of a given pixel value can be assigned as being a member of a particular land cover class (Figure 14).



**Figure 14.** Probability density functions defined by a maximum likelihood classifier (after Lillesand and Kiefer, 1987).

This chapter describes the materials used in the research, and all the procedures to achieve the proposed objectives.

### 3.1. Data

The materials used to conduct this research can be divided into:

a) ancillary data

- . topographic map - scale 1:10,000
- . Forest management map - scale 1:10,000
- . orthophoto map - scale 1:30,000 - August 1984
- . aerial photographs - scale 1:15,000 - 13 May 1992
- . digital maps of the area

b) satellite imagery

- . LANDSAT TM (bands 1, 2, 3, 4, 5 & 7) - 12 May 1992

c) field work data

- . Ground truth data collected during 10 days of field work in June 1993. (see appendix Nº 1)

### 3.2. Methodology

The methodology is an important part of any research to conduct the work successfully and scientifically. In Figure 15, the flowchart summarizes all steps used in this research.

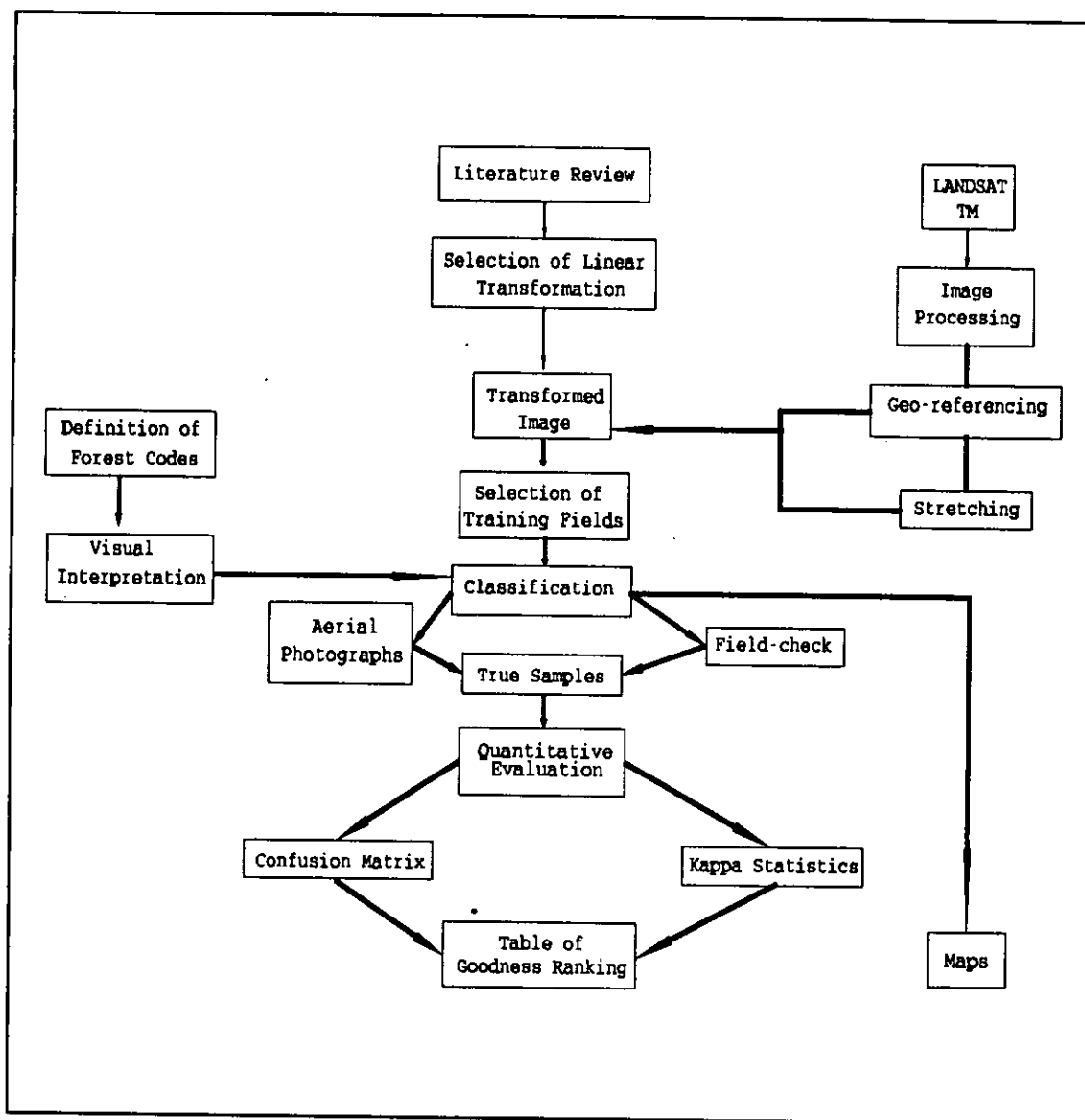


Figure 15. Schematic flowchart representing the main steps followed in the research.



### **3.2.1. Image Pre-processing**

To avoid future problems during the analysis and to get reliable results, the image data set had to be run through a few image pre-processing analysis.

#### **3.2.1.1. Geometrical corrections**

Raw digital images usually contain geometric distortions so significant, that they cannot be used as map. The factors that contribute for the distortions are ranging from altitude of the area which the image involves, velocity of the sensor platform, to a factors such as panoramic distortion, earth curvature, atmospheric refraction and relief displacement. The intention of geometric is to compensate the distortions introduce by these factors, so that the corrected image will have the geometric integrity of a map.

#### **Geo-referencing**

To enable overlaying of digitized maps (i.e. forest units, roads, towns, etc.) with the digital images, ground control points have to be found and be used to geo-reference the different bands of LANDSAT TM to a U.T.M coordinate system. The procedure of geo-referencing can be described in two step procedure:

1. Registering the respective coordinates (row and column) of the ground control points. These values are then submitted to a least square regression analysis. This determines a transformation function to superimpose the grid of one image, "Master Map", on the second one, "Slave Map".

2. Resampling the slave map to determine the new pixel value to be assigned in the new grid system. Three methods for the resampling procedure are available in ILWIS:

1. Nearest neighbour.
2. Bilinear interpolation.
3. Cubic convolution.

If the transformed image is to be classified, it is recommended to select the nearest neighbour resampling. In this case a total of 18 ground control points were found, after deleting the unfavourable ones.

### 3.2.1.2. Stretching

The sensitivity range of LANDSAT TM detectors was designed to record a wide range of terrain brightness from black basalt to white sea ice under a wide range of lighting conditions (Sabins, 1986). TM images seldom utilize the entire possible 0 - 255 range. Instead they tend to be compressed into a rather narrow range.

In order to utilize the entire available range, stretching is a common technique. The simplest stretching is linear stretching, in which the lowest pixel value in the image is assigned to 0 (black), and the highest pixel value is assigned 255 (white).

The remaining pixel values are distributed linearly between the extremes according to the following:

$$PV' = \frac{PV - MIN}{MAX - MIN} * 255$$

where:

PV' - new pixel value.

PV - original pixel value.

MIN - minimum value in the original image.

MAX - maximum value in the original image.

It is also a common practice to omit a certain percentage on both extremes, saturating them into 0 and 255, respectively.

### **3.2.2. Image Processing**

After all above described image pre-processing, a series of selected linear transformations were ran on LANDSAT TM, in order to obtain an improved data set of LANDSAT TM for classification purposes.

#### **3.2.2.1. Band rationing**

##### **A) Ratio Vegetation Index**

Ratio Vegetation index or Vegetation Index was obtained by:

$$VI = \frac{\text{Band 3}}{\text{Band 4}} * 127 + 128$$

The factor 127, was multiplied in order to spread the digital numbers of spectral signature from '0' to '255'.

### **B) Normalized Difference Vegetation Index**

The Normalized Difference Vegetation Index (NDVI) as one of the band rationing for enhance the vegetation was calculated by:

$$\text{NDVI} = \frac{\text{band 4} - \text{band 3}}{\text{band 4} + \text{band 3}} * 128 + 127$$

The factors also are multiplied and added to round up the digital numbers of the spectral signature.

### **C) Transformed Vegetation Index**

The formula for Transformed Vegetation Index (TVI) is:

$$\text{TVI} = \sqrt{\frac{\text{IR} - \text{R}}{\text{IR} + \text{R}}} + 0.5$$

#### **3.2.2.2. Tasseled Cap Transformation**

A set of three tasseled cap were done according to the following formulas derived from Table 2 (Chapter 2):

1. TC1 (brightness) =  $0.3037 * \text{band 1} + 0.2793 * \text{band 2} +$   
 $0.4343 * \text{band 3} + 0.5585 * \text{band 4} + 0.5082 * \text{band 5} +$   
 $0.1863 * \text{band 7}$
2. TC2 (greenness) =  $- 0.2848 * \text{band 1} - 0.2435 * \text{band 2} -$   
 $0.5436 * \text{band 3} + 0.7243 * \text{band 4} + 0.8400 * \text{band 5} -$   
 $0.1800 * \text{band 7}$
3. TC3 (wetness) =  $0.1509 * \text{band 1} + 0.1793 * \text{band 2} +$   
 $0.3299 * \text{band 3} + 0.3406 * \text{band 4} - 0.7112 * \text{band 5} -$   
 $0.4572 * \text{band 7}$

### 3.2.2.3. Principal Component Analysis

Principal Component Analysis (PCA) was done using bands 2, 3, 4, 5, and 7 (green, red, near infrared and two middle infrared), which are the bands that have significant response to vegetation.

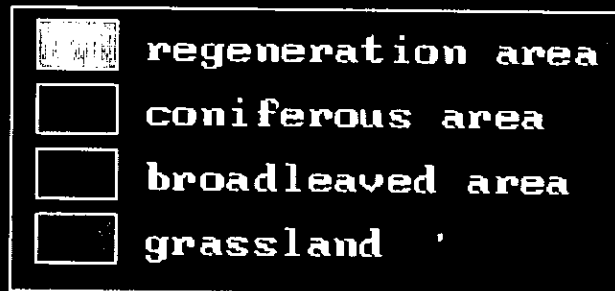
The eigen vectors were calculated through ILWIS.

### 3.2.2.4. Classification

One of the first steps before classifying the image is to enhance the image in a way to provide a rough information of forest land cover differentiation. The simplest way is by combining different bands like infrared, red and green for false colour composite, where the broadleaves species are enhanced from the coniferous with a bright red colour (Figure 16).

# KOBERNAUSSERWALD COVER TYPES

color composite using bands 2, 3 and 4 of LANDSAT TM



0 Km 30 Km

1 : 121485

**Figure 16.** False color composite using bands 2, 3 and 4 of LANDSAT TM representing the study area.

When the forest cover types are roughly differentiated, six different classifications of the area were performed taking in to account the list of classes shown in Table 3.

Table 3 - Forest codes for classification

| CODE | FOREST CLASSES  |
|------|---|
| A    | Cleared areas & regeneration, < 8 metres tall             |
| B    | Pole timber, > 8 metres tall, coniferous, < 41 years old  |
| C    | Pole timber, > 8 metres tall, broadleaves, < 41 years old |
| D    | Coniferous stand, 41 - 80 years old                       |
| E    | Broadleaves stand, 41 - 80 years old                      |
| F    | Coniferous stand, more than 80 years old                  |
| G    | Broadleaves stand more than 80 years old                  |
| H    | Mixed stand, more than 80 years old                       |
| I    | Grassland   |

The forest codes were obtained from forest management maps of the area.

The type of classification to be performed is the maximum likelihood supervised classification, where pixel sampling took place for homogeneous reflectance areas based on aerial photographs and forest management maps of the area (see Appendix 2).

The maximum likelihood classifier is chosen, because, it is the one which can obtain reliable results, since the box classifier would introduce problems when overlapping between two classes occurs while minimum distance classifier is insensitive to the variances in each classes.

The six performed classifications are:

1) Classification of untransformed bands

The set of bands used to perform the classification where 2, 3 and 4 (green, red and infrared, respectively), because these are the classical set of bands for common forest classification.

2) Classification of NDVI, bands 5 and 7

As NDVI uses TM bands 3 and 4, the other bands which have some more response to vegetation are the bands 5 and 7 were selected.

3) Classification of VI, bands 5 and 7.

4) Classification of TVI, bands 5 and 7.

5) Classification of PC1, PC2 and band 5.

The vegetation response in band five is higher than band seven. That is why the band five is selected in classification performance.

6) Classification of TC1, TC2 and TC3 (Tasselled Cap Transformation).



### **3.2.3. Evaluation of the Classification Results**

Based on available resources, a target sample size of approximately 72 points per land cover class was established. Five samples per class had field visits, and the others were based on aerial photographs of same date as the LANDSAT TM (13 May, 1992).

The field samples were established according to the accessibility to the samples to be visited and to verify the possible doubts occurred on the output of the classification.

#### **3.2.3.1. Quantitative Approach**

After the selection and identification of the samples, as true pixels for each land cover classification, it is necessary to quantify the accuracy of the referred classification. The approach should give statistical meaning for better comparison between classifications and to validate the same classifications.

The two statistic techniques described in this research, should be very helpful in comparing and assessing LANDSAT TM classification accuracy data that are in form of error matrices. The approaches allow the LANDSAT TM data user to quantitatively compare the different aspects of the image processing and determine which one perform better under varied conditions.

#### **A) Confusion matrix**

Confusion matrix, is usually used with remotely sensed data applications to find out the accuracy of the image classification. Operationally, the observed forest land cover would usually consist of a set of samples (e.g., test fields) randomly located within

each forest class over the area under study. This enables direct pixel by pixel comparison between the classified maps and the observed forest land cover (Hudson, 1987) and to estimate the classification accuracy.

In Table 4, the number of true pixel per class were achieved through field work (June 1993) and the aerial photographs of the same date of the LANDSAT TM image (13 May, 1992).

The number of true samples should be at least 30. It is believed in infinite references that sample of 30 minimum observations is enough to get a meaning through statistical analysis.

Table 4 - Number of true pixels per class collected to be verified over all area.

| CLASSES | A  | B  | C  | D   | E  | F  | G  | H  | I  |
|---------|----|----|----|-----|----|----|----|----|----|
| PIXELS  | 72 | 72 | 72 | 108 | 68 | 72 | 78 | 66 | 86 |

### **B) Kappa statistics**

The kappa statistics as being another method to assess the accuracy or agreement among the classes, considers all classes for the output statistics results.

Accuracy is expressed by the kappa statistics (Congalton and Mead, 1983), which represents the degree of agreement between the classified forest land cover and the observed forest land cover after chance agreement has been removed.

Congalton (1991) considers kappa statistics a powerful technique in its ability to provide information about a single matrix as well as to statistically compare matrices.

The Kappa factor is given by the formula (Siegel and Castellan Jr., 1988)

$$\text{Kappa (K)} = \frac{P_o - P_e}{1 - P_e},$$

where:

$P_o$  - is the proportion of correctly classified cases

$P_e$  - is the proportion of correctly classified cases  
expected by chance.

Once, Kappa is calculated for each matrix (derived from contingency table), it becomes a measure of how well the classification agrees with the ground truth. The approximate large sample variance of kappa  $\sigma^2$ , can be used to construct a hypothesis for significance between error matrices (Cohen, 1960). This test is possible because the large sample distribution of kappa is normal.

The variance is given by (Siegel and Castellan Jr., 1988):

$$\text{var(K)} = \frac{2}{Nk(k-1)} * \frac{P_e - (2k-3)[P_e]^2 + 2(k-2)\Sigma p_j}{[1 - P(E)]^2}$$

where:

$P_e$  - is the proportion of correctly classified cases  
expected by chance

$N$  - are the number of classes to be compared

$k$  - number of true classes

$p_j$  - proportions of each class

The  $P_o$  and  $P_e$  are calculated by (Siegel and Castellan Jr., 1988):

$$P_e = \sum P_j^2$$

$$P_o = \left( \frac{1}{Nk(k-1)} \sum \sum n_{ij}^2 \right) - \frac{1}{k-1}$$

Therefore, the appropriate analysis is 'Z' test (Siegel et al., 1988), where can be achieved by:

$$Z = \frac{K}{\text{var}(K)}$$

As mentioned above, this statistical test is normally distributed. This means that the values in a standard normal table at varying confidence levels can be used to determine if two matrices are significantly different at determined level of significance.

## CHAPTER 4: PRESENTATION OF THE RESULTS

---

The results of the research are shown in a sequence according to the classifications performed for each linear transformation. Each classification has one table, where the accuracy through contingency table is calculated, and also the agreement between the classes is calculated by kappa statistics.

### 4.1. General Remarks

Measurement of the spectral reflectance in visible spectrum showed no consistent separability among the different age classes for coniferous trees. However, within the high density area of same species, separation between age classes was more evident. For class C (pole timber, > 8 metres tall, broadleaves, < 41 years old), the spectral separation was more difficult because it was mixed with grassland areas.

The number of true pixels distributed per class were not proportional because the area of certain classes was either too big or too small. This hindered in obtaining the reliable results for the classes representing smaller areas.

### 4.2. Image Processing

#### A) Geo-referencing

After geo-referencing the sigma value obtained was 0.45 and sumperimposed with digital map (road network). The result was fairly good.

#### B) Principal Component Analysis

The eigen vectors given by ILWIS for LANDSAT TM, bands 2, 3, 4, 5 and 7 are shown on Table 5.

Table 5 - Variability among the new generated bands.

| New bands | Variability |
|-----------|-------------|
| PC1       | 85.19%      |
| PC2       | 13.77%      |
| PC3       | 0.59%       |
| PC4       | 0.37%       |
| PC5       | 0.07%       |

### 4.3. Classified Maps

Digital image classification is a process of assigning pixels to classes. The training fields (e.g., pixels) is treated as an individual unit composed of values in different bands and in different new originated bands derived by linear transformations. By comparing pixels to one another and to pixels of known identity, it was possible to assemble groups of similar pixels into particular forest land cover. These classes form regions on the map, so that after classification the digital image is presented as a mosaic of uniform parcels, each identified by colour.

## 4.3.1. Untransformed Bands

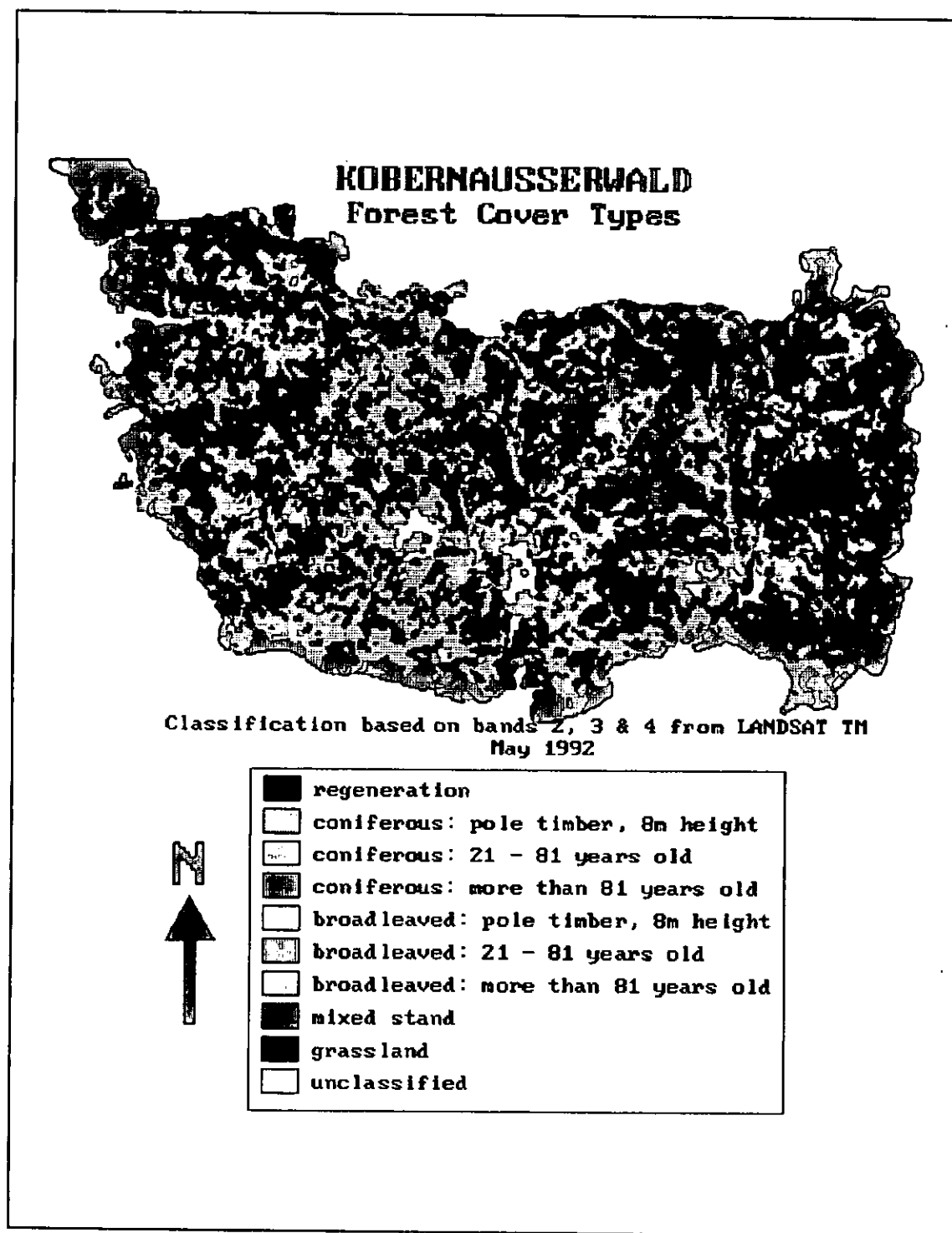


Figure 17. Map A representing forest cover types using untransformed bands of LANDSAT TM.

## 4.3.2. NDVI, Bands 5 and 7

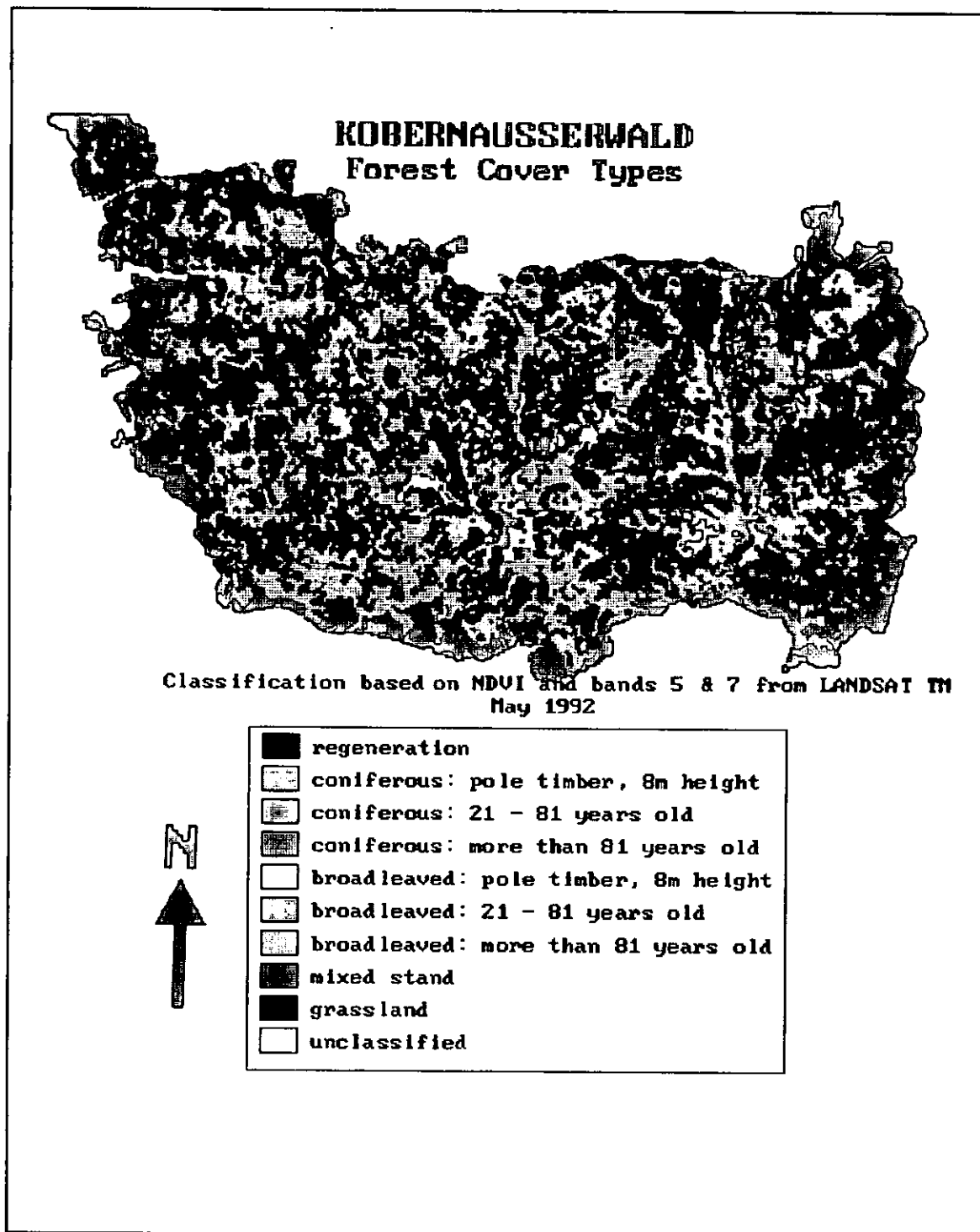
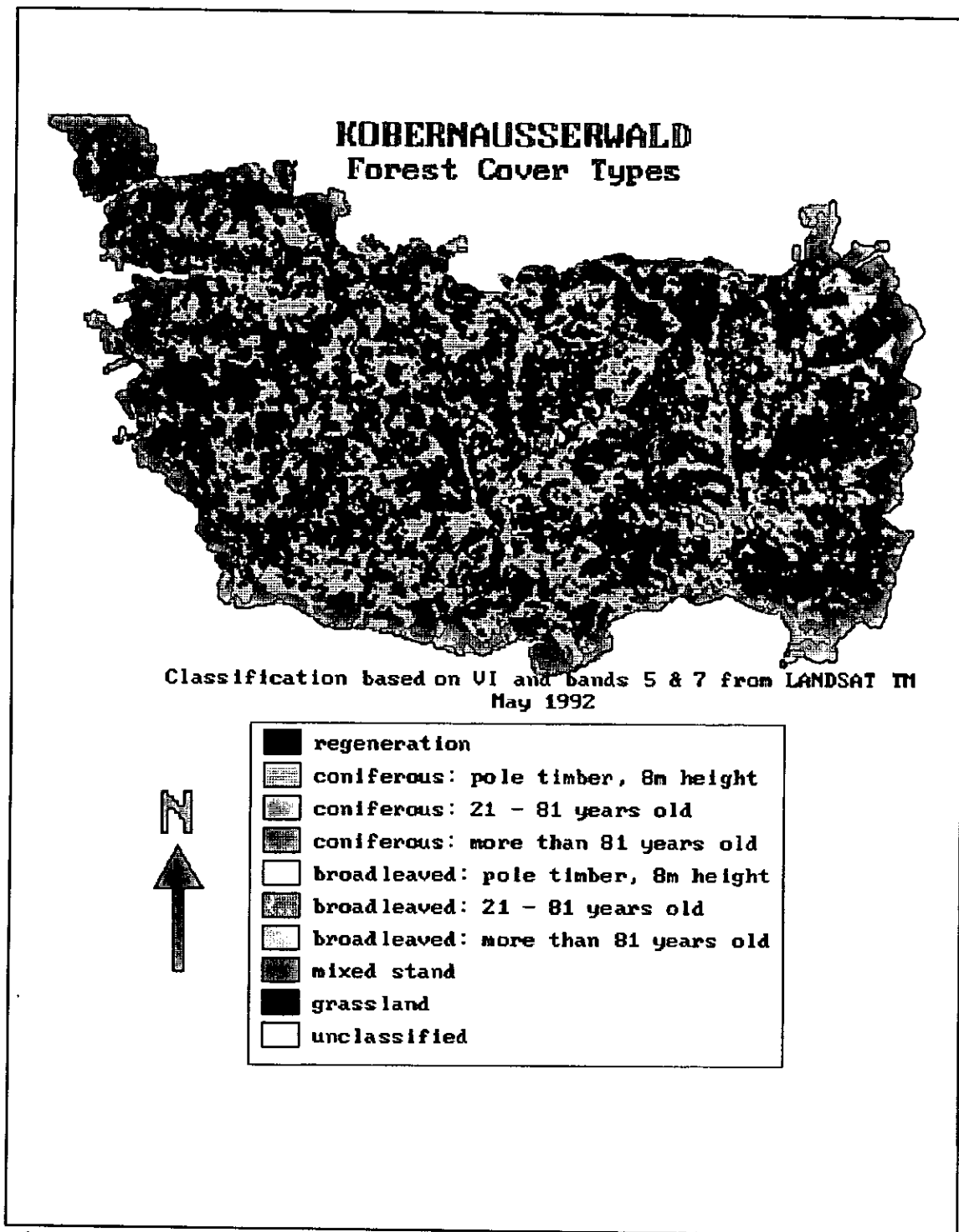


Figure 18. Map B representing forest cover types using NDVI, bands five and seven of LANDSAT TM.

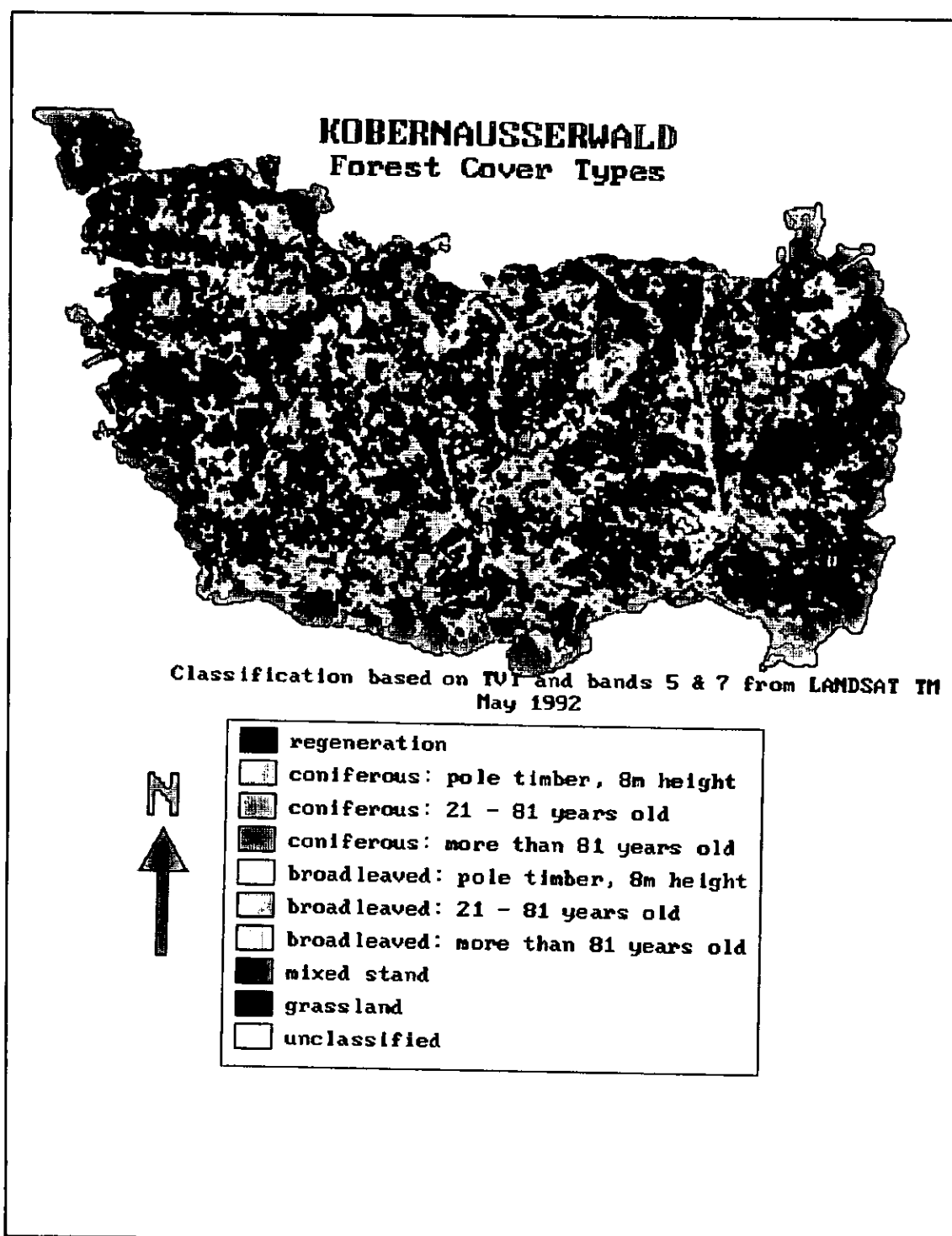


## 4.3.3. VI, Bands 5 and 7



**Figure 19.** Map C representing forest cover types using VI, bands five and seven of LANDSAT TM.

## 4.3.4. TVI, Bands 5 and 7



**Figure 20.** Map D representing forest cover types using TVI, bands five and seven of LANDSAT TM.

## 4.3.5. PC1, PC2 and Band 5

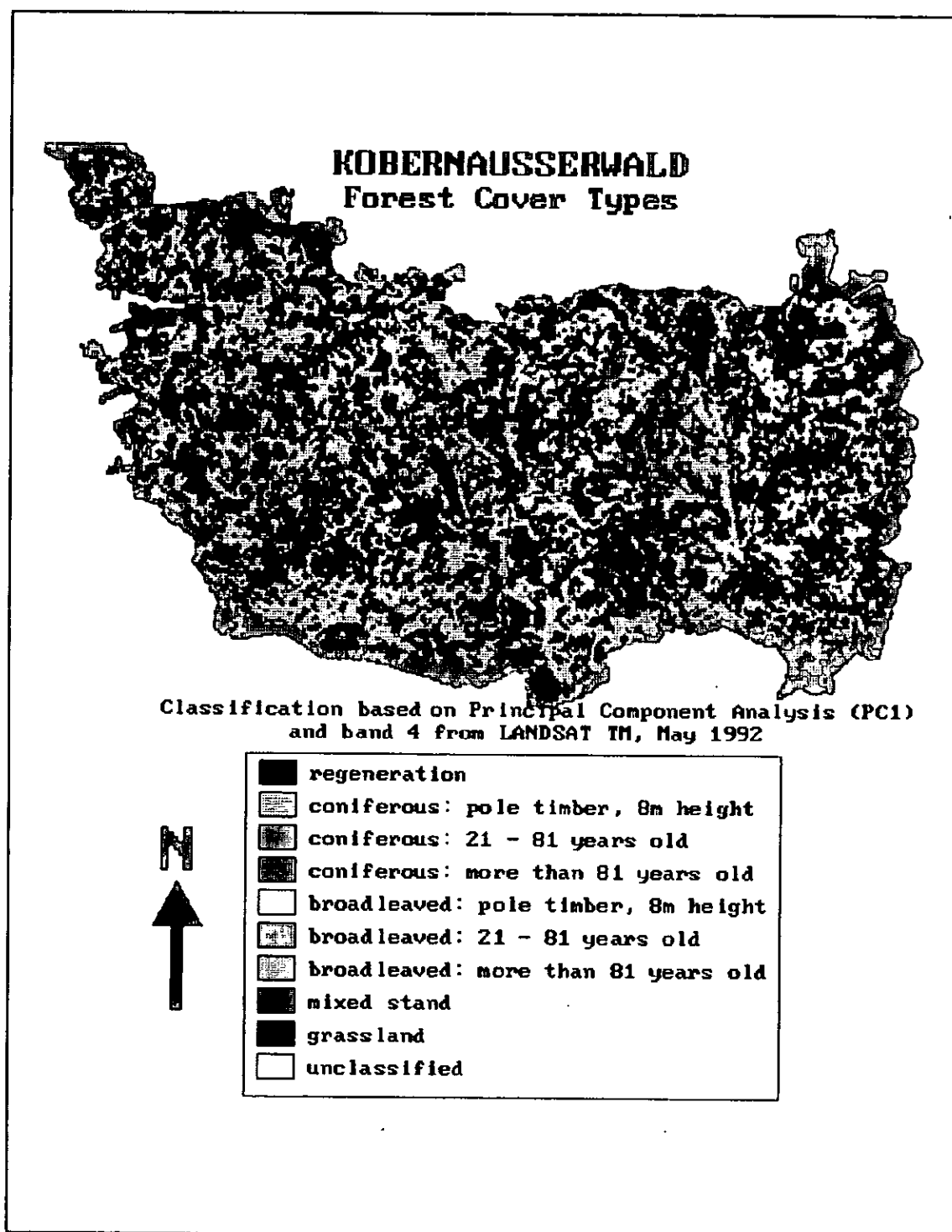
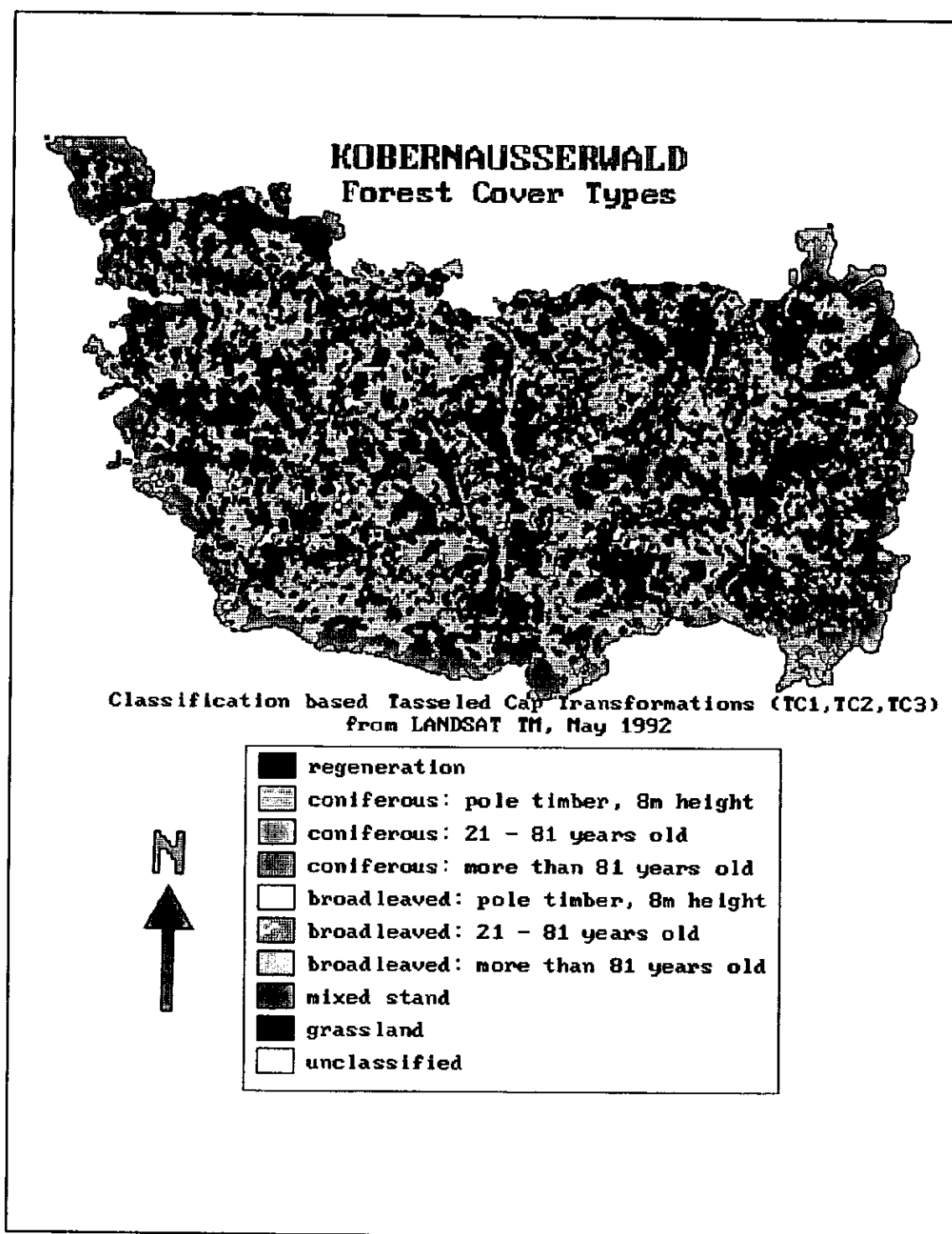


Figure 21. Map E representing forest cover types using PC1, PC2 and band five of LANDSAT TM.

### 4.3.6. Tasseled Cap Transformations



**Figure 22.** Map F representing forest cover types using derived bands of Tasseled Cap Transformations of LANDSAT TM.

#### 4.4. Statistical Results

With the advent of more advanced digital satellite remote sensing techniques, the necessity of performing an accuracy has received renewed interest. With the complexity of digital classification, there is more of a need to assess the reliability of the results. The outcome must have a statistical approach to obtain consistent results. The statistical approach comprises two steps:

- a) Confusion matrix per classification and the respective graphic showing the relationship between percentage accuracy and percentage reliability distribution per class.
- b) Kappa analysis for agreement between different classes.

There are five types of information which can be extracted from confusion matrix:

- a) **accuracy per class** which is calculated by summing the horizontal values of each class and then divided by the total of same horizontal line.
- b) **reliability per class** take in account the vertical values for the same class.
- c) **average accuracy** is calculated by summing all accuracies per class and then dividing by the number of classes.
- b) **overall accuracy** is calculated by dividing the sum of entries that form the major diagonal (i.e. the number of correct pixels) by the total number of samples selected.
- c) **average reliability** is calculated by summing all values of the reliability per class and the divided by the number of classes.

#### 4.4.1. Confusion Matrix for Classification of Untransformed Bands

Table 6a - Accuracy calculated through confusion matrix for classification of untransformed bands of LANDSAT TM (2, 3 & 4).

| CLASS<br>★ | A    | B    | C    | D    | E    | F    | G    | H    | I    | UNC | ACC  | TOTAL |
|------------|------|------|------|------|------|------|------|------|------|-----|------|-------|
| A          | 64   | 8    | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 0   | 0.89 | 72    |
| B          | 4    | 49   | 0    | 9    | 0    | 3    | 0    | 0    | 0    | 9   | 0.60 | 65    |
| C          | 0    | 0    | 62   | 0    | 0    | 0    | 0    | 0    | 10   | 0   | 0.86 | 72    |
| D          | 0    | 0    | 1    | 93   | 1    | 0    | 4    | 2    | 2    | 0   | 0.90 | 103   |
| E          | 0    | 0    | 0    | 4    | 30   | 0    | 5    | 3    | 5    | 0   | 0.64 | 47    |
| F          | 0    | 0    | 8    | 0    | 0    | 53   | 8    | 0    | 0    | 3   | 0.74 | 69    |
| G          | 1    | 0    | 0    | 0    | 5    | 0    | 67   | 0    | 5    | 0   | 0.86 | 78    |
| H          | 0    | 0    | 0    | 1    | 0    | 0    | 0    | 13   | 0    | 0   | 0.93 | 14    |
| I          | 0    | 0    | 0    | 0    | 0    | 0    | 29   | 0    | 52   | 0   | 0.64 | 81    |
| RE/CL      | 0.93 | 0.86 | 0.87 | 0.87 | 0.83 | 0.95 | 0.59 | 0.72 | 0.75 |     |      |       |
| TOTAL      | 69   | 57   | 71   | 107  | 36   | 56   | 113  | 18   | 69   |     |      | 601   |

\*The horizontal classes are the true ones and the vertical classes are from classified image.

Table 6b - The proportion for each class obtained from table 6a, in order to calculate var(K).

| CLASS                       | A      | B      | C      | D      | E      | F      | G      | H        | I      |
|-----------------------------|--------|--------|--------|--------|--------|--------|--------|----------|--------|
| P <sub>j</sub>              | 0.1148 | 0.0948 | 0.1298 | 0.1780 | 0.0516 | 0.1098 | 0.1930 | 0.0299   | 0.0982 |
| P <sub>j</sub> <sup>3</sup> | 0.0015 | 0.0009 | 0.0022 | 0.0056 | 0.0001 | 0.0013 | 0.0072 | 2.69E-05 | 0.0009 |

**Calculations**Accuracy per class (ACC):

$$\cdot \text{ACC}_{(\text{class A})} = (64/72) * 100 = 89\%$$

$$\cdot \text{ACC}_{(\text{class B})} = (49/(9+65)) * 100 = 66\%$$

Reliability per class (RE/CL):

$$\cdot \text{RE/CL}_{(\text{class A})} = (64/69) * 100 = 93\%$$

Overall accuracy (OA):

$$\cdot \text{OA} = \frac{64 + 49 + 62 + 93 + 30 + 53 + 67 + 13 + 52}{601} * 100 = 80.36\%$$

Proportion of correctly classified cases (Po):

$$\cdot \text{Po} = \frac{64 + 49 + 62 + 93 + 30 + 53 + 67 + 13 + 52}{601} = 0.803661$$

Proportion of correctly classified cases expected by chance (Pe):

$$\begin{aligned} \cdot \text{Pe} &= \frac{(64*72) + (49*65) + (62*72) + (93*103) + (30*47) + (53*69) + (67*78) + (13*14) + (52*81)}{601 * 601} \\ &= 0.125753 \end{aligned}$$

kappa statistics (K):

$$K = \frac{0.803661 - 0.125753}{1 - 0.12753} = 0.775419$$

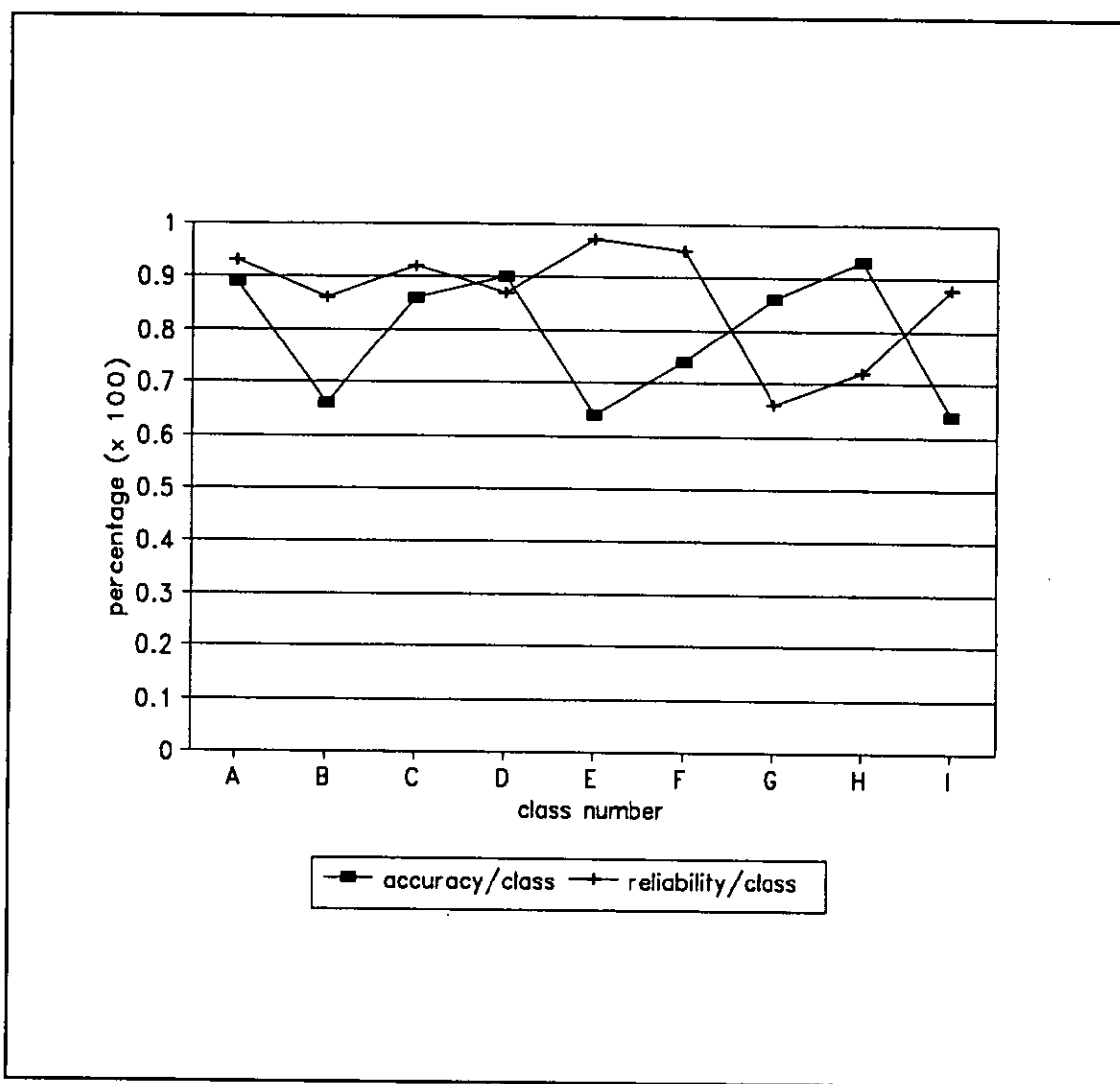
Variance (var(k)):

$$\begin{aligned} \text{var}(K) &= \frac{2}{9 * 9 * 8} * \frac{0.125753 - (2*9-3) * 0.125753^2 + 2(9-2)*0.19223}{(1 - 0.1255753)^2} \\ &= 0.000637 \end{aligned}$$

Z statistics (Z):

$$Z = \frac{0.775419}{0.000637} = 30.73109$$





**Figure 23.** Distribution of the reliability and accuracy percentages among the classes using bands 2, 3 and 4 of LANDSAT TM.

#### 4.4.2. Confusion Matrix for Classification of NDVI, Bands 5 and 7

Table 7a - Accuracy per class for classification of NDVI, bands 5 and 7 of LANDSAT TM.

| CLASS* | A    | B    | C    | D    | E    | F    | G    | H    | I    | UNC | ACC  | TOTAL |
|--------|------|------|------|------|------|------|------|------|------|-----|------|-------|
| A      | 53   | 11   | 0    | 0    | 0    | 0    | 0    | 0    | 8    | 0   | 0.74 | 72    |
| B      | 7    | 42   | 0    | 5    | 0    | 9    | 0    | 0    | 0    | 8   | 0.59 | 63    |
| C      | 2    | 0    | 60   | 0    | 7    | 0    | 0    | 0    | 3    | 0   | 0.83 | 72    |
| D      | 0    | 0    | 0    | 69   | 0    | 10   | 5    | 15   | 4    | 0   | 0.67 | 103   |
| E      | 0    | 0    | 0    | 5    | 31   | 0    | 1    | 6    | 4    | 0   | 0.66 | 47    |
| F      | 0    | 0    | 16   | 0    | 0    | 56   | 0    | 0    | 0    | 0   | 0.78 | 72    |
| G      | 4    | 0    | 0    | 0    | 0    | 0    | 68   | 6    | 0    | 0   | 0.87 | 78    |
| H      | 0    | 0    | 0    | 8    | 0    | 0    | 0    | 6    | 0    | 0   | 0.43 | 14    |
| I      | 0    | 0    | 0    | 5    | 0    | 0    | 32   | 0    | 44   | 0   | 0.54 | 81    |
| RE/CL  | 0.80 | 0.79 | 0.79 | 0.68 | 0.82 | 0.75 | 0.64 | 0.18 | 0.70 |     |      |       |
| TOTAL  | 66   | 53   | 76   | 102  | 38   | 75   | 106  | 33   | 63   |     |      | 602   |

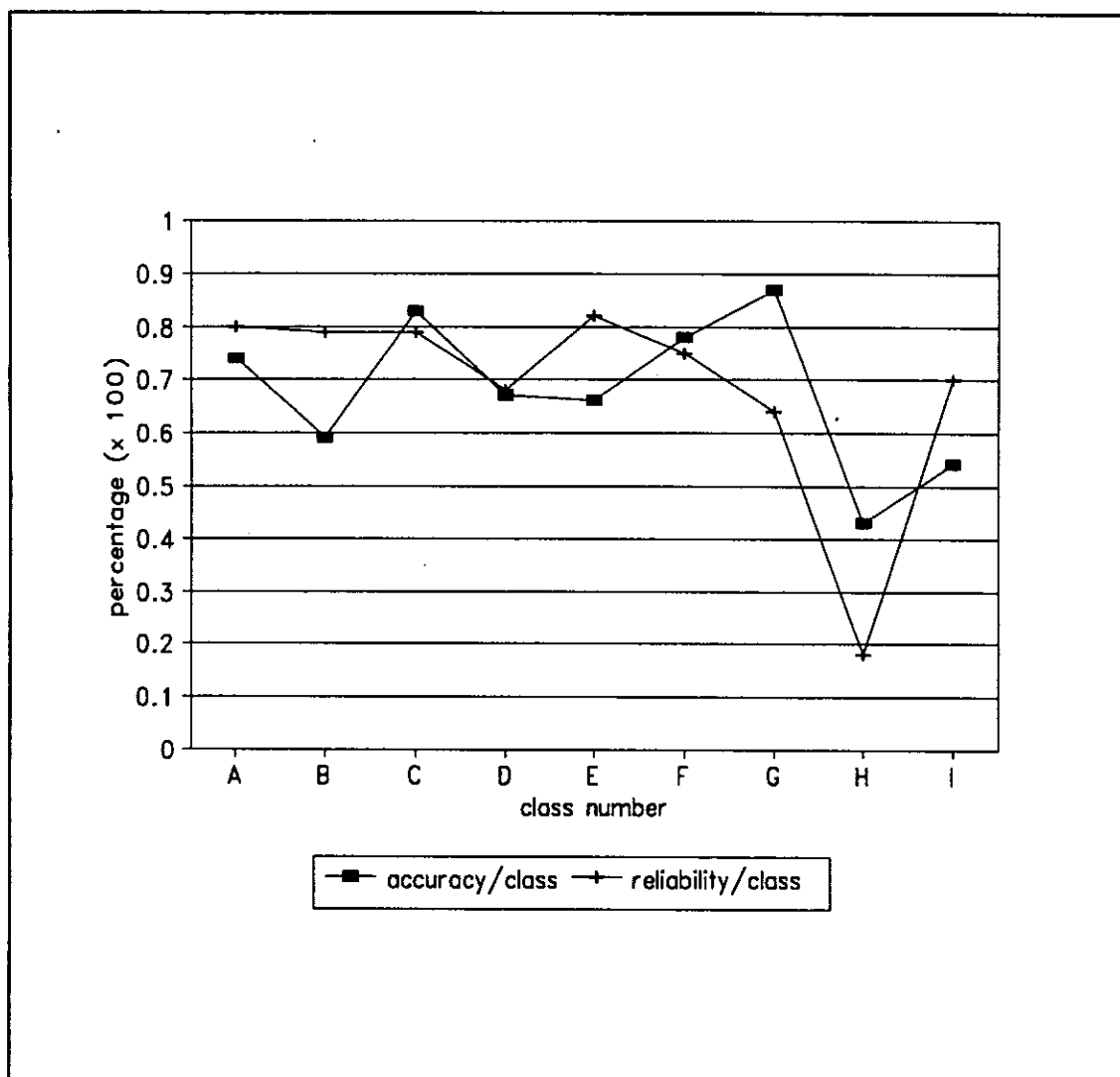
\*The horizontal classes are the true ones and the vertical classes are from classified image.

Average accuracy = 67.89      Overall accuracy = 71.25

Po = 0.712625      Pe = 0.121572

K = 0.67853      var(K) = 0.000526

Z = 29.35043



**Figure 24.** Distribution of the reliability and accuracy percentages among the classes of the classification using NDVI, bands 5 and 7.

### 4.4.3. Confusion Matrix for Classification of VI, Bands 5 and 7

Table 8a - Accuracy per class for classification of VI, bands 5 and 7 of LANDSAT TM.

| CLASS<br>* | A    | B    | C    | D    | E    | F    | G    | H    | I    | UNC | ACC  | TOTAL |
|------------|------|------|------|------|------|------|------|------|------|-----|------|-------|
| A          | 60   | 9    | 0    | 0    | 0    | 0    | 0    | 0    | 3    | 0   | 0.83 | 72    |
| B          | 7    | 40   | 0    | 7    | 0    | 8    | 0    | 0    | 0    | 10  | 0.56 | 62    |
| C          | 2    | 0    | 59   | 0    | 7    | 0    | 1    | 0    | 3    | 0   | 0.82 | 72    |
| D          | 0    | 10   | 0    | 60   | 0    | 10   | 5    | 14   | 4    | 0   | 0.58 | 103   |
| E          | 0    | 0    | 0    | 4    | 32   | 0    | 1    | 5    | 5    | 0   | 0.68 | 47    |
| F          | 0    | 0    | 16   | 8    | 0    | 39   | 0    | 9    | 0    | 0   | 0.54 | 72    |
| G          | 3    | 0    | 0    | 0    | 3    | 0    | 60   | 2    | 10   | 0   | 0.77 | 78    |
| H          | 0    | 0    | 0    | 8    | 0    | 0    | 0    | 6    | 0    | 0   | 0.43 | 14    |
| I          | 0    | 0    | 0    | 5    | 0    | 0    | 27   | 0    | 49   | 0   | 0.60 | 81    |
| RE/CL      | 0.83 | 0.68 | 0.79 | 0.65 | 0.76 | 0.68 | 0.64 | 0.17 | 0.66 |     |      |       |
| TOTAL      | 72   | 59   | 75   | 92   | 42   | 57   | 94   | 36   | 74   |     |      | 601   |

\*The horizontal classes are the true ones and the vertical classes are from classified image.

Average accuracy = 64.56

Overall accuracy = 67.30

Po = 0.662848

Pe = 0.11709

K = 0.618135

var(K) = 0.000497

Z = 27.71773

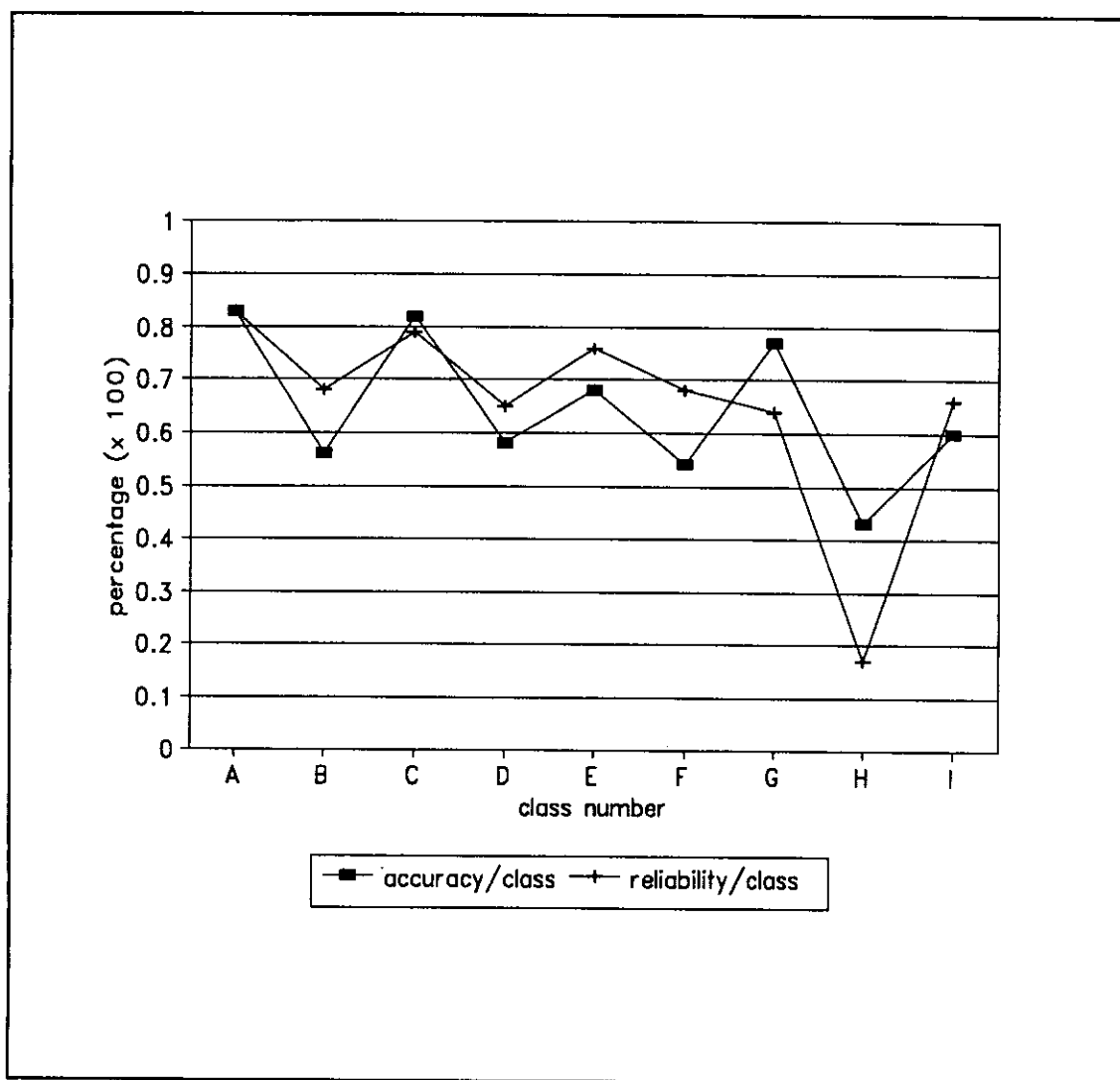


Figure 25. Distribution of the reliability and accuracy percentages among the classes of the classification using VI, bands 5 and 7 from LANDSAT TM.

#### 4.4.4. Confusion Matrix for Classification of TVI, Bands 5 and 7

Table 9a - Accuracy per class for classification of TVI, bands 5 and 7 of LANDSAT TM.

| CLASS<br>* | A    | B    | C    | D    | E    | F    | G    | H    | I    | UNC | ACC  | TOTAL |
|------------|------|------|------|------|------|------|------|------|------|-----|------|-------|
| A          | 62   | 10   | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 0   | 0.86 | 72    |
| B          | 3    | 50   | 0    | 1    | 0    | 10   | 0    | 0    | 0    | 8   | 0.69 | 64    |
| C          | 2    | 0    | 59   | 0    | 10   | 0    | 1    | 0    | 0    | 0   | 0.82 | 72    |
| D          | 0    | 0    | 0    | 80   | 5    | 0    | 5    | 8    | 5    | 0   | 0.78 | 103   |
| E          | 0    | 0    | 0    | 2    | 37   | 0    | 1    | 3    | 4    | 0   | 0.79 | 47    |
| F          | 0    | 0    | 11   | 0    | 0    | 61   | 0    | 0    | 0    | 0   | 0.85 | 72    |
| G          | 3    | 0    | 0    | 0    | 0    | 0    | 72   | 0    | 3    | 0   | 0.92 | 78    |
| H          | 0    | 0    | 1    | 9    | 0    | 0    | 0    | 4    | 0    | 0   | 0.29 | 14    |
| I          | 0    | 0    | 0    | 5    | 0    | 0    | 26   | 0    | 50   | 0   | 0.62 | 81    |
| RL/CL      | 0.88 | 0.75 | 0.83 | 0.82 | 0.71 | 0.85 | 0.69 | 0.27 | 0.81 |     |      |       |
| TOTAL      | 70   | 60   | 71   | 97   | 52   | 71   | 105  | 15   | 62   |     |      | 603   |

\* The horizontal classes are the true ones and the vertical classes are from classified image.

Average accuracy = 73.56

Overall accuracy = 83.14

Po = 0.781095

Pe = 0.123652

K = 0.750207

var(K) = 0.000532

Z = 32.43029

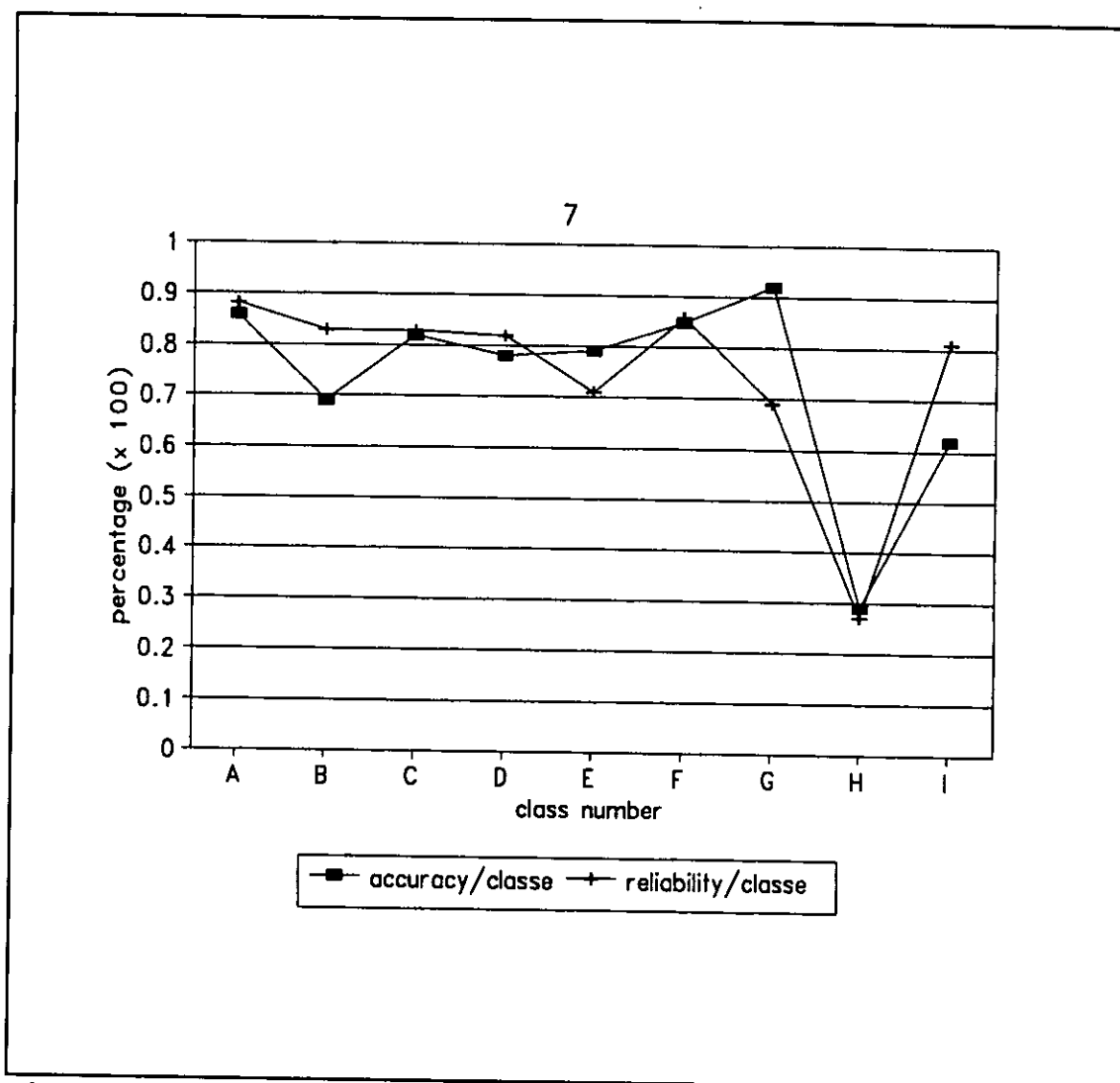


Figure 26. Distribution of the reliability and accuracy percentages of the classification for TVI, bands 5 and 7 from LANDSAT TM.

#### 4.4.5. Confusion Matrix for Classification of PC1, PC2 and Band 5

Table 10a - Accuracy per class for classification of PC1, PC2 and Band 5 of LANDSAT TM.

| CLASS<br>* | A    | B    | C    | D    | E    | F    | G    | H    | I    | UNC | ACC  | TOTAL |
|------------|------|------|------|------|------|------|------|------|------|-----|------|-------|
| A          | 65   | 7    | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 0   | 0.90 | 72    |
| B          | 2    | 60   | 0    | 3    | 0    | 0    | 0    | 0    | 0    | 7   | 0.83 | 65    |
| C          | 0    | 0    | 71   | 0    | 0    | 0    | 1    | 0    | 0    | 0   | 0.99 | 72    |
| D          | 0    | 0    | 0    | 82   | 0    | 0    | 7    | 1    | 13   | 0   | 0.80 | 103   |
| E          | 0    | 0    | 0    | 9    | 27   | 0    | 0    | 8    | 3    | 0   | 0.57 | 47    |
| F          | 0    | 0    | 4    | 0    | 0    | 72   | 0    | 0    | 0    | 0   | 0.94 | 76    |
| G          | 4    | 0    | 16   | 0    | 0    | 0    | 58   | 0    | 0    | 0   | 0.74 | 78    |
| H          | 0    | 0    | 0    | 0    | 6    | 0    | 0    | 8    | 0    | 0   | 0.57 | 14    |
| I          | 0    | 0    | 0    | 1    | 0    | 0    | 13   | 0    | 67   | 0   | 0.83 | 81    |
| RL/CL      | 0.91 | 0.89 | 0.78 | 0.86 | 0.82 | 1.00 | 0.73 | 0.47 | 0.81 |     |      |       |
| TOTAL      | 71   | 67   | 91   | 95   | 33   | 72   | 79   | 17   | 83   |     |      | 608   |

\*The horizontal classes are the true ones and the vertical classes are from classified image.

Average accuracy = 79.67%

Overall accuracy = 83.88

Po = 0.838816

Pe = 0.124302

K = 0.815936

var(K) = 0.000507

Z = 36.22768



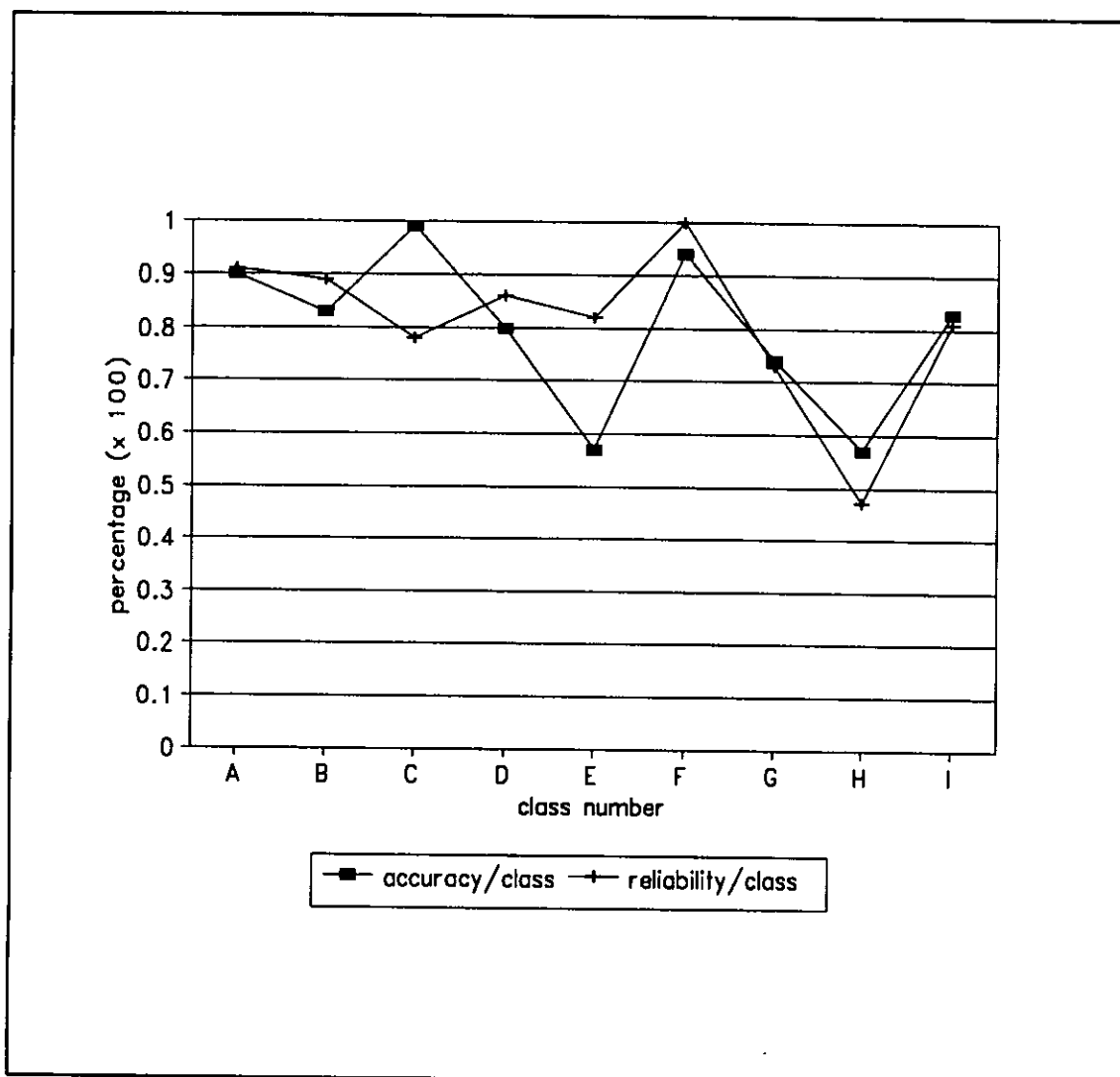


Figure 27. Distribution of the reliability and accuracy percentages among the classes of the classification for PC1, PC2 and band 5 from LANDSAT TM.

#### 4.4.6. Confusion Matrix for Classification of Tasselled Cap Transformations

Table 11a - Accuracy per class for classification of Tasselled Cap Transformations of LANDSAT TM.

| CLASS<br>* | A    | B    | C    | D    | E    | F    | G    | H    | I  | UNC | ACC  | TOTAL |
|------------|------|------|------|------|------|------|------|------|----|-----|------|-------|
| A          | 69   | 3    | 0    | 0    | 0    | 0    | 0    | 0    | 0  | 0   | 0.96 | 72    |
| B          | 1    | 64   | 0    | 1    | 0    | 0    | 0    | 0    | 0  | 6   | 0.90 | 67    |
| C          | 0    | 0    | 72   | 0    | 0    | 0    | 0    | 0    | 0  | 0   | 1.00 | 72    |
| D          | 0    | 0    | 0    | 97   | 0    | 0    | 5    | 0    | 1  | 0   | 0.94 | 103   |
| E          | 0    | 0    | 0    | 3    | 43   | 0    | 0    | 0    | 1  | 0   | 0.91 | 47    |
| F          | 0    | 0    | 2    | 0    | 0    | 68   | 0    | 0    | 0  | 2   | 0.94 | 70    |
| G          | 1    | 0    | 1    | 0    | 0    | 0    | 76   | 0    | 0  | 0   | 0.97 | 78    |
| H          | 0    | 0    | 0    | 0    | 0    | 0    | 0    | 12   | 0  | 2   | 0.86 | 12    |
| I          | 0    | 0    | 0    | 0    | 0    | 0    | 14   | 0    | 67 | 0   | 0.83 | 81    |
| RL/CL      | 0.97 | 0.95 | 0.96 | 0.96 | 1.00 | 0.80 | 1.00 | 0.97 |    |     |      |       |
| TOTAL      | 71   | 67   | 75   | 101  | 43   | 68   | 95   | 12   | 69 |     |      | 602   |

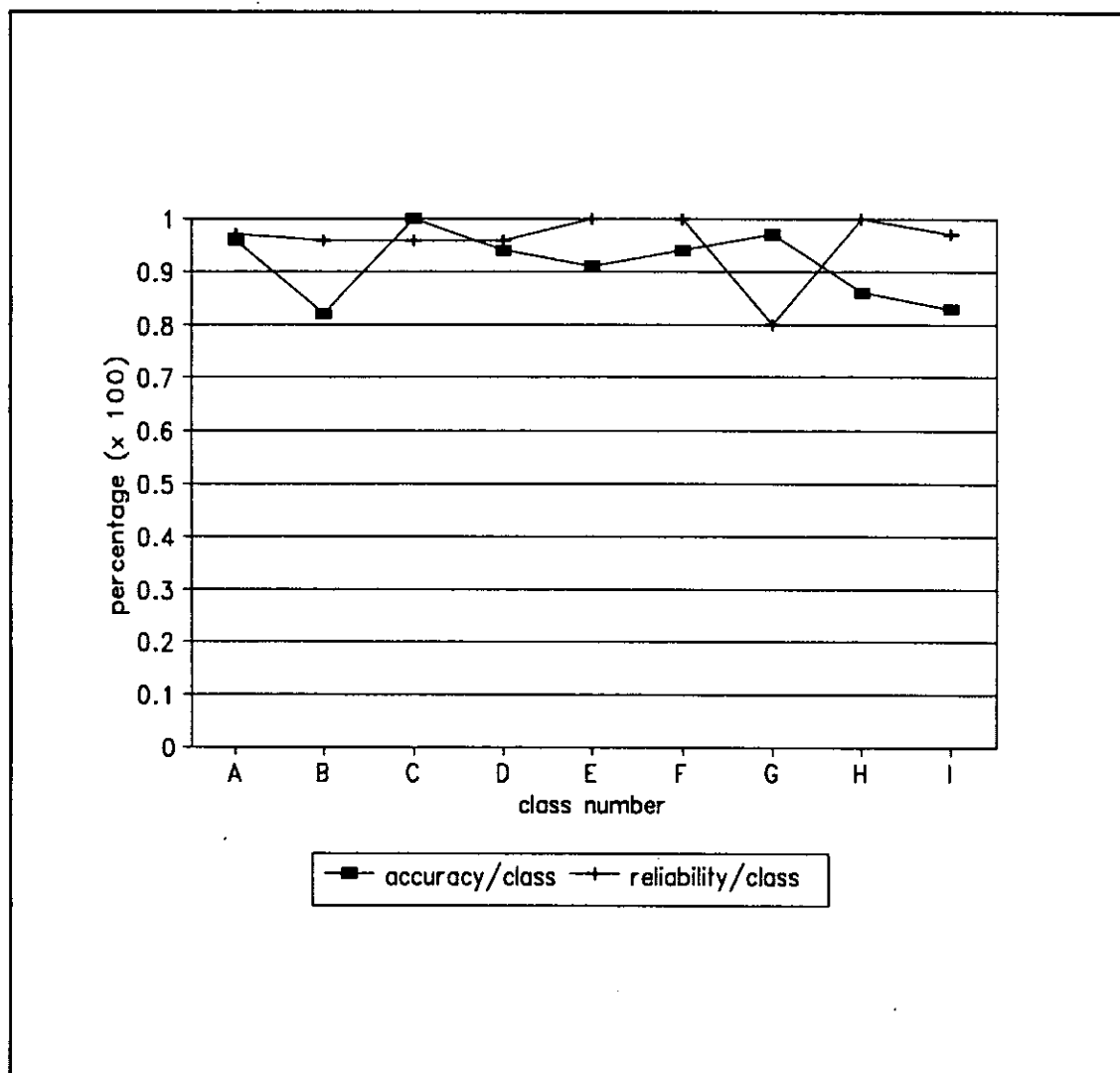
The horizontal classes are the true ones and the vertical classes are from classified image.

Average accuracy = 93.53      Average reliability = 96.07

Overall accuracy = 94.27

Po = 0.945092      Pe = 0.125307      K = 0.937225

var(k) = 0.000516      Z = 41.25927



**Figure 28.** Distribution of reliability and accuracy percentages among the classes of the classification using Tasselled Cap Functions from LANDSAT TM.

## **CHAPTER 5: DISCUSSION AND CONCLUSIONS**

---

### **5.1. Introduction**

The study has been concerned with the application of various remote sensing techniques and the mapping of Kobernausserwald area with respect to different forest cover types.

For the data analysis, some discussion of the statistical output, conclusions and recommendations are put forward in the following paragraphs.

### **5.2. Discussion of the Results**

From the output of eigen vectors generated by principal component analysis, it is possible to say that some bands do not give so much information. Most of the information (e.g., 98.96%) were found in two bands (PC1 and PC2).

The maximum likelihood classifier was applied to six independent linear transformations. The hypothetic relationships between spectral classes and informational classes were checked after classification. Therefore, clusters erroneously labelled, or those who received two possible labels were re-assigned in a matching with a reference data. Spectral classes were grouped according to the pre-established classification scheme.

The Figures 16, 17, 18, 19, 20 and 21 show the output of the classification. Although they look similar, some differences are presented in the size of each category. This differences are due to the different functions of each linear transformation. NDVI, VI and TVI are more related with the amount of vegetation present in a given area. This aspect can be noticed in classes B, C and E, where the type of forest management influence the results.

The Tasseled Cap Transformation is presented as the reliable one, because of the type of categorization performed (species and age classes). This is due to the performance of the Tasseled Cap to be related with different stages of plant growth.

In the untransformed TM bands, the spectral response were more distinguishable in bands 4, 5 and 7, due to high performance of this part of the EM spectrum in respect to the vegetation green leaves.

An ideal vegetation index would be highly sensitive to vegetation, insensitive to soil background changes, and only slightly influenced by atmospheric path radiance. None of the studied indices fully met these criteria.

### **5.2.1. Accuracy Assessment**

The accuracy of classified images was determined in two ways. The first one using a confusion matrix and the second one through kappa statistics.

The results through confusion matrix are shown in 4.3.1., 4.3.2., 4.3.3., 4.3.4., 4.3.5. and 4.3.6. in sequence for different maps. Figures 22, 23, 24, 25, 26 and 27 show the distribution of the accuracy per class. It can be noticed that class B (coniferous, pole timber taller than 8 metres) is the problematic one. The accuracy in this class is considerably lower than the others as it is summarized in Table 12. In same Table 12, the class H (mixed stand more than 81 years old) has low accuracies when using techniques NDVI, VI, TVI and PCA. This is because NDVI, VI and TVI express the amount of green biomass, and do not separate the different species. In the case of class H, PCA should have had higher accuracy comparing to the other indeces, because of the high variability in the selected bands that make the separation between classes more evident.

Table 12 - Performance of the accuracy for different classes for different classifications

| Technique/<br>classes | UNT  | NDVI | VI   | TVI  | PCA  | TCT  |
|-----------------------|------|------|------|------|------|------|
| <b>A</b>              | 0.89 | 0.74 | 0.83 | 0.86 | 0.90 | 0.96 |
| <b>B</b>              | 0.60 | 0.59 | 0.56 | 0.69 | 0.83 | 0.90 |
| <b>C</b>              | 0.86 | 0.83 | 0.82 | 0.82 | 0.99 | 1.00 |
| <b>D</b>              | 0.90 | 0.67 | 0.58 | 0.78 | 0.80 | 0.94 |
| <b>E</b>              | 0.64 | 0.66 | 0.68 | 0.79 | 0.57 | 0.91 |
| <b>F</b>              | 0.74 | 0.78 | 0.54 | 0.85 | 0.94 | 0.94 |
| <b>G</b>              | 0.86 | 0.87 | 0.77 | 0.92 | 0.74 | 0.97 |
| <b>H</b>              | 0.93 | 0.43 | 0.43 | 0.29 | 0.57 | 0.86 |
| <b>I</b>              | 0.64 | 0.54 | 0.60 | 0.62 | 0.83 | 0.83 |

VI - Using Vegetation Index or Ratio Vegetation Index

NDVI - Using Normalized Difference Vegetation Index

UNT - Using Untransformed Bands

TVI - Using Transformed Vegetation Index

PCA - Using Principal Component Analysis

TCT - Using Tasselled Cap Transformation

These low values are because in some portions of the area, the category B is much dense than the others parts, and therefore it was mixed with the other coniferous categories.

Generally speaking, classification results were higher than the values showed by overall accuracy. Problematic categories like class B and A brought down the overall results.

### 5.3. Conclusions

This study attempted to evaluate quantitatively the usefulness of TM data for digital classification for forest resources in temperate conditions. Under the conditions tested, the major findings can be summarized in Tables 13 and 14.

Table 13 - Resume of overall accuracy for different maps using different linear transformations.

| TECHNIQUE | OVERALL ACCURACY |
|-----------|------------------|
| UNT       | 80.30 %          |
| NDVI      | 71.25 %          |
| VI        | 67.30 %          |
| TVI       | 83.14 %          |
| PCA       | 85.78 %          |
| TCT       | 94.27 %          |

The overall accuracy is presented as the best figure of the assessment, because it takes into account all classes and includes the unclassified pixels.

Table 14 - Resume of the results through Kappa statistics.

| TECHNIQUE | KAPPA VALUE | VARIANCE | Z VALUE | RESULT |
|-----------|-------------|----------|---------|--------|
| VI        | 0.73395     | 0.000658 | 28.607  | NS     |
| NDVI      | 0.74028     | 0.000642 | 29.224  | NS     |
| UNT       | 0.83252     | 0.000671 | 32.137  | NS     |
| TVI       | 0.77463     | 0.000571 | 32.430  | NS     |
| PCA       | 0.81470     | 0.000510 | 36.065  | NS     |
| TCT       | 0.77462     | 0.000516 | 41.259  | NS     |

NS - no significance

From Table 13 and 14, the two accuracy approaches show different results for different techniques. This result can not be taken as a standard conclusion for comparison with other case studies. That is because in this case study some of the classes were occupying small areas. Therefore, the number of true pixels for test comparison were not enough.

In Table 14, all techniques performed show no significant differences among the categorized classes. However, some classes show low accuracy as presented in Table 13.

Although the magnitude of the results could be different for other study areas with different conditions, the results presented here should prove useful for determining the amount of information that can be expected from a particular vegetation index for a particular forest cover type and for a given climatic condition.



In conclusion, linear transformation techniques used in this study can support forest classification by enhancing the TM data if ground truth is available for specifying meaningful forest cover categories for specific areas.

The conclusions can be summarized as following:

- o The linear transformation to enhance TM data prior to image classification for forest mapping purposes show very good result as the overall accuracy performed by all classifications are higher than 65%.
- o Among the all linear transformations used, Tasselled Cap Transformation presents the highest and the best overall classification accuracy.
- o The kappa analysis shows that all linear transformations gave acceptable result and much better than the untransformed TM bands, but in the same time there was no significance differences between their performance.
- o It was not possible to rank the different linear transformation techniques for the same reason given by the previous conclusion remark.

## BIBLIOGRAPHIC REFERENCES

---

- Baret, F. and Guyot, G. (1991)**, Potentials and limits of vegetation indices for LAI and APAR assessment. *Remote Sens. Environ.* 35:161 - 173.
- Baret, F.; Guyot, G. and Major, D. (1989b)**, TSAVI: a vegetation index which minimizes soil brightness effects on LAI and APAR estimation, *12th Canadian Symp. on Remote Sensing and IGARSS'90. Vancouver, Canada.* 10 - 14 July 1989, 4 pp.
- Bolstad, P. V. and Lillesand, T. M. (1992)**, Improved classification of forest vegetation in Northern Wisconsin through a rule-based combination of soils, terrain and Landsat thematic mapper data. *Forest Science*, 38(1):2 - 20.
- Budd, J. T. C. and Milton, E. J. (1982)**, Remote sensing of salt marsh vegetation in the first four proposed thematic mapper bands. *Int. J. Remote Sensing*. 3(2):147 - 161.
- Campbell, J. B. (1987)**, Introduction to Remote Sensing. *The Guilford Press*. pp: 379-398.
- Carnegie, D. M.; de Gloria, S. D. and Colwell, R. N. (1974)**, Usefulness of ERTS-1 and supporting aircraft data for monitoring plant development and range conditions in California's annual grassland. *BLM Final Report 53500-CT3-266(N)*.
- Carstensen Jr., L. W. (1987)**, A measure of similarity for cellular maps. *The American Cartographer*. 14(4):345 - 358.

- Cohen, W. B. (1991)**, Response of vegetation indices to change in three measures of leaf water stress. *Photogrammetric Engineering and Remote Sensing*. 57(1), p.195.
- 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 Mead, R. A. (1983)**, A quantitative method to test for consistency and correctness in photointerpretation. *Photogrammetric Engineering and Remote Sensing*. 49:69-74.
- Congalton, R. G.; Oderwald, R. G. and Mead, R. A. (1983)**, Assessing landsat classification accuracy using discrete multivariate analysis statistical techniques. *Photogrammetric Engineering and Remote Sensing*. 49:1671-1678.
- Elvidge, C. D. and Lyon, R. J. P. (1985)**, Influence of rock - soil spectral variation on assessment of green biomass, *Remote Sens. Environ.* 17:221 - 232.
- Gutman, G. (1989)**, On the relationship between monthly mean and maximum value composite normalized vegetation indices. *Int. J. of Remote Sensing*. 12(1,2,3), p.1317.
- Hay, A. M. (1988)**, The derivation of global estimates from a confusion matrix. *Int. J. of Remote Sensing*. 9(8):1395 - 1398.
- Howarth, P. J. and Boasson, E. (1983)**, Landsat digital enhancements for change detection in urban environments. *Remote Sensing of Environment*. 13:149 - 160.

- Huete, A. R.; Jackson, R. D. and Post, D. F. (1985)**, Spectral response of a plant canopy with different soil backgrounds, *Remote Sensing of Environment*. 17:37 - 53.
- Jackson, R. D.; Slater, P. N. and Pinter Jr., P. J. (1983)**, Discrimination of growth and water stress in wheat by various vegetation indices through clear and turbid atmospheres. *Remote Sensing of Environment*. 13:187 - 208.
- Jordan, C. F. (1969)**, Derivation of leaf area index from quality of light on the forest floor. *Ecology*. 50, 663 - 666.
- Jupp, D. L. B. (1989)**, The stability of global estimates from confusion matrices. *Int. J. of Remote Sensing*. 10(9):1563 - 1569.
- Kauth, R. J. and Thomas, G. S. (1976)**, The tasselled cap - a graphic description of the spectral temporal development of agricultural crops as seen by LANDSAT. *Processing of Remote Sensing Data*. LARS, Purdue.
- Karteris, M. A. (1990)**, The utility of digital thematic mapper data for natural classification. *Int. J. of Remote Sensing*. 11(2), p.1589.
- Lillesand, T. M. and Kiefer, R. W. (1987)**, Remote Sensing and Image Interpretation. *Second Edition*. John Wiley & Sons. pp. 669-688.
- Major, D. J.; Baret, F. and Guyot, G. (1990)**, A ratio vegetation index adjusted for soil brightness. *Int. J. of Remote Sensing*. 11(1), p.727.

- Pearson, R. L. and Miller, L. D. (1972)**, Remote mapping of standing crop biomass for estimation of the productivity of the short grass prairie. *Eighth International Symposium on Remote Sensing of Environment*, University of Michigan, Ann Arbor, Mich. 1357 - 1381.
- Pearson, R. L.; Miller, L. D. and Tucker, C. J. (1976)**, Hand - held spectral radiometer to measure gramineous biomass. *Appl. Optics*. 15, 416 - 418.
- Richardson, A. J.; Everitt, J. H. and Gausman, H. W. (1983)**, Short Communication. Radiometric estimation of biomass and nitrogen content of Alicia Grass. *Remote Sensing of Environment*. 13:179 - 184.
- Richardson, A. J. and Wiegand, C. L. (1977)**, Distinguishing vegetation from soil background information. A grey mapping technique allows delineation of Landsat scene into vegetative cover stages, degrees of soil brightness, and water. *Photogrammetric Engineering and Remote Sensing*. 43(12), pp.1541 -1552.
- Rouse, J. W.; Haas, R. H.; Schell, J. A. and Deering, D. W. (1973)**, Monitoring vegetation systems in the great plains with ERTS. *Third ERTS Symposium*, NASA SP-351 1:309 - 317.
- Rouse, J. W.; Haas, R. H.; Schell, J. A.; Deering, D. W. and Harlan, J. C. (1974)**, Monitoring the vernal advancement and retrogration (greenwave effect) of natural vegetation. *NASA/GSFC Type III Final Report, Greenbelt, Md.* 371 pp.c

- Siegel, S. and Castellan Jr., N.J. (1988)**, Nonparametric Statistics for the Behavioral Sciences. *Second Edition. McGraw-Hill International Editions, Statistics Series.* pp. 284-290.
- Spanner, M. A.; Pierce, L. L.; Peterson, D. L. and Running, S.W. (1990)**, Remote sensing of temperate coniferous forest leaf area index. - The influence of canopy closure, understorey vegetation and background reflectance. *Int. J. of Remote Sensing.* 11(1), p.95.
- Townshend, J. R. G. (1984)**, Agricultural land-cover discrimination using thematic mapper spectral bands. *Int. J. Remote Sensing.* 5(4):681 - 698.
- Townshend, J. R. G. and Tucker, C. J. (1984)**, Objective assessment of advanced very high resolution radiometer data for land cover mapping. *Int. J. Remote Sensing.* 5(2), p.497.
- Tucker, C. J. (1979)**, Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of Environment.* 8:127 - 150.
- Vogelmann, J. E. (1990)**, Comparison between two vegetation indices for measuring different types of forest damage in North-eastern United States. *Int. J. of Remote Sensing.* 11(2), p.2281.
- Wardley, N. W. (1984)**, Vegetation index variability as a function of viewing geometry. *Int.J. Remote Sensing.* 5(5):861 - 870.
- Wardley, N. W. and Curran, P. J. (1984)**, The estimation of green-leaf-area index from remotely sensed airborne multispectral scanner data. *Int.J. Remote Sensing.* 5(4):671 - 679.

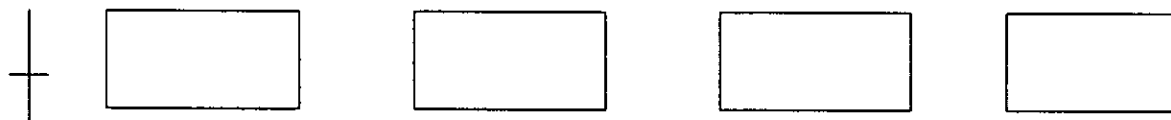
## **APPENDICES**

## APPENDIX 1

### Test Field Data Collection Form

- 1) Test Field Number :
- 2) Test Field Location :
- 3) Date :
- 4) Time :
- 5) Cover Type Expected :
- 6) Cover Type Found :
- 7) Class Results (if differ) :
- 8) Slope :
- 9) Soil Color (if appl) :
- 10) Density (trees/u.a. - 150 m x 150 m) :
- 11) Visible Cover Type Color :
- 12) Color on 234 TM :
- 13) Color on 453 TM :
- 14) Any Strange Phenomenon . :
- 15) Number Photos :
- 16) Direction of the Photos :

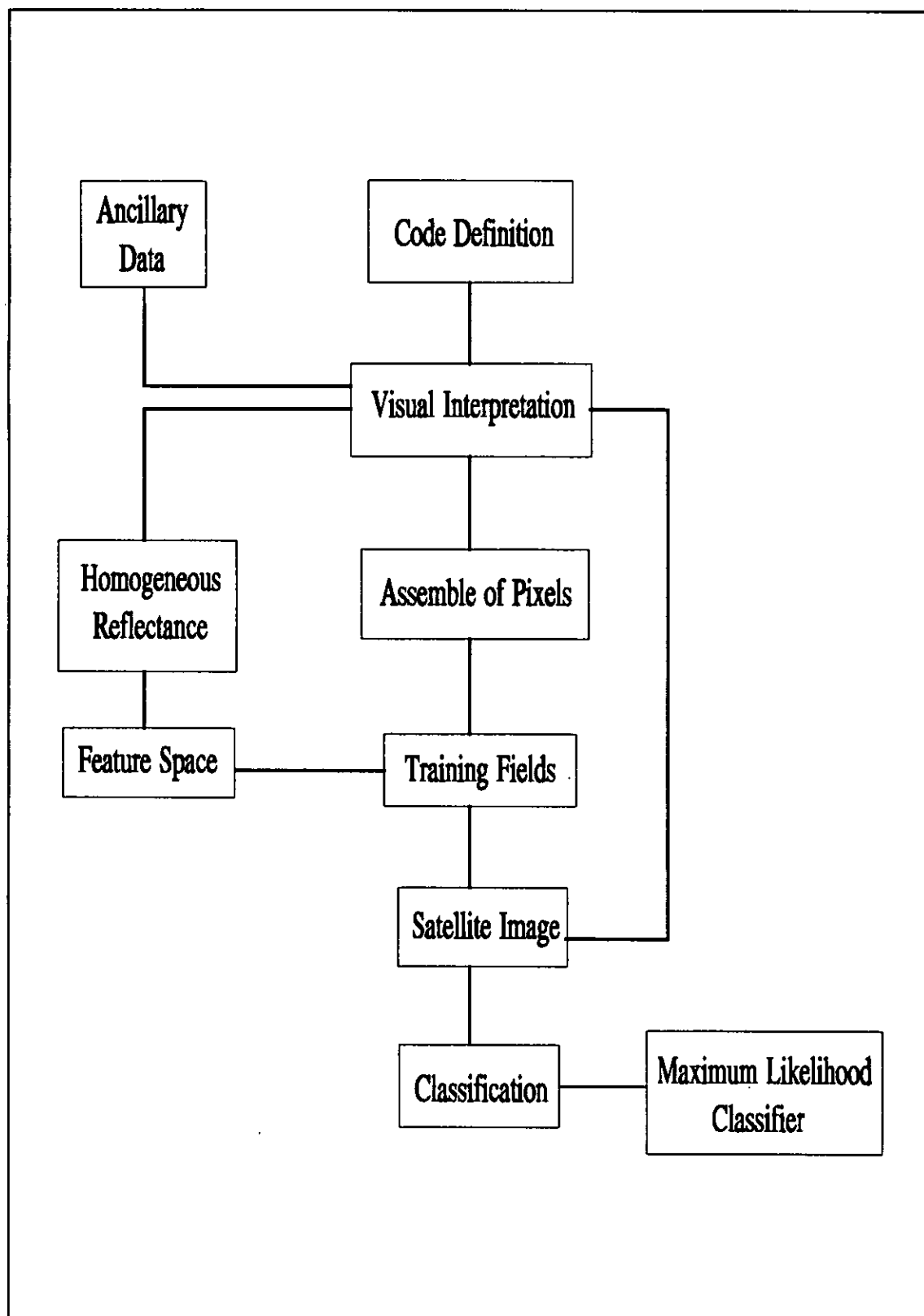
(Draw arrows to show direction)



- 17) Observer Notes :



## APPENDIX 2



Steps followed for digital image classification

### APPENDIX 3

A) The proportions for each class obtained from table 7 in order to calculate  $\text{var}(K)$  for classification using NDVI, bands 5 and 7.

| CLASS   | A      | B      | C      | D      | E      | F      | G      | H       | I      |
|---------|--------|--------|--------|--------|--------|--------|--------|---------|--------|
| $P_j$   | 0.1263 | 0.0847 | 0.1429 | 0.1694 | 0.0515 | 0.1080 | 0.1861 | 0.0449  | 0.0864 |
| $P_j^3$ | 0.0020 | 0.0006 | 0.0029 | 0.0049 | 0.0001 | 0.0013 | 0.0064 | 9.02-05 | 0.0006 |

B) The proportions for each class obtained from table 8 in order to calculate  $\text{var}(K)$  for classification using VI, bands 5 and 7.

| CLASS   | A      | B      | C      | D      | E      | F      | G      | H        | I      |
|---------|--------|--------|--------|--------|--------|--------|--------|----------|--------|
| $P_j$   | 0.1248 | 0.0815 | 0.1414 | 0.1731 | 0.0533 | 0.1032 | 0.1812 | 0.0449   | 0.0965 |
| $P_j^3$ | 0.0019 | 0.0005 | 0.0028 | 0.0052 | 0.0002 | 0.0011 | 0.0060 | 9.07E-05 | 0.0009 |

C) The proportions for each class obtained from table 9 in order to calculate  $\text{var}(K)$  for classification using TVI, bands 5 and 7.

| CLASS   | A      | B      | C      | D      | E      | F      | G      | H        | I      |
|---------|--------|--------|--------|--------|--------|--------|--------|----------|--------|
| $P_j$   | 0.1161 | 0.0995 | 0.1343 | 0.1609 | 0.0697 | 0.1178 | 0.1791 | 0.0249   | 0.0978 |
| $P_j^3$ | 0.0016 | 0.0009 | 0.0024 | 0.0042 | 0.0003 | 0.0016 | 0.0058 | 1.54E-05 | 0.0009 |

D) The proportions of each class obtained from table 10 in order to calculate the  $\text{var}(K)$  for classification using PC1, PC2 and band 5.

| CLASS   | A      | B      | C      | D      | E      | F      | G      | H        | I      |
|---------|--------|--------|--------|--------|--------|--------|--------|----------|--------|
| $P_j$   | 0.1176 | 0.1109 | 0.1507 | 0.1573 | 0.0546 | 0.1126 | 0.1308 | 0.0282   | 0.1374 |
| $P_j^3$ | 0.0016 | 0.0014 | 0.0034 | 0.0039 | 0.0002 | 0.0014 | 0.0022 | 2.23E-05 | 0.0026 |

E) The proportions of each class obtained from table 11 in order to calculate  $\text{var}(K)$  for classification using Tasselled Cap Functions.

| CLASS   | A      | B      | C      | D      | E      | F      | G      | H        | I      |
|---------|--------|--------|--------|--------|--------|--------|--------|----------|--------|
| $P_j$   | 0.1181 | 0.1115 | 0.1248 | 0.1681 | 0.0716 | 0.1132 | 0.1581 | 0.0199   | 0.1148 |
| $P_j^3$ | 0.0017 | 0.0014 | 0.0019 | 0.0048 | 0.0004 | 0.0015 | 0.0039 | 7.96E-06 | 0.0015 |